Πάντα ῥεῖ: Motion compensation strategies for image-guided high-intensity focused ultrasound and external beam radiotherapy in deformable organs

Cornel Zachiu

Cover: A magnetic resonance, an ultrasound and a computed tomography image reveal simultaneously the location of a moving liver tumor, allowing for an accurate and precise treatment delivery by a high-intensity focused ultrasound and external beam radiotherapy machine.
Design: C. Zachiu

Title: Πάντα ῥεῖ (Panta rhei, *eng.* Everything flows) is a quote attributed to the Greek philosopher Heraclitus of Ephesus. The expression alludes to the fact that everything in the Universe is in constant change...including the location of the targeted pathology during lengthy high-intensity focused ultrasound and external beam radiotherapy treatments in deformable organs.
Πάντα ῥεῖ: Motion compensation strategies for image-guided high-intensity focused ultrasound and external beam radiotherapy in deformable organs
PhD Thesis, Utrecht University, The Netherlands

Manuscript

Layout: C. Zachiu
Typeset in: LaTeX
Printed by: ProefschriftMaken || www.proefschriftmaken.nl
ISBN: 978-90-393-6960-9

This work is part of the research programme OnTrack with project number 12813, which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO).

Πάντα ῥεῖ: Motion compensation strategies for image-guided high-intensity focused ultrasound and external beam radiotherapy in deformable organs

Alles vloeit: Bewegingscompensatiestrategieën voor beeldgestuurde hoge-intensiteit gefocust ultrageluid en uitwendige bestraling radiotherapie in vervormbare organen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. G. J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op woensdag 25 april 2018 des ochtends te 10.30 uur

door

Cornel Zachiu

geboren 4 september 1988
te Sebeş, Roemenië

Promotoren:	Prof. Dr. C.T.W. Moonen
	Prof. Dr. B.W. Raaymakers
Copromotoren:	Dr. B. Denis de Senneville
	Dr. M.G. Ries

"One doesn't discover new lands without consenting to lose sight, for a very long time, of the shore."
André Gide, *The Counterfeiters* (1925).

Contents

List of Abbreviations

ANLM Adaptive Non-Local Means

AP Anterior-Posterior

BHTE Bioheat Transfer Equation

CBCT Cone-Beam Computed Tomography

CC Cranio-Caudal

CE Contrast-Enhanced

$\mathbf{CEM_{43}}$ Cumulative equivalent minutes at 43°C

CTV Clinical Tumor Volume

CT Computed Tomography

CUDA Compute Unified Device Architecture

DSC Dice Similarity Coefficient

EBRT External Beam Radiotherapy

EE Endpoint Error

EM Electromagnetic

EPI Echo-Planar Imaging

FA Flip Angle

FEP Flow Endpoint

FOV Field-Of-View

GPU Graphics Processing Unit

H & **S** Horn & Schunck

HFD Hybrid Feature-based Diffeomorphic

HIFU High Intensity Focused Ultrasound

IIR Infinite Impulse Response

Linac Linear Accelerator

LR Left-Right

MI Mutual Information

MIND Modality Independent Neighborhood Descriptor

MLC Multi-Leaf Collimator

MR Magnetic Resonance

MRg-HIFU Magnetic Resonance-guided High Intensity Focused Ultrasound

MRg-EBRT Magnetic Resonance-guided External Beam Radiotherapy

MRI Magnetic Resonance Imaging

MSE Mean Square Error

NCC Normalized Cross-Correlation

NLM Non-Local Means

NPV Non-Perfused Volume

OAR Organ-At-Risk

PID Proportional-Integral-Derivative

PRFS Proton Resonance Frequency Shift

RF Radiofrequency

ROI Region-Of-Interest

RS Registration Step

SNR Signal-to-Noise Ratio

TE Echo Time

TR Repetition Time

TRE Target Registration Error

US Ultrasound

VOI Volume-Of-Interest

CHAPTER 1

Introduction

"There is nothing permanent except change."
- Heraclitus

1.1 General introduction

Cancer is a term attributed to a broad range of diseases characterized by an uncontrolled cell growth.[1–3] The growth usually originates from a single cell that suffers mutation or genetic damage, leading to a production of abnormal cells lacking antiproliferation and apoptosis mechanisms. These cells will continue to divide without control, invading nearby tissues and ultimately form the so-called primary tumor. Some cancer cells may also develop the ability to penetrate the circulatory and/or lymphatic system and migrate to other parts of the body. There they will continue to proliferate and generate further tumors called metastases.[4,5] If left untreated, cancer can lead to severe illness and eventually death.

Cancer is currently a leading cause of mortality, being responsible for approximately 1 in 6 deaths worldwide in 2012.[6,7] By 2030 this number is expected to increase by more than 50%. Moreover, the global economic burden for cancer treatment and research was estimated at 1.16 trillion US dollars in 2010.[7] This sum is estimated to increase to approximately 2.5 trillion US dollars if long-term costs to the patient are added.

The current standard of care for cancer consists of surgery, radiotherapy, chemotherapy or their combination.[2,8] Depending on location and staging, surgical resection of the primary tumor is usually the first course of action in cancer treatment, aiming to achieve local control of the disease.[9] In order to downstage the primary tumor, prevent metastatic disease and/or reduce the risk of disease recurrence, (neo)adjuvant chemotherapy and/or radiotherapy may also be delivered before and/or after surgery.[10] Radiotherapy, as a treatment option, uses ionizing radiation in order to damage the DNA in the cellular nuclei, preventing further division.[11–13] Similar to surgery, radiotherapy also aims to achieve local tumor control being, however, usually more effective in treating local micro-invasions of the tumor. In case the cancer is hematologic or has become metastatic, systemic treatment via chemotherapy is usually employed.[14,15] Chemotherapy consists in the delivery of a combination of cytostatic drugs with the purpose of inhibiting cell growth.

As mentioned, the majority of cancer patients have to undergo open surgery for the resection of the primary tumor, which typically involves pain and long recovery times for the patient. Moreover, in some cases it may cause the suppression of the patient's immune system, increasing the risk of perioperative metastatic tumor dissemination.[16,17] For this reason, ever since the late 1980s there has been a constant tendency to migrate from open surgery towards minimally or non-invasive techniques.[18,19] These are aimed towards reducing patient recovery time and morbidity and increasing healthy tissue preservation without compromising survival. Such techniques are particularly effective for well localized tumors, most of the time ensuring the complete removal of the cancerous tissue.[20–25] While at a first glance this may sound as a severely limiting aspect, the number of cases in which a tumor is discovered before it has

a chance to spread is constantly increasing. Contributing factors to this are both advancements in medical imaging technologies and also the growing number of medical image screenings a person undergoes for non-cancer diagnostic purposes. The latter in particular, has led to an increase in incidental tumor findings at an early stage.[26–29]
Two treatment options for local tumor control have the potential to achieve complete non-invasiveness: treatment by high intensity focused ultrasound[30,31] and external beam radiotherapy.[32,33] These treatment modalities rely on the deposition of acoustic and electromagnetic radiation, respectively, deep inside the human body, without the need of incisions. This implies major benefits for the patients such as shortened stay at the hospital, lower morbidity, decreased healthcare costs and the ability to treat cancer in non-surgical candidates. The current thesis revolves around aspects and problematics related to these two treatment methods, reason for which they are detailed in the following sections.

1.2 High Intensity Focused Ultrasound

The term ultrasound (US) is used to define any sound which has a frequency above the human audible range ($>$ 20 kHz).[34] The existence of US was hypothesized ever since 1794 when Lazzaro Spallanzani demonstrated that bats hunted and navigated by inaudible sound instead of vision.[35] One of the first human-made devices capable of producing US was developed by Francis Galton in 1893. During his studies on the human and animal hearing range he created the Galton whistle (more commonly known as the dog whistle).[36] A major breakthrough in US generation occurred with the discovery of the direct and converse piezoelectric effect in the 1880s.[37] It was observed that certain solid materials produce electric charge when they underwent mechanical stress. The converse was also demonstrated to be true: the application of an electric current creates mechanical vibrations in the material. To this day piezoelectric materials are still predominantly used for US generation.
As US passes through matter, different interactions take place. For example, while crossing boundaries between materials with different densities, part of the US wave is reflected back.[38] The reflections can be detected by a piezoelectric material, transformed into electric signals and subsequently into an US image.[39] However, long before its use for imaging became popular, US was being investigated for its biological effects. While developing a new technique for submarine detection in 1917, french physicist Paul Langevin noticed that operating a powerful US source in a water tank lead to the death of all the fish inside the tank.[40] Moreover, people reported pain when placing their hand in the operation area of the US source. A few years later, Wood and Loomis concluded that indeed high intensity US had biological effects,[41] which led to further investigations into using US for therapeutic purposes. Lynn and Putnam in 1942 were the first to use a beam of focused US in order to thermally

ablate tissues in cortical and subcortical regions of 37 animals.[42] The underlying physical principle behind US thermal ablations is that, as US passes through a material, it is partially absorbed and transformed into heat.[43] For biological tissues in particular, by focusing the US beam in a small region, the temperature of the tissue can be increased to the point that coagulative necrosis occurs.[44] Following the preliminary study of Lynn and Putnam, the principle of thermal ablations using high intensity focused ultrasound (HIFU) was translated into clinics by William and Francis Fry.[45] A 4-element HIFU transducer (see Fig. 1.1(a)) was used as an alternative to open neuro-surgery. A difficulty during treatment delivery, which extended for the following three decades, was the lack of a robust image guidance system. This led to a gradually diminishing interest in medical HIFU.[46] It was only in the 1980s, when B-mode US imaging became widely available, that the potential of HIFU for non-invasive surgery continued to be explored.[47] This paved the way to the HIFU therapeutic systems that are being used today (see Fig. 1.1(b)).

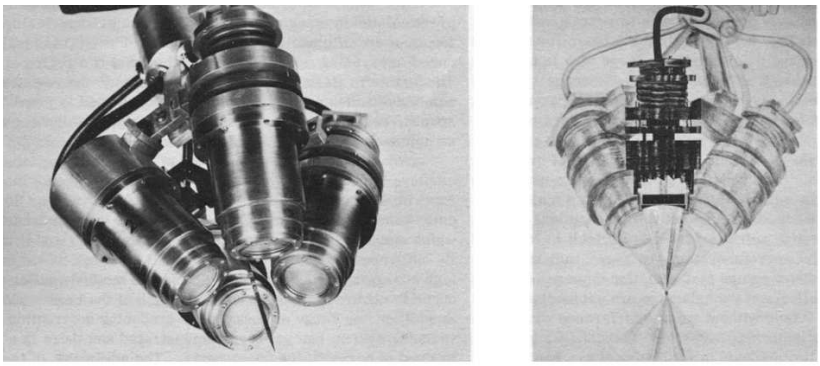
In the past two and a half decades, HIFU thermal ablations have been used to treat several pathologies including uterine fibroids,[49] prostate cancer,[50] liver tumors,[51,52] thyroid nodules,[53] palliation of bone metastases,[54] breast fibroadenoma[55] and more recently tremor-dominant Parkinson's disease.[56] Some of these advancements, however, would have probably not been as successful without the existence of a non-invasive manner to reliably monitor tissue temperature during therapy. A corner-stone in this sense was set by Hynynen an coworkers in 1993, who demonstrated the feasibility of conducting HIFU therapies under magnetic resonance imaging (MRI) guidance.[57] This idea was followed-up by Ishihara *et al*, which showed that, the MR scanner can also be used for non-invasive tissue temperature measurement.[58] Further technical details on how this can be achieved are discussed in the following section.

1.2.1 Proton Resonance Frequency Shift (PRFS) Thermometry

In order to maintain the non-invasive nature of HIFU therapies, a temperature measurement mechanism which does not require sensor implantation is necessary. It was observed by Hindman *et al*[59] that the resonance frequency of a hydrogen proton in water molecules, varies with temperature via the following law:

$$\omega = \gamma(1 - s(T))B_0 \tag{1.1}$$

where ω is the resonance frequency of the hydrogen nucleus, $s(T)$ is the electron screening constant and B_0 is the local magnetic field strength the hydrogen proton experiences. The underlying physical phenomenon of the dependency illustrated in Eq. 1.1 is related to a stronger electronic screening of the hydrogen nucleus in free water molecules compared to bonded water molecules.[59,60] As the water temperature increases, the hydrogen bonds between water molecules



(a)

(b)

Figure 1.1: (a) The 4-element HIFU transducer developed by William and Francis Fry (reprinted from Fry[48] with permission). (b) A modern day hemispherical HIFU array transducer.

start to stretch, bend and break, leading on average to a higher concentration of free water molecules.[61] This produces in turn, stronger screening effects (i.e. a larger $s(T)$ in Eq. 1.1) and therefore a lower proton resonance frequency. Moreover, for the -15 and $100°C$ range, there is a linear relationship between the value of s and temperature:[59]

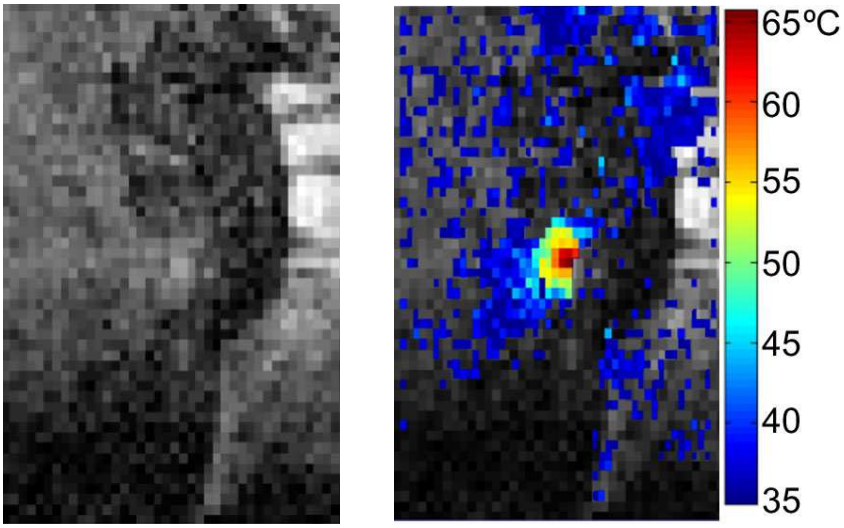
$$s(T) = \alpha \cdot T \tag{1.2}$$

with $\alpha = -1.03 \pm 0.02 \times 10^8\,°\mathrm{C}^{-1}$. Ishihara *et al* exploited the dependencies in Eq. 1.1 and 1.2 in order to determine temperature variations in aqueous materials via measurements of the hydrogen proton resonance frequency shift (PRFS).[58] By making use of a gradient-recalled echo MRI acquisition sequence,[62] temperature variations relative to a baseline were computed in the

following manner:

$$\begin{aligned}\Delta T &= T_{\text{heating}} - T_{\text{base}} \\ &= \frac{\omega_{\text{heating}} - \omega_{\text{base}}}{-\gamma \alpha B_0} \\ &= \frac{\phi(T_{\text{heating}}) - \phi(T_{\text{base}})}{-\gamma \alpha B_0 TE}\end{aligned} \tag{1.3}$$

where $\phi(T_{\text{base}}$ and $\phi(T_{\text{heating}})$ are the phase of the complex MR signal at baseline temperature and during heating, respectively, and TE is the echo time of the gradient-echo MR acquisition sequence. The rest of the terms are either self-explanatory of have already been defined in Eq. 1.1 and 1.2. Fig. 1.2 showcases an example of a PRFS temperature map obtained during a HIFU ablation on a porcine kidney.



(a) (b)

Figure 1.2: (a) Magnitude of a gradient-recalled echo MR image acquired on a kidney during a HIFU ablation. (b) The corresponding PRFS thermal map overlaid on the magnitude image in (a).

1.2.2 Thermal dose

The temperature information provided by the MR temperature maps, can be used in two ways in order to assess thermal damage. First, it was demonstrated that any tissue which is maintained for a few seconds at a temperature above

$50 - 60°C$, can be considered necrotic.[63–65] Since therapeutic response can also be obtained at lower temperatures, a more general metric for quantifying thermal damage has been proposed by Separetto and Dewey: the thermal dose.[44] Mathematically, the thermal dose is the time integral of a non-linear temperature-dependent term and is computed as follows:

$$TD = \int_0^t R^{T(\tau)-43} d\tau \tag{1.4}$$

where $T(\tau)$ is the temperature at time instant τ and with R = 2 for temperatures below 43°C and R = 4 for temperatures above 43°C. The measurement unit for thermal dose is equivalent minutes at 43°C (CEM_{43}), with a thermal dose of 240 CEM_{43} being the threshold for lethal thermal damage.[44] This value, however, may vary depending on the tissue type and the study reporting it.[66]

1.3 External Beam Radiotherapy

While the applicative context of this thesis mainly revolves around HIFU therapies, many of the concepts can be extended to external beam radiotherapy (EBRT) treatments. For this reason, the current section provides a brief summary on the topic of EBRT. Historically, radiotherapy has been used as a cancer treatment method immediately after Wilhem Röntgen discovered the X-Rays in 1895.[67] Early experimenters observed that prolonged exposure of cancerous tissues to X-Rays can lead to remission of the disease and pain relief. However, due to the serious side-effects of X-Ray therapy at that time, interest in such treatments diminished rapidly.[68] It was only after the biological effects of Radium were observed by Becquerell and Marie Currie in the early 1900s that therapy by radiation regained popularity.[69] Although it was a promising treatment modality, until the 1960s radiotherapy was mostly used to treat superficial tumors.[70] This is due to the fact that the early X-Ray and Radium treatment machines were unable to produce high enough energies to treat tumors deep inside the human body. This, however, became possible with the creation of the first linear accelerators (Linacs), which were able to produce energies in the megavoltage range.[71] Today, radiotherapy is a standard cancer treatment option, undergone by more than 50% of the patients.[12]

Radiotherapy relies on the biological effects induced by ionizing radiation on living cells.[72] As radiation traverses a layer of cells, it damages the nuclear DNA either though direct or indirect mechanisms. Direct mechanisms occur when the radiation ionizes the DNA directly, breaking one or both of the sugar phosphate backbones or breaking the base pairs of the DNA. Double strand breaks are considered to be the most effective in inducing cellular death, since they are the most difficult to repair. On the other hand, double strand breaks can also lead to healthy cell mutation, increasing the risk of radiation-induced secondary cancer.[73] Indirect damage to the DNA is caused by the creation of

H_2O^+ and H_2O^- ion radicals, following the interaction between the ionizing radiation and the water molecules inside the cells. These radicals are highly unstable and rapidly dissociate into free hydrogen and hydroxyl radicals. Due to their high reactivity, the latter may end-up re-combining into toxic compounds such as hydrogen peroxide, altering cell function or inducing cell death. Naturally, as radiation traverses the human body, it interacts with both cancerous and healthy cells. Not all cells, however, have the same radiosensitivity. Actively reproducing cells in particular are known to have a higher radiosensitivity than dormant cells.[74] Thus, the fact that cancer cells are dividing at a fast rate, render them vulnerable to radiation. Fast growth is, however, not a property exclusive to cancer cells. White blood cells and bone marrow, for example, have a high regeneration rate, making them sensitive to radiation.[75] Therefore, targeting cancer cells with radiation does not rely exclusively on their fast growth. The reduced expression of the DNA repair genes in the cancer cells also plays an important role. This implies that cancer cells have more difficulty repairing DNA damage than normal healthy cells.[76] For this reason, the total prescribed radiation dose is usually divided into several fractions, each aiming to deliver a sub-lethal amount of radiation.[77] In between fractions, healthy cells are able to repair DNA damage at a higher rate than cancerous cells. Since the effect of radiation is cumulative, cancer cells will eventually accumulate a deadly amount of radiation damage, with the healthy cells having higher survivability due to their more effective repair mechanisms.

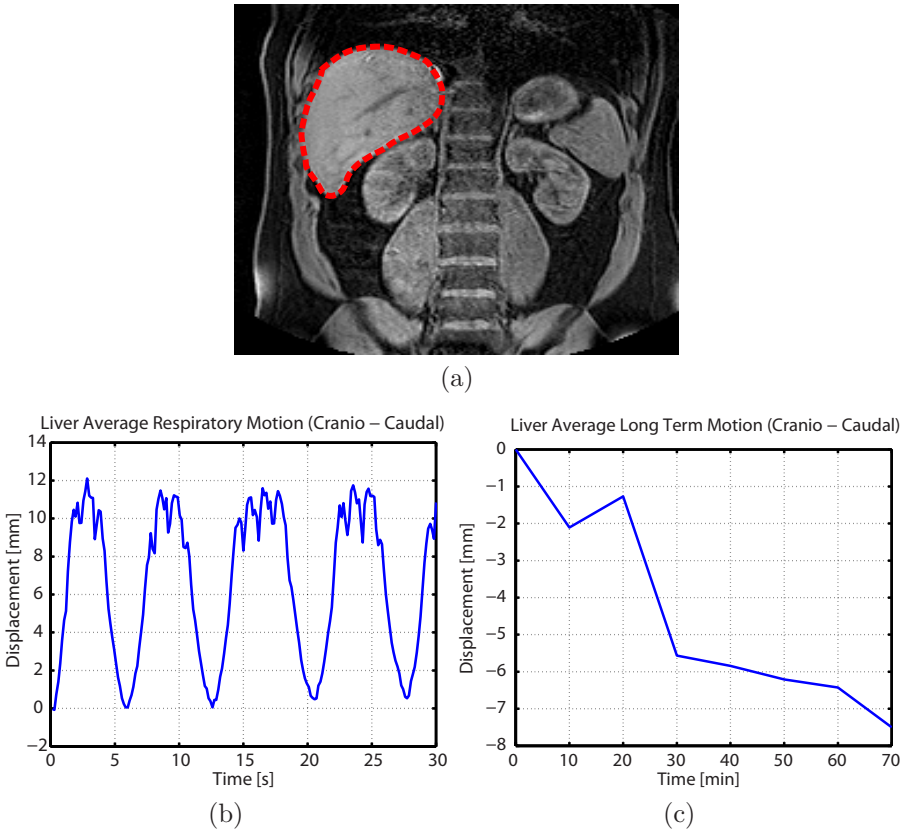
Concerning delivery techniques, radiotherapy can be internal or external. One option to deliver internal radiotherapy involves the implantation of several radioactive seeds inside the tumor and/or in its immediate proximity.[78] This is also known as brachytherapy and is currently performed for several anatomical locations including prostate, uterus, cervix, breast and head-and-neck. A second internal radiotherapy delivery method consists in the administration of radioactive drugs either orally or intravenously.[79] These drugs the travel within the bloodstream and have a systemic effect. This type of treatment is predominantly used for thyroid cancer and imposes strict safety rules for the patient. Since the radiation has a systemic effect, there is a constant risk of the patient contaminating people they come in contact with. External delivery methods or external beam radiotherapy (EBRT) uses a radiation source situated outside of the body.[32,33] For EBRT, radiation is generated either by radioactive metals[80] or by a Linac,[71] with the latter being the more common treatment machine. As a cancer treatment option EBRT is particularly attractive since, similar to HIFU, it is completely non-invasive, with the patient usually being discharged from the hospital immediately after the delivery of a radiation fraction. However, unlike HIFU where ablative energy doses can be safely delivered in a single fraction, such radiation doses in EBRT imply unacceptable toxicity for the organs-at-risk. Therefore, in EBRT it is difficult to assess acute therapeutic response, with the endpoint validation typically being made by term observation of the biological effects (i.e. changes in tumor size and/or physiological parameters).[81–83]

1.4 Tumor motion: A challenge during HIFU and EBRT treatments

While HIFU and EBRT have the advantage of being completely non-invasive, their application in a clinical setting can become challenging due to displacements of the cancer relative to the therapeutic beam.[84,85] Depending on the nature of the displacement, the treated tissue may periodically, partially or permanently move-out of the beam-path. Naturally, this will lead to the therapeutic energy being diverted from the targeted anatomy, increasing the risk of delivering a sub-lethal amount of thermal or radiation dose.[86,87] This can implicitly lead to residual viable pathological tissue and disease re-occurrence. Moreover, in case the displacement is severe, healthy tissues and/or vital structures may move into the beam path, potentially leading to collateral damage and in some instances to serious complications.

Tumor motion during HIFU and EBRT treatments may stem from several sources and can affect therapy at different time-scales. Uncertainties originating from patient positioning errors, while not being related to patient motion per se, it can have the same effect on therapy outcome.[88] As the name suggests, these can occur due to a mismatch between the patient position during treatment planning and at the time of the therapeutic energy delivery. Naturally, if the therapeutic beam would be oriented towards the location of the pathology as established during planning, a partial or a complete miss of the target is highly probable. A second category of displacements can stem from the human body's intrinsic physiological motion. This broadly includes three types: spontaneous, long-term and respiratory.[85] Spontaneous motion is particularly problematic for HIFU and EBRT therapies due to its fast and infrequent nature. It can be voluntary, such as the patient changing position due to discomfort, or involuntary such as coughing or skeletal muscle spasms. Long-term motion is mostly prevalent for organs in the abdomen and lower thorax and stems from digestive activity and/or metabolic processes (e.g. bladder filling). Such displacements manifest themselves as slow temporal drifts, becoming significant over time intervals of several minutes.[89] Finally, respiratory motion induces rapid quasi-periodic displacements with a frequency of 0.2 - 0.3 Hz, affecting to different extents the organs in the thorax and the upper abdomen.[90,91] Fig. 1.3 showcases an example of liver displacements under the effect of both respiratory and long term motion. Besides the extent of the displacements, the figure also illustrates the different time-scales at which the two types of motion become significant.

The outcome and prospect of HIFU and EBRT therapies may, therefore, be considerably improved if motion over the duration of the treatment is taken into consideration. The following sections briefly summarize potential solutions for motion estimation and/or compensation for HIFU and EBRT.



(a) (b) (c)

Figure 1.3: An example of liver displacements under the effect of respiratory and long term motion. (a) Coronal slice of a 3D MR image acquired on the abdomen of a healthy volunteer. The liver is indicated by the red dashed contour. (b) Average displacement of the liver in the cranio-caudal direction under the effect of respiration. (c) Long-term motion of the liver in the cranio-caudal direction.

1.4.1 Motion prevention or reduction in HIFU and EBRT

Depending on the nature of the motion that is being addressed, there are several methods and techniques that can be used to either prevent or reduce the extent of tumor displacements during HIFU and EBRT. Patient spontaneous motion, for example, can be prevented by body fixation via (custom-made) casts, molds and/or masks.[92] Even so, in particular for lengthy therapy sessions (some studies report HIFU therapies of up to 8 hours[93]), the patient may still end-up moving due to the discomfort of resting in the same position for an extended period of time. In such cases, patient sedation is usually necessary.[94,95] Depending on the degree of sedation (light, deep or general anes-

thesia), this can address both voluntary and involuntary spontaneous motion events. However, while sedation is practical for HIFU therapies, since they are usually delivered in a single session, for EBRT this is somewhat unfeasible. The EBRT radiation fractions may be delivered hours, days or weeks apart. Naturally, sedation would have to be performed prior to each fraction. Aside from the possible side-effects of patient sedation with such a high-frequency, this would also imply high financial and personnel demands, since an anesthesia team would have to be present each time.

Long term motion due to digestive and peristaltic activity can be mediated by the administration of digestive motility inhibitors such as butylscopolamine or glucagon.[96,97] However, these drugs have an effective time of 30 - 40 mins, which may be insufficient for long therapy sessions. Moreover, they may induce side-effects in the form of changes in bowel movements well after the end of the therapeutic energy delivery. An alternative or complementary to medication, the diet of the patient could be adjusted prior to the intervention and/or bowel and bladder preparation protocols could be established.[98,99] Naturally, this requires patient cooperation and may imply a certain degree of discomfort. Also, even in the presence of bowel and bladder preparation protocols, reproducibility of bowel and bladder filling is difficult to attain, leading to uncertainty in the location of the surrounding anatomy.[100,101]

A wide-spread solution for reducing the extent of respiration-induced motion during HIFU and EBRT energy deliveries are the so-called breath-hold techniques.[90] This implies that, as the name suggests, while the therapeutic beam is turned-on, the patient's breathing is halted. Breath-hold techniques can be implemented using several approaches, including: 1) Deep inspiration breath-hold - The patient voluntarily holds their breath at deep inspiration;[102,103] 2) Active breath control - A digital spirometer is used to automatically hold the patient's breath at a pre-established phase of the respiratory cycle;[104] 3) Self-held breath-hold without respiratory monitoring - The patient holds their breath at a self-chosen phase of the respiratory cycle, having partial control over the operation of the therapeutic beam;[105] 4) Self-held breath-hold with respiratory monitoring - Same as 3), however, the patient holds their breath during a specific phase of the respiratory cycle.[106] An alternative to complete breath-holds, which can be cumbersome for the patient, is induced shallow breathing. The aim of this technique is rather to reduce the extent of respiratory-induced motion, instead of complete stabilization. This can achieved by applying pressure on the patient's abdomen using pressure plates or corsets, limiting diaphragmatic excursions.[107,108] However, both breath-holding and induced shallow breathing imply discomfort to the patient, to the extent that in some cases it is completely unfeasible due to pain or inability for breath-holding.

1.4.2 Motion estimation techniques in HIFU and EBRT

Since most of the previously described motion prevention and reduction methods are typically cumbersome for the patient and/or have limited capabilities, there have been consistent efforts over the years to develop automatic tumor tracking solutions, followed by a corresponding adaptation of the therapy. The nature of these solutions depends to a large extent on the type of displacement that is being addressed: patient positioning errors, spontaneous motion, long term drifts and/or respiration. At the same time, the adopted approach for motion estimation also depends on whether on-board medical imaging (such as US, CT and/or MRI) is available during treatment. In the following, motion estimation solutions in both the absence and in the presence of image guidance are discussed individually.

1.4.2.1 Motion estimation methods in the absence of medical image guidance

In order to estimate patient positioning inaccuracies prior to the HIFU or EBRT therapeutic sessions, skin markings are typically made during treatment planning at representative locations over the patient's body.[109,110] The coordinates/location of these markings are then recorded relative to a static reference frame provided by a laser positioning system and/or by fixed positioning elements in the treatment table. At the time of treatment, the location of the skin markings is again determined with respect to the static reference frame. In this manner, an estimation of the patient positioning discrepancies with respect to planning can be obtained. Estimating positioning errors in this manner, however, has only limited applicability. While it is effective for static anatomies, for deformable organs the markings are only surrogates for the tumor's true position, with a high probability of relative displacements between the tumor and the markings.[111] Moreover, once the patient has been positioned at the start of the treatment, there is no guarantee that they will maintain the same position over the course of the therapeutic energy delivery.
For the estimation of respiratory motion in particular, several studies suggest the use of indirect tracking methods via external surrogates. Such approaches rely on establishing a correlation model between surrogates such as reflective skin markers,[112] body surface variations,[113] respiratory belts[112] or nasal temperature[114] and the position of the pathology over the respiratory cycle. While such techniques are effective as first order methods, depending on the duration of the therapy, the correlation model may deteriorate over time, affecting the accuracy and precision of the motion estimation process.[115]
A different approach to tumor motion estimation, regardless of its nature, implies the placement of electromagnetic (EM) beacons inside the tumor and/or its close proximity.[116] After implantation, the position of the markers can be detected with sub-millimeter accuracy via a detection array placed next to the patient. Compared to the aforementioned indirect methods this is a more accu-

rate and precise approach, however, it implies a higher degree of invasiveness. Moreover, even the implanted electromagnetic makers are still only surrogates for the tumor position, with relative displacements between the markers and the tumor remaining a possibility.

1.4.2.2 Motion estimation methods in the presence of medical image guidance

The integration of on-board medical imaging systems into HIFU and EBRT treatment machines,[117,118] provides alternative motion estimation solutions to the ones described in section 1.4.2.1. The basic principle of image-based motion tracking relies on comparing two images, where one of the images provides a positional reference.[119] The result of the comparison is usually a set of displacements, indicating the relative motion between the images. Such comparisons are usually performed via image registration, which by definition is the process of aligning two images acquired from different perspectives, using different sensors and/or at different moments in time. Mathematically, the registration process is expressed by the following formulation:[120]

$$\hat{T} = \underset{T \in \mathcal{T}}{\operatorname{argmin}}\ S(I, J(T)) \tag{1.5}$$

where I and J are the images to be registered, S is a metric measuring the degree of dissimilarity between the two images and T is the displacement (or transformation) between the two images. In words, Eq. 1.5 would read: search for the set of displacements (or transformation) T, which maximizes the degree of similarity between the reference image I and the transformed image-to-register J. Medical image registration is still an open problem with a plethora of potential solutions being available in the literature.[119,121]

Several imaging modalities have been historically used for HIFU and EBRT guidance, with US, CT and MR being the most widely spread.[117,118] Compared to CT and MR, US-guidance is a relatively cheap alternative, while at the same time provides the possibility of high-framerate imaging. This is well suited for tracking fast motion events such as respiration. A HIFU system with US-guidance capabilities was, for example, designed by Seo *et al*, who integrated two imaging probes with a parabolic HIFU transducer.[122] The probes provide two orthogonal imaging planes intersecting at the focus of the transducer. This offered the possibility of tracking both in and out-of-plane respiratory motion for the tissue situated at the transducer focus. While it was demonstrated to be a promising motion tracking candidate, sound *in-vivo* validation is still lacking. The feasibility of US tracking was also demonstrated for EBRT by Lattanzi *et al*[123] and Fuss *et al*.[124] In both studies, US imaging was used to estimate daily patient positioning errors for abdominal pathologies. A re-occurring difficulty was establishing a correspondence between the coordinate systems of the US imaging probe and the treatment machine. A general problem with US

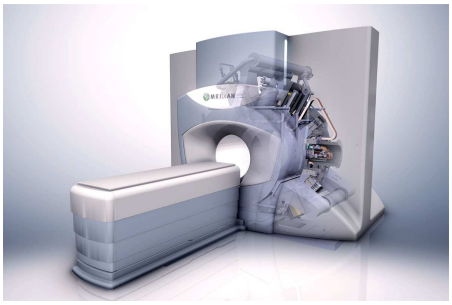
images is the typically low signal-to-noise (SNR) ratio, limited field-of-view, artifacts, poor contrast and orientation-dependent appearance.[125] This makes registration-based tracking particularly challenging for a broad category of existing methods.[126]

In particular for EBRT, a popular approach to image-based patient positioning relies on the acquisition of cone-beam CT (CBCT) images prior to the delivery of the radiation fraction and their subsequent comparison to the master CT image.[127] Current clinical practice, however, does not rely on image registration for the CBCT-CT alignment. Instead, high-Z fiducial markers are implanted in the tumor and its immediate vicinity.[128,129] Due to the high-Z of the marker material, these appear hyper-intense on both the CT and the CBCT image, facilitating the identification and alignment of the markers in the two images. Such an alignment, however, is limited to estimating only rigid displacements, which is sub-optimal for deformable organs.[130] In addition, the implantation of the markers is usually performed in the absence of anesthesia, which can make the procedure particularly uncomfortable for the patients. Therefore, a pure image-based motion estimation is desirable. Unfortunately, CBCT images typically have an intrinsic low contrast and are altered by artifacts, which makes a registration-based alignment difficult for existing approaches.[131]

Due to its high versatility for both anatomical and functional imaging, MRI is often the guidance method of choice for HIFU therapies in Europe and the United States.[132] MRI has also gradually gained popularity for EBRT guidance, where important efforts have been made for a physical integration of MR scanners into EBRT treatment machines and resulted into the development of the MR-Linac.[133,134] An example of both MRg-HIFU and MRg-EBRT treatment machines are illustrated in Fig. 1.4. The diverse MR pulse sequences allow the generation of a large variety of soft tissue contrasts, while at the same time it offers the possibility of measuring several physiological parameters (such as temperature, diffusion, perfusion, etc.)[135] For this reason, it can be employed during most stages of the therapy including planning, positioning, therapy delivery and validation of the therapeutic endpoint. Concerning MR-based motion tracking, most of the studies so far have focused on addressing respiratory motion during the therapeutic energy deliveries.[90,91] A simple, yet effective approach for respiratory motion estimation implies the use of so-called navigator echoes.[136–138] Navigator echoes are fast 1D MR acquisitions, with a field-of-view typically encompassing high-contrast organ boundaries. As the organ moves, this becomes apparent in the navigator temporal trace and, by correlating a current with a reference navigator echo, the displacement of the organ can be determined. Naturally, navigator echoes used on their own usually do not allow the estimation of complex deformations, however, they have proved to be particularly useful for gating energy deliveries for both HIFU and EBRT.[138–140] Aspects related to gating will be discussed in section 1.4.3. In order to estimate more complex deformations of abdominal organs due respiration, several studies propose coupling high-framerate 2D MR-imaging with real-time motion estimation algorithms. In this sense, Denis de Senneville *et al*

(a)

(b)



(c)

Figure 1.4: Examples of MRg-HIFU and MRg-EBRT treatment machines. (a) The Philips Sonalleve MR-HIFU system. (b) The Elekta Unity MR-Linac. (c) Concept picture of the MRIdian MR-Linac. Both systems illustrated in (a) and (b) are installed and fully operational at the University Medical Center Utrecht.

propose an atlas-based real-time motion estimation method, in which dynamic 2D MR imaging is performed during a pre-treatment phase, with each image being stored in an atlas together with a pixel-by-pixel motion field.[141] The motion field is determined relative to a reference 2D MR image by the means of the Horn & Schunck optical flow algorithm. During energy deliveries, rapid MR imaging is performed and an inter-correlation coefficient is computed between each acquired image and the pre-acquired atlas. The motion field associated to the image in the atlas which provides the highest correlation coefficient is then selected as the corresponding displacements for the current MR image. While the approach provides relatively good precision and accuracy it has a downside which is typically specific to atlas-based motion estimation methods: the displacements saved in the atlas together with the images are only surrogates and may not be representative for the current anatomical situation.[142]

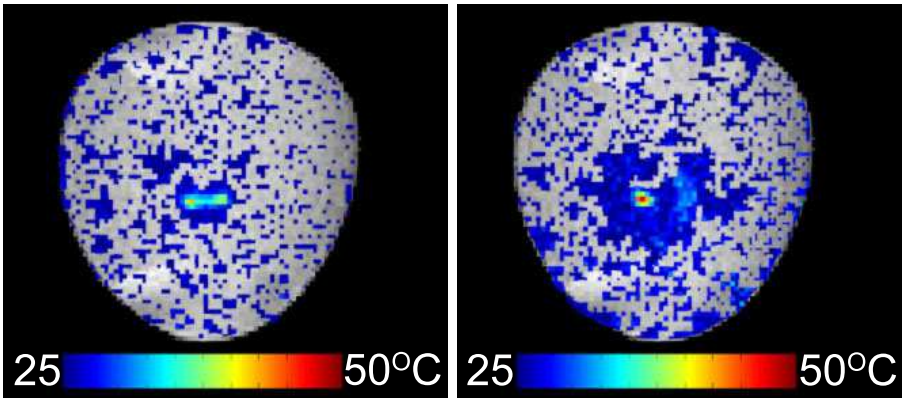
This drawback was addressed in a follow-up study carried-out by Roujol *et al*, who managed to perform an implementation of the Horn & Schunck optical flow algorithm with a latency of under 100 ms.[143] This allowed 2D MR-based real-time motion estimation, without the need of a pre-learned atlas. Naturally, there is a large variety of approaches, either relying on a pre-learned model or not, for real-time 2D MR-based respiratory motion estimation.[91,126] However, all of these approaches share the same drawback: organs generally move in three dimensions while the methods only estimate 2D displacements. Some studies address this issue by aligning the 2D MR imaging plane with the principal axis of respiratory displacements.[137,141,143,144] Even so, through-plane motion still remains a possibility. A 2.5 D approach has been proposed by Ries *et al* where high-framerate 2D MR imaging was used for estimating in-plane displacements, while through plane-motion was addressed via a 1D pencil-beam navigator.[87] Coupled with a motion prediction algorithm, real-time respiratory motion compensation for MRg-HIFU therapies was achieved, although the concept can be easily extended towards MRg-EBRT interventions. Migrating to full 3D real-time MR-based respiratory motion estimation still remains mostly an unsolved problem. This is due to the fact that the acquisition of MR images require a constant trade-off between SNR, spatial resolution and temporal resolution.[135] Since estimating respiratory motion typically demands a high temporal resolution, SNR and/or spatial resolution would have to be compromised. The precision and accuracy of existing image-based motion estimation algorithms generally depends on the SNR and spatial resolution of the images.[145] Therefore, there is a lower SNR and spatial resolution bound before the motion estimation algorithms start to become unreliable. Such limits were explored in a study performed by Glitzner *et al*,[146] where it was demonstrated that for 3D MR images with a voxel size of $5 \times 5 \times 5\text{mm}^3$, dense deformations can still be reliably estimated. However, the achieved temporal resolution was 1 image/s, which usually results in an undersampling of the respiratory cycle. Other ongoing studies for real-time 3D MR-based respiratory motion estimation, follow an approach similar to the one proposed by Denis de Senneville *et al*: a 2D MR image acquired during energy deliveries is correlated to an atlas of 3D volumes pre-acquired during a learning phase.[147] The 3D volumes are again stored together with their corresponding displacements. The motion associated to the volume which provides the highest correlation with the interventional 2D MR image is then selected and established as the current motion.

1.4.3 Motion compensation techniques for HIFU and EBRT

Once the extent of the tumor motion is estimated, measures can be taken in order to compensate for its effects on the therapy outcome. A standard approach for motion compensation in EBRT is to incorporate uncertainties related to the extent of the displacements into the therapeutic margins.[148] In this sense, the

radiation delivery plan is optimized in such a way that, despite tumor motion, the pathology receives 100% of the prescribed radiation dose. While this typically implies a higher radiation dose to the organs-at-risk, such an approach ensures that the therapeutic endpoint is achieved. Aside from including the potential displacements in the treatment margins, several other solutions have been proposed for motion compensation during both HIFU and EBRT treatments. Compensating for patient positioning errors, in particular, is relatively straight-forward. Once these are estimated, based on skin and/or implanted markers, the therapeutic table is adjusted such that the position of the markers at the time of treatment is roughly the same as during planning.[149, 150] Compensating for spontaneous, long-term and/or respiratory motion, on the other hand, is typically addressed via two major approaches: gating or beam-steering. Gating implies the activation of the therapeutic beam only when the targeted organ/pathology is in a pre-defined rage of positions called gating window.[151, 152] The latter is usually defined at locations where the organs are relatively stable for longer periods of time (such as the expiration plateau). Confirmation whether the organ is situated withing the gating window can be obtained by any of the previously discussed motion estimation methods. Even so, the use of pencil-beam navigators, respiratory belts or even optical tracking, in particular, has been identified to be the most practical.[139, 152, 153] The first clinical usage of gating during HIFU treatment of a liver cancer patient has been employed by Okada *et al*.[51] However, instead of using the aforementioned approaches for determining whether the organ was inside the gating window, mechanical ventilation with induced apnea was used instead. While this ensures that the organ is more stable, it involves higher costs and longer recovery times, since the patient was put under general anesthesia. In a more recent study performed by Wijlemans and coworkers it was demonstrated that respiratory-gated HIFU is clinically feasible under Procedural Sedation and Analgesia, instead of general anesthesia, reducing the risk of complications and shortening patient recovery times.[154] The downside of using gating during HIFU and EBRT is that it generally reduces the duty cycle of the therapeutic beam, which implicitly increases the total treatment delivery time.[122, 152, 155] This is especially problematic for HIFU therapies in organs such as the liver and kidneys, where the high perfusion effects can cause strong heat dissipation effects.[156] Thus, continuous operation of the therapeutic beam would be preferable over gating. By steering the beam with respect to the estimated motion, localized energy deposition can be achieved, without losses in terms of the duty cycle. HIFU beam steering has been implemented by Seo *et al* by moving the transducer itself using visual servoing in a three-axis robot.[122] Other studies which employed phased HIFU transducers[87, 157, 158] used electronic beam-steering, by varying the transmit delays of the transducer channels, to dynamically re-adjusting the focus position. Compared to moving the transducer itself, this will naturally put less stress on the system, especially for compensating rapidly evolving motion such as respiration. Fig. 1.5 showcases the effect of electronic steering on the rise in temperature during a sonication on

a phantom undergoing sinusoidal motion. It can be observed that, when steering is enabled (Fig. 1.5(b)), a considerable improvement in energy focusing is attained compared to when steering is disabled (Fig. 1.5(a)). For EBRT on the



(a) (b)

Figure 1.5: Effect of electronic beam steering on the temperature rise during a sonication on a phantom undergoing sinusoidal motion. (a) MR temperature map in the absence of steering. (b) MR temperature map with electronic steering enabled. The temperature maps were overlaid on their corresponding magnitude images.

other hand, modulation of the radiation beam with respect to the estimated motion can be achieved by the usage of a multi-leaf collimator (MLC).[159] An MLC usually consists in an ensemble of interleaved tungsten segments, which can individually move in and out of the beam path. It was demonstrated that motion compensation during EBRT is achievable, by morphing the aperture of the MLC according to the estimated tumor displacements.[160,161] Problematic when performing either HIFU or EBRT beam steering, is the delay between the actual displacement of the tumor and the corresponding adjustments made to the therapeutic beam.[85,87,90,162] This delay incorporates the image acquisition time (in case image guidance is used), the latency of the motion estimation algorithm, the duration of the data transfer through the system and the inertia of the beam steering mechanism. Naturally, the larger this delay is, the higher the chance of missing the targeted pathology. For this reason, most studies which propose beam steering solutions for HIFU and/or EBRT also include a motion prediction algorithm, in order to compensate for the effects of such latencies.[87,163–165]

1.5 Thesis outline

The thesis can be broadly divided into three parts, which partially overlap with each other, both from a content and from a chronological point-of-view. The first part, encompassing **chapters 2** and **3**, connects to the state-of-the-art methodology dedicated to real-time MR - guidance of HIFU therapies, identifies thereby specific shortcomings and proposes strategies to overcome these limitations. **Chapter 2** investigates a methodological improvement to the Horn & Schunck registration algorithm, which has been proposed by several independent studies as a feasible solution to address elastic deformation for real-time respiratory motion estimation during MRg-HIFU therapies in moving organs. While the method demonstrated an overall good performance, it also suffers from specific drawbacks which are rooted in its fundamental functioning principle. The Horn & Schunck registration algorithm relies on the assumption that as a voxel moves, it conserves its gray-level intensity. This becomes problematic during HIFU therapies, where the local increase in tissue temperature leads to higher temperature gradients in the proximity of the focal spot. In turn, during the thermal build-up, this gradually increases the T_2^* relaxation time of the tissue, leading to a transient gray-level intensity drop in the MR-thermometry magnitude images, which play a secondary role as input to the motion estimation process. This gradual intensity drop locally violates the gray-level conservation hypothesis made by the Horn & Schunck algorithm, thus becoming a source of misregistration. A similar scenario occurs if therapy is conducted in the proximity of pulsating arteries. The periodic blood in-flow leads to local gray-level intensity variations, which may again impact registration quality. **Chapter 2** proposes an improvement to the Horn & Schunck algorithm, which robustifies it against such effects. By construction, local gray-level intensity variations only have a linear contribution on the motion estimates provided by the proposed method as opposed to the quadratic impact they have on the Horn & Schunck motion estimates. The proposed method was validated *in-vivo* for real-time MR-based respiratory motion estimation. As a test case, its resilience to in-flow artifacts introduced by pulsating arteries was comparatively evaluated relative to the Horn & Schunck algorithm. Evaluation criteria included both quality of the motion estimates and computational latency. Potential extensions to other application scenarios are also discussed. Motion tracking is, however, a secondary role of fast MR-imaging during HIFU therapies. The primary role of these images is in general MR-thermometry. Problematic in this sense is that fast MR-thermometry typically suffers from low SNR of the images, which in-turn leads to noisy/imprecise temperature measurements. This has a direct impact on therapy control, monitoring and therapeutic endpoint detection. State-of-the art MR-thermometry denoising methods focus on improving the precision of the temperature measurements, often with disregard towards accuracy. Due to their functioning principle, these approaches usually have a pronounced spatial smoothing effect on the MR-thermometry maps,

which can lead to a drop in measurement accuracy. **Chapter 3** investigates as denoising methodology for fast MR-thermometry, the use of a spatio-temporally adaptive filter which includes the maximization of both precision and accuracy in its formulation. The proposed approach was benchmarked against several state-of-the-art methods on test-benches which included simulations, *ex-vivo* and *in-vivo* experiments. The evaluation criteria were the attainable output precision and accuracy and the computational latency of the methods.
The second part of the thesis, which includes **chapter 4** and **5**, shifts focus from methodological advancements which operate on a small time-scale of an MRg-HIFU therapy, to the design and implementation of large-scale frameworks which aim to encompass the entire therapeutic work-flow. The building blocks of these frameworks consist of several state-of-the-art methods, which were fused together to ensure that the displacements of the treated area and its surroundings are accounted for over the entire duration of the therapy. Prior to the work conducted in **chapter 4**, studies have mostly addressed respiratory motion during HIFU energy deliveries. While this may be sufficient on a small time-scale, over the duration of lengthy HIFU therapies, the target and it's surroundings may also manifest significant long-term drifts due to peristaltic and metabolic activity. *Chapter 4* proposes a framework which is capable of compensating for such drifts, which mainly relies on the integration of 3D MR images during the cool-down interval in between sonications. The 3D MR images are then registered to a separate image acquired at the beginning of the therapy, which plays the role of positional reference. The framework was employed on data acquired on the abdomen of 10 healthy volunteers, in order to determine whether indeed such long-term drifts exceed acceptable therapeutic margins (which up to this point was just a hypothesis). In addition, the framework was also tested on *in-vivo* data acquired on a porcine liver, for a preliminary investigation of its capability for online therapy plan adaptation. Since promising results were obtained in **chapter 4**, the framework was further extended in **chapter 5**, such that it also includes respiratory motion compensation capabilities. Moreover, the extended framework allows relating *all* MR images acquired over the duration of the therapy, to a master 3D image acquired at the beginning of the intervention. In effect, with each acquired image, the position of the target anatomy and its surroundings can be determined relative to the master image. This allowed: 1) Online adaptation of the therapeutic plan with respect to the estimated displacements and 2) The projection of the thermal dose delivered by the individual sonications in the reference space of the planning image. This way, therapy progress can be monitored with respect to the therapeutic plan in a spatially consistent way. After offline validation on healthy volunteer data and online validation on a phantom, the framework was ultimately implemented on a clinical system and deployed during 3 *in-vivo* experiments conducted on a porcine liver.
Working on the design and implementation of a motion compensation framework encompassing the entire work-flow of an MRg-HIFU therapy, revealed two major shortcomings of the state-of-the-art: the dual role of the MR-thermometry

images in both motion tracking and temperature measurement and the scarcity of robust and fast cross-contrast/modality elastic registration algorithms. The first shortcoming originates from the need to optimize an MR-acquisition sequence for both real-time tracking and fast thermometry, which typically implies compromises in terms of spatio-temporal resolution, contrast weighting and signal-to-noise ratio. This leads to a sequence that is suitable but not ideal for either purpose. The second shortcoming originates from the lack of a robust and fast cross-modality elastic registration algorithm, which would significantly facilitate the simplification of the framework proposed in **chapter 5**. The latter was encumbered by additional MR scans required to obtain images of the exact same MR contrast weighting in order to be suitable for intensity-based optical flow algorithms for motion tracking. While the improved version of the Horn & Schunck algorithm proposed in **chapter 2** is indeed robust against transient gray-level intensities, it is still not to an extent that would allow cross-contrast registration. Motivated by this, effort was put into the implementation and validation of a novel cross-contrast/modality registration algorithm which suits the technical requirements of tracking deformable anatomies. The idea behind this novel algorithm together with its foundations were already set by Baudouin Denis de Senneville (one of the thesis co-promotors), during the final stages of validating the framework from **chapter 5**. Once the framework was validated, joint effort was put into implementing and validating this novel cross-contrast registration algorithm that we have called *EVolution*. The technical details, implementation and performance evaluation of the EVolution algorithm are the topic of **chapter 6** of the thesis.

The development of the EVolution algorithm bridges the transition between the second and the third part of the thesis. This third part focuses on addressing several limitations of the state-of-the-art image registration methods employed in image-based therapy guidance of both HIFU and EBRT. Naturally, the EVolution algorithm played an important role as a starting point in addressing these shortcomings.

At the same time that the work on the EVolution algorithm was completed, a HIFU system which allows simultaneous MR and US guidance was deployed at the UMC Utrecht. This provided the possibility to migrate the task of image-based motion tracking from the MR scanner to the US imager, which allows the MR scanner to be used exclusively for thermometry or other functional imaging. The B-mode US images provided by this integrated system contain, however, incoherent gray-level intensity variations under the form of speckle noise. This makes tracking via intensity-based registration methods, such as Horn & Schunck or its improved version from **chapter 2**, particularly challenging. The EVolution algorithm, on the other hand, should by construction be robust to such transient artifacts. **Chapter 7** of the thesis investigates this hypothesis in both *ex-vivo* and *in-vivo* scenarios, by employing the EVolution algorithm on B-mode US images acquired on this hybrid MR+USg-HIFU system. The precise application scenario was real-time US-based respiratory motion tracking during HIFU interventions in the abdomen. Since real-time

image registration of B-mode ultrasound images is also of general interest for other applications, within the same study, the tracking capability of the EVolution method on images provided by a contact US-imager was also investigated. At this point, since the identified limitations for interventional guidance during therapy delivery were to a large extent adequately addressed, the following chapters of the thesis focuses on motion-related inaccuracies which occur during the (re-)planning stage of cancer therapy. Since the connection between therapy planning and therapy delivery requires generally the image registration between different image contrasts and/or image modalities, these chapters were motivated by the possibilities of the cross-contrast co-registration methods developed in the scope of this thesis. As a starting point for this evaluation a frequently encountered problem for EBRT treatment was chosen: At the start of an EBRT treatment, a master CT is acquired on the patient, upon which the pathology and the organs-at-risk are delineated. Prior to each radiation fraction, a cone-beam CT (CBCT) image of the treated area is also acquired, in order to ensure that the tumor is approximately at the same location as during planning. The reason behind using CBCT images instead of CT is both due to faster acquisition times and, more importantly, due to the lower imaging-related radiation dose. The alignment between the "daily" CBCT and the planning CT currently relies on positional information provided by implanted high-Z markers in the tumor and/or its immediate vicinity. Besides involving discomfort for the patient, these markers are only surrogates of the true tumor location and can only account for rigid displacements. Automatic image-based elastic registration methods are thus preferred. The low beam energies and low number of X-Ray projections used for the reconstruction of CBCT images, lead to streaking artifacts and an a low signal-to-noise ratio. Since CT - CBCT registration is of interest, streaking artifacts are particularly problematic for intensity-based methods. This is due to the fact that such artifacts manifest themselves as intensity inconsistencies between the two image types. The EVolution algorithm is theoretically robust against such effects, therefore **chapter 8** investigates its capability for both CT - CBCT and CBCT - CBCT online registration. Of interest was also whether there is a dependency of the algorithm performance on the imaged anatomical site. For this reason, the method was evaluated on *in-vivo* data acquired on several patients and for several anatomical sites. As a concluding chapter of this thesis, **chapter 9** provides a critical analysis of the physiological plausibility of the deformations provided by the Horn & Schunck and the EVolution algorithm, both of which played a major role over the course of the thesis. This chapter was motivated by the substantial first-hand experience in applying Horn & Schunck and Evolution on registration problems of a multitude of image contrasts, anatomies and pathologies and a growing in-depth knowledge of their respective limitations in this role. Both algorithms are rooted in the domain of optical image processing and therefore they do not explicitly take into consideration the physical and physiological properties of the tracked anatomical tissues. *Chapter 9* proposes several quality assurance criteria which assist in assessing the physiological plausibility of the deformations

estimated not only by the Horn & Schunck and EVolution, but by deformable registration algorithms in general. In addition, variations of both these algorithms are proposed, which are specifically formulated for tracking biological tissues. The 4 registration methods are then comparatively evaluated using both existing and the proposed quality assurance criteria.
As a guideline meant to assist with the understanding of the thesis structure, Fig. 1.6 showcases its content and chronology in a graphical manner. The timeline is in accordance with the time interval during which the actual work was performed and only loosely related to the publication time of the associated manuscripts.

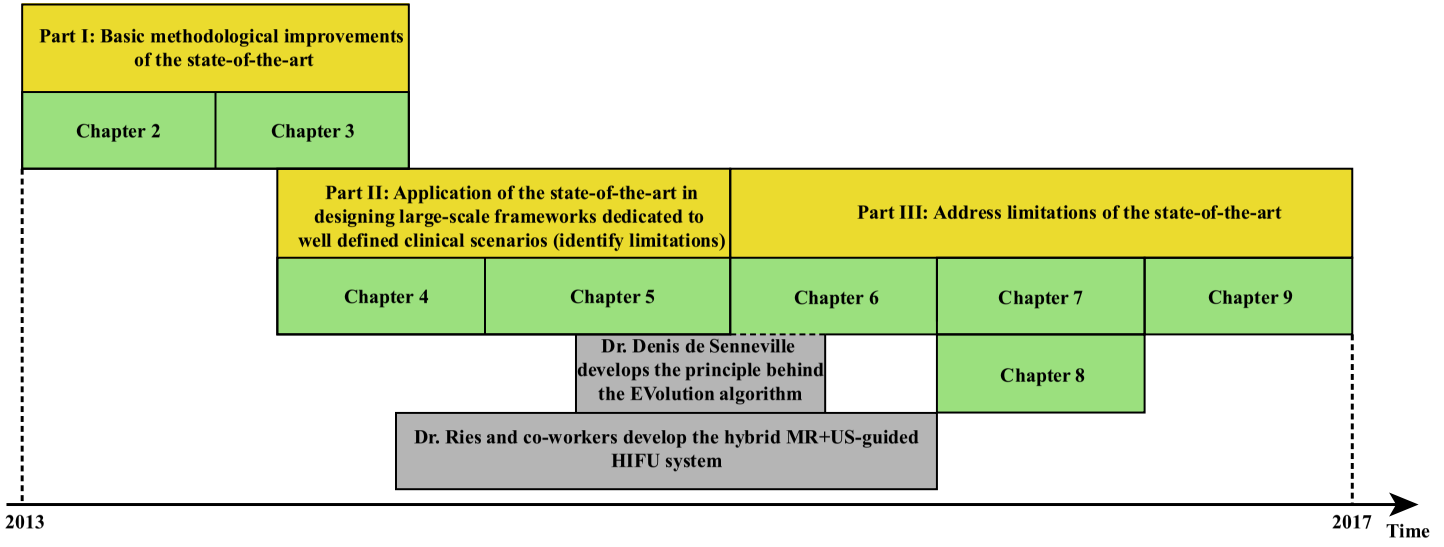


Figure 1.6: Thesis outline, chronology and motivation behind each chapter. Major advancements external to the thesis, which had an important role in its course are also mentioned (see gray boxes).

Bibliography

[1] R. A. Weinberg. *The Biology of Cancer*. Taylor & Francis Group, 2nd edition, 2013.

[2] C. R. Bast Jr., C. M. Croce, W. N. Hait, W. K. Hong, D. W. Kufe, M. Piccart-Gebhart, R. E. Pollock, R. R. Weichselbaum, H. Wang, and J. F. Holland. *Holland-Frei Cancer Medicine*. Wiley - Blackwell, 9th edition, 2008.

[3] D. Hanahan and R. A. Weinberg. Hallmarks of Cancer: The Next Generation. *Cell*, 144(5):646 – 674, 2011.

[4] A. C. Chiang and J. Massagué. Molecular Basis of Metastasis. *N Engl J Med*, 359(26):2814 – 2823, 2008.

[5] C. A. Klein. Cancer: The Metastasis Cascade. *Science*, 321(5897):1785 – 1787, 2008.

[6] J. Ferlay, I. Soerjomataram, M. Ervik, R. Dikshit, S. Eser, and C. Mathers et al. *GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11*. WHO Press, 2013.

[7] B. W. Stewart and C. P. Wild. *World Cancer Report 2014*. WHO Press, 2014.

[8] R. Govindan. *The Washington Manual of Oncology*. Lippinscott Williams & Wilkins, 2008.

[9] F. C. Wright, J. Escallon, M. Cukier, M. E. Tsang, and U. Hameed. *Surgical Oncology Manual*. Springer International Publishing, 2016.

[10] A. J. Neal and P. J. Hoskin. *Clinical Oncology*. CRC Press, 4th edition, 2009.

[11] J. Bernier, E. J. Hall, and A. Giaccia. Radiation Oncology: A Century of Achievements. *Nat Rev Cancer*, 4(9):737 – 747, 2004.

[12] G. Delaney, S. Jacob, C. Featherstone, and M. Barton. The Role of Radiotherapy in Cancer Treatment: Estimating Optimal Utilization from a Review of Evidence-Based Clinical Guidelines. *Cancer*, 104(6):1129 – 1137, 2005.

[13] R. Symonds, C. Deehan, C. Meredith, and J. Mills. *Walter and Miller's Textbook of Radiotherapy: Radiation Physics, Therapy and Oncology.* Churchill Livingstone, 7th edition, 2012.

[14] P. G. Corrie. Cytotoxic Chemotherapy: Clinical Aspects. *Medicine*, 36(1):24 – 28, 2008.

[15] R. Airley. *Cancer Chemotherapy: Basic Science to the Clinic.* Wiley - Blackwell, 2009.

[16] E. Neeman and S. Ben-Eliyahu. The Perioperative Period and Promotion of Cancer Metastasis: New Outlooks on Mediating Mechanisms and Immune Involvement. *Brain Behav Immun*, 30(Suppl):S32 – S40, 2013.

[17] S. Tohme, R. L. Simmons, and A. Tsung. Surgery for Cancer: A Trigger for Metastases. *Cancer Res*, 77(7):1548 – 1552, 2017.

[18] J. E. Wickham. The New Surgery. *Br Med J (Clin Res Ed)*, 295(6613):1581 – 1582, 1987.

[19] L. Wyld, R. A. Audisio, and G. J. Poston. The Evolution of Cancer Surgery and Future Perspectives. *Nat Rev Clin Oncol*, 12(2):115 – 124, 2015.

[20] K. Maeda, M. Maruta, T. Utsumi, H. Sato, K. Masumori, and Y. Koide. Minimally Invasive Transanal Surgery for Localized Rectal Carcinoid Tumors. *Tech Coloproctol*, 6(1):33 – 36, 2002.

[21] M. Noguchi. Minimally Invasive Surgery for Small Breast Cancer. *J Surg Oncol*, 84(2):94 – 101, 2003.

[22] K. M. Slawin, DiBlasio C. J., and M. W. Kattan. Minimally Invasive Therapy for Prostate Cancer: Use of Nomograms to Counsel Patients about the Choice and Probable Outcome of Therapy. *Rev Urol*, 6(Suppl 4):S3 – S8, 2004.

[23] J. K. Mullins, T. Feng, P. M. Pierorazio, H. D. Patel, E. S. Hyams, and M. E. Allaf. Comparative Analysis of Minimally Invasive Partial Nephrectomy Techniques in the Treatment of Localized Renal Tumors. *Urology*, 80(2):316 – 322, 2012.

[24] D. Li, J. Kang, B. J. Golas, V. W. Yeung, and D. C. Madoff. Minimally Invasive Local Therapies for Liver Cancer. *Cancer Biol Med*, 11(4):217 – 236, 2014.

[25] J. Fuchs, L. Schafbuch, M. Ebinger, J. F. Schäfer, G. Seitz, and S. W. Warmann. Minimally Invasive Surgery for Pediatric Tumors - Current State of the Art. *Front Pediatr*, 2(48):1 – 5, 2014.

[26] B. Lumbreras, L. Donat, and I. Hernàndez-Aguado. Incidental Findings in Imaging Diagnostic Tests: A Systematic Review. *Br J Radiol*, 83(988):276 – 289, 2010.

[27] P. Moyle, L. Sonoda, P. Britton, and R. Sinnatamby. Incidental Breast Lesions Detected on CT: What is Their Significance. *Br J Radiol*, 83(987):233 – 240, 2010.

[28] M. S. Al-Marhoon. Small Incidental Renal Masses in Adults. *Sultan Waboos Univ Med J*, 10(2):196 – 202, 2010.

[29] L. Morgan, H. Choi, M. Reid, A. Khawaja, and P. J. Mazzone. Frequency of Incidental Findings and Subsequent Evaluation in Low-Dose Computed Tomographic Scans for Lung Cancer Screening. *Ann Am Thorac Soc*, 14(9):1450 – 1456, 2017.

[30] J. E. Kennedy. High-Intensity Focused Ultrasound in the Treatment of Solid Tumors. *Nat Rev Cancer*, 5(4):321 – 327, 2005.

[31] Y.-S. Kim, H. Rhim, M. J. Choi, H. K. Lim, and D. Choi. High-Intensity Focused Ultrasound Therapy: An Overview for Radiologists. *Korean J Radiol*, 9(4):291 – 302, 2008.

[32] S. Webb. The Future of Photon External-Beam Radiotherapy: The Dream and the Reality. *Physica Medica*, 17:207 – 215, 2001.

[33] M. K. Bucci, A. Bevan, and M. Roach 3rd. Advances in Radiation Therapy: Conventional to 3D, to IMRT, to 4D, and Beyond. *CA Cancer J Clin*, 55(2):117 – 134, 2005.

[34] ANSI/ASA S1.1-2013. Acoustical Terminology. *Acoustical Society of America*, 2014.

[35] S. Dijkgraaf. Spallanzani's Unpublished Experiments on the Sensory Basis of Object Perceptio in Bats. *Isis*, 51(1):9 – 20, 1960.

[36] F. Galton. *Inquiries into Human Faculty and its Development*. Dent & Dutton (Everyman), 2nd edition, 1907.

[37] W. Heywang, K. Lubitz, and W. Wersing. *Piezoelectricity*. Springer, 2008.

[38] D. Ensminger and L. J. Bond. *Ultrasonics*. CRC Press, 3rd edition, 2011.

[39] A. J. Thorsen and G. E. Lakin. Basic Physics of Ultrasonography. *Seminars in Colon and Rectal Surgery*, 21(4):186 – 190, 2010.

[40] K. F. Graff. History of Ultrasonics. In Mason, W. P. and Thurston, R. N., editor, *Physical Acoustics: Principles and Methods*, volume 15, page 392. Academic Press, 1981.

[41] R. Wood and A. Loomis. Physical and Biological Effects of Highfrequency Sound Waves of Great Intensity. *Philos Mag*, 4:417 – 436, 1927.

[42] J. G. Lynn and T. J. Putnam. Histological and Cerebral Lesions Produced by Focused Ultrasound. *Am J Path*, 20:637 – 649, 1944.

[43] G. ter Haar and C. Coussios. High Intensity Focused Ultrasound: Physical Principles and Devices. *Int J Hyperthermia*, 23(2):89 – 104, 2007.

[44] S. A. Sapareto and W. C. Dewey. Thermal Dose Determination in Cancer Therapy. *Int J Radiat Oncol Biol Phys*, 10(6):787 – 800, 1984.

[45] W. J. Fry, W. H Mosberg Jr., J. W. Barnard, and F. J. Fry. Production of Focal Distructive Lesions in the Central Nervous System with Ultrasound. *J Neurosurg*, 11(5):471 – 478, 1954.

[46] J. Jagannathan, N. K. Sanghvi, L. A. Crum, C. P. Yen, R. Medel, A. S. Dumont, J. P. Sheehan, L. Steiner, F. Jolesz, and N. F. Kassel. High Intensity Focused Ultrasound Surgery (HIFU) of the Brain: A Historical Perspective, with Modern Applications. *Neurosurgery*, 64(2):201 – 211, 2009.

[47] D. Tyshlek, J. F. Aubry, G. ter Haar, A. Hananel, J. Foley, M. Eames, N. Kassell, and H. H. Simonin. Focused Ultrasound Development and Clinical Adoption: 2013 Update on the Growth of the Field. *J Ther Ultrasound*, 2:2, 2014.

[48] W. F. Fry. Ultrasound: Its Roles in Basic and Applied Neurologic and Psychologic Research. *Presented at 1959 Alfred Korzybski Memorial Symposium 'In General Semantics Bulletin'*, 24 - 25, 1959.

[49] A. Chapman and G. ter Haar. Thermal Ablation of Uterine Fibroids Using MR-Guided Focused Ultrasound - A Truly Non-Invasive Treatment Modality. *Eur Raiol*, 17(10):2505–2511, 2007.

[50] A. Blana, S. Rogenhofer, R. Ganzer, J. C. Lunz, M. Schostak, W. F. Wieland, and B. Walter. Eight Years' Experience with High-Intensity Focused Ultrasonography for Treatment of Localized Prostate Cancer. *Urology*, 72(6):1329–1333, 2008.

[51] A. Okada, T. Murakami, K. Mikami, H. Onishi, N. Tanigawa, T. Marukawa, and H. Nakamura. A Case of Hepatocellular Carcinoma Treated by MR-Guided Focused Ultrasound Ablation with Respiratory Gating. *Magn Reson Med Sci*, 5(3):167–171, 2006.

[52] Y. Y. Li, W. H. Sha, Y. J. Zhou, and Y. Q. Nie. Short and Long Term Efficacy of High Intensity Focused Ultrasound Therapy for Advanced Hepatocellular Carcinoma. *J Gastroenterol Hepatol*, 22(12):2148–2154, 2007.

[53] R. D. Kovatcheva, J. D. Vlahov, J. I. Stoinov, and K. Zaletel. Benign Solid Thyroid Nodules: US-Guided High-Intensity Focused Ultrasound Ablation-Initial Clinical Outcomes. *Radiology*, 276(2):597–605, 2015a.

[54] B. Liberman, D. Gianfelice, Y. Inbar, A. Beck, T. Rabin, N. Shabshin, G. Chander, S. Hengst, R. Pfeffer, A. Chechick, A. Hanannel, O. Dogadkin, and R. Catane. Pain Palliation in Patients with Bone Metastases Using MR-Guided Focused Ultrasound Surgery: A Multicenter Study. *Ann Surg Oncol*, 16(1):140–146, 2009.

[55] R. D. Kovatcheva, J.-N. Guglielmina, M. Abeshera, L. Boulanger, N. Laurent, and E. Poncelet. Ultrasound-Guided High-Intensity Focused Ultrasound Treatment of Breast Fibroadenoma - A Multicenter Experience. *J Ther Ultrasound*, 3(1), 2015b.

[56] M. Zaaroor, A. Sinai, D. Goldsher, A. Eran, M. Nassar, and I. Schlesinger. Magnetic Resonance-Guided Focused Ultrasound Thalamotomy for Tremor: A Report of 30 Parkinsons Disease and Essential Tremor Cases. *J Neurosurg*, 24:1–9, 2017.

[57] K. Hynynen, C. Damianou, A. Darkazanli, E. Unger, and J. F. Schenck. The Feasibility of Using MRI to Monitor and Guide Noninvasive Ultrasound Surgery. *Ultrasound Med Biol*, 19(1):91 – 92, 2017.

[58] Y. Ishihara, A. Calderon, H. Watanabe, K. Okamoto, Y. Suzuki, K. Kuroda, and Y. Suzuki. A Precise and Fast Temperature Mapping Using Water Proton Chemical Shift. *Magn Reson Med*, 34:814 – 823, 1995.

[59] J. C. Hindman. Proton Resonance Shift of Water in Gas and Liquid States. *J Chem Phys*, 44:4582 – 4592, 1966.

[60] W. G. Schneider, H. J. Bernstein, and J. A. Pople. Proton Magnetic Resonance Chemical Shift of Free (Gaseous) and Associated (Liquid) Hydride Molecules. *J Chem Phys*, 28:601 – 607, 1958.

[61] R. D. Peters, R. S. Hinks, and R. M. Henkelman. Ex-Vivo Tissue-Type Independence in Proton Resonance Frequency Shift MR Thermometry. *Magn Res Med*, 40(3):454 – 459, 1998.

[62] A. D. Elster. Gradient-Echo MR Imaging: Techniques and Acronyms. *Radiology*, 186(1):1 – 8, 1993.

[63] S. N. Goldberg, G. S. Gazelle, E. F. Halpern, W. J. Rittman, P. R. Muller, and D. I. Rosenthal. Radiofrequency Tissue Ablation: Importance of Local Temperature Along the Electrode Tip Exposure in Determining Lesion Shape and Size. *Acad Radiol*, 3(3):212 – 218, 1996.

[64] J. M. Lee, Y. K. Kim, S. W. Kim, J. K. Han, S. H. Kim, and B. I. Coi. Combined Radiofrequency Ablation and Acetic Acid Hypertonic Saline Solution Instillation: An In Vivo Study of Rabbit Liver. *Korean J Radiol*, 5(1):31 – 38, 2004.

[65] A. Kolandaivelu, M. M. Zviman, V. Castro, A. C. Lardo, R. D. Berger, and H. R. Halperin. Non-invasive Assessment of Tissue Heating During Cardiac Radiofrequency Ablation Using MRI Thermography. *Circ Arrythm Electrophysiol*, 13(5):521 – 529, 2010.

[66] P. S. Yarmolenko, E. J. Moon, C. Landon, A. Manzoor, D. W. Hochman, B. L. Viglianti, and M. W. Dewhirst. Thresholds for Thermal Damage to Normal Tissues: An Update. *Int J Hyperthermia*, 27(4):320 – 343, 2011.

[67] J. M. Slater. From X-Rays to Ion Beams: A Short History of Radiation Therapy. In Linz, U., editor, *Ion Beam Therapy*, pages 3 – 16. Springer Verlag, 2012.

[68] G. M. MacKee. *X-Rays and Radium in the Treatment of Diseases of the Skin*. Lea & Febiger, 1921.

[69] A. Dronsfield and P. Ellis. Radium: A Key Element in Early Cancer Treatment. *Education in Chemistry*, 48(2):56 – 59, 2011.

[70] I. D. Kaplan and E. J. Holupka. Tratment of Prostate Cancer: Brachytherapy. In DeVita, V. T. and Hellman, S. and Rosenberg, S. A., editor, *Progress in Oncology 2001*, pages 335 – 346. Jones & Bartlett Learning, 2011.

[71] D. I. Thwaites and J. B. Tuohy. Back to the Future: The History and Development of the Clinical Linear Accelerator. *Phys Med Biol*, 51(13):R343 – R362, 2006.

[72] R. Baskar, J. Dai, N. Wenlong, R. Yeo, and K.-W. Yeoh. Biological Response of Cancer Cell to Radiation Treatment. *Front Med Biosci*, 1(24):1 – 9, 2014.

[73] J. Ng and I. Shuryak. Minimizing Second Cancer Risk Following Raiotherapy: Current Perspectives. *Cancer Manag Res*, 7:1 – 11, 2015.

[74] L. G. Cockerham, T. L. Walden Jr., C. E. Dallas, G. A. Mickley Jr., and M. A. Landauer. Ionizing Radiation. In Hayes, A. W., editor, *Principles and MEthods of Toxicology*, pages 898 – 954. Taylor & Francis Group, 2008.

[75] S. A. Benjamin, B. E. Powers, K. J. Nikula, and F. F. Hahn. Radiation and Heat. In Haschek, W. M. and Rousseaux, C. G., editor, *Handbook of Toxicologic Pathology*, pages 983 – 1045. Academic Press, 1991.

[76] R. D. Kennedy and A. D. D'Andrea. DNA Repair Pathways in Clinical Practice: Lessons from Pediatric Cancer Susceptibility Sindromes. *J Clin Oncol*, 24(23):3799 – 3808, 2006.

[77] G. Mitchel. The Rationale for Fractionation in Radiotherapy. *Clin J Oncol Nurs*, 17(4):412 – 417, 2013.

[78] P. Montemaggi, M. Trombetta, and L. W. Brady. *Brachytherapy*. Springer, 2016.

[79] I. R. McDougall. Systemic Radiation Therapy with Unsealed Radionuclides. *Semin Radiat Oncol*, 10(2):94 – 102, 2000.

[80] G. L. Glasgow, R. G. Kurup, L. Leybovich, S. Wans, and M. Fatyga. Dosimetry and Use of ^{6}0Co 100 cm Source-to-Axis Distance Teletherapy Units. *Current Oncology*, 3:17 – 25, 1996.

[81] E. A. Eisenhauer, P. Therasse, J. Boqaerts, L. H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe, and J. Verwij. New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1). *Eur J Cancer*, 45(2):228 – 247, 2009.

[82] M. Kong and S. E. Hong. Optimal Follow-Up Duration for Evaluating Objective Response to Radiotherapy in Patients with Hepatocellular Carcinoma: A Retrospective Study. *Chin J Cancer*, 34(2):79 – 85, 2015.

[83] H. Cliffe, C. Patel, R. Prestwich, and A. Scarsbrook. Radiotherapy Response Evaluation Using FDG PET-CT-Established and Emerging Applications. *Br J Radiol*, 90(1071):20160764, 2017.

[84] S. S. Korreman. Motion in Radiotherapy: Photon Therapy. *Phys Med Biol*, 57(23):R161 – R191, 2012.

[85] B. Denis de Senneville, C. Moonen, and M. Ries. MRI-Guided HIFU Methods for the Ablation of Liver and Renal Cancers. In Escoffre, J. - M. and Bouakaz, A., editor, *Therapeutic Ultrasound*, pages 43 – 63. Springer, 2015.

[86] T. Bortfeld, S. B. Jiang, and E. Rietzel. Effects of motion on the total dose distribution. *Semin Radiat Oncol*, 14(1):41 – 51, 2004.

[87] M. Ries, B. Denis de Senneville, S. Roujol, Y. Berber, B. Quesson, and C. Moonen. Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues. *Magnetic Resonance in Medicine*, 64:1704 – 1712, 2010.

[88] T. Bortfeld, M. van Herk, and S. B. Jiang. When should systematic patient positioning errors in radiotherapy be corrected. *Phys Med Biol*, 47(23):N297 – 302, 2002.

[89] M. von Siebenthal and G. Szekely. Systematic Errors in Respiratory Gating due to Intrafraction Deformations in the Liver. *Med. Phys.*, 34(9):265 – 278, 2007.

[90] P. J. Keall, G. S. Mageras, J. M. Balter, R. S. Emery, K. M. Forster, S. B. Jiang, J. M. Kapatoes, D. A. Low, M. J. Murphy, B. R. Murray, C. R. Ramsey, M. B. van Herk, S. S. Vedam, J. W. Wong, and E. Yorke. The Management of Respiratory Motion in Radiation Oncology Report of AAPM Task Group 76. *Med. Phys*, 33(10):3874 – 3900, 2006.

[91] A. Muller, L. Petrusca, V. Auboiroux, P. J. Valette, R. Salomir, and F. Cotton. Management of Respiratory Motion in Extracorporeal High-Intensity Focused Ultrasound Treatment in Upper Abdominal Organs: Current Status and Perspectives. *Cardiovasc Intevent Radiol*, 36(6):1464 – 1476, 2013.

[92] L. J. Verhey. Immobilizing and Positioning Patients for Radiotherapy. *Semin Radiat Oncol*, 5(2):100 – 114, 1995.

[93] F. Wu, Z. B. Wang, W. Z. Chen, H. Zhu, J. Bai, J. Z. Zou, K. Q. Li, C. B. Jin, F. L. Xie, and H. B. Su. Extracorporeal High Intensity Focused Ultrasound Ablation in the Treatment of Patients with Large Hepatocellular Carcinoma. *Ann Surg Oncol*, 11(12):1061 – 1069, 2004.

[94] Y. F. Zhou. High Intensity Focused Ultrasound in Clinical Tumor Ablation. *World J Clin Oncol*, 2(1):8 – 27, 2011.

[95] F. Wu, Z. B. Wang, W. Z. Chen, W. Wang, Y. Gui, M. Zhang, G. Zheng, Y. Zhou, G. Xu, M. Li, C. Zhang, H. Ye, and R. Feng. Extracorporeal High Intensity Focused Ultrasound Ablation in the Treatment of 1038 Patients with Solid Carcinomas in China: An Overview. *Ultrason Sonochem*, 11(3 - 4):149 – 154, 2004.

[96] J. Emmott, B. Sanghera, J. Chambers, and W. L. Wong. The Effects of N-Butylscopolamine on Bowel Uptake: An 18F-FDG PET Study. *Nucl Med Commun*, 29(1):11 – 16, 2008.

[97] J. M. Froelich, M. Daenzer, C. von Weymarn, Ertürk S. M., C. L. Zollikofer, and M. A. Patak. Functional Monitoring of Small Bowel Motility: Comparison of Spasmolysis Induced by Glucagon or Buscopan. *Proc Int Soc Magn Reson Med*, 17:1122 – 1123, 2009.

[98] M. H. Smitsmans, F. J. Pos, J. de Bois, W. D. Heemsbergen, J. J. Sonke, J. V. Lebesque, and M. van Herk. The Influence of a Dietary Protocol on Cone Beam CT-Guided Radiotherapy for Prostate Cancer Patients. *Int J Radiat Oncol Biol Phys*, 71(4):1279 – 1286, 2008.

[99] B. Wysocka, Z. Kassam, G. Lockwood, J. Brierley, L. A. Dawson, C. A. Buckley, D. Jaffray, B. Cummings, J. Kim, R. Wong, and J. Ringash. Interfraction and Respiratory Organ Motion During Conformal Radiotherapy in Gastric Cancer. *Int J Radiat Oncol Biol Phys*, 77(1):53 – 59, 2010.

[100] A. Buchali, S. Koswig, S. Dinges, P. Rosenthal, J. Salk, G. Lackner, D. Böhmer, L. Schlenger, and V. Budach. Impact of the Filling Status of the Bladder and Rectum on Their Integral Dose Distribution and the Movement of the Uterus in the Treatment Planning of Gynaecological Cancer. *Radiother Oncol*, 52(1):29 – 34, 1999.

[101] L. Fokdal, H. Honoré, M. Hoyer, P. Meldgaard, K. Fode, and H. der Maase. Impact of Changes in Bladder and Rectal Filling Volume on Organ Motion and Dose Distribution of the Bladder in Radiotherapy for Urinary Bladder Cancer. *Int J Radiat Oncol Biol Phys*, 59(2):436 – 444, 2004.

[102] D. Mah, J. Hanley, K. E. Rosenzweig, E. Yorke, L. Braban, C. C. Ling, S. A. Leibel, and G. Mageras. Technical Aspects of the Deep Inspiration Breath-Hold Technique in the Treatment of Thoracic Cancer. *Int J Oncol Biol Phys*, 48(4):1175 – 1185, 2000.

[103] D. Latty, K. E. Stuart, W. Wang, and V. Ahern. Review of Deep Inspiration Breath-Hold Techniques for the Treatment of Breast Cancer. *J Med Radiat Sci*, 62(1):74 – 81, 2015.

[104] J. W. Wong, M. B. Sharpe, D. A. Jaffray, V. R. Kini, J. M. Robertson, J. S. Stromberg, and A. A. Martinez. The Use of Active Breathing Control (ABC) to Reduce Margin for Breathing Motion. *Int J Radiat Oncol Biol Phys*, 44(4):911 – 999, 1999.

[105] D. J. Kim, B. R. Murray, R. Halperin, and W. H. Roa. Held-Breath Self-Gating Technique for Radiotherapy of Non-Small-Cell Lung Cancer: A Feasibility Study. *Int J Radiat Oncol Biol Phys*, 49(1):43 – 49, 2001.

[106] A. M. Berson, R. Emery, L. Rodriguez, G. M. Richards, T. Ng, S. Sanghavi, and J. Barsa. Clinical Experience Using Respiratory Gated Radiation Therapy: Comparison of Free - Breathing and Breath - Hold Techniques. *Int J Radiat Oncol Biol Phys*, 60(2):419 – 426, 2004.

[107] Y. Negoro, Y. Nagata, T. Aoki, T. Mizowaki, N. Araki, K. Takayama, M. Kokubo, S. Yano, S. Koga, K. Sasai, Y. Shibamoto, and M. Hiraoka. The Effectiveness of an Immobilization Device in Conformal Radiotherapy for Lung Tumor: Reduction of Respiratory Tumor Movement and Evaluation o fthe Daily Setup Accuracy. *Int J Radiat Oncol Biol Phys*, 50(4):889 – 898, 2001.

[108] H. D. Heerkens, O. Reerink, M. P. V. Intven, R. R. Hiensch, C. A. T. van den Berg, S. P. M. Crijns, M. van Vulpen, and G. J. Meijer. Pancreatinc Tumor Motion Reduction by Use of a Custom Abdominal Corset. *Phys Imag Radiat Oncol*, 2:7 – 10, 2017.

[109] K. Wurstbauer, F. Sedlmayer, and H. D. Kogelnik. Skin Markings in External Radiotherapy by Temporary Tattooing with Henna: Improvement of

Accuracy and Increased Patient Comfort. *Int J Radiat Oncol Biol Phys*, 50(1):179 – 181, 2001.

[110] S. Rathod, A. Munshi, and J. Agarwal. Skin Markings Mathods and Guidelines: A Reality in Image Guidance Radiotherapy Era. *South Asian J Cancer*, 1(1):27 – 29, 2012.

[111] M. Youl, K. K. Brock, and L. A. Dawson. Image-Guided Radiation Therapy. In Small Jr, W., editor, *Clinical Radiation Oncology*, pages 83 – 98. Wiley - Blackwell, 2017.

[112] H. D. Kubo, P. M. Len, S. Minohara, and H. Mostafavi. Breathing - Synchronized Radiotherapy Program at the University of California Davis Cancer Center. *Med Phys*, 27(2):346 – 353, 2000.

[113] T. Tada, K. Minakuchi, T. Fujioka, M. Sakurai, I. Kawase, T. Nakajuma, M. Nishioka, T. Tonai, and T. Kozuka. Lung Cancer: Intermittent Iradiation Synchronized with Respiratory Motion – Results of a Pilot Study. *Radiology*, 207(3):779 – 783, 1998.

[114] H. D. Kubo and B. C. Hill. Respiration Gated Radiotherapy Treatment: A Technical Study. *Phys Med Biol*, 41(1):83 – 91, 1996.

[115] K. T. Malinowski, T. J. McAvoy, R. George, S. Dietrich, and W. D. D'Souza. Mitigating Errors in External Respiraotory Surrogate - Based Models of Tumor Position. *Int J Radiat Oncol Biol Phys*, 82(5):e709– e716, 2012.

[116] T. R. Willoughby, P. A. Kupelian, J. Pouliot, K. Shinohara, M. Aubin, M. Roach 3rd, L. L. Skrumeda, J. M. Balter, D. W. Litzenberg, S. W. Hadley, J. T. Wei, and H. M. Sandler. Target Localization and Real - Time Tracking Using the Calypso 4D Localization System in Patients with Localized Prostate Cacner. *Int J Radiat Oncol Biol Phys*, 65(2):528 – 534, 2006.

[117] T. Gupta and C. Anand Narayan. Image - Guided Radiation Therapy: Physiacian's Perspectives. *J Med Phys*, 37(4):174 – 182, 2012.

[118] J. W. Jenne, T. Preusser, and M. Günther. High - Intensity Focused Ultrasound: Principles, Therapy Guidance, Simulations and Applications. *Z Med Phys*, 22(4):311 – 322, 2012.

[119] J. B. Maintz and M. A. Viergever. A Survery of Medical Image Registration. *Med Image Anal*, 2(1):1 – 36, 1998.

[120] T. W. Tang and A. C. Chung. Non - Rigid Image Registration Using Graph - Cuts. *Med Image Comput Comput Assist Interv*, 10(Pt 1):916 – 924, 2007.

[121] V. Mani and S. Arivazhagan. Survey of Medical Image Registration. *J Biomed Eng Technol*, 1:8 – 25, 2013.

[122] J. Seo, N. Koizumi, M. Mithuishi, and N. Sugita. Ultrasound Image - Based Visual Servoing for Moving Target Ablation by High Intensity Focused Ultrasound. *Int J Med Robot*, pages e1793–n/a, 2016.

[123] J. Lattanzi, S. McNeeley, W. Pinover, E. Horwitz, I. Das, T. E. Schultheiss, and G. E. Hanks. A Comparison of Daily CT Localization to a Daily Ultrasound - Based System in Prostate Cancer. *Int Radiat Oncol Biol Phys*, 43(4):719 – 725, 1999.

[124] M. Fuss, Salter. B. J., S. X. Cavanaugh, C. Fuss, A. Sadeghi, C. D. Fuller, A. Ameduri, J. M. Havezi, T. S. Herman, and C. R. Thomas Jr. Daily Ultrasound - Based Image - Guided Targeting for Radiotherapy of Upper Abdominal Malignancies. *Int Radiat Oncol Biol Phys*, 59(4):1245 – 1256, 2004.

[125] G. Jacob, J. A. Noble, M. Mulet Parada, and A. Blake. Evaluating a Robust Contour Tracker on Echocardiographic Sequences. *Med Image Anal*, 3(1):63 – 75, 1999.

[126] V. De Luca, G. Székely, and C. Tanner. Estimation of Large - Scale Organ Motion in B - Mode Ultrasound Image Sequences: A Survey. *Ultrasound Med Biol*, 41(12):3044 – 3062, 2015b.

[127] C. Thilmann, S. Nill, T. Tücking, A. Höss, B. Hesse, L. Dietrich, R. Bendl, B. Rhein, P. Häring, C. Thieke, U. Oelfke, J. Debus, and P. Huber. Correction of Patient Positionning Errors Based on In - Line Cone Beam CTs: Clinical Implementation and First Experiences. *Radiat Oncol*, 1:16, 2006.

[128] N. Kothary, S. Dietrich, J. D. Louie, D. T. Chang, L. V. Hofmann, and D. Y. Sze. Percutaneous Implantation of Fiduacial Markers for Imaging - Guided Radiation Therapy. *Am J Roentgenol*, 192(4):1090 – 1096, 2009.

[129] T. E. Marchant, A. Skalski, and B. J. Matuszewski. Automatic Tracking of Implanted Fiducial Markers in Cone Beam CT Projection Images. *Med Phys*, 39(3):1322 – 1334, 2012.

[130] D. Litzenberg, L. A. Dawson, H. Sandler, M. G. Sanda, D. L. McShan, R. K. ten Haken, K. L. Lam, K. K. Brock, and J. M. Balter. Daily Prostate Tageting Using Implanted Radiopaque Markers. *Int J Radiat Oncol Biol Phys*, 52(3):699 – 703, 2002.

[131] X. Li, Y. Y. Zhang, Y. H. Shi, L. H. Zhou, and X. Zhen. Evaluation of Deformable Image Registration for Contour Propagation Between CT and Cone - Beam CT Images in Adaptive Head and Neck Radiotherapy. *Technol Health Care*, 24(2):S747 – 755, 2016.

[132] D. Schlesinger, S. Benedict, C. Dietrich, W. Geodryc, A. Klibanov, and J. Larner. MR - Guided Focused Ultrasound Surgery, Present and Future. *Med Phys*, 40(8):080901, 2013.

[133] J. J. Lagendijk, B. W. Raaymakers, and M. van Vulpen. The Magnetic Resonance Imaging - Linac System. *Semin Radiat Oncol*, 24(3):207 – 209, 2014.

[134] S. Mutic and J. F. Dempsey. The ViewRay System: Magnetic Resonance - Guided and Controlled Radiotherapy. *Semin Radiat Oncol*, 24(3):196 – 199, 2014.

[135] Brown, R. W. and Cheng, Y. - C. N. and Haacke, E. M. and Thompson, M. R¿ and Venkatesan, R. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Wiley - Blackwell, 2nd edition, 2014.

[136] J. de Zwart, F. C. Vimeux, J. Palussière, R. Salomir, B. Quesson, C. Delalande, and C. T. Moonen. On-line Correction and Visualization of Motion During MRI - Controlled Hyperthermia. *Magn Reson Med*, 45(1):128 – 137, 2001.

[137] M. O. Köhler, B. Denis de Senneville, B. Quesson, C. T. Moonen, and M. Ries. Spectrally Selective Pencil - Beam Navigator for Motion Compensation of MR - Guided High - Intensity Focused Ultrasound Therapy of Abdominal Organs. *Magn Reson Med*, 66(1):102 – 111, 2011.

[138] S. P. Crijns, B. W. Raaymakers, and J. J. Lagendijk. Proof of Concept of MRI - Guided Tracked Radiation Delivery: Tracking One - Dimensional Motion. *Phys Med Biol*, 57(23):7863 – 7872, 2012.

[139] J. W. Wijlemans, M. de Greef, G. Schubert, L. W. Bartels, C. T. Moonen, M. A. van den Bosch, and Ries M. A Clinically Feasible Treatment Protocol for Magnetic Resonance - Guided High - Intensity Focused Ultrasound Ablation in the Liver. *Inves Radiol*, 50(1):24 – 31, 2015.

[140] V. Auboiroux, L. Petrusca, M. Viallon, A. Muller, S. Terraz, R. Breguet, X. Montet, C. D. Becker, and R. Salomir. Respiratory - Gated MRgHIFU in Upper Abdomen Using an MR - Compatible In - Bore Digital Camera. *Biomed Res Intl*, 2014:421726, 2014.

[141] B. Denis de Senneville, C. Mougenot, and C. T. Moonen. Real - Time Adaptive Methods for Treatment of Mobile Organs by MRI - Controlled High - Intensity Focused Ultrasound. *Magn Reson Med*, 57(2):319 – 330, 2007.

[142] C. F. Baumgartner, C. Kolbitsch, J. R. McClelland, D. Rueckert, and A. P. King. Autoadaptive Motion Modelling for MR - Based Respiratory Motion Estimation. *Med Image Anal*, 35:83 – 100, 2017.

[143] S. Roujol, M. Ries, B. Quesson, C. Moonen, and B. Denis de Senneville. Real - Time MR - Thermometry and Dosimetry for Interventional Guidance on Abdominal Organs. *Magn Reson Med*, 63(4):1080 – 1087, 2010.

[144] B. Denis de Senneville, A. El Hamidi, and C. Moonen. A Direct PCA - Based Approach for Real - Time Description of Physiological Organ Deformations. *IEEE Trans Med Imaging*, 34(4):974 – 982, 2015.

[145] S. C. Kale, J. P. Lerch, R. M. Henkelman, and X. J. Chen. Optimization of the SNR - Resolution Tradeoff for Registration of Megnetic Resonance Images. *Hum Brain Mapp*, 29(10):1147 – 1158, 2008.

[146] M. Glitzner, B. Denis de Senneville, J. J. Lagendijk, B. W. Raaymakers, and S. P. Crijns. On-line 3D Motion Estimation Using Low Resolution MRI. *Phys Med Biol*, 60(16):N301 – N310, 2015.

[147] B. Stemkens, R. Tijssen, B. Denis de Senneville, J. J. Lagendijk, and C. A. van den Berg. Image Driven, Model - Based 3D Abdominal Motion Estimation for MR - Guided Radiotherapy. *Phys Med Biol*, 61(14):5335 – 5355, 2016.

[148] M. van Herk. Errors and Margins in Radiotherapy. *Semin Radiat Oncol*, 14(1):52 – 64, 2004.

[149] J. Meyer, J. Wilbert, K. Baier, M. Guckenberger, A. Richter, O. Sauer, and M. Flentje. Positioning Accuracy of Cone - Beam Computed Tomography in Combination with a HexaPOD Robot Treatment Table. *Int J Radiat Oncol Biol Phys*, 67(4):1220 – 1228, 2007.

[150] J. C. Ye, M. M. Quereshi, P. Clancy, L. N. Dise, J. Willins, and A. E. Hirsch. Daily Patient Setup Error in Prostate Image Guided Radiation Therapy with Fiducial - Based Kilovoltage Onboard Imaging and Conebeam Computed Tomography. *Quant Imaging Med Surg*, 5(5):665 – 672, 2015.

[151] K. Ohara, T. Okumura, M. Akisada, T. Inada, T. Mori, H. Yokota, and M. J. Calaguas. Irradiation Synchronized with Resiratory Gate. *Int J Radiat Oncol Biol Phys*, 17(4):853 – 857, 1989.

[152] P. Giraud and A. Houle. Respiratory Gating for Radiotherapy: Main Technical Aspects and Clinical Benefits. *ISRN Pulmonology*, 2013, 2013.

[153] M. K. Stam, S. P. Crijns, B. A. Zonnenberg, M. M. Barendrecht, M. van Vupen, J. J. Lagendijk, and B. W. Raaymakers. Navigators for Motion Detection During Real - Time MRI - Guided Radiotherapy. *Phys Med Biol*, 57(21):6797 – 6805, 2012.

[154] J. Wijlemans, J. van Breugel, M. de Greef, C. Moonen, M. van den Bosch, and M. Ries. Spontaneous Breathing vs Mechanical Ventilation for Respiratory - Gated MR - HIFU Ablation in the Liver. *J Ther Ultrasound*, 3(Suppl 1):O87, 2015.

[155] S. B. Jiang. Technical Aspects of Image - Guided Respiration - Gated Radiation Therapy. *Med Dosim*, 31(2):141 – 151, 2006.

[156] T. S. Koh, C. H. Thng, P. S. Lee, S. Hartono, H. Rumpel, B. C. Goh, and S. Bisdas. Hepatic Metastases: In Vivo Assessment of Perfusion Parameters at Dynamic Contrast - Enhanced MR Imaging with Dual - Input Two - Compartment Tracer Kinetics Model. *Radiology*, 249(1):307 – 320, 2008.

[157] M. Pernot, M. Tanter, and M. Fink. 3-D Real - Time Motion Correction in High - Intensity Focused Ultrasound Therapy. *Ultrasound Med Biol*, 30(9):1239 – 1249, 2004.

[158] J. W. Choi, J. Y. Lee, E. J. Hwang, I. Hwang, S. Woo, C. J. Lee, E. J. Park, and B. I. Choi. Portable High - Intesity Focused Ultrasound System with 3D Electronic Steering, Real - Time Cavitation Monitoring, and 3D Image Reconstruction Algorithms: A Preclinical Study in Pigs. *Ultrasonography*, 33(3):191 – 199, 2014.

[159] P. J. Keall, V. R. Kini, S. S. Vedam, and R. Mohan. Motion Adaptive X - Ray Therapy: A Feasibility Study. *Phys Med Biol*, 46(1):1 – 10, 2001.

[160] A. Sawant, R. Venkat, V. Srivastava, D. Carlson, S. Povzner, H. Cattell, and P. Keall. Management of Three - Dimensional Intrafraction Motion Through Real - Time DMLC Tracking. *Med Phys*, 35(5):2050 – 2056, 2008.

[161] P. Keall, E. Colvill, R. O'Brien, J. A. Ng, P. R. Poulsen, T. Eade, A. Kneebone, and J. T. Booth. The First Clinical Implementation of Electromagnetic Transponder - Guided MLC Tracking. *Med Phys*, 41(2):020702, 2014.

[162] P. R. Poulsen, B. Cho, A. Sawant, D. Ruan, and P. J. Keall. Detailed Analysis of Latencies in Image - Based Dynamic MLC Tracking. *Med Phys*, 37(9):4998 – 5005, 2010.

[163] D. Ruan. Kernel Density Estimation - Based Real - Time Prediction for Respiratory Motion. *Phys Med Biol*, 55(5):1311 – 1326, 2010.

[164] A. Krauss, S. Nill, and U. Oelfke. The Comparative Performance of Four Respiratory Motion Predictors for Real - Time Tumour Tracking. *Phys Med Biol*, 56(16):5303 – 5317, 2011.

[165] P. Arnold, F. Preiswerk, B. Fasel, R. Salomir, K. Scheffler, and P. C. Cattin. 3D Organ Motion Prediction for MR - Guided High Intensity Focused Ultrasound. *Med Image Comput Comput Assist Interv*, 14(Pt 2):623 – 630, 2011.

CHAPTER 2

An Improved Optical Flow Tracking Technique for Real - Time MR - Guided Beam Therapies in Moving Organs

"To find fault is easy; to do better may be difficult."
- Plutarch

This chapter is based on:
C. Zachiu, N. Papadakis, M. Ries, C. Moonen, and B. Denis de Senneville. An Improved Optical Flow Tracking Technique for Real-Time MR-Guided Beam Therapies in Moving Organs. *Phys. Med. Biol.*, 60(23): 9003 - 9029, 2015.

Abstract

Magnetic resonance (MR) guided high intensity focused ultrasound (HIFU) and external beam radiotherapy (EBRT) interventions, to which we shall refer to as beam therapies/interventions, are promising techniques for the non-invasive ablation of tumors in abdominal organs. Therapeutic energy delivery in these areas becomes, however, challenging due to the continuous displacement of the organs with respiration. Previous studies have addressed this problem by coupling high-framerate MR-imaging with a tracking technique based on the algorithm proposed by Horn & Schunck (H & S), which was chosen due to its fast convergence rate and highly parallelizable numerical scheme. Such characteristics were shown to be indispensable for the real-time guidance of beam therapies. In its original form, however, the algorithm is sensitive to local gray-level intensity variations not attributed to motion such as those that occur, for example, in the proximity of pulsating arteries.

In this study, an improved motion estimation strategy which reduces the impact of such effects is proposed. Displacements are estimated through the minimization of a variation of the H & S functional for which the quadratic data fidelity term was replaced with a term based on the linear L^1 norm, resulting in what we have called an L^2-L^1 functional.

The proposed method was tested in the liver and kidney of two healthy volunteers under free-breathing conditions, on a data set comprised of 3000 images equally divided between the volunteers. Results have shown that, compared to the existing approaches, our method demonstrates a greater robustness to local gray-level intensity variations introduced by arterial pulsations. Additionally, the computational time required by our implementation make it compatible with the work-flow of real-time MR-guided beam interventions.

To the best of our knowledge this study was the first to analyze the behavior of an L^1-based optical flow functional in an applicative context: real-time MR-guidance of beam therapies in moving organs.

2.1 Introduction

MR guided high intensity focused ultrasound (MRg-HIFU) together with MR guided external beam radiotherapy (MRg-RT) are attractive techniques for treating tumors deep inside the human body.[1–4] The non-invasive nature of such interventions opens up additional treatment options for patient groups affected by tumors that are inaccessible to current surgical equipment.[5–7] In particular, MRg-HIFU has the potential to ablate primary and metastatic tumors in a single session.[8] It has already met success in treating several conditions

such as symptomatic uterine fibroids, palliative treatment of bone metastases, prostate cancer, breast fibroadenoma and a number of functional neuropathic disorders.[6] Although external beam radiotherapy (EBRT) is already a standard approach for treating pathologies in numerous locations inside the human body,[9] at the time of this study there is just a single clinically-avaliable device which makes use of real-time MR guidance, i.e. the ViewRay system.[3] Nevertheless, there are great perspectives for MRg-RT clinical interventions with the development of the first MR-Linac prototypes.[4,10,11] One of the challenges that might be encountered during EBRT and HIFU interventions, to which for the purpose of this study we shall alternatively refer to as beam therapies/interventions, is the necessity of a tracking technique for the targeted tissue. When therapy is conducted in the upper abdomen or in the thorax, energy delivery is hampered by physiological motion, which can be coarsely categorized in the following sub-types: respiratory motion, cardiac motion, long term motion (e.g. peristalsis) and spontaneous motion (e.g. coughing, twitching).[12] In this chapter the focus will mainly be on respiratory motion, which, during energy deliveries, was identified to be the most problematic.[13–15] In order to prevent unnecessary damage to healthy tissues, therapeutic energy delivery needs to be correlated with the underlying displacements.[14,16] For example, a gating strategy could be employed in which the beam is turned on and off, depending on whether the targeted area is in a predefined fixed location or not.[11,17,18] An alternative would be to continuously re-adjust the focus of the beam according to the motion pattern exhibited by the target tissue.[16] Taking motion into account during energy deliveries will also prevent the therapeutic energy from being diffused along the motion trajectory, which in turn would lead to the under-treatment of the treated tumorous tissue.[14,16]

An additional challenge is the need for a real-time therapy efficiency monitoring strategy. For example, during a MRg-HIFU intervention, the interaction of the ultrasonic beam with the target tissue and the organs at risk (OAR) is observed through a real-time temperature feedback loop. Therefore, high-framerate MR imaging has to be run in parallel with the energy delivery. However, the time varying geometric distortions induced by the moving target might end up degrading the temperature measurements.[19]

Thus, a real-time motion compensation strategy is necessary both for gating and/or steering the energy beam and, in the particular case of a HIFU intervention, correcting temperature maps. Note that the quality of the latter has a direct impact on the thermal dose measurements, which is the main indicator on whether the endpoint of a HIFU therapy has been reached or not.[16,20]

Motion compensation for EBRT was historically achieved by several approaches including the increase of treatment margins, establishing a relation between external and internal motion through surrogate devices usually situated outside the patients body or tracking the target motion directly by implanting MV/kV detectable fiducial markers. Increasing tumor margins, however, often leads to an increased unnecessary toxicity to otherwise healthy tissues. Surrogate external devices were shown to aid in target tracking for several particular cases,[21]

however, for some sites, it becomes difficult to establish a correlation with target motion, since the relationship between external and internal motion may vary during treatment.[22] While the depiction of target motion directly by tracking implanted fiducial markers may be a more reliable solution than external surrogates,[23] such an approach is usually unable to handle deformations and involves a certain degree of invasiveness. All these approaches and several additional methods, together with their advantages and disadvantages are discussed in great detail in.[24] An image-guided intervention, thus, provides a more attractive solution since tracking can be performed in a completely non-invasive manner. In particular, MR imaging can provide real-time target visualization with high soft tissue contrast. Thus, in anticipation to the future release of the MR-Linac, recent research aims at integrating MR-based estimated displacements in the computation of the delivered radiation dose and based on the results, corrections can be made and the interventional plan re-optimized.[11,25] Designing a real-time motion estimation strategy generates, however, additional challenges. The configuration and convergence of the target tracking algorithm must not exceed the sampling time of the image acquisition process. Also, a temporal margin has to be taken into account for the image acquisition itself and for several additional operations (such as thermometry, energy dose reconstruction, etc.). In this chapter the focus will mainly be on the target tracking algorithm itself. It has already been shown in[16] and[20] that, for respiratory motion compensation, the algorithm proposed by Horn & Schunck[26] is compatible with the work-flow and real-time requirements of MRg-HIFU interventions in the abdomen. The optical flow formulation of Horn & Schunck, initially proposed in the context of estimating motion in video sequences in 1981, assumes that pixels conserve their intensity along their trajectory, to which a spatial regularity constraint of the estimated motion is added. This can be mathematically expressed with the following functional involving two quadratic terms:

$$E_{L2L2}(u,v) = \iint_{\Omega} \left(I_x u + I_y v + I_t\right)^2 + \alpha^2 \left(\|\nabla u\|_2^2 + \|\nabla v\|_2^2\right) \mathrm{d}x\mathrm{d}y \qquad (2.1)$$

where $\Omega \subseteq \mathbb{R}^2$ is the image domain, u and v are the components of the 2D displacement vectors and α is a user defined weighting factor designed to link the data fidelity term (first term of the integral in Eq. (2.1)) and the regularity of the estimated motion field (second term of the integral in Eq. (2.1)). In the data fidelity term, assuming pixel gray-level intensity conservation between successive images, $I_{x,y}$ and I_t are the spatial and respectively temporal partial derivatives of the image pixel intensity. The regularization term is given by $\|\nabla u\|_2^2 = u_x^2 + u_y^2$ and $\|\nabla v\|_2^2 = v_x^2 + v_y^2$, with u_x, u_y, v_x and v_y being the partial spatial derivatives of u and v respectively. For the remainder of this chapter $E_{L2L2}(u,v)$ will be referred to as the L^2-L^2 functional. The estimated displacements are given by the minimizers of $E_{L2L2}(u,v)$ with respect to u and v.

Motion estimation based on the L^2-L^2 functional has several properties that make it attractive for respiratory motion compensation during real-time MR-guided beam interventions:

- It requires tuning only one parameter, namely, the regularization parameter α. In practice, the value of α needs to be optimized only once for a particular contrast weighting, after which the same one can be used for an unlimited number of patients. Its value needs to be re-optimized only if the MR-sequence is significantly changed.[27]
- The numerical scheme employed for the minimization of $E_{L2L2}(u, v)$ has a pixel-wise nature which facilitates parallel processing. This, in turn, reduces the computational time of the algorithm, which is advantageous for applications with short processing latency requirements.

In general, respiratory motion compensation during MR-guided EBRT or HIFU interventions requires high-framerate imaging ($>$ 10Hz).[16, 20, 27] At such sampling frequencies, depending on the image contrast weighting, arterial pulsations and peristaltic contractions may become apparent, and manifest themselves as local intensity variations between successive images. Consequently, pixel gray-level intensity is no longer conserved. This will have a direct impact on the quality of the motion estimates, since the assumption made for the construction of the data fidelity term in the H & S functional is locally violated. It is expected, however, that in such cases, these effects are partially reduced by the smoothness constraint imposed on the estimated motion by the regularization term. For example, if tracking is performed for the liver, even if the registration process might be influenced by the pulsating hepatic arteries, the displacements estimated at the edges of the liver will be propagated inwards and reduce the effects caused by the local gray-level intensity variations. This, however, might become problematic when the hepatic tissue appears in the images with a low gray-level intensity, and the registration process has to rely mostly on the information provided by the blood vessels. Such is the case, for example, of an MRg-HIFU intervention for which, during energy delivery, real-time thermometry usually provides T_2^*-*weighted* images. Under such a contrast weighting, the hepatic tissue appears dark, while the blood vessels are hyperintense, thus it is expected that arterial pulsations will have a large impact on the motion estimates. To the authors knowledge such effects have not been previously quantified in the context of image guided beam interventions.
In this chapter we propose solving the motion estimation problem via the following functional:

$$E_{L2L1}(u, v) = \iint_{\Omega} |I_x u + I_y v + I_t| + \beta^2 \left(\|\nabla u\|_2^2 + \|\nabla v\|_2^2 \right) \, \mathrm{d}x\mathrm{d}y \qquad (2.2)$$

Given that the data fidelity term of the above functional is no longer quadratic, it is expected that the estimated motion will be less prone to errors caused by

local gray-level intensity perturbations in cases such as the one described above. For the remainder of this chapter $E_{L2L1}(u, v)$ will be referred to as the L^2-L^1 functional.
In general, quantifying the improvement in the quality of the estimates provided by a new optical flow functional over the ones provided by the original model is challenging since:

- The new functional might compensate for the drawbacks of the original model only in a limited number of scenarios, having a weaker performance otherwise.
- The input parameters of the compared motion estimation models must be set in such a way that the difference in the quality of the estimates is not biased by a sub-optimal configuration.
- In order to keep the intervention non-invasive, *in-vivo* quantification of the quality of the estimates is made difficult due to the lack of a reliable gold standard.

Thus, in the current study, we additionally propose an experimental setup designed to:

- Quantify the effect of arterea pulsations on the motion estimates obtained using the L^2-L^2 model.
- Quantify the improvement in the quality of the estimates provided by the proposed L^2-L^1 functional over the ones provided by the L^2-L^2 criterion.

2.2 Method description

2.2.1 Proposed motion estimation strategy based on the L^2-L^1 functional

For the original L^2-L^2 model, minimization with respect to u and v is achieved by solving the Euler-Lagrange equations. However, the proposed L^2-L^1 functional is no longer differentiable, thus it cannot be minimized by using the same approach. Since $E_{L2L1}(u, v)$ is a convex functional, we propose its minimization via the primal-dual algorithm.[28] Special attention was paid to the configuration parameters of the numerical scheme in order to ensure a fast convergence of the algorithm under various conditions in terms of noise/observed organ displacement amplitudes. The implementation of the algorithm was included as a part of a larger motion estimation library that is freely accessible[1].

[1]The full library can be downloaded from http://bsenneville.free.fr/RealTITracker/.

2.2.1.1 Implemented optimization scheme

We propose the minimization of the L^2-L^1 functional via the primal-dual algorithm. Details concerning the primal-dual algorithm, its adaptation to the L^2-L^1 criterion and the stability of the associated numerical scheme are found in A.

2.2.1.2 Coarse-to-fine scheme

For both the L^2-L^2 and the L^2-L^1 models the data fidelity term is obtained through a linearization by Taylor approximation of the pixel intensity conservation constraint. By ignoring the higher order terms of the expansion, displacements that are larger than the size of one pixel cannot be estimated. In order to overcome this limitation, we adopted an approach similar to the one described in[29] : A a coarse-to-fine strategy was carried out, which iterated the registration algorithm from a 4-fold downsampled image step by step to the original image resolution. In addition, an iterative refinement of the motion estimates was performed within each resolution. This implies running the algorithm several times at the same resolution, initializing the motion fields at the current run of the algorithm with the motion fields that resulted during the previous run. In this manner, the stability of the numerical scheme is improved and at the same time a better quality of the estimates is obtained.

2.2.1.3 Convergence criterion for the numerical scheme

It was considered that the numerical scheme in Eq. (2.9) converged when the average variation of the motion magnitude from one iteration to the next was smaller than 10^{-3} pixels.

2.2.2 Experimental validation

All images (I) were registered to a common reference image (I_{ref}). Image acquisition and processing has been performed in 2D. This is due to the fact that, on currently available MR-systems, the acquisition of a 3D MR image with sufficiently good quality coupled with a registration procedure becomes timewise challenging considering the low latency requirements of a real-time MR-guided beam intervention.
The displacements estimated using the L^2-L^1 model were compared in terms of both accuracy and precision to the ones estimated by the existing techniques based on the L^2-L^2 functional.

2.2.2.1 MR imaging protocol

Dynamic MR imaging was performed under free-breathing conditions on the abdomen of two healthy volunteers. For each volunteer a total number of 1500 dynamics have been acquired over a duration of $\sim$2 min, with an imaging

frame-rate of 12 images/s. The MR sequence was a single-shot gradient recalled echo employing the following parameters: repetition time (TR)=80 ms, echo time (TE)=37 ms, bandwidth in readout direction = 1250 Hz, flip angle=20°, field of view (FOV)=40×40 cm^2, matrix=160×160, in-plane voxel size=2.5 × 2.5mm^2, slice thickness=7 mm, using a 12 element phased array body coil. Each dynamic consisted in a single slice acquired in the coronal plane. Acquisition has been conducted on a Philips Achieva 1.5 T (Philips Healthcare, Best, The Netherlands).

2.2.2.2 Implementation of the motion estimation strategy based on the L^2-L^2 functional

Given that the expression in Eq. (2.1) is differentiable, its minimizers were found by solving the associated Euler-Lagrange equations. These, in turn, generate a linear system which was solved using the Gauss-Seidel method. In order to make a proper comparison with the estimates provided by the L^2-L^1 functional, the same coarse-to-fine scheme and convergence criterion as described in sections 2.2.1.2 and respectively 2.2.1.3 were employed.

2.2.2.3 Quantifying the quality of the estimated motion

Optical flow algorithms are intrinsically sensitive to gray-level intensity variations not attributed to motion. When the intervention is performed for organs in the upper abdomen, such variations might occur due to arterial pulsations:

- In the vicinity of the targeted region. This could happen, for example, when motion estimation is performed in the kidney, case in which the estimated displacements might be affected by the pulsations of the renal aorta.
- Within the targeted region. For example, when the intervention is performed in the liver, the estimated motion may be influenced by the pulsations of the hepatic arteries.

To assess the quality of the estimated motion, we propose a criterion based on the pixel-wise endpoint error (EE):

$$EE(x,y) = \sqrt{(u(x,y) - u_{gold}(x,y))^2 + (v(x,y) - v_{gold}(x,y))^2} \tag{2.3}$$

where (x, y) indicates the location on the image grid, (u, v) is the estimated motion and $(u_{\text{gold}}, v_{\text{gold}})$ is the gold standard motion. The spatial and temporal distribution of the EE was analyzed over two manually defined regions of interest (ROI) encompassing the kidney and the liver respectively. The procedure employed for obtaining u_{gold} and v_{gold} will be detailed in the following paragraphs.
To assess and compare the quality of the motion estimates provided by the L^2-L^2 and the L^2-L^1 models we propose three test benches:

a Reliability test #1 The purpose of this reliability test was to quantify the impact of arterial pulsations on the respiratory displacements estimated by the two compared functionals. Fig. 2.1 summarizes the processes and their succession implied by the current test bench. For the creation of the gold standard, the initial image input stream was temporally sub-sampled by retrospectively selecting solely the images that were acquired during the systolic cardiac phase. The selection procedure will be detailed in section a. This operation will generate a stream of images for which apparent arterial pulsations are minimized, however, the image sampling frequency is reduced from 12 images/s (as in the light green columns) to ∼1 image/s (as in the light blue column), which is the approximate frequency of the cardiac cycle. Displacements were then estimated on the resulting images using both the L^2-L^2 and the L^2-L^1 functional. Since arterial pulsations are minimal, we can assume that their influence on the motion estimates is negligible. Following the registration procedure, positional updates with a frequency of ∼1 image/s are obtained. The resulting motion fields were then upsampled through linear interpolation back to the sampling frequency of the original image sequence. In this test, the resulting interpolated motion fields played the role of gold standard in the computation of the EE. Subsequently, registration was carried out on the original set of images, in which the arterial pulsations are still present, using both the existing and the proposed motion estimation criterion. Finally, the EE was computed between the displacements estimated on the original set of images and the previously created gold standard.

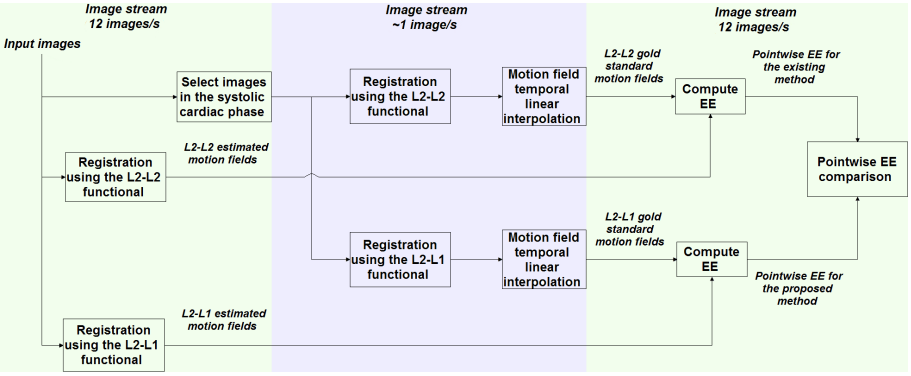


Figure 2.1: Diagram that summarizes the processes and their succession for test bench #1.

b Reliability test #2 This test was designed as a "sub-set" of reliability test #1; i.e. it operated on a more refined dataset in order to emphasize the extent of the potential benefits brought by the proposed L^2-L^1 criterion that might not have been apparent during test #1. For this, in a first step,

both apparent arterial pulsations and respiratory displacements were reduced through a succession of retrospective image selection procedures. Details on these procedures will be discussed in section 2.2.2.4. Only images acquired in the systolic cardiac phase and at full exhalation were further considered for registration (see Fig. 2.2). Note that the initial selection with respect to the cardiac phase will reduce the sampling frequency of the original image stream from 12 images/s (as in the light green column) to ∼1 image/s (as in the light blue column), while the subsequent selection according to the respiratory phase will further reduce it to ∼0.2 images/s (as in the light red column), which is the approximate frequency of the respiratory cycle. The resulting images were then separately registered to two reference images, both acquired at full exhalation but at different stages of the cardiac cycle. Registration was performed using both the both the L^2-L^2 and the L^2-L^1 functional. Since the organs of interest are apparently immobilized, the gold standard for this test was chosen equal to zero. Thus, the EE coincides with the magnitude of the estimated motion vectors.

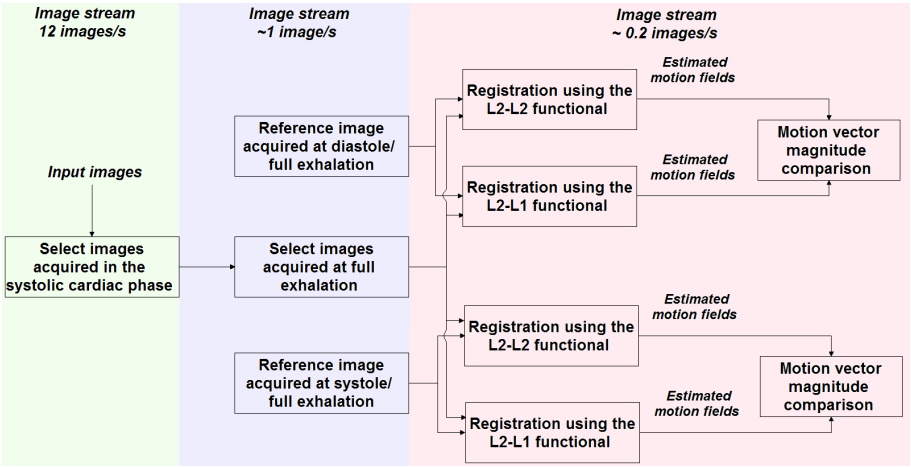


Figure 2.2: Diagram that summarizes the processes and their succession for test bench #2.

c Reliability test #3 The purpose of this test was to establish the robustness to noise of the L^2-L^1 model compared to the L^2-L^2. Fig. 2.3 illustrates a diagram summarizing this test. Precision of the estimated organ displacements was analyzed for both the proposed and the existing motion models under various SNR conditions as follows. In a first step, for each of the volunteers all the images acquired in the systolic cardiac phase and at full exhalation were retrospectively selected and stored in a buffer. An image with a high SNR was created by temporally averaging all images stored inside the buffer. Subse-

quently, increasing levels of Rician noise were added to the resulting noise-free image. Registration was then performed on pairs of images affected by the same level, but resulting from different realizations of the noise. The spatial average of the EE was used as a quality quantification criterion for the motion estimates. The u_{gold} and v_{gold} for the computation of the EE were set to zero, since the registration operated on copies of the same image with the difference that they are affected by independent realizations of the same level of noise. The image sampling frequency in each processing block is indicated at the top of the colored columns in Fig. 2.3.

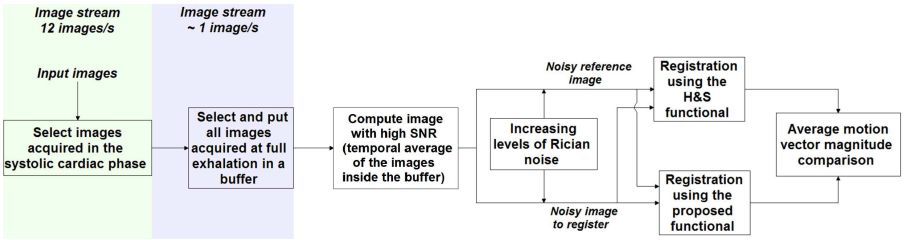


Figure 2.3: Diagram that summarizes the processes and their succession for test bench #3.

2.2.2.4 Retrospective image selection procedures

All three test benches described in section 2.2.2.3 imply selective processing of images depending on the specific time at which they were acquired during the respiratory and/or respectively the cardiac cycle. For this, decision criteria have to be established for the retrospective selection of images acquired at a particular respiratory and/or cardiac phase.

a Image selection with respect to the cardiac phase The slice position of the dynamic MR images acquired on the volunteers was set such that the renal aorta was clearly visible. A ROI encompassing the latter was defined and the averaged gray-level intensity inside the region was analyzed over time. The zeroes of the first order derivative of the obtained time-curve were used to identify the moments at which the signal inside the ROI was either minimal or maximal. This allowed deciding whether the images were acquired during the systolic or respectively diastolic cardiac phase.

b Image selection with respect to the respiratory phase Initially, images acquired during the systolic cardiac phase were selected using the procedure described above. Registration was then performed on the resulting image set using the existing L^2-L^2 motion estimation criterion. Since the previously performed image selection minimizes the intensity variations originating from arterial pulsations, their impact on the estimated motion is expected to be

negligible. A ROI encompassing the liver was defined and the averaged displacements inside the region were analyzed over time. The zeroes of the first order derivative of the obtained time-curve provided the minima and the maxima of the average displacement, which were used to decide whether the images were acquired at the apex of expiration or respectively inspiration.

2.2.2.5 Calibration of the compared functionals

Both the L^2-L^2 and the L^2-L^1 functionals require the fixation of a weighting parameter that links the data fidelity term and the regularization term (see Eq. (2.1) and (2.2)). During the registration process the image gray-level intensities were normalized such that they lie between 0 and 1. Since in this interval $\|x\|_1 \geq \|x\|_2^2$ (where x is an n-dimensional vector) it is expected that $\beta \geq \alpha$ for the functionals to provide similar motion estimates.
In order to reduce the impact of arterial pulsations during the calibration process, images were selected with respect to the cardiac phase in which they were acquired using the procedure detailed in section a. Using solely the images acquired during the systolic phase, the weighting parameters were fixated as follows:

a Calibration of α for the E_{L2L2} functional Two landmarks, one in the liver and one in the kidney, were manually tracked over the resulting set of images. Subsequently, registration of the images was performed using the L^2-L^2 motion model, while varying α between 0 and 1 with a fixed increment of 1/20. The estimated landmark displacements were compared in terms of EE to the ones resulting from the manual tracking. The value of α which provided the smallest time-averaged EE was established as optimal.

b Calibration of β for the E_{L2L1} functional After optimizing the value of α for the L^2-L^2 functional, β was chosen such that the displacements estimated at the previously established landmarks by the L^2-L^1 model are as close as possible to the ones estimated using L^2-L^2. For this purpose, registration was initially performed using the L^2-L^2 functional with the established optimal value for the parameter α. Subsequently, the displacements were estimated using the L^2-L^1 model while varying the value of β between 0 and 1 with an increment of 1/20. The value of β which minimized the temporally averaged EE between the displacements estimated using the L^2-L^2 and respectively the L^2-L^1 functional was established as optimal.

2.2.2.6 Estimation of the true motion induced by arterial pulsations

It is expected that the motion estimated in the vicinity of arteries is also influenced by tissues being displaced by the periodic increasing and decreasing vessel diameter with the cardiac cycle. The extent of these displacements has

been evaluated as follows. A set of 8 multi-slice MR images have been acquired on the abdomen of the same volunteers as in section 2.2.2.1 with the following protocol: repetition time (TR)=1714 ms, echo time (TE)=55 ms, flip angle=90°, field of view (FOV)=40×32×6 cm^3 , slice thickness=5 mm, matrix=268×268×10, with inflow suppression of blood. The images were acquired using respiratory and cardiac gating such that all images reflect the same respiratory phase, but at different stages of the cardiac cycle. The diameter of the two most visible hepatic arteries was then manually measured at the different stages of the cardiac cycle.

2.2.2.7 Hardware

Implementation of the motion estimation algorithms was performed on an Intel 3.2 GHz i7 workstation (8 cores) with 16 GB of RAM. All computationally intensive calculations were offloaded to a dedicated NVIDIA Tesla C2075 graphics processing unit (GPU) with 6 GB of DRAM.

2.2.2.8 Implementation

A multi-threaded C++ implementation was performed for both the Horn & Schunck and the primal-dual algorithm, employed for the minimization of the L^2-L^2 and respectively the L^2-L^1 functional. A total number of 8 threads was found to provide optimal acceleration for the convergence of the numerical schemes. Above 8, the overhead introduced by thread synchronization would prevent further acceleration. An implementation of the algorithms on a GPU using the Compute Unified Device Architecture (CUDA) was also performed.

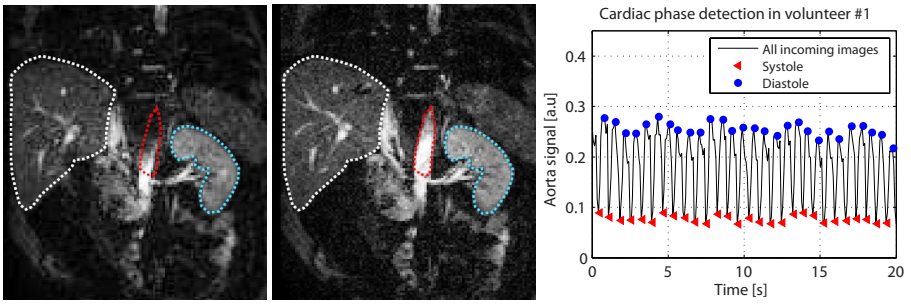
2.3 Results

Initially, in section 2.3.1, we analyze the performance of the image selection techniques described in section 2.2.2.4. Section 2.3.2 reports the values of the regularization parameters α and β resulting from the optimization procedures described in section 2.2.2.5. Both image selection and the calibration of the regularization parameters subsequently allowed analyzing the performance of the two compared functionals against the three proposed test benches (see section 2.2.2.3). The outcome of this analysis is reported in section 2.3.3. In section 2.3.4, we illustrate the measured tissue displacements due to the periodic contraction and dilatation of the arteries. Finally, section 2.3.5 reports the computational time of the numerical scheme implemented for the minimization of the proposed L^2-L^1 functional.

2.3.1 Retrospective image selection with respect to the cardiac and respiratory phase

2.3.1.1 Image selection with respect to the cardiac phase

Fig. 2.4a and 2.4b illustrate two dynamics acquired on volunteer #1 during the systolic and respectively diastolic cardiac phase. The image acquisition protocol is described in section 2.2.2.1. In order to do an automatic selection of the images according to the cardiac phase in which they were acquired, the average gray-level intensity inside a ROI (depicted by the red dotted line in Fig. 2.4) placed on the renal aorta was used. It can be seen that during systole (Fig. 2.4a), the average gray level inside the ROI is lower than during diastole (Fig. 2.4b). Fig. 2.4c displays for volunteer #1, the average gray level inside the ROI for the first 20 s of the study. The zeroes of the first order derivative of the displayed curve indicate the position of its minima (red triangles) and maxima (blue circles), which, in turn, identify the cardiac phase in which the images were acquired: systole or respectively diastole. The white and the blue ROIs depicted in Fig. 2.4a and 2.4b identify the liver and respectively the kidney in the two images.



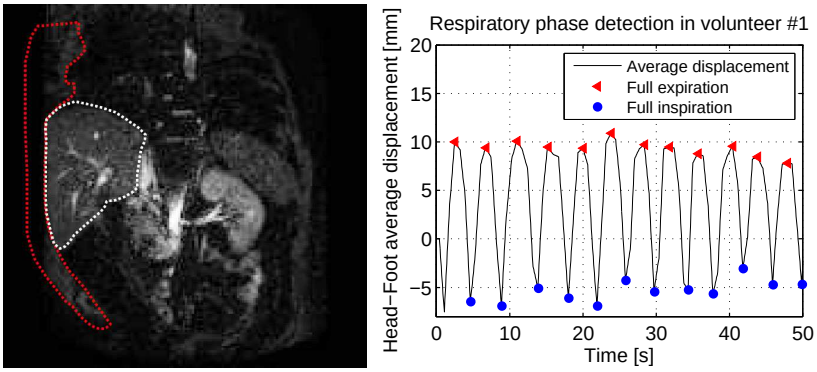
(a) (b) (c)

Figure 2.4: Employed strategy for retrospectively selecting images with respect to the cardiac phase in which they were acquired, exemplified for volunteer #1: (a): Image acquired during systole; (b): Image acquired during diastole; (c): Curve which displays the average gray-level intensity inside a ROI encompassing the renal aorta, delimited by the red dotted line in (a) and (b) during the first 20 s of the study. The red triangles and the blue circles indicate the locations of the minima and respectively the maxima of the curve, as identified by the zeroes of its first order derivative.

2.3.1.2 Image selection with respect to the respiratory phase

Initially, apparent arterial pulsations were minimized using the procedure described in section a. The average displacement inside a ROI encompassing the

liver (depicted by the white dotted line in Fig. 2.5a) was then estimated over time using the L^2-L^2 functional. The resulting curve is displayed in Fig. 2.5b. The zeroes of the first order derivative of this signal were used to find its minima and maxima, which, in turn, will identify the respiratory phase in which the image was acquired: full inhalation or full exhalation respectively.



(a) (b)

Figure 2.5: Employed strategy for selecting images with respect to the respiratory phase in which they were acquired, exemplified for volunteer #1. (a): Image used as reference for estimating the average displacement of the liver for images acquired in the systolic cardiac phase. (b): Curve displaying the evolution of the average displacement of the pixels inside a ROI encompassing the liver depicted by the white dotted line in figure (a). The minima and the maxima of the curve indicate the respiratory phase in which the images were acquired. The red ROI identifies the location of the abdominal wall.

2.3.2 Calibration of the input parameters

2.3.2.1 Calibration of the α parameter for the L^2-L^2 functional

The calibration of the parameter α which links the data fidelity term and the regularization term in the L^2-L^2 functional was performed as described in section a. The two white arrows in Fig. 2.6a indicate the manually tracked landmarks in volunteer #1. Fig. 2.6b displays as a function of α the time averaged EE between the manually tracked displacements and the ones estimated using the L^2-L^2 criterion. In this case, the minimum is attained for $\alpha = 0.3$. This value is in good correspondence with previous reportings.[27,30] Fig. 2.6c depicts the trajectory of the landmark inside the liver of volunteer #1 that resulted from the manual tracking (red dashed line) and respectively from the estimation using the L^2-L^2 criterion for various values of α. It can be noted that for $\alpha = 0.3$ the best approximation of the manually tracked trajectory is

attained. After optimizing the value of α for the second volunteer, the same value was found to generate the lowest EE. This is also reflected by the bar graph illustrated in Fig. 2.6d, where for both of the volunteers, the values of the time-averaged EE were plotted for several values of α.

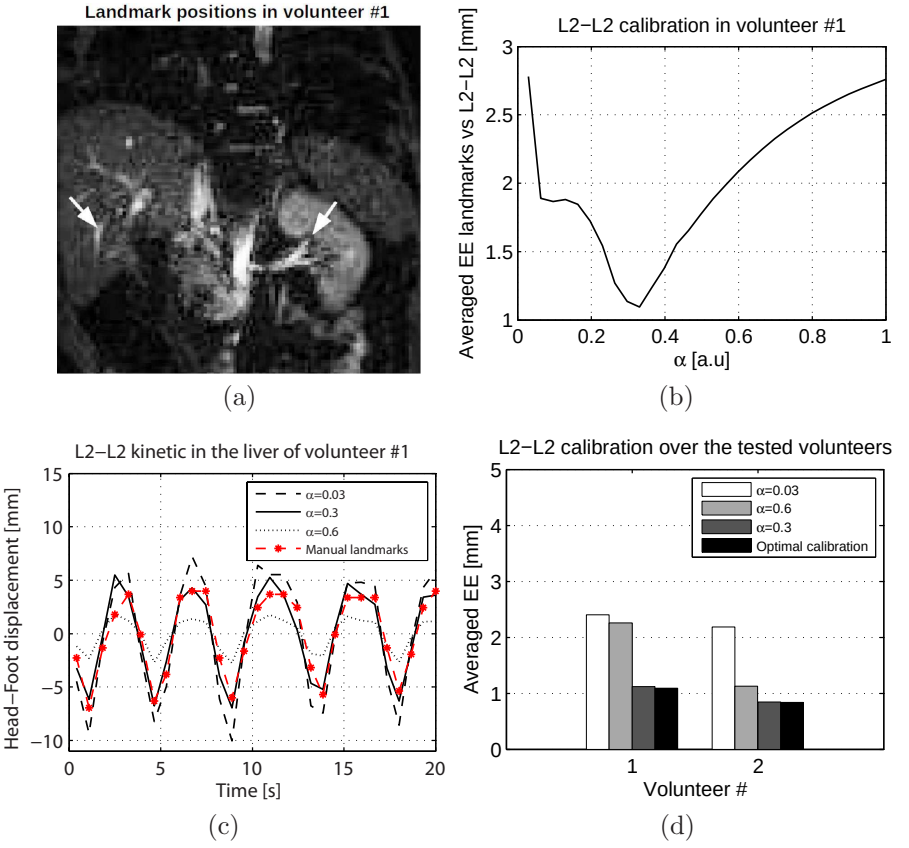
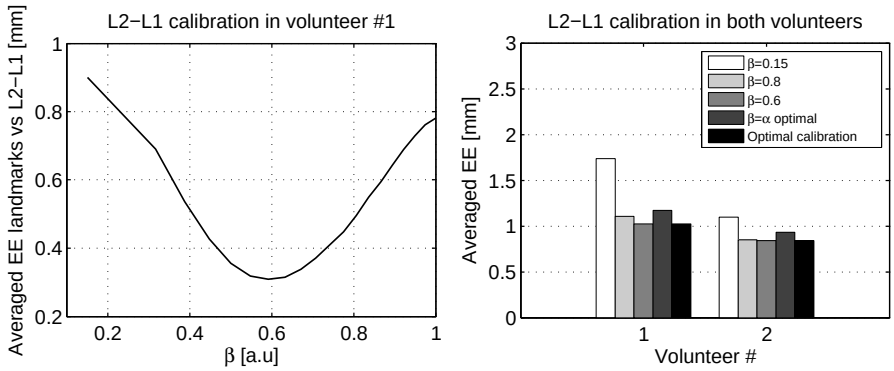


Figure 2.6: Calibration of the α parameter for the L^2-L^2 functional. (a): The reference image upon which the landmarks are indicated by white arrows; (b): Time averaged EE between the displacements resulting from manually tracking the landmarks and the displacements estimated with the L^2-L^2 functional as a function of α; (c): Trajectory of the tracked landmarks. The red dashed line corresponds to the manually tracked trajectory while the black lines are the trajectories estimated by the L^2-L^2 criterion for various values of α; (d): Time averaged EE for both the volunteers for several values of α. The results displayed in (a), (b) and (c) were obtained for volunteer #1.

2.3.2.2 Calibration of the β parameter for the L^2-L^1 functional

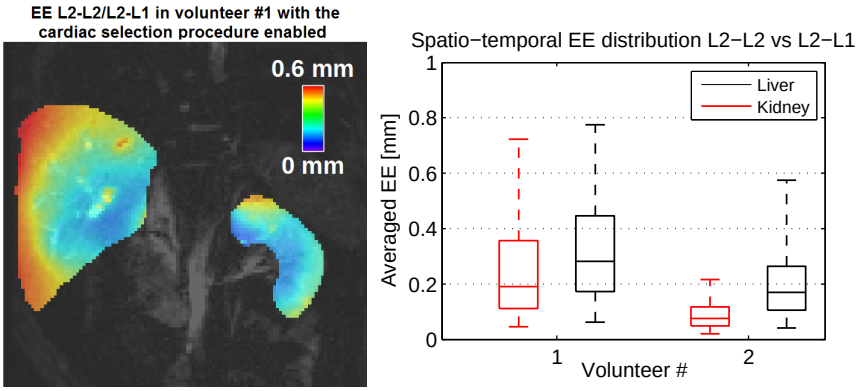
The value of β in the L^2-L^1 functional was optimized using the procedure described in section b. The optimization process was performed in terms of the EE between the displacements estimated at the previously established anatomical landmarks using the L^2-L^2 and respectively the L^2-L^1 criteria. Registration operated on images for which apparent cardiac activity has been minimized. Fig.2.7a illustrates the temporally averaged EE between the displacements estimated by the two criteria as a function of β at the landmark inside the liver of volunteer #1. For $\beta = 0.6$, the L^2-L^1 functional estimated respiratory induced displacements that are most similar to the ones estimated by the L^2-L^2 criterion. Fig.2.7b depicts under the shape of a bar graph the time-averaged EE between the displacements estimated by the L^2-L^1 functional and the manually tracked trajectory in the liver of both volunteers for several values of β. Even though the relative differences in terms of EE as a function of β are not as evident as during the calibration of α, a value of $\beta = 0.6$ was found to be optimal for both volunteers. Fig.2.8a depicts the spatial distribution of the tem-



(a) (b)

Figure 2.7: Calibration of the β parameter in the L^2-L^1 functional. (a): The temporally averaged EE between the motion estimates obtained using the L^2-L^2 and respectively the L^2-L^1 criteria at the landmark in the liver of volunteer #1 (see Fig. 2.6a) as a function of β. (b): Bar graph illustrating the EE between the L^2-L^1 estimates and the manually tracked trajectory for a landmark in the liver of both volunteers for various values of β. Estimation has been conducted only on images acquired during the systolic cardiac phase in order to minimize the influence of arterial pulsations.

porally averaged EE between the respiratory induced displacements estimated in the liver and kidney by the two compared criteria. During the registration process, parameters α and β were set at the previously found optimal values. The boxplots displayed in Fig.2.8b reflect the statistical distribution of the



(a) (b)

Figure 2.8: (a): Spatial distribution of the temporally averaged EE, obtained for volunteer #1, between the respiratory induced displacements estimated with the L^2-L^2 and respectively the L^2-L^1 criteria in the liver and kidney for the optimal values of α and β. (b): The EE from displayed in (a) put under the shape of a boxplot for both volunteers. The relatively small values of the EE is an indicator that the value of β was properly calibrated.

previously computed EE. It is observed that for both the liver and the kidney, errors remain under 0.8 mm for both of the volunteers. This indicates that a proper calibration of the value of β was attained. Additionally, the relatively small overall EE is an indicator that when arterial pulsations are minimized, there is a good match between the motion estimates returned by the two compared functionals, when α and β are properly calibrated. Each boxplot was constructed in the following fashion: the lower limit of the box corresponds to the first quartile, the upper limit of the box is the third quartile, the median is represented by the red horizontal line in the middle of the box, the lower whisker corresponds to the 5^{th} percentile and the upper whisker to the 95^{th} percentile.

2.3.3 Performance comparison between the L^2-L^2 and the L^2-L^1 criteria in the presence of arterial pulsations

2.3.3.1 Reliability test #1

The quality of the estimated motion in the presence of arterial pulsations was initially quantified using the procedure described in section a. Fig.2.9a and 2.9b depict, for volunteer #1, the time averaged EE between the displacements estimated by the two motion estimation criteria and their corresponding gold standard. The construction of the gold standard for each criterion is detailed in section a. It is observed that errors tend to be overall larger when estimation is

performed with the existing L^2-L^2 functional. In particular, error peaks occur in the proximity of the hepatic arteries (indicated by the white arrow in Fig.2.9a) and also in the superior part of the kidney closest to the renal aorta. For the L^2-L^1 criterion the estimation errors do not manifest any clear tendency in the liver. In the kidney, however, errors tend to increase in the area closest to the renal aorta. Fig.2.9c displays the absolute difference between the EE illustrated in Fig.2.9a and 2.9b. This difference further emphasizes the spatial locations in which the two functionals provide motion estimates of different quality in the particular case of reliability test #1. Fig.2.9d and 2.9e illustrate, under the shape of boxplots, the previously computed EE in the liver and kidney of both volunteers. While there is no notable difference in the kidney between the two criteria, there is an obvious improvement in the liver when the estimation is performed using the L^2-L^1 functional. The maximal EE was reduced in the latter from $\sim$2.5 to $\sim$1.5 mm, while the median of the EE decreased from $\sim$2.1 mm to $\sim$1.3 mm. Fig.2.9f and 2.9g display the trajectory of the landmark placed in the liver of volunteer #1 (indicated by a white arrow in figure 2.6a). The red dashed line depicts the trajectory of the landmark given by the gold standard of the two criteria. The black lines correspond to the trajectory of the landmark estimated by the L^2-L^2 and respectively the L^2-L^1 functional. For both criteria there are fluctuations of the landmark position during the positive peak of the curves. Since the landmark was chosen close to one of the hepatic arteries we assume that the source of the fluctuations in the estimated displacement is the local intensity variations due to the pulsations. It can be observed, however, that the amplitude of the fluctuations is much smaller for the L^2-L^1 criterion than for the L^2-L^2.

2.3.3.2 Reliability test #2

This test was designed to evaluate the behavior of the existing and the proposed motion estimation models when registering images for which apparent respiratory motion is minimized, with only arterial pulsations being present. Details concerning this test bench were discussed in section b. All images acquired in the systolic cardiac phase at full expiration were separately registered, using both the motion estimation criteria, to two reference images: one acquired at systole and full expiration, the other at diastole and full expiration. Table 2.1 reports the mean and the standard deviation of the temporally averaged EE between the displacements estimated by the two functionals and their corresponding gold standard (which for this test was established to be zero - see section b). The results were obtained by analyzing the data from both volunteers, but separately for the liver and the kidney. It can be observed that for either of the organs there is no remarkable difference in EE when the reference image is acquired at systole - full exhalation, regardless of the functional which is used for estimation. However, when the reference image is in the diastole - full exhalation phase, both the average and the standard deviation of the EE were reduced with more than 30% by using the proposed L^2-L^1 model instead

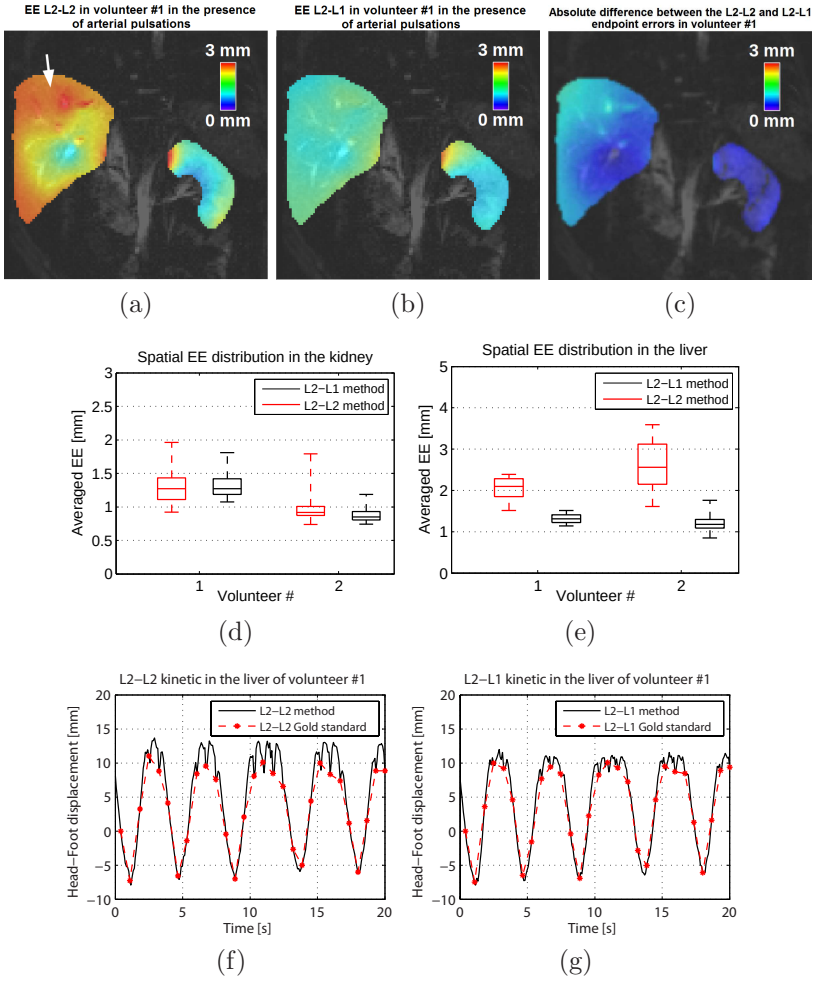


Figure 2.9: Performances of the L^2-L^2 and the L^2-L^1 criteria obtained during reliability test #1. (a), (b): Temporally averaged EE, for volunteer #1, between the displacements estimated with the L^2-L^2 and respectively L^2-L^1 functional and their corresponding gold standards; (c): Absolute difference between the EE displayed in (a) and (b); (d), (e): The EE in (a) and (b) put under the shape of a boxplot separately for the liver and the kidney of both volunteers; (f), (g): Trajectory of a landmark placed close to an artery in the liver of volunteer #1 estimated by the L^2-L^2 and respectively the L^2-L^1 criterion. The red dashed line is the trajectory given by the corresponding gold standard. Note that the fluctuations in the estimated landmark position are lower when the proposed L^2-L^1 functional is used.

of the existing L^2-L^2, for both the liver and the kidney. Note that while no evident improvement could be observed for the latter in test #1, the current test, which is based on a more restricted subset of images, emphasizes the benefit of using the L^2-L^1 functional for estimating kidney motion.

Cardiac phase in which the reference image was acquired	EE [mm]			
	Kidney		Liver	
	L^2-L^2	L^2-L^1	L^2-L^2	L^2-L^1
Systole	0.6 ± 0.5	0.6 ± 0.5	0.7 ± 0.7	0.7 ± 0.6
Diastole	1.2 ± 1.2	0.9 ± 0.7	1.8 ± 1.2	1.0 ± 0.8

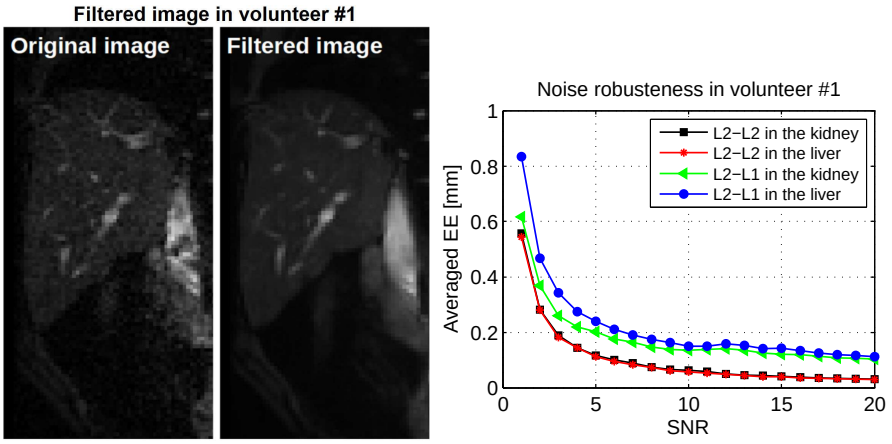
Table 2.1: Mean and standard deviation of the EE obtained during the reliability test #2 by the L^2-L^2 and respectively the L^2-L^1 functionals. The motion estimation procedure operated on images acquired at systole - full exhalation that were independently registered to two reference images acquired at full exhalation, but at the two extreme cardiac phases. Analysis was performed on both volunteers, but separately for the liver and kidney.

2.3.3.3 Reliability test #3

This test had the purpose to evaluate the robustness to noise of the existing and the proposed motion estimation criteria. The evaluation procedure is detailed in section c. Fig.2.10a illustrates an image acquired on volunteer #1 during the systole - full exhalation phase both in its noisy and temporally filtered form. Filtering was performed as described in section c and resulted in an SNR increase from 12.8 in the liver and 20 in the kidney to 20 in the liver and 29.5 in the kidney. Two copies of the filtered image were registered to one another using both motion estimation criteria, while gradually decreasing their SNR by applying the same level but independent realizations of Rician noise. Fig.2.10b depicts the spatially averaged EE between the resulting estimates and their corresponding gold standard (see section c) as a function of the SNR of the images. The analysis has been performed separately for the liver and kidney. Results show that a moderate penalty of 0.1 mm arises for SNR$>$ 5 from the use of the L^2-L^1 metric instead of the L^2-L^2.

2.3.4 Estimation of the true motion due to arterial pulsations

Figure 2.11 displays for volunteer #2 a slice from three of the images acquired using the protocol described in section 2.2.2.6, with focus on the liver. The images illustrated in Fig. 2.11a, 2.11b and 2.11c have been acquired at mid-diastole, 600 ms after the mid-diastole and respectively 1000 ms after the mid-diastole. The arteries indicated by the white arrows had their radius manually



(a) (b)

Figure 2.10: Robustness to noise of the proposed and existing motion estimation criteria. (a): Noisy (SNR_{liver}= 12.8, SNR_{kidney}= 20) and filtered version (SNR_{liver}= 20, SNR_{kidney}= 29.5) of the image used as reference in reliability test #3 for volunteer #1. Only the liver is displayed in the current figure. (b): Spatially averaged EE between the displacements estimated by each of the criteria and their corresponding gold standard as a function of the SNR of the registered images. The analysis has been conducted separately for the liver and the kidney.

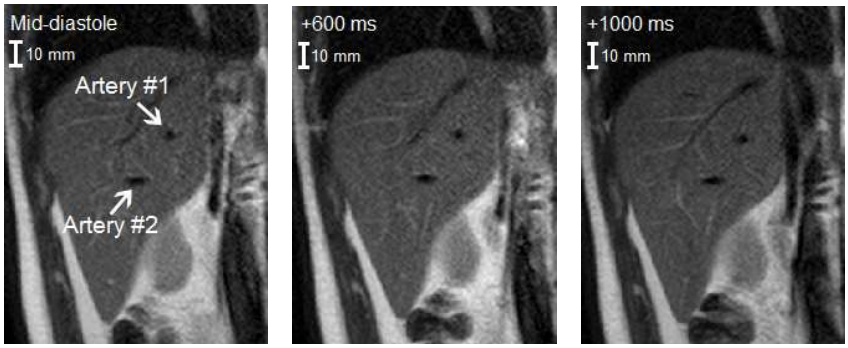
measured in cardiac phases distributed over the entire cardiac cycle. The results were as follows. The radius of artery #1 measured 1.78 ± 0.07 mm (range 1.7 - 1.9 mm) , while artery #2 measured 2.32 ± 0.07 mm (range 2.2 - 2.4 mm).

2.3.5 Computational time of the primal-dual algorithm

The computational time of the primal-dual algorithm used for the minimization of the proposed L^2-L^1 functional was measured for a total of 3000 registrations, i.e. for all the images acquired on the two volunteers. Our C++ implementation converged in 25±1.4 ms per image, with computational times ranging from ∼14 ms to ∼ 29 ms. The GPU implementation did not provide further acceleration.

2.4 Discussion

In recent years, L_1-based variational methods have become increasingly popular for addressing the issue of motion estimation in the video processing domain.[31–33] However, the benefit of such approaches in medical imaging has yet to be shown. To our knowledge, this study is the first to analyze the behavior



(a) (b) (c)

Figure 2.11: Slice from three of the images used to measure the artery radius in volunteer #2 over the cardiac cycle, with focus on the liver. (a): Image acquired at mid-diastole; (b): Image acquired 600 ms after the mid-diastole; (c): Image acquired 1000 ms after mid-diastole. Measurements have been performed for the arteries indicated by the white arrows.

of these techniques in an applicative context: real-time MR-guidance of beam therapies in mobile organs.

In the current study we proposed a motion estimation method that allows real-time tracking of mobile organs which, compared to the existing approaches, has an increased robustness to local gray-level intensity variations not attributed to motion. In order to assess the performance of the proposed approach with respect to previous methods, we have additionally developed a series of tests benches operating on *in-vivo* data sets.

Previous studies have shown that the variational approach for motion estimation proposed by Horn & Schunck has good potential for therapy guidance during MRg-HIFU interventions in moving organs.[16,20,27] We hypothesize that the success of the method is due to the good correspondence between the motion model assumed by the H & S functional and the underlying physical reality. First, the data fidelity term of the functional (see Eq. 2.1) was built on the assumption that the moving image has the same, possibly displaced, content as the reference image. Since the amplitude of the displacements exhibited by the abdominal organs is limited, as long as the FOV covered by the images is properly chosen, this assumption is, in general, satisfied. Second, the regularization term of the functional constrains the estimated motion fields to be spatially continuous. This again is a justified assumption since abdominal organs are incompressible and elastic, thus shearing effects will, in general, not occur. For these reasons, the H & S model was used as a starting point for the construction of the motion estimation functional proposed by the current study.

There are however scenarios in which the H & S model becomes less reliable.

For example, real-time thermometry during MRg-HIFU interventions is usually based on T_2^*-*weighted* images. For such a contrast, the blood might appear more intense than the tracked organs (see, for example, Fig. 2.4a). Since the H & S functional relies on structural information (which provides non-zero spatio-temporal gradients in the data fidelity term) in order to solve the motion estimation problem, blood vessels will have a higher influence on the outcome of the registration process. Given the fast imaging required the real-time thermometry, there is a high chance that the reference and the moving image are acquired at different phases of the cardiac cycle. This leads to arteries having different diameters, and possibly gray-level intensities, in the two images, which violates the pixel intensity conservation hypothesis used for the construction of the data fidelity term. Since the blood vessels have a high influence on the estimated displacement fields, this will lead to erroneous results. In this study, we address this issue by employing a version of the H & S functional for which the quadratic data fidelity term is replaced with an L_1-based term, which is linear. In this manner, the motion estimates become more robust to gray-level intensity variations not attributed to motion.
In their excursion due to respiration, the liver and kidneys slide on the thoracic and respectively abdominal walls which, by comparison, manifest a reduced amount of motion. The arising shearing effects result in two perturbations disturbing the L^2-L^2 functional: one on the data fidelity term and one on the regularization term. The effect on the data fidelity term of local gray-level intensity variations not attributed to motion was reduced using the proposed L^2-L^1 functional. We anticipate that the systematic moderate differences between L^2-L^2 and L^2-L^1 estimates in the vicinity of the thoracic and abdominal walls (see, for example, the red ROI in Fig.2.5a) may be in favor of the L^2-L^1 criterion. However, since the walls are adjacent to the liver and kidneys, the specific impact on the motion regularization term, for its part, may not be negligible, and will need to be addressed in future studies. Due to its quadratic nature, the regularization term will cause the estimated motion fields to be continuous. For this reason, the transition from areas within the liver and kidneys to the thoracic and abdominal walls will be smooth. This implies that false motion will be estimated for the latter due to the propagation of the motion vectors from inside the organs. Inversely, the minimally moving walls will cause a decrease in the magnitude of the motion vectors in adjacent areas from the liver and kidneys. Shearing motion is usually addressed for video sequences by replacing the quadratic regularization term with a linear term (e.g. an L^1 norm). Such techniques are however not straightforward in a clinical context. In particular, a penalty may arise within regions depicting low apparent anatomical structures. In such areas, the low spatio-temporal gradients will cause the data fidelity term to become close to zero. The quadratic regularization would normally propagate the displacements estimated in regions with higher gradients towards these areas. By replacing the quadratic regularization with a linear term the extent to which the motion is propagated will be reduced and thus, in regions with low apparent anatomical structures, the motion estima-

tion problem might become ill-posed. Specific experimental benches will thus be mandatory to assess the benefit of such approaches *in-vivo*. It is, however, important to note that, as long as the assumptions made by the H & S motion model hold, the proposed and the existing motion estimation methods provide similar motion estimates (as shown in Fig. 2.8). Thus, in such a case, it becomes irrelevant which of the methods is chosen to address the motion estimation problem.

A challenging task when estimating displacements for *in-vivo* targets during non-invasive beam interventions is the acquisition of a reliable gold standard. This was of particular interest for reliability test #1 (see section a). At a particular point in the processing chain dedicated to obtaining the gold standard displacement fields, a temporal upsampling procedure was required. This was achieved by the means of a linear interpolation procedure which is, in general, prone to errors. We have evaluated the extent of these errors on an analytic breathing motion pattern (details are found in B). A theoretical maximal error of 0.25 ± 0.2 mm has been estimated for a peak-to-peak motion amplitude of 20 mm, a TR of 80 ms, and the typical breathing and cardiac frequency of 0.2 Hz and 1 Hz, respectively.[34] Given the amplitude of liver and kidney excursions during normal respiration of 11 ± 3 mm and respectively 11 ± 4 mm[14] such errors were deemed to be negligible. However, if the images are acquired with a lower temporal resolution, the interpolation errors are expected to increase. It is worth noting that the resulting interpolation errors are, to a certain degree, sensitive to the simulated breathing motion pattern (see Eq. 2.13) and to the injected model parameters such as the breathing and the cardiac frequency. Nevertheless, the true target motion together with the breathing and cardiac frequency are not expected to vary from the simulated values to such an extent that the interpolation errors exceed sub-millimeter values.

In the current study we have proposed three test benches that aid in comparing the performances of the proposed L^2-L^1 and respectively the existing L^2-L^2 functionals in the presence of local gray-level intensity variations not attributed to motion. In two of the tests, the proposed criterion proved to be more robust to such effects than the existing one. In reliability test #1 respiratory motion was estimated over a series of images containing gray-level intensity variations due to arterial pulsations. By using the L^2-L^1 functional instead of L^2-L^2, while no notable improvement was observed in the kidney, the temporally averaged EE was reduced for almost all pixels inside the liver. Reliability test #2 registers images acquired at the same cardiac and respiratory phase to reference images acquired at the same respiratory phase but at the two extreme phases of the cardiac cycle. This process generates for each image two motion fields for which the magnitude of the contained motion vectors should be close to zero, since we have seen in section 2.3.4 that the true motion due to arterial pulsations is minimal. For this test, the proposed motion estimation criterion decreased the mean and the standard deviation of estimated motion compared to the existing criterion in both the liver and the kidney when the moving and the reference image were acquired at different phases of the cardiac cycle.

Reliability test #3 was designed to test the robustness to noise of the two compared functionals. In this scenario, there are two contributions that impact the gray level intensity variations in the data fidelity term of E_{L2L2} and E_{L2L1}: Spatial variations along anatomical structures present in the FOV, as well as spatio-temporal variations induced by the simulated noise. For an SNR larger than 1, the contribution of anatomical structures exceeds the noise contribution and the quadratic impact of the data fidelity term is an advantage in this case. Nevertheless, since the SNR of the images used for registration is 12.8 in the liver and 20 in the kidney, thus on the right side of the curve depicted by Fig. 2.10, the reduction of the impact of pulsating arteries using the proposed L^2-L^1 criterion is by far more important than the moderate penalty arising from the reduced robustness to noise. Also note from the graph in Fig. 2.10b that, as the SNR drops below 5, the quality of the motion estimates degrades exponentially. This result is in good correspondence with previous reportings.[12] Thus, beyond this threshold, the estimated displacements become unreliable. However, this is not the case for the images used in the current study, since the SNR in the regions of interest lie well above this threshold.
The original Horn & Schunck functional requires a single parameter α as input, the value of which controls the smoothness of the estimated motion fields. This makes it a particularly attractive solution for motion estimation during EBRT and HIFU interventions, since it implies a minimal amount of intervention from the clinician delivering the therapy. Additionally, once the value of α has been optimized for a particular MR contrast weighting, the same value can practically be used for any number of patients as long as the parameters of the sequence remain unchanged. This is also beneficial for the real-time requirements of the aforementioned therapies since α does not have to be re-optimized over the duration of the intervention, thus reducing the overall time delay introduced by the registration process. The same is true for the proposed L^2-L^1 criterion for which parameter β plays the same role as α in the L^2-L^2 functional.
Real-time MR-guided beam interventions usually exploit the MR-image information for a direct retroactive control of the energy delivery device, thus it is necessary to distinguish two key concepts: temporal resolution (i.e. the time interval between two updates/observations) and information latency (i.e. the delay between the actual time of displacement and the availability of motion information). This latency is composed by the sum of the remaining acquisition time after echo-formation, the required data transport time and the image processing time. The data presented in the scope of this chapter has been acquired with 12 images/s and then post-processed off-line. This was necessary since we had to compare the new proposed L^2-L^1 approach with established gold standards (and the L^2-L^2 approach) and to sort images into artefacted/non-artefacted categories. In a real image guided therapy scenario, however, the algorithm would run in real-time with processing on-the-fly. The implementation of the L^2-L^1 algorithm potentially allows average frame rates of $\sim$40 images/s (in practice this is usually limited by the capabilities of the MR scan-

ner) with a typical end-to-end processing latency (from the beginning of the MR-acquisition of the slice to the final output of the motion-vector fields) of under 100 ms. Both are well within the requirements for real-time guidance (see for example[16]).

Both the L^2-L^2 and the L^2-L^1 optical flow algorithms allow the voxel-wise estimation of displacement vectors with sub-voxel precision. While the $2.5 \times 2.5 \times 7\text{mm}^3$ voxel size might appear rather large compared to the typical $1 \times 1 \times 7\text{mm}^3$ HIFU focus size, previous studies have shown that therapy guidance can still be performed with acceptable errors even at such image resolutions.[16,20]

A limiting factor of the current study is the lack of validation of the proposed method under realistic beam therapy scenarios. However, we have shown that, as long as the assumptions made by the H & S algorithm are not violated, the proposed L^2-L^1 motion estimation method has a similar performance to the existing ones based on the L^2-L^2 functional, both in the quality of motion estimates (see Fig. 2.8) and computational time. Moreover, it was proven that, in the eventuality that the pulsating arteries become the dominant structure in the images used for therapy guidance, our method provides motion estimates of superior quality, while maintaining the short latencies required by the real-time nature of the discussed interventions. Thus, given that the L^2-L^2 method was already validated under clinical conditions in previous studies for both MRg-HIFU[16,20] and MRg-RT[35] interventions, the proposed L^2-L^1 method can be easily integrated in the work-flow of such therapies and provide reliable therapy guidance capabilities with increased robustness to arterial pulsations.

This chapter proposes a respiratory motion correction strategy with the effects of cardiac motion/pulsations as a source of artifacts. However, there are additional types of physiological motion might occur during the lengthy EBRT and HIFU interventions, such as spontaneous motion (e.g. coughing or twitching of the patient) and long term drifts (e.g. due to peristaltic activity or slow changes in the respiration baseline). A frame-work for slow 3D organ motion was presented in a previous study, see,[12] which has been validated in-vivo for MRg-HIFU interventions. Additionally, an adaptation of the framework to the work-flow of an MRg-RT intervention is also discussed. Spontaneous motion is a difficult problem due to its infrequent nature and very short timescale. One way to deal with spontaneous motion would be to detect it as fast as possible and to cut the power of the interventional device rapidly and to subsequently to reassess the situation. The respiratory motion correction presented here is fully compatible with the previously published approaches for slow 3D motion and potentially even spontaneous motion.

The proposed motion estimation framework is limited in the fact that it takes into consideration exclusively in-plane displacements. While the trajectory of the kidney and the lower part of the liver could be approximated in a first order through a linear shift, the trajectory of the upper liver is a curve in a 3D space, making it difficult to contain in a 2D slice with static geometry. Dynamic 3D imaging would be ideal, however, on current systems, the acquisition of 3D images coupled with a motion estimation procedure becomes

challenging if the temporal resolution required by real-time beam interventions is to be maintained. One solution is to acquire additional information in the third dimension, such as navigator echoes, and dynamically adapt the slice position/geometry with respect to the through-plane displacements provided by the navigators.[16,36,37] A different full-3D approach, proposed in,[38] consists in dynamically acquiring low-resolution 3D volumes of the target area and its surroundings and register the images, by the means of the H & S algorithm, to a reference volume acquired at the beginning of the intervention. It was shown that such an approach might accelerate the overall image-based therapy guidance process, while maintaining the quality of the motion estimates within acceptable limits. While the method was proven to have only near-real-time capabilities, it shows good potential for future investigations towards real-time guidance of EBRT and HIFU interventions. Additionally, 3D trajectories may be estimated from 2D MRI using one or several volumetric scans obtained before the intervention, as shown in[39,40] and.[35] Note, however, that some of these approaches have not yet underwent pre-clinical or clinical validation.
The 2D real-time motion estimation strategy proposed in this study is not limited to improving the robustness of the motion estimates specifically to arterial pulsations. It also offers good perspectives for improving, in both 2D and 3D, the robustness of the estimates to local gray-level intensity variations originating from other sources, such as: peristaltic contractions, incomplete blood suppression for interventions in the heart or local contrast variations due to the thermal build-up during HIFU ablations.

2.5 Conclusion

The presented study addresses target tracking for MR-guided EBRT and HIFU interventions in the abdomen through an improved real-time optical flow-based motion estimation strategy. The proposed approach based on the L^2-L^1 optical flow functional was shown to be more resilient to local gray level intensity variations not attributed to motion compared to previous approaches which employed the original H & S functional. Our method was shown to be potentially beneficial for interventions in the liver and kidneys, where local gray level intensity variations introduced by pulsating arteries were shown to have a lesser impact on the estimated motion compared to the existing approaches. Additionally, the sub-second temporal resolution of the proposed method renders it compatible with high duty cycle MR-guided beam interventions for treating tumors in the abdomen under free-breathing conditions.

Acknowledgements

This work was supported by the Dutch Technology Foundation (STW) (project OnTrack #12813) and in part by the European Research Council (project ERC-2010-AdG-20100317, Sound Pharma) and ITEA 2 (project SoRTS).

Appendix

A The primal-dual algorithm

A.1 Theoretical background

Let $F : X \to [0; +\infty[$ and $G : X \to [0; +\infty[$ be two proper, convex and lower semi-continuous functions and K a linear operator with $\|K\| = L$. We seek the solution of the problem:

$$\min_{x \in X} F(Kx) + G(x) \tag{2.4}$$

Solving the problem described by Eq. (2.4) is equivalent to finding the solution of the associated dual problem:[28]

$$\max_{y \in Y} -\left(-G^*(-K^*y) + F^*(y)\right) \tag{2.5}$$

where Y is the dual space of X, $F^* : Y \to [0; +\infty[$ and $G^* : Y \to [0; +\infty[$ are the convex conjugates of F and G respectively and K^* is the adjoint operator of K.
Based on Eq. (2.4) and Eq. (2.5) the primal-dual problem can be defined:

$$\min_{x \in X} \max_{y \in Y} \langle Kx, y\rangle + G(x) - F^*(y) \tag{2.6}$$

the solving of which returns the solution of both the primal and its associated dual problem. The application $\langle Kx, y\rangle$ is the inner product between Kx and y. The problem in Eq. (2.6) admits an unique solution given by the fixed point of the following iterative scheme:

$$\begin{cases} y_{n+1} &= \text{prox}_{\sigma F^*}\left(y_n + \sigma K x_n\right) \\ x_{n+1} &= \text{prox}_{\tau G}\left(x_n + \tau K^* y_{n+1}\right)) \end{cases} \tag{2.7}$$

for any σ and τ such that $\sigma\tau L^2 < 1$ and with $x_0 = 0$ and $y_0 = 0$. $\text{prox}_H(z)$ is the proximal operator of function H in point z. Additional technical details regarding the primal-dual algorithm applied for optical flow estimation can be found in.[28]

A.2 Minimization of the L^2-L^1 functional via the primal-dual algorithm

The L^2-L^1 functional was minimized via the primal-dual algorithm, detailed in A.1, with:

$$\begin{cases} F(\nabla w) &= \beta^2 \|\nabla w\|_2^2 \\ G(w) &= \|\left(\nabla I\right)^T \cdot w + I_t\|_1 \end{cases} \tag{2.8}$$

where $\nabla I = (I_x, I_y)^T$ is the 2D spatial gradient of the image being registered, I_t is the difference between the reference image and the image to be registered (i.e. I_t is the temporal gradient), $w = (u, v)^T$ is the 2D displacement vector, F is the L_2 norm multiplied by β^2 and applied on the gradient of w and G is the L_1 norm applied on the data fidelity term. The numerical scheme of the primal-dual algorithm used for the minimization of E_{L2L1} consists in the following:

1. Initialize: $w_0^* = 0$ and $w_0 = 0$.
2. For $n \geq 1$ perform the following iterations:

$$\begin{cases} w_{n+1}^* &= \text{prox}_{\sigma F^*}\left(w_n^* + \sigma \nabla w_n\right) \\ w_{n+1} &= \text{prox}_{\tau G}\left(w_n - \tau(-\text{div}(w_{n+1}^*))\right) \end{cases} \tag{2.9}$$

where w^* is the dual variable of w and *div* is the divergence operator. The remainder of the terms are explained in the following paragraphs. Intuitively, the algorithm consists in four steps:

1. A gradient ascent with a fixed step σ performed in the dual space.
2. A projection of the result obtained in step 1 on σF^* via the proximal operator, where F^* is the convex conjugate of F.
3. A gradient descent with a fixed step τ performed in the primal space.
4. A projection of the result obtained in step 3 on G via the proximal operator.

The proximal operator of a functional H in a point $y \in \mathbb{R}^n$ is given by:

$$\text{prox}_H(y) = \underset{x \in \mathbb{R}^n}{\text{argmin}} \frac{\|x - y\|^2}{2} + H(x) \tag{2.10}$$

It can be shown that the proximal operator of σF^* in a point w^* is given by:

$$\text{prox}_{\sigma F^*}(w^*) = \frac{2\beta^2}{2\beta^2 + \sigma} w^* \tag{2.11}$$

The proximal operator of τG is computed as (see[28]):

$$\text{prox}_{\tau G}(w) = \begin{cases} w + \tau \nabla I & if \quad \rho(w) < -\tau |\nabla I|^2 \\ w - \tau \nabla I & if \quad \rho(w) > \tau |\nabla I|^2 \\ -\rho(w) \frac{\nabla I}{|\nabla I|^2} & if \quad |\rho(w)| \leq \tau |\nabla I|^2 \end{cases} \tag{2.12}$$

where $\rho(w) = (\nabla I)^T w + I_t$. Note that the proximal operators in Eq. (2.11) and (2.12) are computed on a pixel-by-pixel basis.

A.3 Stability of the primal-dual numerical scheme

It was proven in[28] that the numerical scheme described by Eq. (2.9) will converge only if $\tau\sigma\|\nabla\|^2 < 1$, where $\|\nabla\|$ is the norm of the gradient operator. According to[41] for 2D images $\|\nabla\|^2 \leq 8$. Thus, τ and σ need to be chosen such that $\tau\sigma < 1/8$. An additional condition for the convergence of the numerical scheme is that τ itself should be smaller than $\|\nabla\|^2$. Therefore, the parameters τ and σ were established as follows: τ was chosen as close as possible to $1/8$ such that the numerical scheme converges, after which σ was selected as the largest value that fulfills $\tau\sigma < 1/8$.

B Theoretical analysis of the precision of the reliability test #1

In the reliability test #1 (detailed in section a), a temporally under-sampled signal (imaging was retrospectively synchronized to the cardiac activity) was upsampled at the original temporal resolution to build a gold standard data set. To quantify the inherent theoretic uncertainty (denoted by ε) arising from this process, we first define a generic analytic breathing motion pattern as follows:

$$s(t) = A \cdot \sin^4\left(\frac{\pi}{T_r} \cdot t\right) \tag{2.13}$$

where $s(t)$ is the actual displacement amplitude at instant t, A the peak-to-peak displacement, and T_r the period of the breathing activity. Previous studies have shown that such a pattern closely approximates the actual realistic breathing motion.[34,42,43] ε can be mathematically expressed as follows:

$$\varepsilon = \frac{\int_0^{+\infty} [s(t) - S(t)] \cdot \delta_{TR}(t)dt}{\int_0^{+\infty} \delta_{TR}(t)dt} \tag{2.14}$$

where $S(t)$ is the signal reconstructed from the temporally under-sampled data, and δ_{TR} the impulse train corresponding to the original temporal resolution of the acquired MR-images (the time period between impulses is equal to the repetition time TR of employed MR-imaging sequence). $S(t)$ can be mathematically expressed as follows:

$$S(t) = (s(t) \cdot \delta_{T_c}(t)) * \Lambda_{T_c(t)} \tag{2.15}$$

where T_c the period of the cardiac cycle, $\delta_{T_c}(t)$ is the impulse train corresponding to the repetition of the cardiac cycle, and Λ_T is a tent function used in the linear interpolation required by the reconstruction of the temporally under-sampled signal. Λ_T can be expressed as follows:

$$\Lambda_T(t) = \Pi_T(t) * \Pi_T(t) \tag{2.16}$$

where T is the sampling distance, and Π_T is a box function of width T.

Bibliography

[1] H.E. Cline, J.F. Schenck, K. Hynynen, R.D. Watkins, S.P. Souza, and F.A. Jolesz. MR-guided facused ultrasound surgery. *J Comput Assist Tomogr*, 16:956–965, 1992.

[2] K. Hynynen, W.R. Freund, H.E. Cline, A.H. Chung, R.D. Watkins, J.P. Vetro, and F.A. Jolesz. A clinical, noninvasive, MR imaging-monitored ultrasound surgery method. *Radiographics*, 16:185–195, 1996.

[3] S. Mutic and J. F. Dempsey. The viewray system: magnetic resonance-guided and controlled radiotherapy. *Semin Radiat Oncol.*, 24(3):196–199, 2014.

[4] S. P. M. Crijns, B. W. Raaymakers, and J. J. W Lagendijk. Proof of concept of MRI-guided tracked radiation delivery: tracking one dimensional motion. *Phys. Med. Biol.*, 57(23):7863–7872, 2012.

[5] R. Baskar, K. A. Lee, R. Yeo, and K.W. Yeoh. Cancer and radiation therapy: Current advances and future directions. *Int J Med Sci*, 9(3):193–199, 2012.

[6] J.L. Foley, M. Eames, J. Snell, A. Hananel, N. Kassell, and J.-F. Aubry. Image-guided focused ultrasound: state of the technology and the challanges that lie ahead. *Imaging Med.*, 5(4):357–370, 2013.

[7] C.M. Tempany, N.J. McDannold, K. Hynyen, and F.A. Jolesz. Focused ultrasound surgery in oncology: overview and principles. *Radiology*, 36(5):39–56, 2011.

[8] R.O. Illing, J.E. Kennedy, G.R. ter Haar, A.S. Protheroe, P.J. Friend, F.V. Gleeson, D.W. Cranston, R.R. Phillips, and M.R. Middleton. The safety and feasibility of extracorporeal high-intensity focused ultrasound (hifu) for the treatment of liver and kidney tumours in a western population. *Br J Cancer*, 93(8):890–895, 2005.

[9] G. Delaney, S. Jacob, C. Featherstone, and M. Barton. The role of radiotherapy in cancer treatment. *Cancer*, 104(6):1129–1137, 2005.

[10] B.W. Raaymakers, J.J.W. Lagendijk, J. Overweg, J.G.M. Kok, A.J.E. Raaijmakers, E.M. Kerkhof, R.W. van der Put, I. Meijsing, S.P.M. Crijns, F. Benedosso, M. van Vulpen, C.H.W. de Graaff, J. Allen, and K.J. Brown. Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept. *Phys. Med. Biol.*, 54(12):229–237, 2009.

[11] S.P.M. Crijns, J.G.M. Kok, J.J.W. Lagendijk, and B.W. Raaymakers. Towards MRI-gated linear accelerator control: gating on an MRI accelerator. *Phys. Med. Biol.*, 56(15):4815–4825, 2011.

[12] C. Zachiu, B. Denis de Senneville, C. Moonen, and M. Ries. A framework for the correction of slow physiological drifts during MR-guided HIFU therapies: Proof of concept. *Med. Phys.*, 42(7):4137–4148, 2015.

[13] J.T. Booth. *Modelling the impact of treatment uncertainties in radiotherapy.* PhD thesis, University of Adelaide, 2002.

[14] K.M. Langen and D.T.L. Jones. Organ motion and its management. *Int J Radiation Oncology Biol Phys*, 50(1):265–278, 2001.

[15] M. Goitein. Organ and tumor motion: An overview. *Seminars in Radiation Oncology*, 14:2–9, 2004.

[16] M. Ries, B. Denis de Senneville, S. Roujol, Y. Berber, B. Quesson, and C. Moonen. Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues. *Magnetic Resonance in Medicine*, 64:1704–1712, 2010.

[17] K. Ohara, T. Okumura, M. Akisada, T. Inada, T. Mori, H. Yokota, and M.J. Calaguas. Irradiation synchronized with respiratory gate. *Int J Radiation Oncology Biol Phys*, 17(4):853–857, 1989.

[18] A. Okada, T. Murakami, K. Mikami, H. Onishi, N. Tanigawa, T. Marukawa, and H. Nakamura. A case of hepatocellular carcinoma treated by mr-guided focused ultrasound ablation with respiratory gating. *Magn Reson Med Sci.*, 5(3):167–171, 2006.

[19] B. Denis de Senneville, M. Ries, G. Maclair, and C.T.W. Moonen. MR-guided thermotherapy of abdominal organs using a robust PCA-based motion descriptor. *IEEE Transaction on Medical Imaging*, 30(11):1987–1995, 2011.

[20] S. Roujol, M. Ries, B. Quesson, C. Moonen, and B. Denis de Senneville. Real-time MR-thermometry and dosimetry for interventional guidance on abdominal organs. *Magnetic Resonance in Medicine*, 63:1080–1087, 2010.

[21] Y. Seppenwoolde, R. I. Berbeco, S. Nishioka, H. Shirato, and B. Heijmen. Accuracy of tumor motion compensation algorithm from a robotic respiratory tracking system: a simulation study. *Med. Phys.*, 34(7):2774 – 84, 2007.

[22] M. Feng, J. M. Balter, D. Normolle, S. Adusumilli, Y. Cao, T. L. Chenevert, and E. Ben-Josef. Characterization of pancreatic tumor motion using cine MRI: surrogates for tumor position should be used with caution. *Int. J. Radiation Oncology Biol. Phys.*, 74(3):884 – 91, 2009.

[23] A. Y. Minn, D. Schellenberg, P. Maxim, Y. Suh, S. McKenna, B. Cox, S. Dieterich, L. Xing, E. Graves, K. A. Goodman, D. Chang, and A. C. Koong. Pancreatic tumor motion on a single planning 4d-ct does not correlate with intrafraction tumor motion during treatment. *Am. J. Clin. Oncol.*, 32(4):364 – 8, 2009.

[24] P. J. Keall, Gig S. Mageras, J. M. Balter, R. S. Emery, K. M. Forster, S. B. Jiang, J. M. Kapatoes, D. A. Low, M. J. Murphy, B. R. Murray, C. R. Ramsey, M. B. Van Herk, S. S. Vedam, J. W. Wong, and E. Yorke. The management of respiratory motion in radiation oncology report of aapm task group 76a). *Med. Phys.*, 33(10):3874 – 3900, 2006.

[25] M. Guckenberger, A. Richter, J. Boda-Heggemann, and F. Lohr. Motion compensation in radiotherapy. *Crit Rev Biomed Eng.*, 40(3):187–197, 2012.

[26] B.K.P. Horn and B.G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[27] S. Roujol, M. Ries, C. T. W. Moonen, and B. Denis de Senneville. Automatic non-rigid calibration of image registration for real time MR-guided HIFU ablations of mobile organs. *IEEE Transaction on Medical Imaging*, 30(10):1737–1745, 2011.

[28] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J Math Imaging Vis*, 40(1):120–145, 2011.

[29] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. *Computer Vision-ECCV 2004*, pages 25–36, 2004.

[30] B. Denis de Senneville, A. El Hamidi, and C. Moonen. A direct PCA-based approach for real-time description of physiological organ deformations. *IEEE Transaction on Medical Imaging*, 34(4):974–982, 2015.

[31] A. Wedel, T. Pock, C. Zach, H. Bischof, and D. Cremers. An improved algorithm for TV-L1 optical flow. In *Statistical and geometrical approaches to visual motion analysis*, volume 5604 of *Lecture notes in computer science*, pages 23–45. Springer Berlin Heidelberg, 2009.

[32] L.L. Rakêt, L. Roholm, M. Nielsen, and F. Lauze. TV-L1 optical flow for vector values images. In *Energy minimization methods in computer vision and pattern recognition*, volume 6819 of *Lecture notes in computer science*, pages 329–343. Springer Berlin Heidelberg, 2011.

[33] E. d'Angelo, J. Paratte, G. Puy, and P. Vandergheynst. Fast TV-L1 optical flow for interactivity. In *In: IEEE International Conference on Image Processing (ICIP11*, pages 1925–1928, 2011.

[34] Y. Seppenwoolde, H. Shirato, K. Kitamura, S. Shimizu, M. van Herk, J. V. Lebesque, and K. Miyasaka. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int. J. Radiation Oncology Biol. Phys.*, 53(4):822–834, 2002.

[35] B. Stemkens, R. H. N. Tijssen, B. Denis de Senneville, H. D. Heerkens, M. van Vulpen, J. J. W. Lagendijk, and C. A. T. van den Berg. Optimizing 4-dimensional magnetic resonance imaging data sampling for respiratory motion analysis of pancreating tumors. *Int. J. Radiation Oncology Biol. Phys.*, 91(3):571–578, 2015.

[36] D. A. Feinberg, D. Giese, D. A. Bongers, S. Ramanna, M. Zaitsev, M. Markl, and Günther M. Hybrid ultrasound MRI for improved cardiac imaging and real-time respiration control. *Magn. Reson. Med.*, 63(2):2290–296, 2010.

[37] M. O. Köhler, B. Denis de Senneville, B. Quesson, C. T. Moonen, and Ries M. Spectrally selective pencil-beam navigator for motion compensation of MR-guided high-intensity focused ultrasound therapy of abdominal organs. *Magn. Reson. Med.*, 66(1):102–111, 2011.

[38] M. Glitzner, B. Denis de Senneville, J. Lagendijk, B. Raaymakers, and S. Crijns. Real-time 3D anatomy estimation from undersampled MR acquisitions. *Med. Phys.*, 42(6):3320–3320, 2015.

[39] L. Brix, S. Ringgaard, T. S. Sorensen, and P. R. Poulsen. Three-dimensional liver motion tracking using real-time two-dimensional MRI. *Med. Phys.*, 41(4):042302:1–10, 2014.

[40] P. Arnold, F. Preiswerk, B. Fasel, R. Salomir, K. SCheffler, and P. Cattin. 3D organ motion prediction for MR-guided high intensity focused ultrasound. *Proc. MICCAI*, 14(2):623–630, 2011.

[41] A. Chambolle. An algorithm for total variation minimization and applications. *J Math Imaging Vis*, 20(1-2):89–77, 2004.

[42] A. E. Lujan, J. M. Balter, and R. K. Haken. A method for incorporating organ motion due to breathing into 3D dose calculations in the liver: sensitivity to variations in motion. *Med. Phys.*, 30(10):2643–2649, 2003.

[43] J. H. Killoran, V. H. Gerbaudo, M. Mamede, D. Ionascu, S. J. Park, and R. Berbeco. Motion artifacts occuring at the lung/diaphragm interface using 4D CT attenuation correction of 4D PET scans. *J. Appl. Clin. Med. Phys.*, 12(4):261–274, 2011.

CHAPTER 3

An Adaptive Non - Local - Means Filter for Real-Time MR - Thermometry

"Fast is fine, but accuracy is everything."
- Xenophon

This chapter is based on:
C. Zachiu, M. Ries, C. Moonen, and B. Denis de Senneville. An Adaptive Non-Local-Means Filter for Real-Time MR-Thermometry. *IEEE Trans. Med. Imaging*, 36(4): 904 - 916, 2017.

Abstract

Proton resonance frequency shift-based magnetic resonance thermometry is a currently used technique for monitoring temperature during targeted thermal therapies. However, in order to provide temperature updates with very short latency times, fast MR acquisition schemes are usually employed, which in turn might lead to noisy temperature measurements. This will, in general, have a direct impact on therapy control and endpoint detection. In the current study we address this problem through an improved non-local filtering technique applied on the temperature images. Compared to previous non-local filtering methods, the proposed approach takes into account not only spatial information, but also exploits temporal redundancies. The method is fully automatic and designed to improve the precision of the temperature measurements while at the same time maintaining output accuracy. Additionally, the implementation was optimized in order to ensure real-time availability of the temperature measurements while having a minimal impact on latency. The method was validated in three complementary experiments: a simulation, an *ex-vivo* and an *in-vivo* study. Compared to the original non-local means filter and two other previously employed temperature filtering methods, the proposed approach shows considerable improvement in both accuracy and precision of the filtered data. Together with the low computational demands of the numerical scheme, the proposed filtering technique shows great potential for improving temperature measurements during real-time MR thermometry dedicated to targeted thermal therapies.

3.1 Introduction

Magnetic resonance (MR) thermometry using proton resonance frequency shifts (PRFS) is one of the currently used mechanisms for temperature monitoring during thermal ablations of pathologies inside the human body.[1–3] However, such interventions typically require a good spatial coverage of the heated region and/or temperature updates with a high temporal resolution.[4,5] For this reason, fast MR-imaging sequences are in general necessary, which might lead to an acquisition having a low signal-to-noise ratio (SNR), resulting in noisy temperature measurements.[6,7] This is expected to have a direct impact on therapy control and endpoint detection, which might lead to the over or under-treatment of the pathology. In the current study we address this issue through a non-local image filtering technique as a means to improve the SNR of the PRFS-based temperature measurements.

Image denoising through non-local means (NLM) was originally proposed by

Buades *et al.* in the context of SNR restoration for digital images and video sequences altered by white Gaussian noise.[8] A pixel denoised with the NLM method is basically a weighted average of all the pixels in the noisy image. The weights themselves are proportional to the gray-level intensity similarity between the local neighborhoods of the pixel under study and the pixel contributing to the average: the more similar the two neighborhoods are, the larger the weight, hence the greater the impact a pixel has on the resulting denoised value. Since the most important contributions in denoising a pixel are provided by pixels that are most similar to it, the NLM filter is more prone to preserve high-frequency image content such as edges and/or small structures, compared to other denoising techniques.[8] However, the fact that the NLM method is based on spatial similarities is also the source of one of its major drawbacks. In case a particular structure/pixel lacks self-similarity in an image, it might end-up being insufficiently/poorly denoised. Such is the case, for example, of the focal point and its surroundings in the temperature images acquired during targeted thermal ablations. The focal point in particular has a high chance of being unique in the image. In its attempt to denoise the temperature at the focal point, the NLM filter, due to its functioning principle, will perform a weighted average between dissimilar pixels, which might result in a blurred/biased measurement. In addition, since the filter performs on the quantification of the image (i.e the color domain) and not on its sampling support (i.e the pixel domain), the filtering results fully depend on the image content. Therefore, contrary to other existing filtering approaches (such as convolution-based or (in)finite-impulse response techniques), a theoretical characterization of the overall filtered results becomes, from a mathematical point of view, a challenging task. It is thus difficult to evaluate/estimate the inherent resulting bias on the filtered data. Although this consideration is of minor interest for digital photography or video sequences, it becomes of major importance in the field of medical imaging, for which both the accuracy and precision of the quantitative image information are important.

Previous studies have addressed the issue of filtering PRFS-thermometry images through so-called predictive methods. For example, Roujol *et al.*[9] proposed a temperature filtering/prediction technique relying on the model-based Kalman filter.[10] The output of the filter was a linear combination of the measured noisy data and a temperature evolution model (the bioheat transfer equation (BHTE)). While showing great filtering capabilities for high-frame-rate PRFS thermometry (>10 images/s), the performance of the method is expected to deteriorate as the temperature update frequency decreases. Nevertheless, an NLM filter should be seen as complementary to such an approach rather than a replacement. Depending on the application, one might out-perform the other.

In this chapter we propose an improved NLM filtering method for dynamic PRFS thermometry during targeted thermal therapies. Performance analysis of the method was conducted using simulations, *ex-vivo* data acquired during a radiofrequency (RF) thermal ablation experiment on a calf liver and *in-vivo* data acquired on a porcine kidney during an MR-guided high intensity focused

ultrasound (MRg-HIFU) ablative sonication. The contribution of this study is fourfold:

1. We propose denoising PRFS temperature maps using a variation of the NLM filter which during the restoration process takes into account not only spatial information, but also exploits temporal redundancies. More precisely, each new incoming temperature map is also filtered based on information available from previous measurements. This is expected to improve the resulting filtered data compared to the original approach, especially in the focal point and its surroundings, since more reliable temperature information in this area is available in the temporal dimension rather than in the spatial one.

2. The proposed spatio-temporal NLM filter is further enhanced by dynamically adapting its filtration strength, through an optimization process, such that the accuracy of the resulting temperature measurements is maximized (i.e. the blur/bias on the filtered temperature is minimized).

3. The performance of the NLM filter can usually be improved if prior knowledge on the noise distribution in the image is available.[8,11,12] Therefore, in the current study, we propose a fully automatic method that estimates 'on-the-fly' the amount of noise altering the temperature measurements.

4. An implementation with real-time capabilities was performed for the proposed methods in order to provide improved temperature measurements while having a minimal impact on latency. For a smooth MR temperature monitoring work-flow, the computational time of the methods has to lie well beneath the time interval between two temperature measurements.

3.2 Method description

Initially, section 3.2.1 provides a brief mathematical description of the basic NLM filter together with the manner in which it can be applied for denoising MR thermal maps. Section 3.2.2 details the proposed improvements to the NLM filter which are intended to compensate for the shortcomings of the original version in the context of dynamic MR-thermometry. The section also includes a description of the proposed method which allows dynamic characterization of the noise distribution in the temperature measurements. Finally, section 3.2.3 describes the protocol used to evaluate the performance of the proposed variations to the original NLM filter and that of the noise characterization method, together with implementation details of the proposed methods.

3.2.1 The non-local means filter applied to thermal maps

The NLM filter proposed in this study used as a starting point the formulation suggested by Coupé *et al.* in.[11] Compared to the original version of Buades,[8]

Coupé *et al.* proposed an improved approach in which the filtration strength is computed in a deterministic way rather than empirically. While their study focused on 3D diagnostic brain images, the current technical limitations of MR systems and the computational time of NLM filters, notorious for their high processing latency, make real-time 3D PRFS-thermometry paired with a non-local filtering method a challenging task. For this reason, the current study was conducted solely for 2D temperature maps. According to the formulation proposed by Coupé *et al.*, the filtered temperature $T_f(\vec{r_i})$ at pixel $\vec{r_i} = (x_i, y_i)$ is given by the weighted average of all the noisy measurements in the temperature image T_m (principle which is graphically summarized in Fig. 3.1):

$$T_f\left(\vec{r_i}\right) = \sum_{j \in \Omega} w\left(\vec{r_i}, \vec{r_j}\right) T_m\left(\vec{r_j}\right) \tag{3.1}$$

with:

$$w\left(\vec{r_i}, \vec{r_j}\right) = \frac{1}{Z_i} e^{-\frac{\|T_m(N_i) - T_m(N_j)\|^2}{2\beta\hat{\sigma}(\vec{r_i})^2 |N_i|}} \tag{3.2}$$

where Ω is the image grid, Z_i is a normalization constant ensuring that $\sum_j w(\vec{r_i}, \vec{r_j}) = 1$, $|N_i|$ is the size of a local neighborhood around $\vec{r_i}$, $\hat{\sigma}(\vec{r_i})$ is an estimation of the local standard deviation of the noise and β is a constant that needs to be manually tuned. As mentioned in,[11] for Gaussian noise, if the estimation of $\hat{\sigma}$ is correct, the value of β is theoretically close to 1. However, the automatic dynamic adjustment of β will be the keystone of the adaptive NLM strategy proposed in this study. Additionally, note that, while in theory the filtered value is a weighted average of all the noisy measurements in the image, in practice the contributing pixels are limited to a smaller window around the filtered pixel.

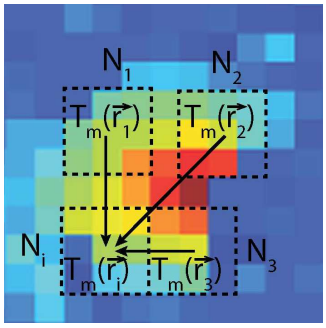


Figure 3.1: Noisy temperature map acquired during a targeted thermal therapy upon which the principle of the NLM filter is illustrated: The denoised value of the pixel at $\vec{r_i}$ is the weighted average of the pixels at $\vec{r_j}$, with the weights computed based on the similarity between their neighborhoods N_i and N_j. For simplicity, the process is illustrated only for three values of j, rather than the entire image domain Ω.

3.2.2 Improvements of the non-local means filter

3.2.2.1 Extension of the NLM filter to dynamic MRI

Instead of relying solely on spatial information in order to filter the noisy temperature maps, in this chapter we propose an extension of the NLM technique also to include the temporal dimension. In effect, a filtered pixel is the weighted average of the noisy measurements not only from the current dynamic, but also from the previous N dynamics. This results in what we have called the NLM-(2D+t) filter, having the following mathematical formulation:

$$T_f\left(\vec{r_i},t\right)=\sum_{n=0}^{N-1}\sum_{j\in\Omega}w\left(\vec{r_i},\vec{r_j},t-n\right)T_m\left(\vec{r_j},t-n\right) \tag{3.3}$$

with:

$$w\left(\vec{r_i},\vec{r_j},t-n\right)=\frac{1}{Z_i}e^{-\frac{\|T_m(N_i,t)-T_m(N_j,t-n)\|^2}{2\beta\hat{\sigma}(\vec{r_i})^2|N_i|}} \tag{3.4}$$

where t is the temporal index of the current dynamic.

3.2.2.2 Adaptive NLM strategy

Due to its underlying principle (i.e. averaging in the color domain), the NLM filter will unavoidably introduce a certain amount of blurring in the filtered data, which might lead to inaccurate temperature measurements. The filtering strength of the NLM is mainly controlled by the denominator in the exponential defined by Eq. (3.2). In particular, the β factor plays the most important role in this regard, since the other parameters are usually fixed. For example, while $\beta=0$ implies that no filtering is performed, a value of β equal to 1 results in filtering performed assuming a standard deviation equal to σ. In this chapter, we propose an optimization scheme for β such that for every pixel of each new incoming temperature map, the filter ensures a maximal output accuracy.
Filter accuracy was evaluated on a pixel-by-pixel basis by computing, for each new incoming image, the temperature bias $\epsilon\left(\vec{r}\right)$ over a temporal window. At instant t, the bias at pixel of coordinates $\vec{r}$ was calculated as follows:

$$\epsilon\left(\vec{r}\right)=\frac{1}{N}\sum_{n=0}^{N-1}\left[T_m(\vec{r},t-n)-T_f(\vec{r},t-n)\right] \tag{3.5}$$

where $T_m(\vec{r},t)$ is the measured (noisy) temperature at location $\vec{r}$ and instant t and $T_f(\vec{r})$ is the filtered output. For ideally filtered thermal maps, $\epsilon\left(\vec{r}\right)$ is a Gaussian centered random variable of standard deviation:

$$\sigma_e(\vec{r})=\frac{1}{N}\sqrt{\sum_{n=0}^{N-1}\hat{\sigma}(\vec{r},t-n)^2} \tag{3.6}$$

In such an ideal case, 95% of the realizations of $\epsilon(\vec{r})$ comply with the following inequality:

$$|\epsilon(\vec{r})| \leq 2\sigma_e(\vec{r}) \tag{3.7}$$

The value $2\sigma_e(\vec{r})$ could thus be taken as an upper bound for $|\epsilon(\vec{r})|$ in order to warrant the output accuracy. For each temperature map, an exhaustive enumeration of β was performed between 0 and 1 with an increment of $\frac{1}{10}$. The resulting filtered data and the associated error maps $\epsilon(\vec{r})$ were stored in a map data structure, for which the entry key was the corresponding tested β value. The optimal value(s) for $\beta(\vec{r})$ was(were) obtained for each pixel individually as follows:

$$\hat{\beta}(\vec{r}) = \underset{\{\beta, |\epsilon(\vec{r})| \leq 2\sigma_e(\vec{r})\}}{\arg\max} |\epsilon(\vec{r})| \tag{3.8}$$

The final filtered temperature map was reconstructed on a pixel-by-pixel basis using the previously stored data corresponding to the values of $\hat{\beta}(\vec{r})$. Intuitively, the proposed adaptive NLM method can be summarized as follows: Search for the value of $\beta(\vec{r})$ which leads to a maximal amount of filtering (denoted by the "arg max" in Eq. (3.8)), while minimizing accuracy losses (imposed by the constraint illustrated in Eq. (3.7)). For the remainder of the chapter we will refer to this filter as the adaptive NLM-(2D+t), abbreviated as ANLM-(2D+t).

3.2.2.3 Dynamic characterization of the temperature noise distribution

As illustrated by Eq. (3.2), the NLM filter requires prior knowledge of the noise distribution in the image. Please note that for a zero-mean Gaussian noise, as assumed in the current study, the noise distribution is fully characterized by its standard deviation. For this reason, throughout this chapter we will use the terms of noise distribution and standard deviation interchangeably. A number of methods have already been proposed to estimate the noise distribution directly from the image itself. However, for the particular case of temperature images, the estimation process might be hampered by several factors such as pixels with low magnitude/signal or phase artifacts due to susceptibility/temporal drift of the magnetic field.[13–15] Since the magnitude component of the MR-signal is less prone to the such artifacts, we decided to use the latter to estimate the noise distribution in the temperature maps. This was achieved by first estimating the standard deviation of the noise in the phase image by making use of the formula proposed by Conturo in:[13]

$$\sigma(\phi(\vec{r})) = \frac{\sigma(M)}{|M(\vec{r})|} \tag{3.9}$$

where ϕ and M are the phase and the magnitude of the complex MR-signal and $\vec{r}$ is the spatial location. Based on the noise distribution estimated in the phase image, the standard deviation of the noise in the temperature maps was then

computed based on the following relationship between phase and temperature:

$$\sigma(T(\vec{r})) = \sqrt{2}\frac{\sigma(\phi(\vec{r}))}{k} \tag{3.10}$$

where k is the PRFS-thermometry constant.[3]
The standard deviation of the noise in the magnitude image can be estimated through several methods, depending on whether the noise is assumed to be homogenous[11,16] or spatially variant.[12] In the current study, since the noise was assumed to be homogeneously distributed over the field of view, the standard deviation of the noise in the magnitude image was estimated via the wavelet-based approach proposed in:[16]

$$\sigma_{\text{est}}(M) = \frac{\text{median}|y_i|}{0.6745} \tag{3.11}$$

where $\sigma_{\text{est}}(M)$ is the estimated value of the noise standard deviation in the magnitude image and y_i are the high frequency coefficients of the first level wavelet decomposition of the image. The method was chosen due to the good compromise between computational time versus quality of the estimation it has shown during our experiments. Note that since the standard deviation of the noise in the magnitude image is not expected to vary significantly over time, the estimation in Eq. (3.11) can be carried-out only for one image in the thermometry sequence.
For the rest of the chapter, unless specified otherwise, σ refers to the standard deviation of the noise in the temperature maps.

3.2.3 Experimental setup

3.2.3.1 Performance assessment of the proposed filtering methods

The following filtering techniques were compared in the context of dynamic MR-thermometry during a targeted thermal therapy:

- An infinite-impulse-response (IIR) temporal low-pass Cauer filter of the 5^{th} order. The pass-band and stop-band were adjusted to result in an overall temporal resolution divided by a factor 2.
- A 2D spatial Gaussian filter (kernel size equal to 3×3, σ=0.5).
- The 2D spatial NLM filter proposed by Coupé *et al.* in.[11] For the remainder of the chapter we will refer to this filter as the NLM-2D.
- The proposed NLM-(2D+t) filter.
- The proposed ANLM-(2D+t) filter.

The first two methods were selected as they represent two previously employed strategies for real-time dynamic MRI-thermometry.[17–19] It is expected that the two will provide a similar SNR improvement by a factor $\sqrt{2}$, since each of them imply a reduction of the resolution by a factor 2, one in the temporal and the other in the spatial domain. The other 3 methods have the purpose to evidence the improvements provided by the proposed NLM filtering methods compared to the existing one. For all the NLM filters, the patch sizes $|N_i|$ and $|N_j|$ were chosen equal to 25 (5×5). Additionally, each pixel was filtered using the contribution of the noisy measurements situated within a 7×7 encompassing window. The size of the latter and also the patch size were chosen as to provide the best compromise between quality of the filtered data and computational time. Following the same reasoning, the temporal window N (see Eq. (3.3)) was chosen equal to 6.
Since the main concern was output accuracy, the performance of the tested filters was assessed quantitatively by evaluating the time averaged absolute difference between the output of the filters and a gold standard. The construction of the latter will be further detailed in future sections. The performance assessment was conducted in two different regions:

- In a region of interest covering 7×7 pixels, centered on the focal point position, over the entire duration of the thermal energy deposition. In this manner, the assessment is performed within an area excluding background pixels which are not of great interest.

- At the focal point itself (i.e in a single pixel located at the focal point position).

The performance of the filters was evaluated in three complementary studies:

a Simulation study Temperature maps resulting from a HIFU energy delivery were simulated using the BHTE model. The simulation extended over a duration of 100 s and 250 W of acoustic power was applied in a single point during the [10 s; 60 s] time interval. The spatial and temporal evolution of the temperature resulting from the acoustic energy delivery was given by the solution of the BHTE, solved using finite elements on a $1 \times 1\text{mm}^2$ lattice and with a temporal resolution of 1 s. The simulated temperature maps were subsequently corrupted by additive white Gaussian noise with increasing standard deviation. The latter was varied between 0°C and 4°C with a regular sampling step of 0.5°C. The noisy temperature maps were then filtered using the approaches enumerated in section 3.2.3.1 and their performance was assessed based on the metric described therein. The gold standard was the simulated, unaltered temperature maps. Note that the assessment was averaged over 10 experiments with different realizations of noise, in order to reduce random impact on the overall results.

b *Ex-vivo* study A heating *ex-vivo* experiment was conducted on a calf liver. The heating itself was the result of delivering 20 W of RF power via a bipolar electrode, for a duration of ∼50 s. The energy delivery was carried-out under MR-guidance with MR-thermometry being performed before, during and after heating the tissue. The MR-acquisition was performed via a dual-shot gradient recalled echo-planar imaging (EPI) sequence with TE = 13 ms, TR = 70 ms, $N_{slices} = 1$, image size 128×128 and a 1.5×1.5×5 mm^3 voxel size, resulting in an image being acquired every 140 ms (∼7.15 Hz). The intense and quick heating induced by the RF energy delivery lead to notable local gray-level intensity drops in the magnitude image in the close vicinity of the heated region.
The purpose of this particular experiment was to analyze the impact of a noise distribution in the temperature maps with a spatio-temporally varying standard deviation. When heating tissues, relaxation times get longer and the magnitude of the MR-signal changes. According to Eq. (3.9) and (3.10) this leads to local spatio-temporal variations in the standard deviation of the noise in the MR-temperature maps. This effect might end-up having an impact on the quality of the filtered data provided by the proposed methods.
A high-SNR dataset was derived from the original sequence by a centered sliding temporal averaging applied on both the magnitude and the phase components of the MR signal, followed by a temporal subsampling, such that the output data had an equivalent temporal frequency of ∼1 Hz. This operation lead to an average SNR of the magnitude image of more than 40 (SNR computation was restricted to the liver). The MR temperature maps were then recalculated based on the averaged phase images and set as a goldstandard for the current experiment. In order to evaluate the performance of the compared denoising methods in this particular experiment, a noisy dataset was constructed based on the temporally averaged magnitude and phase images. The latter were used to recalculate the real and imaginary channels of the MR signal, upon which a Gaussian noise with a standard deviation of 10% of the maximum intensity in the magnitude image was then added. The noisy real and imaginary parts were then re-combined to compute the noisy magnitude and phase images, after which a set of noisy temperature maps was constructed, based on the noisy phase images. The denoising methods were then applied on the resulting temperature maps and the filtered data was compared in terms of accuracy against the goldstandard.

c *In-vivo* study An MRg-HIFU shot was performed *in-vivo* on a porcine kidney, which consisted in the delivery of 250 W of acoustic power over a duration of 50 s. The animal was under general anesthesia, with the targeted area only exhibiting minimal motion. Thus no correction of motion related temperature errors was required. The MR-thermometry acquisition protocol consisted of a multi-slice image, with each slice being acquired via a slice-selective single-shot echo planar imaging sequence employing the following parameters: TE=25

ms, TR=1143 ms, flip angle=35°, FOV=142.5 × 285 mm^2, bandwidth$_{\text{read}}$ = 2078 Hz and voxel size=3 × 3 × 6mm^3. Each dynamic included a total of 9 coronal slices, symmetrically distributed around the focal spot. This would ensure the coverage of the entire kidney and part of the near-field, in order to detect any secondary heating spots. A total of 175 dynamics were acquired over a duration of 200 s, providing temperature information during pre-sonication, the energy delivery itself and respectively post-sonication. The temperature measurements provided by the middle slice (which included the primary focal spot) were filtered using the approaches enumerated in section 3.2.3.1 and their performances were compared based on the metric described therein.
The gold standard was obtained as follows: A pixelwise temporal centered sliding average was performed on the temperature data provided by the middle slice with a span extending over six temperature maps followed by a 6-fold temporal subsampling. This increased the SNR by a factor 2.45 and resulted in a dynamic sequence with a temporal resolution equal to 0.15 Hz. Note that the original noisy image sequence was also temporally subsampled such that it's frequency matches the one of the gold standard. Subsequently, a Mann-Whitney test was carried out in order to study whether the differences between the noisy and the filtered data are statistically significant. A significance threshold of $p = 0.05$ was used.
The animal experiment was conducted in agreement with the European law on animal experimentation and in compliance with the institution's rules for animal care and use.

3.2.3.2 Performance analysis of the noise estimation algorithm

A high-SNR image was initially computed by temporally averaging all the corresponding magnitude images acquired with the protocol described in section c. After normalizing the resulting image to the $[0, 1]$ interval, increasing levels of Gaussian noise were added to it, progressively lowering its SNR. The standard deviation of the added noise was varied between 0.005 and 0.06 with an increment of 0.005. Note that, according to Eq. (3.9) and (3.10), this leads to a standard deviation in the temperature maps that varies between $\sim 0.5°C$ and $\sim 4.5°C$. For each noise level, the standard deviation was then estimated in the magnitude images using the algorithm specified by section 3.2.2.3. The estimated value was then compared to the true standard deviation in terms of the percentage relative error (denoted by δ_σ):

$$\delta_\sigma = 100 \times \left| 1 - \frac{\hat{\sigma}}{\sigma} \right| \tag{3.12}$$

where $\hat{\sigma}$ and σ are the estimated and respectively the true value of the noise standard deviation added to the magnitude images. Subsequently, the percentage relative estimation errors versus the SNR of the image was analyzed. Note that the estimation errors were averaged over 1000 realizations of each noise

level. Note that assessment and processing was performed exclusively for the slice containing the primary focal spot.

3.2.3.3 Implementation

A C++ multi-threaded implementation (8 threads) was performed for all filtering methods enumerated in section 3.2.3.1 on an Intel 3.2 GHz i7 workstation (8 cores) with 16 GB of RAM. In the scope of this study, the filtering process was restricted to a region of interest around the focal point. The size of the region was chosen in accordance with the spatial extent of the heating. This was decided both for computational time considerations and, in the scope of this study, for the lack of interest in the pixels outside the heated area.

3.3 Results

Section 3.3.1 illustrates the performance during the simulation study (see section a for details) of the filtering methods enumerated in 3.2.3.1 including: an example of the spatial distribution of the accuracy errors introduced by each of the methods, the performance of the filters for increasing noise levels and the sensitivity of the proposed ANLM-(2D+t) method to errors of the noise estimation algorithm. Section 3.3.2 displays the results obtained during the *ex-vivo* experiment described in section b. Of interest here were the accuracy errors introduced by the compared methods in the presence of heating-induced spatio-temporal SNR variations in the MR magnitude images. Finally, section 3.3.3 illustrates the performance of the filters for data obtained during the *in-vivo* experiment (detailed in section c) together with an assessment of the noise estimation algorithm (see sections 3.2.2.3 and 3.2.3.2) and the computational requirements of the proposed methods.

3.3.1 Simulation study

3.3.1.1 Performance assessment of the proposed NLM-(2D+t) and ANLM-(2D+t) filtering methods

Fig. 3.2 illustrates the temperature bias introduced by the compared filtering methods (enumerated in section 3.2.3.1), aiming to reflect the accuracy of the resulting filtered data. Fig. 3.2(a) shows the temperature spatial distribution 40 s into the heating simulation detailed in section a. This played the role of gold standard in comparing the performance of the filtering methods. Fig. 3.2(b) displays the gold standard temperature map altered by a Gaussian noise with a 2°C standard deviation, while Fig. 3.2(c)-(g) display the filtered temperature measurements provided by each approach. The resulting temporally averaged absolute errors are illustrated in Fig.3.2(h)-(m). From a visual inspection it already becomes apparent that, in the focal point and its close proximity, the NLM-(2D+t) and respectively the ANLM-(2D+t) methods provide the lowest

errors, observation that will be further confirmed by subsequent results.
Fig. 3.3 and 3.4 compare in terms of accuracy the proposed ANLM-(2D+t)

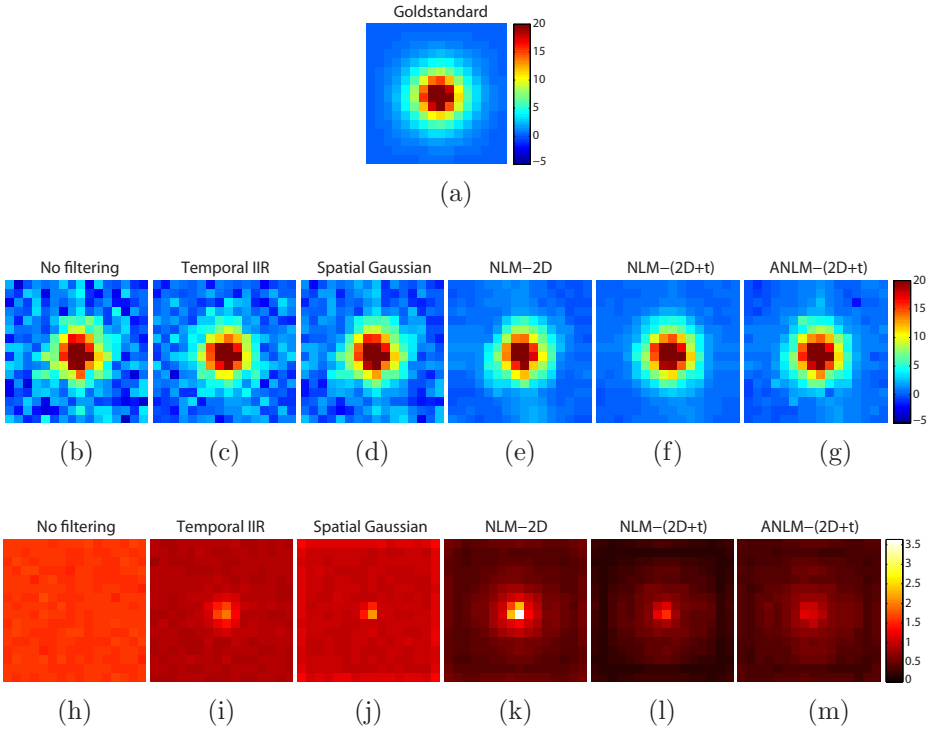


Figure 3.2: Example of thermometry results obtained during the simulation study. (a): The gold standard noise-free temperature map; (b): Corresponding noisy map (standard deviation $2°C$); (c)-(g): Filtered temperature maps; (h)-(m): Spatial distribution of the temporally averaged absolute errors for each filtering method.

against the IIR and respectively the Gaussian filter. Fig. 3.3 in particular illustrates the temperature evolution in the focal point over the duration of the simulation. Fig. 3.3(a) traces the gold standard temperature evolution in the focal point (black curve) overlapped with its noisy version (red curve). The noise standard deviation in this experiment was set to 3°C. The filtered data together with the gold standard is displayed in Fig. 3.3(b). It can be observed that the intrinsic temporal lag of the IIR filter, which becomes evident especially in the proximity of the starting and end point of the energy delivery, leads to a consistent amount of bias in the filtered data. A consistent amount of bias can also be observed for the Gaussian filter, which by construction has a blurring effect on the filtered measurements. A visual analysis of the traces in Fig. 3.3(b) would indicate that the ANLM-(2D+t) filter (blue curve) provides the

most accurate results. This is confirmed in Fig. 3.4, where the mean absolute error of the filtered temperature versus the simulated noise standard deviation is plotted. Note that analysis was performed for the same filters as in Fig. 3.3, but separately for the focal point (Fig. 3.4(a)) and the area surrounding the focal point (Fig. 3.4(b)). While the IIR (red line) and the spatial Gaussian (green line) filters manifest similar tendencies in terms of accuracy of the filtered data, the ANLM-(2D+t) outperforms both for all the analyzed noise levels, especially in the focus point for low noise levels. The black line corresponds to the mean absolute error of the noisy data. As a side-note, the comparable performance of the IIR and the spatial Gaussian filter can be explained by their similar approach in improving the SNR of the noisy image. They both reduce the information resolution by a factor of 2 (see section 3.2.3.1), one however operates in the temporal while the other in the spatial domain, in order to improve the SNR of the image.

Fig. 3.5 shows the accuracy of the temperature filtered using the ANLM-

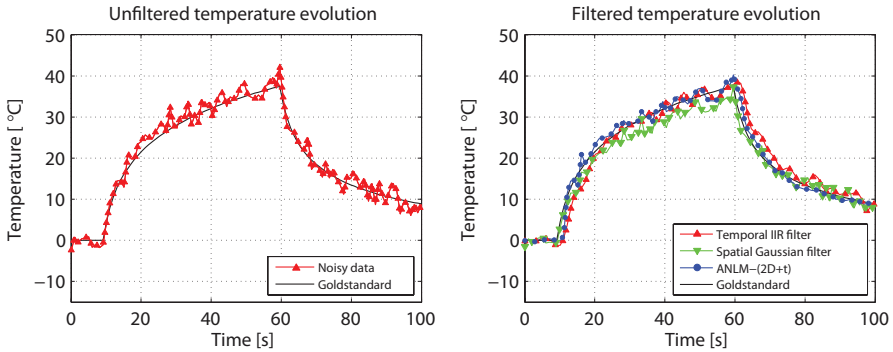
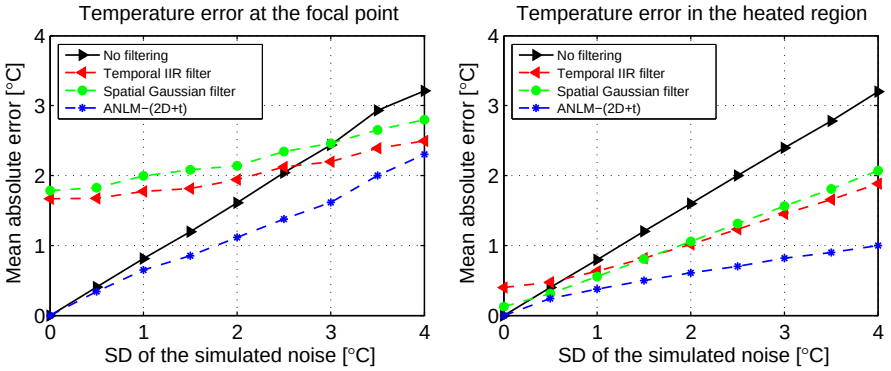


Figure 3.3: Temperature time curve obtained in a single pixel located at the focal point. (a): The noisy time temperature evolution (red curve, standard deviation $3°C$). (b): Filtered temperature values obtained using the proposed ANLM(2D+t) method (blue curve), the spatial Gaussian filter (green curve) and the temporal IIR filter (red curve). The gold standard temperature illustrated in (a) and (b) by the black curve.

(2D+t) (blue line) compared to the original NLM (red line) and to the non-adaptive NLM-(2D+t) (green line). The analysis was conducted for several noise levels in both the focal point (Fig. 3.5(a)) and the area around the focal point (Fig. 3.5(b)). Note that the blue curves corresponding to the ANLM-(2D+t) filter are identically the same as the ones as in Fig. 3.4(a) and 3.4(b) respectively. Except for very low noise levels ($< 0.5°C$), where all three filters have a similar performance, the original NLM is constantly outperformed by the other two filters. As already anticipated from Fig. 3.2(l) and 3.2(m), the



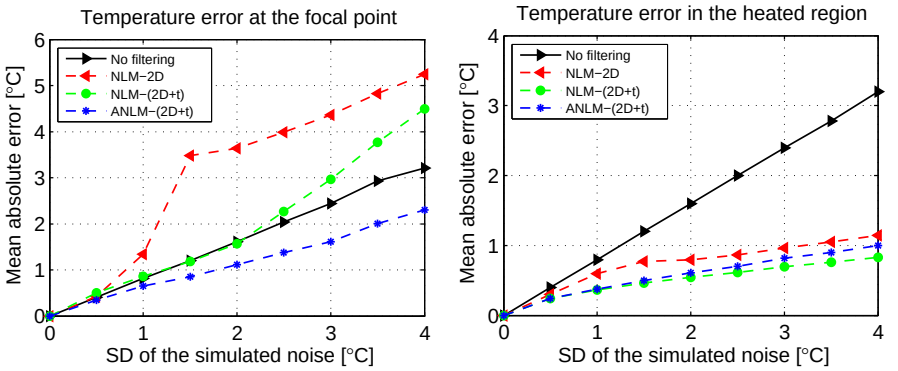
(a) (b)

Figure 3.4: Accuracy of the ANLM-(2D+t) filter (blue line) versus the IIR (red line) and respectively spatial Gaussian (green line) filter. The two figures illustrate the mean absolute errors as a function of the simulated noise standard deviation (a): In the focal point; (b): In the surrounding area of the focal point.

NLM-(2D+t) and the ANLM-(2D+t) have a similar performance, particularly if the errors are averaged over the area surrounding the focal point or if the noise levels are low. However, for a noise standard deviation larger than 2°C, the ANLM-(2D+t) clearly outperforms the NLM-(2D+t) when analysis is conducted in the focal point, especially in the rightmost part of the graph in Fig. 3.5(a). Note that the black line corresponds to the errors in the noisy data. A noteworthy fact that can be observed from both Fig. 3.4 and 3.5 is that the ANLM-(2D+t) filter does not yield higher errors than the noisy data for any of the analyzed noise levels. The rest of the filtering methods, on the other hand, at least in some particular cases, actually end-up altering the temperature measurements even further compared to when no filtering is performed.

3.3.1.2 Robustness of the ANLM-(2D+t) to calibration errors

As illustrated by Eq. (3.2) and (3.4), the NLM, NLM-(2D+t) and the ANLM-(2D+t) filters require the noise standard deviation as an input parameter, which has to be estimated (see section 3.2.2.3). Fig. 3.6 displays for the ANLM-(2D+t) in particular its sensitivity to noise estimation errors. The figure illustrates the mean absolute error of the filtered temperature as a function of the percentage relative estimation error of the noise standard deviation. The analysis was conducted in both the focal point (red curve) and the area around the focal point (blue curve) for temperature maps altered by a noise with a standard deviation of 1°C. It can be observed that an under-estimation of the noise, while the accuracy of the filtered data is still improved, precision is being



(a) (b)

Figure 3.5: Accuracy of the ANLM-(2D+t) filter (blue line) versus the original NLM (red line) and the NLM-(2D+t) (green line) filter. The mean absolute temperature errors are reported for various noise levels (a): In the focal point; (b): In the area surrounding the focal point.

hampered, indicated by the fact that the difference between the filtered and the noisy temperature becomes less evident. Note, however, that the filter still remains reliable for noise estimation errors of up to 50% in either direction.

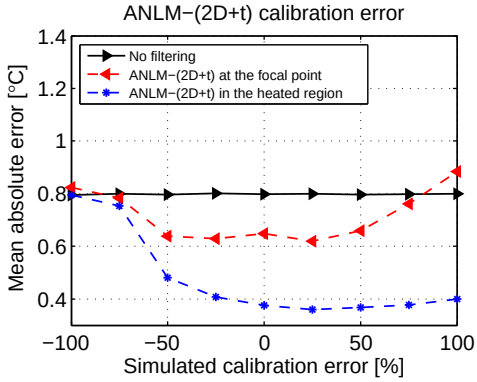


Figure 3.6: Robustness of the proposed ANLM-(2D+t) filter against potential calibration errors for a nominal noise standard deviation of $1°C$. The mean absolute temperature errors are reported for various relative noise estimation errors.

3.3.2 *Ex-vivo* study

Fig. 3.7 displays the performance of the analyzed filtering methods following the MR-guided *ex-vivo* RF thermal ablation study conducted on a calf liver (described in b). Fig. 3.7(a) in particular, illustrates one of the temporally averaged high-SNR MR magnitude images, calculated prior to starting the energy delivery (see section b for details). For better visibility, the treated area (identified by the red dashed square) has been magnified, while at the same time indicating through blue arrows several pixels that were of interest during the *ex-vivo* study. Points (1) and (2), for example, correspond to the approximate position of the tips of the RF-emitting needles, while points (3) and (4) were selected such that the former would be in close proximity to one of the focal spots while the latter is completely outside the heated area. The precise purpose of points (3) and (4) will be detailed in later paragraphs. Fig. 3.7(b) and 3.7(c) display the gold standard temperature distribution and its noisy version, 80 s into the experiment. Fig. 3.7(d)-3.7(h) showcase the effect each of the filtering methods have on the noisy temperature map from Fig. 3.7(c), while Fig.3.7(i)-3.7(n) illustrate the bias introduced in the filtered data by each of the approaches. Consistent with the observations made during the simulation study is the fact that both the NLM-(2D+t) and the ANLM-(2D+t) outperform the other filtering methods, with the ANLM-(2D+t) performing slightly better than the NLM-(2D+t), especially in the proximity of the focal spots. This is further confirmed by Fig. 3.8 where the temperature evolution during the *ex-vivo* experiment is traced for point (3) in the absence (Fig. 3.8(a)) and in the presence (Fig. 3.8(b)) of a filtering method. It can be observed from the latter that compared to the ANLM-(2D+t), the Gaussian and the IIR filter have indeed a tendency to introduce additional bias in the filtered data.

What sets the *ex-vivo* experiment apart from the simulation study is the fact that the temperature maps in the *ex-vivo* experiment are altered by a noise with a spatially variant standard deviation, which was constant during the simulation. The variation of the noise standard deviation stems mainly from the signal drops induced in the MR magnitude images by the heating process, factor which, for the sake of simplicity, was not taken into account during the simulation study. The local signal drops lead to an increase in the temperature standard deviation according to Eq. (3.9) and (3.10). Fig. 3.9 displays the temporal evolution of the gray-level intensity at points (3) and (4) respectively. It can be noted that, as the temperature at point (3) increases, the signal intensity decreases, while at point (4), which was chosen outside the heated area, the signal intensity remains more or less constant. Nevertheless, even in such conditions the proposed filtering methods outperform the existing ones, as shown in Fig. 3.7 and 3.8.

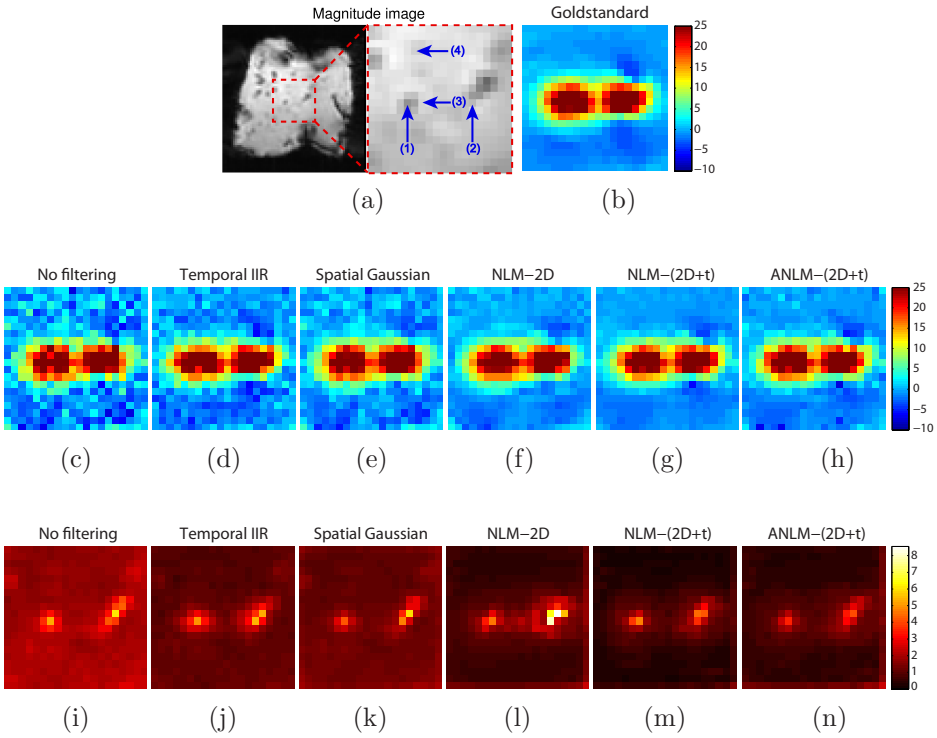


Figure 3.7: Performance of the analyzed filtering methods following the *ex-vivo* study. (a): High SNR MR magnitude computed prior to the energy delivery together with a magnification of the treated area (red dashed square) upon which several points of interest are indicated by blue arrows; (b), (c): The gold standard temperature map together with its noisy version; (d)-(h): Filtered temperature maps; (i)-(n): Spatial distribution of the temporally averaged absolute errors for each filtering method.

3.3.3 *In-vivo* study

3.3.3.1 Performance assessment of the proposed NLM-(2D+t) and ANLM-(2D+t) filtering methods

Fig. 3.10 illustrates the accuracy of each of the compared filtering methods (see section 3.2.3.1) when applied on the *in-vivo* dataset. Details concerning the *in-vivo* experiment can be found in section c. Fig. 3.10(a) and 3.10(b) correspond to the gold standard temperature map and its noisy version, respectively, after 50 s of ultrasonic energy delivery. Fig. 3.10(c)-(g) display the filtered temperature map using each of the compared methods. The accuracy of the filters is shown in Fig. 3.10(h)-(m), where the spatial distribution of the temporally

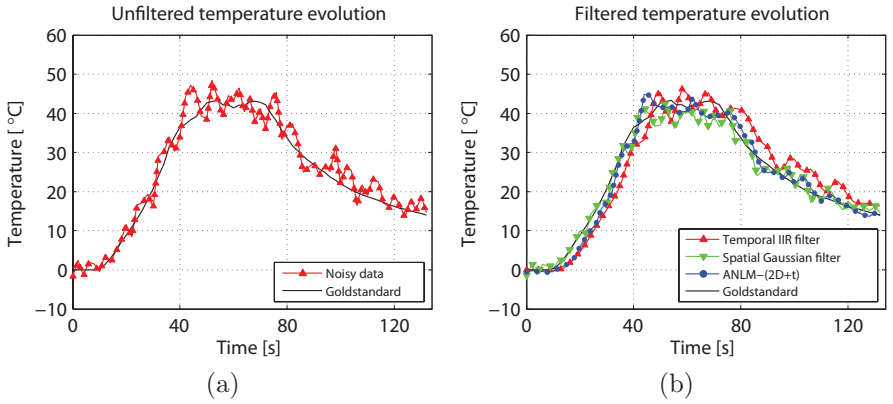


Figure 3.8: Temperature time curve at point (3) from Fig. 3.7(a). (a): The noisy time temperature evolution (red curve). (b): Filtered temperature values obtained using the proposed ANLM(2D+t) method (blue curve), the spatial Gaussian filter (green curve) and the temporal IIR filter (red curve), respectively. The gold standard temperature illustrated in (a) and (b) by the black curve.

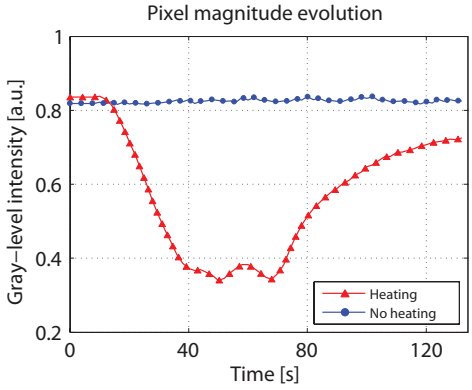


Figure 3.9: The effects of heating on the magnitude of the MR signal during the *ex-vivo* experiment: Temporal evolution of the signal intensity at points (3) (red curve) and (4) (blue curve) from Fig. 3.7(a).

averaged absolute errors is illustrated. Just as in the simulation and the *ex-vivo* experiments, a visual analysis of the results indicates that the NLM-(2D+t) and the ANLM-(2D+t) outperform the other approaches. Fig. 3.11(a) and 3.11(b) illustrate the temperature evolution in the focal point over the duration of the *in-vivo* experiment. The black and the red curves in Fig. 3.11(a) correspond to the gold standard and to the noisy temperature measurements, respectively.

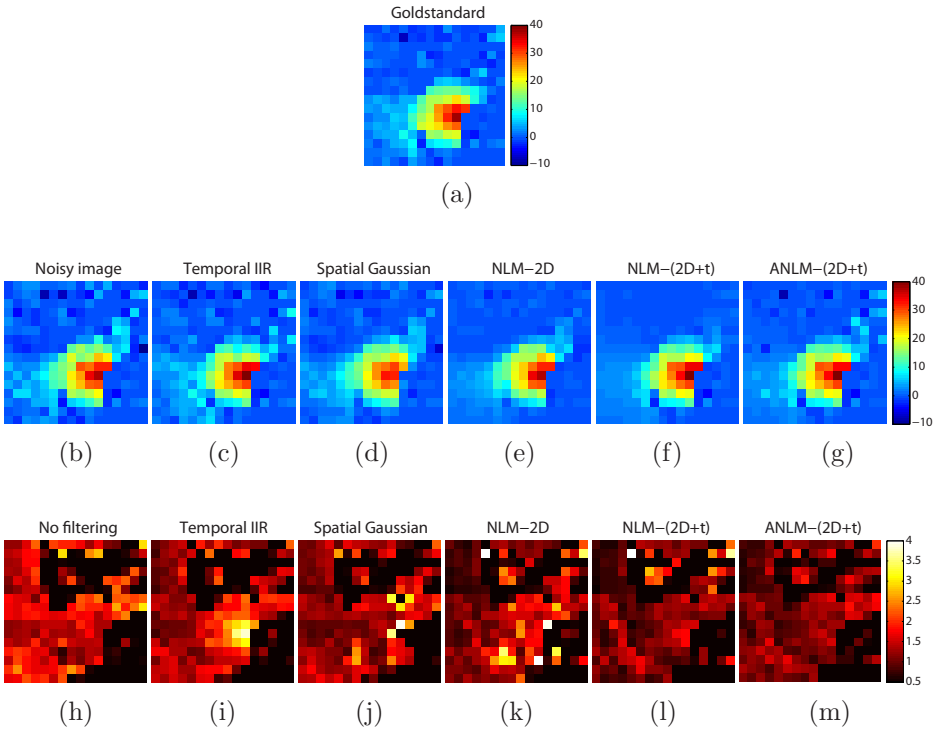
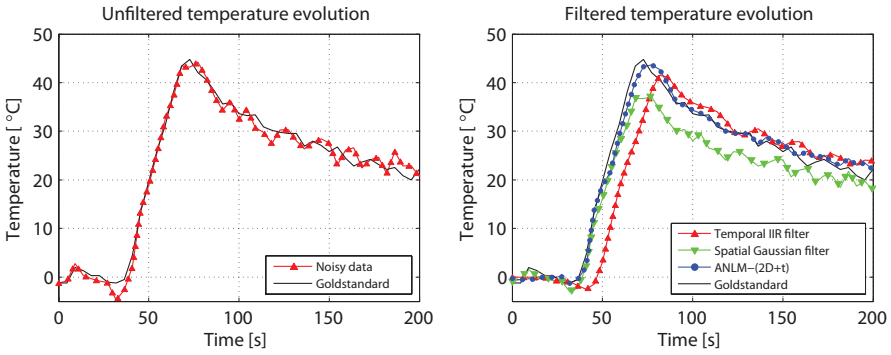


Figure 3.10: Accuracy of the compared filtering methods on the *in-vivo* porcine dataset. (a): The gold standard temperature map; (b): Noisy temperature map; (c)-(g): Filtered temperature maps as provided by the compared methods; (h)-(m): Spatial distribution of the temporally averaged absolute errors for each filtering method.

The filtered values are displayed in Fig. 3.11(b), where the red, green and blue curves correspond to the IIR, spatial Gaussian and the ANLM-(2D+t) filter, respectively. The black curve again represents the gold standard temperature. Similar to the results provided by the simulation and the *ex-vivo* experiment, the temporal lag of the IIR filter introduces consistent accuracy errors in the filtered data. The spatial Gaussian also biases the temperature measurements due to its intrinsic blurring effect. The ANLM-(2D+t) seems to provide the highest fidelity to the gold standard data.

3.3.3.2 Statistical analysis of the filtering errors

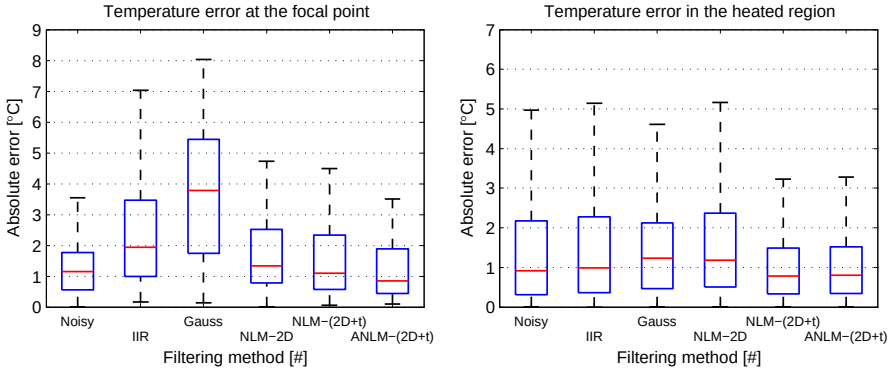
The pixel-wise absolute temperature errors from each acquisition time point were all pooled in a set, separately for each filtering method, and displayed as boxplots in Fig. 3.12. The analysis was performed both in the focal point



(a) (b)

Figure 3.11: Temperature time curve obtained in the focal point during the *in-vivo* experiment. (a): The noisy time temperature evolution (red curve). (b): Filtered temperature values obtained using the proposed ANLM-(2D+t) method (blue curve), the spatial Gaussian filter (green curve) and the temporal IIR filter (red curve). The gold standard temperature is illustrated in (a) and (b) by the black curve.

(Fig. 3.12(a)) and in a 7×7 area around the focal point (Fig. 3.12(b)). It can be observed that, compared to the other filters, the ANLM-(2D+t) and the NLM-(2D+t) provide the best performance in terms of both accuracy and precision. A marginal improvement in the focal point can be observed for the ANLM-(2D+t) over the NLM-(2D+t). Table 3.1 reports the p-values provided



(a) (b)

Figure 3.12: Statistical analysis of the absolute errors introduced by each filtering method for the *in-vivo* experiment (a): In the focal point; (b): In the area surrounding the focal point.

by a Mann-Whitney test applied on the data illustrated in Fig. 3.12. The test indicates whether the improvement or alteration induced by the filters, compared to the case when no filtering is applied, is statistically significant or not. The null hypothesis was that there are no statistically significant differences between the errors found in the noisy and respectively the filtered data. When analysis is conducted in the area around the focal point (right column), of interest are the entries associated to the NLM-(2D+t) and the ANLM-(2D+t) filters. Their p-values are beneath the threshold, confirming that the improvement due to filtration is statistically significant. In the focal point itself (left column), however, the differences between the aforementioned filters and the noisy data are not statistically significant. For the IIR and spatial Gaussian, even though the p-value confirms the alternate hypothesis, it is obvious from Fig. 3.12 that, at least in the focal point, they actually end up deteriorating rather than improving the data.

Filtering method	p-value at the focal point	p-value in the heated region
Temporal IIR filter	0.0155	4.9×10^{-5}
Spatial Gaussian filter	2.2×10^{-6}	3.7×10^{-12}
NLM-2D	0.26	1.3×10^{-18}
NLM-(2D+t)	0.8	9.8×10^{-6}
ANLM-(2D+t)	0.3	2.7×10^{-5}

Table 3.1: Statistical significance of the temperature error distributions obtained for each filtering method tested during the *in-vivo* experiment, as compared to the scenario when no filtering is performed.

3.3.3.3 Computational performance of the ANLM-(2D+t) algorithm

In terms of computational performance, our multi-threaded C++ implementation of the ANLM-(2D+t) filter required 311±10 ms per image, with processing times ranging from 297 ms to a maximum of 385 ms. Note that the reported values were obtained following a total of 1000 executions of the algorithm.

3.3.3.4 Assessment of the noise estimation algorithm

Figure 3.13 illustrates the percentage relative noise estimation errors versus the SNR of the noisy magnitude images. Details concerning the images and the computation of the noise estimation errors can be found in section 3.2.3.2. Noteworthy is the fact that as the SNR of the image increases, the noise estimation errors also have a tendency to increase, exceeding 50% for an SNR of 30.
Concerning computational demands, the algorithm required on average 16 ms per image, evaluated from a total of over 10000 estimations.

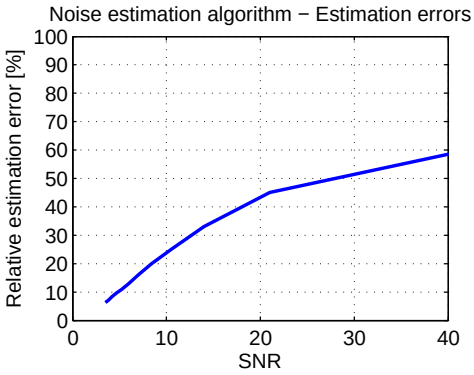


Figure 3.13: Percentage relative noise estimation errors versus the SNR of the magnitude images.

3.4 Discussion

This study proposes an improved non-local filtering method for real-time dynamic PRFS thermometry. The technique was designed using the NLM formulation proposed by Coupé *et al.* in[11] as a starting point, in addition to which, temporal information was also exploited in order to improve SNR. Moreover, an adaptation technique is proposed such that, the amount of filtration is dynamically optimized as to maximize output accuracy. Compared to previously employed filtration methods for real-time thermometry, such as the IIR[17] and the spatial Gaussian,[18,19] the proposed technique has shown improved performance in all tested scenarios. However, this was to be expected since, while the IIR and the Gaussian filters exploit either temporal or spatial information, respectively, the proposed NLM-(2D+t) and ANLM-(2D+t) exploit simultaneously both sources of information.
As specified throughout this chapter, the spatial Gauss filter and the IIR were included in the study due to their prior use for denoising temperature measurements during real-time MR-thermometry. While the exclusion of other filtering methods can be seen as somewhat of a limitation, a criterion had to be established when selecting the filtering methods to include. This is due to the fact that there is an entire branch of digital image processing dedicated to denoising methods, proposing a vast number of algorithms. Thus, it seemed a justified choice to only select methods that have been previously used in the context of MR-thermometry and allow the analysis of other filtering methods to be the object of future studies.
Performance comparison between the analyzed filtering methods was performed in terms of the temporally averaged absolute error, which quantifies the accuracy errors in the (un)filtered temperature measurements. Depending on the application, alternate criteria could be used such as, for example, a thermal

dose-based metric. However, the main purpose of this study was to propose a novel denoising method for real-time MR-thermometry in general. Thermal dose becomes a relevant factor only for applications involving temperatures above $\sim 43°C$,[20] which, for example, is rarely the case for near-field heating surveillance or low-power test-shots. Moreover, some approaches dedicated to quantifying thermal damage rely on detecting a certain temperature threshold rather than on thermal dose.[21,22] In such cases of interest are the precision and accuracy of the temperature measurements themselves. Thus, the impact of noise and filtering on other factors such as the thermal dose is a topic in itself and can be the subject of future studies, dedicated to more specific applications.

The original version of the NLM relies on the redundancies found in an image in order to improve its SNR. Since usually the temperature in the focal point is unique, it is natural to expect the underlying filtration errors to be large. This aspect is reflected very well throughout the majority of the results. This drawback is compensated, however, by using information regarding the temperature in focal point from the previous dynamics, as is the case of the NLM-(2D+t) and the ANLM-(2D+t). Additionally, the NLM filter recovers the signal in a pixel as the weighted average of pixels with similar neighborhoods. This will unavoidably introduce blurring effects, that lead to inaccuracies/biasing of the filtered data. The ANLM-(2D+t) circumvents this issue by optimizing the weights of the pixels used for signal recovery in such a way that accuracy of the filtered temperature is maximized. In the extreme case, when noise levels are low, a pixel might even be left unfiltered.

One of the key points in comparing the performance of the different filtering techniques in the current study was the construction/selection of a proper gold standard. This is a particularly challenging task during *in-vivo* experiments, since obtaining a completely noise-free temperature map is hardly feasible. In the present chapter temperature maps with only improved SNR were adopted as gold standard. Nevertheless, this was sufficient to quantify the performance of the proposed filtering methods and to show the statistical significance of the improvement over the existing ones (see Fig. 3.12 and Table 3.1).

Since temporal data is used in the filtration process, image sampling/acquisition frequency becomes an important issue. In this chapter, the simulation and the *ex-vivo* studies were carried out at a framerate of 1 Hz, while in the *in-vivo* experiment the sampling frequency was 0.15 Hz. Low MRI framerates become problematic especially for pure temporal filters such as the IIR evaluated in this study. As it can also be observed from Fig. 3.3, 3.8 and 3.11, the intrinsic temporal lag and the low temporal resolution will lead to considerable errors in the results provided by the IIR. It is expected that for low MRI framerates, due to low temporal redundancies, the proposed NLM-(2D+t) and the ANLM-(2D+t) will also suffer a decrease in performance. However, as seen in the results corresponding to the *in-vivo* experiments, the proposed filtering methods remain reliable even at frequencies as low as 0.15 Hz. This is an important aspect since, for example, several pre-clinical MRg-HIFU ablation protocols in

moving organs involve respiratory gating, resulting in imaging frequencies of 0.2 - 0.4 Hz.[23,24]

Another important aspect that needs to be taken into consideration when exploiting temporal information in the context of NLM filters is motion. In theory, when denoising a pixel via the NLM filter, the search for redundancies extends over the entire image. In such a case, motion is unlikely to have an impact on filter performance. In practice, however, due to computational considerations, a pixel is usually filtered based only on its neighbors within a particular search window (which is most of the times significantly smaller than the image size). Since the latter is usually fixed in space, motion might lead to similar/identical pixels exiting/entering the search window. Naturally, this is expected to affect filter performance. However, during dynamic MR-thermometry, the acquired images are often registered/aligned to a reference position in order to allow a proper analysis of the temperature evolution in each pixel (which is particularly important in applications where the thermal dose is of interest). Thus, when applied in such a context, motion is not expected to have an impact on the performance of the proposed NLM-(2D+t) and ANLM-(2D+t) filters.

One of the inputs required by the filtering methods proposed in this study is the standard deviation of the noise altering the images. The details concerning its estimation can be found in section 3.2.2.3. While the noise estimator provided good results for low-SNR images, the performance of the algorithm starts to deteriorate as the images become less noisy (see Figure 3.13). We hypothesize that this is due to the fact that the anatomical structures start to become more apparent in the MR-images and the gray-level transitions from one to the other start being interpreted as noise. However, as it can be observed from Figure 3.6, the proposed ANLM-(2D+t) filter remains reliable even for relative noise estimation errors of 50%. The curve in Figure 3.13 indicates that such errors only occur for an SNR higher than 30, which for the magnitude images provided by real-time thermometry is rarely the case. For example, given the MR acquisition sequence employed during the *in-vivo* experiment, a noise standard deviation of $1°C$ in the temperature images would lead, according to Eq. (3.9) and Eq. (3.10), to an SNR in the magnitude images equal to 10. This translates to a percentage relative noise estimation error of $\sim 20\%$ (see Fig. 3.13), value which is within acceptable margins, as previously noted. Any value of the noise standard deviation in the temperature images higher than $1°C$ will translate to an even lower SNR of the magnitude images, further reducing the uncertainty on the estimated noise distribution. Moreover, for low noise levels ($< 1°C$), while the relative estimation errors may be large, the absolute errors will be small. Thus, we anticipate that for a high SNR of the magnitude images (> 30), noise estimation errors of up to one order of magnitude will still have only a minimal impact on the performance of the filter.

In the current study, the noise altering the temperature maps was assumed to follow a Gaussian distribution. In theory, however, the noise distribution is somewhat more complex. PRFS temperature maps depend linearly on the difference between a current phase image and a reference acquired prior to heat-

ing. Thus, it is safe to assume that the noise distribution in the phase images and the temperature maps follow the same statistics. The MR phase image is computed as the inverse tangent of the ratio between the imaginary and the real part of the complex MR-signal. While the two channels are indeed subjected to Gaussian noise,[13] the resulting expression for the noise distribution in the phase image itself becomes fairly complicated. It was shown, however, by Gudbjartsson and Patz in[25] that in practice, as long as the SNR of the magnitude image remains above 3, the noise distribution in the phase images is approximately Gaussian. For MR-thermometry acquisition sequences, this is almost always the case in the heated region, since a sequence that provides a magnitude image with an SNR beneath 3 is practically unusable to begin with. Thus, our assumption that the noise in the temperature images follows a Gaussian distribution is justified. In addition, estimation of the noise standard deviation in the temperature maps was computed based on the result of a noise estimation algorithm applied on the magnitude image. The algorithm is built on the assumption that the noise follows a Gaussian distribution. Theoretically, the noise distribution in the MR magnitude image is Rician, however, it was shown in previous studies that for SNR values higher than 3, it is well approximated by a Gaussian distribution.[25,26] Nevertheless, these SNR limits need to be taken into consideration when denoising MR-temperature maps via the proposed methods, since beyond these boundaries no guarantees can be made on the quality of the filtered data. In case an application involves MR images with a severely low SNR, alternate methods may have to be considered.

Throughout the experiments performed in the current chapter, the noise distribution in the MR magnitude images (based on which the noise in the temperature measurements was calculated via Eq. (3.9) and (3.10)) was assumed to be spatially invariant. However, this hypothesis is likely to be violated in case MR-thermometry is performed using parallel imaging techniques such as SENSE or GRAPPA.[12] Depending on how the MR-signal received by the multi-coil system is re-combined/reconstructed, the noise standard deviation might end-up varying across the MR-magnitude image. Since the noise estimation method employed in our study operates on a global scale, this may lead to the proposed filtering methods to provide sub-optimal results. Thus, in case parallel acquisition is used, the global noise estimation method should be replaced with an approach that estimates the noise standard deviation in the MR-magnitude image on a pixel-by-pixel basis.

As we have shown, the proposed NLM filtering methods can be used on their own to improve the SNR of temperature maps. However, they also show great potential to be used in conjunction with other methods previously employed to improve/enhance PRFS thermometry. For example, as mentioned earlier in this chapter, the output of the Kalman predictor-based filter proposed by Roujol *et al.* in[9] is a linear combination of the measured noisy data and a temperature evolution model, which is obviously noise-free. The resulting filtered data, thus, unavoidably contains a certain amount of noise. We anticipate that the approach can be further enhanced by coupling the output of the Kalman filter

with the NLM methods proposed in the current study. Another instance in which the proposed filters could aid is in the context of the studies conducted by Todd *et al.* in.[27] They propose a method to accelerate MR-image acquisition, with the goal of rendering real-time 3D PRFS-thermometry feasible. While results have shown the great potential of the method, some limitations were encountered due to noisy measurements affecting the precision and accuracy of the resulting temperature data, which prevented further acceleration of the acquisition process. We again speculate that used in conjunction with one of the NLM filters illustrated in our study, the method could be improved.
An important aspect that needs to be taken into consideration is computational latency, i.e. the delay between the actual time of the measurement and the availability of the information. This includes the remaining acquisition time after the passage of the k-space center (equal to half the echo time), the data transport duration and the image processing time (which in the current study is the computational time of the ANLM-(2D+t) filter). Considering the MR-sequence and the hardware employed for the *in-vivo* study, half the echo time together with the data transport delay amount to $\sim$22.5 ms. With an average computational time of $\sim$300 ms for the ANLM-(2D+t) filter, the total latency reaches $\sim$322.5 ms per slice. In order to ensure a smooth temperature monitoring work-flow, this value must lie well beneath the time interval between two MR acquisitions. This constraint is fulfilled with a considerable margin for the application considered in this study, since respiratory gated MR-thermometry usually provides temperature updates with a frequency of 0.2 - 0.4 Hz (corresponding to the respiratory frequency). Note that the ANLM-(2D+t) filter can be even further accelerated by visiting the values of β (see Eq. (3.4)) in parallel, contrary to the exhaustive search performed in the current chapter. We anticipate that this would lead to an average processing time of the filter beneath 30 ms. Under such conditions the total latency would amount to $\sim$52.5 ms, allowing high-frame-rate MR-imaging. Additionally, an implementation of the ANLM-(2D+t) on a graphical processing unit is expected to provide even shorter computational times.

3.5 Conclusion

In this chapter we propose an improved non-local filtering technique for PRFS-based temperature measurements during real-time MR-guided targeted thermal therapies. Accurate and precise temperature maps are highly beneficial for proper therapy control and endpoint detection, potentially avoiding under or over-treatment of the targeted area. Compared to the original NLM filter and two other filtering methods previously employed during MR-thermometry, the proposed approaches have shown a considerable increase in both precision and accuracy. This together with their low computational demands, prove that the proposed filters are potentially beneficial as a means to improve temperature measurements during PRFS-based real-time MR thermometry dedicated to targeted thermal therapies.

Acknowledgements

This work was supported by the Dutch Technology Foundation (STW) (project OnTrack #12813) and in part by the European Research Council (project ERC-2010-AdG-20100317, Sound Pharma) and ITEA 2 (project SoRTS).

Bibliography

[1] Y. Ishihara, A. Calderon, H. Watanabe, K. Okamoto, Y. Suzuki, K. Kuroda, and Y. Suzuki. A precise and fast temperature mapping using water proton chemical shift. *Magn Reson Med*, 34(6):814–823, 1995.

[2] J. De Poorter, De Deene Y. De Wagter, C., Thomsen C., F. Stahlberg, and E. Achten. Noninvasive MRI thermometry with the proton resondance frequency (PRF) method: in vivo results in human muscle. *Magn Reson Med*, 33(1):74–81, 1995.

[3] V. Rieke and K. B. Pauly. MR-thermometry. *J Magn Reson Imaging*, 27(2):376–390, 2008.

[4] C. S. Mei, L. P. Panych, J. Yuan, N. J. McDannold, L. H Treat, Y. Jing, and B. Madore. Combining two-dimensional spatially selective RF excitation, parallel imaging, and UNFOLD for acceleraed MR thermometry imaging. *Magn Reson Med*, 66(1):112–122, 2011.

[5] J. Yuan, C. S. Mei, Panych L. P., McDannold N. J., and B. Madore. Towards fast and accurate temperature mapping with proton resonance frequency-based MR thermometry. *Quant Imaging Med Surg*, 2(1):21–32, 2012.

[6] Q. X. Yang, Demeure R. J., B. J. Dardzinski, B. W. Arnold, and M. B. Smith. Multiple echo frequency-domain image contrast: improved signal-to-noise ratio and T2(T2*) weighting. *Magn Reson Med*, 41(2):423–428, 1999.

[7] D. Erdogmus, Larsson E. G., R. Yan, Principe J. C., and Fitzsimmons J. R. Asymptotic SNR-performance of some image combination techniques for phased-array MRI. *J Sign Proc*, 84(6):997–1003, 2004.

[8] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *J Multiscale Model Simul*, 4(2):490–530, 2005.

[9] S. Roujol, B. Denis de Senneville, S. Hey, C.T.W. Moonen, and M. Ries. Robust Adaptive Extended Kalman filtering for real time MR-thermometry guided HIFU interventions. *IEEE Transaction on Medical Imaging*, 31(3):533–542, 2012.

[10] R. Kalman. A new approach to linear filtering and prediction problems. *J of Basic Engineering*, 82(2):35–45, 1960.

[11] P. Coupe, P. Yger, P. Hellier, C. Kervrann, and C. Barillot. An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Trans Med Imaging*, 27(4):425–441, 2008.

[12] J. V. Manjon, P. Coupe, L. Marti-Bonmati, D. L. Collins, and M. Robles. Adaptive non-local means denoising of MR images with spatially varying noise levels. *J Magn Reson Imaging*, 31(1):192–203, 2010.

[13] T. E. Conturo and G. D. Smith. Signal-to-noise in phase angle reconstruction: Dynamic range extension using phase reference offsets. *Magn Reson Med*, 15(3):420–437, 1990.

[14] S. J. Riederer. Recent technical advances in MR imaging in the abdomen. *J Magn Reson Imaging*, 6(5):822–832, 1996.

[15] R. D. Peters and R. M. Henkelman. Proton-resonance frequency shift MR thermometry is affected by changes in the electrical conductivity of tissue. *Magn Reson Med*, 43(1):62–71, 2000.

[16] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.

[17] M. Ries, B. Denis de Senneville, S. Roujol, Y. Berber, B. Quesson, and C. Moonen. Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues. *Magnetic Resonance in Medicine*, 64:1704–1712, 2010.

[18] B. Denis de Senneville, P. Desbarats, B. Quesson, and Moonen C. T. W. Real-time artifact corrections for quantitative MR temperature mapping. *Journal of WSCG*, 11(1), 2003.

[19] B. Denis de Senneville, A. El Hamidi, and C. Moonen. A direct PCA-based approach for real-time description of physiological organ deformations. *IEEE Transaction on Medical Imaging*, 34(4):974–982, 2015.

[20] S. Separeto and W. Dewey. Thermal dose determination in cancer therapy. *Int J Radiat Oncol Biol Phys*, 10(6):787–800, 1984.

[21] T. J. Dubinsky, C. Cuevas, M. K. Dighe, O. Kolokythas, and J. H. Hwang. High-intensity focused ultrasound: current potential and oncologic applications. *Am J Roentgenol*, 190(1):191–199, 2008.

[22] A. Kolandaivelu, M. M. Zviman, V. Castro, A. C. Lardo, R. D. Berger, and H. R. Halperin. Noninvasive assessment of tissue heating during cardiac radiofrequency ablation using MRI thermography. *Circ Arrhythm Electrophysiol*, 3(5):521–529, 2010.

[23] V. Auboiroux, L. Petrusca, M. Viallon, A. Muller, S. Terraz, R. Breuget, X. Montet, C. D. Becker, and R. Salomir. Respiratory-gated MRgHIFU in upper abdomen using an MR-compatible in-bore digital camera. *Biomed Res Int*, 2014:1–9, 2014.

[24] J. W. Wijlemans, M. de Greef, G. Schubert, L. W. Bartels, C. T. Moonen, M. A. van den Bosch, and M. Ries. A clinically feasible treatment protocol for magnetic resonance-guided high-intensity focused ultrasound ablation in the liver. *Invest Radiol*, 50(1):24–31, 2015.

[25] H. Gudbjartsson and S. Patz. The Rician distribution of noisy MRI data. *Magn Reson Med*, 34(6):910–914, 1995.

[26] R. D. Nowak. Wavelet-based Rician noise removal for magnetic resonance imaging. *IEEE Trans Imag Proc*, 8(10):1408–1419, 1999.

[27] N. Todd, J. Prakesh, H. Odeen, J. de Bever, A. Payne, P. Yalavarthy, and D. L. Parker. Toward real-time availability of 3D temperature maps created with temporally constrained reconstruction. *Magn Reson Med*, 71(4):1394–1404, 2014.

CHAPTER 4

A framework for the correction of slow physiological drifts during MR-guided HIFU therapies: Proof of concept

"It is impossible to begin to learn that which one thinks one already knows."
- Epictetus

This chapter is based on:
C. Zachiu, B. Denis de Senneville, C. Moonen, and M. Ries. A Framework for the Correction of Slow Physiological Drifts During MR-Guided HIFU Therapies: Proof of Concept. *Med. Phys.*, 42(7): 4173 - 4148, 2015.

Abstract

While respiratory motion compensation for MR-guided HIFU interventions has been extensively studied, the influence of slow physiological motion due to, for example, peristaltic activity, has so far been largely neglected. During lengthy interventions, the magnitude of the latter can exceed acceptable therapeutic margins. The goal of the present study is to exploit the episodic work-flow of these therapies to implement a motion correction strategy for slow varying drifts of the target area and organs at risk over the entire duration of the intervention.

The therapeutic work-flow of an MR-guided HIFU intervention is in practice often episodic: Bursts of energy delivery are interleaved with periods of inactivity, allowing the effects of the beam on healthy tissues to recede and/or during which the plan of the intervention is re-optimized. These periods usually last for at least several minutes. It is at this time scale that organ drifts due to slow physiological motion become significant. In order to capture these drifts we propose the integration of 3D MR scans in the therapy work-flow during the inactivity intervals. Displacements were estimated using an optical flow algorithm applied on the 3D acquired images. A preliminary study was conducted on 10 healthy volunteers. For each volunteer 3D MR images of the abdomen were acquired at regular intervals of 10min over a total duration of 80min. Motion analysis was restricted to the liver and kidneys. For validating the compatibility of the proposed motion correction strategy with the work-flow of an MR-guided HIFU therapy, an *in-vivo* experiment on a porcine liver was conducted. A volumetric HIFU ablation was completed over a time span of 2h. A 3D image was acquired before the first sonication, as well as after each sonication.

Following the volunteer study, drifts larger than 8mm for the liver and 5mm for the kidneys prove that slow physiological motion can exceed acceptable therapeutic margins. In the animal experiment motion tracking revealed an initial shift of up to 4mm during the first 10min and a subsequent continuous shift of $\sim$2mm/h until the end of the intervention. This leads to a continuously increasing mismatch of the initial shot planning, the thermal dose measurements and the true underlying anatomy. The estimated displacements allowed correcting the planned sonication cell cluster positions to the true target position, as well as the thermal dose estimates during the entire intervention and to correct the NPV-measurement. A spatial coherence of all three is particularly important to assure a confluent ablation volume and to prevent remaining islets of viable malignant tissue.

This study proposes a motion correction strategy for displacements resulting from slowly varying physiological motion that might occur during an MR-guided HIFU intervention. We have shown that such drifts can lead to a mis-

alignment between interventional planning, energy delivery and therapeutic validation. The presented volunteer study and *in-vivo* experiment demonstrates both the relevance of the problem for HIFU therapies and the compatibility of the proposed motion compensation framework with the work-flow of a HIFU intervention under clinical conditions.

4.1 Introduction

Magnetic resonance guided high intensity focused ultrasound (MRg-HIFU) is an attractive solution for treating tumors inside the human body in a non-invasive manner.[1,2] Such interventions have the potential to target and ablate non-invasively primary and metastatic cancer deep in the human body in a single session.[3] Due to the associated reduced burden for the patient, treatment related patient morbidity and complication rate, HIFU represents a potential alternative to patient groups which are currently not eligible for invasive or minimally invasive therapy.[4,5] MRg-HIFU has already been successfully applied for a number conditions such as symptomatic uterine fibroids, prostate cancer, liver tumors, thyroid nodules, palliative treatment for bone metastases, breast fibroadenoma and functional neurological disorders, including essential tremor, neuropathic pain and tremor-dominant Parkinson's disease.[5] However, MRg-HIFU in clinical applications becomes challenging when the tumor is located in the abdomen or the lower thorax due to continuous physiological motion.[6–10] Uncertainties regarding the exact tumor position might lead to undesired thermal damage to healthy tissue or the delivery of a thermal dose to the tumor that is below the lethal threshold.[6] In addition to shifts caused by respiration, tumors in the abdomen and the lower thorax will also suffer displacements due to digestive and metabolic processes, abdominal and thoracic muscle relaxation and ultimately with the spontaneous motions that the patient might exhibit. Important is to realize that these types of motion do, in general, not occur on the same time-scale:

- Respiratory induced motion of free-breathing patients has a frequency of 0.2-0.4Hz and displays a rapid in and exhalation of a duration of 1-1.5s followed by a rest of the diaphragm of 2-3 times of this duration. The majority of the existing scientific work on motion correction schemes for HIFU interventions focuses on the correction of this type of motion.[11,12]

- Although the time-scale of digestive, metabolic or muscle relaxation motion events depends on the event source (e.g. development of digestive gases or the passage of digestive products in the gastrointestinal tract, bladder filling with urine, changes in patient respiration), the resulting shifts in the position of abdominal organs has in most studies been identified to become significant on a time-scale of several minutes.[13–15] For this reason, this type of motion will be referred to in this chapter as slow

physiological motion or drifts. Digestive activity can clinically be moderated by several measures: Peristaltic bowel motion and the development of digestive gases can be reduced by adjusting the diet of the patient prior to the intervention,[16] or by the administration of butylscopolamine and/or glucagon which act as digestive motility inhibiting agents.[17–19] However, these agents typically have an effective time of approximately 30-40 min, which may not be long enough for a full HIFU treatment. Shifts in the lower abdomen due to bladder filling can be reduced, for example, by the use of Foley catheters.[20,21]

- Finally, spontaneous motion is a difficult problem - it occurs infrequently and on a very short time-scale. It is particularly problematic for long interventions, which require the patient to remain in an uncomfortable position. In the past, this problem has been addressed by either using restraints such as molds or casts[22] or by sedating the patient.[23,24]

Although a motion compensation technique is required in order to reduce tumor displacement effects on the outcome of the therapy, it should imply minimal discomfort caused to the patient. For respiratory and slow physiological motion this can be achieved through the means of an image guided intervention.[6,7,25–27] Beneficial for the development of a motion compensation strategy for HIFU therapies is the fact that the therapeutic work-flow is in practice often episodic: Bursts of energy delivery are interleaved with periods of inactivity. More specifically, interventions usually alternate between short bursts of 30-60s of ultrasonic energy and replanning/cool-down phases of several minutes.[5,28] During energy delivery, respiration is the main source of abdominal organ motion, thus the focus will be on compensating for this type of displacements. Given the larger time scale at which organs shift due to slow physiological drifts, this type of motion can be compensated during the replanning/cool-down phases.

A study was recently conducted by Schmitt *et al.*[29] in which temperature and thermal dose measurements were corrected with respect to susceptibility artifacts arising from bowel motion during MRg-HIFU interventions in the prostate. While in this chapter we propose a correction strategy based on displacements provided by a motion estimation algorithm, their study focuses on developing and implementing a filter that excludes temperature information in regions that manifest a higher temperature fluctuation than the heated area. The approach was validated in five out of six clinical cases showing good potential for further use not only in the prostate but also for targets situated higher in the abdomen.

While substantial work has already been conducted for the compensation of respiratory motion, the current study aims to exploit the episodic work-flow of an MRg-HIFU intervention to propose a motion correction scheme for slow physiological motion. The study is divided in three parts: First, we assess long term shifts of the liver and kidney position in ten healthy volunteers over a

duration of 80min. This is achieved by the periodical examination of the area or the upper abdomen with fast 3D MRI and an on-the-fly 3D co-registration of the image volumes, which are both compatible with the therapeutic workflow of an MRg-HIFU therapy. Subsequently, we assess a lower boundary of the accuracy of the proposed co-registration methods for both organs based on a comparison with manually tracked anatomical landmarks at selected areas. Finally, we demonstrate the clinical benefit of the proposed approach by correcting for the slowly varying organ drifts that occur during an *in-vivo* high intensity focused ultrasound ablation on a porcine liver.

4.2 Materials and methods

4.2.1 General strategy

4.2.1.1 Proposed MR imaging framework

Figure 4.1 details the HIFU ablation protocol employed in our study. Our framework consists of performing a series of n ablative sessions, each one having an approximate duration of $5 - 10$min (which is divided between the sonication and a cool-down interval). Each sonication is preceded by a planning/replanning phase, which consists in establishing/refining the position and size of the area that needs to be ablated by that respective sonication. This step is performed on a separate high-resolution T1/T2 weighted image acquired at the beginning of the therapy. Note that the MR acquisition sequence for the planning image might differ from the one of the subsequent 3D T1-weighted abdominal scans (see Figure 4.1). During the actual sonications, MR thermometry is performed in order to monitor the temperature of the sonicated tissue and detect when the lethal thermal dose has been delivered. In this study we were interested in tracking the liver and kidneys over the entire duration of the intervention. At this timescale the organs may exhibit significant displacements in all three directions. Thus, organ tracking should be performed in 3D. With current technology this becomes, however, a challenging task during sonications due to the real-time feedback loop required for temperature monitoring.[6] Such a time constraint does currently not allow the acquisition of 3D images with reasonable quality coupled with a 3D image based tracking procedure. For this reason imaging and tracking during sonications has to be performed in 2D. During this time interval the focus is mainly on compensating for displacements caused by respiratory motion. In order to extend liver and kidney tracking to the duration of the entire therapy, we propose interlacing 3D T1-weighted abdominal scans in between sonications. These scans are evenly distributed over the time span of the therapy with a time gap T_a in between them. Registration of the acquired 3D volumes to a reference position allows compensation of the 3D displacements exhibited during sonication by the targeted organs due to

slow physiological motion. The acquisition and registration of the 3D volumes in between sonications is timewise feasible. This is due to the fact that each sonication is followed by a time interval within which the tissue situated in the near and far field of the HIFU transducer is allowed to cool down. The displacements estimated during the 3D registration process are then used to locate the ablated tissue on a 3D reference volume.

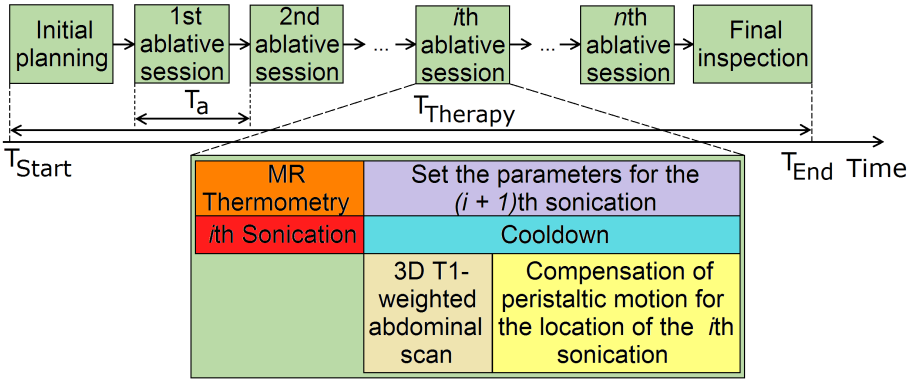


Figure 4.1: HIFU ablation protocol employed in our study. The 3D T1-weighted abdominal scans performed at regular time intervals allow the estimation of the 3D displacements exhibited by liver and kidneys due to slow physiological motion.

4.2.2 Estimation of 3D organ deformation

The 3D organ deformations were estimated after each sonication using an optical flow algorithm applied on the T1-weighted volumetric scans. The deformation estimation is based on the optical flow algorithm proposed by Horn & Schunck.[30] In its original form the algorithm is unable to estimate displacements that are larger than the size of one voxel. In order to overcome this limitation, a multi-resolution approach with an iterative refinement of the optical flow has been employed.[31] A multi-threaded C++ implementation of this algorithm has been provided by Team MR-HIFU, UMC Utrecht, The Netherlands. The algorithm requires as input two images: a reference image and an image to register. For the volunteer study and also for the animal experiment, the first image in their corresponding series of 3D T1-weighted abdominal scans was established as reference. Additionally, the optical flow algorithm requires fixating a parameter α which reflects the elasticity of the tracked organs. All 3D images were registered using $\alpha = 0.05$, value which was established empirically.

Given that the optical flow algorithm is gray-level intensity based,[30] it is important to evaluate its robustness to the noise that might superimpose on the images during the acquisition process. For this purpose, registration has

been performed between two un-altered 3D T1-weighted abdominal scans and the resulting motion vector field was established as gold standard. Afterwards, noise was added to the images, gradually reducing their signal-to-noise ratio (SNR). Registration was then performed for each noise level and a mean square error(MSE) between the current estimated motion vector field and the gold standard was computed. Such an approach will aid in detecting the lowest value of the SNR for which the estimated motion is still reliable.

4.2.3 Experimental setup

All MR imaging was performed on a 1.5T Philips Achieva scanner (Philips Healthcare, Best, The Netherlands).

4.2.3.1 Healthy volunteer study

The slowly varying drifts exhibited by the human liver and kidneys have been assessed for ten healthy volunteers. All volunteers went inside the MR scanner in a head first - prone position and were instructed not to move over the duration of the image acquisition process. A series of nine 3D T1-weighted scans were performed on the abdomen of each volunteer. Timewise the scans were evenly distributed over a duration of approximately $T_{Therapy} = 80$min with an average time gap of $T_a = 10$min in between the scans. The MR acquisition sequence was a fat suppressed spoiled gradient recalled echo employing the following parameters: TE = 2ms, TR = 4.3ms, image matrix 192 × 192 × 75, 10° flip angle, bandwidth in phase-encoding direction 373.4Hz, an isotropic voxel size of 2 × 2 × 2 mm^3, resulting an acquisition time of $\sim$ 90s. Fat suppression was achieved through the spectral presaturation with inversion recovery technique. For the remainder of this chapter the images acquired using the MR sequence with the aforementioned parameters will be referred to as 3D anatomical images. An exception to this nomenclature will be made for the 3D anatomical images that play the role of reference in the registration process. These images will be refered to as 3D anatomical anchors.

4.2.3.2 Animal experiment

HIFU ablation has been performed *in-vivo* on a porcine liver using a modified Sonalleve (Philips Healthcare, Helsinki, Finland) ablation system. The pig was anesthetized by an initial intramuscular injection containing ketamine (13 mg/kg), midazolam (0.7 mg/kg), atropine (0.02 mg/kg) and meloxicam (0.4 mg/kg). Subsequently, anesthesia was maintained by intravenous administration of sufentanil (11.3 μg/kg/h), midazolam (1 mg/kg/h) and nimbex (0.09 mg/kg/h). A total number of $n = 10$ sonications were carried out during an intervention that extended over a duration of $T_{\text{Therapy}} \approx 2$h. Sonication cells with a 4mm diameter were shot at a power of 450W for 30-35s with 8 - 9min cool-down intervals. After each sonication a 3D anatomical image has been

acquired. An additional 3D anatomical image was acquired before the first sonication, which will play the role of reference during the registration process. Closely after the acquisition of the last 3D anatomical image, a gadolinium based contrast agent has been administered to the animal. Subsequently a contrast enhanced T1-weighted image was acquired. In this manner, a visual feedback of the non-perfused tissue resulting from the 10 sonications is obtained. After the completion of the experiment, the animal was euthanized by intravenous injection of pentobarbital.
The animal experiment was conducted in agreement with the European law on animal experimentation and in compliance with the institution's rules for animal care and use.

4.2.4 Analysis and validation of the 3D estimated motion

4.2.4.1 Motion analysis and validation on healthy volunteers

In this study we were interested exclusively in the motion exhibited by the liver and kidneys. In order to restrict motion analysis to the mentioned organs of interest, a semi-automatic segmentation procedure was performed prior to the analysis. An initial segmentation was obtained by the means of active contours which was subsequently refined through a manual correction. The software tool used for segmentation was ITK-Snap v 2.4.0.
The validity of the estimated motion was tested by manually tracking blood vessel bifurcations in the liver and kidneys. For each of the volunteers three vessel bifurcations were tracked: one in the upper part of the liver (lobes 7 and 8), one in the lower part of the liver (lobes 5 and 6) and one in the kidneys. The manually tracked landmarks were then compared to the displacements estimated by the optical flow algorithm.
The manually measured displacements were also used to quantify the positional errors caused by the slowly varying drifts in the absence and respectively in the presence of a motion compensation strategy. This was achieved by comparing the displacements obtained through manual tracking (which correspond to the positional errors we would have in the absence of a motion compensation scheme) with the difference between these displacements and the displacements estimated by the optical flow algorithm (which corresponds to the residual positional errors when a motion compensation scheme is integrated in the therapy work-flow). Note that for computing the positional errors, the magnitude of the 3D displacement vectors was used instead of separately evaluating the errors on each of the three directions of motion.

4.2.4.2 Motion analysis and validation for a real HIFU therapy on an animal

As a result of slow physiological motion, the porcine liver might suffer displacements during the HIFU ablation session. By using the acquired 3D anatomical

images, the position of the sonicated areas was retrospectively tracked over the time span of the session. This allowed identifying on the planning image, acquired at the beginning of the HIFU intervention, the actual anatomy that was ablated, which, due to the organ drifts, might differ from the one originally planned. The estimated motion was also used to bring the temperature maps acquired during the 10 sonications to a common reference space determined by the planning image. In this reference space, the delivered thermal dose[32] was re-calculated. This resulted in a correction such that the newly calculated thermal dose will match the anatomy captured by the planning image. Additionally, the position of the non-perfused volume (NPV) resulting from the intervention was also brought in the reference space of the planning image.
Similar to the healthy volunteer study, motion analysis for the animal experiment has been restricted to the organ of interest. For this purpose a segmentation procedure as the one described in section 4.2.4.1 has been conducted.

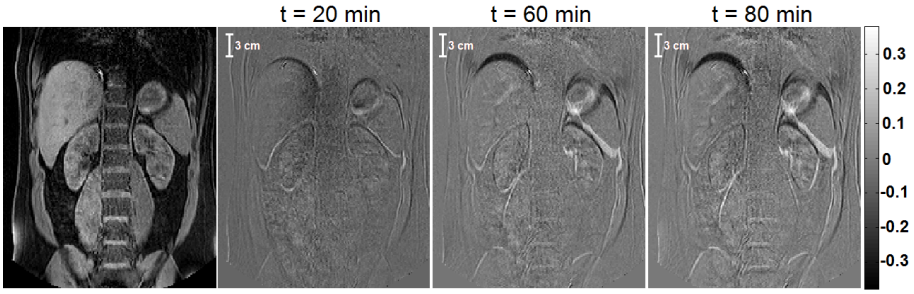
4.3 Results

4.3.1 Measured displacement for liver and kidneys caused by slow physiological drifts

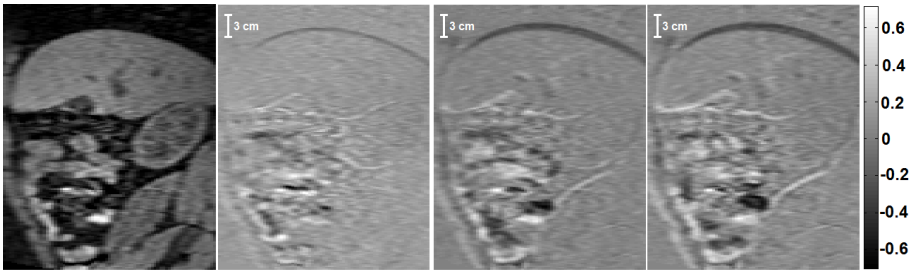
Figure 4.2 illustrates the evolution of the anatomy changes induced by slow physiological motion in one of the volunteers abdomen. The displayed figures correspond to a coronal (figure 4.2a) and respectively a sagittal (figure 4.2b) slice extracted from a set of images obtained by subtracting from the 3D anatomical anchor subsequent 3D anatomical images. The leftmost images in the two rows depict the anatomy present in the two slices as captured by the 3D anatomical anchor. In this particular data set, considering just the liver and kidneys, the superior part of the liver experiences an obvious displacement in the cranio-caudal direction. The peak of this displacement takes place at the acquisition time of the last 3D anatomical image. Given the image's voxel size of $2\times2\times2\text{mm}^3$, this displacement is measured to be approximately 13.5mm. In comparison, from a visual point of view, the displacements in the left-right and anterior-posterior directions were deemed to be negligible.

4.3.2 Robustness to noise of the optical flow algorithm

Figure 4.3 displays the relationship between the SNR of the registered images and the MSE of the corresponding estimated vector field with respect to the gold standard. It can be observed that for SNR values less than 5 errors start increasing exponentially. Thus, for images with an SNR value below this threshold, the estimated motion vectors are to be considered unreliable.



(a) Coronal slice



(b) Sagittal slice

Figure 4.2: Set of images that result from the subtraction between the 3D anatomical anchor and the 3D anatomical image acquired t minutes after the acquisition of the anatomical anchor in one of the volunteers. The figures correspond to a (a) coronal (b) sagittal slice from the resulting 3D difference images. The leftmost images give a visual depiction of the anatomy present in the two slices.

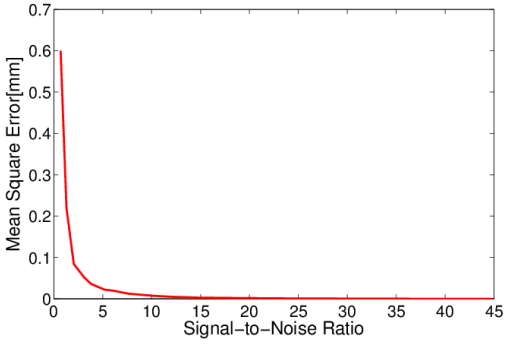


Figure 4.3: The relationship between the SNR of the registered images and the MSE of the corresponding estimated vector field with respect to the gold standard.

4.3.3 Quantitative analysis of the motion amplitude

For each volunteer the output of the registration process was a set of eight motion vector fields, each corresponding to the displacements estimated at the acquisition time point of their associated 3D anatomical image. The motion fields from the individual volunteers were then pooled such that a joint set containing the displacements from all volunteers at each time point is obtained. Figures 4.4a-(f) display this new joint set as an ensemble of boxplots. The displacements have been illustrated separately for the liver and kidneys in the cranio-caudal (CC), left-right (LR) and respectively anterior-posterior (AP) directions. The values on the abscissa correspond to the time delay in minutes since the acquisition of the reference image. The values on the ordinate indicate the estimated signed displacements in millimeters. Each boxplot is constructed in the following fashion: the lower limit of the box corresponds to the first quartile, the upper limit of the box is the third quartile, the median is represented by the red horizontal line in the middle of the box, the red cross indicates the mean of the set, the lower whisker corresponds to the 5^{th} percentile and the upper whisker to the 95^{th} percentile. For all the organs of interest, the overall median of the displacement in the CC direction has a tendency to increase over time reaching ~ 3 mm. The maximum liver displacement in the CC direction is ~ 8 mm, while for the kidneys this measures ~ 5 mm. For the LR and AP directions, the median of the displacements does not manifest any observable tendency. For the liver the maximum displacement in the LR direction is ~ 3 mm, while for the kidneys it reaches at a value of ~ 4 mm. In the AP direction the maximum displacement measures ~ 2.5 mm for the liver and ~ 4 mm for the kidneys. Figures 4.4g and 4.4h depict the time evolution of the median motion vector magnitude estimated in the liver and respectively kidneys of each volunteer. These graphs reconfirm the tendency of the displacements to increase over time, however, this might happen in non-monotone manner, with the organs randomly moving closer or further from their reference position, and with rather large inter-individual variations.

4.3.4 Validation of the 3D estimated motion

Validation of the estimated motion was made using the approach described in section 4.2.4.1. Figure 4.5 displays, for all of the volunteers, the correspondence between the manually determined landmark displacements and the landmark displacements estimated by the optical flow algorithm. The assessment was made in the CC (blue circles), LR (red squares) and AP (black crosses) directions respectively. In the case of a perfect correspondence between the manually tracked landmarks and the estimated landmark positions, the points from the graph would lie on the first bisector (depicted by the purple line in figure 4.5). The precision of the manually tracked displacements is however limited by the discrete nature of the acquired images. Therefore tracking errors of up to one voxel (2mm) are to be expected. Nevertheless, a paired T-test applied on the

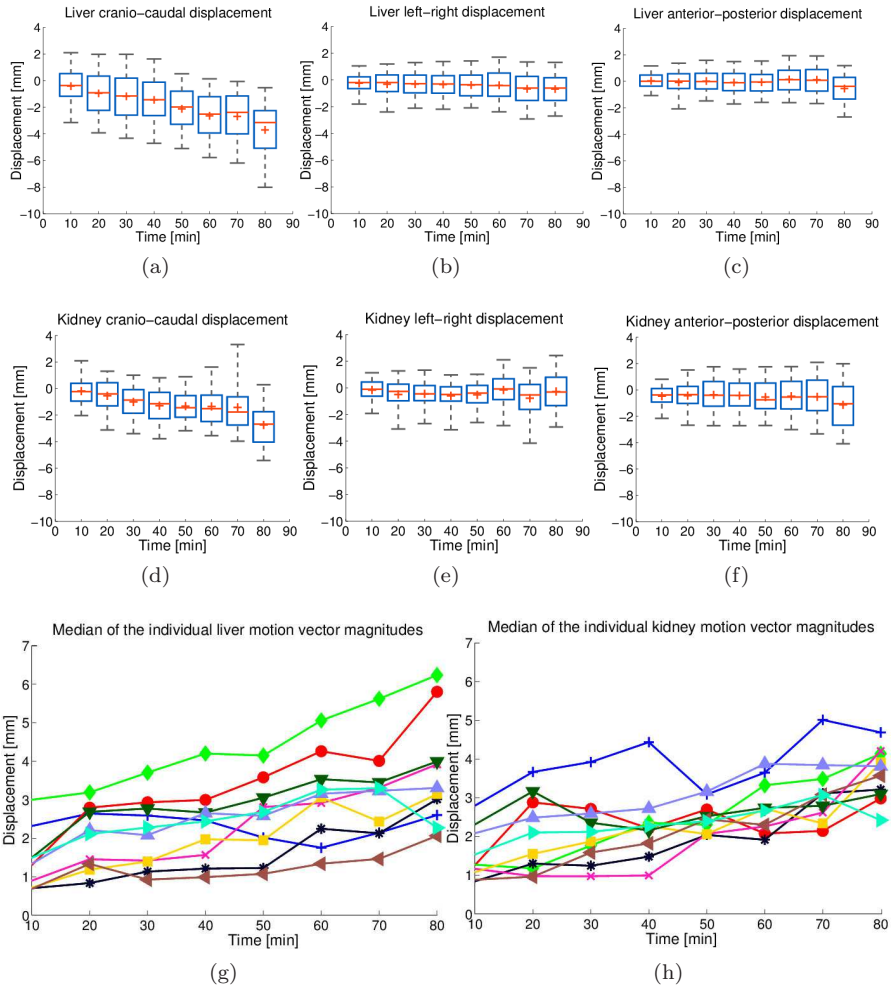


Figure 4.4: Evolution of the liver and kidneys displacement caused by slow physiological motion over a duration of 80min in ten healthy volunteers. (a) - (f) Each boxplot is a representation of a joint set containing the displacements estimated on all volunteers after $k \times 10$min since the acquisition of the 3D anatomical anchor, with $k = 1, \ldots, 8$. (g), (h) The median of the magnitudes of the displacement vectors estimated for the liver and respectively the kidneys of each volunteer. The values on the abscissa correspond to the time in minutes since the acquisition of the 3D anatomical anchor. The values on the ordinate are the signed displacements in millimeters in the cranio-caudal, left-right and anterior-posterior directions. Each color corresponds to a different volunteer.

displacements obtained through the two techniques indicates that there is no significant difference between the values at 5 % significance level.
Figure 4.6 displays positional errors in millimeters that would affect the intervention in the selected landmarks, in the presence and respectively in the absence of a motion correction scheme. The errors were computed as described in section 4.2.4.1. The boxplots are built in the same manner as the ones illustrated in Figure 4.4. More than 75% of the errors become less than 2 mm (one voxel) when motion compensation is present. The median of the errors also drops from $\sim$ 2.8 mm to $\sim$ 1.3 mm.

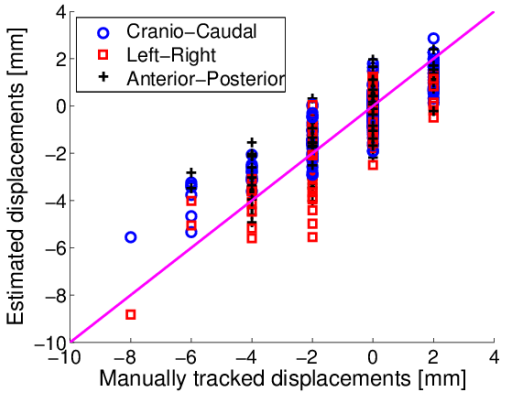


Figure 4.5: Validation of the optical flow algorithm. The graph illustrates the correspondence between the displacements obtained by manually tracking a set of vessel bifurcations in the liver and kidneys (on the abscissa) and the displacements estimated by the optical flow algorithm (on the ordinate). The displacements were determined for all volunteers in three directions: CC (blue circles), LR (red squares) and AP (black crosses). The purple line corresponds to the ideal location of the plotted points.

4.3.5 *In-vivo* experiment on a porcine liver

Figure 4.7 illustrates the slow physiological drifts experienced by the porcine liver over the duration of the HIFU intervention. The displayed graphs were constructed in a similar fashion as the ones presented in Figure 4.4. In contrast with the observations made during the volunteer study, there is no clear tendency of the displacements to increase over time in either of the directions. Nevertheless, displacements of up to 4mm can be observed in the CC direction.
A magnification of the planning image, acquired at the beginning of the therapy, with focus on the originally planned cell cluster overlaid as yellow circles is illustrated in figure 4.8a. The estimated displacements will allow identifying on the same image the motion corrected sonication cell locations, indicated by

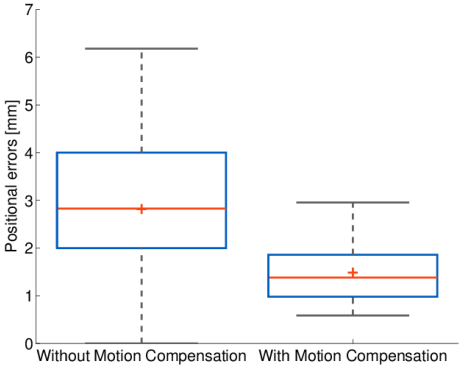


Figure 4.6: Positional errors that would affect the intervention in the selected landmarks, in the presence and respectively in the absence of a motion correction scheme. The values on the ordinate correspond to the positional errors in millimeters computed according to 4.2.4.1.

the red circles which, due to the organ drifts, will differ from the ones originally planned. The distance between the centers of mass of the planned and the corrected sonication cell clusters measures ~ 2 mm. The high intensity regions from figures 4.8b and 4.8c depict the anatomy that was determined to have received a lethal thermal dose[32] without and respectively with motion correction. Without registering the thermometry images with respect to the estimated organ drifts, a positional error will occur between the lethal dose and the corrected location of the sonications. This phenomenon is illustrated in figure 4.8b. Figure 4.8c displays the location of lethal thermal dose after registering the thermometry images. The distance between the centers of mass of the position corrected sonication cell cluster and the region determined to have received a lethal thermal dose is reduced from ~ 2 mm without motion correction to ~ 0.3 mm with motion correction.

Figure 4.9a displays a slice from the contrast enhanced image acquired at the HIFU therapy endpoint. The NPV resulting from the 10 sonications is enclosed in a red rectangle and magnified for better visibility. A mask covering the NPV was manually extracted from this image and overlaid on the planning scan as a high intensity region as depicted in figure 4.9b. For positional reference, the corrected sonication locations were also overlaid as red circles. A positional discrepancy between the cell cluster and the NPV can be observed. By using the displacements estimated between the last 3D anatomical image and the 3D anatomical anchor, the contrast enhanced image was registered to match the planning image anatomy. A mask covering the motion corrected NPV was then manually extracted and superimposed on the planning image as a high intensity region as depicted by figure 4.9c. Note that due to observer subjectivity

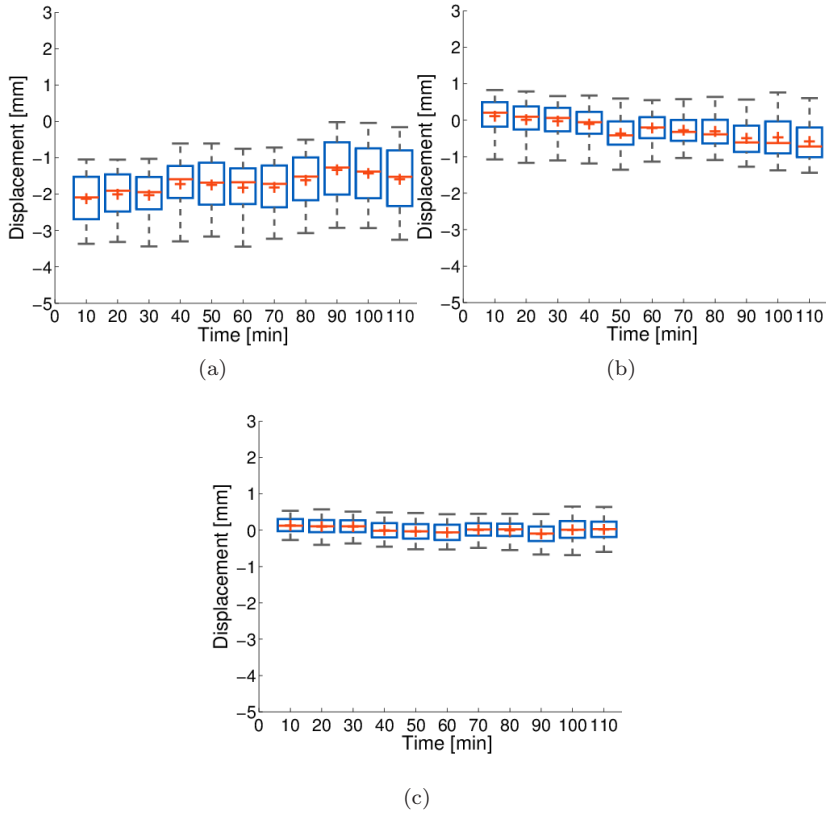
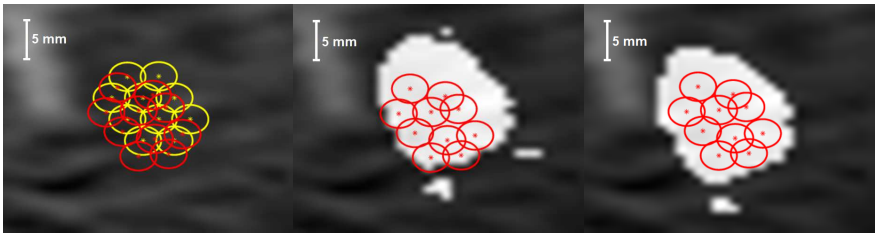


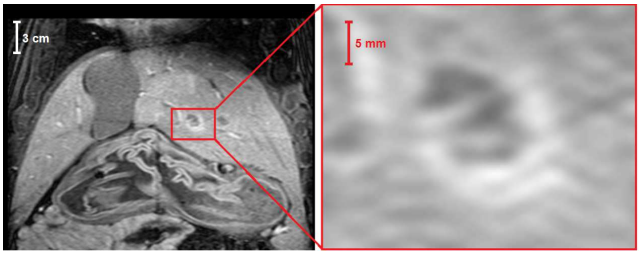
Figure 4.7: Evolution of the liver displacements caused by slow physiological motion during the *in-vivo* HIFU pig experiment. The values on the abscissa correspond to the time delay in minutes since the acquisition of the 3D anatomical anchor. The values on the ordinate are the signed displacements in millimeters in the (a) CC, (b) LR and (c) AP directions.

during the manual NPV delineation, the non-corrected and the corrected NPV masks might differ. Without motion compensation, the distance between the centers of mass of the position corrected cell cluster and the NPV amounts to ~ 3.5 mm. By correcting the NPV position the distance is reduced to ~ 1 mm.

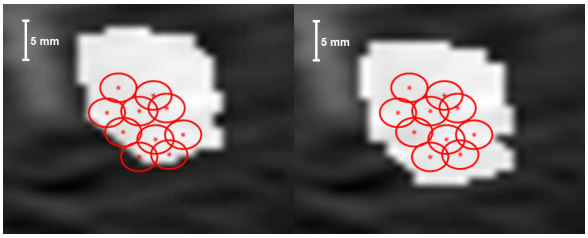


(a) (b) (c)

Figure 4.8: (a) Planned cluster of sonication cells (yellow circles) together with the motion corrected location of the sonications (red circles) overlaid on the corresponding anatomy from the planning image acquired at the beginning of the therapy. (b), (c) The anatomy that was determined to have received a lethal thermal dose (depicted by the high intensity region): (b) before registering the thermometry images (c) after registering the thermometry images.



(a)



(b) (c)

Figure 4.9: (a) Slice of the contrast enhanced scan acquired at the end of the HIFU intervention upon which the resulting NPV is visible. The NPV resulting from the sonications is enclosed in a red rectangle and magnified for better visibility. (b), (c) Corrected sonication locations overlaid as red circles on the corresponding anatomy from the planning image. The high intensity region in the images corresponds to the NPV: (b) without motion correction (c) with motion correction.

4.4 Discussion

The goal of the presented study was to exploit the episodic work-flow of an MR-guided HIFU therapy to implement a motion correction strategy for the displacement of abdominal organs under the effect of slow physiological drifts. The strategy involves tracking both the target area and organs at risk (OAR) over the entire duration of the intervention.
In the first part of the study we investigated the slowly varying drifts exhibited by the human liver and kidneys during a time period of 80min in 10 healthy volunteers. The obtained displacements of up to 8mm for the liver and 5mm for the kidneys are in good correspondence with previous reportings.[13, 27] Please note that respiratory motion was compensated by a navigator gated acquisition scheme for all acquired images. Therefore the obtained shifts are also likely to include shifts of the diaphragm with respect to the navigator.
The findings that result from the volunteer study have shown that large inter-individual variations have to be expected. For the liver of volunteer #9 the median of the displacements remained close to 1mm over the entire 80min duration of the study, while for volunteer #3 the median was already close to 4mm halfway through the experiment (see Figure 4.4h). The measured displacements are also expected to vary with patient position: prone or supine. Thus, this reconfirms that a continuous supervision and correction of the displacements is clearly beneficial for the therapeutic success. For example for the noninvasive treatment of stage 1 tumors in the liver and kidneys the reported shifts represent $\sim$ 20-50% of the typical tumor dimensions. If the therapy is conducted without a motion compensation scheme, undesired shifts of this magnitude need to be included in the surgical margin and lead to a substantially longer intervention.[33, 34]
One of the challenges that might occur while evaluating the performance of motion estimation algorithms dedicated to non-invasive *in-vivo* interventions is the lack of a reliable gold standard.[35, 36] However, the multi-resolution Horn & Schunck optical flow algorithm with iterative refinement of the displacement fields, which we employed in our study, has already been successfully validated in previous works for both in-vivo and ex-vivo scenarios. For example, Denis de Senneville et *al.*[37] made use of a similar algorithm for steering the focus of the HIFU beam during an ex-vivo experiment performed on a piece of porcine muscle tissue. A known periodical motion pattern was induced to the target tissue while performing 2D MR imaging. Based on the resulting images, real-time optical flow-based motion estimation was used to determine the target position. The errors between the true and the estimated trajectories remained on average under 0.4 mm. For minimally invasive beam therapies the issue of validating motion estimation algorithms could be addressed, for example, by tracking implanted fiducial markers in/or in the proximity of the targeted region. Such an approach is, however, not feasible for the applications discussed in the current study since it would imply a certain degree of invasiveness. In this

chapter, the validation was performed by manually tracking a group of landmarks and established the resulting displacements as a gold standard. More than 75% of the errors between the estimated displacements and the gold standard remained under 2mm which is actually the image resolution and thus the precision limit for manual tracking. The source of these errors is the combined effect of misregistration by the motion estimation algorithm and the subjective nature of the landmark manual tracking procedure. A validation of the Horn & Schunck optical flow algorithm using a similar approach was also performed by Ostergaard et *al*.[38] In their work, the validation was based on 3D CT lung images acquired at different phases of the respiratory cycle. The images were part of an open database,[39] which also contains the location of 41 manually identified anatomical landmarks during each of the respiratory phases. The position of the manually tracked landmarks was then compared against the one estimated by the algorithm. This resulted in errors which on average remained under 2mm, which is consistent with the results obtained in our study. It is also expected that as long as the assumptions made by the optical flow algorithm are not violated, including voxel gray-level intensity conservation from the reference to the moving image and the continuous nature of the displacements, the same results can be extrapolated to the organs as a whole. Obtaining a gold standard through the proposed approach is however limited by the presence and/or visibility of specific anatomical landmarks (e.g. vessel bifurcations) in the images considered for manual tracking. If this condition is not met, alternative methods need to be considered. Nevertheless, our findings are in good correspondence with previous reportings.[38,40–42]

The best way to improve the accuracy and precision of the estimated motion is to increase the resolution of the 3D anatomical images used for registration. This however becomes challenging due to the time constraints imposed by the therapy work-flow. If image resolution is increased while keeping the acquisition time at acceptable limits, the SNR of the image might drop to values that deteriorate the estimated motion to the point that the improvement in image resolution becomes practically useless. Accuracy of the estimated motion is also limited by the image distortions intrinsic to the employed MR acquisition sequence. The distortions mainly originate from the inhomogeneities of the MR scanner's static magnetic field and from chemical shifts (ex. water-fat shift) that might occur at organ and body borders. For example, Dragonu et *al*[43] have reported spatial distortions due to B_0 inhomogeneities larger than 1 cm in the liver. The study has been performed, however, for rapid gradient recalled MRI sequences using an echo-planar readout, which is not the case in the present chapter (see section 4.2.3.1). Nevertheless, the results in their study indicate that B_0 inhomogeneities of ~ 1.3 ppm are to be expected in the human abdomen. This would cause the sequence used in our study to generate spatial distortions of ~ 0.3mm. Such a value is well beneath the image resolution (which is 2mm), thus its effects on the estimated motion were deemed to be negligible. Nevertheless, spatial distortions due to B_0 inhomogeneities must be addressed for MR images used for motion estimation, either by optimizing the

MR acquisition sequence or by choosing an appropriate correction scheme.[44–47] In order to eliminate the effects of the water-fat shift, the MR-sequence used in our study has been optimized to be fat suppressing. This will not only remove fat shifts (and thus image artifacts) at the organ and image boundaries but it will also increase the contrast-to-noise ratio of the entire image, which is beneficial for the accuracy of the estimated motion. An additional factor that affects the quality of the estimated motion is the SNR of the images themselves, fact which is to be expected given that the optical flow algorithm employed for registration is gray-level intensity based. According to Figure 4.3, as the SNR drops bellow 5, the motion estimation errors start increasing exponentially. Thus, it can be considered that beyond this threshold the estimated motion vector field becomes unreliable. This however does not affect in any way the success rate of our method since the 3D anatomical images used for registration have a SNR value of ~ 44, therefore in the right extrema of the curve depicted by Figure 4.3.

The Horn & Schunck motion estimation algorithm employed in our study was chosen mainly due to its previous validation for the *in-vivo* correction of respiratory motion during HIFU ablations in the liver.[6,26] There are, however, a few aspects concerning the algorithm that could be further improved. For example, since it relies on the assumption that voxels conserve their gray-level intensity on the followed trajectory, the Horn & Schunck algorithm is sensitive to gray-level intensity variations not attributed to motion. Such a situation might occur when, for instance, anatomical structures appear and/or disappear from the field-of-view. In addition, the algorithm imposes that the estimated displacements are spatially continuous. Since abdominal organs are elastic and incompressible, this is a justified assumption. However, an exception from this rule arises at the interface between the organs and the abdominal and/or the thoracic wall. In these regions shearing effects will occur, thus violating the continuity assumption. In order to overcome these limitations, versions of the optical flow algorithm that reduce the impact of such effects could be employed. Such algorithms have already been applied successfully in the video domain,[48] however, to our knowledge, the validity of such approaches for *in-vivo* medical imaging applications has yet to be proven.

The focus of this study was on the analysis of the displacements exhibited by the human liver and kidneys due to slow physiological drifts, over the entire duration of a HIFU therapy session. Thus, the MR sequence used to acquire the 3D anatomical images, upon which the tracking procedure is based, was optimized to cover the entire upper abdomen. This resulted in a rather lengthy image acquisition time of $\sim$90s. Since in the scope of our study this was sufficient, we considered this acceptable. However, this might not always be the case. During the animal experiment conducted for this study, 4mm sonication cells were shot with a power of 450W for 30-35s. This translates to a delivered energy of $\sim$15kJ per shot. In effect, the HIFU system imposed cool-down intervals of 8-9min. More typical values for clinical treatments are 2-3min cool-down with $\sim$5kJ of delivered energy per shot. Thus, for a really smooth clinical

work-flow the acquisition time for the 3D anatomical images should preferably be accelerated to 30-40s. During a HIFU intervention, only the displacements experienced by the tumor and its immediate surroundings are usually of importance, thus the most obvious way to achieve this acceleration would be to limit the 3D coverage to the organ/area of interest.
The HIFU experiments performed *in-vivo* on a porcine liver have shown that the method proposed in our study can be applied for a real HIFU ablation therapy. In the investigated case, the animal was not fed 24h prior to the intervention. During the intervention, the animal was under full anesthesia and thus the thoracic and abdominal muscles were under temporary paralysis. Similar to the volunteer experiments, all respiratory induced displacements were compensated by a gating of both all MR-acquisitions and the therapeutic energy delivery. As a consequence, the observed displacements can be largely attributed to shifts due to digestive gas development in the stomach and the intestines. As the timecourse displays (see Figure 4.7), the majority of the displacement occurred during the first 10 minutes of the intervention. This demonstrates that even under ideal conditions, the occurrence of shifts of 3.5 - 4mm can occur and are problematic for therapy. The interventional planning was carried out on a T2-weighted high resolution image obtained at the beginning of the intervention. The subsequent energy deliveries were placed on the target delineation. Therefore, the shift after 10min would have the following consequences for the subsequent therapy :

- Parts of the tumor would have not been ablated, or the surgical margin would have not been sufficient.
- Shifts in-between shots potentially lead to small areas of viable tissue in between the ablation cells. Without motion compensation these have to be accounted for by a large overlap of the sonication cells, which significantly decreases the achievable ablation speed and thus the overall duration of the intervention.

It is important to note that during the *in-vivo* experiment no motion estimation or correction has been performed. The estimation of the displacements was made in post-processing using the 3D anatomical images acquired during therapy. The resulting motion vectors were used to retrospectively update the position and configuration of the sonication cell cluster such that it will match the actual ablated anatomy. The small residual shift of $\sim$ 0.3mm between the updated cell cluster and the motion corrected thermal dose and of $\sim$ 1mm between the updated cell cluster and the motion corrected NPV is a good indicator that the proposed motion estimation procedure can be successfully applied *in-vivo* with acceptable error margins.
The current multithreaded C++ implementation of the 3D registration algorithm used in our study has an execution time of approximately 30s. Although motion compensation was performed in post-processing, the time interval between energy deliveries allows an integration. Furthermore, acceleration using

an implementation via the compute unified device architecture (CUDA) platform on a graphical processing unit (GPU)[26] allows to bring the motion estimation down to $\sim$ 10s, which would further contribute to a smooth overall work-flow.
The motion compensation framework proposed in this study is not limited to HIFU interventions. With the development of the first MR linear accelerator (MR-LinAc) prototypes,[49,50] our motion correction strategy could potentially be integrated in the work-flow of an MR-guided radiotherapy (MRg-RT). Similar to HIFU interventions, energy delivery during external beam radiotherapy is usually episodic, with the total prescribed radiation dose being divided into multiple fractions. The time interval in between fractions, which can sometimes take up to several days, allows the healthy tissues to partially recover from the effects of the absorbed radiation. These time intervals can be seen as the equivalent of the cool-down periods during HIFU therapies. Obviously, there is a high probability that due to the physiological drifts discussed in this chapter or due to patient positional errors, the target location will differ from one fraction to the next. If these displacements are not taken into consideration, they might lead to the under-treatment of the tumor or to additional damage to the OAR's. To address this problem, a 3D anatomical image could be acquired, for example, after each fraction and registered to a 3D anatomical anchor acquired before delivering the first fraction. The resulting motion vector fields could then be integrated into the therapy planning algorithm and the plan could be re-optimized prior to the next fraction. The proposed framework can also be adapted to operate during the delivery of a single fraction. In order to reduce the damage caused to the OAR's, a single fraction is usually delivered from several positions of the LinAc gantry. A 3D anatomical image could be acquired while changing gantry positions and registered to a 3D anatomical anchor acquired prior to starting dose delivery. The estimated displacements could be used, for example, to re-optimize the plan on-the-fly. However, such an approach is timewise challenging since the gantry positional changes usually take place over time intervals of only several seconds.

4.5 Conclusion

So far, studies concerned with motion estimation during MRg-HIFU interventions in the abdomen were focused on tracking the targeted area only over the time course of the therapeutic energy delivery. At this level, the interest was compensating for motion mainly caused by respiration. In the presented study, we demonstrated that motion compensation only during energy delivery is conceptually not sufficient : HIFU therapies rely on a non-invasive target identification and delineation, the subsequent interventional planning of the ablation areas, the therapeutic energy delivery and the validation of the therapeutic endpoint. One of the strengths of MR-guided interventions is that except the initial target identification and delineation, all subsequent steps can be car-

ried out on-the-fly with the patient in place on the interventional table. The advantage of the proposed approach is that MR-guided interventions allow to adjust planning and therapy directly to the current anatomical position for the particular position of the patient on the interventional table. While static parts of the human anatomy such as the cranium or the limbs are relatively easy to immobilize (stereotactic frames, limb molds), the thoracic and the abdominal area are subject to both respiratory motion and slow physiological drifts. As we have demonstrated, this can lead to a misalignment between interventional planning, energy delivery and therapeutic validation. Since this can potentially lead to both undesired tissue damage of organs at risk and viable residual tumor tissue, a suitable motion correction strategy should encompass all of the essential steps of the intervention: Planning, monitoring of the therapeutic energy delivery and the validation of the therapeutic endpoint. The presented method is a first step towards such a motion correction strategy for interventions in the abdomen.

Acknowledgements

This work was supported by the Dutch Technology Foundation (STW) (project OnTrack #12813) and in part by the European Research Council (project ERC-2010-AdG-20100317, Sound Pharma) and ITEA 2 (project SoRTS).

Bibliography

[1] H.E. Cline, J.F. Schenck, K. Hynynen, R.D. Watkins, S.P. Souza, and F.A. Jolesz. MR-guided facused ultrasound surgery. *J Comput Assist Tomogr*, 16:956–965, 1992.

[2] K. Hynynen, W.R. Freund, H.E. Cline, A.H. Chung, R.D. Watkins, J.P. Vetro, and F.A. Jolesz. A clinical, noninvasive, MR imaging-monitored ultrasound surgery method. *Radiographics*, 16:185–195, 1996.

[3] R.O. Illing, J.E. Kennedy, G.R. ter Haar, A.S. Protheroe, P.J. Friend, F.V. Gleeson, D.W. Cranston, R.R. Phillips, and M.R. Middleton. The safety and feasibility of extracorporeal high-intensity focused ultrasound (hifu) for the treatment of liver and kidney tumours in a western population. *Br J Cancer*, 93(8):890–895, 2005.

[4] C.M. Tempany, N.J. McDannold, K. Hynyen, and F.A. Jolesz. Focused ultrasound surgery in oncology: overview and principles. *Radiology*, 36(5):39–56, 2011.

[5] J.L. Foley, M. Eames, J. Snell, A. Hananel, N. Kassell, and J.-F. Aubry. Image-guided focused ultrasound: state of the technology and the challanges that lie ahead. *Imaging Med.*, 5(4):357–370, 2013.

[6] M. Ries, B. Denis de Senneville, S. Roujol, Y. Berber, B. Quesson, and C. Moonen. Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues. *Magnetic Resonance in Medicine*, 64:1704–1712, 2010.

[7] S.P.M. Crijns, B.W. Raaymakers, and J.J.W. Lagendijk. Proof of concept of MRI-guided tracked radiation delivery: tracking one-dimensional motion. *Physics in Medicine and Biology*, 57:7863–7872, 2012.

[8] S. Diez, J. Garcia, and F. Sendra. Analysis and evaluation of periodic physiological organ motion in radiotherapy treatments. *Radiotherapy and Oncology*, 73:325–329, 2004.

[9] M. Goitein. Organ and tumor motion: An overview. *Seminars in Radiation Oncology*, 14:2–9, 2004.

[10] S. Vijayan, S. Klein, E.F. Hofstad, F. Lindseth, B. Ystgaard, and T. Lango. Validation of a non-rigid registration method for motion compensation in 4D ultrasound of the liver. *IEEE 10th International Symposium on Biomedical Imaging*, pages 792–795, 2013.

[11] J.T. Booth. *Modelling the impact of treatment uncertainties in radiotherapy.* PhD thesis, University of Adelaide, 2002.

[12] K.M. Langen and D.T.L. Jones. Organ motion and its management. *Int J Radiation Oncology Biol Phys*, 50(1):265–278, 2001.

[13] M. von Siebenthal and G. Szekely. Systematic errors in respiratory gating due to intrafraction deformations in the liver. *Med. Phys.*, 34(9):265–278, 2001.

[14] K.M. Langen, T.R Willoughby, S.L. Meeks, A. Santhanam, A. Cunningham, L. Levine, and P.A. Kupelian. Observations on real-time prostate gland motion using electromagnetic tracking. *Int J Radiation Oncology Biol. Phys.*, 71(4):1084–1090, 2008.

[15] P. Arnold, F. Preiswerk, B. Fasel, R. Salomir, K. Scheffler, and P.C. Cattin. 3D organ motion prediction for MR-guided high intensity focused ultrasound. *Med Image Comput Comput Assist Interv*, 14(Pt 2):623–630, 2011.

[16] M.H.P. Smitmans, F.J. Pos, J. de Bois, W.D. Heemsbergen, J.-J. Sonke, J.V. Lebesque, and M. van Herk. The influence of a dietary protocol on a cone beam CT-guided radiotherapy for prostate cancer patients. *Int J Radiation Oncology Biol Phys*, 71(4):1279–1286, 2008.

[17] J. Emmott, B. Sanghera, J. Chambers, and W.L. Wong. The effects of n-butylscopolamine on bowel uptake: an F-FDG PET study. *Nuclear Medicine Communications*, 29:11–16, 2008.

[18] A.R. Padhani, V.S. Khoo, J. Suckling, P. Minst, J.E. Husband, M.O. Leach, and D.P. Dearnaley. Evaluating the effect of rectal distension and rectal movement on prostate gland position using cine MRI. *Int J Radiation Oncology Biol Phys*, 44(3):525–533, 1999.

[19] J. M. Froelich, M. Daenzer, C. von Weymarn, Ertürk S. M., C. L. Zollikofer, and M. A. Patak. Functional Monitoring of Small Bowel Motility: Comparison of Spasmolysis Induced by Glucagon or Buscopan. *Proc Int Soc Magn Reson Med*, 17:1122 – 1123, 2009.

[20] R. Mirabell, P. Nouet, M. Rouzaud, A. Bardina, N. Heijira, and D. Schneider. Radiotherapy of bladder cancer: relevance of bladder volume changes in planning boost treatment. *Int J Radiation Oncology Biol Phys*, 41(4):741–746, 1998.

[21] J. Thariat, S. Aluwini, and M. Housset. Adaptive partial-boost stereotactic radiation therapy for muscle-invasive carcinoma of the urinary bladder. In L. E. Ponsky, D.B. Fuller, R.M. Meier, and C. Ma, editors, *Robotic Radiosurgery. Treating prostate cancer and related genitourinary applications*, pages 185–200. Springer Berlin Heidelberg, 2012.

[22] L.J. Verhey. Immobilizing and positioning patients for radiotherapy. *Seminars in radiation oncology*, 5(2):100–114, 1995.

[23] Y.-F. Zhou. High intensity focused ultrasound in clinical tumor ablation. *World J Clin Oncol*, 2(1):8–27, 2011.

[24] L. Zhang, W.-Z. Chen, Y.-J. Liu, X. Hu, K. Zhou, S. Chen, L. Peng, H. Zhu, H.-L. Zou, J. Bai, and Z.-B. Wang. Feasibility of magnetic resonance imaging-guided high intensity focused ultrasound therapy for ablating uterine fibroids in patients with bowel lies anterior to uterus. *Eur J Radiol*, 73(2):396–403, 2010.

[25] B. Denis de Senneville, M. Ries, G. Maclair, and C. Moonen. Mr-guided thermotherapy of abdominal organs using a robust PCA-based motion descriptor. *IEEE Trans Med Imaging*, 30(11):1987–1995, 2011.

[26] S. Roujol, M. Ries, B. Quesson, C. Moonen, and B. Denis de Senneville. Real-time MR-thermometry and dosimetry for interventional guidance on abdominal organs. *Magnetic Resonance in Medicine*, 63:1080–1087, 2010.

[27] R.B. Case, J.-J. Sonke, D.J. Moseley, J. Kim, K.K. Brock, and L.A. Dawson. Inter- and intrafraction variability in liver position in non-breath-hold stereotactic body radiotherapy. *Int J Radiation Oncology Biol Phys*, 75(1):302–308, 2009.

[28] C. Damianou and K. Hynynen. Focal spacing and near-field heating during pulsed high-temperature ultrasound therapy. *Ultrasound Med Biol*, 19(9):777–787, 1993.

[29] A. Schmitt, C. Mougenot, and R. Chopra. Spatiotemporal filtering of MR-temperature artifacts arising from bowel motion during transurethral MR-HIFU. *Med Phys*, 41(11), 2014.

[30] B.K.P. Horn and B.G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[31] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In *Proc. 8th European Conference on Computer Vision*, volume 4, pages 25–34. Springer, May 2004.

[32] S.A. Sapareto and W.C. Dewey. Thermal dose determination in cancer therapy. *Int. J. Radiation Oncology Biol. Phys.*, 10(6):787–800, 1984.

[33] F. Wu. Extracorporeal high intensity focused ultrasound in the treatment of patients with solid malignancy. *Minimally Invasive Therapy*, 15(1):26–35, 2006.

[34] S.N. Goldberg, G.S. Gazelle, and P.R. Mueller. Thermal ablation therapy for focal malignancy: A unified approach to underlying principles, techniques, and diagnostic imaging guidance. *American Journal of Roentgenology*, 174(2):323–331, 2000.

[35] T. S. Cook, N. Tustison, J. Biederer, R. Tetzlaff, and J. Gee. How do registration parameters affect quantitation of lung cinematics ? *Med Image Comput Assist Interv*, 10:817–824, 2007.

[36] X. Papademetris, A. J. Sinusas, D. P. Dione, R. T. Constable, and J. S. Duncan. Estimating 3D strain from 4d cine-MRI and echocardiography: In-vivo validation. In *Medical Image Computing and Computer-Assisted Intervention MICCAI 2000*, volume 1935 of *Lecture Notes in Computer Science*, pages 678–686. Springer Berlin Heidelberg, 2000.

[37] B. Denis de Senneville, C. Mougenot, and C. T. W. Moonen. Real-time adaptive methods for treatment of mobile organs by MRI-controlled high-intensity focused ultrasound. *Magn Reson Med*, 57(2):319 – 330, 2007.

[38] N. K. Ostergaard, B. Denis de Senneville, U. V. Elstrom, K. Tanderup, and T. S. Sorensen. Acceleration and validation of optical flow based deformable registration for image-guided radiotherapy. *Acta Oncol.*, 47(7):1286 – 1293, 2008.

[39] J. Vandermeulebroucke, D. Sarrut, and P. Clarysse. The popi-model, a point-validated pixel-based breathing thorax model. *In XVth International Conference on the Use of Computers in Radiation Therapy (ICCR), Toronto, Canada*, 2007.

[40] N. Kadoya, Y. Fujita, Y. Katsuta, S. Dobashi, K. Takeda, K. Kishi, M. Kubozono, R. Umezawa, T. Sugawara, H. Matsushita, and K. Jingu. Evaluation of various deformable image registration algorithms for thoracic images. *Journal of Radiation Research*, pages 1–8, 2013.

[41] R. Castillo, E. Castillo, R. Guerra, V. E. Johnson, T. McPhail, A.K. Garg, and T. Guerrero. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Phys. Med. Biol*, 54:1849–1870, 2009.

[42] K.K. Brock. Results of a multi-institution deformable registration accuracy study(MIDRAS). *Int J Radiation Oncology Biol. Phys.*, 76(2):583–596, 2010.

[43] J. Dragonu, B. Denis de Senneville, B. Quesson, C. Moonen, and M. Ries. Real-time geometric distorsion correction for interventional imaging with

echo-planar imaging (EPI). *Magnetic Resonance in Medicine*, 61:994–1000, 2009.

[44] H. Chang and J. M. Fitzpatrick. A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities. *IEEE Trans Med Imaging*, 11:319–329, 1993.

[45] T.S. Sumanaweera, G.H. Glover, T.O. Binford, and J.R. Adler. MR susceptibility misregistration correction. *IEEE Trans Med Imaging*, 12:251–259, 1993.

[46] T.S. Sumanaweera, J.R. Adler, S. Napel, and G.H. Glover. Characterization of spatial distorsion in magnetic resonance imaging and its implications for stereotactic surgery. *Neurosurgery*, 35:696–703, 1994.

[47] S.A. Reinsberg, S.J. Doran, E.M. Charles-Edwards, and M.O. Leach. A complete distorsion correction for MR images: Ii. rectification of static field inhomogeneities by similarity-based profile mapping. *Med Phys Biol*, 50:2651–2661, 2005.

[48] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J Math Imaging Vis*, 40(1):120–145, 2011.

[49] S.P.M. Crijns, J.G.M. Kok, J.J.W. Lagendijk, and B.W. Raaymakers. Towards MRI-gated linear accelerator control: Gating on an MRI accelerator. *Phys. Med. Biol.*, 56(15):4815–4825, 2012.

[50] B.W. Raaymakers, J.J.W. Lagendijk, J. Overweg, J.G.M. Kok, A.J.E. Raaijmakers, E.M. Kerkhof, R.W. van der Put, I. Meijsing, S.P.M. Crijns, F. Benedosso, M. van Vulpen, C.H.W. de Graaff, J. Allen, and K.J. Brown. Integrating a 1.5 T MRI scanner with a 6 mv accelerator: Proof of concept. *Phys. Med. Biol.*, 54(12):229–237, 2009.

CHAPTER 5

A Framework for Continuous Target Tracking During MR-Guided High Intensity Focused Ultrasound Thermal Ablations in the Abdomen

"There is no smallest among the small and no largest among the large, but always something still smaller and something still larger."
- Anaxagoras

This chapter is based on:
C. Zachiu, B. Denis de Senneville, I.D. Dmitriev, C.T.W. Moonen, and M. Ries. A Framework for Continuous Target Tracking During MR-Guided High Intensity Focused Ultrasound Thermal Ablations in the Abdomen. *J Ther Ultrasound*, 5: 27, 2017.

Abstract

During lengthy magnetic resonance-guided high intensity focused ultrasound (MRg-HIFU) thermal ablations in abdominal organs, the therapeutic work-flow is frequently hampered by various types of physiological motion occurring at different time-scales. If left un-addressed this can lead to an incomplete therapy and/or to tissue damage of organs-at-risk. While previous studies focus on correction schemes for displacements occurring at a particular time-scale within the work-flow of an MRg-HIFU therapy, in the current chapter we propose a motion correction strategy encompassing the entire work-flow.

The proposed motion compensation framework consists of several linked components, each being adapted to motion occurring at a particular time-scale. While respiration was addressed through a fast correction scheme, long term organ drifts were compensated using a strategy operating on time-scales of several minutes. The framework relies on a periodic examination of the treated area via MR scans which are then registered to a reference scan acquired at the beginning of the therapy. The resulting displacements were used for both on-the-fly re-optimization of the interventional plan and to ensure the spatial fidelity between the different steps of the therapeutic work-flow. The approach was validated in three complementary studies: an experiment conducted on a phantom undergoing a known motion pattern, a study performed on the abdomen of 10 healthy volunteers and during 3 *in-vivo* MRg-HIFU ablations on porcine liver.

Results have shown that, during lengthy MRg-HIFU thermal therapies, the human liver and kidney can manifest displacements that exceed acceptable therapeutic margins. Also, it was demonstrated that the proposed framework is capable of providing motion estimates with sub-voxel precision and accuracy. Finally, the 3 successful animal studies demonstrate the compatibility of the proposed approach with the work-flow of an MRg-HIFU intervention under clinical conditions.

In the current study we proposed an image-based motion compensation framework dedicated to MRg-HIFU thermal ablations in the abdomen, providing the possibility to re-optimize the therapy plan on-the-fly with the patient on the interventional table. Moreover, we have demonstrated that even under clinical conditions, the proposed approach is fully capable of continuously ensuring the spatial fidelity between the different phases of the therapeutic work-flow.

5.1 Introduction

Percutaneous thermal ablation of tumors has emerged as an alternate treatment option for patient groups affected by unresectable pathologies and/or are

not eligible for surgical interventions.[1] Such therapies rely on locally increasing the temperature of the pathological tissue to an extent that induces irreversible cell injury and eventually apoptosis and/or coagulative necrosis.[2] In particular, high intensity focused ultrasound (HIFU)[3–5] is currently the only percutaneous thermal ablation modality capable of non-invasive treatment delivery.[1,6] It has already met success in treating several medical conditions such as: palliation of painful bone metastases, uterine fibroids, prostate malignancies, liver tumors and several neurological diseases such as tremor-dominant Parkinson's or neuropathic pain.[7] However, HIFU thermal ablations still remain challenging when the target pathology is situated in the abdomen or lower thorax. The challenge mainly stems from the fact that therapy delivery in such areas is hampered by various types of physiological motion, occurring at different time scales:[8–10]

- Respiratory motion, for example, leads to a rapid quasi-periodic displacement of the organs in the upper abdomen and thoracic cage, with a typical frequency of 0.2 - 0.3 Hz (3 - 5 s per respiratory cycle).[11] Previous studies have addressed this type of motion through different compensation techniques such as breath-holding, gating and/or beam-steering, with each approach involving their own set of advantages and drawbacks.[12]

- Digestive activity, metabolic processes and muscle relaxation have been identified to lead to significant displacements of abdominal organs on time scales of several minutes.[13–15] For this reason, in the scope of this chapter, such motion will be referred to as slow physiological/long term drifts. The term "drift" was chosen due to the fact that contrary to respiration, these types of motion are generally of progressive nature and irreversible. The problem of digestive and peristaltic activity, in particular, can be alleviated by adjusting the patient's diet prior to the intervention[16] or by the administration of drugs such as butylscopolamine and/or glucagon acting as digestive motility inhibitors.[17,18] Long term drifts originating from other physiological sources such as bladder filling are usually addressed by the use of Foley catheters.[19,20]

- Finally, spontaneous motion due to, for example, muscle spasms is fast and infrequent, making it difficult to predict and to compensate for. It becomes particularly problematic for long interventions in absence of sedation or anesthesia, requiring the patient to lie in an uncomfortable position for lengthy periods of time. This is usually addressed by using molds or casts or by putting the patient under sedation.[21,22]

If left un-addressed, motion can lead to the therapeutic energy being diverted from the anatomy due for ablation. This not only increases the risk of undertreating the pathology but can also lead to unnecessary damage to otherwise healthy tissue.[9,23,24] Thus, a motion compensation strategy dedicated to HIFU thermal ablations in the abdomen can be beneficial for reducing the probability of such developments.

Regardless of the approach used for thermal ablation, there are multiple imaging modalities that can be used for therapy guidance,[12,25–28] with magnetic resonance imaging (MRI) possibly being the most versatile. Besides allowing precise delineation and identification of the pathology due to its superior soft tissue contrast,[29] it also provides a means to non-invasively monitor in real-time the temperature of the treated anatomy and its surroundings through a technique called MR-thermometry.[30] The MR temperature measurements are typically the main observable during MR-guided HIFU (MRg-HIFU) thermal ablations. By computing the time integral of a non-linear temperature dependent term at a particular anatomical location, thermal dose measurements can be obtained,[31] providing a mean to quantify thermal damage. The unit of measurement for thermal dose is equivalent minutes at 43°C (CEM_{43}), with an anatomy being regarded as necrotic once it exceeds 240 CEM_{43}.[31,32] However, motion-induced spatial misalignments between the MR temperature maps will most likely lead to miscalculations of the delivered thermal dose, since the associated time integral at a point in space actually includes temperature measurements from different anatomical locations. Thus, a motion compensation strategy which ensures the spatial alignment between the MR temperature maps is expected to improve the monitoring of therapy progress and effectiveness. Moreover, MR-thermometry acquisition sequences that are optimized for acquisition speed, which is usually the case for MRg-HIFU therapies, are often subject to geometric distortions.[33] This leads to a spatial inconsistency between the apparent location of a voxel in the acquired image and its true position in the imaged anatomy. Thus, a spatial misalignment between the temperature maps and the true underlying anatomy may occur. Therefore, it would be preferable that the geometric distortions, which can potentially affect the MR-thermometry images, are addressed prior to the calculation of the thermal dose.

While previous studies concerned with motion compensation during MRg-HIFU thermal therapies focused on displacements occurring at a particular time-scale,[12,34] in the current study we propose a motion correction framework that encompasses the entire intervention. Our solution consists of several linked motion compensation modules, each addressing a particular type of displacement/deformation, including:

1. A correction scheme for slow physiological motion, allowing on-the-fly adaptation of the interventional plan according to the displacements exhibited by the target anatomy.

2. A respiratory motion compensation scheme operating during HIFU energy deliveries, which ensures the spatial alignment between temperature measurements. This, in turn, is expected to lead to a more accurate evaluation of the delivered thermal dose.

3. A feature which allows therapy progress to be evaluated on-the-fly, in a spatially consistent way, on the interventional planning image(s) acquired

at the beginning of the therapy.

4. A method for correcting the geometric distortions that frequently affect the MR-thermometry images, ensuring a spatial consistency between the estimated thermal dose maps and the true underlying anatomy.

The above features were integrated in the work-flow of an MRg-HIFU thermal ablation, ensuring continuous availability of the target position over the duration of the intervention. For the remainder of this chapter percutaneous thermal ablations by the means of HIFU will be simply referred to as HIFU therapies/interventions/ablations.

5.2 Methods

5.2.1 General strategy

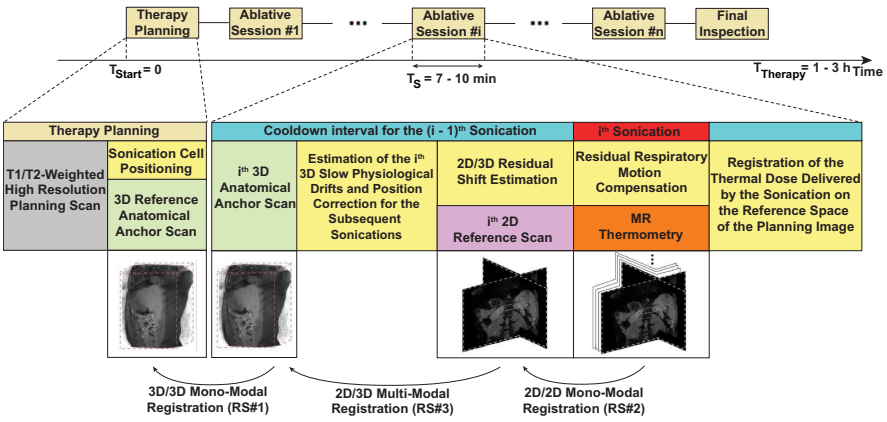


Figure 5.1: Proposed motion estimation framework for continuous target tracking during MRg-HIFU interventions.

Fig. 5.1 illustrates schematically the typical work-flow of an MRg-HIFU therapy, together with the proposed motion correction framework. The work-flow of an MRg-HIFU thermal therapy is in practice frequently episodic: bursts of energy deliveries (called sonications) are interleaved with periods of inactivity, during which the tissues in the near- and far-field are allowed to cool down. This allowed various MR-scans to be integrated at different stages of the therapy, which had the purpose of sampling the position of the treated anatomy and its surroundings over the duration of the intervention. The scans include 3D acquisitions several minutes apart (green boxes in Fig. 5.1), for sampling long-term drifts, while motion during sonications was sampled using more rapidly

acquired 2D MR images, primarily used for thermometry (orange box in Fig. 5.1). Motion estimation was achieved by comparing the acquired MR-images, using image registration,[35] to a reference image acquired at the beginning of the therapy. In order to relate all the acquired MR images to a single reference, several registration steps were employed: 1) A 3D-to-3D registration step (RS#1), for estimating long term drifts; 2) A 2D-to-2D registration step (RS#2), for estimating motion during sonications; 3) A 2D-to-3D registration step (RS#3), for estimating any residual displacement between a 2D reference image (pink box in Fig. 5.1) and its preceding 3D volume. All of these steps and the manner in which they link to each other will be described in detail during later sections.
The design of the framework includes images of two different MR contrasts: one for the 3D and one for the 2D images. For this reason, image registration was performed using two classes of algorithms: mono- and multi-modal.[36] In the scope of this study, a mono-modal registration method, namely the optical flow algorithm,[37] was employed when the compared images had the same contrast weighting (more specifically during (RS#1) and (RS#2)). For images acquired with different contrasts, a multi-modal method based on the modality independent neighborhood descriptors (MIND)[38] was used (more specifically during (RS#3)). An important feature of both the optical flow and the MIND algorithm is their capability of providing dense and elastic deformations. This is particularly beneficial for estimating the complex deformations underwent by abdominal and thoracic organs.
The resulting estimated displacements/deformations provided by the framework were used for two purposes:

1. A "down-stream" propagation of the planned sonication positions such that their updated location match the initial anatomy due for ablation.

2. An "up-stream" propagation of the thermal dose delivered by each individual sonication such that therapy progress evaluation can be made in a common reference-frame.

Both the MR-acquisitions and the energy deliveries were performed using respiratory gating via a pencil-beam navigator placed on the diaphragm.[39] This implied that images were acquired and/or the HIFU beam was turned on only when the diaphragm is close to a predefined range of locations, referred to as the gating window.

5.2.1.1 Estimation of the 3D slow physiological drifts

In order to measure the long term drifts of the target area and its surroundings, a T1-weighted 3D scan is acquired after each sonication (green boxes in Fig. 5.1), during the cool-down intervals. The drifts are then estimated by registering the 3D images via the optical flow algorithm, to a reference scan of the same size and contrast acquired at the beginning of the therapy (RS#1).

The 3D scans employed the following MR acquisition protocol: TE = 2ms, TR = 4.3ms, image matrix $192 \times 192 \times 75$, 10° flip angle, with an isotropic voxel size of $2 \times 2 \times 2\,\text{mm}^3$, resulting in an acquisition time of 60 - 90 s, depending on the frequency and reproducibility of the subject's breathing cycle. For the remainder of the chapter, this type of images will be referred to as 3D anatomical anchors.

5.2.1.2 Real-time compensation of respiratory motion

Respiratory gating was used as a first-order method for respiratory motion compensation and is expected to considerably reduce the side-effects of respiration during individual sonications, since energy deposition becomes more localized. However, due to the finite size of the gating window, residual respiratory motion might still be present. This can lead to spatial misalignments between the MR temperature maps and in turn to miscalculations of the thermal dose.
In order to estimate the residual displacements between the temperature maps, the magnitude images provided by MR-thermometry during a particular sonication (orange box in Fig. 5.1) were registered in real-time (RS#2), using the optical flow algorithm, to a common reference scan (pink box in Fig. 5.1). The resulting motion estimates were then used to spatially align the temperature measurements, thus more accurate thermal dose estimates are expected.
MR-thermometry was performed using the proton resonance frequency shift (PRFS),[30] with the acquisition sequence employing the following parameters: each scan was a single shot gradient-recalled echo, TE = 15 ms, TR = 72.5 ms, 20° flip angle, image size 160×160, voxel size $2.5 \times 2.5 \times 7\,\text{mm}^3$. Concerning spatial coverage, each dynamic consisted of one coronal and one sagittal slice intersecting in the focal point. All the scans in the MR-thermometry series were acquired with a 5 mm gating window, except the 2D reference scan used in the registration process (pink box in Fig. 5.1). For the latter, the gating window was reduced to 2 mm for increased precision, which will be explained later in this chapter.

5.2.1.3 Registration of the thermal dose measurements to a common reference

Correcting the temperature maps with respect to respiratory motion should provide improved thermal dose measurements for a particular sonication. At this point, however, each MR-thermometry series is registered to their own dedicated reference image. This leads to the resulting individual thermal dose maps being represented in their own frame-of-reference. It is, however, preferable to have all thermal dose measurements mapped into the reference frame of the planning image. In this manner, therapy progress can be monitored relative to the interventional plan, in a spatially consistent way. This mapping of the thermal dose measurements was achieved in two steps. First, the 2D reference scan was registered to its preceding 3D anatomical anchor (RS#3), in order to

account for any potential residual displacements. This was followed by a mapping of the thermal dose with respect to both these residual displacements and the 3D drifts estimated on that particular anatomical anchor, thus projecting the dose into the reference space of the planning image.
The 2D reference scan and the 3D anatomical anchor have different dimensionality, geometry and MR-contrast weighting. In effect, their registration was achieved via a two-step process. Initially, the geometry and dimensionality issue was addressed by re-formatting the 3D anatomical anchor into the coordinate system of the 2D reference scan. The re-formatting was performed by relying on geometrical information extracted from the imaging parameters (field-of-view position, orientation, size, etc.). Once the 3D anchor was re-formatted, the planes covered by the 2D reference scan were extracted from the 3D re-formatted image and the alignment was further refined via the MIND multi-modal registration algorithm.[38]

5.2.1.4 Extensions to the proposed motion correction framework

Two additional components were included in the proposed framework, which do not directly contribute towards motion compensation, having instead validation purposes. The extended framework is displayed in Fig. 5.2. Note that it includes two additional scans: one called a non-distorted EPI (purple box) and the other a sparse anatomical anchor (brown box). The non-distorted EPI scan is used in the process of estimating the geometric distortions that frequently hamper fast MR thermometry acquisitions, while the purpose of the sparse anatomical anchor is in the validation of the MIND multi-modal registration algorithm. The two added components are described in more detail within the following two paragraphs.

a Estimation of geometric distortions In the proposed protocol, MR-thermometry is based on images acquired using an echo-planar imaging (EPI) readout train. In order to have fast temperature updates, the sequence was optimized for acquisition speed, which leads to the images containing geometric distortions.[33,40] The latter manifest themselves as a mismatch between the apparent position of the anatomy in the acquires images and its true position. Thus, when registering the 2D reference scans to their preceding 3D anatomical anchors (RS#3 in Fig. 5.1), the estimated displacements will also include the geometric distortions.
In order to differentiate motion from geometric distortions we propose acquiring an additional image, which was called a non-distorted EPI scan (purple box in Fig. 5.2), having the same contrast and geometry as the 2D reference scans, but re-optimized such that distortions are minimized. An estimation of the geometric distortions is then obtained by registering via the optical flow algorithm the 2D reference scan to the non-distorted EPI (RS#4 in Fig. 5.2). The actual residual motion between the 2D reference scans and their preceding 3D anatomical anchor is now estimated using the non-distorted EPI (RS#3 in

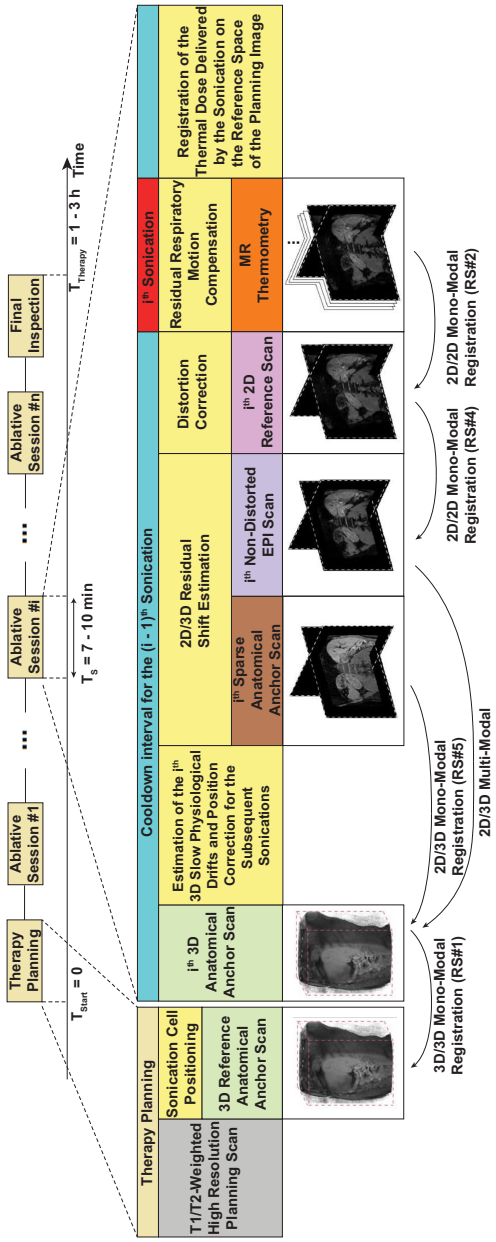


Figure 5.2: Extended motion estimation framework. In addition to the original, the extended version also includes a distortion correction scheme and a validation component for the multi-modal registration algorithm.

Fig. 5.2). The non-distorted EPI was acquired using a 2 mm gating window, immediately prior to the 2D reference scan. This also justifies why the latter was also acquired using a 2 mm gating window. Since the goal is to estimate solely geometric distortions, the potential residual motion between the 2D reference and the non-distorted EPI scans was minimized by using a narrow gating window.

b Validation of the MIND multi-modal registration algorithm Multi-modal algorithms tend to be more complex and error prone than mono-modal methods. Therefore, an independent validation procedure is proposed in this chapter for the MIND algorithm. This procedure consists in the acquisition of an additional image immediately prior to the non-distorted EPI scan, in the same geometry, but with the contrast of the 3D anatomical anchors. The newly acquired image, called a sparse anatomical anchor (brown box in Fig. 5.2), is then registered to the 3D anatomical anchor via the optical flow algorithm (RS#5 in Fig. 5.2). This is performed, however, after re-formatting the 3D anatomical anchor into the coordinate system of the sparse anatomical anchor, and the selection of the corresponding planes. By design, the motion fields obtained during RS#3 and RS#5 should be identical. The deformations provided by the optical flow algorithm during RS#5 were established as a silver standard, with the errors associated to the MIND algorithm during RS#3 being quantified in terms of the endpoint error (EE):

$$\mathrm{EE}(\vec{r}) = \|\mathbf{u}_{\mathrm{OF}}(\vec{r}) - \mathbf{u}_{\mathrm{MIND}}(\vec{r}))\|_2 \tag{5.1}$$

where $\vec{r}$ is the pixel position, $\|\cdot\|_2$ is the Euclidean distance and $\mathbf{u}_{\mathrm{OF}}$ and $\mathbf{u}_{\mathrm{MIND}}$ are the motion vectors estimated by the optical flow and MIND algorithms, respectively.

5.2.1.5 Registration algorithms

As previously mentioned, depending on whether the images being registered were acquired with the same MR-contrast weighting or not, one of two registration algorithms were employed: the optical flow[37] or the MIND[38] algorithm. The methods were chosen due to their fast numerical schemes, low number of input parameters and their capability to provide deformations on a voxel-by-voxel/pixel-by-pixel basis, characteristics which make them particularly attractive for medical image registration.
The optical flow algorithm followed the implementation described by Zachiu et *al.*[37] The approach provides the deformation between two images I and J, as the minimizer of the following functional:

$$E_{\mathrm{OF}}(\mathbf{u}) = \sum_{\vec{r}\in\Omega} \left(|I(\vec{r}) - J(\vec{r} + \mathbf{u}(\tilde{\mathbf{r}}))| + \alpha \|\vec{\nabla}\mathbf{u}(\tilde{\mathbf{r}})\|_2^2 \right) \tag{5.2}$$

where $\mathbf{u}$ is the 2D or 3D displacement, depending on the image dimensionality, Ω is the image domain, $\vec{r}$ is a pixel/voxel spatial location, $\vec{\nabla}$ is the gradient

operator, $\|\cdot\|_2$ is the Euclidean norm and α is a parameter linking the two terms of the functional. The optimization scheme and method validation are discussed at large in the original paper.[37]
The MIND algorithm, initially proposed by Heinrich et *al*,[38] is a deformable multi-modal/cross-contrast registration algorithm relying on the concept of self-similarity introduced by Buades et *al*.[41] The method associates to each pixel/voxel of an image I, a descriptor based on local similarities defined by:

$$\mathrm{MIND}(I,\vec{r},\gamma) = \frac{1}{Z}\exp\left(-\frac{D_p(I,\vec{r},\vec{r}+\gamma)}{V(I,\vec{r})}\right) \tag{5.3}$$

where Z is a normalization constant, D_p is the Euclidean distance between the local neighborhoods of size p around the pixels/voxels at positions $\vec{r}$ and $\vec{r}+\gamma$, $\gamma \in \Gamma$ is a search region of the pixels/voxels included in the descriptor and $V(I,\vec{r})$ is a local variance estimation accounting for noise perturbations. In effect, a MIND descriptor associates to each pixel/voxel of the image I a vector of size Γ. According to the MIND algorithm, the deformations between two images are found as the minimizer of the following functional:

$$E_{MIND}(\mathbf{u}) = \sum_{\vec{r}\in\Omega}\left(S(I(\vec{r}),J(\vec{r}+\mathbf{u}(\tilde{\mathbf{r}})))) + \alpha\|\vec{\nabla}\mathbf{u}(\tilde{\mathbf{r}})\|_2^2\right) \tag{5.4}$$

with

$$S(I,J) = \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}|\mathrm{MIND}(I,\vec{r},\gamma) - \mathrm{MIND}(J,\vec{r},\gamma)| \tag{5.5}$$

where I and J are the images to be registered. The optimization scheme together with implementation and validation details can be found in the original paper.[38]

5.2.2 Validation of the proposed motion estimation framework

The proposed motion estimation framework was validated in three complementary studies:

1. An experiment carried-out on a phantom undergoing a known motion pattern.

2. A study conducted on the abdomen of 10 healthy volunteers.

3. An *in-vivo* study involving MRg-HIFU thermal ablations on porcine liver, which included a total of 3 animal experiments.

5.2.2.1 Phantom experiment

The motion estimates provided by the proposed framework were initially validated on a dataset acquired during a phantom study. Motion estimation algorithms typically rely on structural information in order to estimate displacements, having to interpolate/extrapolate/infer motion in more homogenous regions. For this reason, the phantom consisted of a set of two grid-like structures (positioned in the coronal and the sagittal plane) placed inside a cylindrical plastic casing. The casing itself was then filled with a polymer gel and one of the bases was provided with an acoustic-transparent mylar membrane. Fig. 5.3(a) and 5.3(b) illustrate a coronal and a sagittal slice of a T1-weighted 3D MR image acquired on the phantom, showcasing the two integrated structures surrounded by the polymer gel. The 3D rendering in Fig. 5.3(c) displays the shape and the manner in which the two structures are positioned with respect to one-another.

Known displacements were induced to the phantom and used as gold-standard

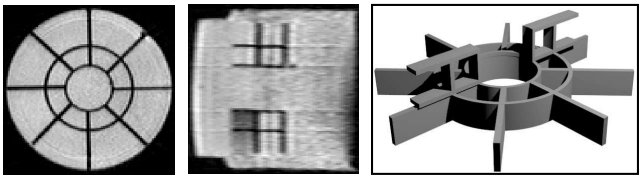


Figure 5.3: The custom-built phantom. (a) A coronal and (b) a sagittal slice of a T1-weighted 3D MR image acquired on the phantom. The two structures integrated in the phantom have a dark appearance in the images, while the polymer gel appears as bright. (c): A 3D computer-generated rendering of the two structures inside the phantom.

during the validation process. The known displacements were applied to the phantom via a motorized platform linked to an in-house developed interface which allowed the injection of custom-designed motion patterns. For the purpose of making the experiment more realistic, a pre-recorded free-breathing pattern was induced to the phantom. The pattern was recorded during a separate experiment and consisted of the average head-foot liver displacement of a healthy volunteer.

The phantom, moving according to the pre-recorded breathing pattern, underwent the extended MR image acquisition protocol illustrated in Fig. 5.2. A total of 9 3D anatomical anchors (including the reference) were acquired on the phantom, with all the other scans in between, over a duration of $\sim$1 h. During the experiment, two 5 mm drifts were induced in the breathing pattern after the acquisition of the 3$^{\text{rd}}$ and respectively 6$^{\text{th}}$ anatomical anchor, in order to simulate the effect of slow physiological motion. Note that the FOV of the 3D anatomical anchors was fixed in such a way that the phantom did not leave it

following the two induced drifts.
Besides validating the motion estimates provided by the optical flow and the MIND algorithms, the phantom experiment also allowed evaluating the extent of the geometric distortions present in the 2D thermometry images and the performance of the proposed distortion correction scheme. For this purpose, the position of the grid points contained by the two structures placed inside the phantom, was manually identified in the non-distorted EPI scans and compared, in terms of the Euclidean distance, to their position in the succeeding 2D reference scans, before and after correction. The distances before and after correction, were then placed in an individual set, and the statistical distribution of the resulting two sets were compared in order to determine the extent to which the geometric distortions were corrected by the proposed scheme.

5.2.2.2 Healthy volunteer study

A study satisfying the required standards and in conformity with regulatory requirements was carried-out on 10 healthy volunteers. The main purpose of this experiment was to evaluate the displacements underwent by the human liver and kidneys over the typical duration of an MRg-HIFU intervention. Five of the volunteers were subjected to the original MR-protocol illustrated in Fig. 5.1, while the other five were put through the extended protocol displayed in Fig. 5.2. The volunteers were placed in the MR-scanner in a prone-head-first position and were instructed not to move over the duration of the study. Each experiment lasted for a maximum of 60 min, with a minimum duration imposed by the volunteer. During this time interval, the original or the extended (depending on the volunteer) MR-acquisition protocol was run continuously, with a time gap of 7-8 min between the 3D anatomical anchors.
The 3D slow physiological drifts and the residual respiratory motion present within the MR-thermometry series were quantified for the liver and kidneys of each individual volunteer. The quantification was performed in terms of the spatial and/or statistical distribution of the following set:

$$M = \{\|\mathbf{u}(\vec{r})\|_2^2 \,|\, \vec{r} \in \text{ROI}\} \tag{5.6}$$

where $\mathbf{u}$ are the estimated displacements, $\vec{r}$ is the spatial position, $\|\cdot\|_2$ is the Euclidean norm and ROI is a region encompassing an organ of interest. Basically, M is a set containing the magnitude of the motion vectors estimated in all pixels/voxels belonging to the organs of interest. Such a metric is meant to provide the extent of the deformations undergone by the organs. Additionally, by using the metric in Eq. 5.1, validation of the motion estimates provided by the MIND cross-contrast registration algorithm was performed for the volunteers that underwent the extended MR-protocol.
In order to define the ROI in Eq. 5.6 encompassing the organs of interest, an active contour-based segmentation procedure was employed, which was subsequently manually refined. The segmentation was performed using ITK-Snap v3.0.[42]

5.2.2.3 *In-vivo* experiments

The current animal study was performed in agreement with the European law on animal experimentation and in compliance with the institution's rules for animal care and use.
In order to evaluate the compatibility of the proposed motion estimation framework with the work-flow of an MRg-HIFU intervention under clinical conditions, 3 animal experiments were conducted, consisting of *in-vivo* ablations on porcine liver. Similar to the study performed, for example, by Wijlemans *et al*, the MRg-HIFU ablations were carried-out on female Dalland pigs of 60 - 70 kg. The animals were anesthetized by an initial intramuscular injection containing ketamine (13 mg/kg), midazolam (0.7 mg/kg), atropine (0.02 mg/kg) and meloxicam (0.4 mg/kg) and placed under mechanical ventilation. Subsequently, general anesthesia was maintained by continuous intravenous administration of sufentanil (11.3 μg/kg/h), midazolam (1 mg/kg/h) and cisatracurium (0.09 mg/kg/h). A total number of 6 sonications were performed on one of the animals and 9 on the other two. The sonication cells had a 4 mm diameter, and each sonication consisted in delivering 450 W of acoustic power over a duration of $\sim$30 s, resulting in $\sim$13.5 kJ of energy per shot. The sonications were carried-out at a frequency of 1.2 MHz and a depth of $\sim$10 cm, using a modified Philips Sonalleve ablation system (Philips Healthcare, Vantaa, Finland). Time-wise, the experiments extended over a duration of $\sim$1-2h each.
In order to prevent rather large delays during the experiments that would allow naturally occurring slow physiological drifts to become significant, artificial motion was induced in the abdominal area of the animals. This was achieved by varying the volume of water within a cooling cushion placed between the mylar membrane of the Philips Sonalleve system and the skin of the animals. Technical details related to the cushion can be found in[43] and Chapter 6 of.[44] Throughout each experiment, the volume of water inside the cooling cushion was varied twice, reducing the cushion's height by $\sim$5 mm each time.
During animal experiment # 3, immediately after the last sonication, an additional 3D anatomical anchor was acquired, followed by a contrast-enhanced (CE) 3D T1w scan. The additional 3D anchor was registered to the reference 3D anchor and the resulting deformations were used to map the CE 3D T1w image into the reference space of the planning image. In this manner, the non-perfused volume (NPV) visible on the CE image, can be compared to the initial volume due for ablation and the up-stream propagated thermal dose, in a spatially consistent way. The CE 3D T1w acquisition sequence employed the following parameters: TE = 2.6ms, TR = 5.4ms, image matrix $512 \times 512 \times 53$, 10° flip angle, with a voxel size of $0.48 \times 0.48 \times 1.5\,\mathrm{mm}^3$. Prior to its mapping into the reference space of the planning image, the CE 3D T1w scan was reformatted onto the grid of the 3D anatomical anchors. This facilitated the consistent application of the deformation estimated on the last 3D anatomical anchor.
Once the experiments were finished, the animals were euthanized using an overdose of sodium pentobarbital.

5.2.3 Hardware and implementation

A multi-threaded (8 threads) C++ implementation was performed for all registration algorithms included by the proposed motion correction framework. The data resulting from the healthy volunteer study and the phantom experiment was processed in retrospect, with motion estimation and analysis being conducted on an Intel 3.2 GHz i7 workstation (8 cores) with 16 GB of RAM.
During the animal experiments, the calculations associated to the proposed framework were offloaded on a dedicated custom-build node with 32 cores and 64 GB of RAM. The implementation was performed as an additional module directly into the clinical software dedicated to the interventional radiologist delivering the therapy. The node together with access to the clinical software code was provided by Philips Healthcare, Vantaa, Finland.

5.3 Results

5.3.1 Phantom study

5.3.1.1 Validation of the optical flow mono-modal registration algorithm

A total of 9 volumes (including the reference 3D anchor) were acquired on the phantom undergoing a known motion pattern, with a 5 mm drift injected in the pattern after every third scan. In order to validate the optical flow mono-modal registration algorithm, the magnitude of the motion vectors estimated on the 3D anchors were compared to the injected drifts. The resulting EE between the estimated and the injected displacements are displayed in Fig. 5.4. The illustrated boxplots correspond to the statistical distribution of the errors estimated on each of the 3D images (except the reference scan itself). The boxplots were constructed as follows: the box limits are the 25^{th} and the 75^{th} percentiles, the whiskers correspond to the 5^{th} and the 95^{th} percentiles, the red cross indicates the average of the set and the red line is the set median. Note that, for most of the volumes, 95% of the errors remain sub-voxel (< 2 mm). It is only for two of the volumes that the 95^{th} percentile of the registration errors marginally exceeds this threshold. Also, with one exception (the first of the volumes), the average EE remains sub-millimeter. The analysis was restricted to the structures included in the body of the phantom.

5.3.1.2 Validation of the multi-modal registration algorithm

A validation of the motion estimates provided by the MIND multi-modal registration algorithm was also performed on the phantom dataset. However, instead of registering the non-distorted EPI image to its preceding 3D anatomical anchor, it was registered directly to the 3D reference anchor (see Fig. 5.2 for reference). The estimated motion vectors were then compared in terms of the

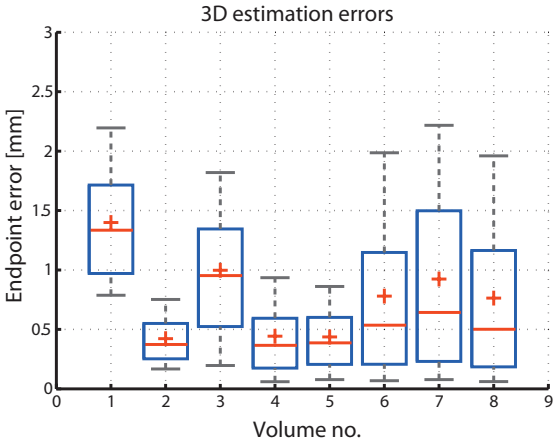


Figure 5.4: Validation of the optical flow algorithm on the phantom dataset. The boxplots correspond to the statistical distribution of the EE between the estimated motion vectors and the injected drifts for each 3D anatomical anchor acquired during the phantom experiment. The abscissa provides the index of the 3D anatomical anchor, for which the errors are displayed, within the corresponding time-series.

EE to the drifts induced in the motion pattern. The spatial distribution of the temporally averaged EE maps is displayed in Fig. 5.5(a) and 5.5(b) for a coronal and a sagittal slice respectively. It can be observed that the EE remain below the in-plane voxel size (< 2.5 mm). This is further confirmed by analyzing their statistical distribution illustrated in Fig. 5.5(c). Moreover, the two boxplots in Fig. 5.5(c) show that 95% of the errors remain sub-millimeter with an average of $\sim$0.5 mm. Note that the analysis was restricted to the two structures embedded in the phantom. The borders of the phantom were also excluded from the analysis, due to a signal drop in the non-distorted EPI images, particularly visible in Fig. 5.5(b).

5.3.1.3 Quantification and correction of the geometric distortions present in the MR-thermometry images

Fig. 5.6(a) and 5.6(b) display a coronal and a sagittal slice from an MR-thermometry image acquired on the phantom, before distortion correction. It can be observed that some segments of the structures inside the phantom appear to be bent, when in reality the structures are made of straight elements. This effect is notably reduced after distortion correction, as shown in Fig. 5.6(c) and 5.6(d).
The grid-like shape of the structures integrated in the body of the phantom, in both the coronal and sagittal plane, allowed the quantification of the geo-

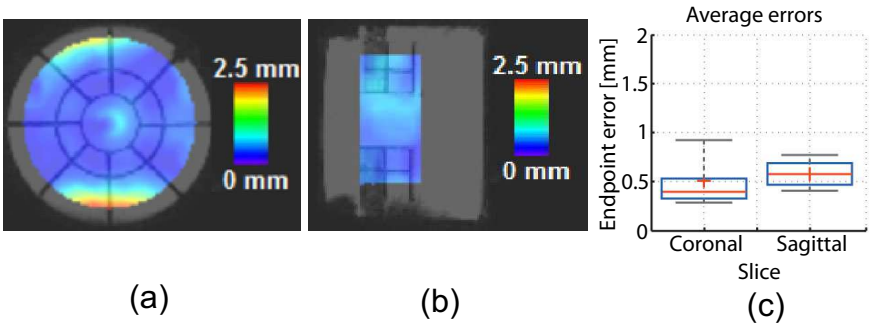


Figure 5.5: Validation of the multi-modal registration algorithm on the moving phantom data set. The spatial distribution of the temporally averaged EE in a (a): coronal (b): sagittal plane through the phantom (c): Statistical distribution of the errors illustrated in (a) and (b).

metric distortions present in the MR-thermometry images and also the extent to which these are corrected by the proposed method. The boxplots in Fig. 5.6(e) display the extent of the geometric distortions before and after applying the proposed correction scheme. Measurements were performed for the grid points marked with "x" in Fig. 5.6(a) and 5.6(b), for all the 2D reference - non-distorted EPI pairs acquired on the phantom. It can be observed that, in the absence of correction, distortions extend up to ∼6 mm. After correction, however, distortions were reduced to in-plane voxel size values (< 2.5 mm), with an average reduced from ∼5 mm to ∼1 mm.

5.3.2 Volunteer study

5.3.2.1 Analysis of the 3D slow physiological drifts

Fig. 5.7 illustrates the temporal evolution of the long term drifts estimated on the 10 healthy volunteers. Fig. 5.7(a) and 5.7(b) show, separately for the liver and kidneys, the statistical distribution of the magnitude of the 3D motion vectors, pooled from all volunteers, at each 3D anatomical anchor acquisition time point. It can be observed, for both the liver and kidneys, that there is a tendency of the displacements to increase over time, with a magnitude of the motion vectors exceeding 7 mm at the acquisition time point of the last 3D anatomical anchor. Fig. 5.7(c) and 5.7(d) display separately for the liver and kidneys, the average magnitude of the displacement vectors over time, individually for each of the 10 volunteers. The different length of some curves compared to others is due to the respective volunteers terminating the experiment before the 60 min maximum observation period. A rather large inter-subject variability can be observed in both the liver and kidneys. In volunteer #4, for example, the average displacement remained under 2 mm for the entire duration of the

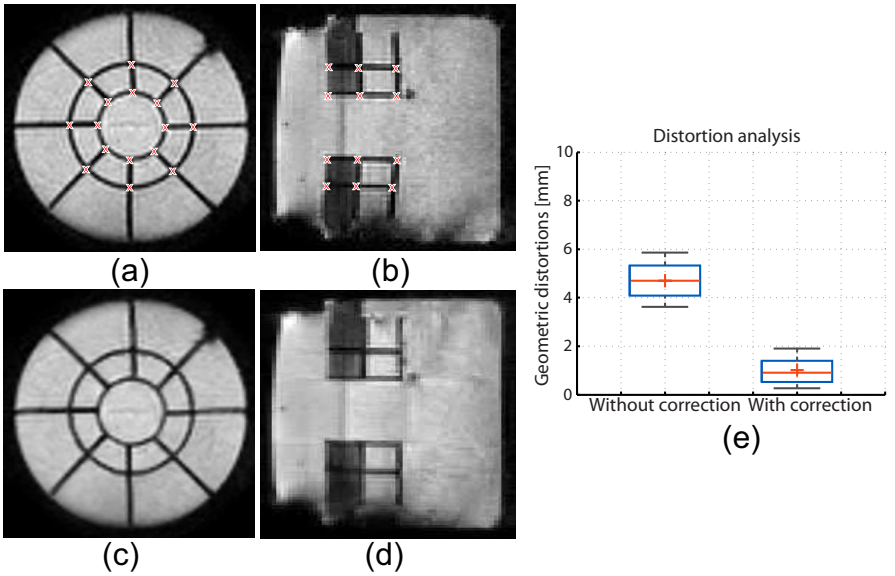


Figure 5.6: Validation of the proposed distortion correction scheme employed for the MR-thermometry images. Example of a coronal (left) and a sagittal (middle) MR-thermometry magnitude image acquired on the phantom (a), (b): before and (c), (d): after distortion correction. (e): Extent of the geometric distortions measured for the grid points marked with "x" in (a) and (b), before (left boxplot) and after (right boxplot) correction.

study, for all organs of interest. However, in volunteer #1, already halfway through the experiment, the average displacement exceeded 6 mm.

5.3.2.2 Assessment of the residual respiratory motion present in the MR-thermometry series

For each dynamic of the multiple MR-thermometry series acquired on the healthy volunteers, the average liver and kidney residual respiratory displacement was estimated and pooled in a separate set for each volunteer. Fig. 5.8(a) and 5.8(b) display, per individual, the statistical distribution of the average magnitude of the 2D motion vectors corresponding to the residual respiratory displacements, separately for the liver and the kidneys. While the average displacements remain close to 1 mm, they occupy a rather large range of values, in some instances exceeding 4 mm. The extent of the residual motion and the inter-individual variations are most likely determined by the stability/reproducibility of the breathing cycle of each individual volunteer.

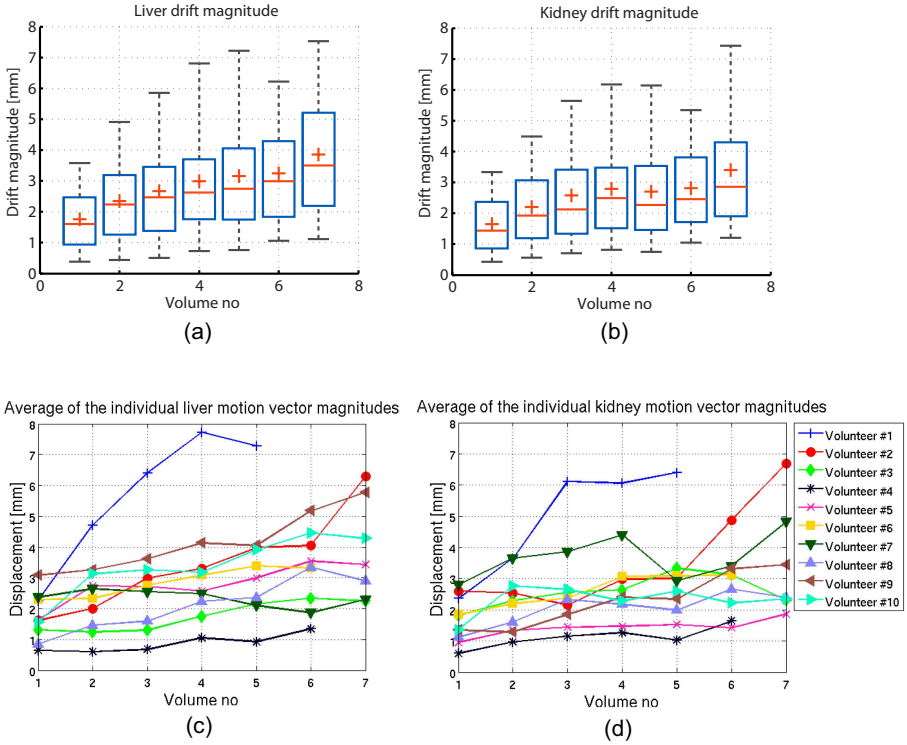


Figure 5.7: Temporal evolution of the 3D slow physiological drifts estimated on the 10 healthy volunteers. (a), (b): Statistical distribution of the magnitude of the 3D motion vectors estimated on the liver and kidneys, pooled from all volunteers, at each 3D anatomical anchor acquisition time point. (c), (d): Time evolution of the average magnitude of the motion vectors estimated on the liver and kidneys, illustrated individually for each of the 10 healthy volunteers. The abscissa in Fig. (a) - (d) provides the index of the 3D anatomical anchor, for which the displacements are displayed, within the time-series.

5.3.2.3 Validation of the multi-modal registration algorithm

As specified in the methods section, in order to validate the MIND registration algorithm, an endpoint error map (see Eq. 5.1) was calculated for each sparse anatomical anchor - non-distorted EPI image pair acquired on the volunteers (see Fig. 5.2 for scan nomenclature). Fig. 5.9(a) and 5.9(b) illustrate for one of the volunteers, the spatial distribution in the organs of interest of the temporally averaged EE maps. It can be observed that, for this particular volunteer, the estimation errors for the MIND multi-modal algorithm remain under 2.5 mm, which corresponds to the in-plane voxel size. The pixel-wise EE were

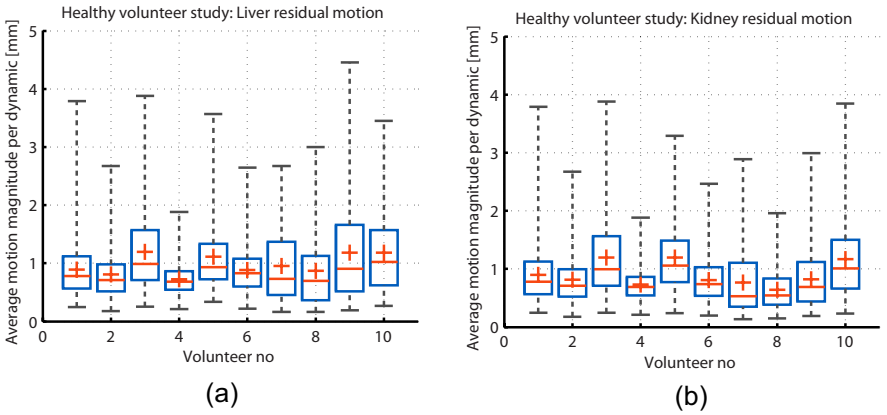


Figure 5.8: Quantification of the residual respiratory motion for the 10 healthy volunteers. Statistical distribution of the average magnitude of the motion vectors estimated for (a): liver (b): kidneys.

pooled separately for the liver and kidneys of each volunteer and illustrated under the shape of a boxplot in Fig. 5.9(c) and 5.9(d). For all volunteers, the estimation errors associated to the MIND algorithm reside beneath the in-plane voxel size.

5.3.3 Animal experiments

The proposed motion correction framework was validated under clinical conditions during 3 separate *in-vivo* MRg-HIFU thermal ablations conducted on porcine liver. Fig. 5.10(a) illustrates the therapy planning for animal experiment #3, showcasing a coronal (left) and respectively a sagittal (right) slice through the 3D planning image, upon which the 9 sonication cells are overlaid. For better visibility, the interventional plan was magnified and shown in Fig. 5.10(b). For this particular experiment, an artificial motion event was induced in the abdominal area of the animal after every 3 sonications. In effect, the initially planned sonication positions were updated twice during the experiment, according to the displacements estimated on the 3D anatomical anchors after each artificial motion event. Fig. 5.11(a) illustrates the "down-propagated" sonication cells overlaid as blue ellipses on a coronal and a sagittal slice from the planning image. Following the two motion events, displacements of over 5 mm can be observed for all initially planned sonication cells. The red overlay represents the lethal thermal dose accumulated from all sonications, without mapping the thermal dose delivered by the individual sonications into the reference space of the planning image. This provides the means to evaluate the effects of motion on the outcome of the therapy, in the absence of the proposed motion compensation framework. This scenario is better illustrated in

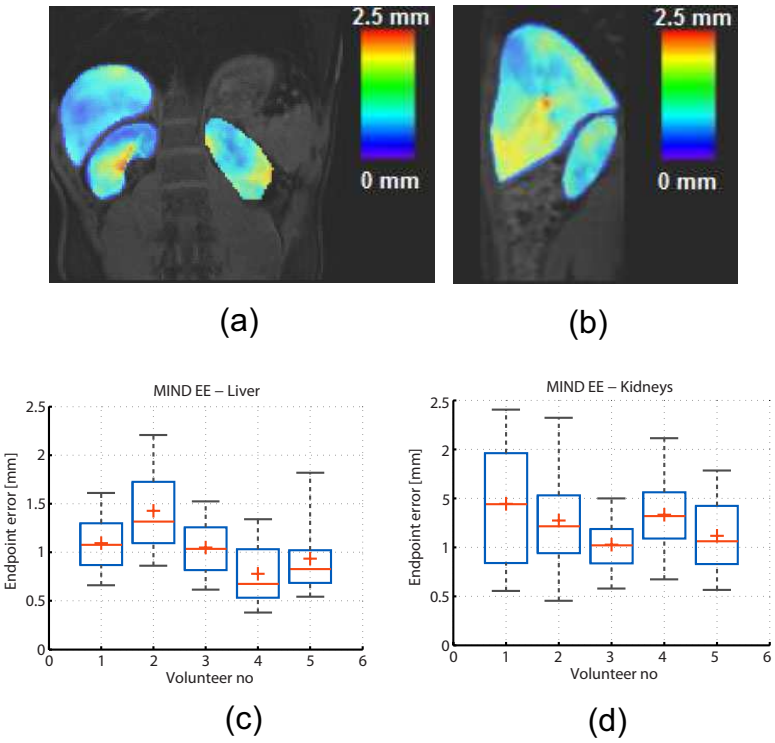
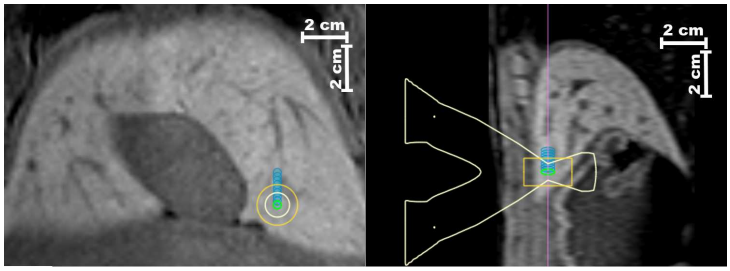


Figure 5.9: Validation, on the healthy volunteer dataset, of the MIND registration algorithm. (a), (b): Spatial distribution of the temporally averaged registration errors in the liver and kidneys of one of the volunteers, in a coronal and a sagittal plane respectively. (c), (d): Statistical distribution of the multi-modal registration errors in the liver and kidneys of each volunteer.

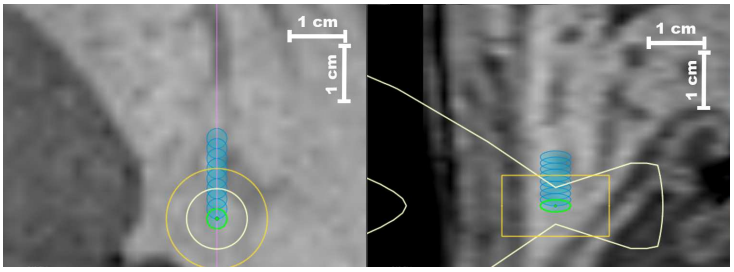
Fig. 5.11(b) where the initial plan is overlaid as blue ellipses on the planning image, together with the non-registered lethal thermal dose in red. In the absence of a motion compensation strategy, a large part of the anatomy initially due for ablation would have been left untreated, while at the same time the therapy would have resulted in considerable collateral damage. Following the "up-stream" propagation and accumulation of the thermal dose delivered by each individual sonication, the result shown in Fig. 5.11(c) was obtained. The initially planned location of the sonications is displayed as blue ellipses overlaid on the interventional planning image, together with the motion corrected lethal thermal dose. The good overlap between the latter and the initial plan indicates that the proposed motion compensation framework performed as intended. Fig. 5.11(d) displays the initially planned sonication cells overlaid as blue ellipses on a CE T1w image (acquired exclusively on animal #3), after its registration to the planning image. A good correspondence can be observed between the

NPV, the initial volume due for ablation and the registered lethal thermal dose map displayed in Fig. 5.11(c). Note that the contrast of the Fig. 5.11(d) was digitally enhanced for improved NPV visibility. In all images from Fig. 5.10 and 5.11, the white overlay traces an approximation of the HIFU beam cone, while the yellow overlay defines a search region within which the HIFU system checks for the existence of a focal spot.

Table 5.1 reports, for the animal experiments, the percentage of the anatom-



(a)



(b)

Figure 5.10: Therapy plan for animal experiment #3. Coronal (left) and sagittal (right) slice through the 3D planning image together with the 9 sonication cells (a): overall (b): magnified for better visibility.

ical volume due for ablation that was estimated to receive a lethal amount of thermal dose, with and without enabling the proposed motion compensation framework. Although to different extents, when motion compensation is enabled, improvements in coverage were observed in all reported cases.

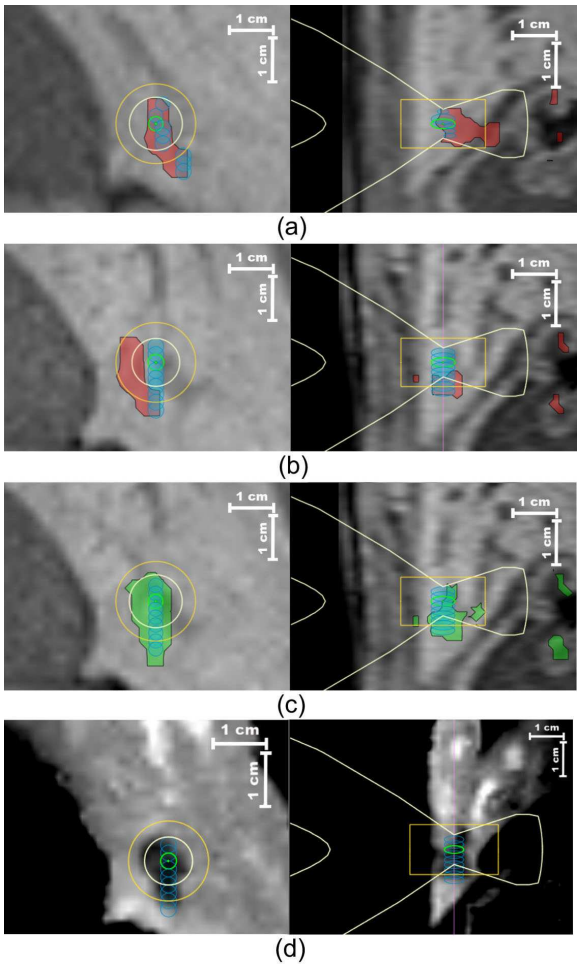


Figure 5.11: Validation of the proposed motion compensation framework under clinical conditions: Outcome of the MRg-HIFU liver ablation in animal experiment #3. (a): Down-propagated/Motion corrected sonication locations (blue ellipses) overlapped with the non-registered lethal thermal dose (in red). (b): Originally planned sonication locations (blue ellipses) overlapped with the non-registered lethal thermal dose (in red). (c): Originally planned location of the sonications (blue ellipses) overlapped with the motion corrected lethal thermal dose (in green). (d): The initially planned sonication cells (blue ellipses) overlaid on a coronal (left) and a sagittal (right) slice of the registered contrast-enhanced T1w image. Figures (a), (b) and (c) all showcase a coronal (left image) and a sagittal (right image) slice through the 3D planning image acquired at the beginning of the therapy as background.

	Animal #1	Animal #2	Animal #3
With motion correction	88%	74%	70%
Without motion correction	36%	63%	40%

Table 5.1: Validation of the proposed motion compensation framework under clinical conditions. Percentage of the volume due for ablation estimated to have received a lethal amount of thermal dose with (first row) and without (second row) the proposed motion compensation framework enabled.

5.3.4 Computational performance of the registration algorithms

In order to ensure a smooth work-flow for the MRg-HIFU intervention, the registration algorithms included by the proposed framework should provide positional information with minimal latency. Table 5.2 reports the average convergence time for each of the registration steps included by the framework. See Fig. 5.1 for details concerning the purpose of each algorithm.

Step	Average convergence time [ms]
Estimation of 3D drifts (RS#1 in Fig. 5.1)	15000
Estimation of residual respiratory motion (RS#2 in Fig. 5.1)	80
Registration between the 2D reference image and its preceding 3D anchor (RS#3 in Fig. 5.1)	400

Table 5.2: Average convergence time per pair of images of the registration algorithms included by the proposed motion compensation framework.

5.4 Discussion

Lengthy MRg-HIFU thermal therapies in the abdomen are usually hampered by various types of physiological motion occurring at different time-scales. So far, studies have focused on developing correction schemes for displacements arising at a particular time-scale[12, 34] which, as we have demonstrated in the present chapter, may be conceptually insufficient. In effect, the present study proposes a motion compensation framework that encompasses the entire work-flow of an MRg-HIFU intervention. The framework consists of several linked components, each dedicated to estimating a particular type of motion/deformation, with the resulting displacements being used for two purposes: 1) Down-stream propagation of the initially planned sonication locations such that they match the current position of the anatomy and 2) Up-stream propagation of the thermal dose delivered by each individual sonication such that therapy progress

can be evaluated in a single frame of reference (namely the reference space of the planning image(s)). Due to its modular nature, the proposed approach has increased flexibility, facilitating the addition, modification, replacement and/or removal of individual components. The proposed motion compensation strategy was tested and validated in three complementary experiments: 1) An experiment carried-out on a phantom undergoing a known motion pattern; 2) A study conducted on the abdomen of 10 healthy volunteers and 3) An *in-vivo* study involving HIFU ablations on porcine liver.

The healthy volunteer study reconfirmed that over a duration of 1h, the human liver and kidneys can manifest slow physiological drifts of up to 7 - 8 mm, exceeding acceptable therapeutic margins. This is in good correspondence with previous reportings.[13,45,46] Additionally, a rather large inter-subject variability was observed. Concerning the proposed correction scheme for respiratory motion, the study conducted on the 10 volunteers demonstrated that gating during energy deliveries, as a first order method, has good motion compensation capabilities, with estimated average residual displacement of ∼1-1.5 mm. Of importance are, however, the instances in which the estimated average residual displacements extend up to 4 mm or more (see Fig. 5.8). Due to the misalignments induced between the temperature maps, such displacements during MR-thermometry can impact the thermal dose measurements to an extent that they become unreliable for that particular sonication. A simple solution to reduce the range of residual respiratory displacements is to narrow-down the size of the gating window. However, depending on the reproducibility of the patient's breathing cycle, this can lead to a poor duty cycle of the HIFU beam, affecting the overall therapy efficiency. The proposed dedicated registration scheme, on the other hand, allows energy deliveries with a wide gating window ($\geq 5mm$), facilitating a higher duty cycle not only for the HIFU beam, but also for the MR-thermometry.

The reliability of the motion estimates provided by the proposed framework was analyzed in both the phantom and the volunteer experiment. Following the phantom study, both the mono- and the multi-modal registration algorithms have proven on average sub-voxel accuracy and precision. In order to avoid a bias due to outliers during the validation of the multi-modal registration algorithm, the borders of the phantom were excluded from the analysis (see Fig. 5.5). This was performed due to a signal drop in these areas in the non-distorted EPI images, signal drops which were not present in the reference 3D anatomical anchor. Due to a violation of the basic assumption made by the MIND algorithm, that all structures in the reference image have a counterpart in the moving image, the reliability of the motion estimates was poor on the phantom borders. Since this aspect is known a priori, these low-signal areas are not representative for the algorithm's performance. Instead they simply emphasize some of its limitations. The phantom experiment, however, has only limited validation capabilities since the phantom was able to undergo motion with fewer degrees of freedom than an actual abdominal organ. On the other hand, while the known motion patterns induced to the phantom can be used

as a robust gold standard when analyzing the performance of the registration algorithms, obtaining a gold standard for *in-vivo* studies is a challenging task. In particular for the cross-contrast registration algorithm, this issue was addressed by comparing the *in-vivo* motion estimates against a silver standard. The latter was constructed based on motion estimates provided by the optical flow algorithm. This decision was made due to its prior successful *in-vivo* validation in previous independent studies.[47–50] However, the drawback of such an approach is that the optical flow algorithm has its own shortcomings which affect its performance (discussed at length in[37]). Thus, the *in-vivo* errors reported for the multi-modal algorithm stem from both its own mis-registrations and the errors in the silver standard itself. Nevertheless, the overall estimation errors remain sub-voxel for 3D registration and lower than the in-plane voxel size for the 2D registration methods, which is in good correspondence with previous studies.[38,47–50] Potential errors may also occur during the registration of the 2D reference scan to its preceding 3D anatomical anchor (RS#3 in Fig. 5.1) or during the registration of the non-disorted EPI scan and the sparse anatomical anchor, again, to their preceding 3D anatomical anchor (RS#3 and RS#5 in the extended framework from Fig. 5.2). This is due to the fact that the elastic refinement of the registration between the 2D scans and the corresponding planes from the reformatted 3D anatomical anchor (following the initial rigid alignment step), was only performed in 2D. In case of severe through-plane motion, misregistration may occur, since through-plane motion might be interpreted as in-plane motion. The risk of such a development is, however, considerably reduced since the 3D anatomical anchor and the 2D reference scan, sparse anatomical anchor and the non-distorted EPI scan are respiratory gated and acquired in rapid succession. In the scope of this study, this led to residual displacements predominantly in the cranio-caudal direction, with the anterior-posterior component being well under the voxel size. For the estimation of through-plane deformations, a 3D dense and elastic registration between the 2D scans and their preceding 3D anatomical anchor would be necessary. This is, from a mathematical point-of-view, a severely ill-posed problem and a topic in itself, making it the object of future studies.

Since the proposed motion compensation framework consists of several linked components, with the estimated displacements in some instances being successively added to one another, error accumulation becomes an important aspect. While slow physiological drifts and residual respiratory displacements are the result of independent registrations between a scan and its corresponding reference image, with sub-voxel/sub-pixel estimation errors, projecting the thermal dose delivered by a particular sonication onto the reference space of the planning image relies on a chain of up to 4 registration algorithms (see Fig. 5.2). However, even so, the accumulated estimation errors remain within $10 - 20\%$ of the total average displacement.

An effective way to improve the precision and accuracy of the registration algorithms is to increase the spatial resolution of the acquired MR images. A higher spatial resolution typically implies a greater level of detail and structural infor-

mation in the images, which facilitates a better performance of the registration algorithms. Note, however, that MR imaging usually implies a tradeoff between spatial resolution, temporal resolution and signal-to-noise ratio (SNR). For a smooth work-flow of an MRg-HIFU therapy, constrains may have to be imposed on the image acquisition times. Therefore, a higher spatial resolution may result in a lower SNR of the acquired images. Particular attention is thus required, in order to ensure that the losses in terms of SNR do not counteract the gains in terms of precision and accuracy facilitated by a higher spatial resolution.

The overall performance of the proposed motion compensation framework together with its compatibility with the typical work-flow of an MRg-HIFU intervention under clinical conditions was validated during 3 animal experiments. In all 3 cases, the mapping of the thermal dose in the reference space of the planning image, with respect to the displacements estimated by the framework, resulted in different amounts of improvement between the planned sonication locations and the lethal thermal dose. Moreover, as shown in Fig. 5.11(c) and 5.11(d), for the animal for which a CE 3D T1w image was acquired at the end of the HIFU ablation session, a good correspondence can be observed between the planned location of the sonication cells, the registered lethal thermal dose and the registered NPV. This further confirms the success the proposed motion correction framework. Although in the scope of this study a CE T1w image was acquired only for the third animal, the visible NPV upon such images can generally be used as a metric for evaluating acute therapeutic response. Furthermore, in the context of motion correction schemes for MRg-HIFU, the NPV pattern can be used for additional validation, as demonstrated in animal experiment #3.

The majority of the MR-scans and registration algorithms included by the framework were integrated during the cool-down intervals between successive sonications. In order to ensure a smooth therapeutic work-flow, the acquisition times of the scans together with the computational requirements of the registration algorithms must not exceed typical cool-down durations. For the animal experiments, each sonication consisted in the delivery of $\sim$13.5 kJ of energy, which led to the HIFU system imposing cool-down intervals of 2-5 min, determined by the perfusion effects in the near-field. In practice, more typical values are 5 kJ of energy per sonication with 2-3 min cool-down. For the proposed framework, the duration of the MR-scans together with the registration algorithms integrated during the cool-down intervals resulted in average delays of 2-3 min, depending on the subject's respiratory frequency. For the purpose of this study this was sufficient since such a duration is well in accordance with the cool-down threshold imposed by the HIFU system. Nevertheless, the protocol can be further accelerated if necessary. For example, this can be achieved by re-optimizing the acquisition sequence of the 3D anatomical anchors for speed rather than resolution and spatial coverage, which were favored in the current study.

In interventional oncology, pathology identification and delineation, interven-

tional planning, therapy monitoring and therapy response evaluation are preferably performed in the same frame of reference. Especially in moving organs, this can become problematic for HIFU thermal ablations, since physiological motion frequently induces spatial mismatches between these steps. The current study aims to render the work-flow of an MRg-HIFU therapy in mobile organs compatible with pre-existing work-flows from interventional oncology by proposing a suitable motion estimation/correction strategy that encompasses all the previously specified phases of an intervention. Results have shown that physiological drifts of 7-8 mm have to be expected when therapy is conducted in the liver or kidneys, displacements which, if left unaddressed, can have severe consequences. For example, as illustrated in Fig. 5.11(b), in case the ablation area is situated in the proximity of the gallbladder, there is a high risk that the latter and/or the associated structures (such as the bile duct) are perforated/damaged, leading to complications due to the possible release of emulsifying enzymes into the bloodstream. Or, if therapy is conducted in the kidney, damage to the pelvis or the ureter might occur. Moreover, such displacements might lead to large areas of residual pathological tissue. Such a case is again depicted in Fig. 5.11(b), where only 40 % of the initially planned anatomy would have been ablated (see also table 5.1). However, the animal experiments have demonstrated that the risk of such developments can be considerably reduced when therapy is conducted with the proposed motion compensation framework active. As shown in table 5.1, in all analyzed cases the framework led to improvements of the overlap between the planned anatomy due for ablation and the delivered lethal thermal dose, compared to the scenario when no corrections are performed. Moreover, the framework demonstrated good compatibility with the typical work-flow of an MRg-HIFU thermal therapy, with latencies introduced by the integrated MR scans and the registration algorithms that allowed a smooth progress of the intervention. Additionally, the fact that the proposed motion compensation strategy was implemented directly into the clinical software stack granted the possibility of on-the-fly re-optimization of the therapeutic plan according to the estimated displacements and the projection of the thermal dose delivered by each individual sonication in the same frame of reference (see Fig. 5.11). Particular attention was paid during the implementation of the framework such that the work-flow of an intervention conducted on (quasi-)static anatomies remains unchanged, with the additional motion compensation features being hidden to the radiologist. Noteworthy is also the fact that the implementation and execution of the proposed framework was performed using commercially available hardware, all being integrated on an existing HIFU platform in conjunction with a standard 1.5 T Achieva MRI. One of the drawbacks of the proposed motion correction framework is the lack of a component that addresses spontaneous motion. Thus, the latter is considered to be circumvented by some other means such as patient sedation. Problematic might also be the fact that the framework does not include an objective quality evaluation criterion for neither the MR images used for tracking, nor for the estimated displacements. While in the scope of this study, during the *in-vivo*

experiments, a visual inspection of the MR images and the resulting estimated displacements was deemed sufficient, future studies need to address this issue by establishing an objective quality evaluation criterion. In addition, note that all deformations estimated and corrected by the proposed framework rely entirely on image registration algorithms. However, in particular for correcting geometric distortions, more specialized correction schemes can be employed. For example, the problem can be entirely avoided by an optimization of the acquisition parameters, such that it still allows thermometry with a high update rate, while at the same time minimizing geometric distortions. It is, however, difficult to predict whether such alternative approaches perform better than the proposed registration-based method, making this again the object of future studies.

5.5 Conclusion

The present study proposes a motion correction framework encompassing the entire work-flow of an MRg-HIFU thermal therapy, ensuring spatial coherence between the different stages of the therapeutic work-flow. It was demonstrated that the framework allows both the adjustment of the interventional plan and projection of the therapy observables (e.g. temperature and thermal dose measurements) in a common frame-of-reference on-the-fly, with the patient on the interventional table. Furthermore, the proposed motion correction strategy was implemented as an additional feature directly into the clinical software stack, while at the same time maintaining compatibility with MRg-HIFU therapies for static anatomies.

Acknowledgements

This work was supported by the Dutch Technology Foundation (STW) (project OnTrack no. 12813) and in part by the European Research Council (project ERC-2010-AdG-20100317, Sound Pharma) and ITEA 2 (project SoRTS).

Bibliography

[1] K.F. Chu and D. E. Dupuy. Thermal ablation of tumours: biological mechanisms and advances in therapy. *Nature Reviews Cancer*, 14:199–208, 2014.

[2] E. G. Moros. *Physics of Thermal Therapy: Fundamentals and Clinical Applications*, chapter 2, pages 23–41. CRC Press Taylor and Francis Goup, Boca Raton, Florida, 1 edition, 12 2012.

[3] H.E. Cline, J.F. Schenck, K. Hynynen, R.D. Watkins, S.P. Souza, and F.A. Jolesz. MR-guided focused ultrasound surgery. *J Comput Assist Tomogr*, 16:956–965, 1992.

[4] K. Hynynen, W.R. Freund, H.E. Cline, A.H. Chung, R.D. Watkins, J.P. Vetro, and F.A. Jolesz. A clinical, noninvasive, MR imaging-monitored ultrasound surgery method. *Radiographics*, 16:185–195, 1996.

[5] R.O. Illing, J.E. Kennedy, G.R. ter Haar, A.S. Protheroe, P.J. Friend, F.V. Gleeson, D.W. Cranston, R.R. Phillips, and M.R. Middleton. The safety and feasibility of extracorporeal high-intensity focused ultrasound (HIFU) for the treatment of liver and kidney tumours in a western population. *Br J Cancer*, 93(8):890–895, 2005.

[6] J. E. Kennedy. High-intensity focused ultrasound in the treatment of solid tumours. *Nature Reviews Cancer*, 5:321–327, 2005.

[7] C.M. Tempany, N.J. McDannold, K. Hynyen, and F.A. Jolesz. Focused ultrasound surgery in oncology: Overview and principles. *Radiology*, 36(5):39–56, 2011.

[8] M. Goitein. Organ and tumor motion: An overview. *Seminars in Radiation Oncology*, 14:2–9, 2004.

[9] M. Ries, B. Denis de Senneville, S. Roujol, Y. Berber, B. Quesson, and C. Moonen. Real-time 3D target tracking in MRI guided focused ultrasound

ablations in moving tissues. *Magnetic Resonance in Medicine*, 64:1704–1712, 2010.

[10] S. Vijayan, S. Klein, E.F. Hofstad, F. Lindseth, B. Ystgaard, and T. Lango. Validation of a non-rigid registration method for motion compensation in 4D ultrasound of the liver. *IEEE 10th International Symposium on Biomedical Imaging*, pages 792–795, 2013.

[11] Y. Seppenwoolde, H. Shirato, K. Kitamura, S. Shimizu, M. van Herk, J. V. Lebesque, and K. Miyasaka. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int. J. Radiation Oncology Biol. Phys.*, 53(4):822–834, 2002.

[12] J.-M. Escoffre and A. Bouakaz. *Therapeutic Ultrasound*, volume 880, chapter 3, pages 43–65. Springer, Switzerland, 2016.

[13] M. von Siebenthal and G. Szekely. Systematic errors in respiratory gating due to intrafraction deformations in the liver. *Med. Phys.*, 34(9):265–278, 2007.

[14] K.M. Langen, T.R Willoughby, S.L. Meeks, A. Santhanam, A. Cunningham, L. Levine, and P.A. Kupelian. Observations on real-time prostate gland motion using electromagnetic tracking. *Int J Radiation Oncology Biol. Phys.*, 71(4):1084–1090, 2008.

[15] P. Arnold, F. Preiswerk, B. Fasel, R. Salomir, K. Scheffler, and P.C. Cattin. 3D organ motion prediction for MR-guided high intensity focused ultrasound. *Med Image Comput Comput Assist Interv*, 14(Pt 2):623–630, 2011.

[16] M.H.P. Smitmans, F.J. Pos, J. de Bois, W.D. Heemsbergen, J.-J. Sonke, J.V. Lebesque, and M. van Herk. The influence of a dietary protocol on a cone beam CT-guided radiotherapy for prostate cancer patients. *Int J Radiation Oncology Biol Phys*, 71(4):1279–1286, 2008.

[17] J. Emmott, B. Sanghera, J. Chambers, and W.L. Wong. The effects of n-butylscopolamine on bowel uptake: An F-FDG PET study. *Nuclear Medicine Communications*, 29:11–16, 2008.

[18] J. M. Froehlich, M. Daenzer, C. von Weymarn, S. M. Eturk, C. L. Zolikofer, and M. A. Patak. Functional monitoring of small bowel motility: Comparison of spasmolysis induced by glucagon or buscopan. In *International Society for Magnetic Resonance in Medicine*, pages 1122–1123, 2009.

[19] R. Mirabell, P. Nouet, M. Rouzaud, A. Bardina, N. Heijira, and D. Schneider. Radiotherapy of bladder cancer: relevance of bladder volume changes in planning boost treatment. *Int J Radiation Oncology Biol Phys*, 41(4):741–746, 1998.

[20] J. Thariat, S. Aluwini, and M. Housset. Adaptive partial-boost stereotactic radiation therapy for muscle-invasive carcinoma of the urinary bladder. In L. E. Ponsky, D.B. Fuller, R.M. Meier, and C. Ma, editors, *Robotic Radiosurgery. Treating prostate cancer and related genitourinary applications*, pages 185–200. Springer Berlin Heidelberg, 2012.

[21] L.J. Verhey. Immobilizing and positioning patients for radiotherapy. *Seminars in radiation oncology*, 5(2):100–114, 1995.

[22] Y.-F. Zhou. High intensity focused ultrasound in clinical tumor ablation. *World J Clin Oncol*, 2(1):8–27, 2011.

[23] V. Auboiroux, L. Petrusca, M. Viallon, T. Goget, C. D. Becker, and R. Salomir. Ultrasonography-based 2d motion-compensated HIFU sonication integrated with reference-free mr temperature monitoring: A feasibility study ex vivo. *Phys Med Biol*, 57(10):N159–71, 2012.

[24] A. B. Holbrook, P. Ghanouni, J. M. Santos, C. Dumoulin, Y. Medan, and K. B. Pauly. Respiration based steering for high intensity focused ultrasound liver ablation. *Magn Reson Med*, 71(2):797–806, 2014.

[25] J. E. Kennedy, F. Wu, G. R. ter Haar, F. V. Gleeson, R. R. Phillips, M. R. Middleton, and D. Cranston. High-intennsity focused ultrasound for the treatment of liver tumours. *Ultrasonics*, 42(1-9):931–935, 2004.

[26] D. Liu and E S. Ebbini. Real time 2-d temperature imaging using ultrasound. *IEEE Transactions on Bio-Medical Enginnering*, 57(1):12–16, 2010.

[27] B. J. Wood, J. K. Locklin, A. Viswanathan, J. Kruecker, D. Haemmerich, J. Cebral, A. Sofer, R. Cheng, E. McCreedy, K. Cleary, M. J. McAuliffe, N. Glossop, and J. Yanof. Technologies for guidance of radiofreequency ablation in the multimodality interventional suite of the future. *J Vasc Interv Radiol*, 18(1):9–24, 2007.

[28] M. Cheng, M. Fay, and K. Steinke. Percutaneous CT-guided thermal ablation as salvage therapy for recurrent non-small cell lung cancer after external beam radiotherapy: A retrospective study. *International Journal of Hyperthermia*, 32(3):316–323, 2016.

[29] D. V. Sahani, M. A. Bajwa, Y. Andrabi, S. Bajpai, and J. C. Cusack. Current status of imaging and emerging techniques to evaluate liver metastases from colorectal carcinoma. *Annals of Surgery*, 259(5):861–872, 2014.

[30] V. Rieke and K. B. Pauly. MR-thermometry. *J Magn Reson Imaging*, 27(2):376–390, 2008.

[31] S. Separeto and W. Dewey. Thermal dose determination in cancer therapy. *Int J Radiat Oncol Biol Phys*, 10(6):787–800, 1984.

[32] C. Damianou and K. Hynynen. The effect of various physical parameters on the size and shape of necrosed tissue volume during ultrasound surgery. *J Acoust Soc Am*, 95(3):1641–1649, 1994.

[33] J. Dragonu, B. Denis de Senneville, B. Quesson, C. Moonen, and M. Ries. Real-time geometric distorsion correction for interventional imaging with echo-planar imaging (EPI). *Magnetic Resonance in Medicine*, 61:994–1000, 2009.

[34] A. Muller, L. Petrusca, V. Auboiroux, P. J. Valette, R. Salomir, and F. Cotton. Management of respiratory motion in extracorporeal high-intensity focused ultrasound treatment in upper abdominal organs: current status and perspectives. *Cardiovasc Intervent Radiol*, 36(6):1464 – 1476, 2016.

[35] V.R.S. Mani and S. Arivazhagan. Survey of medical image registration. *Journal of Biomedical Engineering and Technology*, 1(2):8–25, 2013.

[36] D.L. Hill, P.G. Batchelor, M. Holden, and D.J. Hawkes. Medical image registration. *Phys Med Biol.*, 46(3):R1–R45, 2001.

[37] C. Zachiu, N. Papadakis, M. Ries, C. T. W. Moonen, and B. Denis de Senneville. An improved optical flow tracking technique for real-time mr-guided beam therapies in moving organs. *Physics in Medicine and Biology*, 60(23):9003, 2015.

[38] M.P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F.V. Gleeson, S.M. Brady, and J.A. Schnabel. MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical Image Analysis*, 16(7):1423–1435, 2012.

[39] R. L. Ehman and J. P. Felmlee. Adaptive technique for high-definition MR imaging of moving structures. *Radiology*, 173(1):255 – 263, 1989.

[40] M. Poutschi-Amin, S. A. Mirowitz, J. J. Brown, R. C. McKinstry, and Li T. Principles and applications of echo-planar imaging: A review for the general radiologist. *Radiographics*, 21(3):767–779, 2001.

[41] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *J Multiscale Model Simul*, 4(2):490–530, 2005.

[42] Paul A. Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C. Gee, and Guido Gerig. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128, 2006.

[43] J. Wijlemans, M. de Greef, G. Schubert, M. van den Bosch, C. Moonen, and M. Ries. Near-field management during MR-HIFU ablation in highly perfused organs. In *International Society for Magnetic Resonance in Medicine*, page 2340, 2014.

[44] J. Wijlemans. *Towards MR-HIFU ablation of liver tumors.* PhD thesis, Utrecht University, 2014.

[45] C. Zachiu, B. Denis de Senneville, C. Moonen, and M. Ries. A framework for the correction of slow physiological drifts during MR-guided HIFU therapies: Proof of concept. *Med Phys*, 42(7):4137–4148, 2015.

[46] R.B. Case, J.-J. Sonke, D.J. Moseley, J. Kim, K.K. Brock, and L.A. Dawson. Inter- and intrafraction variability in liver position in non-breath-hold stereotactic body radiotherapy. *Int J Radiation Oncology Biol Phys*, 75(1):302–308, 2009.

[47] N. K. Ostergaard, B. Denis de Senneville, U. V. Elstrom, K. Tanderup, and T. S. Sorensen. Acceleration and validation of optical flow based deformable registration for image-guided radiotherapy. *Acta Oncol.*, 47(7):1286 – 1293, 2008.

[48] R. Castillo, E. Castillo, R. Guerra, V. E. Johnson, T. McPhail, A.K. Garg, and T. Guerrero. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Phys. Med. Biol*, 54:1849–1870, 2009.

[49] K.K. Brock. Results of a multi-institution deformable registration accuracy study(MIDRAS). *Int J Radiation Oncology Biol. Phys.*, 76(2):583–596, 2010.

[50] N. Kadoya, Y. Fujita, Y. Katsuta, S. Dobashi, K. Takeda, K. Kishi, M. Kubozono, R. Umezawa, T. Sugawara, H. Matsushita, and K. Jingu. Evaluation of various deformable image registration algorithms for thoracic images. *Journal of Radiation Research*, pages 1–8, 2013.

CHAPTER 6

EVolution: an edge-based variational method for non-rigid multi-modal image registration

"Necessity is the mother of invention."
- Plato

This chapter is based on:
B. Denis de Senneville, C. Zachiu, M. Ries, and C. Moonen. EVolution: an edge-based variational method for non-rigid multi-modal image registration. *Phys Med Biol*, 61(20): 7377 - 7396, 2016.

Abstract

Image registration is part of a large variety of medical applications including diagnosis, monitoring disease progression and/or treatment effectiveness and, more recently, therapy guidance. Such applications usually involve several imaging modalities such as ultrasound, computed tomography, positron emission tomography, X-Ray or magnetic resonance imaging, either separately or combined. In this chapter, we propose a non-rigid multi-modal registration method (namely EVolution: an Edge-based VariatiOnaL method for non-rigid mUlti-modal image registraTION) that aims at maximizing edge alignment between the images being registered. The proposed algorithm requires only contrasts between physiological tissues, preferably present in both image modalities, and assumes deformable/elastic tissues. Given both is shown to be well suitable for non-rigid co-registration across different image types/contrasts (T1/T2) as well as different modalities (CT/MRI). This is achieved using a variational scheme that provides a fast algorithm with a low number of control parameters. Results obtained on an annotated CT data set were comparable to the ones provided by state-of-the-art multi-modal image registration algorithms, for all tested experimental conditions (image pre-filtering, image intensity variation, noise perturbation). Moreover, we demonstrate that, compared to existing approaches, our method possesses increased robustness to transient structures (i.e that are only present in some of the images).

6.1 Introduction

Image registration is the process of aligning two or more images of the same scene acquired at different time instants, using different sensors and/or from a different point-of-view. Although initially proposed in the field of digital photography and video sequence processing, as medical imaging technologies progressed, it has become an increasingly important pre-processing step for medical image analysis. Currently, image registration is part of a large variety of medical applications including diagnosis, monitoring disease progression and/or treatment effectiveness and, more recently, therapy guidance.[1] Such applications usually involve several imaging modalities such as ultrasound (US), computed tomography (CT), positron emission tomography (PET), X-Ray or magnetic resonance imaging (MRI), either separately or combined. Whether one or more imaging modalities are employed, image registration algorithms can be divided in two categories: mono-modal and respectively multi-modal. A variety of each has already been proposed in the medical image processing literature.[2] Mono-modal registration algorithms typically rely on the assumption that an anatom-

ical structure is present in all images included in the registration process and also that the structures preserve, to a certain extent, their gray-level intensities. Such algorithms are well suited, for example, for tumor growth monitoring and intervention verification.[3] Multi-modal algorithms relax the gray-level intensity conservation constraint. Moreover, a certain variety do not even require a structure to be present in all images. Cross-modality registration algorithms are more suitable for diagnostic purposes, since different imaging modalities usually provide complementary information. For example, the excellent anatomical information provided by an MRI image could be complemented by the physiological information present in a PET scan, each modality having, in general, more vast capabilities than the other in terms of the specified provided information.[1] Compared to mono-modal algorithms, multi-modal methods are usually more complex and require additional computational resources.

A particular variety of registration algorithms are the so-called variational methods which aim at minimizing a cost function in order to find the transformation between the registered images.[4] Since rigid/affine transformations have a limited (although important) purpose in medical image registration, the focus in this chapter will be on deformable transformations. While different in nature, all variational methods usually employ cost functions that include two components: a data fidelity term and respectively a regularization term. The data fidelity term measures the similarity between the images undergoing the registration process, while the regularization term imposes constraints on the estimated deformation, since the minimization of the data fidelity term alone is usually an ill-posed problem. Variational methods are a particularly attractive solution for medical image registration since they require a low number of input/control parameters and typically involve fast numerical schemes.[5,6] Whether they are able to cope with images acquired through different modalities depends mostly on the similarity measure used by the data fidelity term. While for mono-modal registration algorithms the sum of squared differences (SSD) applied directly on the images might suffice,[7,8] such a measure is highly unsuitable for registering across modalities. Modality independent similarity measures are thus necessary.

Previous studies have addressed the cross-modality registration problem by pre-processing the input images into modality-independent scalar representations of themselves. Once pre-processed, the SSD was used as a similarity metric between the images under the new representation. Examples of modality independent scalar representations include the local phase, local entropy or gradient orientation.[9–11] While indeed showing a good performance in both registration quality and computational demands, these methods are hampered by the fact that for more challenging multi-modal registration tasks, the features present in the images under the new representation might not be discriminative enough.

Another broad category of multi-modal registration algorithms makes use of the mutual information (MI) (directly or variations) as an image similarity measure. Initially proposed independently by Viola and Wells[12] and respectively by Maes et *al*,[13] mutual information-based multi-modal registration al-

gorithms were met with great interest by the medical image registration community, triggering a long series of studies showcasing the great potential of such approaches.[14] However, MI is inherently a global measure, making its local estimation a rather challenging task. For this reason, while having excellent performance for multi-modal rigid/affine realignment tasks, it becomes more difficult to apply when estimating local deformations. In addition, optimizing the mutual information between two or more images is computationally more complex and slow compared to an SSD-based similarity measure.

Recently, a novel cross-modality registration algorithm has been proposed by Heinrich et *al* based on so-called modality independent neighborhood descriptors (MIND).[15] The method transforms the input images into a vectorial representation such that a descriptor is associated to each voxel, reflecting its relationship to the surrounding voxels. The descriptors are constructed based on the notion of self-similarity as initially proposed by Buades et *al* in the design of their non-local filtering method.[16] This results in a registration algorithm that is sensitive to structural information rather than intensity. For the test cases considered in the original paper, the MIND algorithm has shown increased performance over several other popular multi-modal registration methods. Moreover, since the similarity criterion uses the SSD to measure the distance between the descriptor images, it allows for fast optimization numerical schemes to be employed. However, the method is prone to the same limitations as the non-local means filter:[16] it is sensitive to several parameters such as the local estimate of the noise variance, the patch size around the central voxel used to compute each element of the descriptors, the number of voxels included in the computation of the descriptors, and it is affected by low contrasts. In addition, if the structural information is different from one image to the other, for example, if an area that is smooth in one image and textured in the other, then the descriptors cannot be compared reliably. More precisely, the method hypothesizes that each anatomical structure in one image has its counterpart in the other. This may not be the case in several situations, including transient structures occurring in the field-of-view as a result of peristaltic activity (in case the analysis is conducted in the abdomen), the appearance/disappearance of pathologies within the analyzed organ and/or parts of the region-of-interest or its surroundings entering/exiting the field-of-view.

A different approach to cross-contrast registration are algorithms that maximize the edge alignment between the registered images. Various metrics that quantify the extent to which the edges are aligned have already been proposed in the literature. The methods were dedicated to registering either video sequences acquired with different sensors or medical images acquired through different modalities.[17–19] In this approach, the underlying assumption that needs to be fulfilled is that the boundaries and details of physiologic structures show image contrast to the surrounding tissue with both image modalities. However, the studies have only gone so far as to estimate rigid/affine or coarsely deformable transformations (using splines). In the current chapter, by using the principle behind the edge alignment methods, we propose EVolution: an

Edge-based Variational algorithm for multi-modal image registration. Our contribution is four-fold:

- By construction, the algorithm is designed to increase the robustness of the registration process against structural information variations from one image to the other.
- A patch-based approach is designed to leverage limitations arising from the above mentioned scalar representation, especially for highly challenging multi-modal scans.
- Since a variational approach is employed, the method requires a reduced number of input parameters that need to be calibrated. Moreover, the cost function we propose also renders itself compatible with fast numerical schemes, while providing a dense voxel-by-voxel deformation field.
- The benefit of using multi-CPU and GPU (graphics processing unit) architectures is evaluated.

6.2 Method description

6.2.1 Proposed EVolution method

The equations provided in the current chapter refer to the 3D implementation of the algorithm. An image J is registered to the reference position given by I using a variational image registration method as follows.

6.2.1.1 Proposed data fidelity term

Let $\vec{\nabla} I$ and $\vec{\nabla} J$ be the gradient of the reference image I and the image to register J, respectively. We defined the following patch-based criterion (a patch consists in a cubic subset of the image domain, denoted by Γ, centered on one single voxel):

$$C(T) = \frac{\int\limits_{\Gamma} \left| \vec{\nabla} I\left(T(\vec{r})\right) \cdot \vec{\nabla} J\left(\vec{r}\right) \right| \mathrm{d}\vec{r}}{\int\limits_{\Gamma} \| \vec{\nabla} I\left(T(\vec{r})\right) \|_2 \| \vec{\nabla} J\left(\vec{r}\right) \|_2 \mathrm{d}\vec{r}} \tag{6.1}$$

with $T = (u, v, w)$ the spatial transformation from I to J, u, v and w the displacement vector components, and $\vec{r}$ the spatial location.

The expression $C(T)$ can also be rewritten under the following form:

$$C(T) = \frac{\int\limits_{\Gamma} w_T(\vec{r}) \left| \cos\left(\Delta\theta_T(\vec{r})\right) \right| \mathrm{d}\vec{r}}{\int\limits_{\Gamma} w_T(\vec{r}) \mathrm{d}\vec{r}} \tag{6.2}$$

where $w_T(\vec{r})$ and $\Delta\theta_T(\vec{r})$ are calculated from the magnitude M and the orientation θ of the image gradient at location $\vec{r}$ as follows:

$$\begin{aligned} w_T(\vec{r}) &= M_I\left(T(\vec{r})\right) M_J(\vec{r}) \\ \Delta\theta_T(\vec{r}) &= \theta_I\left(T(\vec{r})\right) - \theta_J(\vec{r}) \end{aligned} \tag{6.3}$$

Intuitively, the term $|\cos\left(\Delta\theta_T(\vec{r})\right)|$ in Eq. (6.2) favors the transformations that align the edges, regardless any possible contrast reversals: due to the absolute value of the cosine, both parallel and anti-parallel edges are considered to coincide irrespective of the gradient direction. In addition, when dealing with multi-modal images, some discontinuities may only appear in one of the two modalities, so the weight $w_T(\vec{r})$ favors strong edges that occur in both modalities. The denominator of Eq. (6.2) performs a weighted average of the score obtained for each edge.

Since, for image registration using variational methods, a minimization of the functional is mandatory, we defined the following patch-based similarity criterion $D(T)$:

$$D(T) = e^{-C(T)} \tag{6.4}$$

$D(T)$, which is computed individually at each spatial location $\vec{r}$, is a strictly positive number, which decreases as long as the alignment of I and J is improved within the local neighborhood Γ. In this manner, $D(T)$ can be employed as a data fidelity term for the proposed variational registration method.

6.2.1.2 Optimized variational functional

We propose minimizing the energy E given by:

$$E(T) = \int_{\Omega} D(T) + \frac{\alpha}{2}\left(\| \vec{\nabla} u \|_2^2 + \| \vec{\nabla} v \|_2^2 + \| \vec{\nabla} w \|_2^2\right) \mathrm{d}\vec{r} \tag{6.5}$$

where Ω is the image coordinates domain, α a weighting factors designed to link both the data fidelity term $D(T)$ and the motion field regularity (right part of Eq. (6.5)).

At this point, it is important to underline that two user-defined parameters may impact the performance of the registration process:

- The parameter α which infers the regularity of the estimated motion field.
- The patch size p (noted p) which may infer the robustness against different structural information from the image to register to the reference (for example, on an area that is smooth on one image and textured on the other). Simultaneously, p also infers the regularity of the estimated motion field.

Throughout the rest of the chapter, a special attention will be paid to the impact of these two parameters on the overall registration results.

6.2.1.3 Implemented optimization scheme

By applying the Euler-Lagrange equations on a voxel-by-voxel basis, one can derive the following system of equations for each $\vec{r} \in \Omega$:

$$\begin{cases} \frac{\partial D}{\partial u}(T) - \alpha \Delta u = 0 \\ \frac{\partial D}{\partial v}(T) - \alpha \Delta v = 0 \\ \frac{\partial D}{\partial w}(T) - \alpha \Delta w = 0 \end{cases} \tag{6.6}$$

where Δ denotes the Laplacian operator.

From here, we have a set of $3 \times |\Omega|$ non-linear equations with common unknowns u, v and w. The latter can be found iteratively through the following explicit fixed-point scheme:

$$\begin{cases} u_{k+1} = u_k + \Delta u_k - \alpha^{-1} D(T_k) \frac{\partial D}{\partial u}(T_k) \\ v_{k+1} = v_k + \Delta v_k - \alpha^{-1} D(T_k) \frac{\partial D}{\partial v}(T_k) \\ w_{k+1} = w_k + \Delta w_k - \alpha^{-1} D(T_k) \frac{\partial D}{\partial w}(T_k) \end{cases} \tag{6.7}$$

where $k+1$ denotes the new iteration and $T_k = (u_k, v_k, w_k)$. Each operator $\Delta(.)$ was numerically computed using a $3 \times 3 \times 3$ Laplacian kernel.[20] For each voxel, the partial derivatives of the data fidelity term along each displacement component were computed using the following finite difference schemes:

$$\begin{cases} \frac{\partial D}{\partial u}(T_k) = \frac{D(T_k \circ S_x^+) - D(T_k \circ S_x^-)}{2} \\ \frac{\partial D}{\partial v}(T_k) = \frac{D(T_k \circ S_y^+) - D(T_k \circ S_y^-)}{2} \\ \frac{\partial D}{\partial w}(T_k) = \frac{D(T_k \circ S_z^+) - D(T_k \circ S_z^-)}{2} \end{cases} \tag{6.8}$$

where S_x^+, S_y^+ and S_z^+ denote the forward shifted image transformations by one voxel along x-, y- and z- directions, respectively. S_x^-, S_y^- and S_z^- being the corresponding backward shifted image transformations. For example, $T_k \circ S_x^+$ reads as: "a translation of $(x, y, z) \rightarrow (x+1, y, z)$ followed by the spatial transformation T_k".

It was considered that the numerical scheme in Eq. (6.7) converged when the average variation of the motion magnitude from one iteration to the next was smaller than 10^{-3} voxels.

6.2.1.4 Coarse-to-fine scheme

A coarse-to-fine strategy was carried out, which iterated the registration algorithm from a downsampled image (4-, 8- and 16-fold downsampled images were employed for nominal original image dimensions of 128, 256 and 512 voxels, respectively), step by step to the original image resolution. To achieve a global motion regularization, the estimate arising from lowest resolution levels (noted $T_g = (u_k^g, v_k^g, w_k^g)$) was added in the regularization term (i.e in the right part of Eq. (6.5)). The system of Eq. (6.7) was thus simply rewritten as follows:

$$\left\{\begin{array}{rcll} u_{k+1} = & u_k & +\Delta u_k + \Delta u_k^g & -\alpha^{-1} D\left(T_k\right) \frac{\partial D}{\partial u}\left(T_k\right) \\ v_{k+1} = & v_k & +\Delta v_k + \Delta v_k^g & -\alpha^{-1} D\left(T_k\right) \frac{\partial D}{\partial v}\left(T_k\right) \\ w_{k+1} = & w_k & +\Delta w_k + \Delta w_k^g & -\alpha^{-1} D\left(T_k\right) \frac{\partial D}{\partial w}\left(T_k\right) \end{array}\right. \tag{6.9}$$

6.2.2 Experimental setup

The proposed EVolution method was evaluated on three data sets:

1. one set of thorax inhale and exhale CT-scans, for which gold-standard landmark displacements are available,
2. one set of MR-scans capturing the expansion of the bladder in a healthy volunteer, for which a silver-standard motion is constructed,
3. one set of paired head CT/MR-scans, for which gold-standard spatial transformations are available.

6.2.2.1 Thorax inhale and exhale CT-scans

We used a set of ten CT-scan pairs, freely provided by the DIR-Lab[1] at the University of Texas.[21–23] Each scan pair was acquired on the thorax and upper abdomen of patients treated for esophagian cancer, between inhale and exhale phase of the breathing cycle. For each CT-scan, 300 anatomical landmarks have been annotated by experts (inter-observer errors below 1 mm).

Four reliability tests were conducted in order to evaluate the registration performance under various experimental conditions:

a Reliability test #1 Registration was directly applied between each pair of the original CT-scans. Major challenges arise from possible contrast variations between tissue and air induced by lung compression, motion discontinuities at the lung/rib cage interface, as well as large deformations of small features such as lung vessels, airways.[22]

b Reliability test #2 Registration was applied on the original data set, each image being previously filtered using Gaussian filter (kernel $3 \times 3 \times 3$, $\sigma = 0.5$). In this manner, the impact of filtered anatomical structures on the overall registration accuracy is assessed.

c Reliability test #3 A prior intensity perturbation was applied between I and J in order to challenge the registration against intensity contrast variation. For this purpose, the image to register J was replaced by its negative contrast in the registration process, the reference image remaining identical to the original one.

[1] http://www.dir-lab.com

d Reliability test #4 The robustness against noise was investigated by adding a Gaussian white noise ($\sigma = 10\%$ of the maximum image intensity) on both I and J, prior to the registration process.

6.2.2.2 Bladder filling MR-scans

A series of 3D MR-images were acquired on a healthy volunteer, capturing the expansion of the urinary bladder over time. In order to sample the shape and size of the bladder from void to partially full in a relatively short amount of time, a volunteer preparation protocol was devised. The volunteer was instructed not to consume any liquids 6 hours prior to the MR-acquisition session such that, at the time of the study, they would be mildly dehydrated. Before the start of the experiment, the volunteer was allowed to drink a self-chosen quantity of water (∼350 ml). It was speculated that the mild dehydration would stimulate renal activity, leading to a rapid expansion of the bladder. Approximately every 7 minutes, a pair of 3D T1-T2-weighted images was acquired on the lower abdomen of the volunteer, over a total duration of ∼40 minutes. The MR-acquisition sequences were optimized to mitigate the fact that each anatomical structure in the T1-weighted images has its counterpart in the T2-weighted ones. The T1-weighted sequence employed the following parameters: TE = 2ms, TR = 4.3ms, image matrix $192 \times 192 \times 100$, 10° flip angle, with an isotropic voxel size of $2 \times 2 \times 2\text{mm}^3$. For the T2-weighted acquisition, the employed parameters were as follows: TE = 130ms, TR = 1000ms, reconstructed image size $512 \times 512 \times 133$, 90° flip angle, with a voxel size of $0.5 \times 0.5 \times 1.5\text{mm}^3$. Out of the entire data set, two pairs of T1-T2-weighted images (∼30 min apart) were selected for the role of reference and respectively image to register. A silver-standard motion field was estimated using a common contrast for both the reference and the image to register (the T1-weighted images were employed) using the following mono-modal optical flow (OF) functional:

$$E_{OF}(T) = \int_{\Omega} |I_x u + I_y v + I_z w + I_t| + \alpha \left(\| \vec{\nabla} u \|_2^2 + \| \vec{\nabla} v \|_2^2 + \| \vec{\nabla} w \|_2^2 \right) \mathrm{d}\vec{r} \tag{6.10}$$

where $I_{x,y,z,t}$ are the spatio-temporal partial derivatives of the image voxel intensity. A complete description of the associated non-rigid registration framework can be found in.[24] The obtained OF-motion field was taken as a reference for the evaluation of the proposed EVolution method applied on cross-contrast images (the T1- and respectively the T2-weighted images were employed for the reference and the image to register).

6.2.2.3 Paired head CT/MR-scans

We used a set of two CT/MR-scan pairs, freely provided by the Retrospective Image Registration Evaluation Project (RIRE) [2] at the National Institutes of Health, Vanderbilt University.[25,26] For this purpose, three 3D images of the head of a common volunteer were employed: a CT-scan (voxel size of $0.65 \times 0.65 \times 4\text{mm}^3$, image matrix $512 \times 512 \times 29$) and a pair of T1-T2-weighted MR images (voxel size of $1.25 \times 1.25 \times 4\text{mm}^3$, image matrix $256 \times 256 \times 26$). The performance of the proposed EVolution method for CT/T1-MRI and CT/T2-MRI registrations was assessed using gold standard spatial transforms defined using a prospective, marker-based technique, as described in.[27]

6.2.2.4 Implementation of the MIND non-rigid registration algorithm

The identical experimental setup was also applied on the MIND non-rigid registration method.[15] We recall that the latter is a modality invariant local image descriptor based on the notion of self-similarity derived from the NL-means algorithm. A descriptor, based on the local similarities, is associated to each voxel $\vec{r}$ of the image I :

$$\text{MIND}(I, \vec{r}, \gamma) = \frac{1}{n} \exp\left(-\frac{D_p(I, \vec{r}, \vec{r}+\gamma)}{V(I, \vec{r})}\right) \tag{6.11}$$

where n is a normalization constant, $\gamma \in \Gamma$ defines a search region in which the patches are compared, D_p is the voxel-wise square distance between patches of size p, and $V(I, \vec{r})$ is an estimation of the local variance in order to account for noise perturbations. The MIND descriptor associates a vector of size $|\Gamma|$ to each pixel of the image I. The descriptors of each modality can subsequently be compared in order to provide the following voxel-wise data fidelity metric:

$$S(T) = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} |\text{MIND}(I, \vec{r}, \gamma) - \text{MIND}(J, \vec{r}, \gamma)| \tag{6.12}$$

This data-fidelity term is then combined to a L^2 regularization term, similarly to Eq. (6.5). The implemented non-rigid registration framework was identical to the one described in.[15] However, in order to make a proper comparison with the estimates provided by our proposed EVolution algorithm, we employed the same coarse-to-fine scheme and convergence criterion as described in section 6.2.1.4. Note that, for the purpose of this study, $\Gamma = [-1;1]^3 \cap \mathbb{Z}^3$ and $p = 3 \times 3 \times 3$ were found to be a good compromise between algorithm computational performance and quality of the motion estimates (the interested reader is referred to[15] for a complete analysis of the two parameters Γ and p).

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

6.2.2.5 Assessment of the motion estimation process

Concerning the data set of thorax CT-scans, the performance were assessed using the target registration error (TRE) of anatomical landmarks as follows:[28]

$$\mathrm{TRE} = \sqrt{\left(x_I + u(\vec{r}_I) - x_J\right)^2 + \left(y_I + v(\vec{r}_I) - y_J\right)^2 + \left(z_I + w(\vec{r}_I) - z_J\right)^2} \tag{6.13}$$

where $\vec{r}_I = (x_I, y_I, z_I)$ and $\vec{r}_J = (x_J, y_J, z_J)$ denote the anatomical landmark coordinates in I and J, respectively.

Concerning the data set of bladder MR-scans and the paired head CT/MR-scans, the quality of the estimated motion was assessed using the voxel-wise error in flow endpoint (FEP) computed as follows:[29]

$$\mathrm{FEP} = \sqrt{\left(u - u_{ref}\right)^2 + \left(v - v_{ref}\right)^2 + \left(w - w_{ref}\right)^2} \tag{6.14}$$

$(u_{ref}, v_{ref}, w_{ref})$ being the reference OF-motion estimate.

6.2.3 Hardware and implementation

Our test platform was an Intel 2.5 GHz i7 workstation (8 cores) with 32 GB of RAM. The GPU was a Quadro K2100M card with 2 GB of dynamic random-access memory (NVIDIA, Santa Clara, CA, USA). Both CPU and GPU implementations of the proposed EVolution algorithm were realized and tested. The CPU implementation was performed in C++ and parallelized through multi-threading. The GPU implementation was realized using the compute unified device architecture (CUDA).[30]

6.3 Results

6.3.1 Thorax inhale and exhale CT-scans

A visualization of a registration result can be seen in Fig. 6.1 (empirically established values of $\alpha = 0.5$ and $p = 11$ were employed). In this visualization, the source image is shown in magenta while the reference image is shown in green. Where the images align a gray scale image emerges. In the unregistered case on the top, magenta and green areas can clearly be seen indicating that the morphology is not aligned. In the registered case to the bottom, these colored areas have almost disappeared indicating that the images have been successfully registered.

Fig. 6.2 shows cumulative distributions of the target registration error (TRE) over 3000 landmarks (300 landmarks for 10 CT-scans) obtained before and after registration, during the "Reliability test #2" (common default values of $\alpha = 0.5$ and $p = 11$ were again employed for the EVolution method, and $\alpha = 0.1$ was

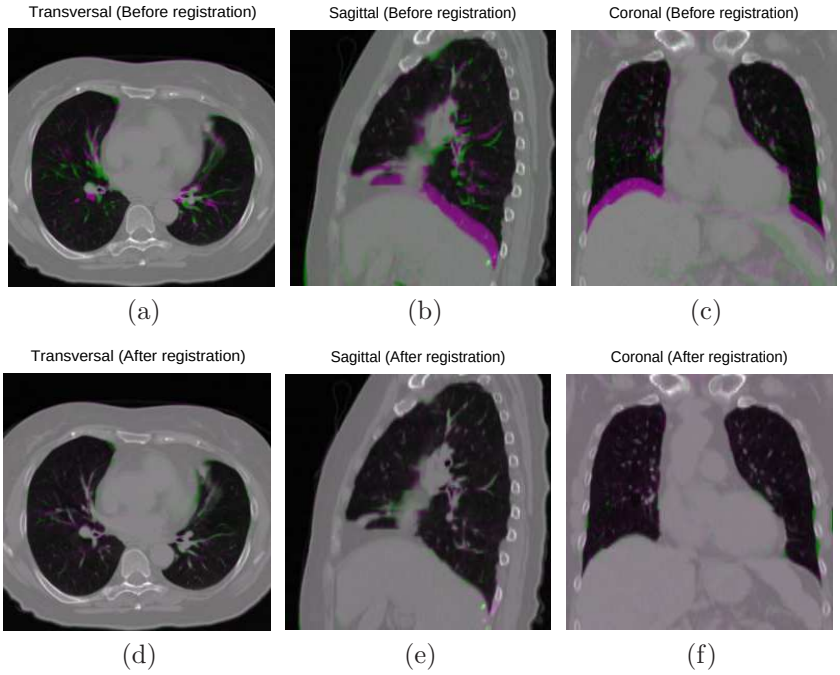


Figure 6.1: Example of the registration of a thorax CT-scan (CT-scan #5) using the proposed EVolution method. Top: before registration and bottom: after registration using the proposed EVolution method. The target image is displayed in green and the source image in magenta (complementary colour).

used for the MIND algorithm, as advised by[15]). Staircasing effects are observable on the curve obtained before registration, due to the voxel precision of the expert landmark positioning. However this was not the case for the curves associated to the MIND and the proposed EVolution methods, due to the real number precision of the motion estimates. Only a marginal difference can be observed between the MIND and the proposed EVolution methods for all tested experimental conditions. This tendency is confirmed for all tested experimental conditions, as shown in Table 6.1. For both algorithms, no negative impact was observed when the image to register was replaced by its corresponding negative contrast image ("Reliability test #3"). It is also noticeable that both algorithms performed better on pre-filtered images ("Reliability test #4"). For comparison purposes, Table 6.2 provides several findings reported in[15] using various existing registration methods and the identical set of CT-scans (the interested reader is also referred to[15] for a complete description of the associated non-rigid registration framework employed to generate the results). It can be observed that the TRE obtained using our implementation of the MIND

algorithm are on average almost identical to the ones reported in the original paper (2.18 mm for our implementation against 2.14 mm for the original). The standard deviation of the TRE was even marginally improved using our implementation (3.45 mm against 3.71 mm). We believe that this is due to the fact that our convergence criterion, which had to be similar to the one employed for the proposed EVolution method (see section 6.2.1.3), is more severe in our implementation.

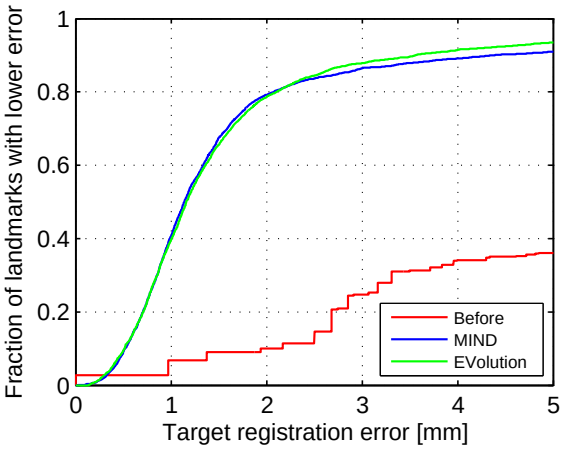


Figure 6.2: Registration of 10 cases of CT-scans. The plot shows the cumulative distribution of target registration error.

Fig. 6.3 analyzes the sensitivity to parameter α of the MIND algorithm and respectively the proposed EVolution method. It is important to indicate that, for $\alpha < 0.1$, high instabilities of the numerical schemes were observed for both the MIND and the EVolution approaches. It can be observed that the performance of the MIND algorithm deteriorates as α increases. $\alpha = 0.1$ was thus found the optimal calibration for the MIND algorithm, which matches the findings reported in.[15] $\alpha = 0.1$ was thus employed as a default value throughout the chapter for the MIND algorithm. Using the EVolution method, the TRE, as a function of α, exhibits a flat zone: a range of values for α allowing an accurate registration could thus be determined. This justifies our choice to set a value of 0.5 as a default parameter for α for the proposed EVolution algorithm in the scope of this study.

Fig. 6.4 analyzes the impact of the patch size on the performance of the proposed approach. The TRE was calculated over the 10 cases of the CT-scan data set (300 expert landmarks per case). The mean and standard deviation of the TRE can thus be directly compared to the values reported in Tables 6.1 and 6.2. It can be observed that the performance of the EVolution method could be further improved for all tested experiment conditions by simply reducing

Reliability test	Target registration error [mm]	
	MIND	EVolution
Original data	2.33 ± 3.54	2.07 ± 3.23
(*Reliability test* #1)	[0.81, 1.18, 1.91]	[0.78, 1.17, 1.85]
Filtered data	2.18 ± 3.45	1.96 ± 2.99
(*Reliability test* #2)	[0.78, 1.14, 1.77]	[0.78, 1.16, 1.80]
Cross-contrast	2.33 ± 3.55	2.08 ± 3.24
(*Reliability test* #3)	[0.81, 1.18, 1.91]	[0.78, 1.18, 1.85]
Noisy data	2.33 ± 3.55	2.01 ± 3.14
(*Reliability test* #4)	[0.81, 1.18, 1.91]	[0.77, 1.16, 1.82]

Table 6.1: Target registration error (in millimeters) obtained over the 10 cases of thorax CT-scans for all tested experimental conditions (Reliability tests #1−4). The first line of each cell reports the mean ± standard deviation of the TRE, and the second line provides the first, second and third quantiles ([0.25, 0.5, 0.75]).

Registration method	Target registration error [mm]
Before	8.46 ± 6.58
	[3.11, 6.97, 12.55]
Sum of squared differences[14]	2.73 ± 3.72
	[0.89, 1.44, 2.85]
SSD of entropy images[14]	2.86 ± 4.91
	[0.86, 1.33, 2.33]
Normalised mutual information[14]	2.97 ± 4.22
	[0.91, 1.42, 2.67]
Conditional mutual information[31]	3.06 ± 4.10
	[1.00, 1.59, 2.85]
MIND[15]	2.14 ± 3.71
	[0.77, 1.16, 1.79]

Table 6.2: Summary of several findings reported in the literature using the employed set of 10 CT-scans.

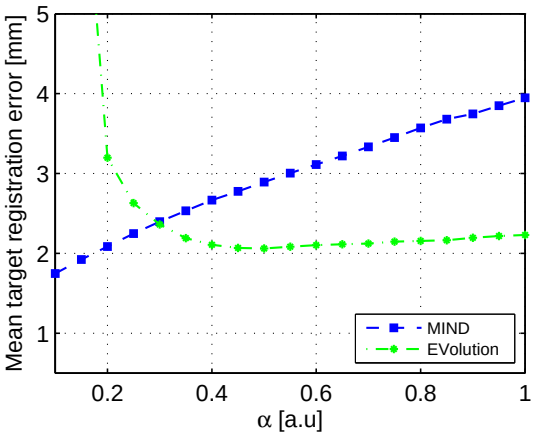
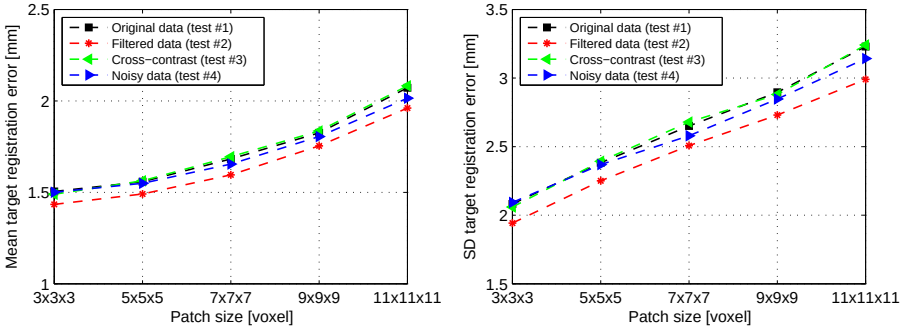


Figure 6.3: Example (CT-scan #5) of the sensitivity to the α parameter of the MIND and the EVolution algorithms. We recall that the size of the patch for the EVolution method was fixed to cover a dimension of $11 \times 11 \times 11$ voxels.

the patch size. It is also important to indicate that a patch restrained to the dimension of one voxel resulted in high instabilities in the employed numerical scheme.



(a) (b)

Figure 6.4: Analysis of the impact of the patch size on the performance of the EVolution method. We recall that the regularization parameter α was here fixed to a value of 0.5.

6.3.2 Bladder filling MR-scans

Fig. 6.5 illustrates the first step of the experimental protocol designed to assess the proposed EVolution algorithm on the bladder filling MR-data set: a sliver-standard motion is constructed using an optical flow (OF) algorithm (Eq. (6.10)) applied on the T1-weighted data for both the reference and the image to register. The inserts in Fig. 6.5d-f show a spatially regular motion field. Moreover, a good correspondence has been found between the bladder on the registered volumetric image of Fig. 6.5g-i and the region of interest manually defined on the reference data.

A visualization of cross-contrast registration results can be seen in Fig. 6.6. Default calibration parameters were employed for both the MIND and the proposed EVolution methods.

This visualization shows a close correspondence of the contour of the co-registered bladder with the manually defined bladder boundary when the proposed algorithm is employed (see Fig. 6.6a-c). The averaged FEP in the bladder decreased from 6.33 mm before registration to 3.12 mm and 4.34 mm using the proposed EVolution and the MIND algorithm, respectively (see Fig. 6.6d-f).

Fig. 6.7 shows the impact of the parameter α on the registration within the bladder (blue curve) as well as the complete image field-of-view (green curve), using the EVolution method (6.7a) and the MIND algorithm (6.7b). It can be observed that the behavior of both algorithms in the bladder was similar than the one reported in Fig. 6.3. Using the MIND algorithm, the default value of $\alpha = 0.1$ optimized locally the results in the bladder but also provided the worst global performance within the complete image field-of-view. Moreover, while high α values improved the accuracy on the complete image field-of-view (wherein the high motion smoothness then coped nicely with the small global motion of the abdomen), this also had an opposite impact in the bladder (wherein the mean FRE necessarily converged toward the one obtained before registration). On the other hand, using the proposed EVolution method, a range of values could be found which optimized results in both the complete field-of-view as well as locally in the bladder (see arrows (1) and (2) in Fig. 6.6b and corresponding locations in the error map of Fig. 6.6e).

Fig. 6.8 shows the impact of the patch size on the performance of the proposed EVolution method. The behavior of these two curves match closely the behavior of the green and the blue curves displayed in Fig. 6.7a. It is important to mention that a patch restrained to the dimension of one voxel resulted in high instabilities in the employed numerical scheme.

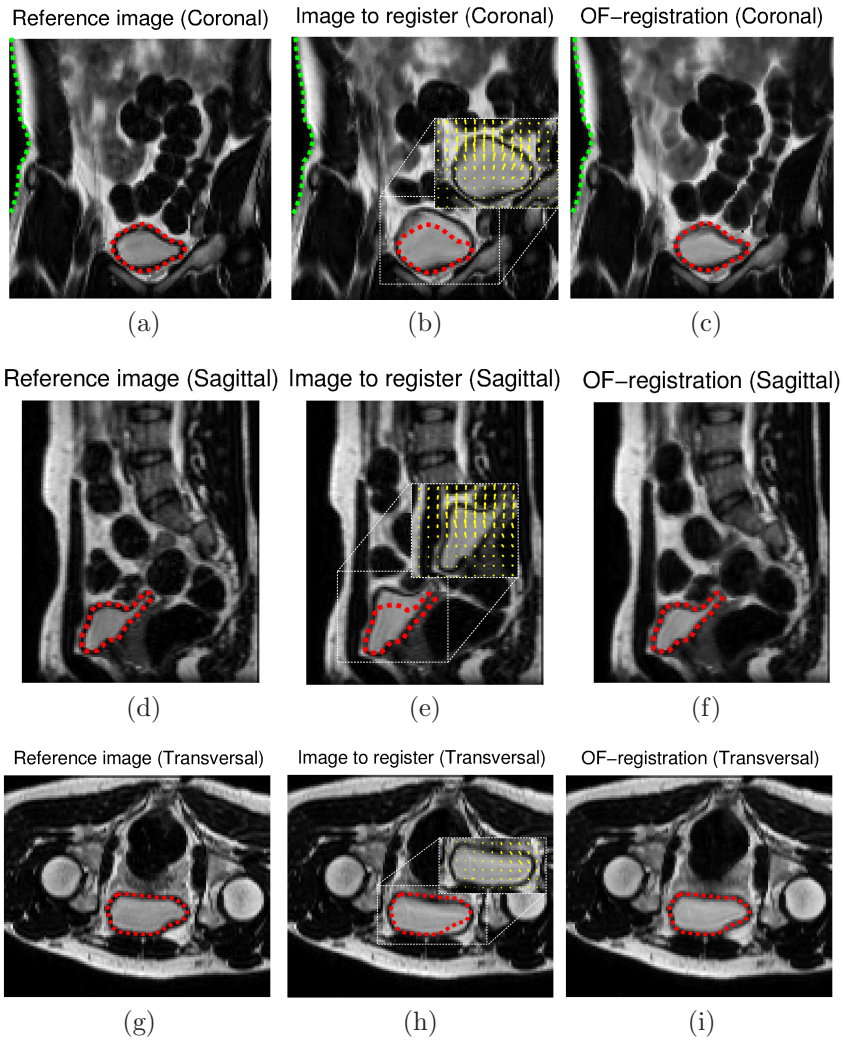


Figure 6.5: Construction of a silver-standard motion for the cross-contrast MR-data set: an OF-based algorithm was applied on the T1-weighted data for both the reference and the image to register. Manually defined ROI encompassing the bladder and the thorax are illustrated using red and green dashed lines in order to improve the visual inspection of the quality of the OF registration. The inserts in (b), (f) and (h) depict the estimated OF-motion field vectors within the bladder.

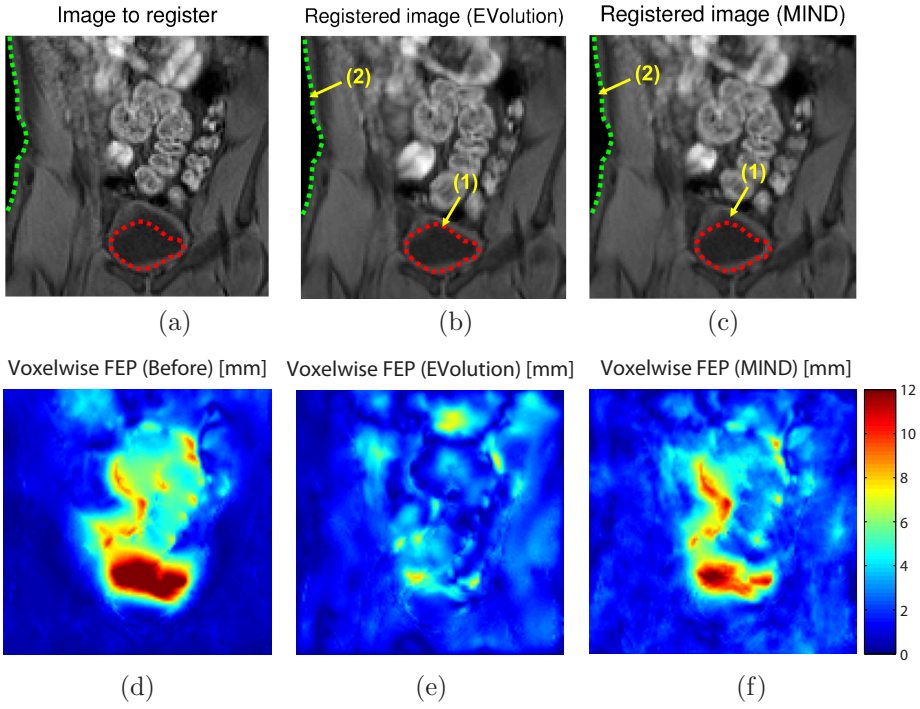


Figure 6.6: Cross-contrast registration of the T2-weighted bladder image (the reference image being the T1-weighted image displayed in Fig. 6.5a-c. The coronal image is displayed before registration (left column), using the EVolution method (middle column, $\alpha = 0.5$, patch size=$11 \times 11 \times 11$ voxels) and using the MIND algorithm (right column, $\alpha = 0.1$). (a-c) display the anatomical images. Manually defined ROI encompassing the bladder (arrow (1)) and the thorax outline (arrow (2)) are shown using red and green dashed lines in order to improve the visual inspection of the quality of the cross-contrast registration. (d-f) display the corresponding voxel-wise FEP maps.

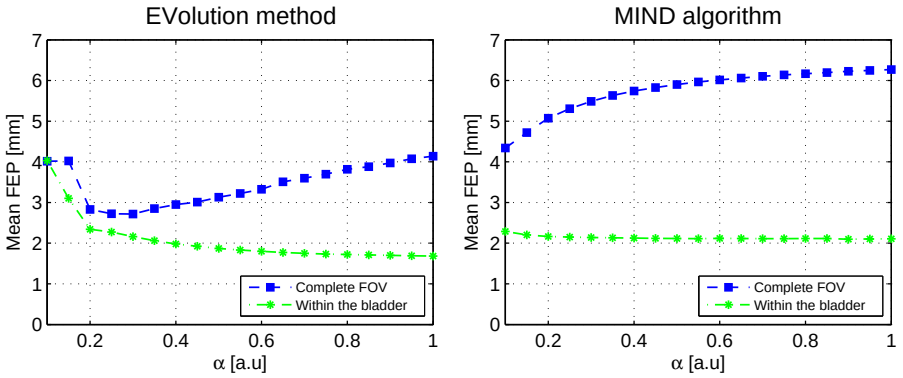


Figure 6.7: Impact of the α parameter on the performance of the proposed EVolution (a) and the MIND (b) algorithms for the bladder filling MR-scans.

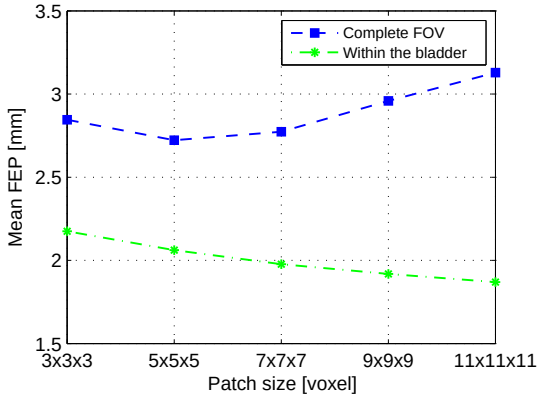


Figure 6.8: Impact of the patch size on the performance of the proposed EVolution method for the bladder filling MR-scans.

6.3.2.1 Paired head CT/MR-scans

A visualization of a CT/MRI registration result can be seen in Fig. 6.9 (default values of $\alpha = 0.5$ and $p = 11$ were used). The employed visualization is identical to that described in section 6.3.1. In the registered case shown on the right, it can be observed that the skull, which features a hyper-intense signal in the CT-scan (magenta), matches closely the spacial location of the scalp and the brain, the latter being visible only in the T1-weighted image (green). This indicates that the images have been successfully registered. This visual

Registered image modalities (Modality #1 / Modality #2)	Motion estimation error [mm]	
	Before registration	EVolution
CT-scan / T1-weighted MRI	41.24 ± 2.66 [39.27, 41.27, 43.27]	3.30 ± 1.84 [1.93, 2.98, 4.24]
CT-scan / T2-weighted MRI	49.32 ± 2.54 [47.34, 49.25, 51.24]	3.87 ± 2.24 [2.04, 3.42, 5.36]

Table 6.3: Motion estimation error (in millimeters) obtained for the registration of the two paired head CT/MR-scans (i.e CT/T1-MRI and CT/T2-MRI). The first line of each cell reports the mean ± standard deviation of the FEP, and the second line provides the first, second and third quantiles ([0.25, 0.5, 0.75]).

Patch size [voxels]	Thorax CT-scan #5 [$256 \times 256 \times 106$]		Bladder MR-scan [$128 \times 128 \times 128$]	
	CPU	GPU	CPU	GPU
$3 \times 3 \times 3$	106	68	34	19
$5 \times 5 \times 5$	195	98	56	22
$7 \times 7 \times 7$	331	134	114	29
$9 \times 9 \times 9$	693	219	179	41
$11 \times 11 \times 11$	1094	320	300	59

Table 6.4: Computation time in seconds obtained using our implementation of the proposed EVolution method on our test platform (multi-CPU and GPU implementations) for different patch size (see section 6.2.3).

inspection is confirmed by the Table 6.3, which analyzes the voxel-wise error in the flow endpoint for the two CT/MRI paired scans. For each CT/T1 and CT/T2 scenario, an average FEP below the voxel-size was obtained, indicating a sub-voxel registration accuracy.

6.3.2.2 Benchmark

A benchmark of our CPU and GPU implementations of the proposed EVolution method for different patch size is provided in Table 6.4. As compared to the CPU implementation, a great acceleration could be obtained using the GPU, especially for high patch size. It is interesting to note that less than a minute was necessary to register two volumes of $128 \times 128 \times 128$ voxels (i.e the Bladder MR-scan).

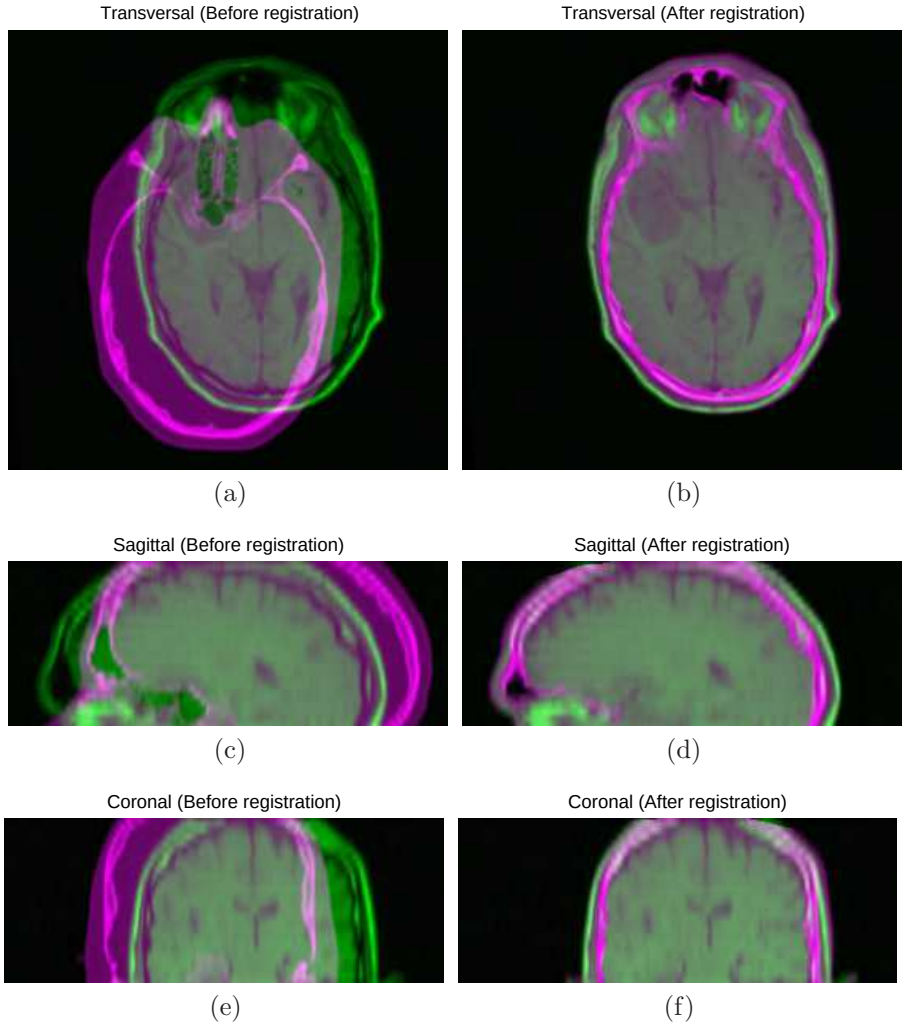


Figure 6.9: Example of CT/T1-MRI registration of the brain using the proposed EVolution method. Left: before registration and right: after registration using the proposed EVolution method. The CT-image is displayed in magenta and the T1-weighted MRI in green (complementary color).

6.4 Discussion

Using variational image-based registration methods, the negative impact of appearing/disappearing anatomical structures (for example induced by specific anatomical structures featured by the involved imaging modalities, peristaltic activity or any potential image artifacts) can typically be circumvented by increasing the weight of a spatial motion regularization term (see the improvement of the FEP with increasing α values on the green curves of Fig. 6.7a and 6.7b). This, however, may have a negative impact on locally moving structures (see the deterioration of the FEP with increasing α values on the blue curve of Fig. 6.7a and 6.7b). Using our method, the robustness against structure appearance/disappearance is, by construction, not only taken into account by the motion regularization term, but also in the proposed data fidelity term. The reason is that our criterion converts the image into an "edge representation", and only strong edges that occur in both modalities are favored, due to the use of the weight $w_T(\vec{r})$ in Eq. (6.2).

It can be observed in Fig. 6.9 that the skull is properly registered by the EVolution algorithm, although it features hyper- and hypo-intense signals in CT- and MR- images, respectively. This arises from the fact that the absolute value of the cosine in Eq. (6.2) favors the alignment of both parallel and anti-parallel edges, irrespective of the gradient direction. As a consequence, in this particular example, the strong contrast of the cortical bone in the CT-image and the hyper-intense signal of the subcutaneous fat layers in the MR-images dominate the overall convergence process of the variational algorithm.

In order to mitigate the fact that the image features under the new representation might not be discriminative enough (especially in the calculation of the partial derivatives of Eq. (6.8)), the defined normalized criterion performs a weighted average of the score obtained for each edge that occur in both modalities within patches. The patch size is thus a free parameter which can be increased to improve the robustness against anatomical structure without counterpart between the image to register and the reference one. However, it must be also underlined that, similarly to the parameter α, the patch size also impacts the spatial regularity of the estimated motion field. Decreasing these two parameters generally results in a poor estimation of the displacement, due to numerical instabilities. Identically, for values increasing toward infinity, the tested reliability criteria indicated a poor estimation of the displacement, as the smoothness of motion constrains the velocity amplitude estimation. As a consequence, a good compromise of the choice of both parameters is essential for a reliable and accurate co-registration. Using the proposed default user-defined parameters (i.e $\alpha = 0.5$ and a patch size of $11 \times 11 \times 11$ voxels), results obtained using the proposed method were comparable to those obtained using the MIND algorithm, and this for all tested experimental conditions (image pre-filtering, image intensity variation, noise perturbation). It is interesting to note in Fig. 6.4 that the proposed EVolution method could systematically be

further improved, however for this specific 10 CT-scans data set, by simply decreasing the patch size. Using default calibration parameters, the proposed method outperforms the existing MIND algorithm on the bladder filling scenario, for which structures in the image to register haven't counterparts in the reference image.

It can be noticed that the addition of noise perturbation somewhat improved registration results, especially for high patch size (see the blue curve in Fig. 6.4a and 6.4b). The presence of noise indeed alleviated possible divisions by values close to 0 in Eq. (6.1) which, in turn, slightly stabilized the numerical process. A typical alternative method to get around this would be to add a very small positive constant ϵ in the denominator of Eq. (6.1) ($\epsilon = 0.1$ provided results similar to the ones reported by the blue curves of Fig. 6.4).

6.5 Conclusion

An image registration tool box including complementary multi-modal algorithms is a necessary prerequisite in the field of medical imaging, with respect to the ever growing need to align images of a same scene acquired at different time instants, using different sensors and/or from a different point-of-view.

In this chapter we proposed a non-rigid multi-modal registration method that aligns edges in both modalities using a variational scheme that provides a fast algorithm with a low number of parameters to tune. Using the proposed default user-defined parameters (i.e $\alpha = 0.5$ and a patch size of $11 \times 11 \times 11$ voxels), results obtained using the proposed EVolution method were comparable to those obtained using the MIND algorithm, and this for all tested experimental conditions (image pre-filtering, image intensity variation, noise perturbation). This demonstrates that both methods performed with comparable accuracy and precision for the general case. Using identical calibration parameters, it was also demonstrated that the proposed method outperforms existing approaches on a specific scenario for which structures in the image to register do not have counterparts in the reference image.

It must be underlined that, in their excursion due to respiration, the lung, liver and kidneys slide on the thoracic and abdominal walls which, by comparison, manifest a reduced amount of motion. The impact of the arising shearing effects on the motion regularization term may not be negligible, and will need to be addressed in future works. Moreover, future studies will challenge the proposed method for various non-rigid multi-sensor registration scenarios (e.g MRI/echography, PET/MRI, etc...).

Acknowledgements

This work was financially supported by ITEA (project 12026, SoRTS) and the European Research Council (project ERC-2010-AdG-20100317, Sound Pharma).

Bibliography

[1] V.R.S. Mani and S. Arivazhagan. Survey of medical image registration. *Journal of Biomedical Engineering and Technology*, 1(2):8–25, 2013.

[2] D.L. Hill, P.G. Batchelor, M. Holden, and D.J. Hawkes. Medical image registration. *Phys Med Biol.*, 46(3):R1–R45, 2001.

[3] J.B. Maintz and M.A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[4] J. Weickert, A. Bruhn, N. Papenberg, and T. Brox. Variational optic flow computation: From continuous models to algorithms. In *International Workshop on Computer Vision and Image Analysis (ed. L. Alvarez), IWCVIA03, Las Palmas de Gran Canaria*, 2003.

[5] M. Ries, B. Denis de Senneville, S. Roujol, Y. Berber, B. Quesson, and C. T. W. Moonen. Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues. *Magnetic Resonance in Medicine*, 64(6):1704–1712, 2010.

[6] B. Denis de Senneville, A. El Hamidi, and C. T. W. Moonen. A direct PCA-based approach for real-time description of physiological organ deformations. *IEEE Transactions on Medical Imaging*, 34(4):974–982, 2015.

[7] B.K.P. Horn and B.G. Schunck. Determining optical flow. *Artificial intelligence*, 17:185–203, 1981.

[8] B. Denis de Senneville, M. Ries, G. Maclair, and C.T.W. Moonen. MR-guided thermotherapy of abdominal organs using a robust PCA-based motion descriptor. *IEEE Transactions on Medical Imaging*, 30(11):1987–1995, 2011.

[9] M. Mellor and M. Brady. Phase mutual information as a similarity measure for registration. *Medical Image Analysis*, 9:330–343, 2005.

[10] E. Haber and J. Modersitzki. Intensity gradient based registration and fusion of multi-modal images. *MICCAI*, pages 726–733, 2006.

[11] C. Wachinger and N. Navab. Entropy and laplacian images: structural representations for multi-modal registration. *Medical Image Analysis*, 16:1–17, 2012.

[12] P. Viola and W. Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24:137–154, 1997.

[13] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.

[14] J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. Mutual-information-based registration of medical images: A survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004, 2003.

[15] M.P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F.V. Gleeson, S.M. Brady, and J.A. Schnabel. MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical Image Analysis*, 16(7):1423–1435, 2012.

[16] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *J Multiscale Model Simul*, 4(2):490–530, 2005.

[17] C. Sutour, J.-F. Aujol, C.-A. Deledalle, and B. Denis de Senneville. Edge-based multi-modal registration and application for night vision devices. *Journal of Mathematical Imaging and Vision*, 53(2):131–150, 2015.

[18] Y. Sun, M.P. Jolly, and J.F. Moura. Integrated registration of dynamic renal perfusion MR images. In *ICIP*, pages 1923–1926. IEEE, 2004.

[19] J. Pluim, J. Maintz, and M. Viergever. Image registration by maximization of combined mutual information and gradient information. *IEEE Transactions on Medical Imaging*, 19(8):809–814, 2000.

[20] R.C. Gonzalez and R.E. Woods. *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., 2006.

[21] E. Castillo, R. Castillo, R. Guerra, V. Johnson, T. McPhail, A. Garg, and T. Guerrero. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Physics in Medicine and Biology*, 54:1849–1870, 2009.

[22] E. Castillo, R. Castillo, J. Martinez, M. Shenoy, and T. Guerrero. Four-dimensional deformable image registration using trajectory modeling. *Physics in Medicine and Biology*, 55:305–327, 2010a.

[23] E. Castillo, R. Castillo, J. Martinez, and T. Guerrero. Ventilation from four-dimensional computed tomography: density versus jacobian methods. *Physics in Medicine and Biology*, 55(16), 2010b.

[24] C. Zachiu, N. Papadakis, M. Ries, C. T. W. Moonen, and B. Denis de Senneville. An improved optical flow tracking technique for real-time MR-guided beam therapies in moving organs. *Physics in Medicine and Biology*, 60(23):9003, 2015.

[25] J. West, J.M. Fitzpatrick, M.Y. Wang, B.M. Dawant, C.R. Maurer, R.M. Kessler, R.J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P.A. Van den Elsen, P.F. Hemler, S. Napel, T.S. Sumanaweera, B. Harkness, P.F. Hemler, D. Hill, C. Studholme, G. Malandain, X. Pennec, M.E. Noz, G.Q. Maguire, M. Pollack, C.A. Pelizzari, R.A. Robb, D. Hanson, and R.P. Woods. Comparison and evaluation of retrospective intermodality image registration technique. *SPIE Medim'96, SPIE 2710*, pages 332–347, 1996.

[26] J. West, J.M. Fitzpatrick, M.Y. Wang, B.M. Dawant, C.R. Maurer, R.M. Kessler, R.J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P.A. Van den Elsen, S. Napel, T.S. Sumanaweera, B. Harkness, P.F. Hemler, D. Hill, D. Hawkes, C. Studholme, J.B.A. Maintz, M.A. Viergever, G. Malandain, X. Pennec, M.E. Noz, G.Q. Maguire, M. Pollack, C.A. Pelizzari, R.A. Robb, D. Hanson, and R.P. Woods. Comparison and evaluation of retrospective intermodality image brain image registration techniques. *Journal of Computer Assisted Tomography*, 21:554–566, 1997.

[27] J. M. Fitzpatrick, J. B. West, and C. R. Maurer. Predicting error in rigid-body, point-based registration. *IEEE Transactions on Medical Imaging special issue on Image Registration*, 17:694–702, 1998.

[28] C.R.J. Maurer, J. Fitzpatrick, M. Wang, R.L.J. Galloway, R. Maciunas, and G. Allen. Registration of head volume images using implantable fiducial markers. *IEEE Transactions on Medical Imaging*, 16:447–462, 1997.

[29] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski. A database and evaluation methodology for optical flow. *International Journal of Computer Vision*, 92(1):1–31, 2011.

[30] NVIDIA. NVIDIA CUDA: compute unified device architecture. *Programming guide, NVIDIA Corporation, 2.0 edition*, 2008.

[31] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens. Non-rigid image registration using conditional mutual information. *IEEE Transactions on Medical Imaging*, 29:19–29, 2010.

CHAPTER 7

Real-Time Non-Rigid Target Tracking for Ultrasound-Guided High-Intensity Focused Ultrasound Therapies

"Since we cannot change reality, let us change the eyes which see reality."
- Nikos Kazantzakis

This chapter is based on:
C. Zachiu, M.G. Ries, P. Ramaekers, J.L. Guey, C.T.W. Moonen, and B. Denis de Senneville. Real-Time Non-Rigid Target Tracking for Ultrasound-Guided Clinical Interventions. *Phys. Med. Biol.*, 62(20): 8154-8177, 2017.

Abstract

Biological motion is a problem for high-intensity focused ultrasound (HIFU) therapies when conducted in mobile/deformable organs due to the targeted pathology moving/deforming with the organ. This may lead to high miss rates and/or incomplete treatment of the pathology. Therefore, real-time tracking of the target anatomy during the intervention would be beneficial for such applications.

Previous studies which conduct HIFU therapies under magnetic resonance imaging (MRI) - guidance, propose using the magnitude of the MR-thermometry images for the purpose of target tracking, via image registration. This, however, leads to MR-acquisition sequences which are suitable, but sub-optimal for both tracking and thermometry. A system which allows simultaneous MR and B-mode ultrasound (US) guidance was developed at the UMC Utrecht. This provides the possibility to delegate imaging in the scope of tracking to the US system, while the MR-scanner can be used exclusively for thermometry or other functional imaging. However, B-mode US images are intrinsically altered by speckle noise, introducing incoherent gray-level intensity variations. This may prove problematic for existing intensity-based registration methods. In the current study we address US-based target tracking by employing the recently proposed EVolution registration algorithm. The method is, by construction, robust to transient gray-level intensities. Instead of directly matching image intensities, EVolution aligns similar contrast patterns in the images. Moreover, the displacement is computed by evaluating a matching criterion for image sub-regions rather than on a point-by-point basis, which typically provides more robust motion estimates. However, unlike similar previously published approaches, which assume rigid displacements in the image sub-regions, the EVolution algorithm integrates the matching criterion in a global functional, allowing the estimation of an elastic dense deformation.

The approach was validated for soft tissue tracking under free-breathing conditions on the abdomen of 7 healthy volunteers. Contact echography was performed on all volunteers, while 3 of the volunteers also underwent parallel fast MR-imaging/standoff echography on the hybrid MR-USg-HIFU system. The method demonstrated on average an accuracy of $\sim$1.5 mm and submillimeter precision. This, together with a computational performance of 20 images/s make the EVolution algorithm an attractive solution for real-time target tracking during US-guided HIFU therapies.

7.1 Introduction

Therapy by high-intensity focused ultrasound (HIFU) is currently the only percutaneous thermal ablation procedure that is completely non-invasive.[1] It has already been applied successfully for the treatment of numerous pathologies such as uterine fibroids,[2] prostate cancer,[3] liver tumors,[4,5] thyroid nodules,[6] palliation of bone metastases,[7] breast fibroadenoma[8] and tremor-dominant Parkinson's disease.[9] Treatment becomes challenging, however, when therapy is conducted in the abdomen and the lower thorax. Physiological motion such as respiration, peristalsis and/or spontaneous motion, may cause the targeted area to continuously change position over the duration of the intervention. If left un-addressed, this can lead to the therapeutic energy dose to be diverted from the target area and in turn to the under-treatment of the pathology. If motion effects are severe, damage to organs at risk may occur.[10–12] Previous studies conducting HIFU therapies under magnetic resonance imaging (MRI) - guidance, propose using the magnitude images resulting from dynamic MR-thermometry[13] as a means to track the target area and its surroundings.[10,14–16] This was achieved by establishing one of the images as positional reference, with all the subsequent acquisitions being aligned to this reference scan via image registration.[17] The resulting displacements were then used to steer the HIFU beam according to the target's motion. A downside of such an approach is that registration algorithms in general, impose additional constraints on the MR-thermometry acquisition sequence. This typically leads to a compromise on the MR acquisition parameters, resulting in a series of images which are suitable, but sub-optimal for either tracking or thermometry. Migrating tracking-related imaging to an independent imaging system is thus preferable.

Tracking based on B-mode ultrasound (US) images has been previously proposed in several independent studies. Seo *et al*, for example, designed a system consisting of two US imaging probes integrated with a HIFU transducer.[18–20] One probe was placed at the center of the parabolic transducer, while the second was placed on the edge of the transducer at a 45° angle with respect to the first one. This arrangement of the US probes provided two orthogonal US imaging planes intersecting at the transducer focal point, offering the possibility to track both in-plane and out-of-plane motion. The displacements were estimated either by matching the intra-operative target surface to a pre-operative model[18,19] or by tracking the discriminative speckle pattern produced in the US images by the HIFU-induced lesion.[20] Beam steering with respect to the estimated motion was achieved by visual servoing implemented in a three-axis robot. Promising results were obtained with tracking errors of 1.5 - 2 mm, however, *in-vivo* validation is yet to be performed. A different integrated system capable of providing simultaneous magnetic resonance imaging (MRI) and US guidance during HIFU therapies was developed by Auboiroux *et al* and is described in several related studies.[11,21,22] US guidance was provided by a probe placed on a custom-built support, allowing 5 degrees of freedom for movement

around the patient. The support was attached to the MR patient table, which also contained the HIFU transducer. US-based motion estimation was achieved both by an optical flow approach[11] and also by a more advanced estimation-prediction algorithm relying on a population-based statistical model.[22] This lead to motion estimates with a 1-3 mm accuracy, available with near real-time latency. A system capable of providing simultaneous MR and US guidance was also developed at the UMC Utrecht, Netherlands. This would hypothetically allow motion tracking via the US imager, while the MR scanner is dedicated exclusively to MR-thermometry or other functional imaging.

A large variety of US-based motion tracking methods have already been proposed in the literature.[23] According to the classification made in the review work of de Luca *et al*,[23] registration methods can be coarsely divided into three categories: intensity-based, feature-based and hybrid. Hybrid approaches, in particular, have the benefit of an improved accuracy and precision of the estimated deformation compared to intensity-based and feature-based methods.[23] Cifor *et al*,[24] for example, propose a so-called Hybrid Feature-based Diffeomorphic (HFD) registration, inspired by the diffeomorphic Log-Demons registration framework. The method estimates the deformation between two US images by the means of a diffeomorphic transformation, which in turn is computed based on a set of demon-like forces[25] driven by local image features, including: image intensity, local phase[26] and phase congruency.[27] The method was used to track liver tumors on B-mode US images acquired on 7 patients. Results have shown an improvement of the Dice similarity coefficient[28] from an average of 73.9 ± 19.5, in the absence of tracking, to 90.4 ± 5.5. A drawback of the approach is that it requires the optimization and/or computation of several input parameters to which the method is more or less sensitive, encumbering its application on a case-by-case basis. Moreover, the method is computationally demanding, requiring 5 min for registering one pair of 2D images. Hybrid methods, making use of both image intensity and contour information, have also been proposed in the past for contour tracking in B-mode US images.[29–31] The approach proposed by Li et *al*[29] relies on the active contour framework[32] in order to track from frame-to-frame the deformation of a manually-selected image contour. As an application scenario, the method was employed for automatic US-based tracking of tongue movements during speech. It was demonstrated to have a good accuracy, with an average sum of distances between the automatically tracked tongue contour and the delineations performed by experts of 0.54 - 1.06 mm. While robust, the method requires a moderate amount of user interaction. This includes manual initialization of the contour, its re-initialization in case of large inter-frame displacements and the tuning of several input parameters. Moreover, since the method uses the state of the contour from the previous frame as an initialization for the current frame, it is susceptible to error accumulation. The approaches proposed by Wang and Huang,[30,31] on the other hand, make use of prior information learned on training datasets in order to perform automatic contour detection and/or motion characterization. Both methods were used in the context of echocardiography, for automatic tracking

of heart contours. The solution proposed by Wang *et al* was shown to have an accuracy and precision of 2.68 mm and 2.63 mm, respectively, for tracking "end - systole" - "end - diastole" heart deformations. Similarly, the approach of Huang *et al* demonstrated promising results, with a Hausdorff distance[33] between automatically-tracked and expert-defined contours of 2.95 ± 0.62 mm for the endocardium and 3.03 ± 0.76 mm for the epicardium. The method was also evaluated in terms of the Dice similarity coefficient, with an average value after registration of 93 and 97 for the endocardium and epicardium, respectively. Such approaches, however, have complex numerical schemes and require moderate user interaction. Moreover, due to the fact that they are learning-based, the prior information has to be updated on a case-by-case basis, having also difficulties in handling deformation patterns which are not included in the learned atlas/dictionary. As previously mentioned, many other US-based tracking solutions are available in prior art, dedicated to various applications. While individual validation of the methods was performed using several different criteria, it was concluded in the review work of de Luca *et al*[23] that, on average, existing US-based tracking methods can achieve an accuracy of 1 - 2 mm.

For the interventional guidance of HIFU therapies, a motion tracking solution with real-time capabilities is of particular importance. In the scope of this chapter, a registration method is defined as "real-time" if it is capable of aligning two images with a sufficiently low latency such that the monitoring and control of the intervention is not hampered by the associated delay. Naturally, the term "sufficient" depends on the particularities of the application. In this sense the so-called variational registration methods[34] are a particularly attractive option for real-time US-based motion tracking.[10] This is due to both their fast numerical schemes and the reduced number of required input parameters, with the latter leading to an increased ease-of-use in a clinical setting on a case-by-case basis.

In the present study we propose a real-time US-based motion tracking solution based on the EVolution algorithm, recently proposed by Denis de Senneville *et al*.[35] EVolution is a variational registration method which, intuitively speaking, in order to estimate the deformation between two images, searches for similar local contrast patterns in the two images, which it then tries to align. With respect to the classification made by de Luca *et al* in,[23] EVolution falls into the category of hybrid registration methods. This is due to the fact that the method computes a deformation field on a point-by-point basis, iteratively minimizing a global matching criterion between the images (which is specific to intensity-based methods), however, it does not operate directly on image intensities, matching instead local contrast patterns in the images (which is historically more used in feature-based approaches). Similar to the methods proposed in[36] and,[37] the EVolution algorithm operates on square sub-regions (called patches or windows) rather than on a point-by-point basis. This typically provides a robust local displacement, since the motion in a point-of-interest is calculated based on several pixels/voxels. However, while the aforementioned approaches assume a rigid motion inside the sub-regions, EVolution integrates the patch-

based matching criterion in a global variational framework, which allows estimating an elastic deformation on a point-by-point basis.
The current study validates the EVolution algorithm with both *ex-vivo* and *in-vivo* US-based motion estimation experiments. The method was tested hereby for both contact echography as well as for standoff echography. In particular for the standoff case, the evaluation was conducted directly on the hybrid MR-US-guided HIFU system developed at our institute, which ultimately imposed additional constraints on obtainable image quality.

7.2 Method description

7.2.1 Proposed solution for B-mode US-based motion tracking

7.2.1.1 The EVolution algorithm

Motion estimation via the EVolution algorithm, implies finding the minimizer of the following functional:

$$E(T) = \int_{\Omega} e^{-C(T)} + \frac{\alpha}{2}\left(\|\vec{\nabla}u\|^2 + \|\vec{\nabla}v\|^2\right) d\vec{r} \tag{7.1}$$

with

$$C(T) = \frac{\int_{\Gamma}\left|\vec{\nabla}_I(T(\vec{r}))\cdot\vec{\nabla}_J(\vec{r})\right| d\vec{r}}{\int_{\Gamma}\|\vec{\nabla}_I(T(\vec{r}))\|_2\|\vec{\nabla}_J(\vec{r})\|_2 d\vec{r}} \tag{7.2}$$

where $T = (u, v)$ is the 2D displacement to be estimated, I and J are the reference and the moving image, α is a parameter linking the two terms of the functional, $\vec{\nabla}$ is the spatial gradient operator, $\|\cdot\|_2$ is the Euclidean norm, Ω is the image domain, $\vec{r}$ is a spatial location and Γ is a square image patch centered on the pixel of interest. The term $C(T)$ in Eq. 7.2 can be rewritten as:

$$C(T) = \frac{\int_{\Gamma} w_T(\vec{r})\,|\cos(\Delta\theta_T(\vec{r}))|\,d\vec{r}}{\int_{\Gamma} w_T(\vec{r})d\vec{r}} \tag{7.3}$$

with the terms $w_T(\vec{r})$ and $\Delta\theta_T(\vec{r})$ being computed based on the magnitude M and the phase θ of the reference and moving image spatial gradients:

$$\begin{aligned} w_T(\vec{r}) &= M_I(T(\vec{r}))M_J(\vec{r}) \\ \Delta\theta_T(\vec{r}) &= \theta_I(T(\vec{r})) - \theta_J(\vec{r}) \end{aligned} \tag{7.4}$$

The term $C(T)$ can be interpreted as follows: $|\cos(\Delta\theta_T(\vec{r}))|$ in Eq. 7.3 favors transformations that align edges in the reference and the moving image, regardless of any possible contrast reversals. This basically makes the algorithm capable of both mono and multi-modal registration. On the other hand, $w_T(\vec{r})$

favors strong edges that exist in both images. Thus the method is, by construction, robust to transient structures.
Minimizing $e^{-C(T)}$ alone, however, is a severely ill-posed problem with infinite solutions. For this reason, the functional in Eq. 7.1 contains a regularization term weighted by α, constraining the estimated displacements to be spatially differentiable, which "enforces"/mimics plastic deformations.[38] This renders the optimization problem overdetermined and therefore solvable with variational methods.

7.2.1.2 Implemented optimization scheme

The EVolution functional was minimized via the iterative procedure described in.[35] Similarly, the iterative process was stopped when the average absolute difference between the displacements at the current and the previous iteration was smaller or equal to 10^{-3} pixels.
Due to the strong non-linearity of the functional, the algorithm has difficulties estimating large displacements. For this reason, a coarse-to-fine strategy was adopted in which motion is estimated step-by-step starting from a 16-fold downsampled version of the images up to their original resolution. The displacements estimated at a particular resolution level are then used as an initialization for the next level.

7.2.2 Experimental setup

In the current study, the capabilities of the EVolution algorithm for B-mode US tracking, were validated for both contact and standoff echography, with a separate experiment dedicated to each of the two US-acquisition modalities. The two experiments are detailed in the following sections.

7.2.2.1 *In-vivo* contact echography study

The current healthy volunteer study was conducted in agreement with the required standards and regulatory requirements. Ethical approval was provided by the Ethics Board of the University Medical Center Utrecht. B-mode US imaging was performed under free-breathing conditions on the liver and one of the kidneys of 7 healthy volunteers. The images were acquired using a Telemed C3.5/60/128Z convex array transducer (number of elements = 128, radius of curvature = 65 mm, field-of-view = 59°) connected to a LS128 EXT-1Z Telemed beamformer. The acquisition sequence employed the following parameters: Frequency = 5 MHz, Depth = 80 mm, FPS = 18 and pixel size = 0.25×0.25 mm^2. The images were saved on disk and processed offline. Since approximately half of the far field-of-view of the images contained either anatomical structures that were not of interest in the scope of this study, either exclusively noise due to signal attenuation, these parts were manually discarded from the images. Note,

however, that this step was performed only in the scope of the present chapter and can be omitted in practice. For computational purposes, the manual discarding of parts of the field-of-view was followed by a spatial sub-sampling of the images by a factor of 2. These two pre-processing steps resulted in each image in the series having a size of 256×256 with a pixel size of 0.5×0.5 mm^2. For both the liver and the kidney datasets, three identifiable landmarks were selected, for which the tracking capabilities of the EVolution algorithm was evaluated. More specifically, for the liver this consisted in the lower boundary of segment # 4 (landmark # 1) and two branches of the hepatic artery/portal vein (landmarks # 2 and # 3). The two blood vessels were chosen in approximately the same liver regions across all of the 7 volunteers. For the kidney, the selected landmarks were the caudal boundary (landmark # 1) and the two extremities of the major calyx in the cranio-caudal direction (landmarks # 2 and # 3). In case not all of the landmarks would be visible in a single sequence, additional imaging would be performed on the volunteer, with the FOV focused on the missing landmark(s). Note that, in some cases, the breathing dynamics of the volunteer would change in between the sequences (i.e. becoming deeper or shallower). The landmarks were manually tracked on a time-frame of $\sim$ 8 s (which usually included 2-3 breathing cycles), and the coordinates of the each landmark were stored and used as a gold standard. The displacements in each liver and kidney dataset were then estimated on a pixel-by-pixel basis via the EVolution algorithm and the resulting coordinates of the same landmarks were compared against the manually established gold standard. The comparison was performed both visually, by plotting the estimated and the manually determined trajectory for the landmarks, and by the evaluating the target registration error (TRE)[39] defined as:

$$\text{TRE[mm]} = \sqrt{(x_I + u(\vec{r}_I) - x_J)^2 + (y_I + v(\vec{r}_I) - y_J)^2} \tag{7.5}$$

where (x_I, y_I) and (x_J, y_J) are the manually determined landmark coordinates in the reference and the moving image, and $(u(\vec{r}_I), v(\vec{r}_I))$ is the displacement estimated by the EVolution algorithm for that particular landmark.

7.2.2.2 Standoff echography study

A system was developed at our institute that allows clinical HIFU interventions under MR and/or US guidance (see Fig. 7.1(a)). The system is a modified Philips Sonalleve abdominal HIFU platform (Philips Healthcare, Vantaa, Finland), which is integrated in a 1.5T Philips Achieva scanner (Philips Healthcare, Best, Netherlands). Therapeutic HIFU energy in this modified system is provided by a custom-built sparse spherical phased transducer (IMASONIC, Voray sur l'Ognon, France), which replaces the original spherical transducer and is described in detail here.[40] The custom-built transducer is composed of 256 elements, operates at a frequency of 1.3 MHz and has an aperture diameter and radius of curvature of 16cm. Embedded in the therapeutic matrix of the transducer is a linear 128-element single-crystal linear phased array, which

operates at a frequency of 3.5 MHz (see Fig. 7.1(b)). The imaging array provides a co-axial image of the focus area of the HIFU transducer at 16 cm. Both the therapeutic array as well as the embedded US-imaging array are passively and actively MR-compatible and thus allow full simultaneous operation with real-time MRI in absence of mutual interference. A passively RF-shielded Verasonics beamformer,[41,42] which can thus also be operated in the Faraday cage of the MR-scanner, is used over a shielded cable connector to drive the imaging probe. The beamformer is operated via a fiber-optical PCIe link from a computer located adjacent to the MR-scanner console outside the Faraday cage. Time synchronization between the MR-acquisition level and the Verasonics beamformer was achieved by the means of a dedicated TTL line between the two systems and a Cortex M3 microcontroller as an intermediary. A third TTL-connection from the microcontroller allowed additional time synchronization of the HIFU-generator. This setup allowed the following two configurations, which were used in the scope of this chapter: 1) MRI as the master-clock and the US-imager and the HIFU system slaved (with independent delays) and 2) The microcontroller as the master-clock with both HIFU and US-imager slaved, while the MRI runs asynchronously.
One purpose of the standoff echography study was to analyze the tracking capabilities of the EVolution algorithm as a means for B-mode US-based motion compensation during hybrid MR-US-guided HIFU therapies. Two separate experiments were conducted in this regard: one on a gel phantom undergoing a known motion pattern and an *in-vivo* study on three healthy volunteers. A third independent experiment was carried-out on a gel phantom in order to ensure that HIFU ablations are possible with minimal interference between the HIFU therapeutic signal and the US imaging signal.

(a) (b)

Figure 7.1: The custom-built HIFU system developed at our institute, which can provide MR and/or B-mode US guidance. (a): Overall setup. (b): The HIFU transducer together with the integrated US imaging probe.

a *Ex-vivo* standoff echography phantom experiment Since motion estimation algorithms typically rely on structural information in order to provide accurate displacements, a custom phantom was built for the current experiment. The phantom consisted in two perpendicular grid-like structures placed inside a cylindrical plastic casing, provided with a lid on one end and with an acoustic-transparent mylar membrane at the other. Prior to sealing the plastic case, the phantom was filled with an acoustically semi-transparent polymer gel. The phantom was then placed on a motorized platform, which allowed to move the US-phantom in a linear fashion over the acoustic window of the HIFU platform. The motor of the platform (USR60-E3T, Shinsei Corporation, Tokyo Japan) was controlled by an embedded microcontroller (ARM Cortex M3), which provided full closed-loop proportional-integral-derivative (PID) control over the linear motion pattern. In the scope of the experiment, a motion pattern corresponding to the average head-foot liver displacement of a free-breathing volunteer from previous studies[43] was loaded into the controller algorithm (duration $\sim$ 90 s, with an original sampling density of 6 Hz, up-sampled by sinc interpolation to 100 Hz). Both the motion platform and the acoustic phantom were designed such that the HIFU transducer (and implicitly the US imager) had acoustic access to the phantom during the entire motion cycle.
Simultaneous MR and B-mode US imaging was performed on the moving phantom. The MR acquisition sequence was a multi-shot (4 shots per image) gradient recalled echo with an echo planar read-out, employing the following parameters: TR = 45 ms, TE = 11 ms, FA = 25°, BW_{read} = 1132 Hz, image size = 176×176, pixel size = 2×2×7 mm^3, acquiring a total of 500 images over a duration of ∼90 s. Synchronously with the MRI, a B-mode US-image series was acquired using a 3-cycle pulse with a frequency of 3.5 MHz, image size 1024×256, a 0.25×0.25 mm^2 pixel size and at the same imaging rate as the MR sequence. The images were obtained using a spatial compounding over 6 angles (45° angle range) and synthetic aperture focusing, with the virtual source placed at a depth of 125 mm.[44,45] Both the MR and the US images were saved on disk and processed offline. Timing-wise, for the synchronized MR-US acquisition, the US-imager is triggered with the acquisition of the first k-space segment of an MR-slice (which acquires the k-space center). After the trigger signal is received, the US-imager immediately acquires the 6 compounding angles, while the MRI acquires simultaneously the first k-space segment. The US-acquisition is completed with the first k-space segment acquisition.
During the *ex-vivo* experiment, the motion pattern induced to the phantom by the motorized platform can be used as a gold standard. In comparison, for *in-vivo* scenarios, obtaining a reliable ground-truth *in-vivo* in a non-invasive manner can be a difficult task. For this reason, in anticipation to the *in-vivo* study described in the following section, the phantom experiment was used to validate both the EVolution algorithm applied on the US image series and the optical flow algorithm[43] applied on the simultaneously acquired MR images. The output provided by the optical flow algorithm on MR images will be used as a validation mechanism during the *in-vivo* study on the three healthy vol-

unteers.
The phantom trajectory estimated in both the MR and the US image series was compared to the injected motion pattern. However, note that the MR and the US images were acquired in a different coordinate system. Thus, in order to ensure consistency between the displacements estimated on the two modalities, prior to registration, the MR and the US images were re-sampled on a common frame-of-reference. The axes of the new coordinate system were aligned with the axes of the MR images, whereby, the size of the the common reference frame was established at 256×256 with a 1×1 mm^2 pixel size. Moreover, prior to re-sampling, the far field of the US images was cropped-out, since for the current experiment it did not contain useful information. The US images also underwent a histogram equalization procedure. The average phantom displacements estimated from the two modalities were analyzed in relation to the injected motion pattern both in terms of followed trajectory and the pixel-wise error in flow endpoint (FEP):[46]

$$\mathrm{FEP[mm]} = \sqrt{(u_{\mathrm{gold}} - u_{\mathrm{est}})^2 + (v_{\mathrm{gold}} - v_{\mathrm{est}})^2} \tag{7.6}$$

where $(u_{\mathrm{gold}}, v_{\mathrm{gold}})$ and $(u_{\mathrm{est}}, v_{\mathrm{est}})$ are the gold standard and the estimated motion fields, respectively. For the current experiment, u_{gold} is the displacement induced to the moving phantom, while v_{gold} was set to 0, since the moving platform only underwent a 1D translational motion. Thus, unlike the contact echography study, the gold standard here is a set of displacements, rather than a set of coordinates providing the location of a landmark over the US image series. By definition, the FEP relates two motion vector fields to one-another, reason for which, in this case, it was deemed to be a more suitable metric than the TRE.

b *In-vivo* standoff echography study The current healthy volunteer study was conducted in agreement with the required standards and regulatory requirements. Ethical approval was provided by the Ethics Board of the University Medical Center Utrecht. Simultaneous MR and B-mode US imaging was carried-out on the abdomen of three healthy volunteers. The MR acquisition sequence was a multi-shot (9 shots per image) gradient recalled echo with echo planar readout employing the following parameters: TR = 36 ms, TE = 11 ms, FA = 30°, $\mathrm{BW_{read}}$ = 1131 Hz, image size 176×176, voxel size = 2×2×7mm^3, acquiring a total of 300 images over a duration of ∼100 s. The US acquisition sequence employed the same parameters and had the same timing as in the phantom experiment described in the previous section. The exact same preprocessing and processing steps as in the phantom experiment were carried-out on the data acquired on the healthy volunteers. However, as anticipated from the previous section, the displacements estimated by the EVolution algorithm on the US images were evaluated with respect to the ones estimated by the optical flow on the simultaneously-acquired MR images.[43] Motion analysis was

conducted in a region of interest surrounding the lower boundary of liver segment # 5, using the same motion quality evaluation criteria as in the phantom experiment. However, for the computation of the FEP, $(u_{\text{gold}}, v_{\text{gold}})$ from Eq. 7.6 were replaced with the 2D displacements estimated by the optical flow algorithm on the MR series.

c Interference avoidance scheme between the HIFU therapeutic signal and the B-mode US imaging signal For the system used in the standoff echography study, in case the HIFU transducer and the US imager operate simultaneously, the high amplitude of the the HIFU signal may completely saturate the low intensity electronics of the imaging system receiver. This will most likely lead to severe interference artifacts in the acquired B-mode US images, potentially affecting the tracking capabilities of the EVolution algorithm. In order to address this issue, previous studies propose implementing a synchronization scheme between the HIFU transducer and the US imager such that the HIFU beam is off during US-acquisitions.[47] Another solution, which allows continuous operation of both the HIFU transducer and the US imager, is to filter the composite signal received by the US imager. In this sense, the HIFU therapeutic signal is attenuated to an extent that it has only minimal impact on the B-mode image quality.[48–52] In the scope of this study, the more technically simple solution of a time-synced acquisition was adopted, where US-imaging and HIFU are rapidly interleaved. HIFU-US synchronization was achieved using the setup described in the first paragraph of section 7.2.2.2.
In order to study the effectiveness of the proposed interference avoidance scheme, a HIFU heating experiment was conducted on a moving gel phantom. The phantom was placed on the same motorized platform as in the previous phantom experiment, however, the motion pattern induced here consisted in a simple sinusoidal displacement. A 300 W sonication was performed on the moving phantom, while continuously acquiring B-mode US images (frequency 20 images/s) over a duration of 55 s. MR-thermometry[13] was simultaneously performed, in order to ensure that a lethal amount of thermal damage[53] is achieved at the focus point. Both HIFU energy delivery and MR-Thermometry were gated using an MR pencil-beam navigator, however, the MR scanner and the US imager operated asynchronously. The quality of the B-mode US images was then evaluated both in the presence and in the absence of the proposed interference avoidance method.

7.2.3 Configuration of algorithm input parameters

The EVolution algorithm requires two parameters as input: the regularization parameter α and the patch size Γ (see Eq. 7.1 and 7.2). An exhaustive optimization procedure was performed for both parameters. For both the contact and standoff echography studies, the value of α was varied between 0.1 and 1.0 with an increment of 0.05, while the size of Γ was varied between 3×3 and 21×21 with an increment of 2×2. The combination of α and Γ which

provided the minimum average TRE (for contact US) or FEP (for standoff US) was selected for use. For contact echography, the optimization procedure was carried-out per US sequence, with the average TRE including all three landmarks. For the standoff data, the average FEP was computed over the three volunteers.
The procedure described above also allowed the study of the algorithm's sensitivity to the two input parameters. This was achieved by evaluating the average TRE and FEP computed across all the contact and standoff sequences, respectively, for each combination of α and Γ.

7.2.4 Implementation

A computer unified device architecture (CUDA) implementation was performed for both the EVolution and the optical flow algorithms used to register the US and the MR images, respectively. The implementation was then executed on an nVidia Tesla K20 graphical processing unit (GPU).

7.3 Results

7.3.1 Tracking quality evaluation for contact B-mode echography

Fig. 7.2 showcases the tracking capabilities of the EVolution algorithm in one of the volunteers included in the contact echography study (see section 7.2.2.1). More precisely, Fig. 7.2(a) and 7.2(d) illustrate for the liver and kidney, the images used as positional reference during the tracking process. The red dashed lines in the two figures delineate the contour of the organs of interest, while the yellow numbered arrows indicate the location of the tracked landmarks. Fig. 7.2(b) and 7.2(c) plot the vertical and horizontal displacements of landmark # 1 in the liver, resulting from both manual tracking (red curve) and the EVolution algorithm (blue curve). The same curves were plotted in Fig. 7.2(e) and 7.2(f) for landmark # 2 in the kidney. A good similarity can be observed between manual tracking and the EVolution motion curves.
A quantitative evaluation of the algorithm's performance is provided in Table 7.1 for all 7 volunteers. The TRE between the manually determined and the estimated position of the tracked landmarks is reported for both the liver and kidney. The TRE was evaluated under the format *mean* $\pm$ *standard deviation* both in the presence and in the absence of a registration procedure via the EVolution algorithm. The two statistics were computed on the TREs pooled from all time samples. The TRE in the absence of registration was evaluated by setting $u(\vec{r}_I)$ and $v(\vec{r}_I)$ in Eq. 7.5 to 0, while keeping (x_I, y_I) and (x_J, y_J) unaltered. Since the latter were established as gold standard (thus it is assumed that their values are in correspondence with the true motion), by setting $u(\vec{r}_I)$ and $v(\vec{r}_I)$ equal to 0, the TRE provides the tracking errors/misalignments that

would occur in the absence of a registration procedure. When tracking is enabled, a considerable improvement is attained in all analyzed cases, with a mean and standard deviation of the registration errors under 1.6 mm for most of the landmarks. It is only in a few isolated instances in which the accuracy drops to 1.6 - 1.7 mm.

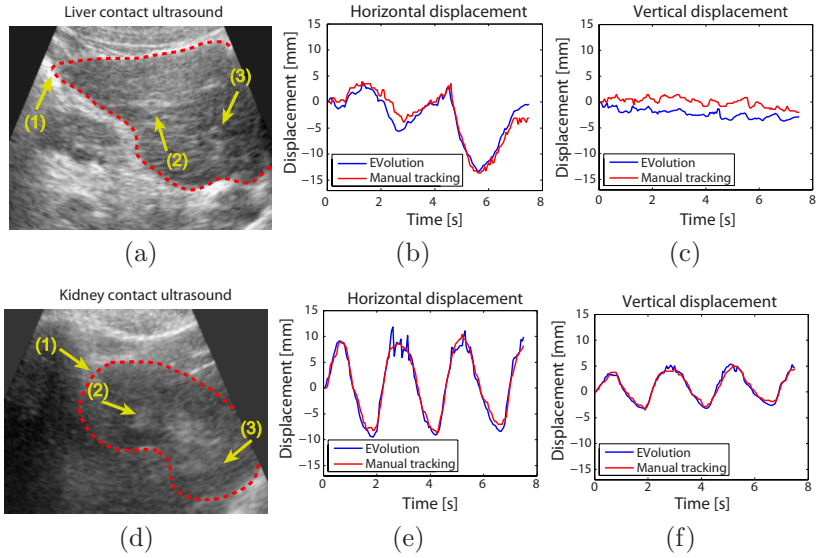


Figure 7.2: Tracking capabilities of the EVolution method for B-mode contact echography: (a), (d): US images used as positional reference for liver and kidney tracking, in one of the volunteers included in the contact echography studies. The red dashed line emphasizes the contour of the organs of interest while the yellow numbered arrows indicate the tracked landmarks. (b), (c): Horizontal and vertical displacements of landmark # 1 in the liver. (e), (f): Horizontal and vertical displacements of landmark # 2 in the kidney. The red curves are the manually determined displacements (used as a gold standard), while the blue curves correspond to the displacements estimated by the EVolution method.

7.3.2 Tracking quality evaluation for standoff B-mode echography

7.3.2.1 Phantom study

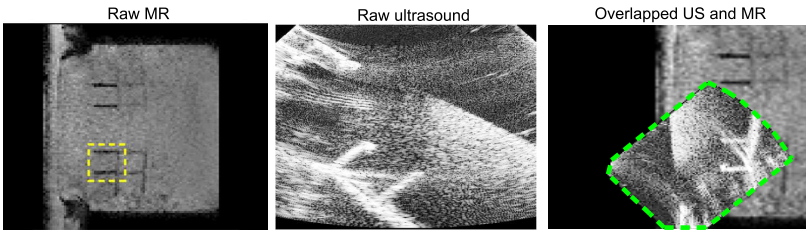
The purpose of this experiment was to validate both the optical flow and the EVolution algorithm against a gold standard consisting in a known motion pattern induced to a phantom. Fig. 7.3 illustrates the images used as positional

Vol. no.	Landmark no.	Target registration error [mm]			
		Liver		Kidney	
		Without registration	With registration	Without registration	With registration
1	# 1	4.34 ± 2.18	1.35 ± 0.88	4.47 ± 2.74	1.58 ± 0.76
	# 2	4.51 ± 2.52	1.09 ± 1.00	5.33 ± 2.98	1.54 ± 0.65
	# 3	4.00 ± 2.00	1.77 ± 1.07	3.75 ± 2.07	1.23 ± 0.64
2	# 1	4.28 ± 3.94	1.54 ± 0.86	5.97 ± 3.17	1.45 ± 0.95
	# 2	5.90 ± 2.82	1.56 ± 0.92	6.01 ± 2.91	1.10 ± 0.62
	# 3	6.31 ± 3.27	1.52 ± 0.83	5.37 ± 2.39	1.37 ± 0.73
3	# 1	4.63 ± 2.27	1.61 ± 0.88	6.05 ± 2.65	1.52 ± 0.65
	# 2	11.80 ± 5.12	1.47 ± 1.06	5.85 ± 2.82	0.90 ± 0.43
	# 3	13.56 ± 8.51	1.43 ± 0.60	6.42 ± 2.72	1.49 ± 1.05
4	# 1	3.61 ± 2.00	1.05 ± 0.58	2.86 ± 1.20	1.29 ± 0.65
	# 2	3.55 ± 2.87	1.03 ± 0.62	4.95 ± 3.99	1.53 ± 0.91
	# 3	4.11 ± 2.63	0.85 ± 0.47	5.23 ± 3.17	0.97 ± 0.48
5	# 1	3.51 ± 2.59	1.19 ± 0.80	6.90 ± 3.45	1.50 ± 0.98
	# 2	3.83 ± 2.60	1.77 ± 0.97	7.48 ± 3.51	1.25 ± 0.73
	# 3	7.01 ± 3.85	1.42 ± 1.05	6.30 ± 2.99	1.24 ± 0.53
6	# 1	5.41 ± 3.06	1.39 ± 0.82	5.58 ± 3.32	1.52 ± 0.89
	# 2	6.76 ± 3.55	1.20 ± 0.76	3.67 ± 2.02	1.70 ± 0.97
	# 3	6.34 ± 2.94	1.70 ± 0.85	4.96 ± 3.08	1.30 ± 0.56
7	# 1	2.88 ± 1.79	1.16 ± 0.55	4.42 ± 1.88	1.57 ± 0.67
	# 2	3.20 ± 2.16	1.53 ± 0.66	5.62 ± 2.34	1.45 ± 1.04
	# 3	7.25 ± 4.22	1.10 ± 0.60	4.71 ± 2.11	1.35 ± 0.66

Table 7.1: Accuracy and precision of the EVolution algorithm following the contact echography study on the 7 healthy volunteers. The table reports the TRE for the tracked landmarks in the liver and kidney, in the absence and in the presence of registration via the EVolution algorithm. Reporting is made under the format: *mean* ± *standard deviation*.

reference during the phantom tracking study (detailed in section a). Fig. 7.3(a) and 7.3(b) showcase the reference images for the MR and US series, in their original frame-of-reference. The yellow dashed line in Fig. 7.3(a) indicates the tracked structure. Fig. 7.3(c) displays the two images in the same coordinate system. In order to show how the content of the two images relate to one-another, the US imaging plane was indicated on the MR image by a green dashed contour in Fig. 7.3(c). Also, the part where the two images overlap was replaced with the content of the US image, after its transformation in the new coordinate system.

Fig. 7.4 depicts the trajectory of the structure indicated by the yellow contour in Fig. 7.3(a) during the first ∼60 s of the phantom study. The red curve corre-



(a) (b) (c)

Figure 7.3: Example of data used during the phantom tracking study: (a), (b): Reference MR and its corresponding US image in their original coordinate system. The yellow dashed contour indicates the structure which was tracked during the experiment. (c): The MR and the US images in (a) and (b), projected onto a common reference frame. The green dashed contour delineates the US imaging plane in the illustrated composite image.

sponds to the known motion pattern induced to the phantom by the motorized platform, while the blue and the green curves are the average displacements estimated for the structure of interest by the optical flow and the EVolution algorithm, respectively, on the MR and the simultaneously acquired B-mode US image series. The trajectory is displayed only for the phantom's direction of travel, since the estimated magnitude of the perpendicular component was negligible. From a visual perspective, a good similarity between the red (used as a gold standard) and the other two curves can be observed. An evaluation of the FEP reveals a reduction of the alignment errors from 3.00 mm $\pm$ 3.35 mm in the absence of tracking to 0.33 mm $\pm$ 0.31 mm for the optical flow and 0.37 mm $\pm$ 0.48 mm for the EVolution algorithm. The statistics on the registration errors were computed based on the pooled errors from all temporal sampling points.

7.3.2.2 In-vivo healthy volunteer study

The simultaneous MR-US healthy volunteer study had the purpose of showcasing the performance of the EVolution tracking method for *in-vivo* standoff B-mode echography. Similar to Fig. 7.3, Fig. 7.5 displays the reference MR and US images acquired on one of the healthy volunteers, both in their original (Fig. 7.5(a) and 7.5(b)) and in a common coordinate system (Fig. 7.5(c)). The red dashed lines contour the visible part of the liver in the two images, while the blue lines delineate the abdominal wall. Again, in order to show how the content of the two images relate to one-another, the US imaging plane is indicated on the MR image in Fig. 7.5(c) by a green dashed line and the overlapping region was replaced with the transformed US image.

Fig. 7.6(a) illustrates, for volunteer #3, the spatial distribution of the tempo-

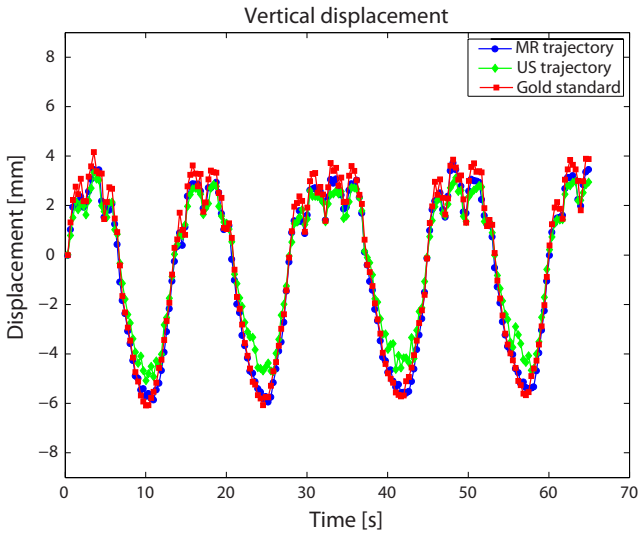
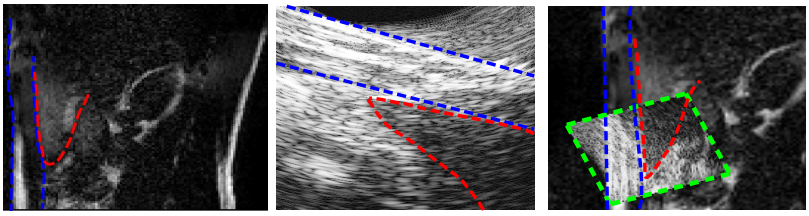


Figure 7.4: Performance evaluation of the optical flow and the EVolution algorithms following the phantom tracking experiment: The average spatial displacement in the direction of travel of the structure of interest during the first ∼60 s of the study. The blue curve corresponds to the displacements estimated by the optical flow algorithm on the MR-image series, while the green curve showcases the displacements estimated by the EVolution method on the simultanously acquired US images. The red curve illustrates the motion pattern induced to the phantom via the motorized platform.

rally averaged FEP between the motion vectors estimated by the optical flow and the EVolution algorithm, respectively, on the simultaneously acquired MR and US image series. The average FEP has been overlaid on the image used as reference during the registration of the standoff US image series. Note that the analysis was limited to a region of interest (ROI) (indicated by the dashed yellow contour in Fig. 7.6(a)) around the lower bound of liver segment # 5. It can be observed that, for the majority of the pixels, the temporally averaged FEP rests below 2 mm. Fig. 7.6(b) plots the temporal evolution of the spatially averaged displacements estimated for the liver of volunteer #3. The average was calculated in the same ROI indicated by the yellow contour in Fig. 7.6(a). The analysis was conducted only for the head-foot component, since the displacement in the anterior-posterior direction was negligible. The blue curve corresponds to the displacements estimated by the optical flow algorithm on the MR image series, which played the role of silver standard in the current experiment, while the green curve is the trajectory estimated by the EVolution method on the simultaneously acquired US image sequence.

Table 7.2 reports the FEP for all 3 volunteers, in the absence and in the pres-



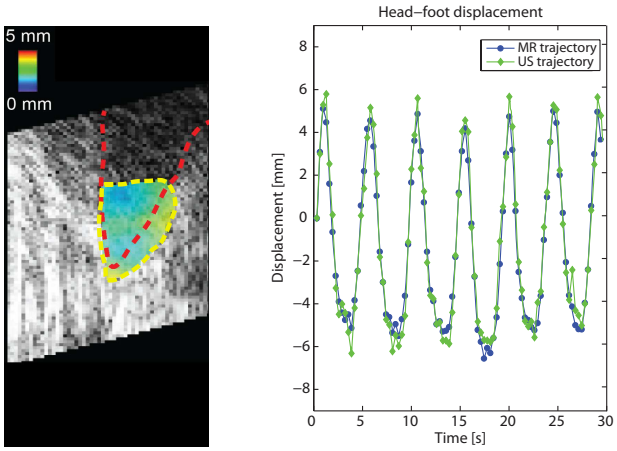
(a) (b) (c)

Figure 7.5: Example of data used for tracking during the *in-vivo* standoff echography study: (a), (b): The MR and its corresponding B-mode US image used as positional reference for volunteer #3, in their original coordinate system. In both (a) and (b) the red dashed line delineates the visible part of the liver, while the blue line delineates the abdominal wall. (c): The two images projected onto a common frame-of-reference. The green contour delineates the US imaging plane on the composite MR-US image. The red and the blue contours maintain the significance they had in (a) and (b).

ence of the EVolution tracking method. Since (u_{gold}, v_{gold}) in Eq. 7.6 stand for the true motion of the tracked structure, the FEP in the absence of the EVolution algorithm was computed by setting (u_{est}, v_{est}) to 0. The values in Table 7.2 report the *mean* ± *standard deviation* of the spatially averaged FEP in a region of interest covering the inferior part of the liver, with the statistics computed on the pooled values from all temporal sampling points. Improvements can be observed in all cases, with both the accuracy and precision of the proposed tracking method residing beneath 1.5 mm. Noteworthy is also the inter-subject variability in tracking performance.

	Error in flow endpoint [mm]	
	Without registration	With registration
Volunteer #1	3.68 ± 3.16	1.3 ± 1.2
Volunteer #2	2.55 ± 2.87	1.25 ± 1.26
Volunteer #3	3.72 ± 3.92	0.8 ± 1.01

Table 7.2: Accuracy and precision of the EVolution method following the standoff B-mode US study on the three healthy volunteers. Evaluation was performed for an ROI at the lower bound of liver segment # 5, both in the presence (third column) and in the absence (second column) of the EVolution tracking method. The FEP in the absence of the EVolution algorithm was computed by setting (u_{est}, v_{est}) to 0 in Eq. 7.6. The errors are reported under the format *mean* ± *standard deviation* of the spatially averaged FEP pooled from all sampling time points.



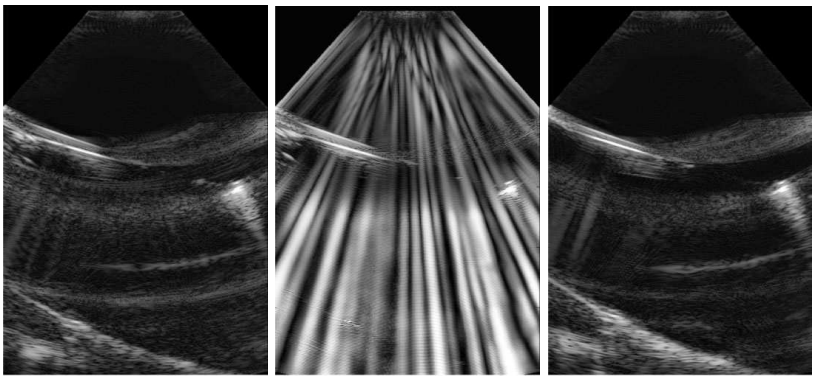
(a) (b)

Figure 7.6: Performance evaluation of the EVolution method for *in-vivo* standoff B-mode echography. (a): Spatial distribution of the temporally averaged FEP calculated for volunteer #3 and overlaid on the corresponding reference image of the standoff B-mode US image series. (b): The average liver displacement in the region-of-interest indicated by the yellow contour in (a), in the head-foot direction during the first 30 s of the study conducted on volunteer #3. The blue curve illustrates the estimation provided by the optical flow algorithm on the MR image series (established as silver standard during the current experiment), while the green curve plots the one provided by the EVolution algorithm on the simultaneously acquired US image series.

7.3.2.3 Efficiency of the proposed interference avoidance scheme

Sections 7.3.2.1 and 7.3.2.2 report the achievable accuracy and precision of the EVolution algorithm for standoff US-based tracking. However, both the experiments were conducted with the HIFU beam off. As pointed-out in section c, the HIFU therapeutic signal may oversaturate the signal received by the US imager and affect the quality of the US images and implicitly the performance of the tracking algorithm. Fig. 7.7 showcases such saturation effects in US images acquired during the experiment described in c, paragraph 3. Fig. 7.7(a) illustrates one of the US images acquired while the HIFU transducer is off. Fig. 7.7(b) displays an US image acquired during a 50 W sonication in absence of an interference avoidance scheme. Note the strong interference artifacts that are present in the image, obscuring the imaged phantom. Fig. 7.7(c) displays an US image acquired during a 300 W sonication, with the proposed synchronization scheme between the HIFU transducer and the US imager enabled. It can be observed that the saturation artifacts become visually unnoticeable.

Fig. 7.8(a) and 7.8(b) showcase a coronal and a sagittal MR-Thermometry



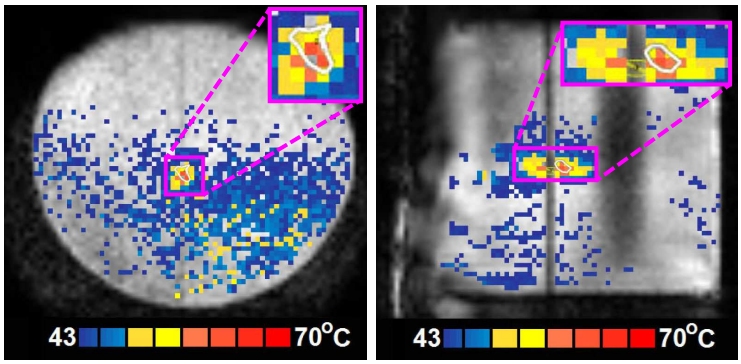
(a) (b) (c)

Figure 7.7: Efficiency of the proposed interference avoidance scheme between the HIFU therapeutic signal and the US imaging signal. (a) US image acquired while the HIFU transducer is off. (b): US image acquired during a 50 W HIFU sonication, without synchronization between the HIFU transducer and the US imager. (c): US image acquired during a 300 W HIFU sonication, with the interference avoidance scheme enabled.

image acquired on the moving phantom during the HIFU heating experiment described in section c, paragraph 3. The illustrated images were acquired at peak temperature, with the focal region magnified for better visibility. The white curve around the focal point indicates the region estimated to have received a lethal amount of thermal damage. This demonstrates that the proposed interference avoidance scheme does not hamper the system's ability for HIFU ablations. Moreover, no degradation in MR and US image quality was observed, despite the fact that the MR scanner and the US imager operated asynchronously during this experiment.

7.3.3 Algorithm sensitivity to the choice of input parameters

Fig. 7.9(a) and 7.9(b) illustrate the average TRE and FEP as a function of α and Γ for the contact and standoff echography studies, respectively. For details on how these figures were obtained, please see section 7.2.3. It can be observed that for contact echography, the maximum accuracy is attained for $(\alpha, \Gamma) = (0.3, 15 \times 15)$, while for standoff echography the optimal combination is $(\alpha, \Gamma) = (0.35, 3 \times 3)$. For both the contact and the standoff studies, the errors tend to be notably larger for small values of α and Γ, with a sudden drop at a value of $\sim (0.1, 5 \times 5) - (0.15, 7 \times 7)$. As the values of α and Γ increase towards their maximum, the registration errors also show a slight tendency to increase.



(a) (b)

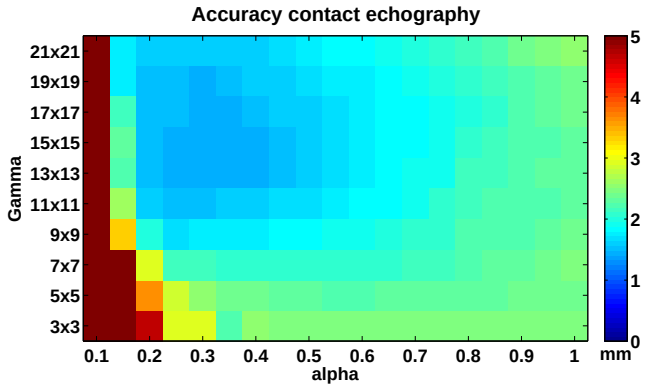
Figure 7.8: A (a): coronal (b): sagittal MR-Thermometry image acquired at peak temperature on the moving phantom during the heating experiment. The focal region was magnified for better visibility, with the white contour indicating the area estimated to have received a lethal amount of thermal damage.

7.3.4 Algorithm computational performance evaluation

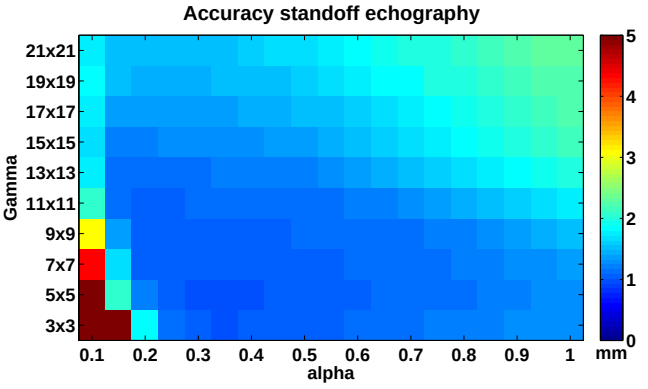
The GPU implementation of the EVolution algorithm converged on average in ∼50 ms. The convergence time was measured for images of size 256×256, which was the case for all data processed in the current study.

7.4 Discussion

The current study evaluates the real-time tracking capabilities of the recently proposed EVolution algorithm[35] for both contact and standoff B-mode echography. The tracking method was chosen due to its intrinsic robustness to transient image structures (such as speckles in US images), fast convergence and reduced number of input parameters, while at the same time providing displacements on a pixel-by-pixel basis. According to the classification made in,[23] the EVolution algorithm is a hybrid registration method. However, compared to existing hybrid methods for B-mode US registration, the EVolution algorithm has low computational demands, requires the configuration of a small number of input parameters, with no user interaction being necessary. Also, it does not require a learning step prior to its application, therefore its performance does not depend on the reproducibility of the estimated motion.
The evaluated tracking method was validated in three independent experiments, which were chosen as representative mock-ups of US-based motion guidance of HIFU therapies. The experiments encompassed an *in-vivo* contact echography study conducted separately on the liver and kidney of 7 healthy volunteers, a standoff echography study conducted on a phantom undergoing



(a)



(b)

Figure 7.9: Sensitivity of the EVolution algorithm to input parameters α and Γ. Average (a): TRE and (b): FEP as a function of α and Γ for the contact and standoff echography studies, respectively.

a known motion pattern and an *in-vivo* standoff echography experiment performed on the liver of three healthy volunteers. The low signal-to-noise (SNR) ratio, artifacts, poor contrast and orientation-dependent appearance of B-mode US images make tracking via gradient-based methods a challenging task.[54] It should also be noted that, in the scope of this study, with respect to the choice of the images, the employed equipment and acquisition schemes, no particular selection with respect to image quality has been made. Despite all this, the EVolution algorithm, which is gradient-based, provided in the majority of the analyzed cases, for both contact and standoff echography, motion estimates with an accuracy and precision higher than 1.6 mm, which is on par with state-of-the-art algorithms dedicated to B-mode US image registration.[23,55] However,

while comparable, there is still a particular amount of inter-experiment and inter-subject variability in terms of tracking accuracy and precision. The cause to this is most likely variations in image quality (i.e. SNR, artifacts, etc.) between the different US image series and the US images themselves. Moreover, during the *in-vivo* studies, due to the limited acoustic window of the employed B-mode US imager, parts of the organs-of-interest would periodically enter and leave the field-of-view (FOV). In the scope of this chapter, the tracked landmarks/ROIs were purposefully maintained in the FOV during the acquisition process. However, this aspect requires special attention when tracking is performed in a clinical setting, such that the region-of-interest does not leave the FOV. Also, note that the evaluation in the standoff echography study was limited to the lower bound of liver segment # 5. The rather large distance between the US probe and the organ-of-interest led to a poor signal in regions towards the interior of the liver. Due to a lack of visible structures (such as organ boundaries and/or blood vessels), the algorithm performed poorly in these areas. Therefore, future work should should focus on improving the quality of the standoff US images acquired on the hybrid MR-US HIFU system.

One of the challenges that occurred during the validation of the proposed tracking method in the *in-vivo* scenarios was the choice of a reliable ground truth. For the *in-vivo* contact echography study the gold standard was determined by the means of manual tracking of identifiable anatomical landmarks. While a reliable approach, it is still sensitive to the observer's subjectivity, especially given the time-varying visibility of the structures in the US images over the respiratory cycle. For the standoff case, since a time-synchronized MR-image series was also acquired, a natural choice seemed the deformation fields estimated by the optical flow algorithm[43] on the MR-images. This was motivated by the fact that the optical flow algorithm was previously validated for MR-based motion tracking by several independent studies in both *ex-vivo* and *in-vivo* scenarios.[56–59] Furthermore, the good MR-based tracking capabilities of the optical flow algorithm was further confirmed in the current study during the phantom tracking experiment (see section 7.3.2.1). Another aspect that needs to be taken into consideration when evaluating the *in-vivo* tracking errors of the proposed method is that the imaging (and implicitly tracking) was performed in 2D. While the imaging planes were optimized such that they are in good correspondence with the principal axis of respiratory displacements, out-of-plane motion still remains a possibility. In case the latter becomes significant, its effects on the image content may be interpreted by the algorithm as in-plane motion, leading to estimation errors. In such instances 3D US-tracking may be necessary (which has not been investigated in the scope of this study) or a method that includes through-plane motion detection capabilities could be employed, as described in.[60]

The algorithm performance may also depend on the local validity of the smoothness assumption made by the second term of the functional in Eq. 7.1. For abdominal organs in particular, such a constraint is generally justified, since their composing tissues are incompressible and elastic. Thus, from a physical

point-of-view, the displacement of the organs under discussion is indeed spatially smooth. Problematic are, however, the interfaces between the moving organs and the quasi-static abdominal/thoracic wall, where sliding/shearing motion may occur. Since this implies a local violation of the smoothness assumption, such areas are prone to motion estimation errors. More specifically, the displacements estimated for the mobile structures will get spatially propagated towards the static ones. Such a phenomenon was observed, for example, in the results associated to the *in-vivo* data displayed in Fig. 7.5. Due to the propagation of the displacements estimated for the liver, false motion was estimated for the quasi-static peritoneal wall. Area-wise, this effect extended ∼2-3 pixels into the wall. A similar scenario can arise due to ghost image (aliasing) artifacts caused by acoustic multiple reflection. The latter can lead to some objects/structures appearing at several locations in the US image: once at their true spatial location and the rest, due to the additional reflections, at false locations, overlapping with the real structures that are situated at those particular sites. Such an effect is visible, for example, in Fig. 7.3(b), where a ghost of the HIFU transducer membrane partially overlaps with the tracked structure. The issue is that, in general, the ghost image of an object may or may not move relative to the real structure that it overlaps with. Thus, a smoothness assumption on the estimated displacements may not be valid. In the aforementioned example from Fig. 7.3(b), the reflection of the transducer membrane was static relative to the tracked structure. The constraint to provide a smooth deformation in the overlap region between the static membrane reflection and the moving structure, lead to a local decrease in the magnitude of the estimated motion vectors, especially for large displacements. This phenomenon may explain the systematic bias at the negative peaks of the green curve in Fig. 7.4. While in the scope of this study, neither sliding motion nor ghosting artifacts notably hampered the estimated displacements, as a source of mis-registration, they should be carefully monitored when tracking is performed via the EVolution method.

A feature that makes the EVolution algorithm an attractive tracking solution is that it requires the input of only two parameters: α and Γ (see Eq. 7.1). Moreover, as shown in Fig. 7.9, registration errors remain close to their global minimum for a wide range of α and Γ combinations. This is the case especially for the standoff experiments. Note, however, that the accuracy rapidly decreases as α and Γ approach 0.1 and 3×3, respectively. On the other hand, the impact on accuracy is less important as α and Γ increase towards their maximum. Therefore, in practice it is favorable to overestimate the values of α and Γ.

An important factor for motion tracking in general is the computational time required by the employed algorithm. In particular for interventional guidance of mini- or non-invasive procedures, the real-time performance of such an algorithm is paramount. With respect to this, two fundamental aspects have to be considered and differentiated: latency and throughput. In the scope of this study, latency is defined as the computational time required by the algorithm

to register one pair of images. Registration methods that provide a dense deformation field between two images, are known to have high computational demands, making an implementation with real-time capabilities a challenging task.[23] However, an implementation with short associated latencies for the EVolution algorithm is facilitated by the method's pixel-wise numerical scheme, which can be massively parallelized. As stated in section 7.3.4 this resulted in an average latency of 50 ms per registration, for images of size 256×256. For respiratory-related HIFU beam steering, such latencies were established to be acceptable.[10,15] It is, however, important to keep in mind that this latency may vary, depending on the extent of the deformation between the images, image size and the computational capabilities of the available hardware. Image throughput, on the other hand, refers to the number of images per unit time that a system is capable of processing. The implementation performed in the current study allowed for an image throughput of ~ 20 images per second, which is suitable for most state-of-the-art B-mode US imagers. The throughput can, however, be increased if necessary. For example, if two or more graphic cards of the same model as used in the present chapter are included in the system, the image throughput can be scaled and adapted to imaging systems with high frame-rates (such as, for example, plane-wave US imagers[61]).
From a more practical side, the current chapter also proposes a scheme for addressing the interference between the US imaging and the HIFU signals. As illustrated in Fig. 7.7(b), if no interference avoidance measures are taken, a 50 W sonication may already lead to saturation artifacts that render the US images unusable for tracking. The proposed solution consists in a temporal multiplexing of the US imager and the HIFU transducer. As shown in Fig. 7.7(c), this removes the interference artifacts to an extent that they are no longer observable. A concern that arises with such a technique is that, since the HIFU transducer is turned off every time an US image is acquired, the thermal damage may be insufficient to have a therapeutic effect. This was, however, not the case in the current study since it was demonstrated that lethal thermal damage can still be achieved in the focal area (see Fig. 7.8). Nevertheless, such an approach, while efficient, it increases the time required to achieve this effect. This aspect can be addressed by employing more technically advanced solutions, which allow the continuous operation of the HIFU transducer and US imager. The principle behind such approaches involves encoding and/or filtering the composite signal received by the US imager, such that the HIFU therapeutic signal is strongly attenuated prior to image reconstruction.[48–52] Such techniques improve the duty cycle of the therapeutic energy delivery, leading to shorter sonication/therapy times compared to a temporal multiplexing solution. Nevertheless, in the scope of this chapter, temporal synchronization between the two systems was deemed to be an acceptable solution, with more efficient approaches for interference avoidance being the topic of future studies.
The main focus of the current study was motion estimation. However, in particular for the hybrid MR-USg-HIFU system, a solution for motion compensation with respect to the estimated displacements would also be of interest. Motion

compensation can be achieved, for example, by delivering the therapeutic energy in a gated manner. In effect, the HIFU transducer would be activated only when the targeted tissue is inside a pre-defined range of locations. Whether the tissue is situated in this range would be provided by the displacements estimated by the EVolution algorithm on the US images. Another solution for motion compensation would be electronic steering of the HIFU beam. In this sense, an approach similar to the one described by Ries *et al*[10] could be used, with steering being performed according to the motion vectors estimated by EVolution on the US image series.

7.5 Conclusion

A solution is proposed for target tracking in B-mode US sequences, capable of providing real-time guidance for HIFU therapies in mobile/deformable organs. The method is based on the recently proposed EVolution registration algorithm, a dynamic contour tracking method mathematically formulated in a variational setting. EVolution is a hybrid registration method which, as the current study demonstrates, showcases the advantages provided by both intensity-based and feature-based methods. It was also shown to be robust against their individual drawbacks. Similar to intensity-based registration methods, EVolution estimates a dense deformation between two images and does not require the selection/detection of image features. However, while intensity-based methods providing a dense deformation field usually imply high computational demands, in the current study it was shown that the EVolution algorithm has an image throughput and latency which are compatible with real-time B-mode US-guidance. Moreover, EVolution was shown to be capable of estimating rather large motion amplitudes of more than 10 mm and, by construction, is robust against transient structures, qualities which are specific to feature-based registration methods. However, since EVolution does not involve the selection/detection of a set of features, the registration process is not severely hampered by a low number of distinctive features in the images and/or the loss of features from one image to the next due to noise, image artifacts or through plane motion. In addition, even though B-mode US images are subject to low SNR, artifacts, poor contrast, and orientation-dependent appearance, which renders the usage of gradient-based registration methods challenging, results have demonstrated that the EVolution algorithm is resilient against such effects, having on average an accuracy of $\sim$1.5 mm and submillimeter precision.

Acknowledgements

This work was supported by the Dutch Technology Foundation (STW) (project OnTrack no. 12813) and by the Seventh Framework Programme, grant no. 603028 (iPaCT).

Bibliography

[1] K. F. Chu and D. E. Dupuy. Thermal ablation of tumours: biological mechanisms and advances in therapy. *Nat Rev Cancer*, 14(3):199–208, 2014.

[2] A. Chapman and G. ter Haar. Thermal ablation of uterine fibroids using MR-guided focused ultrasound - A truly non-invasive treatment modality. *Eur Raiol*, 17(10):2505–2511, 2007.

[3] A. Blana, S. Rogenhofer, R. Ganzer, J. C. Lunz, M. Schostak, W. F. Wieland, and B. Walter. Eight years' experience with high-intensity focused ultrasonography for treatment of localized prostate cancer. *Urology*, 72(6):1329–1333, 2008.

[4] A. Okada, T. Murakami, K. Mikami, H. Onishi, N. Tanigawa, T. Marukawa, and H. Nakamura. A case of hepatocellular carcinoma treated by MR-guided focused ultrasound ablation with respiratory gating. *Magn Reson Med Sci*, 5(3):167–171, 2006.

[5] Y. Y. Li, W. H. Sha, Y. J. Zhou, and Y. Q. Nie. Short and long term efficacy of high intensity focused ultrasound therapy for advanced hepatocellular carcinoma. *J Gastroenterol Hepatol*, 22(12):2148–2154, 2007.

[6] R. D. Kovatcheva, J. D. Vlahov, J. I. Stoinov, and K. Zaletel. Benign solid thyroid nodules: US-guided high-intensity focused ultrasound ablationinitial clinical outcomes. *Radiology*, 276(2):597–605, 2015a.

[7] B. Liberman, D. Gianfelice, Y. Inbar, A. Beck, T. Rabin, N. Shabshin, G. Chander, S. Hengst, R. Pfeffer, A. Chechick, A. Hanannel, O. Dogadkin, and R. Catane. Pain palliation in patients with bone metastases using MR-guided focused ultrasound surgery: a multicenter study. *Ann Surg Oncol*, 16(1):140–146, 2009.

[8] R. D. Kovatcheva, J.-N. Guglielmina, M. Abeshera, L. Boulanger, N. Laurent, and E. Poncelet. Ultrasound-guided high-intensity focused ultrasound

treatment of breast fibroadenoma-A multicenter experience. *J Ther Ultrasound*, 3(1), 2015b.

[9] M. Zaaroor, A. Sinai, D. Goldsher, A. Eran, M. Nassar, and I. Schlesinger. Magnetic resonance-guided focused ultrasound thalamotomy for tremor: a report of 30 parkinsons disease and essential tremor cases. *J Neurosurg*, 24:1–9, 2017.

[10] M. Ries, B. Denis de Senneville, S. Roujol, Y. Berber, B. Quesson, and C.T.W. Moonen. Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues. *Magnetic Resonance in Medicine*, 64(6):1704–1712, 2010.

[11] V. Auboiroux, L. Petrusca, M. Viallon, T. Goget, C. D. Becker, and R. Salomir. Ultrasonography-based 2D motion-compensated HIFU sonication integrated with reference-free MR temperature monitoring: A feasibility study ex vivo. *Phys Med Biol*, 57(10):N159–71, 2012.

[12] A. B. Holbrook, P. Ghanouni, J. M. Santos, C. Dumoulin, Y. Medan, and K. B. Pauly. Respiration based steering for high intensity focused ultrasound liver ablation. *Magn Reson Med*, 71(2):797–806, 2014.

[13] V. Rieke and K. B. Pauly. MR thermometry. *J Magn Res Imaging*, 27(2):376 – 390, 2008.

[14] B. Denis de Senneville, C. Mougenot, and C. T. Moonen. Real - Time Adaptive Methods for Treatment of Mobile Organs by MRI - Controlled High - Intensity Focused Ultrasound. *Magn Reson Med*, 57(2):319 – 330, 2007.

[15] S. Roujol, M. Ries, B. Quesson, C. Moonen, and B. Denis de Senneville. Real - Time MR - Thermometry and Dosimetry for Interventional Guidance on Abdominal Organs. *Magn Reson Med*, 63(4):1080 – 1087, 2010.

[16] B. Denis de Senneville, A. El Hamidi, and C. Moonen. A Direct PCA - Based Approach for Real - Time Description of Physiological Organ Deformations. *IEEE Trans Med Imaging*, 34(4):974 – 982, 2015.

[17] V.R.S. Mani and S. Arivazhagan. Survey of medical image registration. *Journal of Biomedical Engineering and Technology*, 1(2):8–25, 2013.

[18] J. Seo, N. Koizumi, T. Funamoto, N. Sugita, K. Yoshinaka, A. Nomiya, Y. Homma, Y. Matsumoto, and M. Mitsuishi. Visual servoing for a us-guided therapeutic HIFU system by coagulated lesion tracking: A phantom study. *Int J Med Robot*, 7(2):237 – 247, 2011a.

[19] J. Seo, N. Koizumi, T. Funamoto, N. Sugita, K. Yoshinaka, A. Nomiya, Y. Homma, Y. Matsumoto, and M. Mitsuishi. Biplane US-guided real-time volumetric target pose estimation method for theragnostic HIFU system. *Journal of Robotics and Mechatronics*, 23(3):400–407, 2011b.

[20] J. Seo, N. Koizumi, M. Mithuishi, and N. Sugita. Ultrasound image-based visual servoing for moving target ablation by high intensity focused ultrasound. *Int J Med Robot*, pages e1793–n/a, 2016.

[21] L. Petrusca, P. Cattin, V. De Luca, F. Preiswerk, Z. Celicanin, M. Viallon, P. Arnold, F. Santini, S. Terraz, K. Scheffler, C. D. Becker, and R. Salomir. Hybrid ultrasound/magnetic resonance simultaneous acuisition and image fusion for motion monitoring in the upper abdomen. *Invest Radiol*, 48(5):333 – 340, 2013.

[22] F. Preiswerk, V. De Luca, P. Arnold, Z. Celicanin, L. Petrusca, C. Tanner, O. Bieri, R. Salomir, and P. C. Cattin. Model-guided respiratory organ motion prediction of the liver from 2D ultrasound. *Med Image Anal*, 18(5):740–751, 2014.

[23] V. De Luca, G. Székely, and C. Tanner. Estimation of large-scale organ motion in B-mode ultrasound image sequences: A survey. *Ultrasound Med Biol*, 41(12):3044–3062, 2015b.

[24] A. Cifor, L. Risser, D. Chung, E. M. Anderson, and J. A. Schnabel. Hybrid feature-based diffeomorphic registration for tumor tracking in 2-d liver ultrasound images. *IEEE Trans Med Imaging*, 32(9):1647 – 1656, 2013.

[25] J. P. Thirion. Image matching as a diffusion process: An analogy with Maxwell's demons. *Med Image Anal*, 2(3):243 – 260, 1998.

[26] J. Woo, B. W. Hong, C. H. Hu, K. K. Shung, C. C. Jay Kuo, and P. J. Slomka. Non-rigid ultrasound image registration based on intensity and local phase information. *J Sign Process Sys*, 54(1):33 – 43, 2009.

[27] P. Kovesi. Phase congruency: A low-level image invariant. *Psychol Res*, 64(2):136 – 148, 2000.

[28] L. R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297 – 302, 1945.

[29] M. Li, C. Kambhamettu, and M. Stone. Automatic contour tracking in ultrasound images. *Clin Linguist Phon*, 19(6 - 7):545 – 554, 2005.

[30] Y. Wang, B. Georgescu, D. Comaniciu, and H. Houle. Learning-based 3D myocardial motion flow estimation using high frame rate volumetric ultrasound data. In *Int Symp Biomed Imaging*, pages 1097 – 1100, Piscataway, NJ, USA, 2010. IEEE.

[31] X. Huang, D. P. Dione, C. B. Compas, X. Papademetris, B. A. Lin, A. Bregasi, A. J. Sinusas, L. H. Staib, and J. S. Duncan. Contour tracking in echocardiographic sequences via sparse representation and dictionary learning. *Med Image Anal*, 18(2):253 – 271, 2014.

[32] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int J Comp Vis*, 1(4):321 – 331, 1988.

[33] D. P. Huttenlocher, G. A. Klanderman, and W. Rucklidge. Comparing images using the hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(9):850 – 863, 1993.

[34] J. Weickert, A. Bruhn, N. Papenberg, and T. Brox. Variational optic flow computation: From continuous models to algorithms. In *International Workshop on Computer Vision and Image Analysis (ed. L. Alvarez), IWCVIA03, Las Palmas de Gran Canaria*, 2003.

[35] B. Denis de Senneville, C. Zachiu, M. Ries, and C. Moonen. EVolution: an edge-based variational method for non-rigid multi-modal image registration. *Phys Med Biol*, 61(20):7377–7396, 2016.

[36] L. König, T. Kipshagen, and J. Rühaak. A non-linear image registration scheme for real-time liver ultrasound tracking using normalized gradient fields. In *MICCAI Challenge on Liver Ultrasound tracking, Boston, USA*, 2014.

[37] Y. Kubota, A. Matsumura, M. Fukahori, S. Minohara, S. Yasuda, and H. Nagahashi. A new method for tracking organ motion on diagnostic ultrasound images. *Med Phys*, 41(9):092901, 2014.

[38] B.K.P. Horn and B.G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[39] C.R.J. Maurer, J. Fitzpatrick, M. Wang, R.L.J. Galloway, R. Maciunas, and G. Allen. Registration of head volume images using implantable fiducial markers. *IEEE Transactions on Medical Imaging*, 16:447–462, 1997.

[40] P. Ramaekers, M. de Greef, R. Berriet, C. T. W. Moonen, and M. Ries. Evaluation of a novel therapeutic focused ultrasound transducer based on Fermat's spiral. *Phys Med Biol*, 62(12):5021–5045, 2017.

[41] P. J. Kaczkowski and R. E. Daigle. The Verasonics ultrasound system as a pedagigic tool in teaching wave propagation, scattering, beamforming, and signal processing concepts in physics and engineering. *J Acoust Soc Am*, 129(4):2648, 2011.

[42] Verasonics. The vantage research ultrasound systems. `http://verasonics.com/vantage-systems/`, 2017. Accessed: 2017-10-02.

[43] C. Zachiu, N. Papadakis, M. Ries, C. T. W. Moonen, and B. Denis de Senneville. An improved optical flow tracking technique for real-time MR-guided beam therapies in moving organs. *Physics in Medicine and Biology*, 60(23):9003, 2015.

[44] S. Nikolov and J. A. Jensen. Virtual ultrasound sources in high-resolution ultrasound imaging. In *Medical Imaging: Ultrasonic Imaging and Signal Procesing*, volume 4687, International Society for Optics and Photonics, pages 395–405, 2002.

[45] C. H. Frazier and W. R. O'Brien. Synthetic aperture techniques with a virtual source element. *IEEE Trans Ultrason Ferroelectr Freq Control*, 45(1):196 – 207, 1998.

[46] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski. A database and evaluation methodology for optical flow. *International Journal of Computer Vision*, 92(1):1–31, 2011.

[47] S. Vaezy, X. Shi, R. W. Martin, E. Chi, P. I. Nelson, M. R. Bailey, and L. A. Crum. Real-time visualization of high-intensity focused ultrasound treatment using ultrasound imaging. *Ultrasound Med Biol*, 27(1):33 – 42, 2001.

[48] J. S. Jeong, J. H. Chang, and K. K. Shung. Ultrasound transducer and system for real-time simultaneous therapy and diagnosis for noninvasive surgery of prostate tissue. *IEEE Trans Ultrason Ferroelectr Freq Control*, 56(9):1913 – 1922, 2009.

[49] J. S. Jeong, J. M. Cannata, and K. K. Shung. Adaptive HIFU noise cancellation for simultaneous therapy and imaging using an integrated HIFU/imaging transducer. *Phys Med Biol*, 55(7):1889 – 1902, 2010.

[50] J. S. Jeong, J. H. Chang, and K. K. Shung. Pulse compression technique for simultaneous HIFU surgery and ultrasonic imaging: A preliminary study. *Ultrasonics*, 52(6):730 – 739, 2012.

[51] J. S. Jeong, Y. Yoo, T. K. Song, and J. H. Chang. Real-time monitoring of HIFU treatment using pulse inversion. *Phys Med Biol*, 58(15):5333 – 5350, 2013.

[52] R. Takagi, H. Jimbo, R. Iwasaki, K. Tomiyasu, S. Yoshizawa, and S.-I. Umemura. Feasibility of real-time treatment feedback using novel filter for eliminating therapeutic ultrasound noise with high-speed ultrasonic imaging in ultrasound-guided high-intensity focused ultrasound treatment. *Jap J App Phys*, 55(7S1):07KC10: 1 – 8, 2016.

[53] S. A. Sapareto and W. C. Dewey. Thermal dose determination in cancer therapy. *Int J Radiat Oncol Biol Phys*, 10(6):787 – 800, 1984.

[54] G. Jacob, J. A. Noble, M. Mulet-Parada, and A. Blake. Evaluating a robust contour tracker on echocardiographic sequences. *Med Image Analisys*, 3(1):63–75, 1999.

[55] V. De Luca, T. Benz, S. Konodo, L. König, D. Lübke, S. Rothlübbers, O. Somphone, S. Allaire, M. A. Lediju Bell, D. Y. Chung, A. Cifor, C. Grozea, M. Günther, J. Jenne, T. Kipshagen, M. Kowarschik, N. Navab, J. Rühaak, J Schwaab, and C. Tanner. The 2014 liver ultrasound tracking benchmark. *Phys Med Biol*, 60(14):5571–5599, 2015a.

[56] N. K. Ostergaard, B. Denis de Senneville, U. V. Elstrom, K. Tanderup, and T. S. Sorensen. Acceleration and validation of optical flow based deformable registration for image-guided radiotherapy. *Acta Oncol.*, 47(7):1286 – 1293, 2008.

[57] R. Castillo, E. Castillo, R. Guerra, V. E. Johnson, T. McPhail, A.K. Garg, and T. Guerrero. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Phys. Med. Biol*, 54:1849–1870, 2009.

[58] K.K. Brock. Results of a multi-institution deformable registration accuracy study(MIDRAS). *Int J Radiation Oncology Biol. Phys.*, 76(2):583–596, 2010.

[59] N. Kadoya, Y. Fujita, Y. Katsuta, S. Dobashi, K. Takeda, K. Kishi, M. Kubozono, R. Umezawa, T. Sugawara, H. Matsushita, and K. Jingu. Evaluation of various deformable image registration algorithms for thoracic images. *J Radiat Res*, 55(1):175–182, 2014.

[60] C. N. Ta, M. Eghtedari, R. F. Mattrey, Y. Kono, and A. C. Kummel. 2-tier in-plane motion correction and out-of-plane motion filtering for contrast-enhanced ultrasound. *Invest Radiol*, 49(11):707–719, 2014.

[61] G. Montaldo, M. Tanter, J. Bercoff, N. Benech, and M. Flink. Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography. *IEEE Trans Ultrason Ferroelectr Freq Control*, 53(3):489–506, 2009.

CHAPTER 8

Non-rigid CT/CBCT to CBCT registration for online external beam radiotherapy guidance

"Force has no place where there is need of skill."
- Herodotus

This chapter is based on:
C. Zachiu, B. Denis de Senneville, R.H.N. Tijssen, A.N.T.J. Kotte, A.C. Houweling, L.G.W. Kerkmeijer, J.J.W. Lagendijk, C.T.W. Moonen, and M. Ries. Non-rigid CT/CBCT to CBCT registration for online external beam radiotherapy guidance. *Phys. Med. Biol.*, 63(1): 015027, 2017.

Abstract

Image-guided external beam radiotherapy (EBRT) allows radiation dose deposition with a high degree of accuracy and precision. Guidance is usually achieved by estimating the displacements, via image registration, between cone beam computed tomography (CBCT) and computed tomography (CT) images acquired at different stages of the therapy. The resulting displacements are then used to reposition the patient such that the location of the tumor at the time of treatment matches its position during planning. Moreover, ongoing research aims to use CBCT-CT image registration for online plan adaptation.
However, CBCT images are usually acquired using a small number of X-Ray projections and/or low beam intensities. This often leads to the images being subject to low contrast, low signal-to-noise ratio and artifacts, which ends-up hampering the image registration process. Previous studies addressed this by integrating additional image processing steps into the registration procedure. However, these steps are usually designed for particular image acquisition schemes, therefore limiting their use on a case-by-case basis.
In the current study we address CT to CBCT and CBCT to CBCT registration by the means of the recently proposed EVolution registration algorithm. Contrary to previous approaches, EVolution does not require the integration of additional image processing steps in the registration scheme. Moreover, the algorithm requires a low number of input parameters, is easily parallelizable and provides an elastic deformation on a point-by-point basis. Results have shown that relative to a pure CT-based registration, the intrinsic artifacts present in typical CBCT images only have a sub-millimeter impact on the accuracy and precision of the estimated deformation. In addition, the algorithm has low computational requirements, which are compatible with online image-based guidance of EBRT treatments.

8.1 Introduction

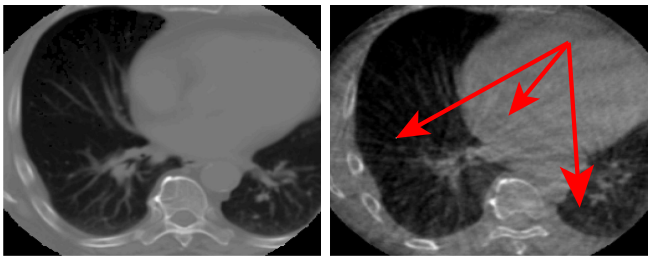
Image-guided external beam radiotherapy (EBRT) provides the possibility of radiation cancer treatment delivery with a high degree of geometric precision and accuracy.[1] This is facilitated by the capability of the current treatment machines to provide onboard 3D imaging of the target pathology and its surroundings. Due to its cost effectiveness and typically low radiation dose to the patient, cone beam computed tomography (CBCT) has been integrated for therapy guidance purposes, in most state-of-the-art therapeutic linear accelerator (linac) solutions.[2] During EBRT treatments, a CBCT image is usually acquired at the start of a radiation fraction and compared to a high-resolution

computed tomography (CT) image acquired during the planning phase of the treatment.[3] The outcome of this comparison is a set of displacements indicating the anatomical situation at the time of treatment relative to therapy planning. These displacements are currently used to reposition the patient such that the pathology at the time of treatment is approximately at the same location as during planning.[3–5] Also, in case of high fraction doses or long treatment times, an additional CBCT image is acquired after the radiation fraction delivery, in order to determine whether significant motion occurred during treatment delivery.

However, depending on the imaged anatomy and rotation speed of the linac gantry, the acquisition of both CT and CBCT images may be hampered by physiological motion.[6,7] Respiratory motion, in particular, can induce large displacements of the organs situated in the thorax and upper abdomen. If not taken into consideration, this can lead to motion-induced blurring effects in the acquired images, increasing the uncertainties related to the location of the pathology and organs-at-risk.[6,8] In this sense, a common practice is to continuously acquire X-ray projections over a longer period of time, as the linac gantry rotates around the patient, and reorder them according to the respiratory phase in which they were acquired.[9–11] The reordered X-ray projections are then used to reconstruct a 4D (3D + time) image series, providing the configuration of the imaged anatomy over the respiratory cycle. Estimation of respiratory displacements based on a 4D CT/CBCT series could be useful, for example, for providing patient-specific motion models which, in turn, can be used to adapt treatment margins on a case-by-case basis.[12–14] This usually results in a more conformal radiation dose for the target pathology while at the same time improving the sparing of healthy tissues.

The displacements between CT and CBCT images acquired during the different phases of EBRT can be estimated using a process called image registration.[15] The output of the registration process between two images is a transformation that allows mapping one of the images, called the moving image, to the other image, called reference, leading to a spatial consistency between the two images. A different number of degrees-of-freedom can be imposed on the estimated transformation, which is usually chosen depending on the mobility of the imaged anatomical site. A rigid transformation, which has a low number of degrees-of-freedom, is for example, more suitable for static anatomies such as the brain or parts of the skeletal system.[16] For mobile and/or deformable organs, on the other hand, a rigid transformation may be sub-optimal. Thus, the number of degrees-of-freedom of the estimated transformation may have to be increased in order to capture the complex deformation underwent by these organs. Regardless of the nature of the estimated transformation, however, registration of CBCT images is a challenging task, especially for registration methods which rely on image intensity. This is due to the fact that, in order to reduce the radiation dose related to imaging and the required acquisition time, CBCT images are acquired/reconstructed using a low number of X-ray projections and/or low beam intensities. This, in turn, leads to the images con-

taining so-called streaking artifacts, while at the same time being affected by a low contrast and a low signal-to-noise ratio (SNR) (see Fig. 8.1).[17,18] Streak-



(a) (b)

Figure 8.1: Streaking artifacts in CBCT images: A transverse slice selected from a (a) CT (b) CBCT image acquired on a lung cancer patient. Compared to the CT image, the CBCT is visibly altered by streaking artifacts (indicated by the red arrows), which obstruct/degrade finer details of the anatomy and introduce intensity variations.

ing artifacts become especially problematic when CT-CBCT intensity-based registration is of interest, due to the intrinsic intensity inconsistencies they introduce between the CT and the CBCT image.[19] Previous studies address this issue by incorporating intensity correction/matching procedures into existing intensity-based registration algorithms such as demon registration[20–22] or optical flow.[19,23] The efficiency of such approaches, however, depends on the specific particularities of the artifacts hampering the CBCT images on a case-by-case basis. Even though CT and CBCT images are acquired using the same modality, previous studies have also addressed CT-CBCT registration using methods which are typically used for aligning images acquired with different modalities. Due to their functioning principle, such methods relax the assumptions made on image intensity and align the images based on other criteria such as mutual information (MI) or normalized cross correlation (NCC).[24–26] However, both MI and NCC are evaluated on a global level. Therefore, while such approaches may perform well for estimating rigid displacements, they may become suboptimal for elastic deformations. Moreover, such methods typically imply high computational demands and optimization of several input parameters, limiting their use in a clinical setting with the patient still on the interventional table. A different approach to CT-CBCT registration involves aligning the images based on a set of common features/landmarks. It is, for example, current clinical practice that several CT/CBCT-visible fiducial markers are implanted inside the tumor and its proximity.[27–29] The markers are then either automatically or manually identified and re-aligned in order to estimate a displacement between the images. Besides the implantation procedure involving discomfort for the patient (since in some cases it is performed without anesthesia), this

only allows a rigid alignment between the images. Other methods which rely on naturally occurring anatomical markers[30,31] are sensitive to the number and the discriminative properties of the landmarks. The number of landmarks also limits the degrees-of-freedom of the estimated transformation. In addition, both implanted and natural markers are only surrogates for the tumor position. There is a possibility that the tumor shifts relative to the markers, leading to misalignments between the images.
In the current study, the issue of both CT-CBCT and CBCT-CBCT registration is addressed via the EVolution registration algorithm, recently proposed by Denis de Senneville *et al* in.[32] The EVolution method estimates a deformation between two images, by locally matching similar contrast patterns. Thus, contrary to intensity-based methods, the algorithm is by construction robust to local and/or global intensity changes from one image to the other. Also, it does not require the implantation, detection and/or identification of a set of features/markers. At the same time it is capable of providing an elastic transformation between the registered images on a voxel-by-voxel basis. The method requires a small number of input parameters and implies a fast numerical scheme rendering it as an attractive solution for online therapy guidance.
The contribution of the present chapter is fourfold:

1. A validation procedure is proposed, which allows quantifying the impact of the artifacts and low SNR of the CBCT images on the motion estimates, relative to a high-resolution CT-based alignment.

2. The performance of the EVolution algorithm is evaluated qualitatively and quantitatively for CT to CBCT registration.

3. The algorithm's performance for CBCT-based 3D respiratory motion estimation was investigated.

4. An implementation with low computational demands is performed for the EVolution method, making it suitable for use in a clinical setting for online therapy guidance.

8.2 Method description

8.2.1 The EVolution registration algorithm

8.2.1.1 Functioning principle

EVolution is a variational multi-modal registration algorithm,[33] capable of estimating a non-rigid deformation between two images. It was recently proposed by Denis de Senneville *et al*[32] and it provides the displacement between two images as the optimizer of the following functional:

$$E(T) = \int_{\Omega} e^{-C(T)} + \frac{\alpha}{2} \left(\|\vec{\nabla} u\|_2^2 + \|\vec{\nabla} v\|_2^2 + \|\vec{\nabla} w\|_2^2 \right) d\vec{r} \qquad (8.1)$$

with

$$C(T) = \frac{\int_\Gamma \left| \vec{\nabla}_I(T(\vec{r})) \cdot \vec{\nabla}_J(\vec{r}) \right| d\vec{r}}{\int_\Gamma \|\vec{\nabla}_I(T(\vec{r}))\|_2 \|\vec{\nabla}_J(\vec{r})\|_2 d\vec{r}} \tag{8.2}$$

where $T = (u, v, w)$ is the 3D displacement, I and J are the two images to be aligned, Ω is the image domain, Γ is a cubic local neighborhood around the voxel being processed, $\vec{r}$ is a spatial location, $\vec{\nabla}$ is the 3D gradient operator, $\|\cdot\|_2$ is the Euclidean norm and α is an algorithm parameter whose purpose will be detailed shortly. Notice that the functional optimized by the EVolution algorithm (Eq. 8.1) is defined by two terms. The first term of the integral, also known as the data fidelity term (or simply data term), is a measure of the similarity between the images. The rationale behind using this term as a similarity metric becomes apparent if $C(T)$ is re-written under the following form:

$$C(T) = \frac{\int_\Gamma w_T(\vec{r}) \left|\cos(\Delta\theta_T(\vec{r}))\right| d\vec{r}}{\int_\Gamma w_T(\vec{r}) d\vec{r}} \tag{8.3}$$

with $w_T(\vec{r})$ and $\Delta\theta_T(\vec{r})$ given by:

$$\begin{aligned} w_T(\vec{r}) &= M_I(T(\vec{r}))M_J(\vec{r}) \\ \Delta\theta_T(\vec{r}) &= \theta_I(T(\vec{r})) - \theta_J(\vec{r}) \end{aligned} \tag{8.4}$$

where M_I and M_J are the magnitudes of the spatial gradients of the two images and $\Delta\theta$ is the phase difference between the gradient orientations. Therefore, the term $w_T(\vec{r})$ favors the alignment of edges/gradients that are present in both images, while $|\cos(\Delta\theta)|$ provides a measure for the angle between the gradient orientations, favoring parallel and anti-parallel gradients and implicitly similar local contrast patterns. The optimization of the data term alone is generally an ill-posed problem with an infinity of solutions. This was addressed by adding a regularization term to the functional in Eq. 8.1, which constrains the estimated deformation to be spatially smooth/differentiable. The amount of smoothness is controlled by the parameter α: the larger the value of α, the smoother the resulting deformation will be.

8.2.1.2 Numerical scheme

In this chapter, the functional in Eq. 8.1 was optimized using an approach similar to the one proposed in the seminal paper of Denis de Senneville *et al.*[32] In effect, its minimizer was found by solving the associated Euler - Lagrange equations. This ultimately implied finding the solution of a non-linear system. In order to reduce the non-linearity of the functional, a coarse-to-fine strategy was employed, which iterated the algorithm from a 32-fold downsampled version of the images step-by-step to the original resolution. The motion field estimated at each step was used as an initialization for the next higher resolution level. At each resolution level, the algorithm was considered to have converged when the average difference between the displacements estimated at the current and

the previous iteration was smaller than 10^{-3} voxels. For a more detailed and technical description of the numerical scheme, please consult.[32]

8.2.2 Experimental setup

In order to test the capability of the EVolution algorithm for CBCT-based motion estimation, several experiments were conducted:

1. The performance of the EVolution algorithm was initially evaluated for CT to CT registration and put in perspective with respect to a state-of-the-art method (namely the optical flow algorithm[34]). This evaluation was of interest due to the fact that CT to CT registration played an important role for quantifying the accuracy and precision of the EVolution method for both CT-CBCT and CBCT-CBCT registration.

2. A validation procedure was developed, based on CBCT images synthesized from high resolution CT images. Accuracy and precision of the method was then evaluated for CT - synthetic CBCT registration, with the synthesized CBCT containing an increasing amount of streaking artifacts.

3. A qualitative and quantitative evaluation of the algorithm was performed for registering CT to CBCT images, each acquired with their respective imaging hardware.

4. Finally, the algorithm's performance for 3D respiratory motion estimation based on synthesized CBCT images was investigated.

The following sections will provide details on each of the 4 experiments.

8.2.2.1 Validation of the EVolution algorithm for CT to CT registration

As it will be detailed in the following sections, CT to CT registration plays an important role, in the current chapter, for quantifying the accuracy and precision of the EVolution algorithm for CT to CBCT and CBCT to CBCT registration. For this reason, a preliminary experiment was conducted on CT data acquired on 5 head-and-neck cancer patients. Each data set contained a pair of CT images acquired at different time instants (more than 1 week apart). On each of the CT images, an experienced radiation therapist delineated both pathological areas and organs-at-risk (OAR). The 5 pairs of CT images were then registered using both EVolution and the optical flow algorithm[34] (used for comparison). For both registration methods, the Dice Similarity Coefficient (DSC),[35] normalized cross correlation (NCC)[36] and structural similarity index (SSIM)[37] were evaluated for several anatomical sites. The voxel-wise error in

flow endpoint (FEP) between the motion fields provided by EVolution and the optical flow methods was also evaluated:[38]

$$FEP(\vec{r}) = \|\mathbf{u_{EVo}}(\vec{r}) - \mathbf{u_{OF}}(\vec{r})\|_2 \tag{8.5}$$

where $\mathbf{u_{EVo}} = (u_{EVo}, v_{EVo}, w_{EVo})$ and $\mathbf{u_{OF}} = (u_{OF}, v_{OF}, w_{OF})$ are the 3D motion fields estimated by the EVolution and the optical flow algorithms, $\|\cdot\|_2$ is the Euclidean norm and $\vec{r}$ is the voxel position.
For computational purposes, the images and organ delineations were projected on a $256 \times 256 \times 256$ lattice with a voxel size of $1.5 \times 1.5 \times 1.5$ mm^3. The axes of the lattice were aligned with the axes of the DICOM patient coordinate system.[39]

8.2.2.2 Validation of the EVolution algorithm using synthetic CBCT images

a Gold standard selection In order to evaluate the precision and accuracy of the EVolution algorithm for CT to CBCT and CBCT to CBCT registration, the selection/establishment of a ground truth deformation was necessary. In general, obtaining an *in-vivo* gold standard is a challenging task, especially for deformable/elastic anatomies. In this chapter, a particular validation strategy was developed in this sense. In a first step, two high resolution CT images acquired at different time points are registered to one another using the EVolution algorithm, with the resulting displacements being established as gold standard. One or both of the images are then used to synthesize CBCT volumes (the synthesis procedure will be described in the next section) and then the images are registered to one another a second time. The resulting displacements are then compared to the gold standard. This process will provide the additional impact the undersampling of the sinogram has on CT-CBCT and CBCT-CBCT registrations performed using EVolution, compared to a registration based on high-resolution CT images.

b Synthesis of CBCT images from high-resolution CT volumes In order to synthesize a CBCT from a high resolution CT image, the freely available TIGRE CBCT reconstruction toolbox was employed.[40,41] The software allows a basic simulation of a CBCT acquisition and reconstruction, using any image as input, while at the same time providing access to the associated acquisition and reconstruction parameters. In particular, this allows control over the number of X-Ray projections used to reconstruct the CBCT volume. Therefore, the quality of the motion estimates as a function of the number of projections of can be investigated.

c Performance evaluation of the EVolution algorithm for CT to synthetic CBCT registration The capability of the EVolution algorithm for registering CT to synthesized CBCT images was evaluated on data acquired

on a head-and-neck cancer patient. The data initially consisted in a pair of CT images acquired on different occasions. In a first step towards the evaluation, the dependency of the estimation errors on the amount of the streaking artifacts present in the synthetic CBCT image was investigated. For this purpose, one of the CTs was used to synthesize a CBCT image with an increasing amount of streaking artifacts. The CBCT image was reconstructed using 360, 120, 72, 50 and 36 projections, which corresponds to an acquisition with an angular increment of 1°, 3°, 5°, 7°, and 10°, respectively. The other CT from the pair was then registered via the EVolution algorithm to each of the synthesized CBCT images and the resulting motion field was compared to the established gold standard (see sections 8.2.2.1 and a for details on the gold standard). The comparison between the gold standard and the estimated motion field was performed in terms of the spatial distribution of the voxel-wise error in flow endpoint (FEP):

$$FEP(\vec{r}) = \|\mathbf{u}(\vec{r}) - \mathbf{u_{gold}}(\vec{r})\|_2 \tag{8.6}$$

where $\mathbf{u} = (u, v, w)$ is the estimated 3D displacement, $\mathbf{u_{gold}} = (u_{gold}, v_{gold}, w_{gold})$ is the gold standard 3D displacement. Following this initial part of the experiment, the performance of the algorithm was investigated in more detail for a CBCT image synthesized using 72 projections. This case was of particular interest, since a reconstruction using 72 projections provides an image quality towards the lower end of what a typical CBCT system would provide.
For computational purposes the images were projected, prior to registration, on a $256 \times 256 \times 256$ lattice with a voxel size of $1.5 \times 1.5 \times 1.5$ mm^3. The axes of the lattice were aligned with the axes of the DICOM patient coordinate system.

8.2.2.3 CT to clinical CBCT image registration

This experiment aimed to evaluate the algorithm's capabilities for registering CT to clinical CBCT images. In effect, the algorithm was employed for the registration of a high resolution planning CT image to one of the daily CBCT images. This was carried-out for two pairs of CT - CBCT images: one acquired on a kidney cancer patient and the other on a lung cancer patient. The performance of the algorithm was assessed both by visual inspection and quantitative evaluation. The latter was achieved by manual annotation of 10 landmarks in both the kidney and the lung data sets. The resulting manually determined displacements were then compared in terms of the FEP, to the displacements estimated by the EVolution algorithm. As an additional registration error metric, the normalized cross correlation (NCC) between the images before and after registration was also calculated.
Before registration, all the images were projected on a $256 \times 256 \times 256$ lattice with a voxel size of $0.9 \times 0.9 \times 0.9$ mm^3. This was carried-out not only for computational purposes, but also, more important, due to the fact that the original CT and the CBCT images were reconstructed using different voxel sizes. Moreover, for both patients, part of the CT images was masked-out since

their field-of-view was considerably larger than the one of the CBCTs, which increases the risk of the algorithm converging towards a local minimum.

8.2.2.4 Synthetic CBCT-based respiratory motion estimation using EVolution

A 4D CT dataset was acquired on a renal cell carcinoma patient. The images sample the anatomy of the upper abdomen at 10 phases of the patient's respiratory cycle, between 0% and 90% of the cycle with a 10% increment. The 4D series was registered using the EVolution algorithm to a reference position provided by the CT image acquired at the 0% phase of the respiratory cycle. The resulting displacements were then stored and used as a gold standard. A 4D CBCT series was then synthesized based on the original CT series, and registered to the CBCT image corresponding to the 0% respiratory cycle phase. The resulting displacements were then compared to the aforementioned gold standard, in terms of the temporally averaged FEP.
Additional validation was performed by analyzing the CBCT-estimated 3D trajectory of two landmarks over the respiratory cycle. The first of the landmarks was one of the fiducial markers implanted in one of the kidneys, while the second marker corresponded to the centroid of the contralateral kidney. For the implanted marker, the ground truth was considered to be the trajectory provided by the EVolution algorithm on the original 4D CT series. However, for the centroid of the contralateral kidney, the trajectory estimated on the synthetic 4D CBCT series was compared to a manually determined trajectory. The estimated and the gold standard trajectories were again compared in terms of the FEP for the selected landmarks.
The 4D CBCT images were synthesized using 120 and 72 projections respectively, which produced images of average to low quality compared to what most clinical systems provide. Three scenarios were investigated: 1) All CBCT images were synthesized using 120 projections, 2) The reference image was reconstructed using 120 projections while the rest of the images were synthesized using 72 projections and 3) All CBCT images were reconstructed using 72 projections.
For computational purposes, the original 4D CT images were downsampled from $512 \times 512 \times 123$ to a size of $256 \times 256 \times 123$ with a $2 \times 2 \times 3$ mm^3 voxel size.

8.2.3 Hardware, implementation and algorithm configuration

The EVolution algorithm was implemented on a graphical processing unit (GPU) using the compute unified device architecture (CUDA) platform. The implementation was then executed on an nVidia Tesla K20 graphics card.
EVolution requires as input two parameters, namely α and Γ (see Eq. 8.1 and 8.2). A procedure was developed in the scope of this chapter in order to cali-

brate the two parameters. Three landmarks were manually tracked in the 4D CT series described in section 8.2.2.4. These landmarks consisted in two of the fiducial makers implanted at different locations in the ipsilateral kidney and the centroid of the contralateral kidney. The 4D series was registered to the CT image acquired at the 0% phase of the respiratory cycle, while varying the value of α in the 0.05 - 1.0 range, with a 0.05 increment and the size of Γ in the $3\times3\times3$ - $21\times21\times21$ range with an increment of $2\times2\times2$. The combination of α and Γ which provided the smallest average FEP between the estimated and the manually determined displacements for the three landmarks was chosen for use. A similar procedure was carried-out in order to calibrate α and Γ for CT - CBCT registration, with the difference that the reference image was replaced with a synthetic CBCT volume, generated from the CT image acquired at the 0% phase of the respiratory cycle. The CBCT image was synthesized using 72 projections.
For the optical flow algorithm, used in the experiment described by section 8.2.2.1, an implementation similar to the one described in[34] was carried-out. The optical flow algorithm provides motion estimates as the minimizers of the following functional:

$$E_{\mathrm{OF}}(\mathbf{u}) = \sum_{\vec{r}\in\Omega} \left((I(\vec{r}) - J(\vec{r} + \mathbf{u}(\vec{r})))^2 + \beta^2 \|\vec{\nabla}\mathbf{u}(\vec{r})\|_2^2 \right) \tag{8.7}$$

where I and J are the reference and the moving image, $\mathbf{u} = (u, v, w)$ is the 3D set of displacements, $\vec{r}$ is a spatial location, Ω is the image domain, $\vec{\nabla}$ is the gradient operator, $\|\cdot\|_2$ is the Euclidean norm and β is a parameter linking the two terms of the functional. Details related to the numerical scheme used to solve Eq. 8.7 can be found in.[34] Similar to α and Γ, the input parameter β was calibrated via an exhaustive search for the value which provided the smallest FEP between the estimated and the manually determined displacements of the same three landmarks specified above. During the search, the value of β was varied in the 0.01 - 1.0 interval, with an increment of 0.05 (except for the first incrementation which was 0.04).

8.3 Results

Section 8.3.1 provides an evaluation of the EVolution algorithm's capability for registering high resolution CT images. In addition, the associated results are put in perspective with respect to the optical flow algorithm. Sections 8.3.2 and 8.3.3 showcase the algorithm's performance for registering CT to CBCT images for both synthetic and clinical CBCT images. The potential of the method to perform CBCT-based 3D respiratory motion estimation is showcased in section 8.3.4. The results associated to the calibration of the algorithm input parameters are illustrated in section 8.3.5, followed by a report on the computational requirements of the EVolution method in section 8.3.6.

8.3.1 Assessment of the EVolution algorithm for CT to CT registration

Table 8.1 reports the DSC, NCC and SSIM for several volumes-of-interest, after registering pairs of CT images acquired at different time instants on 5 head-and-neck cancer patients. Registration was performed using both the EVolution and the optical flow algorithm. Note that, with very few exceptions, the DSC for both methods remains above 0.7 for all anatomies under consideration, with a consistent value of ∼0.98 for the body contour itself. Moreover, it can be observed that in the majority of cases, the two registration algorithms provide comparable DSC values. Concerning the NCC and SSIM global criteria, the two methods again provide similar results.
Table 8.2 reports the FEP between the motion fields provided by the EVolution and the optical flow algorithms. Evaluation was performed for the same CT data analyzed in Table 8.1. In the majority of cases, the mean and the standard deviation of the FEP between the two methods remains sub-millimeter. Note that the higher FEP values in Table 8.2 also correspond to a larger DSC difference between the two methods in Table 8.1.

Criterion	VOI	Method	P#1	P#2	P#3	P#4	P#5
DSC	Body	OF	0.98	0.98	0.98	0.97	0.97
		EVo	0.98	0.98	0.98	0.98	0.98
	Spinal Cord	OF	0.77	-	0.84	0.7	0.78
		EVo	0.8	-	0.83	0.71	0.8
	Right Parotid	OF	0.85	0.85	0.85	0.71	0.73
		EVo	0.86	0.84	0.86	0.81	0.63
	Left Parotid	OF	0.84	0.83	0.88	0.74	0.77
		EVo	0.85	0.82	0.9	0.82	0.76
	CTV	OF	0.69	0.83	0.87	0.89	0.77
		EVo	0.7	0.82	0.87	0.89	0.73
	Lymph Node 1	OF	0.83	0.82	0.89	0.76	0.8
		EVo	0.84	0.82	0.87	0.75	0.83
	Lymph Node 2	OF	0.8	-	0.87	0.66	0.8
		EVo	0.8	-	0.86	0.67	0.88
NCC	Entire Volume	OF	0.98	0.99	0.99	0.97	0.99
		EVo	0.98	0.98	0.99	0.97	0.98
SSIM	Entire Volume	OF	0.93	0.96	0.97	0.92	0.93
		EVo	0.92	0.95	0.96	0.91	0.92

Table 8.1: DSC, NCC and SSIM after the registration of pairs of CT images acquired on 5 head-and-neck cancer patients. The images were registered using both the optical flow (OF) and the EVolution (EVo) algorithm. Results are reported for several volumes-of-interest (VOI). The missing values indicate that delineations were not available for the respective anatomies.

VOI	FEP between OF and EVo [mm]				
	P#1	P#2	P#3	P#4	P#5
Body	0.8 ± 0.6	0.9 ± 0.8	0.8 ± 0.7	0.9 ± 0.6	0.8 ± 0.7
Spinal Cord	0.3 ± 0.1	-	0.3 ± 0.1	0.3 ± 0.1	0.3 ± 0.2
Right Parotid	0.8 ± 0.4	0.3 ± 0.1	0.4 ± 0.2	1.4 ± 0.6	1.4 ± 0.6
Left Parotid	0.6 ± 0.3	0.5 ± 0.2	0.4 ± 0.1	1.0 ± 0.6	0.9 ± 0.5
CTV	0.7 ± 0.4	0.4 ± 0.2	0.6 ± 0.3	0.3 ± 0.2	1.1 ± 0.4
Lymph Node 1	0.6 ± 0.3	0.3 ± 0.2	1.0 ± 0.4	0.6 ± 0.2	0.8 ± 0.3
Lymph Node 2	0.6 ± 0.2	-	0.7 ± 0.3	0.7 ± 0.4	1.2 ± 0.7

Table 8.2: FEP between the motion fields provided by the EVolution and the optical flow methods. Reporting is made under the format *mean* ± *standard deviation*, for the same CT data used to generate Table 8.1.

8.3.2 Algorithm performance for CT to synthetic CBCT registration

The accuracy and precision of the EVolution algorithm following the alignment of a CT to synthesized CBCT images was evaluated for a head-and-neck cancer patient (more precisely Patient #1 from Table 8.1). Methodological details related to the evaluation can be found in section 8.2.2.2.
Fig. 8.2(a) - 8.2(e) illustrate a transverse slice from the synthetic CBCT images. The images were synthesized using 360, 120, 72, 50 and 36 projections, respectively. From a visual analysis it can be observed that as the number of projections decreases, the streaking artifacts in the synthesized CBCT images become stronger. Fig. 8.2(f) - 8.2(j) on the other hand, showcase the spatial distribution of the FEP associated to each of the images from Fig. 8.2(a) - 8.2(e). Notice that, especially in homogeneous areas, the alignment errors have a tendency to increase with the amount of artifacts present in the images, coming close to 5 mm when 36 projections are used for the reconstruction of the CBCT image.
The registration errors were analyzed in more detail for a CBCT image reconstructed using 72 projections (see section c for details). Fig. 8.3(a) - 8.3(c) displays a coronal, a sagittal and a transverse slice selected from this CBCT image. The red overlay on the three images corresponds to the location of the clinical tumor volume (CTV). Fig. 8.3(d) - 8.3(f) illustrate the spatial distribution of the FEP prior to registration. This was evaluated by setting $\mathbf{u}(\vec{r})$ in Eq. 8.6 to zero. Since $\mathbf{u_{gold}}(\vec{r})$ is assumed to be the true motion, by setting $\mathbf{u}(\vec{r})$ equal to zero the FEP provides the misalignments that would occur in the absence of a registration procedure. It can be observed that, locally, the alignment errors come close or even exceed 25 mm. The spatial distribution of the FEP after registration is illustrated in Fig. 8.3(g) - 8.3(i). With a few exceptions, the alignment errors rest well beneath 1.5 mm, which corresponds to the image voxel size.

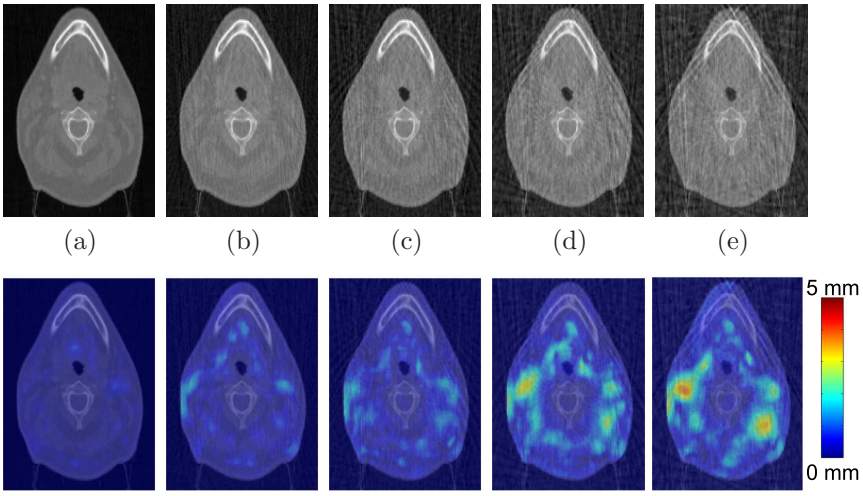


Figure 8.2: (a) - (e) Effect of the number of projections on the CBCT images: Transverse slice from a CBCT image synthesized/reconstructed using (a) 360 (b) 120 (c) 72 (d) 50 (e) 36 projections. The original CT image used to synthesize the illustrated CBCT images, was acquired on a head-and-neck cancer patient. (f) - (j) Performance of the EVolution algorithm when registering a high resolution CT to CBCT images synthesized using a decreasing number of projections: The spatial distribution of the FEP between the estimated displacements and the gold standard, following the registration of a CT to a synthetic CBCT image reconstructed using (f) 360 (g) 120 (h) 72 (i) 50 (j) 36 projections. The alignment errors are illustrated as an overlay for the transverse slice in Fig. (a) - (e).

A statistical analysis of the FEP for the pathological tissues and several organs-at-risk was also performed. This is reported, before and after registration, in the second and third columns of Table 8.3. The error statistics are provided under the format *mean*±*standard deviation* of a set that includes the pooled FEP in all the voxels contained by the analyzed anatomical structures. On average, after registration, the alignment errors remain sub-voxel. The same observation can be made for the precision of the estimated motion, indicated by the FEP standard deviation. The DSC values before and after registration are reported in columns 4 and 5 of Table 8.3. Improvements can be observed after registration, for all VOIs, with DSC values ranging from 0.7 for the smaller structures such as the CTV, up to 0.98 for the body itself. An evaluation of the NCC before and after image registration also showcases an increase from 0.89 to 0.96.

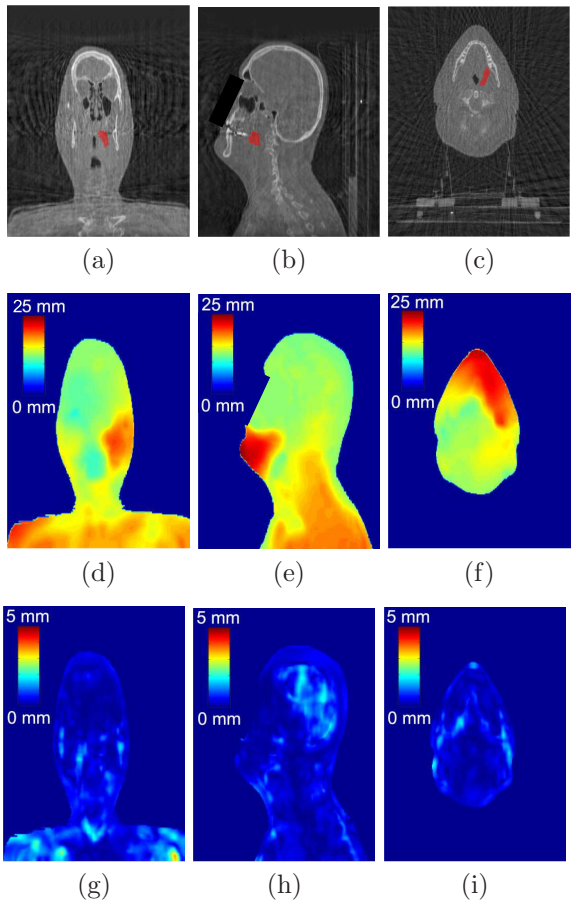


Figure 8.3: Performance of the EVolution algorithm for registering a CT image from a head-and-neck cancer patient to a synthesized CBCT image: (a) - (c) A coronal, a sagittal and a transverse slice selected from the synthetic CBCT image. The red overlay indicates the location of the CTV. (d) - (f) The spatial distribution of the FEP before registration in the slices displayed in (a) - (c). (g) - (i) The spatial distribution of the FEP in the slices from (a) - (c) after registration. Note that for illustration purposes, two different color scales were used in the error maps from (d) - (f) and (g) - (i). Also, the patient's face was masked in the sagittal slice for privacy reasons.

8.3.3 Algorithm performance for registering CT to clinical CBCT images

Fig. 8.4 displays a coronal (first row), a sagittal (second row) and a transverse (third row) slice selected from a planning CT - daily CBCT pair of images

VOI	FEP before[mm] (mean±stdev)	FEP after[mm] (mean±stdev)	DSC before	DSC after
CTV	16.58 ± 2.2	0.31 ± 0.18	0.4	0.7
Lymph node 1	15.85 ± 1.27	0.7 ± 0.37	0.27	0.84
Lymph node 2	14.38 ± 1.68	0.65 ± 0.33	0.56	0.8
Body	16.1 ± 3.45	0.66 ± 0.6	0.91	0.98
Spinal cord	16.21 ± 1.74	0.15 ± 0.09	0	0.8
Left parotid	16.38 ± 1.63	0.73 ± 0.42	0.38	0.84
Right parotid	12.7 ± 1.12	0.65 ± 0.35	0.58	0.85

Table 8.3: Qualitative and quantitative evaluation of the algorithm performance when registering a CT to a CBCT image synthesized using 72 projections. The statistical distribution of the FEP is reported in columns 2 and 3 for the anatomical sites indicated in the first column, before and after registration. The fourth and fifth column of the table report the DSC for the indicated anatomies before and after registration.

acquired on a renal cell carcinoma patient. Each column contains the following (from left to right): a slice from the CBCT image, a slice from the CT image, checkerboard overlap between the images before registration and checkerboard overlap between the images after registration. An improvement in image alignment can be observed, especially for the implanted fiducial markers (whose coherence is restored after registration), for the skeletal structures and at organ boundaries. As specified in section 8.2.2.3, 10 landmarks were manually identified in both the CBCT and the CT image. Comparing the manually determined displacements for the landmarks to the ones provided by the EVolution algorithm led to the observation that the FEP in the 10 landmarks was reduced from 9.57 ± 1.61 mm in the absence of registration, to 1.18 ± 0.3 mm after registration. In addition, after registration, the NCC showcased an increase from 0.86 to 0.94.

A set of images following the same arrangement as the ones in Fig. 8.4 is illustrated in Fig. 8.5, however, the kidney case was replaced with images acquired on a lung cancer patient. Similar to the kidney case, a notable improvement in image alignment can be observed after registration, especially in the lung parenchyma, the skeletal structures and the organ boundaries. Manual tracking of 10 landmarks revealed a decrease of the alignment errors from 6.58 ± 1.54 mm in the absence of registration, to 0.94 ± 0.3 mm with registration. After registration, the NCC also increased from 0.81 to 0.96.

8.3.4 Synthetic CBCT-based 3D respiratory motion estimation

As detailed in section 8.2.2.4, for this particular experiment, CBCT images were synthesized using as a base a 4D CT series covering one respiratory cycle of a

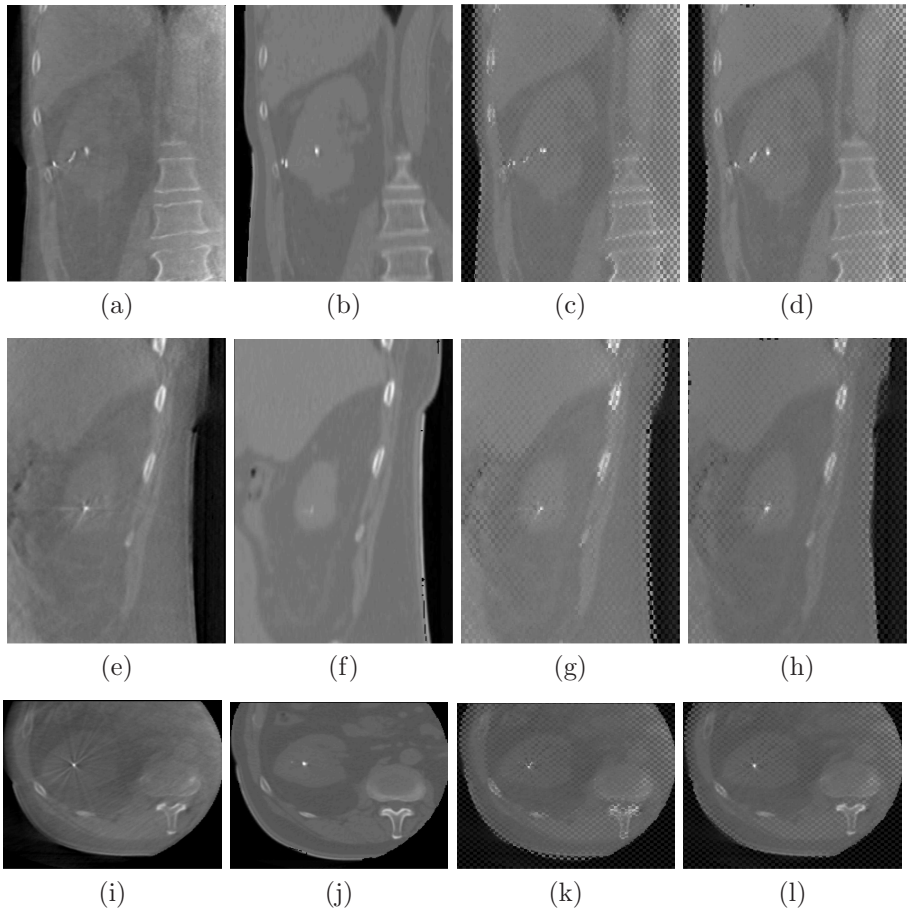


Figure 8.4: (a) - (b) A coronal, (e) - (f) a sagittal and (i) - (j) a transverse slice selected from a CBCT (first column) and a CT image (second column), acquired on a renal cell carcinoma patient. (c) - (d) A coronal, (g) - (h) a sagittal and (k) - (l) a transverse slice selected from a checkerboard overlap between the CBCT and the CT images before (third column) and after registration (fourth column).

renal cell carcinoma patient. Fig. 8.6(a) - 8.6(c) illustrate a coronal, a sagittal and a transverse slice, selected from the CBCT image used as reference during the registration process. Fig. 8.6(d) - 8.6(f) and Fig. 8.6(g) - 8.6(i) display for the slices illustrated in Fig. 8.6(a) - 8.6(c) the temporally averaged FEP, before and after registration. It can be observed that errors which exceed 5 mm in moving organs such as the liver and kidneys, are reduced, in most areas, to values smaller than $\sim$1.5 - 2 mm. Note that the results illustrated in Fig. 8.6

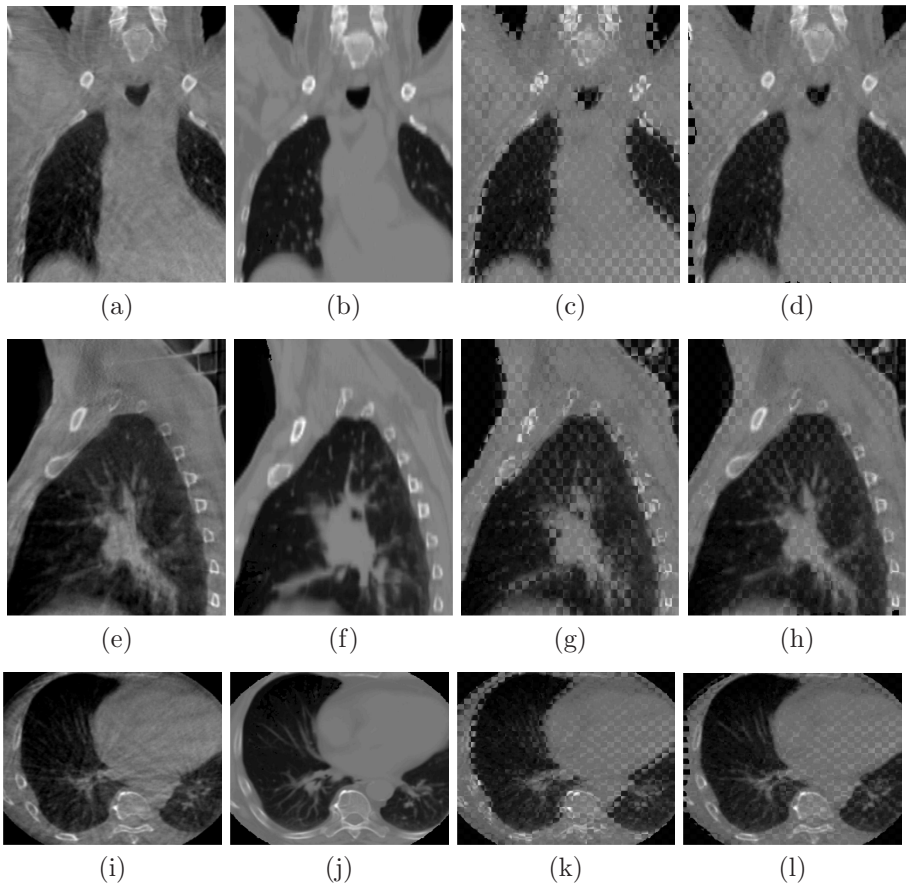


Figure 8.5: (a) - (b) A coronal, (e) - (f) a sagittal and (i) - (j) a transverse slice selected from a CBCT (first column) and a CT image (second column) acquired on a lung cancer patient. (c) - (d) A coronal, (g) - (h) a sagittal and (k) - (l) a transverse slice selected from a checkerboard overlap between the CBCT and the CT images, before (third column) and after registration (fourth column).

were obtained from a 4D CBCT series synthesized using 120 projections.

The performance of the EVolution algorithm for CBCT-based respiratory motion estimation was further evaluated by analyzing the estimated 3D trajectory of two landmarks, over the respiratory cycle. One of the landmarks was a fiducial gold marker (indicated by the red arrow in Fig. 8.6), while the second landmark was the centroid of the contralateral kidney (delineated by the blue dashed line in Fig. 8.6). The gold standard trajectory for the fiducial marker was provided by an estimation on the original 4D CT image series, while for

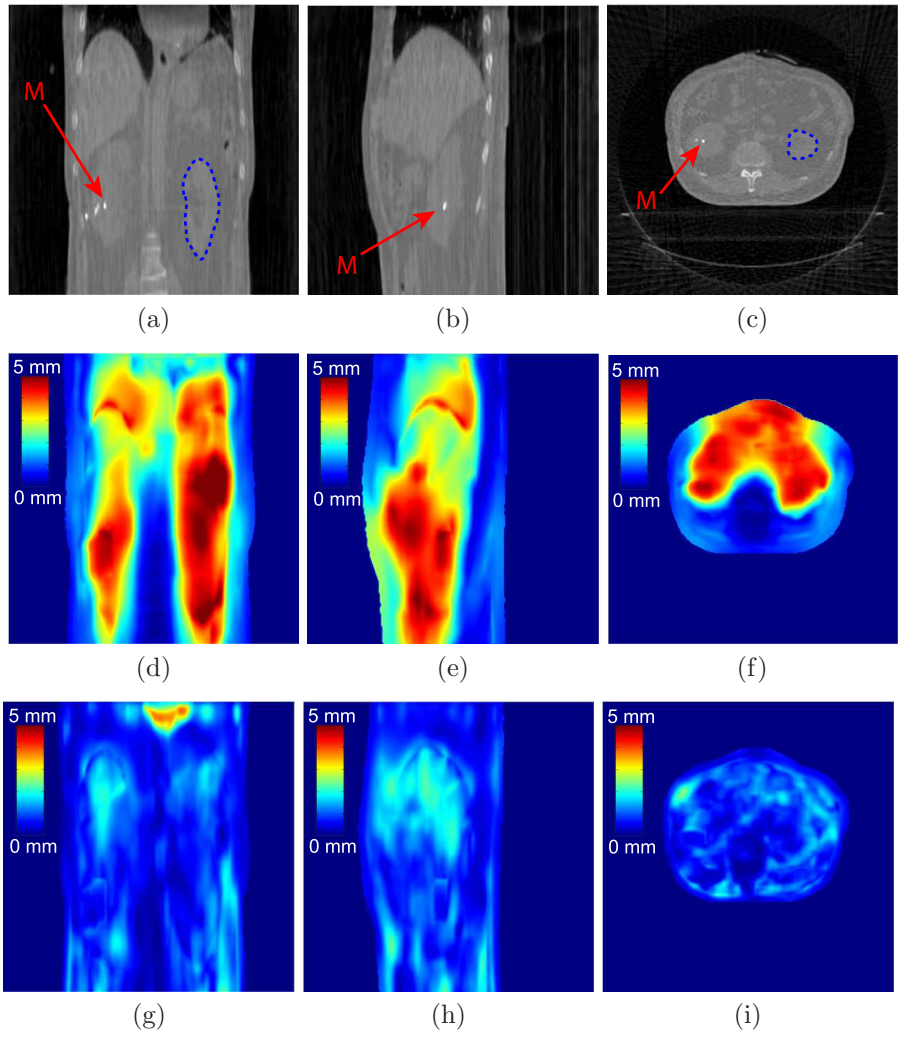


Figure 8.6: Performance of the EVolution algorithm for CBCT-based respiratory motion estimation: (a) - (c) A coronal, a sagittal and a transverse slice selected from the CBCT image used as reference during the registration process. Point M indicates one of the implanted fiducial gold markers, while the blue contour delineates the contralateral kidney. (d) - (f) The spatial distribution of the temporally averaged FEP in the slices from (a) - (c), before aligning the synthetic 4D CBCT series to the reference image. (g) - (i) The spatial distribution of the temporally averaged FEP in the slices from (a) - (c), after registering the 4D series to the reference position.

the contralateral kidney centroid, manual tracking was performed. Fig. 8.7 illustrates the trajectory of the two landmarks in the head - foot, anterior - posterior and the left - right directions. It can be observed that, with very few exceptions, the distance between the estimated curves and the gold standard remains sub-millimeter. Note that the CBCT images were synthesized from the original 4D CT series using 120 projections.

Table 8.4 reports the errors between the trajectories estimated for the landmarks and their corresponding gold standard trajectory, for three scenarios: 1) Both the reference and the moving CBCT images were synthesized using 120 projections (second row); 2) The reference image was synthesized using 120 projections, while the moving images were synthesized using 72 projections (third row); 3) Both the reference and the moving images were synthesized using 72 projections (fourth row). The first row in the table reports the FEP for the two landmarks, without registering the 4D synthetic CBCT series. The FEP is reported in the *mean* $\pm$ *standard deviation* format, with the statistics computed on the pooled FEPs from all the sampled phases of the respiratory cycle. Notice that for the two scenarios in which the reference CBCT image is reconstructed using 120 projections, the errors remain approximately in the sub-millimeter range, regardless of the amount of the streaking artifacts in the moving images. However, when both the reference and the moving images were synthesized using 72 projections, the quality of the motion estimates in the contralateral kidney is affected.

Registration type	Registration errors [mm] *mean* $\pm$ *stdev*	
	Marker	Contralateral kidney centroid
No registration	4.65 ± 3.8	5.01 ± 2.9
$CBCT^{120}$ to $CBCT^{120}$	0.55 ± 0.38	0.89 ± 0.6
$CBCT^{72}$ to $CBCT^{120}$	0.44 ± 0.4	0.85 ± 0.5
$CBCT^{72}$ to $CBCT^{72}$	0.32 ± 0.19	2.03 ± 1.24

Table 8.4: FEP between the trajectories estimated for the implanted marker and the contralateral kidney, for different amounts of streaking artifacts added to the synthesized 4D CBCT images. The first row indicates the FEP for the two landmarks, in the absence of a registration scheme. The subsequent rows report the errors between the estimated and the gold standard trajectories for the cases in which both the reference and the moving images were synthesized using 120 projections (second row), the reference image was synthesized using 120 projections and the moving images were synthesized using 72 projections (third row) and when both the reference and the moving images were reconstructed using 72 projections (fourth row).

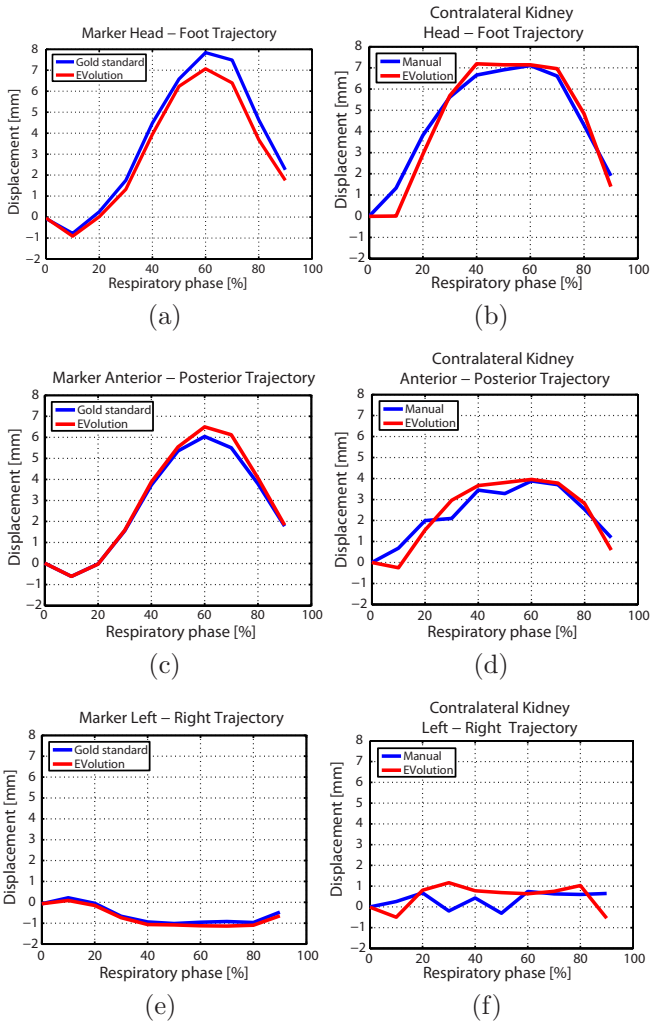


Figure 8.7: Performance of the EVolution algorithm for CBCT-based respiratory motion estimation: Trajectory of a fiducial gold marker and the centroid of the contralateral kidney over the respiratory cycle in the (a) - (b) head - foot, (c) - (d) anterior - posterior and (e) - (f) left - right directions. The red curves correspond to the trajectories estimated on the synthetic 4D CBCT series, while the blue curves plot their respective gold standards. The ordinate of the graphs indicate the displacement in millimeters, while the abscissa indicates the percentage phase within the respiratory cycle.

8.3.5 Input parameter calibration for the EVolution and optical flow algorithm

The EVolution algorithm requires α and Γ as input parameters (see Eq. 8.1 and 8.2), while the optical flow algorithm requires the calibration of the reg-

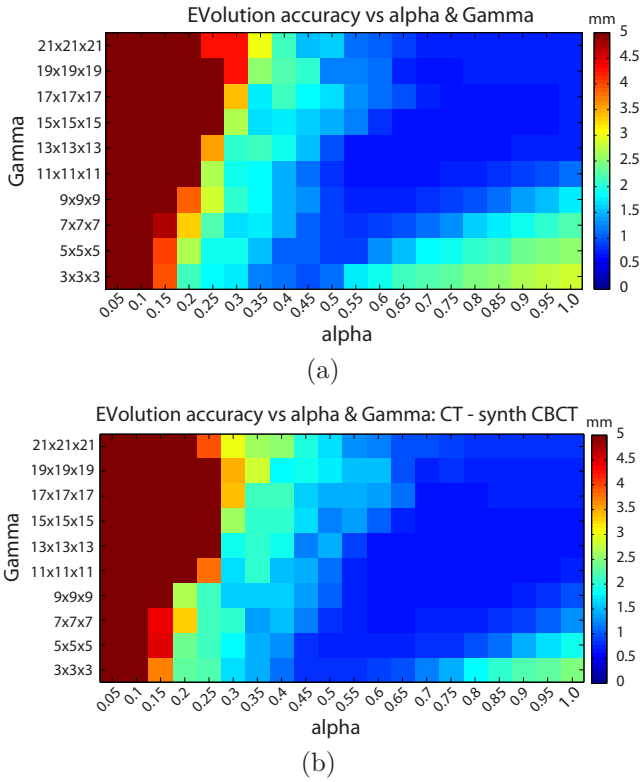


Figure 8.8: Accuracy of the EVolution algorithm as a function of the input parameters α and Γ: (a) For CT - CT registration. (b) For CT - synthetic CBCT registration.

ularization parameter β (see Eq. 8.7). Details related to the manner in which these input parameters were optimized can be found in section 8.2.3.
The dependency between estimation accuracy and the parameters α and Γ for both CT to CT and CT to synthetic CBCT registration via the EVolution algorithm is displayed in Fig. 8.8. The highest accuracy of 0.63 mm and 0.67 mm, respectively, is achieved for the combination $(\alpha, \Gamma) = (0.6, 11 \times 11 \times 11)$ for both calibrations. It can also be observed that the accuracy remains close to its optimal value for a rather wide range of α and Γ values. A rapid decrease in algorithm accuracy can be noticed as the value of α drops under 0.25 - 0.3. Fig. 8.9 illustrates, in the scope of this study, the relationship between the accuracy of the optical flow algorithm and the regularization parameter β. It can be observed that the best accuracy is $\sim$1.2 mm, which is achieved for a value of β equal to 0.05.

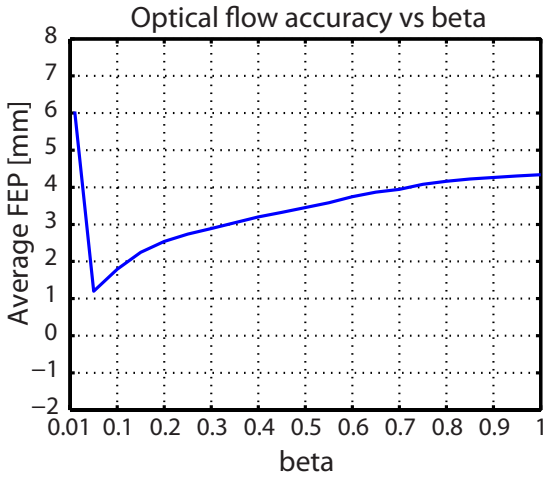


Figure 8.9: Accuracy of the optical flow algorithm as a function of the regularization parameter β.

8.3.6 Computational performance of the algorithm

Following the benchmarking of the computational requirements of the EVolution algorithm it was found that for images of size 256×256×256, the algorithm converges on average in approximately 60 s.

8.4 Discussion

Several concepts in image guided radiotherapy, such as day-to-day positioning compensation,[3–5] "virtual couch shifts",[42] dose accumulation[43,44] or delineation propagation[25] often rely on establishing a spatial coherence, using image registration, between a planning high-resolution CT and a daily CBCT image and/or between CBCT images acquired at different time instants. In order to limit the radiation dose associated to imaging, which is of particular importance for pediatrics,[45] the CBCT volumes are frequently reconstructed using a low number of X-Ray projections and/or a low amount of beam intensity. This leads to low SNR images altered by streaking artifacts, which is problematic for state-of-the-art registration algorithms relying on gray-level intensity conservation. Moreover, the registration process is also often hampered by the low soft-tissue contrast, which is intrinsic to the CT and CBCT images. This chapter demonstrates, that despite such constraints on image quality, the recently proposed EVolution algorithm[32] is fully capable of providing reliable estimates of the deformation between the images.
In an initial step, however, prior to investigating the performance of the EVolution algorithm for CT-CBCT and CBCT-CBCT registration, the current study

evaluates its capabilities for re-aligning CT images, since CT to CT registration played a key role during the validation procedures employed in this chapter (see, for example, section 8.2.2.2). This was achieved by evaluating the post-registration DSC, NCC and SSIM provided by the EVolution algorithm, for several anatomical structures on CT data acquired on 5 head-and-neck cancer patients. During this experiment, the use of the DSC, NCC and SSIM as registration quality evaluation criteria is argued by their previous use in several independent studies focusing on CT and CBCT image registration,[22, 35, 46, 47] providing a means to relate the obtained results to prior art. In the current chapter, the obtained DSC, NCC and SSIM values were also put in perspective with the ones provided by a pre-existing optical flow algorithm. The latter was selected as a point-of-comparison due to its prior evaluation and validation for CT to CT registration in the study of Ostergaard *et al*.[23] Therein it was demonstrated that the optical flow algorithm can achieve an accuracy of ∼1.1 mm, value which was concluded to be acceptable for registering CT volumes. Such results are also in good correspondence with the ones obtained in the present chapter (see section 8.3.5). In the scope of this study, both the EVolution and the optical flow algorithm provided similar DSC, NCC and SSIM values. The DSC in particular ranged from ∼0.7 for smaller structures up to 0.98 for the body contour. These values are overall comparable with the ones provided by other state-of-the-art methods.[46] Also, an evaluation of the FEP between the motion fields provided by the EVolution and the optical flow applied on the same CT datasets, resulted on average in sub-millimeter values (see Table 8.2). There are, however, a few isolated cases with rather large discrepancies between the DSC values provided by the EVolution and the optical flow algorithm (see, for example, left and right parotid in patients #4 and #5 from Table 8.1). These discrepancies are also in good correspondence with large FEP values in Table 8.2. This can be explained by the different fundamental principles on which the two algorithms rely for performing motion estimation. While EVolution assumes the conservation of contrast between the reference and the moving image, the optical flow algorithm assumes voxel intensity conservation. Therefore, depending on which of the two assumptions is locally violated, one or the other algorithm will perform poorer in a particular area. This may also explain the rest of the smaller differences between the DSC values provided by the two compared algorithms. It is also worth noting that, in this chapter, validation of the EVolution algorithm for CT to CT registration was performed solely for confirmation purposes, since this aspect was already addressed in the seminal work of Denis de Senneville *et al*.[32]

Over the course of this chapter, two criteria played a major role for evaluating the outcome of image registration: the DSC and the FEP. The reason for which one may be chosen over the other depends on the available gold standard. For example, for CT to CT registration on the head-and-neck cancer patient data, the only available ground truth were the structure delineations made on the reference and the moving image. This does not allow the calculation of the FEP, since gold standard displacements were not available for these data sets.

Therefore, the DSC is a more suitable quality evaluation metric in this case. The downside of the DSC is that it only provides a qualitative measure of registration performance. On the other hand, for the experiments where displacements or point-to-point correspondences are available as a gold standard, the FEP is preferred, since it provides a quantitative measure of registration accuracy and precision. Nevertheless, it should be noted that, in case of a severely low number of gold standard displacements or point-to-point correspondences between the reference and the moving image, the DSC may be a better choice (since it usually encompasses a larger number of voxels).
In the scope of this chapter, the performance of the EVolution algorithm for CT to CBCT registration was evaluated using three gold standards: the motion fields provided by the registration of high resolution CT images, the DSC for several anatomical structures, including pathological tissues and OARs and manually annotated landmarks. Qualitative visual criteria were also employed for the evaluation, in particular for the study described in section 8.2.2.3. Fig. 8.2 showcases the impact on the motion estimates of an increasingly coarse sampling of the sinogram for the reconstruction of the analyzed CBCT image. Note that even for CBCT images reconstructed using only 72 projections, leading to an image quality at the lower end of what a typical clinical system would provide, the additional registration errors remain predominantly sub-millimeter. Even for less than 72 projections, only a few error peaks start occurring in the more homogeneous areas of the image. The body contour and skeletal muscle are still correctly re-aligned. These observations were further confirmed during the more in-depth analysis conducted on one of the head-and-neck cancer patients. Fig. 8.3 showcases the spatial distribution of the alignment errors in the aforementioned patient, following the registration of a CT to a CBCT image reconstructed using 72 projections. The errors are illustrated in three planes intersecting approximately midway into the CTV. Noteworthy is that the alignment errors remain mostly sub-millimeter, including for the nasal and air cavities, which have proven problematic in previous related studies.[21, 22, 30]
The only locations in which the errors of the EVolution method have a tendency to increase is in areas devoid of contrast, such as deep into the brain and isolated into the shoulder muscles. There are basically two factors which contribute to this: For one, the lack of intrinsic contrast in the CTs in these areas shifts the burden of the registration process towards the regularization term. Second, since the CBCT artifacts in such areas are dominant compared to "true" anatomical contrast, they counteract the effect of the regularization, which as a net-effect degrades the quality of the registration in these domains. Nevertheless, even in these regions, with few exceptions, the errors remain close to the voxel size ($\sim$1.5 mm). The statistical distribution of the alignment errors for the pathology and several of the organs-at-risk of the same patient is reported in Table 8.3. Following registration, sub-millimeter precision and accuracy was achieved for all analyzed anatomical structures. Also, by comparing the post-registration DSC values reported in Table 8.3 to the ones in Table 8.1 (Patient #1) it can be observed that, despite the additional artifacts and low

SNR altering the CBCT image, the DSC remains approximately the same ($\geq$0.7 for smaller anatomies up to 0.98 for the body contour) as if a CT acquisition would have been used instead. The good performance of the EVolution algorithm for CT-CBCT registration is further confirmed by the kidney and lung cases illustrated in Fig. 8.4 and 8.5, respectively. Following registration, in both cases, a high degree of spatial continuity from one square to the next in the composite images can be observed for the visible features. The most obvious are the implanted fiducials in the kidney cancer patient, the parenchyma of the lung cancer patient and the organ boundaries and skeletal structures in both patients. Comparing the manually determined displacements for 10 landmarks in each of the two data sets to the ones provided by the EVolution algorithm revealed an accuracy and precision comparable to the voxel size. In addition, both cases have demonstrated after registration an increase of the NCC to values which are in good correspondence with the state-of-the-art.[22]

Concerning its capability for CBCT-based 3D respiratory motion estimation (i.e. for CBCT to CBCT image registration), the EVolution algorithm was validated against two gold standards: the motion fields resulting from the alignment of the CT images used to synthesize the 4D CBCT series and the trajectory of two landmarks over a respiratory cycle. Fig. 8.6 displays the spatial distribution of the temporally averaged alignment errors for the 4D CBCT series. In particular for the kidneys, which were the organs of interest in the illustrated data, it can be observed that after the alignment of the 4D series, the errors remain mostly sub-millimeter. On a global scale, except for a few isolated areas, the post-registration misalignments rest under the in-plane voxel size (2 mm). The good capabilities for respiratory motion estimation of the EVolution algorithm are further emphasized in Fig. 8.7, where it can be observed that the estimated 3D trajectory of the selected landmarks closely follows the gold standard trajectory. A more quantitative evaluation of the errors between the estimated and the gold standard trajectories is provided in Table 8.4. In the table it is shown that, as long as the sinogram sampling of both CBCT images (reference and moving image) resides above a particular threshold, sub-millimeter precision and accuracy can be achieved. Also, asymmetric registrations with a high-quality reference image and a "faster" CBCT image of lower quality maintains an accuracy and precision below one millimeter. It is only when both the reference and the moving images of the 4D CBCT series are reconstructed using 72 projections that the algorithm performance starts to degrade below the native image resolution. An interesting aspect is that, as the quality of the synthetic CBCT images degrades, only the registration errors for the centroid of the contralateral kidney have a tendency to increase (see Table 8.4). This can be explained by the presence of the implanted fiducial markers in the ipsilateral kidney, which appear as hyper-intense features in the images. Therefore, despite an increase in the magnitude of the streaking artifacts, the fiducial markers provide sufficient contrast in the synthetic CBCT images, facilitating and accurate and precise tracking.

An issue which was generally overlooked by previous related studies is that the

CT and the CBCT images are often acquired using different geometries (e.g. FOV, image size, voxel size, etc.). If not taken into account, this can frequently lead to unreliable motion estimates. In the current study this was addressed by regridding the images into a common coordinate system prior to registration. The dimensions and voxel size of the new lattice were chosen such that the entire FOV of the original images is also covered by the regridded images. Obviously, the two lattice parameters (i.e. image and voxel size) can also be optimized with respect to computational performance since, in general, the computational time required by the EVolution algorithm depends on the size of the images. In addition, for the kidney and lung cancer patients (see Fig. 8.4 and 8.5) the FOV of the CBCT images was considerably smaller than that of the CT images, which is often the case in clinical practice. If not taken into consideration, this may impair the convergence of the algorithm to the global minimum. In the scope of this chapter, this was addressed by imposing boundary conditions on the CBCT image such that all voxels that are outside or on the border of the FOV are excluded from the registration process.
One of the traits that makes EVolution an attractive option for motion estimation in a clinical setting is the low number of input parameters it requires: the regularization parameter α and the patch size Γ (see Eq. 8.1 and 8.2). A noteworthy fact is that, once Γ and α were optimized, their value was maintained constant during all experiments, with a consistent good quality of the motion estimates. Moreover, both the calibration for CT to CT registration and for CT to synthetic CBCT registration provided the same optimal values for α and Γ (see Fig. 8.8). So it is fair to say that for the problem at hand, no inter-individual adjustments of the parameters were necessary. Therefore, we hypothesize, that once the acquisition scheme of the CBCT images is fixed, the values of α and Γ could be optimized prior to the intervention, making algorithm configuration completely transparent to the clinician delivering the therapy. One elegant way to optimize the value of the input parameters prior to the intervention is, for example, to register two CT images acquired on a patient and establish the resulting motion field as gold standard. One or both of the images (depending on whether CT-CBCT or CBCT-CBCT registration is of interest) could then be used to synthesize CBCT volumes using the same acquisition parameters as the clinical system. Offline CT - synthetic CBCT or synthetic CBCT - synthetic CBCT registration is then repeatedly performed while exhaustively searching for the optimal values of α and Γ. The search is then stopped when the configuration that provides the best similarity between the estimated displacements and the gold standard is found.
An aspect which may prove to be slightly problematic for the EVolution algorithm is estimating motion in the proximity of sliding interfaces. The quadratic nature of the regularization term included in the functional from Eq. 8.1, constrains the estimated displacements to be spatially smooth. For soft tissues in particular, such an assumption is physically justified since they are incompressible and elastic and thus, shearing motion will in general not occur. On the other hand, at interfaces between organs and the abdominal or thoracic

wall or between the organs themselves, sliding motion usually occurs, which locally violates the smoothness assumption. For example, during the experiment addressing CBCT-based respiratory motion compensation, this lead to a propagation of the displacements estimated for the kidneys and liver, ∼2 - 3 voxels into the quasi-static abdominal and thoracic wall. However, this only had a local effect without notably affecting the overall performance of the algorithm. Nevertheless, as a source of misregistration this aspect may have to be monitored and can make the object of future studies.
An important feature for registration algorithms in general is the required computational time. This is of particular interest when registration is employed for online therapy guidance, with the patient still on the interventional table. Already from a numerical point-of-view, the EVolution algorithm facilitates an implementation with low computational demands since its optimization scheme is easily and massively parellelizable. This is due to the fact that the functional in Eq. 8.1 is an integral over a set of convex functions (one for each voxel of the image). It is a know result from convex optimization theory that the independent optimization of each of these functions leads to a global optimum of the functional. Therefore, the optimization of each function can and was delegated to a separate processing core, with all voxels being processed in parallel. Moreover, while not illustrated in this study, all operations required by the numerical scheme of the EVolution algorithm can be implemented by using convolutions, which again render themselves easily parallelizable. As reported in section 8.3.6, for images of size $256 \times 256 \times 256$ the algorithm requires ∼60 s per registration. This duration may scale with the image size, the extent of the displacements between the registered images and the available hardware.
In order to correct for physiological motion and/or positioning errors, the current clinical standard for CBCT-guided radiotherapy implies the alignment of CT/CBCT images acquired at different stages of the treatment based on implanted fiducial markers. Besides causing discomfort to the patient during the implantation process, the marker-based alignment only accounts for rigid displacements, which may be sub-optimal for deformable anatomies. By employing the EVolution algorithm instead, elastic deformation fields become available for both pathological tissues and organs-at-risk, providing the possibility of improved corrections due to the additional positional information. Moreover, this circumvents the implantation of fiducial markers, since the EVolution algorithm is purely image-based. We also anticipate that the dense elastic deformation fields could be used for adaptive replanning in the sense that the daily radiation dose could be delivered in accordance with the position of the pathology and the organs-at-risk at the time of treatment. Also, the daily dose can be projected with respect to the estimated displacements into the reference space of the planning CT image. This way, therapy progress could be monitored with respect to the initial plan and, if necessary, the remaining therapeutic dose can be adapted accordingly.

8.5 Conclusion

For an accurate and precise therapeutic dose delivery, image-guided radiotherapy treatments often rely on establishing a spatial coherence via image registration between CBCT and CT or CBCT images acquired at different stages of the therapy. Due to imaging-related radiation constraints, CBCT images are generally acquired using a low number of X-Ray projections and/or low beam intensities. This leads to acquisitions altered by streaking artifacts and a low SNR, which renders their registration using intensity-based methods challenging. The current study addresses both CT to CBCT and CBCT to CBCT registration by the means of the recently proposed EVolution algorithm. The method was validated in several complementary experiments where it was demonstrated that the artifacts associated to the sparse sampling of the CBCT sinogram only have a minimal impact, of sub-millimeter order, on the quality of the motion estimates. The EVolution algorithm requires the configuration of only two input parameters which, in practice, can be optimized prior to the therapy. Their values can then be maintained across different patients and anatomies, making the configuration of the method completely transparent to the clinician. In addition, the investigated method implies low computational demands, which is rarely the case for registration algorithms providing an elastic deformation between two images on a point-by-point basis. It can be therefore concluded that the EVolution algorithm is potentially an attractive asset for online motion compensation of CBCT-guided external beam radiotherapy treatments.

Acknowledgements

This work was supported by the Dutch Technology Foundation (STW) (project OnTrack #12813).

Bibliography

[1] T. Gupta and C. Anand Narayan. Image-guided radiation therapy: Physician's perspectives. *J Med Phys*, 37(4):174–182, 2012.

[2] M. Guckenberger. Image-guided radiotherapy based on kilovoltage cone-beam computed tomography - A review of technology and clinical outcome. *Eur Onc Haema*, 7(2):121–124, 2011.

[3] C. Thilmann, S. Nill, T. Tücking, A. Höss, B. Hesse, L. Dietrich, R. Bendl, B. Rhein, P. Häring, C. Thieke, U. Oelfke, J. Debus, and P. Huber. Correction of patient positionning errors based on in-line cone beam CTs: Clinical implementation and first experiences. *Radiat Oncol*, 1:16, 2006.

[4] M. Oldham, D. Létourneau, L. Watt, G. Hugo, D. Yan, D. Lockman, L. H. Kim, Y. Chen P., A. Martinez, and J. W. Wong. Cone-beam-CT guided radiation therapy: A model for on-line application. *Radiother Oncol*, 75(3):271–278, 2005.

[5] J. Lehmann, J. Perks, S. Semon, R. Harse, and J. A. Purdy. Commissioning experience with cone-beam computed tomography for image-guided radiation therapy. *J Appl Clin Med Phys*, 8(3):2354, 2007.

[6] P. J. Keall, G. S. Mageras, J. M. Balter, R. S. Emery, K. M. Forster, S. B. Jiang, J. M. Kapatoes, D. A. Low, M. J. Murphy, B. R. Murray, C. R. Ramsey, M. B. van Herk, and S. S. Vedam. The management of respiratory motion in radiation oncology report of AAPM Task Group 76. *Med Phys*, 33(10):3874–3900, 2006.

[7] T. E. Marchant, G. J. Price, B. J. Matuszewski, and C. J. Moore. Reduction of motion artefacts in on-board cone beam CT by wapring of projection images. *Br J Rdiol*, 84(999):251–264, 2011.

[8] T. Li, E. Schreibmann, Y. Yang, and L. Xing. Motion correction for improved target localization with on-board cone-beam computed tomography. *Phys Med Biol*, 51(2):253–267, 2006.

[9] J. J. Sonke, L. Zijp, P. Remeijer, and M. van Herk. Respiratory correlated cone beam CT. *Med Phys*, 32(4):1176–1186, 2005.

[10] L. Dietrich, S. Jetter, T. Tücking, S. Nill, and U. Oelfke. Linac-integrated 4D cone beam CT: first experimental results. *Phys Med Biol*, 51(11):2939–2952, 2006.

[11] X. Jia, Z. Tian, Y. Lou, J. J. Sonke, and S. B. Jiang. Four-dimensional cone beam CT reconstruction and enhancement using a temporal nonlocal means method. *Med Phys*, 39(9):5592–5602, 2012.

[12] C. Coolens, S. Webb, H. Shirato, K. Nishioka, and P. M. Evans. A margin model to account for respiration-induced tumous motion and its variability. *Phys Med Biol*, 53(16):4317–4330, 2008.

[13] S. Dhou, M. Hurwitz, P. Mishra, W. Cai, J. Rottman, R. Li, C. Williams, M. Wagar, R. Berbeco, D. Ionascu, and J. H. Lewis. 3D fluoroscopic image estimation using patient-specific 4DCBCT-based motion models. *Phys Med Biol*, 60(9):3807–3824, 2015.

[14] W. Cai, S. Dhou, F. Cifter, M. Myronakis, M. H. Hurwitz, C. L. Williams, R. I. Berbeco, J. Seco, and J. H. Lewis. 4D cone beam CT-based dose assessment for SBRT lung cancer treatment. *Phys Med Biol*, 61(2):554–568, 2016.

[15] V.R.S. Mani and S. Arivazhagan. Survey of medical image registration. *Journal of Biomedical Engineering and Technology*, 1(2):8–25, 2013.

[16] D. L. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes. Medical image registration. *Phys Med Biol*, 46(3):R1–45, 2001.

[17] M. Ahmad, P. Balter, P Munro, J. Lu, and T. Pan. Optimization of 4D cone-beam CT: Evaluation of streaking artifacts and noise with various simulated gantry rotation speeds. In *Medical Imaging 2009: Physics of Medical Imaging*, volume 7258, Lake Buena Vista, FL, USA, 2009. SPIE Proceedins.

[18] R. Schulze, U. Heil, D. Gross, D. D. Bruellmann, E. Dranischnikow, U. Scwanecke, and E. Schoemer. Artefacts in CBCT: A review. *Dentomaxillofac Radiol*, 40(5):265–273, 2011.

[19] X. Li, Y. Y. Zhang, Y. H. Shi, L. H. Zhou, and X. Zhen. Evaluation of deformable image registration for contour propagation between CT and cone-beam CT images in adaptive head and neck radiotherapy. *Technol Health Care*, 24(2):S747–755, 2016.

[20] S. Nithiananthan, S. Schafer, A. Uneri, D. J. Mirota, J. W. Stayman, W. Zijewski, Brock K. K., M. J. Daly, J. C. Chan, H. Irish, and J. H. Siewerdsen. Demons deformable registration of CT and cone-beam CT using an iterative intensity matching approach. *Med Phys*, 38(4):1785–1798, 2011.

[21] X. Zhen, X. Gu, L. Zhou, X. Jia, and S. B. Jiang. CT to cone-beam CT deformable registration with simultaneous intensity correction. *Phys Med Biol*, 57(21):6807–6826, 2012.

[22] S. Park, W. Plishker, H. Quon, J. Wong, R. Shekhar, and J. Lee. Deformable registration of CT and cone-beam CT with local intensity matching. *Phys Med Biol*, 62(3):927–947, 2017.

[23] N. K. Ostergaard, B. Denis de Senneville, U. V. Elstrom, K. Tanderup, and T. S. Sorensen. Acceleration and validation of optical flow based deformable registration for image-guided radiotherapy. *Acta Oncol*, 47(7):1286–1293, 2008.

[24] W. H. Greene, S. Chelikani, K. Purushothaman, J. P. Knisely, Z. Chen, X. Papademetris, L. H. Staib, and J. S. Duncan. Constrained non-rigid registration for use in image-guided adaptive radiotherapy. *Med Image Anal*, 13(5):806–817, 2009.

[25] M. Rubeaux, A. Simon, K. Gnep, J. Colliaux, O. Acosta, R. de Crevoisier, and P. Haigron. Evaluation of non-rigid constrained CT/CBCT registration algorithms for delineation propagation in the context of prostate cancer radiotherapy. In *Medical Imaging 2013: Image-guided procedures, robotic interventions and modeling*, volume 8671, Lake Buena Vista, FL, USA, 2009. SPIE Proceedins.

[26] Y. Lou, T. Niu, X. Jia, P. A. Vela, L. Zhu, and A. R. Tannenbaum. Joint CT/CBCT deformable registration and CBCT enhancement for cancer radiotherapy. *Med Image Anal*, 17(3):387–400, 2013.

[27] N. Kothary, S. Dietrich, J. D. Louie, D. T. Chang, L. V. Hofmann, and D. Y. Sze. Percutaneous implantation of fiduacial markers for imaging-guided radiation therapy. *Am J Roentgenol*, 192(4):1090–1096, 2009.

[28] R. D. Timmerman and L. Xing. *Image-guided and adaptive radiation therapy.* Lippincott Williams and Wilkins, 2010.

[29] T. E. Marchant, A. Skalski, and B. J. Matuszewski. Automatic tracking of implanted fiducial markers in cone beam CT projection images. *Med Phys*, 39(3):1322–1334, 2012.

[30] D. Paquin, D. Levy, and L. Xing. Multiscale registration of planning CT and daily cone beam CT images for adaptive radiation therapy. *Med Phys*, 36(1):4–11, 2009.

[31] V. Kearney, S. Chen, X. Gu, T. Chiu, H. Liu, L. Jiang, J. Wang, J. Yordy, L. Nedzi, and W. Mao. Automated landmark-guided deformable image registration. *Phys Med Biol*, 60(1):101–116, 2015.

[32] B. Denis de Senneville, C. Zachiu, M. Ries, and C. Moonen. EVolution: an edge-based variational method for non-rigid multi-modal image registration. *Phys Med Biol*, 61(20):7377–7396, 2016.

[33] J. Weickert, A. Bruhn, N. Papenberg, and T. Brox. Variational optic flow computation: From continuous models to algorithms. In *International Workshop on Computer Vision and Image Analysis (ed. L. Alvarez), IWCVIA03, Las Palmas de Gran Canaria*, 2003.

[34] C. Zachiu, N. Papadakis, M. Ries, C. T. W. Moonen, and B. Denis de Senneville. An improved optical flow tracking technique for real-time MR-guided beam therapies in moving organs. *Phys Med Biol*, 60(23):055010, 2015.

[35] X. Li, Y. Zhang, Y. Shi, S¿ Wu, Y. Xiao, X. Gu, X. Zhen, and L. Zhou. Comprehensive evaluation of ten deformable image registration algorithms for controu propagation between ct and cone-beam ct images in adaptive hean & neck radiotherapy. *PLoS One*, 12(4):e0175906, 2017.

[36] J. Luo and E. E. Konofagou. A fast normalized cross-correlation calculation method for motion estimation. *IEEE Trans Ultrason Ferroelectr Freq Control*, 57(6):1347 – 1357, 2010.

[37] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans Image Process*, 13(4):600 – 612, 2004.

[38] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski. A database and evaluation methodology for optical flow. *Int J Comp Vis*, 92(1):1–31, 2011.

[39] NEMA PS3 / ISO 12052. Digital imaging and communications in medicine (DICOM) standard. *National Electrical Manufacturers Association, Rosslyn, VA, USA*, 2017.

[40] A. Biguri, M. Dosanjh, S. Hancock, and M. Soleimani. TIGRE: A MATLAB-GPU toolbox for CBCT image reconstruction. *Biomed Phys Eng Express*, 2(5):4–11, 2009.

[41] L. A. Feldkamp, L. C. Davis, and J. W. Kress. Practical cone-beam algorithm. *J Opt Soc Am*, 1(6):612–619, 1984.

[42] G. H. Bol, J. J. Lagendijk, and B. W. Raaymakers. Virtual couch shift (VCS): Accounting for patient translation and rotation by online IMRT reoptimization. *Phys Med Biol*, 58(9):2989–3000, 2013.

[43] A. Richter, Q. Hu, D. Steglich, K. Baier, J. Wilbert, M. Guckenberger, and M. Flentje. Investigation of the usability of conebeam CT data sets for dose calculation. *Radiat Oncol*, 3:42, 2008.

[44] J. R. Sykes, R. Lindsday, G. Iball, and D. I. Thwaites. Dosimetry of CBCT: methods, doses and clinical consequences. *J Phys Conf*, 444(1):012017, 2013.

[45] A. J. Olch. *Pediatric radiotherapy: Planning and treatment*, chapter 2. CRC Press, Boca Raton, Florida, 2013.

[46] A. Kumarasiri, F. Siddiqui, C. Liu, R. Yechieli, M. Shah, D. Pradhan, H. Zhong, I. J. Chetty, and J. Kim. Deformable image registration vased automatic CT-to-CT contour propagation for head and neck adaptive radiotherapy in the routine clinical setting. *Med Phys*, 41(12):121712, 2014.

[47] B. Rigaud, A. Simon, J. Castelli, M. Gobeli, J.-D. Ospina Arango, G. Cazoulat, O. Henry, P. Haigron, and R. de Crevoisier. Evaluation of deformable image registration methods for dose monitoring in the head and neck radiotherapy. *Biomed Res Int*, 2015:726268, 2015.

CHAPTER 9

Anatomically plausible models and quality assurance criteria for online mono- and multi-modal medical image registration

"It's not what happens to you, but how you react to it that matters."
- Epictetus

This chapter is based on:
C. Zachiu, B. Denis de Senneville, C. T. W. Moonen, B. W. Raaymakers and M. G. Ries. Anatomically plausible models and quality assurance criteria for online mono- and multi-modal medical image registration. *Phys. Med. Biol.*, 2018. [In press]

Abstract

Medical imaging is currently employed in the diagnosis, planning, delivery and response monitoring of cancer treatments. Due to physiological motion and/or treatment response, the shape and location of the pathology and organs-at-risk may change over time. Establishing their location within the acquired images is therefore paramount for an accurate treatment delivery and monitoring.
A feasible solution for tracking the pathology and the organs-at-risk over the course of an image-guided cancer treatment, can be provided by image registration algorithms. Such methods are, however, frequently adopted from optical imaging and applied on medical data with little to no modification. Since many of these algorithms are not adapted to anatomical imaging, the anatomical plausibility of the estimated deformations is frequently not guaranteed. In this chapter we adapt two existing variational registration algorithms, namely Horn-Schunck and EVolution, to soft tissue tracking. This is achieved by enforcing an incompressibility constraint on the estimated deformations during the registration process. The existing and the modified registration methods were comparatively tested against several quality assurance criteria on abdominal *in-vivo* MR and CT data. These criteria included: the Dice Similarity Coefficient (DSC), the Jaccard index, the target registration error (TRE) and three other criteria proposed in the scope of this study, evaluating the anatomical plausibility of the estimated deformations.
Results demonstrated that both the original and the modified registration methods have similar registration capabilities in high-contrast areas, with DSC and Jaccard index values predominantly in the 0.8 - 0.9 range and an average TRE of 1.1 - 1.2 mm. In contrast-devoid regions of the liver and kidneys, however, the three proposed quality assurance criteria have indicated a considerable improvement of the anatomical plausibility of the deformations estimated by the incompressibility-constrained methods. In addition, the proposed registration models maintain the potential of the original methods for online image-based guidance of cancer treatments.

9.1 Introduction

Advances in medical imaging technology have led to an improvement of cancer treatment, by offering the possibility to non-invasively localize the tumor and analyze its physical and functional characteristics.[1] With the introduction of imaging devices within the operating room in particular, medical images also began to play an important role for treatment guidance.[1,2] For treatments of long duration or repetitive interventions, the pathology and the organs-at-risk

(OARs) may change their position and shape over the course of the treatment, due to physiological motion and/or the effects of the treatment.[3–5] Therefore, in order to ensure an accurate treatment delivery and monitoring, tracking the displacement and deformation of the regions-of-interest within the images is particularly important.[3,6,7] In clinical practice, such tracking is frequently performed manually by experienced clinicians.[4,8] With the widespread availability of imaging systems, resulting in a growing volume of image data acquired using several modalities, manual tracking becomes increasingly cumbersome, time consuming and error-prone. Moreover, for online therapy guidance manual tracking is hardly feasible, since it requires tracking within time intervals of several minutes down to hundreds or tens of milliseconds.[5,9,10] Thus, an automatic solution which is capable of tracking displacements/deformations across the images acquired over the course of an image-guided cancer treatment would be highly beneficial.

A feasible solution for automatic tracking of the pathology and OARs is provided by image registration,[11] initially proposed for tracking objects within optical video sequences (see[12] for an overview). A broad range of these methods were also adopted for medical image registration, most with little to no modification.[11–13] A category of registration methods particularly suitable for medical image registration are variational methods.[14] This is due to their fast numerical schemes, low number of input parameters and their capability of estimating dense and elastic deformations. As a functioning principle, variational registration methods aim to find the deformation which optimizes a cost function dependent on the similarity between two images.[15] The most elementary term of cost function is typically known as a similarity metric and compares the difference of common image features during the registration process. Optimizing the similarity metric alone, however, is usually an ill-posed problem with an infinity of solutions. In order to address this, constraints on the estimated deformations of the tracked structures have to be included in the registration model, referred to as regularization.[15,16] An example of such a cost function is the conservation of image intensity under deformation as the similarity metric combined with the constraint of local smoothness of the deformation as the regularization, proposed in the seminal work of Horn & Schunck.[17] By penalizing large spatial gradients in the estimated deformations with the regularization term, the underlying deformations of the tracked structures are assumed to be spatially smooth. This basically hypothesizes that neighboring points within an image usually move together (or at least in a similar fashion) and penalizes motion patterns such as shearing. The Horn-Schunck algorithm has been adopted successfully for medical image registration in several independent studies.[18–21] Although the algorithm suggested by Horn & Schunck was originally not developed for medical image registration and thus does not specifically take into account the material properties of biological tissues, it has been shown to be a reliable performer for the mono-modal image registration of soft-tissue structures. However, the Horn-Schunck algorithm suffers from several limitations: The similarity metric of image intensity conservation employed by the algorithm

is only directly suitable for images acquired with the same image modality and the same image contrast. Furthermore, even for mono-modal image registration, intensity fluctuations not related to deformation/displacement, such as image artifacts or noise, in-flow artifacts due to arterial pulsations, contrast changes due to the presence/arrival of contrast agents, [etc.], are frequently incorrectly interpreted by the algorithm as genuine displacement.[18,22,23] This is due to the hypothesis that any spatial location of an image point does not change it's gray-level intensity except if it is displaced or deformed.

From an application point of view, the inability to perform a medical image registration across different imaging modalities and/or contrasts is particularly penalizing for the planning and the guidance of image-guided cancer treatments, which frequently require the registration of several different imaging modalities (such as CT, cone-beam CT, MRI, PET) and/or image contrasts to a common frame of reference.[2,24] This shortcoming motivated recently the development of several new algorithms that have this capability, such as MIND,[25] SeSaMI[26] and EVolution.[27]

The recently proposed EVolution algorithm is hereby derived directly from previously suggested variational image registration algorithms based on optical-flow, but uses in the cost function a similarity metric that hypothesizes the conservation of contrast patterns between the images rather than image intensity. The rationale is hereby that while image intensities are generally not conserved across different image modalities/contrasts, the principal contrast patterns of soft-tissue structures, tissue-bone and tissue-air interfaces are generally preserved. It is important to understand that for both the Horn-Schunck and the EVolution algorithm, high contrast areas (e.g. organ boundaries) provide sufficient structural information for the respective similarity metric, while the deformation constraint present in the regularization is playing only a secondary role. In homogenous iso-intensity areas however, the registration process is entirely driven by the regularization as the assumed constraint of the underlying true anatomical deformation. While the seminal work on both the Horn-Schunck and the Evolution algorithm evidenced that the constraint of a smooth deformation as the regularization for the registration process is generally a sensible choice, it should be noted that it is neither the only sensible choice, nor does it necessarily represent a generally valid assumption: While for example smoothness as a constraint of the estimated motion pattern correctly penalizes shearing motion within organs and muscles, it also does so at organ boundaries, where this type of motion patterns frequently occurs, in particular in-between soft-tissue interfaces subject to cardiac or respiratory motion. In the context of the original Horn-Schunck algorithm, the problem to address motion discontinuities at object boundaries did motivate several proposed amendments to the constraints[28–32] or to employ a spatio-temporal regularization.[33–35]

A rather interesting approach was thereby be pursued by Yang *et al*,[36] in the context of analyzing satellite image observations of ocean currents, with a variational based on the optical-flow principle. In this work, the regularization of the original optical flow implementation of Horn & Schunck, also included a

term enforcing incompressibility of the matter under observation. The idea to use incompressibility as a sensible constraint for medical image registration has also been employed for a variety of different algorithms. Rohlfing *et al*,[37] for example, enforce an incompressibility constraint on a B-spline-based deformation model and employed in the context of automatic estimation of breast tumor deformations. The intended incompressibility effect was attained, with a typical post-registration volume change of the tumors of $\sim 1\%$. A drawback of the approach is that the B-spline transformation model itself may have limited anatomical justification for estimating deformations of biological soft-tissues. This model also only allows the estimation of locally affine deformations for a limited set of control points, with the displacements in between the control points being interpolated according to the deterministic B-spline functions. Moreover, the similarity metric employed in the study of Rohlfing *et al* aims to find the deformation which maximizes the mutual information (MI) between the registered images. MI is inherently a global measure, which makes it particularly suitable for estimating rigid displacements, becoming challenging on the other hand to estimate deformations on a local level. A different approach was proposed by Mansi *et al*,[38] which relied on a pre-existing logDemons framework.[39] The former method, called iLogDemons, integrated tissue incompressibility in its formulation by constraining the demon-based deformation fields to be divergence-free. The approach was employed for monitoring 3D cardiac strain on cine magnetic resonance (MR) images. Compared to the original logDemons method, results have shown an improvement in the accuracy and precision of the deformation estimates when employing the iLogDemons approach. One of the downsides of the method is that demon-based registration methods are suitable for registering images acquired with the same imaging modality and contrast. Also, the method implied rather large computational latencies, limiting its employment for online and/or (near-) real-time treatment/therapy guidance. Additional similar solutions have been proposed in prior art, which further build upon the frameworks established by Rohlfing *et al* and Mansi *et al*.[40–42]

In this chapter, (I) we propose to evaluate incompressibility as a constraint to regularize the Horn-Schunck variational for medical image processing, which has to the authors knowledge so far not been thoroughly investigated in the context of medical imaging. In addition, while the original Horn-Schunck algorithm with its preservation of image intensity inherently leads to motion vector fields which conserve the material/structure under observation, the derived EVolution algorithm, which relies on the assumption of image contrast conservation, does not necessarily do so. As shown in the seminal paper,[27] this can be an advantage, since this allows the EVolution algorithm also to address registration problems of structures, which clearly change in volume, for example the bladder. On the other hand, this might potentially be disadvantageous for the registration of soft-tissue structures like the liver or the kidneys, which generally preserve their volume in-between observations. As a consequence, (II) we also propose a modified variational algorithm based on EVolution, which uses tissue

incompressibility as the constraint for the regularization. All four algorithms, the modified Horn-Schunck variational, the modified EVolution variational and their seminal counterparts are compared with respect to their performance for both mono-modal and cross-modal medical image registration. Particular attention was paid to the differences in performance between the smoothness-regularized and the incompressibility-regularized formulations.
Furthermore, it has to be emphasized that for a thorough evaluation of these registration algorithms previously established matching criteria, such as the dice similarity coefficient (DSC), Jaccard index or the target registration error (TRE), are certainly mandatory but not necessarily sufficient. The DSC, Jaccard index or the TRE, are very sensitive to evidence image differences in high contrast regions such as organ boundaries or embedded vessel structures, but lack sensitivity in areas which are essentially devoid of contrast. As a consequence, we hypothesize that the impact of the modified regularization terms, which are particularly dominant in these contrast devoid regions, require additional criteria for benchmarking. Therefore, we (III) propose the mapping of the Jacobian determinant and the curl magnitude of the resulting motion fields and an independent evaluation with spin-tagged MRI images[43,44] as additional evaluation criteria, to assess the anatomical plausibility of the resulting image registrations.

9.2 Method description

9.2.1 Description of the investigated registration algorithms

9.2.1.1 The original Horn-Schunck and EVolution algorithms

The Horn-Schunck algorithm estimates the deformation between two images I_1 and I_2 as the minimizer of the following functional:[17]

$$E_{HSO}(\mathbf{u}) = \sum_{\vec{r}\in\Omega} \left(I_1(\vec{r}) - I_2(\vec{r}+\mathbf{u}(\vec{r}))\right)^2 + \alpha\|\vec{\nabla}\mathbf{u}(\vec{r})\|_2^2 \tag{9.1}$$

where $\mathbf{u}$ is the 2D/3D displacement, Ω is the image domain, $\vec{r}$ is the image location, $\vec{\nabla}$ is the spatial gradient operator, $\|\cdot\|_2$ is the Euclidean norm and α is an algorithm input parameter. As previously mentioned, the Horn-Schunck functional is composed of 2 terms: a data fidelity term and a regularization term. The data fidelity term is a measure of similarity between the images, which for this particular algorithm is defined as the sum of squared differences between the image gray-level intensities. From a physical point-of-view, such a data fidelity term is built on the assumption that a voxel, as it moves, preserves its gray-level intensity. The regularization term on the other hand, consists of the square of the Euclidean norm applied on the gradient of the displacements, constraining the estimated deformations to be spatially smooth.

The input/regularization parameter α controls the degree of smoothness: the higher the value of α, the smoother the estimated deformations will be.
The EVolution algorithm[27] replaces the intensity-based data fidelity term of the Horn-Schunck functional, with a term dependent on the spatial gradients of the registered images. Therefore, the deformation between two images I_1 and I_2 is obtained as the minimizer of the following functional:

$$E_{EVO}(\mathbf{u}) = \sum_{\vec{r}\in\Omega} e^{-C(\mathbf{u}(\vec{r}))} + \alpha\|\vec{\nabla}\mathbf{u}(\vec{r})\|_2^2 \tag{9.2}$$

where:

$$C(\mathbf{u}(\vec{r})) = \frac{\sum_{\vec{s}\in\Gamma} |\vec{\nabla} I_1(\vec{s}) \cdot \vec{\nabla} I_2(\vec{s}+\mathbf{u}(\vec{s}))|}{\sum_{\vec{s}\in\Gamma} \|\vec{\nabla} I_1(\vec{s})\|_2 \|\vec{\nabla} I_2(\vec{s}+\mathbf{u}(\vec{s}))\|_2} \tag{9.3}$$

with Γ being a symmetric neighborhood around $\vec{r}$. All the other terms preserve their meaning from Eq. 9.1. For a better understanding of its principle, the term in Eq. 9.3 can be re-written under the following shape:

$$C(\mathbf{u}(\vec{r})) = \frac{\sum_{\vec{s}\in\Gamma} w_{\mathbf{u}}(\vec{s}) |cos(\Delta\phi_{\mathbf{u}}(\vec{s}))|}{\sum_{\vec{s}\in\Gamma} w_{\mathbf{u}}(\vec{s})} \tag{9.4}$$

where $w_{\mathbf{u}}(\vec{s})$ and $\Delta\phi(\vec{s})$ are computed based on the magnitude M and orientation ϕ of the image spatial gradients:

$$\begin{aligned} w_{\mathbf{u}}(\vec{s}) &= M_{I1}(\vec{s}) M_{I2}(\vec{s}+\mathbf{u}(\vec{s})) \\ \Delta\phi_{\mathbf{u}}(\vec{(}s)) &= \phi_{I1}(\vec{s}) - \phi_{I2}(\vec{s}+\mathbf{u}(\vec{s})) \end{aligned} \tag{9.5}$$

Basically speaking, the weight $w_{\mathbf{u}}(\vec{s})$ favors strong gradients that occur in both of the images. As a consequence, the registration process is driven only by structures which are common in the two images. The term $|cos(\Delta\phi_{\mathbf{u}}(\vec{s}))|$, on the other hand, favors deformations which align both parallel and anti-parallel gradients. In effect, this renders the EVolution algorithm robust to any possible contrast reversals between the aligned images. Note that the regularization term of the EVolution functional remains unchanged compared to the Horn-Schunck algorithm. Therefore, the spatial smoothness constraint is maintained on the estimated deformations.

9.2.1.2 The Horn-Schunck and EVolution algorithms, with modified regularization

It is a known fact from continuum mechanics that, when an incompressible material is subjected to an external force, the Jacobian determinant of its deformation is equal to 1.[37,45,46] Thus, in order to constrain the deformations estimated by the Horn-Schunck and the EVolution algorithms to be incompressible, deviations from unity of the local Jacobian determinant of the deformations are

penalized within the regularization term of their respective functionals. Registration is hereby performed via the following optimization problems:

$$E_{HSI}(\mathbf{u}) = \sum_{\vec{r}\in\Omega} \left(I_1(\vec{r}) - I_2(\vec{r}+\mathbf{u}(\vec{r}))\right)^2 + \alpha\|J(\vec{r}+\mathbf{u}(\vec{r})) - 1\|_2^2 \tag{9.6}$$

for the incompressibility-regularized Horn-Schunck algorithm and:

$$E_{EVI}(\mathbf{u}) = \sum_{\vec{r}\in\Omega} e^{-C(\mathbf{u}(\vec{r}))} + \alpha\|J(\vec{r}+\mathbf{u}(\vec{r})) - 1\|_2^2 \tag{9.7}$$

for the incompressibility-regularized EVolution algorithm. In both Eq. 9.6 and 9.7, J is defined as:

$$J(\vec{r}+\mathbf{u}(\vec{r})) = \begin{vmatrix} 1+\frac{\partial u_1(\vec{r})}{\partial x} & \frac{\partial u_1(\vec{r})}{\partial y} & \frac{\partial u_1(\vec{r})}{\partial z} \\ \frac{\partial u_2(\vec{r})}{\partial x} & 1+\frac{\partial u_2(\vec{r})}{\partial y} & \frac{\partial u_2(\vec{r})}{\partial z} \\ \frac{\partial u_3(\vec{r})}{\partial x} & \frac{\partial u_3(\vec{r})}{\partial y} & 1+\frac{\partial u_3(\vec{r})}{\partial z} \end{vmatrix} \tag{9.8}$$

where $\mathbf{u} = (u_1, u_2, u_3)$ are the 3D displacements.

9.2.1.3 Optimization of the motion estimation functionals

Since the functionals illustrated in Eq. 9.1, 9.2, 9.6 and 9.7 are differentiable, their optimization was carried-out by solving the associated Euler-Lagrange equations. This resulted in all cases in a non-linear system of equations which was resolved using a fixed-point scheme similar to the one described in.[27]

9.2.1.4 Coarse-to-fine strategy

In order to reduce the risk of the optimization scheme converging towards a local minima, the input images were processed using a coarse-to-fine strategy. This basically iterated the optimization scheme from a 16-fold downsampled version of the images up to the original resolution, with a resolution upsampling factor of 2. The displacements estimated at one resolution level were used as initial values for the next resolution level.

9.2.1.5 Employed convergence criterion

The fixed-point scheme used to solve the Euler-Lagrange equations, is an iterative procedure. Similar to,[27] the iterations were stopped when the average absolute difference between two successive iterates of the motion field was smaller than 10^{-3} voxels/pixels.

9.2.2 Experimental setup

The original and the re-formulated Horn-Schunck and EVolution registration methods were evaluated in two complementary studies: one conducted on a dataset comprised of 5 pairs of 3D MR and/or CT images, and one on spin-tagged MR images. The upcoming sections provide details on these two studies, including the quality assessment criteria used to evaluate the investigated registration methods.

9.2.2.1 Evaluation of the registration methods on 3D CT and MR data

a Description of the selected datasets The Horn-Schunck and the EVolution algorithms, in both their original and modified versions, were initially employed for aligning five pairs of 3D medical images. This included the following: (1) A pair of phase-cycled gradient-recalled turbo field echo T1-weighted (T1w) MR images, (2) A pair of CT images, (3) A dataset comprised of a T1w and a T2-enhanced (T2e) steady-state gradient echo MR image, (4) A CT and a cone-beam CT (CBCT) image and (5) A dataset which included a CT and a turbo spin echo T2-weighted (T2w) MR image. Datasets (1) and (3) were acquired on the abdomen of a healthy volunteer, while datasets (2), (4) and (5) were obtained from a patient suffering from kidney cancer [1]. Except for dataset (5), whose field-of-view only includes the kidney, all the other images include both the liver and the kidneys. The acquisition parameters for the MR images are reported in Table 9.1. Note that the T1w MR images from datasets (1) and (3) were acquired using the same MR sequence. The CT and the CBCT images had a matrix and a voxel size of $512 \times 512 \times 123$ and $1.1 \times 1.1 \times 3\text{mm}^3$, respectively. In order to ensure spatial consistency, all images were mapped prior to registration onto common lattice of size $256 \times 256 \times 128$ with an isotropic voxel volume of $1.5 \times 1.5 \times 1.5\text{mm}^3$.

In a complementary experiment, the four algorithms were also used to register one of the datasets from the DIR-Lab database.[47] The selected dataset was a pair of CT images acquired on an esophagus cancer patient, at the two extreme phases of the respiratory cycle. Each image had a size of $256 \times 256 \times 94$ and a $0.97 \times 0.94 \times 2.5\text{mm}^3$ voxel volume.

b Quality assessment criteria for the estimated deformations The five datasets described in the previous paragraph were registered using the four investigated registration methods. Since the data fidelity term of the Horn-Schunck algorithm in both its original and modified version is incompatible with cross-contrast and cross-modality registration, the two methods were only employed for the first two datasets (i.e. for T1w MR to T1w MR and CT to CT registration).

[1] Ethical approval for the acquisition, usage and publication of the data was provided by the Ethics Board of the UMC Utrecht, Utrecht, The Netherlands.

Sequence	T1w-MR [Dataset (1)]	T2e-MR [Dataset (3)]	T2w-MR [Dataset (5)]
TR [ms]	4.3	3	523
TE [ms]	2	1.5	80
FA$^\circ$	10	40	90
Volume size	$192 \times 192 \times 75$	$192 \times 192 \times 75$	$288 \times 288 \times 53$
Voxel size [mm^3]	$2 \times 2 \times 2$	$2 \times 2 \times 4$	$1.4 \times 1.4 \times 3$
$\mathrm{BW}_{\mathrm{readout}}$ [Hz]	373	723	478

Table 9.1: Acquisition parameters for the 3D MR images used in the evaluation of the four investigated registration methods. Distinct MR acquisition sequences were used for the T1w image from dataset (1) (2^{nd} column), the T2e image from dataset (3) (3^{rd} colum) and the T2w image from dataset (5) (4^{th} colum). The used abbreviations for the acquisition parameters are as follows: TR - repetition time, TE - echo time, FA - flip angle and BW - bandwidth.

The performance of the registration methods was evaluated against several criteria. In an initial step, two typically-employed quality assessment metrics were analyzed: the Dice Similarity Coefficient (DSC)[48] and the Jaccard index.[49] The two criteria were used to determine the contour overlap for the liver and kidneys in the images, before and after registration. Such an analysis showcases the capability of the investigated registration algorithms for organ contour alignment, which is representative for their performance in high-contrast areas. Mathematically, the DSC and the Jaccard index are defined as follows:

$$DSC(A,B) = \frac{2 * |A \cap B|}{|A| + |B|}; \qquad \mathrm{Jaccard}(A,B) = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{9.9}$$

where A and B are the two contours whose overlap is determined and $|\cdot|$ corresponds to the number of voxels within a contour. Note that for all images in the initial datasets, delineations of the liver and kidneys were made available by an experienced radiologist.
The performance of the algorithms in high contrast areas was further evaluated by comparing the estimated and the manually determined locations of 1280 features identified by the DIR-Lab database provider in the dataset analyzed in the scope of this study. The evaluation was performed in terms of the target registration error (TRE):[50]

$$\mathrm{TRE[mm]} = \sqrt{(x_I + u_1(\vec{r}_I) - x_J)^2 + (y_I + u_2(\vec{r}_I) - y_J)^2 + (z_I + u_3(\vec{r}_I) - z_J)^2} \tag{9.10}$$

where (x_I, y_I, z_I) and (x_J, y_J, z_J) are the gold standard landmark coordinates in the reference and the moving image, and $(u_1(\vec{r}_I), u_2(\vec{r}_I), u_3(\vec{r}_I))$ is the estimated 3D displacement for that particular landmark.
In order to evaluate the validity of the deformations in contrast-devoid regions, the local Jacobian determinant (see Eq. 9.8) and the local curl magnitude of the

estimated deformations were analyzed with respect to the underlying anatomy. The Jacobian determinant is used as an indicator whether the estimated deformations are in good correspondence with the incompressibility property of biological soft tissues. The local vorticity of the deformations, provided by the curl magnitude, is used as an independent error criterion which is not explicitly enforced by any of the investigated motion models. Mathematically, the curl of a 3D vector field is computed as:

$$\mathrm{curl}(\vec{\mathbf{u}}) = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ u_1 & u_2 & u_3 \end{vmatrix} \tag{9.11}$$

where $\vec{\mathbf{u}} = (u_1, u_2, u_3)$ is the 3D vector field and $(\vec{i}, \vec{j}, \vec{k})$ are the versors of the 3D Euclidean space.

9.2.2.2 Evaluation of the investigated registration methods on spin-tagged MR images

MR-tagging is a technique, which has been historically introduced to investigate the dynamics of the myocardium for diagnostic purposes.[43,44] RF-saturation pulses are used to place low-signal stripes or grids on the images, which move together with the underlying tissue. In this chapter, this technique was used for validation purposes, since it provides a non-invasive way to acquire an *in-vivo* gold standard for organ/tissue displacement, in particular in image regions devoid of anatomical contrast.

Two pairs of MR images were acquired during this experiment: one without and one with spin-tagging. The tagging consisted in a horizontal stripe pattern with a 5 mm distance between the stripes. Each pair contained two sagittal 2D MR slices acquired on the abdomen of a healthy volunteer at the two extreme phases of the respiratory cycle, via respiratory gating.

The four investigated algorithms were initially used to register the pair of un-tagged MR images. Similar to the analysis described in section 9.2.2.1, the pre- and post-registration DSC and Jaccard index were analyzed for the liver, kidney and, for this particular experiment, also the abdominal wall. The Jacobian determinant and the curl magnitude of the deformations were also comparatively analyzed between the original and the modified versions of the Horn-Schunck and the EVolution algorithms. The deformations resulting from the registration of the un-tagged pair of images were then used to align the spin-tagged images. In the ideal case, the post-alignment tagging patterns in the two images should match. For a quantitative evaluation of the registration errors, the distance between the tag lines pre- and post-alignment was also calculated.

The MR-acquisition sequence, for both the un-tagged and tagged images, employed the following parameters: TR = 90 ms, TE = 8.8 ms, FA = 35°, image size = 224×224, voxel size = $1.8 \times 1.8 \times 5\mathrm{mm}^3$ and $\mathrm{BW}_{\mathrm{readout}} = 996$ Hz.

9.2.3 Algorithm parameter calibration and implementation

The original and the modified Horn-Shunck algorithm require the calibration of the regularization parameter α (see Eq. 9.1 and 9.6). Besides this, the EVolution algorithm in both its formulations also requires configuring the size of the additional parameter Γ (see Eq. 9.2, 9.3 and 9.7). For the experiment described in 9.2.2.1 an exhaustive search was conducted for these parameters and the values which provided the highest average post-registration DSC for all delineated structures were then chosen for use. The rationale behind such an approach is that, in a clinical setting, the output of registration algorithms is often evaluated based on a visual inspection of the post-registration organ boundary alignment between the images. Thus, maximizing the DSC is representative for such an evaluation criterion. On the other hand, for the spin-tagged datasets, the values of the parameters which provided the best match between the tag patterns after registration, were selected for use.
The numerical schemes of all four registration algorithms were implemented and executed on an nVidia Tesla K20 graphics card, using the compute unified device architecture (CUDA).

9.3 Results

9.3.1 Performance of the investigated registration methods in high-contrast areas

Table 9.2 reports the DSC and the Jaccard coefficients for the liver and the kidneys before and after registration of each of the 5 initial datasets described in section 9.2.2.1. For the remainder of this chapter the four algorithms were abbreviated as follows: HSO - the original Horn-Schunck algorithm with a constraint on smoothness, HSI - the modified Horn-Schunck algorithm with a constraint on the Jacobian determinant, EVO - the original EVolution algorithm with a constraint on smoothness and EVI - the modified EVolution algorithm with a constraint on the Jacobian determinant. All the algorithms provide high values of the post-registration DSC and the Jaccard index, predominantly in the 0.8 - 0.9 range. It can also be observed that the HSI and the EVI provide similar results with respect to their smoothness-regularized counterparts.
Table 9.3 showcases the TRE obtained on the DIR-Lab dataset (see section 9.2.2.1 for details). It can be observed that improvements were achieved in all cases, with a registration accuracy of 1.1 - 1.2 mm and a sub-millimeter precision. Only marginal differences can be noticed between the TREs calculated for the original Horn-Schunck and EVolution methods compared to their versions with modified regularization.

Criterion	Dataset	Organ	Algorithm				
			None	HSO	HSI	EVO	EVI
DSC	T1w MR - T1w MR	Liver	0.89	0.96	0.96	0.95	0.95
		Kidneys	0.87	0.92	0.90	0.91	0.90
	CT - CT	Liver	0.9	0.91	0.91	0.91	0.91
		Kidneys	0.84	0.88	0.87	0.91	0.91
	T1w MR - T2e MR	Liver	0.85	-	-	0.91	0.91
		Kidneys	0.63	-	-	0.85	0.85
	CT - CBCT	Liver	0.9	-	-	0.91	0.91
		Kidneys	0.84	-	-	0.91	0.91
	CT - MR	Liver	-	-	-	-	-
		Kidneys	0.74	-	-	0.91	0.91
Jaccard Index	T1w MR - T1w MR	Liver	0.80	0.92	0.92	0.90	0.90
		Kidneys	0.77	0.85	0.81	0.83	0.83
	CT - CT	Liver	0.82	0.84	0.84	0.83	0.83
		Kidneys	0.73	0.79	0.77	0.83	0.83
	T1w MR - T2e MR	Liver	0.73	-	-	0.84	0.84
		Kidneys	0.46	-	-	0.74	0.74
	CT - CBCT	Liver	0.82	-	-	0.83	0.83
		Kidneys	0.73	-	-	0.83	0.83
	CT - MR	Liver	-	-	-	-	-
		Kidneys	0.59	-	-	0.83	0.83

Table 9.2: Evaluation of the performance of the four investigated algorithms in high contrast areas. The table reports the pre- and post-registration DSC and the Jaccard index for the liver and the kidneys following the registration of the five datasets containing 3D MR and CT images. The missing entries either indicate the absence of a delineation for a particular anatomical structure or the fact that a method was incompatible with a particular data set. The abbreviations for the algorithms are provided in the text.

Method	TRE [mm]
None	4.01 ± 2.91
Horn-Schunck Original	1.13 ± 0.68
Horn-Schunck Incompressible	1.21 ± 0.70
EVolution Original	1.12 ± 0.65
EVolution Incompressible	1.19 ± 0.83

Table 9.3: Accuracy and precision of the four investigated algorithms with respect to manually tracked landmarks. The table reports the target registration errors associated to the methods, when applied on the DIR-Lab dataset. The errors are reported under the format *mean* $\pm$ *standard deviation*.

9.3.2 Compression/Expansion and vorticity of the deformations provided by the investigated registration methods

Figure 9.1 displays the spatial distribution of the Jacobian determinant of the deformations provided by the four registration methods, in a sagittal slice selected from each of the five registered datasets. Following a visual inspection, strong fluctuations from unity of the Jacobian determinant can be observed for the HSO and for EVO, in particular within homogeneous regions of the liver and kidneys. These fluctuations are considerably reduced by HSI and EVI, for which the Jacobian determinant is notably closer to 1, with only slight deviations. This is further confirmed by the statistical analysis conducted for the Jacobian determinant of the joint liver and kidney deformations illustrated in Figure 9.2. The average range of the Jacobian determinats of the liver and kindey is reduced from [0.09; 2.29] for the original methods to [0.73; 1.33] for their incompressibility-regularized counterparts. A 1-tailed Mann-Whitney-Wilcoxon test applied for each dataset, in between the original and the modified versions of the algorithms, revealed a statistically significant difference in all cases at 0.05 significance level.

Fig. 9.3 showcases the vorticity of the deformations estimated by the four registration algorithms. The magnitude of the curl of the motion vector fields is displayed for the same sagittal slices selected from the five 3D datasets, as in Fig. 9.1. Within the boundaries of the liver and kidney, except for a few problematic areas, the vorticity of the estimated deformations remains predominantly close to zero. A tendency to increase of the deformation vorticity can be observed, however, in the proximity of the thoracic, peritoneal and back wall, where sliding motion is typically expected. Figure 9.4 depicts the Jacobian determinant, the motion vectors and the curl magnitude of the deformations estimated by HSO and HSI inside in ROI 1 from figures 9.1 and 9.3. Again, it can be observed that there is a link between the values in the Jacobian and the curl maps and the motion vector field estimated by HSO, which showcases a regional tendency to diverge, with a strong vortex in the ventral part of the kidney. This indicates that the HSO estimated both stretches and swirls of the liver and kidney tissues within the selected ROI. Both of these effects are dampened by the HSI algorithm, with the estimated motion fields displaying a higher degree of uniformity.

Figure 9.5 illustrates the Jacobian determinant, the motion vectors and the curl magnitude of the deformations estimated by EVO and EVI inside ROI 2 from figures 9.1 and 9.3. The illustrated motion field provides a more refined interpretations of the Jacobian determinant and the curl magnitude maps. Unnatural divergences and convergences can be observed within the motion vector field, in areas corresponding to deviations of the Jacobian determinant from unity. These, in turn, also introduce a particular degree of vorticity within the deformations, which are the cause of the peak values observed in the illustrated curl map. The deformations estimated by EVI, on the other hand, showcase

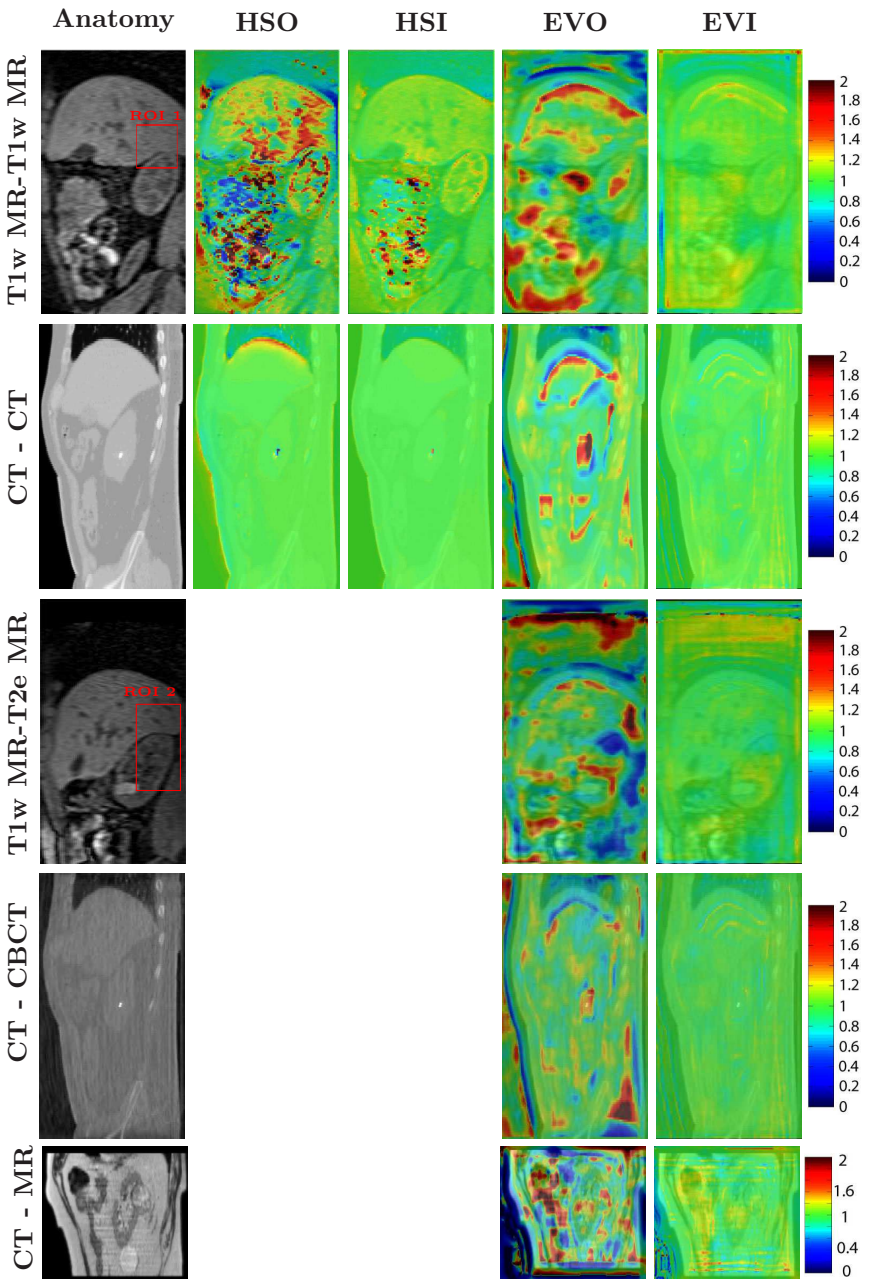


Figure 9.1: Spatial distribution of the Jacobian determinant. First column: A sagittal slice selected from each of the five registered datasets. Second to fifth column: Spatial distribution of the Jacobian determinant of the deformations provided by each of the four investigated methods overlapped with the slice from the first column. ROI 1 and 2 are two selected problematic areas, for which a more detailed analysis will be conducted.

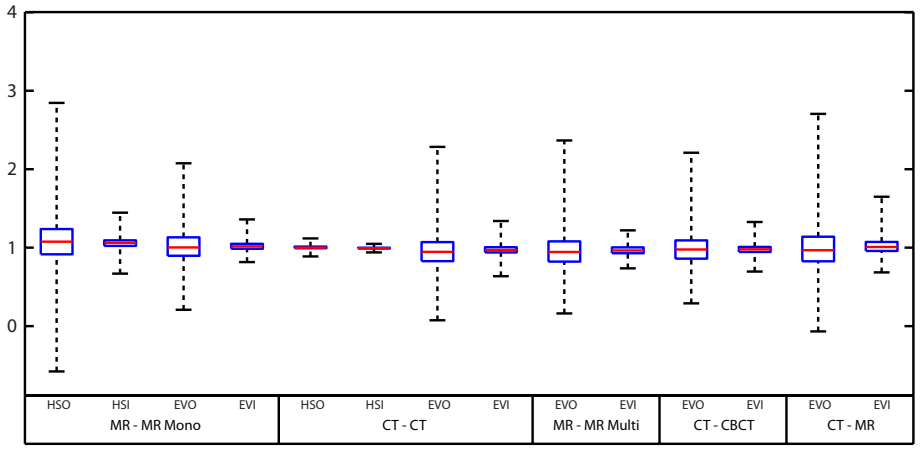


Figure 9.2: Statistical distribution of the Jacobian determinant of the liver and kidney deformations estimated by the four registration methods. Evaluation was performed for each of the five registered 3D datasets. The illustrated boxplots are constructed as follows: the box limits are the 25^{th} and the 75^{th} percentiles, the whiskers correspond to the 5^{th} and the 95^{th} percentiles and the red line is the set median. The differences between the original and modified versions of the Horn-Schunck and EVolution algorithms, respectively, are statistically significant at 0.05 significance level.

a higher degree of uniformity, with no notable peaks in either the Jacobian determinant nor the curl magnitude map.

9.3.3 Absolute registration errors of the investigated algorithms

The images in figure 9.6 illustrate the two pairs of images described in section 9.2.2.2, used to determine the accuracy of the four registration algorithms. Figures 9.6(a) and 9.6(b) showcase the first pair of images acquired on the healthy volunteer, without applying spin-tagging. Figure 9.6(c) and 9.6(d) illustrate the other pair of 2D scans acquired at the same respiratory phases as the images in 9.6(a) and 9.6(b), in the presence of tagging. The tag pattern allows to visually determine the trajectory followed by the spin-tagged tissue, as the volunteer went from inhalation to exhalation.

As described in section 9.2.2.2, all of the four registration algorithms were used to estimate the displacements between the images in figures 9.6(a) and 9.6(b). Table 9.4 reports the pre- and post-registration DSC and Jaccard index for three ROIs: liver, kidney and the thoracic/peritoneal wall. Overall high values

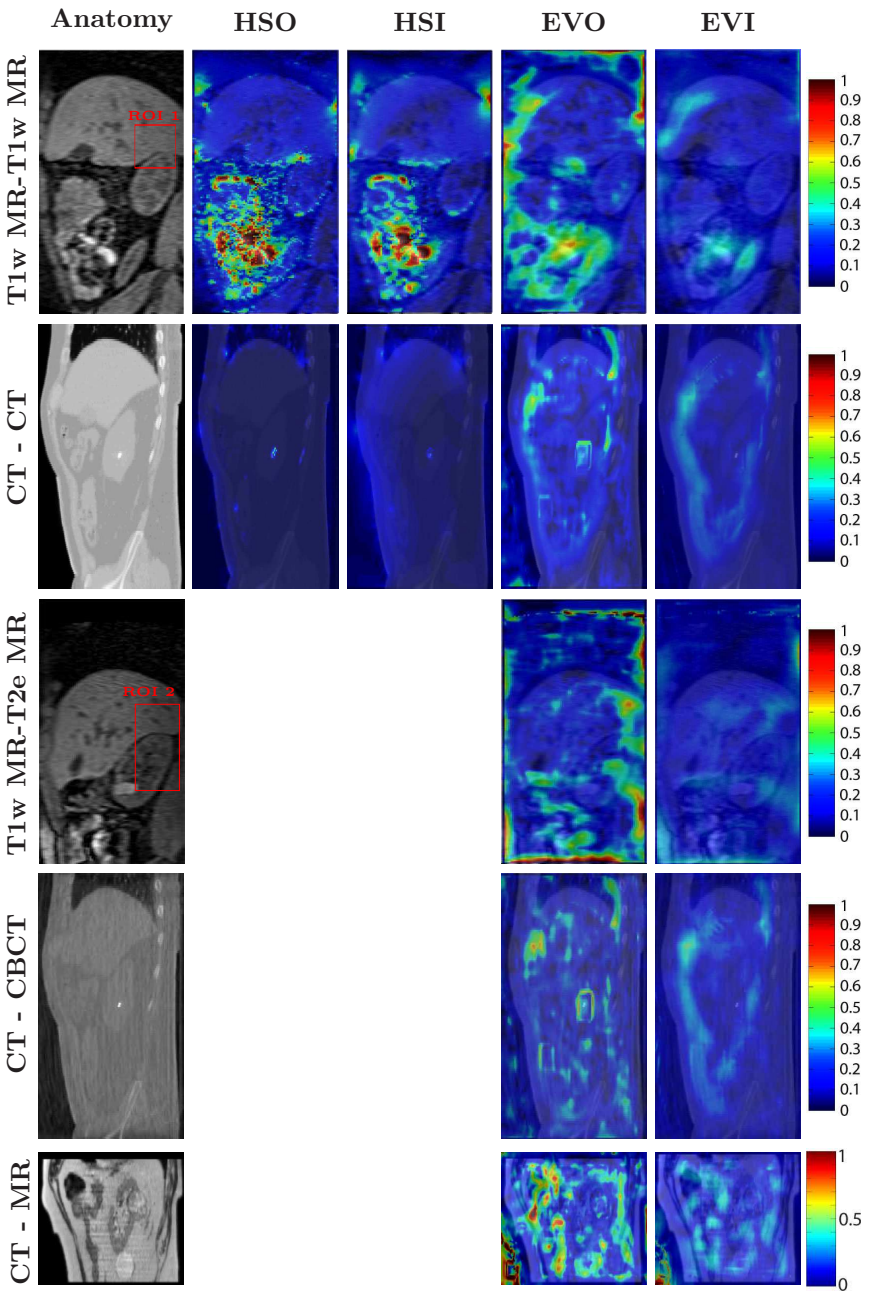


Figure 9.3: Vorticity of the deformations estimated by the 4 registration algorithms. The first column depicts a sagittal slice selected from each of the 5 registered 3D datasets. Columns 2 to 4 showcase the curl magnitude of the motion vector fields estimated by the four methods overlaid on the sagittal slice displayed in the first column. ROI 1 and 2 are the same selected problematic areas indicated in Fig. 9.1.

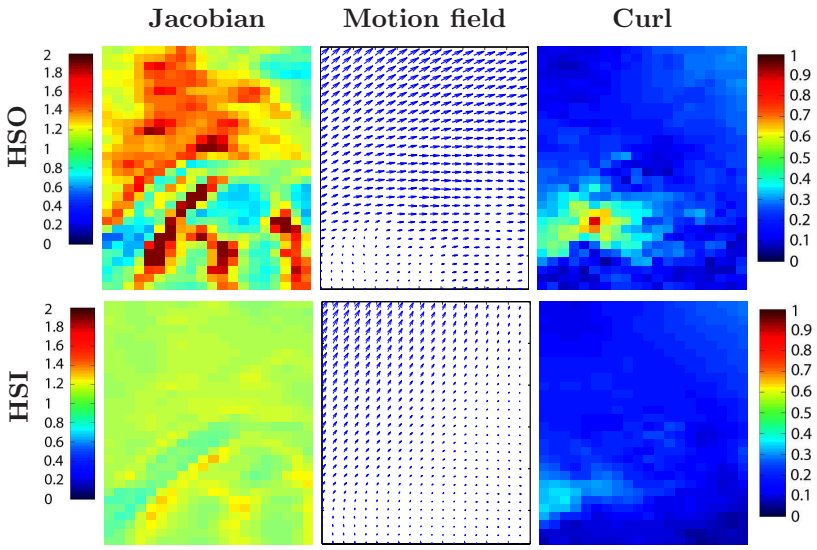


Figure 9.4: The Jacobian determinant (first column), motion vector field (second column) and curl magnitude map (third column) of the deformations estimated by HSO (first row) and HSI (second row) in ROI 1 from Fig. 9.1 and 9.3.

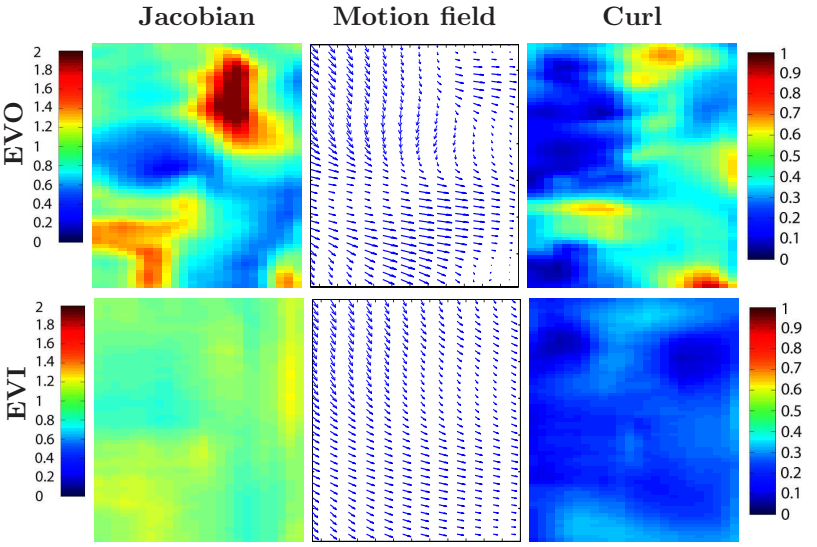
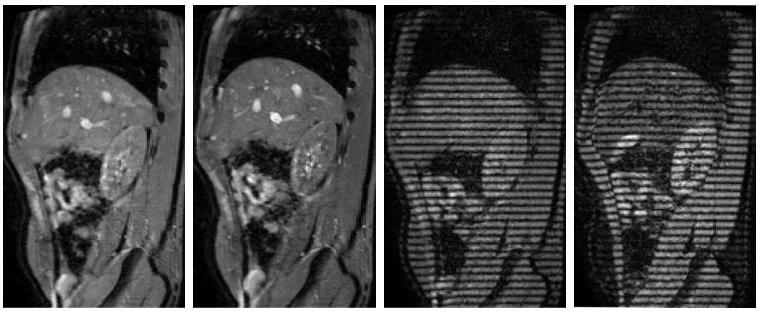


Figure 9.5: The Jacobian determinant (first column), motion vector field (second column) and curl magnitude map (third column) of the deformations estimated by EVO (first row) and EVI (second row) in ROI 2 from Fig. 9.1 and 9.3.



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

Figure 9.6: The two pairs of MR images acquired on the abdomen of a healthy volunteer during the study described in section 9.2.2.2. (a), (b): In the absence of spin-tagging. (c), (d): With spin-tagging applied. The images in each pair were acquired at the two extreme phases of the respiratory cycle.

(≥ 0.9) were obtained for all three structures by all of the registration methods. Only marginal differences can be observed between the original Horn-Schunck and EVolution algorithms and their respective incompressibility-constrained counterparts.

The resulting deformations were used to align/map the tagged image in figure

Criterion	ROI	Registration method				
		None	HSO	HSI	EVO	EVI
DSC	Liver	0.90	0.98	0.98	0.97	0.97
	Kidney	0.91	0.98	0.99	0.97	0.98
	Wall	0.82	0.95	0.94	0.94	0.94
Jaccard index	Liver	0.82	0.96	0.96	0.95	0.95
	Kidney	0.83	0.96	0.97	0.95	0.96
	Wall	0.70	0.90	0.89	0.89	0.89

Table 9.4: Pre- and post-registration contour overlap for three regions-of-interest (ROIs) in the images from Fig. 9.6. The DSC and the Jaccard index is reported for the liver, kidney and thoracic/abdominal wall, for all of the four investigated registration methods (HSO, HSI, EVO and EVI). The pre-registration DSC and Jaccard index is also included.

9.6(d) to the image in figure 9.6(c). In the ideal case, the tag pattern of this mapped image should match the one in figure 9.6(c). Figure 9.7 illustrates the following: the mapped tagged images (first row), the Jacobian determinant and the curl magnitude maps of the estimated deformations (second and third row) and spatial distribution of the absolute registration errors (fourth row). Details related to the manner in which these errors were calculated can be found in

section 9.2.2.2. Each column corresponds to one of the four algorithms. It can be observed that for all of the evaluation criteria showcased in figure 9.7, the HSO and HSI provide similar results, with moderate deviations of the Jacobian determinant from unity and moderate values of the curl magnitude. On the other hand, more prominent differences can be distinguished between EVO and EVI, in particular in terms of the Jacobian determinant of the estimated deformations, with EVI showcasing Jacobian determinants considerably closer to 1 than EVO. Both EVO and EVI have curl magnitude maps containing overall larger values compared to HSO and HSI, manifesting a tendency to increase in the proximity of organ-thoracic/abdominal/back wall interfaces. A peak in the curl magnitude maps can also be observed in the cranio-ventral part of the liver, peak which is coupled with increased registration errors.
In terms of the absolute registration errors, except for EVO, all of the algorithms showcase a good accuracy of 1 - 2 mm within the liver and kidney boundaries. The EVO algorithm, while showcasing predominately low errors within the organ boundaries, it has difficulties in estimating displacements at the interface between the liver and the thoracic/peritoneal wall, with the errors exceeding 5 mm. It can also be noted that all of the methods showcase, to different extents, difficulties in estimating the displacements of the thoracic and peritoneal wall, indicated both by the high registration errors and the mapped/registered images in the first row (the tag lines should be horizontal). Out of the four algorithms, it is EVO which has the best accuracy in this area, with EVI coming-in second best.

9.3.4 Computational performance of the investigated algorithms

Table 9.5 reports the average computational performance of the four registration methods for both 2D and 3D images. Reporting is performed for images of size 224×224 and $256 \times 256 \times 128$, which correspond to the dimensions of the 2D and 3D images registered in the scope of this chapter.

Method	2D $[224 \times 224]$	3D $[256 \times 256 \times 128]$
Horn-Schunck Original	27 ms	2.7 s
Horn-Schunck Incompressible	105 ms	10 s
EVolution Original	150 ms	67 s
EVolution Incompressible	250 ms	65 s

Table 9.5: Computational time required by the four investigated registration methods.

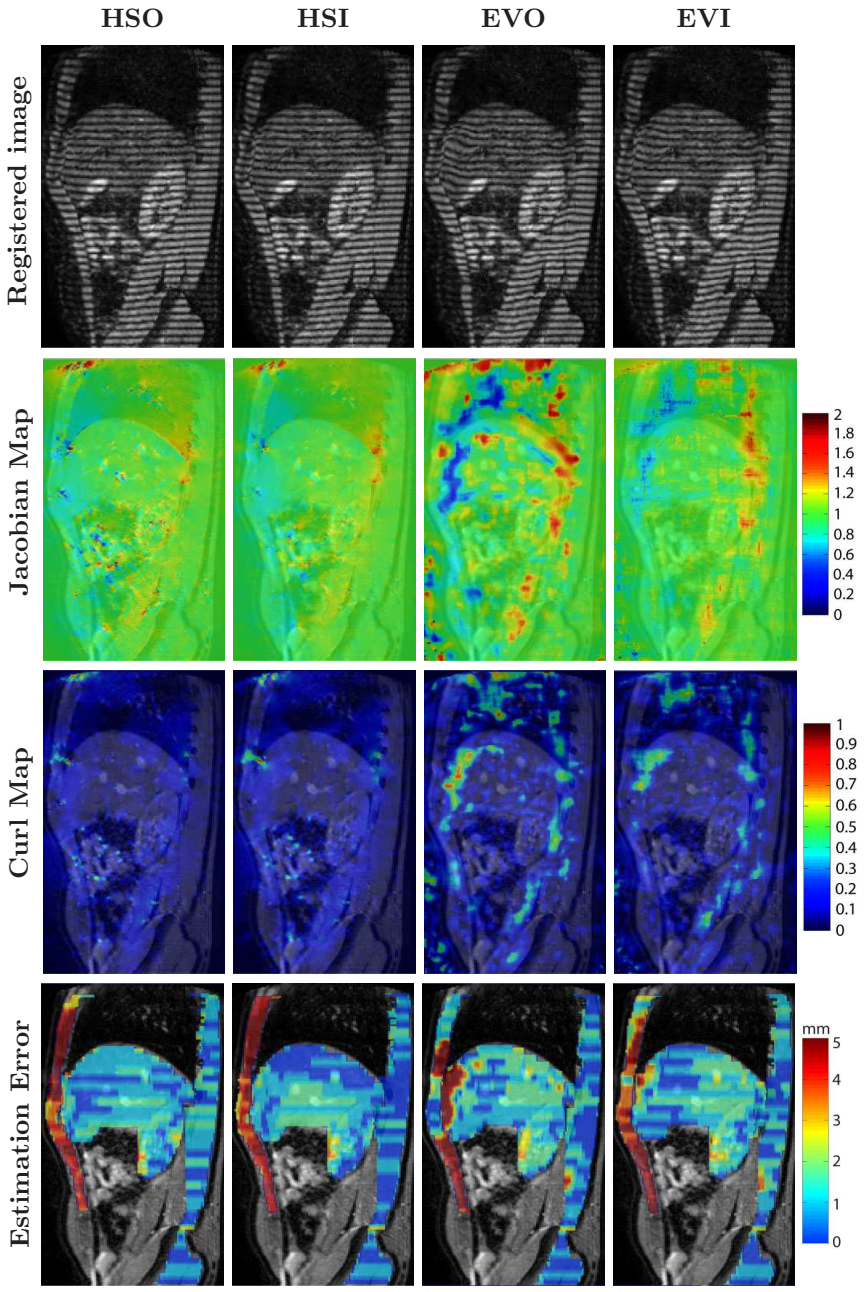


Figure 9.7: Outcome of the study conducted on the spin-tagged MR images. The four rows of images illustrate: the mapped/registered spin-tagged images, the Jacobian determinant maps, the curl magnitude maps and the spatial distribution estimation errors for each of the four registration methods.

9.4 Discussion

Evaluation criteria for the image registration algorithms
One fundamental problem that complicates the development as well as the evaluation and validation of different medical image registration algorithms is the ground truth dilemma: The true anatomical deformation of soft tissues can frequently only be assessed in contrast rich regions such as organ boundaries or in the vicinity of organ features, while the true physiological deformation in contrast devoid regions is not known. This poses a problem, since in a historical context, medical image registration developed as a side-branch of the more general field of computer vision. As a consequence, many established algorithms and evaluation criteria in the medical image registration domain as well as the associated benchmarking criteria are rooted/derived from methods developed for optical object tracking.
Three criteria that are often used for evaluating the performance of medical image registration algorithms are the Dice Similarity Coefficient (DSC), the Jaccard index and the target registration error (TRE). In clinical usage of registration algorithms, a visual inspection of the post-registration alignment of identifiable anatomical features such as organ boundaries and/or embedded vessel structures is often used to validate the estimated deformations.[51] Therefore, the DSC, the Jaccard index and the TRE are representative for the validity of the estimated deformations, in case such a visual criterion is used. Given the high 0.8 - 0.9 values obtained for the DSC and the Jaccard index on the five 3D data sets (see Table 9.2) and the 1.1 - 1.2 mm average TRE obtained on the DIR-Lab CT images (see Table 9.3), if a contour-based validation criterion would be used, one would conclude that all four registration methods provide satisfactory deformations. Note that the obtained DSC, Jaccard index and TRE values are also in good correspondence with previous reportings.[5,27,52–55]
The most obvious way to evaluate whether the estimation of the tissue deformation in contrast-devoid regions corresponds to the true physiological deformation, is to repeat the observation of the deformation process and to introduce non-invasively contrast patterns across the entire image. MR-tagging is essentially an old and well established method, which has been used for many years in the field of cardiac MRI for this purpose. MR-tagging methods, such as C-SPAMM allow to saturate the MR magnetization of a tissue under observation prior to the motion event with a contrast pattern, such as lines or a grid, and to track the subsequent elastic deformation of this pattern through the tissue deformation. The principal limitations of this MR-tagging approach are two-fold: The tagging-pattern disappears with the longitudinal relaxation time of tissue under observation. This means the tag must be applied immediately before the motion event and - considering muscle tissue - can only be observed over a duration of approx 700 - 1000 ms. The second limitation arises from the fact that tagged liquids, such as for example arterial blood, exits the

"tagged stripes" during the observation period and mixes with untagged tissue. This additionally reduces the contrast of the tag-pattern, in particular in high perfused organs such as liver and kidney. In the presented experiments, this effect is main the reason why the presented data was only MR-tagged in one dimension (i.e. a line pattern in the direction of the principal deformation).
Here, respiratory gated MR tagging has been used to evaluate the precision and accuracy of the four investigated registration methods in the abdomen under the assumption that the respiratory cycle is periodic. After registering the un-tagged pair of images, high post-registration values of the DSC and the Jaccard index were obtained in the liver, kidney and the abdominal/thoracic wall for all four of the algorithms. This indicates a good apparent performance of the algorithms in all of these areas. However, the error maps illustrated in the fourth row of figure 9.7 showcases precisely the areas in which registration was successful and areas which are problematic. While the tag lines deep inside the liver and the kidney do not evidence any obvious signs of misregistration, the thoracic wall and its interface with the liver appears to be particularly problematic for all four algorithms. This finding validates the hypothesis, that while the DSC and the Jaccard index are certainly part of a comprehensive assessment of the image registration, neither is sufficient to evaluate if the estimated deformation really corresponds to the physiological tissue deformation. However, while the tagged MR-images as the "gold-standard" permit a differential analysis of complex respiratory kinetics as shown in figure 9.6, the untagged images neither display sufficient anatomical contrast between the different tissue types, nor sufficient contrast within the tissue layers to provide the required information to analyze the complex situation correctly. Similar to the employed motion estimation algorithms, even a careful visual inspection of the entire respiratory cycle as a movie does not allow human observers to analyze the underlying tissue deformations correctly, while C-SPAMM depicts the motion pattern clearly.
Unfortunately, not all physiological forms of motion and not all medical imaging technologies lend themselves to tagging techniques like (or similar to) C-SPAMM and therefore, for many applications, the ground truth dilemma essentially remains. As stated in the introduction, we propose as the next best measure to evaluate the quality of the registration results beyond the established indices such as DSC and/or Jaccard, to investigate the properties of the estimated motion vector fields for anatomical plausibility.
Figure 9.1, 9.4, 9.5 and 9.7 display examples of the spatial distribution of the Jacobian determinant. While several parts of the human anatomy such as lung volume, heart volume, bladder volume, large parts of the digestive tract display volume changes during physiological motion processes/events, most of the inner organs, such as liver, spleen, kidney, muscle and fat layers are liquid filled tissue structures and as a consequence incompressible. Both HSO and EVO preserve to a large extent the overall volume of the liver and kidneys. However, within the organ interiors, the high deviations from unity of the Jacobian determinant are physically implausible, since the composing tissues are incompressible. Also,

as shown in figures 9.4 and 9.5 such fluctuations correspond to sinks and/or sources in the estimated deformations, which basically indicates the creation and disappearance of tissue. This is aberrant given the material properties of the observed organs and the short time interval between the acquired images.
The counter-part of the Jacobian determinant is an analysis of the vorticity of the estimated deformations under consideration of the underlying local anatomy. Although not quite as intuitive to interpret as the former, variations in vorticity are generally an indicator for shearing motion, local rotations and - similar to the Jacobian determinants - divergence of the motion fields. As a quantitative measure for vorticity, the examples in figure 9.3 showcase the spatial distribution of the curl magnitude of the deformations estimated on the five abdominal data sets by the four investigated registration methods. High values of the curl magnitude can be predominately observed in the proximity of shearing interfaces (e.g. abdominal and back wall), as expected, but also within regions in which a high vorticity is unlikely from an anatomical point-of-view. While rotations of the liver and kidney as a whole are expected, they are highly improbable to occur at finer scales.
While clearly inferior to local error maps derived from a "true goldstandard" as shown in figure 9.7, both the local Jacobian determinant as well as the vorticity are useful indicators with respect to areas where the registration process has potentially failed. For example in the registration with EVO, the low Jacobian determinant as well as the increased vorticity in the ventral part of the liver corresponds well with the independently assessed area of misregistration depicted in the bottom row. For the incompressibility-constrained methods, however, the map of the local Jacobian determinant is only useful as an indicator if convergence of the functional has been achieved. Thus, the local vorticity maps become an increasingly important indicator of local misregistrations. For example, as shown in figure 9.7, the elevated local vorticity maps for EVI clearly indicate the small area in the cranio-ventral part of the liver, which as depicted by the independent error estimation, is not correctly registered.

Performance of the original Horn-Schunck algorithm
The original Horn-Schunck algorithm, as one of the most established variational mono-modal registration methods, has been shown to be a reliable performer by several independent studies.[10,18–21,23] This reliability mostly stems from the good correspondence between the motion estimation model it employs and the physical properties of elastic soft tissues: the conservation of image-intensity as the similarity metric of the cost function maps into the conservation of matter during elastic soft-tissue deformation, while the smoothness constraint as the regularization of the cost function maps into the stiffness of soft-tissue structures.[56–59] In the scope of this chapter, the DSC and the Jaccard index provided by HSO, were among the highest of all investigated algorithms, displaying an overall robust performance. The assumption made by the HSO that a material point conserves its intensity as it moves, however, can also be a source of misregistrations, as already established in previous studies.[18,22,23] In the current

chapter, this is best illustrated for the MR - MR mono-modal registration in figure 9.1 and 9.4. Several superimposed effects and artifacts including arterial pulsations, noise and transient fold-over artifacts of the bowels, have led to violations of the voxel intensity conservation hypothesis in the dorsal part of the liver and the kidney cortex. In its attempt to redistribute image intensities such that the data fidelity term of the HSO is minimized, the algorithm ended-up estimating non-anatomical deformations which translate to local tissue expansion/creation. Similarly, for the CT-CT registration, the superior part of the lung is expanded by the algorithm due to a CT reconstruction artifact at the lung-liver interface, which varies due to a changing curvature during respiration. On the other hand, the smoothness constraint imposed by the regularization term has the tendency to dampen vortices, reason for which the curl magnitudes remain overall low to moderate.
With respect to practical considerations, the HSO algorithm allows a very efficient implementation. This is due both to the point-wise nature of the numerical scheme employed for the minimization of the cost function and the overall mathematical simplicity of the optimization process. This resulted in the HSO having the best convergence time among all four investigated registration methods.

Performance of the Horn-Schunck algorithm with modified regularization
The incompressibility-regularized formulation of the Horn-Schunck algorithm (HSI) replaces the smoothness constraint imposed by the HSO with a constraint which penalizes deviations from unity of the local Jacobian determinant of the deformations. In terms of the DSC and the Jaccard index, HSI provides comparable values to the HSO. However, the HSI significantly dampens the effect of intensity variations not attributed to motion on the quality of the estimated deformations. As illustrated in figure 9.4, the Jacobian determinant in an area which was problematic for the HSO, remains close to unity in the entire region-of-interest, with the associated motion fields uniformly pointing in the general direction of motion of the observed organs. From a physical point-of-view, the incompressibility constraint leads to a liquid-like behavior of the estimated deformations, imposing a lower penalty on shearing motion. This leads to an improved performance in the vicinity of organ/tissue boundaries. This effect can be best observed in figure 9.3, for CT - CT registration using HSI, where the curl magnitude showcases a tendency to increase in the proximity of the thoracic and abdominal wall. Note that this effect is less pronounced for HSO. On the other hand, during the experiment conducted on the tagged MR images, the differences between the HSO and the HSI methods are less noticeable. A reason for this may be that the registered images did not include violations of the intensity conservation hypothesis to an extent that would induce a differentiated behavior between the two algorithms (like, for example, in figure 9.1). In terms of computational time, the more complex numerical scheme of the HSI implied a threefold penalty compared to the HSO. Nevertheless, the resulting

computational latency is still acceptable, if necessary, in the context of online therapy guidance.
One of the drawbacks of the HSI method, as well as the HSO, is that the data fidelity term relies on the voxel intensity conservation hypothesis, property which is generally not conserved across imaging modalities. This is a severe practical limitation, given the fact that planning and guidance of image-guided cancer treatments, often rely on the registration of images acquired using different modalities and/or image contrasts.

Performance of the original EVolution algorithm
The principal advantage of the original Evolution (EVO) algorithm compared to HSO and HSI lies in its capability of performing both mono and multi-modal registration. This capability stems from the formulation of the data fidelity term, which no longer relies on the conservation of image intensity, but rather on the conservation of contrast patterns. The algorithm demonstrated good contour matching capabilities, with DSC and Jaccard indices predominantly in the 0.8 - 0.9 range. This indicates that the original EVolution algorithm is well suitable for estimating deformations of volume-changing anatomical areas, such as the lung or the bladder (as demonstrated in[27]). For incompressible organs, on the other hand, the EVO algorithm may be less favorable. In both CT and MR images, image intensity is the density weighted signal average over a given voxel. Thus, from a physical point-of-view, image intensity in such images can be seen equivalent to mass. The HSO and HSI algorithms imposed the conservation of gray-level intensity and therefore the conservation of mass in between the registered images. The EVO data fidelity term loses this property, which leads to the algorithm resolving "boundary matching" by creating sinks or sources (and therefore to the disappearance and creation of mass) in contrast devoid regions, which is anatomically implausible in incompressible organs. This effect can be observed to different extents in all of the registered data sets, where strong deviations from unity of the Jacobian determinant are present in the homogeneous image regions (see for example the cranial part of the liver in figure 9.1). A more refined illustration of this tendency of aligning structure boundaries at the expense of creating sinks and sources in the estimated deformations is provided in figure 9.5.
In terms of the absolute registration error illustrated in figure 9.7, the EVO algorithm provided the best result among all four methods in registering the abdominal and the thoracic wall. This can be mostly attributed to the algorithm prioritizing the alignment of high contrast areas. However, the transition towards the contrast-devoid area is resolved with high divergences in the estimated deformations, leading to strong misregistrations in the ventral part of the liver. This behavior is also well indicated by the Jacobian and the curl maps, which both showcase moderate to large deviations from their expected values in these regions. Nevertheless, if sufficient contrast is available within the images, the EVolution algorithm was demonstrated to be capable of high accuracy and precision in several independent studies.[27, 60, 61]

Concerning the required computational time, the EVO algorithm introduces a rather large relative penalty, compared to the HSO and HSI. This is mainly due to the time-consuming calculation of the EVO data fidelity term.

Performance of the EVolution algorithm with modified regularization

The incompressibility-regularized EVolution (EVI) algorithm maintains the same data fidelity term as EVO, replacing the smoothness constraint with a penalty on deviations of the deformation Jacobian determinant from unity. This conceptually means that the EVI algorithm preserves the capability for performing cross-modality and cross-contrast co-registration, while restoring the volume conservation property lost by EVO. Since the data fidelity term is still aimed at aligning similar contrast patterns, the excellent edge alignment capabilities of EVO are maintained by EVI, providing mostly identical DSC and Jaccard indices. On the other hand, maximizing edge-alignment no longer leads to strong compressions and expansions of the tissues in homogeneous areas (see for example figure 9.5). At the same time, as showcased in both figure 9.3 and 9.5, the strong peaks in the curl magnitude maps are also attenuated, with a predominant tendency to increase in the proximity of sliding interfaces. There are, however, a few instances in which the EVI estimates deformations with a moderate magnitude of the curl within the organ boundaries. An example can be seen in the ventral part of the liver in the MR - MR mono-modal dataset (figure 9.3) and in the cranio-ventral part of the liver in figure 9.7. Since the reduction of the deviations of the Jacobian determinant from unity is simply a proof of convergence of the algorithm, the curl magnitude map can be a good indicator for misregistrations. This aspect is well emphasized by the spatial correspondence between an increase in the curl magnitude and of the absolute registration error in the cranio-ventral part of the liver in figure 9.7. Overall, for cross-contrast or cross-image modality registration of soft-tissue structures, EVI provided superior results compared to EVO, with no penalty in terms of computational latency.

Selection of a motion estimation model for image-based therapy guidance in moving organs

The selection of a motion estimation model or registration algorithm should primarily depend on the particularities of the application and the intended usage. For example, if mapping of organ labels/segmentations between images is of interest, all four of the algorithms discussed in this chapter would be suited for such a task (since all four demonstrated good contour matching capabilities). On the other hand, if the aim is to map quantitative data such as radiation dose, contrast agent absorption, diffusion/perfusion coefficients, etc., the precise displacement of the tissues within the organ boundaries becomes particularly important. In such applications, the occurrence of non-anatomical sinks and sources within the estimated deformations, would lead to a set of inaccurate post-registration quantitative measurements.

The current study showcases how adapting a branch of registration methods, such as variational algorithms, for a particular image modality/contrast and/or a particular anatomical location, may work towards the detriment of a different registration scenario. For example, in order to address the inability of the Horn-Schunck optical flow formulation for cross-contrast and cross-modality registration, the gray-level intensity conservation constraint was replaced with a more relaxed constraint on the conservation of contrast, imposed by the original EVolution algorithm. This, however, leads to the loss of the mass conservation property, imposed by the data fidelity term of the Horn-Schunck algorithm, thus decreasing the penalty on the formation of sinks and sources in the estimated motion field. This property is restored by the EVI by replacing the smoothness constraint imposed by the regularization term, with a penalty on deviations of the deformation Jacobian determinant from unity. At a first glance, after addressing these shortcomings, it may seem that the EVI method is best among all of the four investigated algorithms. One has to take into consideration, however, that the incompressibility constraint is optimal only for anatomical structures which posses this particular physical property. In case the tracked anatomy changes volume in between the images to be registered, EVO would be the better option. Also, for mono-modal registration scenarios, EVO (and implicitly EVI) tend to have a lower accuracy compared to HSO (and implicitly HSI). This is due to the fact that the data fidelity term of the HSO and HSI relies on information from voxel intensities, which is usually more abundant than image gradients upon which the data fidelity term of EVO and EVI is constructed. Therefore, it would be overall preferable that such variational registration methods are selected depending on the specific application, acquisition modalities, acquisition parameters, image contrasts and the physical properties of the tracked anatomy.

9.5 Conclusion

Current validation procedures for medical image registration algorithms in a clinical setting, often rely on visual inspection and/or contour-based criteria. The current chapter demonstrates that this is likely insufficient to guarantee the anatomical plausibility of the estimated deformations. Here, three additional criteria were proposed, which can assist such evaluations on a voxel-by-voxel basis: the local Jacobian determinant of the deformations, the local vorticity of the deformations and a criterion relying on MR-tagging. Four image registration algorithms were evaluated against these criteria: two pre-existing and two proposed in the scope of this study. The latter two replaced the smoothness constraint on the estimated deformations imposed by the former two with a constraint on compressibility. It was demonstrated that while both the original and incompressible methods provide high values of the DSC and the Jaccard index (0.8 - 0.9) and low average TRE (1.1 - 1.2 mm), the pre-existing methods may estimate physiologically inaccurate compressions and expansions inside

the liver and kidney. The proposed methods significantly reduced such effects, without a negative impact on deformation vorticity, estimation accuracy determined from MR-tagged images or increase in computational latencies. While improvements were obtained for incompressible tissues, the current study also demonstrates that each optimization of variational registration algorithms represents a compromise, which should be chosen consciously while considering the particular anatomical and physiological properties of the tissue under observation.

Acknowledgements

This work was supported by ITEA 3, project no. 16016 (STARLIT).

Bibliography

[1] L. Fass. Imaging and cancer: A review. *Molecular Oncololgy*, 2(2):115 – 152, 2008.

[2] M. Roach 3rd, J. L. Alberini, A. P. Pecking, A. Testori, F. Verrecchia, J. Soteldo, U. Ganswindt, J. L. Joyal, J. W. Babich, R. S. Witte, E. Unger, and R. Gottlieb. Diagnostic and therapeutic imaging for cancer: Therapeutic considerations and future directions. *J Surg Oncol*, 103(6):578 – 601, 2011.

[3] P. J. Keall, G. S. Mageras, J. M. Balter, R. S. Emery, K. M. Forster, S. B. Jiang, J. M. Kapatoes, D. A. Low, M. J. Murphy, B. R. Murray, C. R. Ramsey, M. B. van Herk, S. S. Vedam, J. W. Wong, and E. Yorke. The Management of Respiratory Motion in Radiation Oncology Report of AAPM Task Group 76. *Med. Phys*, 33(10):3874 – 3900, 2006.

[4] E. A. Eisenhauer, P. Therasse, J. Bogaerts, L. H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe, and J. Verweij. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur J Cancer*, 45(2):228 – 247, 2009.

[5] C. Zachiu, B. Denis de Senneville, C. Moonen, and M. Ries. A framework for the correction of slow physiological drifts. *Med Phys*, 42(7):4137 – 4148, 2015.

[6] R. D. Timmerman and L. Xing. *Image-Guided and Adaptive Radiation Therapy*. Wolters Kluwer — Lippincott Williams & Wilkins, 2009.

[7] L. M. Fanchon, A. Apte, C. R. Schmidtlein, E. Yorke, Y. C. Hu, S. Dogan, M. Hatt, D. Visvikis, J. L. Humm, S. B. Solomon, and A. S. Kirov. Evaluation of the tumor registration error in biopsy procedures performed under real-time PET/CT guidance. *Med Phys*, 44(10):5089 – 5095, 2017.

[8] A. J. Mundt and J. C. Roeske. *Image-Guided Radiation Therapy: A Clinical Perpective*. PMPH - USA, 2011.

[9] B. Denis de Senneville, C. Moonen, and M. Ries. MRI-Guided HIFU Methods for the Ablation of Liver and Renal Cancers. In Escoffre, J. - M. and Bouakaz, A., editor, *Therapeutic Ultrasound*, pages 43 – 63. Springer, 2015.

[10] M. Ries, B. Denis de Senneville, S. Roujol, Y. Berber, B. Quesson, and C. Moonen. Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues. *Magnetic Resonance in Medicine*, 64:1704 – 1712, 2010.

[11] V. Mani and S. Arivazhagan. Survey of Medical Image Registration. *J Biomed Eng Technol*, 1:8 – 25, 2013.

[12] B. Zitová and J. Flusser. Image Registration Methods: A Survey. *Image and Vision Computing*, 21(11):977 – 1000, 2003.

[13] J. B. Maintz and M. A. Viergever. A Survery of Medical Image Registration. *Med Image Anal*, 2(1):1 – 36, 1998.

[14] J. Weickert, A. Bruhn, N. Papenberg, and T. Brox. Variational optic flow computation: From continuous models to algorithms. In *International Workshop on Computer Vision and Image Analysis (ed. L. Alvarez), IWCVIA03, Las Palmas de Gran Canaria*, 2003.

[15] L. A. Vese and C. Le Guyader. *Variational Methods on Image Processing*. CRC Press, 2016.

[16] A. B. Osman, M. Ovinis, H. Osei, and F. M. Hashim. A Review of Optical Flow Models Applied for Fluid Motion Estimation. *ARPN J Eng App Sci*, 11(20):12181 – 12187, 2016.

[17] B. P. Horn and B. G. Schunck. Determining the Optical Flow. *Artificial Intelligence*, 17(1 - 3):185 – 203, 1981.

[18] N. K. Ostergaard, B. Denis de Senneville, U. V. Elstrom, K. Tanderup, and T. S. Sorensen. Acceleration and validation of optical flow based deformable registration for image-guided radiotherapy. *Acta Oncol*, 47(7):1286–1293, 2008.

[19] S. Roujol, M. Ries, B. Quesson, C. Moonen, and B. Denis de Senneville. Real - Time MR - Thermometry and Dosimetry for Interventional Guidance on Abdominal Organs. *Magn Reson Med*, 63(4):1080 – 1087, 2010.

[20] M. Glitzner, B. Denis de Senneville, J. J. Lagendijk, B. W. Raaymakers, and S. P. Crijns. On-line 3D Motion Estimation Using Low Resolution MRI. *Phys Med Biol*, 60(16):N301 – N310, 2015.

[21] D. Yang, H. Li, D. A. Low, O. Deasy, J., and I. El Naqa. A Fast Inverse Consistent Deformable Image Registration Method Based on Symmetric Optical Flow Computation. *Phys Med Biol*, 53(21):6143 – 6165, 2016.

[22] A. L. Martel, M. S. Froh, K. K. Brock, D. B. Plewes, and D. C. Barber. Evaluating an optical-flow-based registration algorithm for contrast-enhanced magnetic resonance imaging of the breast. *Phys Med Biol*, 52(13):3803 – 3816, 2007.

[23] C. Zachiu, N. Papadakis, M. Ries, C. Moonen, and B. Denis De Senenville. An improved optical flow tracking technique for real - time mr - guided beam therapies in moving organs. *Phys Med Biol*, 60(23):9003 – 9029, 2015b.

[24] D. E. Dupuy, M. McCabe, and W. N. McMullen. *Image - Guided Cancer Therapy: A Multidisciplinary Approach*. Springer, 2013.

[25] M.P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F.V. Gleeson, S.M. Brady, and J.A. Schnabel. MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Med Image Anal*, 16(7):1423–1435, 2012.

[26] H. Rivaz, Z. Karimaghaloo, and D. L. Collins. Self - similarity weighted mutual information: A new nonrigid image similarity metric. *Med Image Anal*, 18(2):343 – 358, 2014.

[27] B. Denis de Senneville, C. Zachiu, M. Ries, and C. Moonen. EVolution: An Edge - Based Variational Method for Non - Rigid Multi - Modal Image Registration. *Phys Med Biol*, 61(20):7377 – 7396, 2016.

[28] D. Terzopoulos. Image analysis using multigrid relaxation methods. *IEEE Trans Pattern Analysis and Machine Learning*, PAMI - 8(2):129 – 139, 1986.

[29] P. Nessi. Variational approach to optical flow estimation managing discontinuities. *Imag Comp Vis*, 11(7):419 – 439, 1993.

[30] J. Weickert and C. Schnör. A theoretical framework for convex regularizers in PDE-based computation of image motion. *Int J Comp Vis*, 45(3):245 – 264, 2001a.

[31] Y. Lei, L. Jinzong, and L. Dongdong. Discontinuity - preserving optical flow algorithm. *J Sys Eng Electronics*, 18(2):347 – 354, 2007.

[32] N. Monzón, A. Salgado, and J. Sánchez. Robust discontinuity - preserving optical flow methods. *IPOL Journal*, 6:165 – 182, 2016.

[33] H. H. Nagel. Extending the 'oriented smoothness constraint' into the temporal domain and the estimation of derivatives of optical flow. In O. Faugeras, editor, *Computer Vision — ECCV 90: First European Conference on Computer Vision Antibes, France, April 23–27, 1990 Proceedings*, pages 139–148. Springer Berlin Heidelberg, Berlin, Heidelberg, 1990.

[34] M. J. Black and P. Anandan. Robust dynamic motion estimation over time. In *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 296–302, 1991.

[35] J. Weickert and C. Schnör. Variational optic flow computation with a spatio - temporal smoothness constraint. *J Math Imag Vis*, 14(3):245 – 255, 2001b.

[36] Q. Yang, B. Parvin, and A. Mariano. Singular features in sea surface temperature data. In *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, volume 1, pages 516 – 520, 2000.

[37] T. Rohlfing, C. R. Jr Maurer, D. A. Bluemke, and M. A. Jacobs. Volume - Preserving Nonrigid Registration of MR Breast Imaged Using Free - Form Deformation with an Incompressibility Constraint. *IEEE Trans Med Imaging*, 22(6):730 – 741, 2003.

[38] T. Mansi, X. Pennec, M. Sermesant, H. Delingette, and N. Ayache. iLogDemons: A Demons - Based Registration Algorithm for Tracking Incompressible Elastic Biological Tissues. *Int J Comp Vis*, 92(1):92 – 111, 2011.

[39] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache. Symmetric log-domain diffeomorphic registration: A demons-based approach. In D. Metaxas, L. Axel, G. Fichtinger, and G. Székely, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2008: 11th International Conference, New York, NY, USA, September 6-10, 2008, Proceedings, Part I*, pages 754 – 761. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[40] Y. Yin, E. A. Hoffman, and C. L. Lin. Mass Preserving Nonrigid Registration of CT Lung Images Using Cubic B - Spline. *Med Phys*, 36(9):4213 – 4222, 2009.

[41] A. Wilczewska, S. Cygan, and J. Żmigrodzki. Displacement field estimation for echocardiography strain imaging using b-spline based elastic image registration: Synthetic data study. In R. Jabłoński and R. Szewczyk, editors, *Recent Global Research and Education: Technological Challenges: Proceedings of the 15th International Conference on Global Research and Education Inter-Academia 2016*, pages 309 – 315. Springer International Publishing, 2017.

[42] F. Xing, J. Woo, A. D. Gomez, D. L. Pham, P. V. Bayly, M. Stone, and J. L. Prince. Phase Vector Incompressible Registration Algorithm for Motion Estimation From Tagged Magnetic Resonance Images. *IEEE Trans Med Imaging*, 36(10):2116 – 2128, 2017.

[43] S. E. Fischer, G. C. McKinnon, S. E. Maier, and P. Boesiger. Improved myocardial tagging contrast. *Magn Res Med*, 30(2):191 – 200, 1993.

[44] E.-S. H. Ibrahim. Myocardial tagging by cardiovascular magnetic resonance: Evolution of techniques–pulse sequences, analysis algorithms, and applications. *J Cardiovasc Magn Reson*, 13(1):36, 2011.

[45] A. Bistoquet, J. Oshinski, and O. Škrinjar. Myocardial deformation recovery from cine MRI using a nearly incompressible biventricular model. *Med Image Anal*, 12(1):69 – 85, 2008.

[46] D. A. Stout, E. Bar-Kochba, J. B. Estrada, J. Toyjanova, H. Kesari, J. S. Reichner, and C. Franck. Mean deformation metrics for quantifying 3D cell-matrix interactions without requiring information about matrix material properties. *PNAS*, 113(11):2898 – 2903, 2016.

[47] E. Castillo, R. Castillo, J. Martinez, M. Shanoy, and T. Guerrero. Four-dimensional deformable image registration using trajectory modeling. *Phys Med Biol*, 55(1):305 – 327, 2010.

[48] L. R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297 – 302, 1945.

[49] P. Jaccard. The distribution of the flora in the alpine zone. *New Pathologist*, 11(2):37 – 50, 1912.

[50] C.R.J. Maurer, J. Fitzpatrick, M. Wang, R.L.J. Galloway, R. Maciunas, and G. Allen. Registration of head volume images using implantable fiducial markers. *IEEE Transactions on Medical Imaging*, 16:447–462, 1997.

[51] K. K. Brock, S. Mutic, T. R. McNutt, H. Li, and M. L. Kessler. Use of Image Registration and Fusion Algorithms and Techniques in Radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132. *Med Phys*, 44(7):e43 – e76, 2017.

[52] A. Kumarasiri, F. Siddiqui, C. Liu, R. Yechieli, M. Shah, D. Pradhan, H. Zhong, I. J. Chetty, and J. Kim. Deformable image registration vased automatic CT-to-CT contour propagation for head and neck adaptive radiotherapy in the routine clinical setting. *Med Phys*, 41(12):121712, 2014.

[53] B. Rigaud, A. Simon, J. Castelli, M. Gobeli, J.-D. Ospina Arango, G. Cazoulat, O. Henry, P. Haigron, and R. de Crevoisier. Evaluation of deformable image registration methods for dose monitoring in the head and neck radiotherapy. *Biomed Res Int*, 2015:726268, 2015.

[54] X. Li, Y. Zhang, Y. Shi, S¿ Wu, Y. Xiao, X. Gu, X. Zhen, and L. Zhou. Comprehensive evaluation of ten deformable image registration algorithms for controu propagation between ct and cone-beam ct images in adaptive hean & neck radiotherapy. *PLoS One*, 12(4):e0175906, 2017.

[55] C. Zachiu, B. Denis de Senneville, I. D. Dmitriev, C. T. W. Moonen, and M. Ries. A framework for continuous target tracking during MR-guided high intensity focused ultrasound thermal ablations in the abdomen. *J Ther Ultrasound*, 5:27 – 45, 2017c.

[56] J. Ophir, I. Céspedes, H. Ponnekanti, Y. Yazdi, and X. Li. Elastography: A quantitative mrthod for imaging the elasticity of biological tissues. *Ultrason Imaging*, 13(2):111 – 134, 1991.

[57] R. Akhtar, M. J. Sherratt, J. K. Cruickshank, and B. Derby. Characterizing the elastic properties of tissues. *Mater Today (Kidlington)*, 14(3):96 – 105, 2011.

[58] K. Arda, N. Ciledag, B. K. Aribas, E. Aktas, and K. Köse. Quantitative assessment of the elasticity values of liver with shear wave ultrasonic elastography. *Indian J Med Res*, 137(5):911 – 915, 2013.

[59] B. Selmi, G. Engelmann, U. Teufel, S. El Sakka, M. Dadrich, and J. P. Schenk. Normal values of liver elasticity measured by real-time tissue elastography (RTE) in healthy infants and children. *J Med Ultrason*, 41(1):31 – 38, 2014.

[60] C. Zachiu, M. Ries, P. Ramaekers, J.-L. Guey, C. T. W. Moonen, and B. Denis de Senneville. Real-time non-rigid target tracking for ultrasound - guided clinical interventions. *Phys Med Biol*, 62(20):8154 – 8177, 2017a.

[61] C. Zachiu, B. Denis de Senneville, R. H. N. Tijssen, A. N. T. J. Kotte, A. C. Houweling, L. G. W. Kerkmeijer, J. J. W. Lagendijk, C. T. W. Moonen, and M. G. Ries. Non - rigid ct/cbct to cbct registration for online external beam radiotherapy guidance. *Phys Med Biol, [Epub ahead of print]*, 2017b.

CHAPTER 10

Summary

"Time is the most valuable thing a man can spend."
- Theophrastus

Image-guided HIFU and EBRT have the potential to deliver local and regional cancer treatment in a completely non-invasive manner. Such therapies are, however, subject to uncertainties induced by patient and/or tumor motion. These can stem from several different sources such as patient positioning errors, spontaneous motion, long term organ drifts and/or respiration. If not taken into consideration, such displacements can lead to the therapeutic energy being diverted from the target pathology, increasing the risk of under-treatment and implicitly of disease recurrence. Moreover, in case the motion effects are severe, the absence of a compensation scheme can lead to serious complications due to collateral damage to healthy tissues. The contribution of the current thesis, towards technological and methodological advances in the field of image-guided HIFU and EBRT treatments, is threefold:

1. It initially connects to the state-of-the-art methodology dedicated to real-time MR - guidance of HIFU therapies, identifies thereby specific shortcomings and proposes strategies to overcome these limitations.

2. State-of-the-art methodology is then selectively and jointly used to design and implement large-scale frameworks, encompassing the entire interventional work-flow, dedicated to motion compensation during non-invasive beam therapies. The design and validation of these frameworks around pre-existing clinical work-flows has allowed further identifying state-of-the-art limitations in terms of image registration methodology.

3. Provide novel solutions for the previously identified limitations, specifically tailored for the particular clinical demands of image-guided HIFU and EBRT interventions.

In the following, the topic, the main results and the conclusions that can be drawn from each thesis chapter is briefly summarized.
The Horn & Schunck optical flow algorithm has been previously proposed as a feasible solution by several independent studies for elastic real-time MR-based respiratory motion compensation. The choice in these studies was justified by the capability of the method to provide dense and elastic deformations, its fast numerical scheme and low number of input parameters. However, the high confidence the algorithm places in the voxel intensity conservation hypothesis, may be a source of misregistrations in areas where this assumption is locally violated. **Chapter 2** addresses this limitation by proposing an improved version of the original Horn & Schunck algorithm, with increased resilience to local gray-level intensity variations not attributed to motion. This was achieved by replacing the quadratic data fidelity term of the Horn & Schunck method with a term based on the L_1 linear norm. The proposed approach was validated on a dynamic 2D MR data set comprised of 3000 images acquired on the abdomen of two free-breathing healthy volunteers. Compared to the Horn & Schunck algorithm, results have demonstrated an improved robustness of the proposed method to in-flow effects introduced by arterial pulsations. The

increased robustness was obtained while maintaining capabilities for real-time MR-based motion estimation. To the author's knowledge, this was the first study to analyze the behavior of an L_1-based optical flow functional in the context of real-time MR-guided HIFU. Fast imaging during such applications, however, plays only a secondary role as input to motion tracking algorithms. Its primary role lies in providing temperature maps of the treated area via MR-thermometry. A side-effect of fast MR-imaging is that the resulting images are frequently hampered by low signal-to-noise ratio (SNR). This, in turn, leads to imprecise temperature measurements, having a direct impact on therapy control, monitoring and therapeutic endpoint detection. Existing solutions for MR-thermometry denoising improve precision, often at the cost of accuracy. **Chapter 3** proposes a novel non-local denoising method dedicated to real-time MR-thermometry. Contrary to previous approaches, the proposed method improves MR-thermometry precision, while at the same time maximizing output accuracy. Particular attention was paid to computational latency, such that the real-time availability of the temperature measurements is maintained. The approach was validated in three complementary experiments: a simulation, an *ex-vivo* and an *in-vivo* study. Compared to existing MR-thermometry denoising methods, the proposed approach has demonstrated a statistically significant improvement at $p = 0.05$ significance level, in the quality of the denoised data. **Chapters 4** and **5** "zoom-out" from sonication-level-related methodological shortcomings and focus on large-scale motion compensation solutions encompassing the entire work-flow of MR-guided non-invasive beam interventions. Prior to the work conducted in **chapter 4**, studies addressing motion compensation in abdominal organs during both HIFU and EBRT, have focused on mitigating respiratory displacements during therapeutic energy deliveries. **Chapter 4** demonstrates that this may be conceptually insufficient, since during lengthy therapies abdominal organs may also undergo long-term drifts due to digestive activity, muscle relaxation and/or metabolic processes. As such, a solution is proposed, dedicated to compensating long term drifts occurring over the entire duration of the intervention. This was achieved by integrating in an existing clinical work-flow a series of 3D MR scans, quasi-uniformly distributed over the course of the treatment. The scans were subsequently registered to a reference volume acquired at the beginning of the therapy. The resulting displacements demonstrated that, during lengthy HIFU therapies, the human liver and kidneys can undergo drifts of over 5 mm, exceeding acceptable therapeutic margins. Based on *in-vivo* data acquired on a porcine liver, the study also demonstrated that the estimated displacements are sufficiently accurate and are provided with a sufficiently low temporal latency, to allow on-the-fly adaptation of the interventional plan and aggregation of the thermal dose in a common reference space. The promising nature of these results have lead to the design of the extended framework described and validated in **chapter 5**. This new framework addressed both respiratory displacements during the individual energy deliveries and long-term drifts occurring over the entire duration of the therapy. Additional components were included in the framework

both for quality assurance and linkage purposes. The linking of the several motion/deformation estimation components allowed to spatially relate each MR image acquired over the course of the therapy to the master planning image acquired at the beginning of the intervention. The resulting displacements were used for two purposes: 1) "Down-stream" propagation of the initially planned cluster of cells, such that the new location of the sonications matches the initial volume due for ablation and 2) "Up-stream" propagation of the thermal dose, such that therapy progress can be monitored on the planning image in a spatially consistent way. This extended framework was validated in three complementary experiments: a study on a gel phantom undergoing a known motion pattern, a study on the abdomen of 10 healthy volunteers and 3 *in-vivo* ablations on porcine liver. While the phantom and the volunteer data were processed offline, the estimations and corrections during the 3 animal experiments were performed on-the-fly. The phantom and the healthy volunteer experiments demonstrated that the proposed motion estimation/compensation framework is capable of providing estimates with sub-voxel precision and accuracy. Results have also re-confirmed what was observed in **chapter 4**, that during lengthy HIFU therapies, the human liver and kidneys can manifest displacements that exceed acceptable therapeutic margins. Finally, the 3 successful animal studies demonstrate the compatibility of the proposed approach with the work-flow of an MRg-HIFU intervention under clinical conditions.

The timely development of the principle behind the EVolution registration algorithm and the MR+USg-HIFU system, laid the foundation for addressing limitations of the state-of-the-art image registration methods, identified while designing, evaluating and validating the frameworks from **chapters 4** and **5**. These include: 1) The scarcity of robust and fast elastic multi-modal registration algorithms and 2) The dual role of fast MR imaging during HIFU sonications, in both tracking and thermometry, which requires a compromise on the acquisition parameters of the employed MR-sequence. The focus of **chapter 6** is therefore exclusively on the technical details and evaluation of the EVolution multi-modal registration algorithm. While overall the algorithm was on-par with the state-of-the-art methods, EVolution demonstrated an improved resilience to transient structures. In addition, it also maintains the advantages provided by variational registration methods, namely: it provides elastic and dense deformations, it implies fast numerical schemes and it requires a low number of input parameters. The properties of the EVolution algorithm together with the independent stream of B-mode US images provided by the hybrid MR+USg-HIFU system, presented the opportunity to migrate the task of real-time tracking during sonications from the MR-scanner to an US-imager. It is a known fact that B-mode US images are usually altered by transient gray-level intensity variations under the form of speckle noise. This was already proven to be problematic for existing intensity-based registration methods. **Chapter 7** evaluates the EVolution algorithm's capabilities for real-time US-based respiratory motion tracking. Due to its intrinsic robustness to transient structures, EVolution was hypothesized to be resilient to speckle noise. The approach was

evaluated on both contact and standoff echography images acquired on the abdomen of 7 healthy volunteers. The method demonstrated on average an accuracy of ∼1.5 mm and sub-millimeter precision. Together with a computational performance of 20 images/s, this made the EVolution method a feasible solution for the problem-at-hand. With this, the identified limitations for interventional guidance during therapy delivery were considered to be, to a large extent, adequately addressed. Therefore, **chapter 8** of the thesis shifts focus toward motion-related inaccuracies which occur during the (re-)planning stage of cancer therapy. Image-guided EBRT typically relies on the registration of the planning CT to CBCT images acquired at different stages of the treatment. This is currently achieved in clinics by the implantation of high-Z markers in the tumor and/or its close proximity. The CT and the CBCT images are then aligned based on the location of the implanted markers. Besides the marker implantation involving a particular degree of discomfort for the patient, such an alignment procedure only allows for estimating/correcting rigid displacements. Naturally, this is usually sub-optimal for deformable organs. In addition, the implanted markers are only surrogates for the true tumor position, with relative drifts between the markers and the tumor being a possibility. Therefore, an elastic pure image-based registration would be desirable. A difficulty in this sense is that CBCT images are usually acquired using a small number of X-Ray projections and/or low beam intensities. This often leads to the images being subject to low contrast, low signal-to-noise ratio and streaking artifacts. The latter, in particular, introduce intensity inconsistencies between CT and CBCT images, making registration problematic for a wide variety of methods. Due to its functioning principle, the EVolution algorithm was hypothesized to be robust against such effects. **Chapter 8** investigates this hypothesis by evaluating the method on several test benches involving CT/CBCT data acquired on several cancer patients afflicted by various pathologies. Results have demonstrated that the additional artifacts present in the CBCT images only have a sub-millimeter impact on the accuracy and precision of the EVolution algorithm. In addition, the method required on average 60 s to converge for $256 \times 256 \times 256$ images, which makes it applicable in online scenarios.

The Horn & Schunck and the EVolution algorithms played an important role in the motion estimation/compensation solutions provided over the course of this thesis. Both methods are, however, rooted in the optical image processing domain. Thus, no guarantee can be made on the anatomical plausibility of the estimated deformations, when applied to medical images. **Chapter 9** investigates this aspect by analyzing the methods through the perspective of two existing contour-based criteria and three additional indicators for anatomical plausibility. Two new versions of the two algorithms are also proposed, which replace the smoothness regularization with a constraint on deformation compressibility. This replacement is justified by the fact that the majority of biological tissues are incompressible. This should render the new formulations better adapted to tracking such tissues. Results have revealed that while all four methods have similar contour matching capabilities, the proposed incom-

pressible algorithms provide deformations which are in better correspondence with the material properties of biological soft tissues. Despite these improvements, **chapter 9** demonstrates that such methods should be used selectively, depending on the physical properties of the observed anatomy.

CHAPTER 11

General Discussion

"There are two sides to every question."
- Protagoras

11.1 Introduction

HIFU and EBRT are currently the only two cancer treatment options which have the potential to treat tumors deep inside the human body, in a completely non-invasive manner. Due to physiological motion, however, therapy delivery via these two modalities becomes challenging, especially for organs in the thorax and abdomen. In case the displacements of the target pathology are not taken into consideration, the therapeutic energy may end-up being diverted from it's intended delivery location, with two possible consequences: 1) A sub-lethal amount of therapeutic energy is delivered to the pathology, increasing the risk of under-treatment and implicitly of disease recurrence; 2) In case motion effects are severe, healthy tissues may move into the beam-path, resulting in potential collateral damage to critical structures. Therefore, motion compensation schemes are crucial for achieving the therapeutic end-point of HIFU and EBRT interventions in moving organs, while at the same time maximizing healthy tissue sparing. Medical image-based motion compensation is particularly attractive since it provides a non-invasive means for therapy guidance. Moreover, the fact that an increasing number of therapeutic systems are fitted with on-board medical imaging devices, further facilitate the widespread usage of such approaches. A large selection of image-based motion correction solutions for HIFU and EBRT are already available in the literature, however, the development and adoption of a method which definitively addresses this issue is particularly difficult. This is due to a variability in physiological properties across different pathologies and anatomies, the large degree of freedom in selecting imaging devices and/or image acquisition parameters and a high variability in inter-patient anatomical motion characteristics. For this reason, adapted motion estimation/compensation schemes are necessary, optimized for a particular application/scenario. Therefore, if a standard motion compensation strategy is to be translated into clinics, a tendency towards specialized multi-component frameworks, addressing a broad category of displacements and encompassing the entire work-flow of the intervention is desirable. This thesis makes a step towards such a strategy, providing an ensemble of novel motion estimation and compensation solutions, which when fused together, become a large-scale framework encompassing the therapeutic work-flow of HIFU and EBRT interventions.

11.2 Selection of a motion estimation model

Prior to addressing a medical image-based motion correction problem, a decision has to be made on a particular motion estimation model or formulation. There is much debate in the scientific community over which algorithm or approach would be suited best for addressing medical image-based motion tracking.[1,2] What is sometimes omitted or is overlooked in practice is that, the choice of a registration method should depend on a combination of several

criteria, including: the imaging modality, the image acquisition parameters, the anatomical region which is imaged, the physical characteristics of the tumor and the range of motions underwent by the tracked structure. To the authors knowledge, there is currently no standard method which is optimal for all possible combinations of these criteria. Therefore, only once the details related to imaging and anatomy are fixed, a motion estimation model should be decided upon. To a certain extent, an inverse chain of reasoning is also possible in the sense that, if a motion model is already fixed, the parameters of the image acquisition sequence could be optimized, for example, such that the basic assumptions made by the motion model are valid.

This thesis employs at different points several motion estimation algorithms: the Horn & Schunck optical flow algorithm,[3] the $L^2 - L^1$ method proposed in **chapter 2**, the MIND algorithm proposed by Heinrich *et al*,[4] EVolution (see **chapter 6**) and the incompressible versions of Horn & Schunck and EVolution from **chapter 9**. While all have showcased a high degree of accuracy and precision in particular scenarios, no conclusions can be drawn on which algorithm is the best overall. For example, the Horn & Schunck (original and incompressible) and the L^2-L^1 are only suitable for mono-modal image registration tasks. In this sense, MIND and EVolution should be considered superior, since they are capable of both mono-modal and multi-modal image registration. However, as showcased in Fig. 6.5 from **chapter 6**, for mono-modal registration of 3D MR bladder images, the $L^2 - L^1$ (and most likely also the original Horn & Schuck method) has a better performance than both the MIND and the EVolution algorithms. Compared to EVolution, MIND has increased sensitivity to transient structures, since the latter relies on the assumption that all objects in the reference image have a counterpart in the moving image (assumption which is not made by EVolution). On the other hand, especially in homogenous areas, the EVolution algorithm is sensitive to transient spatial gradients, which are present in both the reference and the moving image at the same spatial location. An example of such a scenario is illustrated in Table 8.4 from **chapter 8** where the performance of the EVolution algorithm for CBCT-based respiratory motion estimation was investigated. Note that, when a high resolution CT volume is used as a reference image, the errors remain considerably lower inside the liver and kidneys, compared to when both the reference and the moving images are CBCT volumes. This can be explained by the fact that, since the CT image does not contain streaking artifacts, the confidence in the data term of the EVolution functional is low, with the regularization term being dominant. As soon as temporally variant streaking artifacts are present in both the reference and the moving image, the data fidelity term will erroneously rely on the gradients induced by the artifacts in order to estimate motion. It is likely that the MIND algorithm is not sensitive to such effects, since the associated motion descriptors in these areas would be sufficiently distinctive such that the data fidelity is low in these areas.

An aspect which deserves particular attention for all the algorithms employed over the course of the thesis, is the regularization term. Unlike the data fidelity

term, which makes assumptions on the image content, the regularization term makes assumptions on the properties of the true underlying motion. Except for the two methods proposed in **chapter 9**, the regularization term of the rest of the algorithms assumes that the displacements of the tracked structures are spatially smooth. Besides being in some cases unsuitable for incompressible organs (see **chapter 9**), local violations of this assumption can also occur at sliding interfaces. Vishnevskiy *et al*[5] and Szmul et al[6] have addressed this by replacing the quadratic regularization on the motion gradients, with a more relaxed constraint on the total variation of the motion fields. The resulting solution space also includes motion fields containing discontinuities and therefore sliding and shearing motion is permitted. However, in their efforts to address downsides of the existing medical image-based motion estimation methods, such new studies often overlook that ultimately, the goal is to estimate physiologically plausible deformations. Penalizing the total variation allows discontinuities in the motion fields due to the fact that such a constraint provides a *piece-wise constant* solution, resulting in a so-called staircasing effect[7] (see Fig. 11.1). Such an assumption is, however, often not valid for deformable organs in the thorax and abdomen. The incompressibility constraint proposed in **chapter 9** for the Horn & Schunck and the EVolution algorithm provides better capabilities for the estimation of sliding motion for free. Since the estimated deformations are allowed to behave like liquid flows, shearing motion is not penalized and is solved by the model through the formation of vortices (see Fig. 9.3 and 9.7 from **chapter 9**). However, as argued in the proposing chapter, such a constraint is only suitable for biological tissues and not for volume-changing cavities. Therefore, it cannot be stressed enough that when a motion estimation model is selected, validation test benches should primarily demonstrate that the estimated deformations are in accordance with the physiological properties of the imaged anatomy.

11.3 Analogies between the work-flow of an image - guided HIFU and EBRT treatment

Most of the motion estimation/compensation methods described in this thesis are exemplified for the MR-guidance of HIFU therapies. The approaches, however, can be directly extended to image-guided EBRT. This is facilitated by the similarities shared by the typical work-flows of the two treatment modalities. Fig. 11.2 provides an abstract illustration of a typical work-flow which is applicable for both HIFU and EBRT. The procedures associated to each phase of the work-flow are enumerated in Table 11.1 for both HIFU and EBRT, further emphasizing the analogies between the two modalities. As long as on-board MR, CT and/or US imaging is available, robust motion compensation is achievable at different time-scales via the methods proposed over the course of the thesis. The adaptation of the framework described in **chapter 4** and its extension proposed in **chapter 5** to image-guided EBRT is relatively direct. Estimation

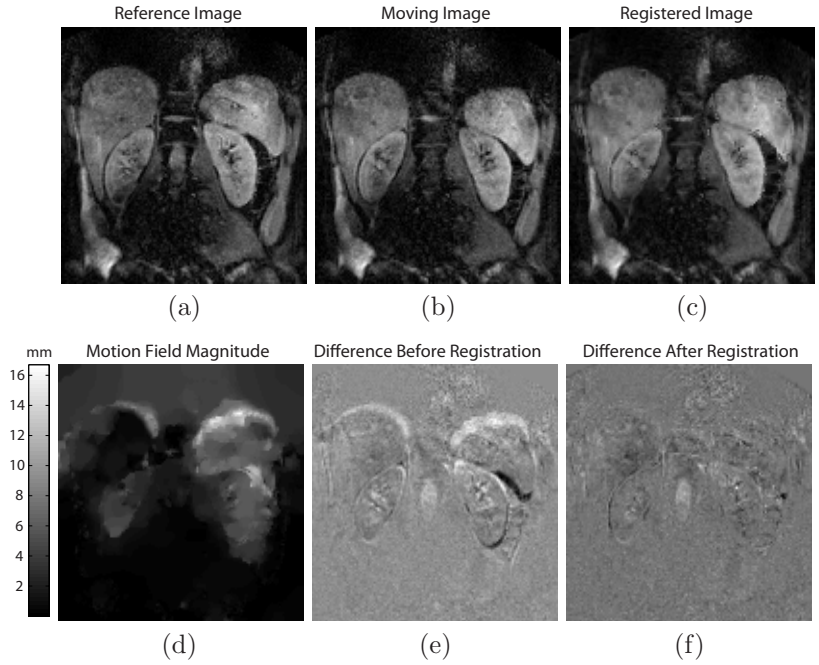


Figure 11.1: Illustration of the staircasing effect when using total variation as a constraint in the motion estimation model. (a) Reference image. (b) Moving image. (c) Registered image. (d) Magnitude of the estimated displacements. (e) The difference between the reference and the moving image. (f) The difference between the reference and the registered image. The piecewise-constant nature of the estimated displacements can be observed in Fig. (d). Despite the good alignment between the images after registration, the resulting deformation is not physiologically plausible.

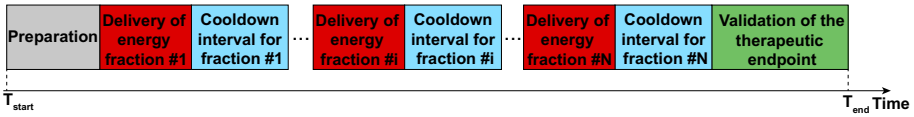


Figure 11.2: Typical workflow of HIFU and EBRT treatments.

and compensation of the slow physiological drifts can be achieved by the integration of 3D scans in between fractions or gantry rotations (blue boxes in Fig. 11.2). This is naturally followed by the registration of the acquired 3D images to a reference scan. During the delivery of the radiation fractions/segments (red boxes in Fig. 11.2), real-time 2D imaging can be performed for respiratory motion estimation and subsequent compensation via gating or tracking. The

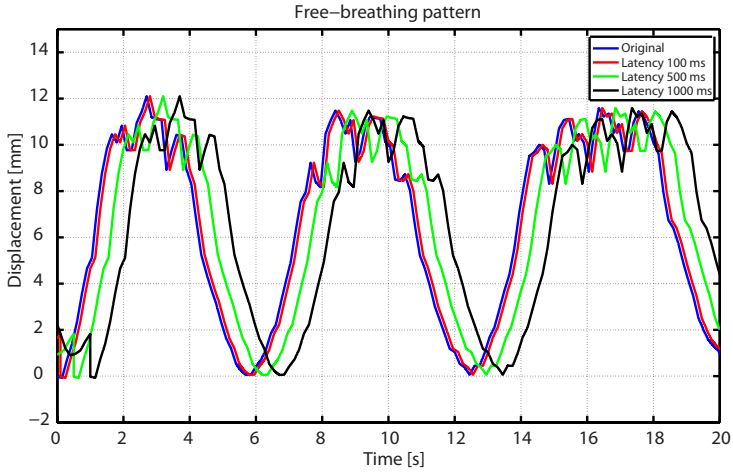
Treatment phase	Possible procedures		
	HIFU	EBRT	
		Intra-fraction	Inter-fraction
Preparation	• Patient positioning • Acquisition of T_1/T_2 weighted images • Setup of the sonication cells	• Patient positioning • Anatomical imaging for position validation	• Acquisition of high-resolution CT and/or MR images • Delineation of pathological tissues and organs-at-risk • Optimization of the treatment delivery plan
Deliver energy fraction $\#i$	• Perform sonication $\#i$ • Real-time motion compensation • Real-time thermometry	• Deliver radiation segment $\#i$ • Real-time motion compensation (still under research) • Real-time dose mapping (far future)	• Deliver radiation fraction $\#i$
Cooldown interval for fraction $\#i$	• Allow tissues in the near field to cool-down • Inter-sonication motion compensation	• Rotate gantry • Motion estimation for aperture adaptation	• Allow OARs to recover from the effects of radiation
Validation of the therapeutic endpoint	• Perfusion imaging for NPV validation • T_2 imaging for edema detection	• Anatomical imaging for position re-validation	• Anatomical imaging for response monitoring

Table 11.1: Example of operations which are/can be performed during each phase of the typical workflow of image-guided HIFU and EBRT treatments.

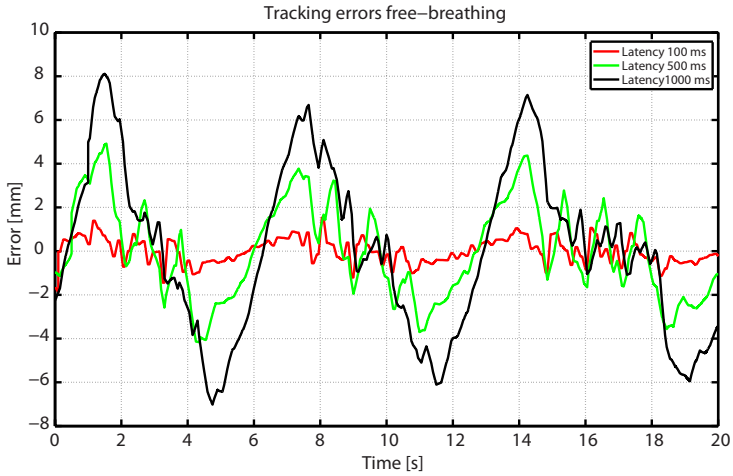
rest of the scans included by the framework described in **chapter 5** can also easily be integrated in the work-flow of an image-guided EBRT intervention, simply by analogy with the typical work-flow of and MRg-HIFU therapy (see Table 11.1).

11.4 Real-time 3D motion compensation

The solutions proposed in this thesis for respiratory motion estimation, consistently relied on 2D images (see **chapter 2**, **chapter 5** and **chapter 7**). While the imaging planes were aligned as much as possible to the principal axis of respiratory displacements, through-plane motion still remains a possibility. In this sense 3D tracking of respiratory motion would be preferable. Achieving this during image-guided HIFU and EBRT therapies with real-time performance is, however, a technical challenge. It imposes severe constraints on both the image acquisition process and the latencies introduced by the motion tracking algorithm. From an imaging perspective, one of the constraints is that the image acquisition frequency should be sufficiently high to resolve in 3D all the deformations induced by respiration over the respiratory cycle. Limiting real-time imaging to this alone is, however, insufficient. The typical anatomical situation over a respiratory cycle, for example, could be densely resolved in 3D by partial acquisition of the image in the same respiratory phase, over several respiratory cycles.[8,9] The partial images are then sorted according to the respiratory phase in which they were acquired, with the full image being subsequently reconstructed. In this manner, the anatomical configuration over the respiratory cycle can be sampled with a custom - defined density. A keyword which is often omitted in studies related to real-time motion compensation is *latency*. Whether a method is real-time or not depends on the temporal latency between the actual anatomical motion event and the reaction to it (e.g. gating or beam-steering). Obviously, it is unacceptable to wait several respiratory cycles for a 3D image to be acquired and subsequently decide whether the organ is in the gating window or adapt the HIFU beam focus/MLC aperture. Therefore, in order to fulfill the real-time constraint for respiratory motion compensation, the total latency associated with the image acquisition time, the computational time of the motion estimation algorithm, the (potential) inertia of the beam steering system and the data transfer latency should together reside below a particular threshold, which has been identified by previous studies to be ≤ 500 ms for EBRT[10] and ≤ 100 ms for HIFU.[11–13] Fig. 11.3 showcases an example of respiratory motion tracking errors as a function of latency. It can be observed in Fig. 11.3(b) that for a latency of 100 ms, maximum tracking errors mostly remain sub-millimeter. Depending on the application, this may be considered acceptable. Latencies of 500 ms are already debatable, since the maximum errors already exceed 50 % of the amplitude of the original breathing pattern. In such a case, motion prediction algorithms may be necessary. For latencies of 1000 ms, the magnitude of the tracking errors exceed the amplitude of the original pattern. Therefore, in this case, the absence of tracking would have less of a negative impact. Note, however, that this is just an example of a breathing pattern. In general, the extent of the tracking errors as a function of latency, depends on the frequency and amplitude of the respiratory displacements In 3D, already from an imaging point-of-view, real-time tracking is



(a)



(b)

Figure 11.3: Respiratory motion tracking errors as a function of latency. (a) The average cranio-caudal displacements of the liver of a free-breathing volunteer. The figure showcases both original and delayed versions of the motion pattern. The delayed versions illustrate the effect of 100 ms (red curve), 500 ms (green curve) and 1000 ms (black curve) latencies on the original pattern. (b) The resulting motion tracking errors as a function of the three simulated latencies.

hardly achievable with an SNR, spatial resolution and/or field-of-view (FOV) that would be suitable for image-based motion estimation. For 3D MR imaging, a temporal resolution of ~ 1 s was achieved in the study of Glitzner *et al* by

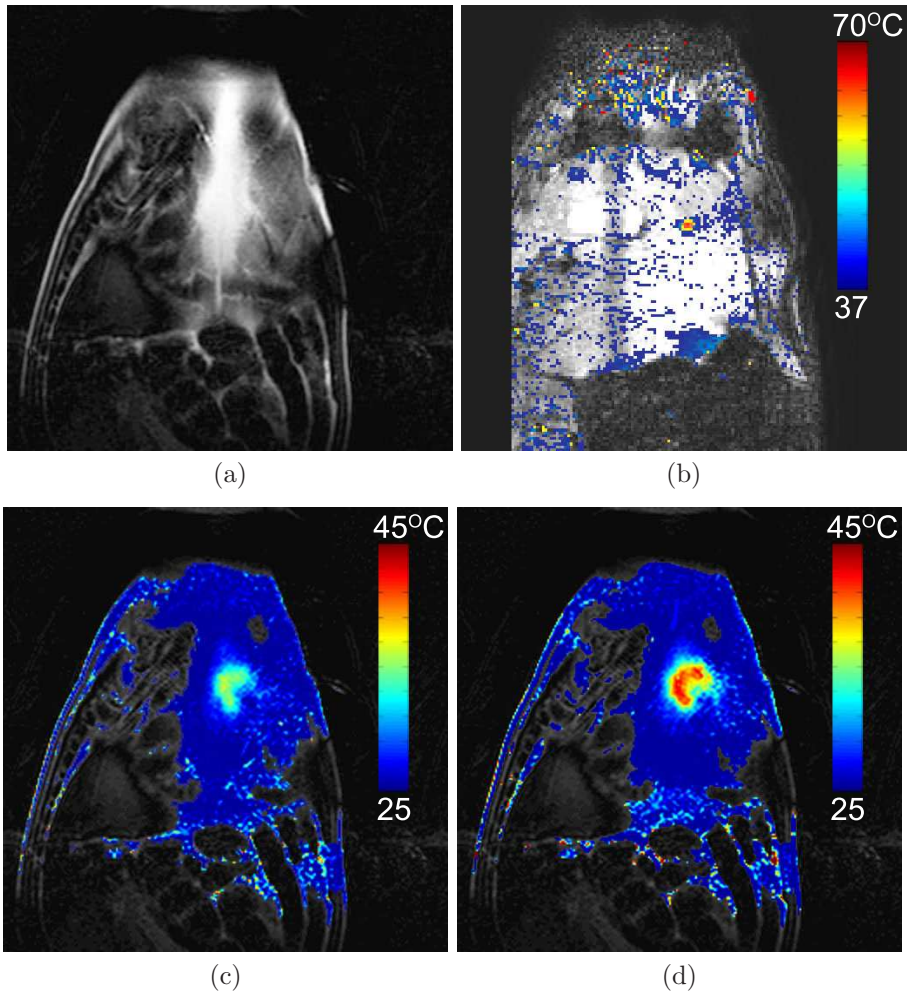
acquiring undersampled volumes.[14] It was also demonstrated that the Horn & Schunck optical flow algorithm still provides reliable motion estimates for 3D MR-images acquired with a spatial resolution of $5 \times 5 \times 5$ mm^3. Beyond this threshold, however, the confidence in the motion estimates diminishes to unacceptable values. Therefore, further acceleration of the acquisition process via undersampling was no longer of practical usage. Similar acquisition times for 3D images were also obtained by Plathow[15] and Sawant,[16] where fast imaging was achieved via the fast low angle shot (FLASH) and partial Fourier acquisition techniques, respectively. Therein it was also indeed concluded that further work is required in order to achieve 3D MR imaging with a sufficiently high temporal resolution to allow real-time respiratory motion compensation. On the other hand, recent advancements in medical imaging have lead to the development of 3D US systems. While compared to MRI the soft-tissue contrast is lower, 3D US images of diagnostic quality can be acquired with framerates of over 10 Hz.[17] To the authors knowledge, neither HIFU nor EBRT were to date conducted under 3D US guidance. Therefore, this may be a potential topic for future studies, since it was already demonstrated in **chapter 7** that reliable US-based motion estimation is achievable via the EVolution algorithm. Naturally, besides fast acquisitions, a fast convergence of the tracking algorithm is also required in order to achieve real-time 3D respiratory motion estimation. For 3D images covering an FOV of a clinically-relevant size, none of the implementations performed in the scope of this thesis were capable of sub-second convergences on the available hardware. However, if the algorithms were to be executed on several linked GPUs, the author anticipates that such convergence speeds could be achieved. Besides hardware accelerations, faster optimization schemes for the proposed algorithms are also an option, especially with the recent advancements in deep learning methods for medical image processing.[18]

11.5 The need for (real-time) MR-Thermometry denoising

Chapter 3 proposes a novel method for real-time denoising of PRFS temperature maps acquired during thermal ablations. While at first glance it may seem somewhat disconnected from the core topic of the thesis, real-time MR-Thermometry and motion compensation can benefit from the availability of a fast and robust temperature denoising method. MRg-HIFU therapies in general, rely on fast MR acquisition sequences for high duty-cycle thermometry, allowing an adequate therapy control and monitoring. Real-time respiratory motion estimation for beam-steering during HIFU sonications, impose even stricter constraints on the imaging temporal resolution. In order to reach such high MR imaging framerates, a tradeoff between spatial resolution, FOV and/or SNR is necessary. Out of the three, it is typically the SNR that is compromised, since a thermal map with a high spatial resolution is particularly important

for HIFU therapies, while the FOV in general cannot be directly reduced without the introduction of fold-over artifacts. Therefore, given also the tendencies towards 3D MR-Thermometry[19,20] and 3D real-time motion compensation,[11,14,21] which impose even further constraints on the achievable SNR, a fast and robust MR-thermometry denoising method should be readily available. While the denoising method proposed in **chapter 3** has been used in the context of PRFS thermometry during HIFU and RF ablations, it can be further extended towards other applications. The PRFS can be used as a temperature measuring mechanism in aqueous tissues, being impractical in fatty tissues due to their low water content. For HIFU therapies in the breast, this can prove particularly problematic, since the breast only contains isolated isles of glandular tissue, with the rest being predominantly fat. In this sense, a new method for MR-thermometry has been proposed by Baron *al*,[22] which relied on the sensitivity of the T_2 relaxation time on temperature, in order to perform MR-thermometry in fatty tissues. Breast cancer primarily develops in the glandular tissue, therefore T_2 temperature maps are useful for ensuring that the fatty tissue situated in the near-field does not suffer collateral damage. Tissues in the near-field tend to have a smoother temperature distribution than the immediate area surrounding the focal point, with rather large homogenous regions (see Fig. 11.4). Therefore, the non-local filter proposed in **chapter 3** would prove particularly effective in this scenario, since the T_2 temperature maps contain a high degree of redundancy. Moreover, in the study conducted by Baron *et al*, the acquisition of a T_2 temperature map required on average $\sim$ 16 s. While in the scope of their study this was sufficient, in a clinical setting this would be rather impractical, since an early detection of a near-field overheating event is imperative for preventing collateral damage. Therefore, the temporal resolution of the T_2 thermometry could be increased by an SNR tradeoff, which would then be subsequently addressed by the fast NLM filter proposed in **chapter 3**.

Up to date, the solutions proposed for real-time motion estimation during MRg-HIFU therapies typically rely on the magnitude images provided by MR-Thermometry. Therefore, the image acquisition parameters have to be optimized to suit both applications, with a compromise being usually necessary. For example, thermometric updates at rates of $\sim$2 Hz are typical and acceptable in a clinical setting,[23] while for the estimation and compensation of respiratory motion via beam steering this would be sub-optimal.[11–13] In effect, the MR-thermometry acquisitions need to be accelerated, with a cost in SNR. It would be therefore preferable to decouple MR-Thermometry and real-time motion compensation. One solution is to temporally multiplex two MR-sequences: one optimized for MR-thermometry and the other optimized for real-time respiratory motion tracking. A second solution, which is available at UMC Utrecht, it to assign imaging for the purpose of tracking to a different imaging system altogether. Such a solution was described in **chapter 7** and consisted in an integration of both and MRI scanner and an US imager into a HIFU system. Therein it was also proved that real-time US-based respiratory motion tracking



(a) (b)

(c) (d)

Figure 11.4: Spatial temperature distribution in the focal plane versus near-field. (a) Coronal near-field MR slice acquired on a pig during an *in-vivo* liver ablation. (b) PRFS temperature map in the focal plane. The low level of spatial redundancy in the temperature map makes the application of non-local filtering methods particularly challenging. (c) T2 temperature map acquired on the fatty tissue in the near-field. Mild heating can be observed. (d) T2 near-field temperature map, showcasing high near-field heating. Compared to the focal plane, the temperature maps acquired on the near-field showcase a higher degree of spatial redundancy, facilitating the application of non-local denoising methods.

is achievable on this system, with an average accuracy of $\sim$ 1.5 mm and sub-millimeter precision. By re-allocating the task of motion tracking to the US imager, the MR-thermometry acquisition sequence could then be optimized, for example, for higher SNR, partially or completely eliminating the necessity of a denoising method.

11.6 A step towards completely non-invasive EBRT

A direct leap from the current state-of-the-technology to completely non-invasive guidance of EBRT treatments would be to a large extent disruptive and highly unlikely to be clinically adopted in the near future. Such a transition should be preferably made in increments, leaving intact most of the elements with which the clinicians are familiar, an gradually introducing a new technology. A step in this direction was made in **chapter 8** of the thesis, where it was demonstrated that CBCT guidance during EBRT treatments may be possible without the implantation of high-Z markers.
Current clinical practice aims to detect and to correct patient positioning errors, by ensuring that the markers implanted in the tumor and/or in its immediate proximity, are at roughly the same location at the time of treatment as they were during planning.[24,25] This is achieved by acquiring CBCT images immediately prior to the delivery of a radiation fraction and comparing them to a planning high-resolution CT image. Besides only allowing the estimation of rigid displacements, the implantation process of the markers can imply discomfort to the patient, since the procedure is usually performed in the absence of anesthesia. Moreover, the markers are surrogates for the tumor location, with relative displacements between the markers and the tumor being a possibility. All these shortcomings can be addressed by replacing the marker-based alignment procedure with the purely image-based approach proposed in **chapter 8**. Similar to the other solutions proposed in the thesis, particular effort has been invested such that the method proposed in **chapter 8** is compatible with the existing clinical work-flow. Therefore, a clinical translation of the method is facilitated by the fact that no modifications are required to the typical work-flow of the intervention: it simply involves the replacement of the currently used alignment method. Naturally, additional validation in a clinical setting is required prior to its adoption. This can be achieved, for example, by clinical usage of EVolution in parallel with the existing marker-based approach as a gold standard. By evaluating how the EVolution algorithm compares to the existing one over a reasonable number of cases, the potential of clinical usage of the proposed solution could be investigated.

11.7 Clinical translation of image-based motion compensated HIFU and EBRT therapies

The ultimate goal of the methods and frameworks developed in this thesis is their translation into clinical practice. This may prove, however, rather problematic. Despite the benefits it offers, HIFU is still far from a widespread adoption. One of the likely reasons is that, under the shape that they are currently promoted, HIFU therapies are a *disruptive* technology. To date, cancer standard-of-care typically implies the combined effort of surgeons, radiation therapists and oncologists. The surgeon ensures the surgical removal of the bulk of the primary tumor and/or its metastases, allowing also their pathological analysis. The radiation therapist is responsible for de-bulking the tumor prior to surgical intervention and to ensure the regional control of micro-invasions. In case the cancer is systemic, the oncologist will prescribe chemotherapeutic drugs for curative or adjuvant purposes. Therefore, there is a well established clinical protocol for cancer treatment. HIFU therapies aim to achieve in a non-invasive manner what is currently the task of the surgeon: local tumor control. Also, since the tumor is no longer removed, its pathological analysis is disrupted, making its classification and choice of the optimal cytostatic drugs difficult. Naturally, all this unbalances the current clinical work-flow for cancer treatment, which demotivates the widespread adoption of HIFU. Another factor which contributes to this is limited validation in clinics, in particular for moving organs. Currently, HIFU has been FDA approved in the United States and CE marked in Europe for the treatment of uterine fibroids, prostate cancer, bone metastases and essential tremor.[26] In China, however, the clinical usage of HIFU has been employed for a larger array of conditions including: uterine fibroids, bone tumors, breast cancer, liver cancer, renal cancer and pancreatic cancer.[27] Motion compensation for mobile organs was predominately achieved by putting the patient under general anesthesia and respiratory control via intubation. A purely image-based motion compensation strategy as the ones proposed over the course of this thesis would improve patient recovery times, complication rates and cost efficiency of such interventions. Additional validation and stringent quality assurance (QA) tests of the methods, would be indeed required in pre-clinical experiments prior to clinical translation. Also, objective intra-operative criteria for the quality of information extracted from the images, either being deformations or temperature, would have to be established in order to ensure patient safety. **Chapter 9** already makes a step in this direction, by providing QA criteria for the physiological plausibility of the estimated deformations. For example, the availability of a Jacobian and/or a curl map is quasi-instantaneous once the deformations are estimated. Thus, a clinician or an automated system could analyze them immediately and decide whether the estimations are in good correspondence with the properties of the underlying anatomy. Nevertheless, a more widespread of HIFU as a treatment modality is necessary before even considering the choice of methodology.

Clinical translation of motion compensation methods for EBRT should prove somewhat more straight-forward than for HIFU. This is due to the fact that EBRT is already a part of the standard-of-care for cancer treatment. The downside is that the benefit of using novel methods in EBRT can only be made through long term observation (e.g. 3, 5, 10 year survivability). Moreover, there is a strong momentum to overcome, precisely due to the fact that EBRT treatment delivery is already standardized. Therefore, the inclusion of advanced motion compensation strategies in the therapeutic work-flow is somewhat disruptive and would require costly QA procedures. A good opportunity for employing such methods in EBRT is with the newly developed MR-Linacs. Since such systems have to be certified as a whole, a battery of tests can be developed for motion compensation strategies in particular. This circumvents the disruptive nature of novel motion compensation methods in classical EBRT, since the MR-Linacs were already designed to include motion compensation capabilities.[28] This does not come without complications, however, since online motion compensation during EBRT may imply online replanning. Thus, an overall work-flow which includes both motion compensation and online replanning will have to be designed and clinically validated.

11.8 Future perspectives

Following the methodological and technological advances provided by the work in this thesis, there are several directions which can be further pursued.
A goal that is constantly pursued in motion-related guidance for both HIFU and EBRT is real-time 3D motion estimation coupled with beam steering. As discussed in section 11.5 this is technically challenging, since it requires image acquisition, motion estimation and beam steering with relatively extreme constraints on temporal latencies for current technology. Nevertheless, this would facilitate a 100% duty cycle of the therapeutic beam during energy deliveries, considerably reducing the overall therapy times. Such a development would also enhance the framework proposed in **chapter 5**, where respiratory motion estimation was addressed by a combination of gating and residual 2D respiratory motion compensation. Furthermore, an acceleration of the image-based motion compensation strategies would potentially even allow addressing patient spontaneous motion, which in the scope of this thesis was considered to be circumvented by fixation devices and/or patient sedation. All this would make the interventions increasingly less cumbersome for both the patient and the clinical staff. Therefore, future research in this sense should focus on developing faster image acquisition schemes, motion estimation algorithms and actuators for beam steering, with a few options already being discussed in section 11.5.
The motion estimation framework proposed in **chapter 5**, while it addresses displacements stemming from different sources and occurring at different timescales, it does so mostly at the therapy delivery and guidance level. The overall

treatment work-flow typically includes two additional components: planning and therapy response monitoring. In particular for EBRT (and to a certain extent for HIFU), during the planning phase, a master high resolution CT image is acquired upon which the pathological tissues and the OARs are delineated. In addition, several other CT scans (e.g. contrast enhanced) and MR images of different contrasts weightings might be acquired, for differential diagnostic purposes. A clinician would then have to manually propagate the delineations (and potentially the therapeutic plan) from the master CT onto the subsequent scans. Similarly, the patient's therapeutic response monitoring is performed by the acquisition of CT and MR scans, employing a range of acquisition schemes/parameters. Following treatment, the tumor may regress, progress or stagnate. Therefore, a clinician would have to delineate the pathological tissues on the newly acquired images in order to decide on the tumor status. With the growing number of medical image screenings a patient undergoes, this manual delineation and treatment response assessment becomes highly time-consuming and error prone. This aspect could be potentially addressed by expanding the framework proposed in **chapter 5** to also include therapy planning and response monitoring and relate, via image registration, all the images acquired on the patient to a single master CT. The resulting deformations could then be used for propagating the delineations performed on the master CT on any subsequent image. Moreover, the therapeutic dose could be warped as it is being delivered onto the master CT and compared to the initial delivery plan. Such linking can be hypothetically achieved by the use of the EVolution algorithm (used selectively in its original and incompressible form), which is still in its infancy with its potential being far from being fully developed.

An element which is often absent from most of the online image-based motion estimation solutions is an objective intra-operative quality evaluation criterion for the estimated motion/deformation. The proposed methods are usually validated beforehand and promoted for clinical usage with the guarantee that they will perform similarly at all time. In some instances, visual contour-based inspection of the registration outcome is adopted as a validation criterion, which is a subjective and error-prone approach. In order to reduce risks for the patients undergoing registration-based motion compensated HIFU or EBRT therapies, objective online quality evaluation criteria for the estimated deformation need to be established. This aspect has been to a particular extent addressed in **chapter 9**, where in three objective QA criteria were proposed. Future work could also focus on further expanding or refining this list of criteria, by including for example, aside incompressibility and vorticity, also tissue stiffness and/or elasticity. Similar to the work performed in **chapter 9**, these parameters can then be extracted from the estimated deformation and compared against nominal values. In case the discrepancy between the true and the estimated value of the parameter exceeds a particular threshold, the estimated deformation is to be considered unreliable. This idea can be extended a step even further. These could be integrated in the motion estimation model as constraints on the es-

timated deformation. This will presumably lead to even more physiologically meaningful deformations provided by the tracking methods.
One of the drawbacks of HIFU therapies is the typically lengthy therapy duration, with some groups reporting interventions extending up to 8 hours.[29] This is both due to perfusion effects which increase the required times for thermal build-up at the focal spot and due to the formation of secondary high temperature ares in the prefocal zone, imposing lengthy cool-down intervals. While not detailed in **chapter 7**, since the emphasis was on US-based motion estimation, the HIFU transducer integrated in the hybrid MR-US-guided HIFU system, was designed with this issue in mind. The Voronoi tessellated transducer with the piezoelectric elements placed in a Fermat spiral, was demonstrated via numerical simulations to have improved focal point quality and more favorable prefocal pressure distribution compared to the existing hemisperical designs.[30] This could considerably reduce therapy durations, since thermal ablation at the focal spot could be achieved in shorter time intervals, with a higher duty-cycle of the therapeutic beam, due to the lower thermal build-up in the near-field. If a real-time motion estimation strategy as described in **chapter 7** would be coupled with beam steering during energy deliveries, the therapy duration times can potentially be reduced even further. Therefore, pre-clinical and clinical validation of this system during *in-vivo* scenarios is critical to further motivate the adoption of HIFU on larger scales, since currently it is drawn back by both hardware and software limitations.
Particular effort has been recently invested into providing alternatives to the already standard photon EBRT. These alternatives rely on delivering the therapeutic radiation dose via energetic beams of protons or heavy ions, with proton therapy being the most popular as of 2012.[31] The usage of such heavier particles leads to a more conformal radiation dose deposition, delivering their maximum lethal dosage at or near the tumor. This makes the necessity of a motion/deformation compensation solution of an even higher importance compared to photon therapy. The advancements achieved by this thesis could therefore also be employed for motion compensation in the context of such emerging therapies.
This thesis discusses HIFU and EBRT as individual treatment options, with an evaluation of the analogies between the therapeutic work-flows of the two modalities. There is, however, research towards clinical usage of combined HIFU - EBRT treatments.[32,33] It was demonstrated that prolonged hyperthermia at temperatures above 43°C, considerably increases the radiosensitivity of biological tissues by hampering DNA repair mechanisms. Therefore, by inducing HIFU hyperthermia immediately after EBRT, the effectiveness of the latter could be significantly enhanced. The work conducted in the scope of the current thesis could serve, for example, as a step towards clinical translation of a joint HIFU - EBRT work-flow.

11.9 Concluding remarks

This thesis brings several methodological and technological advances to the state-of-the-art, dedicated to motion compensation for image-guided HIFU and EBRT therapies. The tendency was towards large-scale frameworks encompassing the entire work-flow of such treatments. The components of these frameworks offer motion estimation solutions for several types of motion, hampering different anatomical sites which were imaged using several imaging modalities and image contrasts. Together with these solutions, criteria which provide objective quality assurance of the estimated deformations are also included, facilitating the process of clinical translation. While the content of this thesis focuses mostly on interventional therapy guidance, the proposed methodological advances pave the way for a direct extension to frameworks also including treatment planning and response monitoring.
The developments presented in this thesis also provide a good basis for future studies, which should address the further identified limitations. Some of the most important include a lack of a real-time 3D image-based respiratory motion compensation solution, the need for additional objective criteria for online QA of the motion estimates and, perhaps the most important, a shortage of clinical studies demonstrating the potential of the methods in clinical scenarios. With the development of novel image-guided HIFU systems and transducers, plus the first successful clinical deployments of the MR-Linac, it becomes increasingly clear that such technologies will play an important role in the future of cancer treatment. Regardless whether the widespread adoption of these treatment options is short-term or long-term, image-based motion compensation frameworks are going to be indispensable to their success.

Bibliography

[1] J. B. Maintz and M. A. Viergever. A Survery of Medical Image Registration. *Med Image Anal*, 2(1):1 – 36, 1998.

[2] V. Mani and S. Arivazhagan. Survey of Medical Image Registration. *J Biomed Eng Technol*, 1:8 – 25, 2013.

[3] B. P. Horn and B. G. Schunck. Determining the Optical Flow. *Artificial Intelligence*, 17(1 - 3):185 – 203, 1981.

[4] M. P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. V. Gleeson, S. M. Brady, and J. A. Schabel. MIND: Modality Independent Neighbourhood Descriptor for Multi - Modal Deformable Registration. *Med Image Anal*, 16(7):1423 – 1435, 2012.

[5] V. Vishnevskiy, T. Gass, G. Szekely, C. Tanner, and O. Goskel. Isotropic Total Variation Regularization of Displacements in Parametric Image Registration. *IEEE Trans Med Imaging*, 36(2):385 – 395, 2017.

[6] A. Szmul, B. W. Papiez, R. Bates, A. Hallack, J. A. Schnabel, and V. Grau. Graph Cuts-Based Registration Revisited: A Novel Approach for Lung Image Registration Using Supervoxels and Image-Guided Filtering. In *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 592 – 599, June 2016.

[7] S. Wei and H. Xu. Staircasing Reduction Model Applied to Total Variation Based Image Reconstruction. In *2009 17th European Signal Processing Conference*, pages 2579 – 2583, Aug 2009.

[8] E. Rietzel, T. Pan, and G. T. Chen. Four - Dimensional Computed Tomography: Image Formation and Clinical Protocol. *Med Phys*, 32(4):874 – 889, 2005.

[9] J. Cai, Z. Chang, Z. Wang, W. Paul Segars, and Yin F. F. Four - Dimensional Magnetic Resonance Imaging (4D - MRI) Using Image - Based Respiratory Surrogate: A Feasibility Study. *Med Phys*, 38(12):6384 – 6394, 2011.

[10] P. J. Keall, G. S. Mageras, J. M. Balter, R. S. Emery, K. M. Forster, S. B. Jiang, J. M. Kapatoes, D. A. Low, M. J. Murphy, B. R. Murray, C. R. Ramsey, M. B. van Herk, S. S. Vedam, J. W. Wong, and E. Yorke. The Management of Respiratory Motion in Radiation Oncology Report of AAPM Task Group 76. *Med. Phys*, 33(10):3874 – 3900, 2006.

[11] M. Ries, B. Denis de Senneville, S. Roujol, Y. Berber, B. Quesson, and C. Moonen. Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues. *Magnetic Resonance in Medicine*, 64:1704 – 1712, 2010.

[12] S. Roujol, M. Ries, B. Quesson, C. Moonen, and B. Denis de Senneville. Real - Time MR - Thermometry and Dosimetry for Interventional Guidance on Abdominal Organs. *Magn Reson Med*, 63(4):1080 – 1087, 2010.

[13] P. L. Oliveira, B. Denis de Senneville, I. Dragonu, and C. T. Moonen. Rapid Motion Correction in MR - Guided High - Intensity Focused Ultrasound Heating Using Real - Time Ultrasound Echo Information. *NMR Biomed*, 23(9):1103 – 1108, 2010.

[14] M. Glitzner, B. Denis de Senneville, J. J. Lagendijk, B. W. Raaymakers, and S. P. Crijns. On-line 3D Motion Estimation Using Low Resolution MRI. *Phys Med Biol*, 60(16):N301 – N310, 2015.

[15] C. Plathow, M. Schoebinger, C. Fink, S. Ley, M. Puderbach, M. Eichinger, M. Bock, H. P. Meinzer, and H. U. Kauczor. Evaluation of Lung Volumetry Using Dynamic Three - Dimensional Magnetic Resonance Imaging. *Invest Radiol*, 40(3):173 – 179, 2005.

[16] A. Sawant, P. Keall, K. B. Pauly, M. Alley, S. Vasanawala, B. W. Loo Jr, J. Hinkle, and S. Joshi. Investigating the Feasibility of Rapid MRI for Image - Guided Motion Management in Lung Cancer Radiotherapy. *BioMed Res Int*, 2014:1 – 7, 2014.

[17] Q. Huang and Z. Zeng. A Review on Real - Time 3D Ultrasound Imaging Technology. *BioMed Res Int*, 2017:1 – 20, 2017.

[18] S. Miao, J. Z. Wang, and R. Liao. Chapter 12 - Convolutional Neural Networks for Robust and Real - Time 2-D/3-D Registration. In Zhou, S. K. and Greenspan, H. and Shen, D., editor, *Deep Learning for Medical Image Analysis*, pages 271 – 296. Academic Press, 2017.

[19] S. Fielden, L. Zhao, W. Miller, X. Fend, M. Wintermark, K. B. Pauly, and C. Meyer. Spiral - Based 3D MR Thermometry. *J Ther Ultrasound*, 3(Suppl 1):18, 2015.

[20] C. G. Overduin, Fütterer. J. J., and T. W. Scheenen. 3D MR Thermometry of Frozen Tissue: Feasibility and Accuracy During Cryoablation at 3T. *J Magn Reson Imaging*, 44(6):1572 – 1579, 2016.

[21] B. Stemkens, R. Tijssen, B. Denis de Senneville, J. J. Lagendijk, and C. A. van den Berg. Image Driven, Model - Based 3D Abdominal Motion Estimation for MR - Guided Radiotherapy. *Phys Med Biol*, 61(14):5335 – 5355, 2016.

[22] P. Baron, M. Ries, R. Deckers, M. de Greef, J. Tanttu, M. Köhler, M. Viergever, C. T. Moonen, and L. W. Bartels. In Vivo T2 - Based MR Thermometry in Adipose Tissue Layers for High - Intensity Focused Ultrasound Near - Field Monitoring. *Magn Reson Med*, 72(4):1057 – 1064, 2014.

[23] J. W. Wijlemans, M. de Greef, G. Schubert, L. W. Bartels, C. T. Moonen, M. A. van den Bosch, and Ries M. A Clinically Feasible Treatment Protocol for Magnetic Resonance - Guided High - Intensity Focused Ultrasound Ablation in the Liver. *Inves Radiol*, 50(1):24 – 31, 2015.

[24] N. Kothary, S. Dietrich, J. D. Louie, D. T. Chang, L. V. Hofmann, and D. Y. Sze. Percutaneous Implantation of Fiduacial Markers for Imaging - Guided Radiation Therapy. *Am J Roentgenol*, 192(4):1090 – 1096, 2009.

[25] T. E. Marchant, A. Skalski, and B. J. Matuszewski. Automatic Tracking of Implanted Fiducial Markers in Cone Beam CT Projection Images. *Med Phys*, 39(3):1322 – 1334, 2012.

[26] J.L. Foley, M. Eames, J. Snell, A. Hananel, N. Kassell, and J.-F. Aubry. Image - Guided Focused Ultrasound: State of the Technology and the Challanges That Lie Ahead. *Imaging Med.*, 5(4):357 – 370, 2013.

[27] L. Zhang and Z. B. Wang. High - Intensity Focused Ultrasound Tumor Ablation: Review of Ten Years of Clinical Experience. *Front Med China*, 4(3):294 – 302, 2010.

[28] S. P. Crijns, B. W. Raaymakers, and J. J. Lagendijk. Proof of Concept of MRI - Guided Tracked Radiation Delivery: Tracking One - Dimensional Motion. *Phys Med Biol*, 57(23):7863 – 7872, 2012.

[29] F. Wu, Z. B. Wang, W. Z. Chen, H. Zhu, J. Bai, J. Z. Zou, K. Q. Li, C. B. Jin, F. L. Xie, and H. B. Su. Extracorporeal High Intensity Focused Ultrasound Ablation in the Treatment of Patients with Large Hepatocellular Carcinoma. *Ann Surg Oncol*, 11(12):1061 – 1069, 2004.

[30] P. Ramaekers, M. de Greef, R. Berriet, C. T. W. Moonen, and M. Ries. Evaluation of a Novel Therapeutic Focused Ultrasound Transducer Based on Fermat's Spiral. *Phys Med Biol*, 62(12):5021 – 5045, 2017.

[31] M. Durante, R. Orecchia, and J. S. Loeffler. Charged - Particle Therapy in Cancer: Clinical Uses and Future Perspectives. *Nat Rev Clin Oncol*, 14(8):483 – 495, 2017.

[32] H. H. Kampinga and E. Dikomey. Hyperthermic Radiosensetization: Mode of Action and Clinical Relevance. *Int J Radiat Biol*, 77(4):399 – 408, 2001.

[33] G. Borasi, G. Russo, F. Vicari, A. Nahum, and M. C. Gilardi. Experimental Evidence for the Use of Ultrasound to Increase Tumor - Cell Radiosensitivity. *Trans Cancer Res*, 3(5):512 – 520, 2014.

CHAPTER 12

Samenvatting

Beeldgeleide HIFU en EBRT hebben de potentie om lokaal en regionaal kanker te behandelen op een volledig niet-invasieve manier. Dergelijke therapien zijn echter onderhevig aan onzekerheden veroorzaakt door beweging van de patint en/of de tumor. Deze kunnen afkomstig zijn van verschillende bronnen zoals positioneringsfouten van de patint, spontane beweging, lange termijn drifts van de organen en/of ademhaling. Als met dergelijke verplaatsingen geen rekening wordt gehouden, kan het ertoe leiden dat de therapeutische energie wordt afgeleid van de doelpathologie, waardoor het risico van onderbehandeling en impliciet van ziekterecidief toeneemt. Bovendien kan, in het geval van grote bewegingseffecten, het ontbreken van een compensatiestrategie leiden tot ernstige complicaties door collaterale schade aan gezonde weefsels. De bijdrage van dit proefschrift aan technologische en methodologische vooruitgang op het gebied van beeldgeleide HIFU en EBRT behandelingen is drievoudig:

1. Het sluit in eerste instantie aan op de state-of-the-art methodologie van real-time MR-geleide HIFU therapien, identificeert daarbij specifieke tekortkomingen en stelt strategien voor om deze beperkingen op te lossen.

2. State-of-the-art methodologie wordt vervolgens selectief en gezamenlijk gebruikt om grootschalige frameworks te ontwerpen en te implementeren die de gehele interventionele workflow omvatten en toegespitst zijn op bewegingscompensatie gedurende niet-invasieve stralings therapien. Het ontwerp en de validatie van deze frameworks rondom reeds bestaande klinische workflows heeft verdere identificatie mogelijk gemaakt van state-of-the-art limitaties in termen van beeldregistratie methodologie.

3. Het geven van nieuwe oplossingen voor de eerder aangeduide beperkingen, in het bijzonder afgestemd op de specifieke klinische eisen gesteld door beeldgeleide HIFU en EBRT interventies.

In het vervolg wordt het onderwerp, belangrijkste resultaten en conclusies die volgen uit elk hoofdstuk kort samengevat.
Het optical flow algoritme van Horn & Schunck is eerder voorgesteld als een haalbare oplossing voor elastische realtime MR-gebaseerde ademhaling bewegingscompensatie door verschillende onafhankelijke studies. De keuze werd door deze studies gerechtvaardigd door de mogelijkheid van deze methode tot het geven van dichte en elastische vervormingen, haar snelle numerieke schema en haar klein aantal invoerparameters. Echter, de grote waarde die het algoritme hecht aan de behouden-voxel-intensiteit hypothese, kan leiden tot misregistraties op plekken waar deze aanname lokaal geschonden wordt. **Hoofdstuk 2** adresseert deze limitatie door een verbeterde versie van het orginele Horn & Schunck algoritme voor te stellen met een hogere tolerantie voor lokale intensiteitsvariaties van grijswaarden die niet veroorzaakt worden door beweging. Dit is bereikt door het vervangen van de kwadratische data-fidelity term van de Horn & Schunck methode door een term gebaseerd op de L_1 lineaire norm. De voorgestelde aanpak was gevalideerd op een dynamische 2D MR

dataset bestaande uit 3000 beelden van de buikholte van twee vrij-ademende gezonde vrijwilligers. Vergeleken met het Horn & Schunck algoritme hebben de resultaten van de voorgestelde methode een verbeterde robuustheid voor instroom effecten door arterile pulsaties laten zien. De verhoogde robuustheid werd verkregen met behoud van mogelijkheden tot realtime MR-gebaseerde bewegingsestimatie. Voor zover de auteur weet was dit de eerste studie om het gedrag te analyseren van een L_1-gebaseerde optical flow functionaal in de context van realtime MR-geleide HIFU. Snelle beeldvorming tijdens dergelijke toepassingen, speelt echter slechts een secundaire rol als invoer voor bewegingsestimatie algoritmen. Zijn primaire rol is het aanleveren van temperatuurbeelden van het behandelde gebied via MR-thermometrie. Een neveneffect van snelle MR-beeldvorming is dat de resulterende beelden vaak worden gehinderd door een lage signaal-ruisverhouding (SNR). Dit, op zijn beurt, leidt tot onnauwkeurige temperatuurmetingen met een directe invloed op de therapiecontrole, monitoring en therapeutische eindpuntherkenning. Bestaande oplossingen voor ruisonderdrukking in MR-thermometrie verbeteren de precisie maar gaan vaak ten koste van de nauwkeurigheid. **Hoofdstuk 3** stelt een nieuwe niet-lokale ruisonderdrukkingsmethode voor die is gericht op realtime MR-thermometrie. In tegenstelling tot eerdere benaderingen, verbetert de voorgestelde methode de precisie van MR-thermometrie, terwijl tegelijkertijd de output nauwkeurigheid wordt gemaximaliseerd. Bijzondere aandacht werd besteed aan computationele latentie, zodat de realtime beschikbaarheid van de temperatuurmetingen werd gehandhaafd. De aanpak werd gevalideerd in drie complementaire experimenten: een simulatie, een ex-vivo en een in vivo onderzoek. Vergeleken met bestaande ruisonderdrukkingsmethoden van MR-thermometrie levert de voorgestelde aanpak een statistisch significante verbetering op $p = 0.05$ significantieniveau van de kwaliteit van de ontruisde data. **Hoofdstukken 4** en **5** "zoomen uit" van sonicatie-niveau gerelateerde methodologische tekortkomingen en focussen op grootschalige bewegingscompensatie oplossingen die de volledige workflow van MR-geleide niet-invasieve stralingsinterventies omvatten. Voorafgaand aan het werk uitgevoerd in **hoofdstuk 4**, hebben onderzoeken naar bewegingscompensatie in organen in de buikholte tijdens zowel HIFU als EBRT, gefocust op het omgaan met ademhalingsverplaatsingen gedurende de therapeutische energie-afgifte. **Hoofdstuk 4** laat zien dat dit conceptueel onvoldoende kan zijn omdat tijdens langdurige therapien buikorganen ook een langdurige drift kunnen ondergaan vanwege spijsvertering, spierontspanning en/of metabolische processen. Als zodanig, is een oplossing voorgesteld die gericht is op het compenseren van langetermijnverschuivingen die zich voordoen gedurende de volledige duur van de interventie. Dit werd bereikt door een reeks 3D MR scans quasi-uniform verspreid over de behandeling te integreren in een bestaande klinische workflow. De scans werden vervolgens geregistreerd op een referentievolume verkregen aan het begin van de therapie. De resulterende verplaatsingen toonden aan dat tijdens langdurige HIFU therapien de menselijke lever en nieren drifts kunnen ondergaan van meer dan 5 mm, waardoor geaccepteerde therapeutische marges worden over-

schreden. Gebaseerd op in-vivo verkregen data van een varkenslever, toonde de studie ook aan dat de geschatte verplaatsingen voldoende nauwkeurig zijn voorzien van een voldoende lage tijdslatentie om on-the-fly aanpassingen van het interventieplan mogelijk te maken en aggregatie van de thermische dosis in een gemeenschappelijke referentieruimte. Het veelbelovende karakter van deze resultaten hebben geleid naar het ontwerp van het uitgebreide framework beschreven en gevalideerd in **hoofdstuk 5**. Dit nieuwe framework richtte zich zowel op ademhalingsverplaatsingen tijdens de individuele energie-afgifte als op lange termijnafwijkingen die zich voordoen over de volledige duur van de therapie. Extra componenten zijn opgenomen in het framework voor kwaliteitsborging en koppelingsdoeleinden. Het koppelen van de verschillende bewegings-/vervormingsestimatie componenten maakte het mogelijk om elk gedurende de therapie opgenomen MR beeld ruimtelijk te relateren aan een hoofdplanningsbeeld dat was opgenomen aan het begin van de interventie. De resulterende verplaatsingen zijn gebruikt voor twee doeleinden: 1) "Down-stream" propagatie van het oorspronkelijk geplande cluster van cellen, zodat de nieuwe locatie van de sonicaties overeenkomt met het initiele volume dat geableerd moest worden, en 2) "Up-stream" propagatie van de thermische dosis, zodat de voortgang van de therapie kan worden gevolgd op het planningsbeeld op een ruimtelijk consistente manier. Dit uitgebreide framework is gevalideerd in drie complementaire experimenten: een onderzoek op een gel-fantoom dat een bekend bewegingspatroon onderging, een onderzoek op de buikholte van 10 gezonde vrijwilligers en 3 in vivo ablaties op varkenslevers. De fantoom- en de vrijwilligersdata werden offline verwerkt terwijl de schattingen en correcties tijdens de 3 dierexperimenten on-the-fly werden uitgevoerd. De fantoom- en gezonde vrijwilligersexperimenten toonden aan dat het voorgestelde framework voor de bewegings-schatting/compensatie in staat is tot het geven van schattingen met sub-voxel precisie en nauwkeurigheid. De resultaten bevestigden ook opnieuw wat werd waargenomen in **hoofdstuk 4**, namelijk dat gedurende lange HIFU therapien, de menselijke lever en nieren verplaatsingen kunnen vertonen die acceptabele therapeutische marges overschrijden. Tenslotte tonen de 3 succesvolle dierstudies de compatibiliteit van de voorgestelde aanpak met de workflow van een MR-geleide HIFU interventie onder klinische omstandigheden aan.

De ontwikkeling van het principe achter het EVolution-registratie algoritme en het MR+USg-HIFU systeem, legden de basis voor het adresseren van beperkingen van state-of-the-art beeldregistratie methoden. Deze werden gedentificeerd tijdens het ontwerpen, evalueren en valideren van de frameworks beschreven in **hoofdstukken 4** en **5**. Deze beperkingen bestaan onder andere uit: 1) De schaarste aan robuuste en snelle elastische multimodale registratie-algoritmen en 2) De dubbele rol van snelle MR-beeldvorming tijdens HIFU-sonicaties, van zowel tracking als thermometrie, waarvoor een compromis vereist is tussen de acquisitieparameters van de gebruikte MR-sequentie. De focus van **hoofdstuk 6** is daarom uitsluitend gericht op de technische details en evaluatie van het multimodal registratie-algoritme EVolution. Hoewel het algoritme over

het algemeen op gelijke voet stond met state-of-the-art methoden, toonde EVolution een verbeterde robuustheid voor veranderlijke structuren. Bovendien behoudt het ook de voordelen van variatie registratiemethoden, te weten: het geeft elastische en dichte vervormingen, het impliceert snelle numerieke schema's en het vereist een laag aantal invoerparameters. De eigenschappen van het EVolution-algoritme samen met de onafhankelijke stroom van B-mode US beelden, verschaft door het hybride MR+USg HIFU systeem, bood de mogelijkheid de realtime tracking gedurende sonicaties over te hevelen van de MR scanner naar de US scanner. Het is een bekend gegeven dat B-mode US beelden normaal gesproken verandert worden door vluchtige intensiteitsvariaties van de grijswaarden in de vorm van spikkelruis. Het was al aangetoond dat dit problematisch is voor bestaande intensiteitsgebaseerde registratiemethoden. **Hoofdstuk 7** evalueert de mogelijkheden van het EVolution algoritme tot realtime US gebaseerde ademhalingstracking. Vanwege zijn intrinsieke robuustheid voor veranderlijke structuren werd geopperd dat EVolution ongevoelig is voor spikkelruis. De aanpak was gevalueerd voor zowel contact als afstandelijke echografische beelden, opgenomen in de buikholte van 7 gezonde vrijwilligers. De methode gaf gemiddeld een nauwkeurigheid van 1.5 mm en submillimeter precisie. Samen met een computationele prestatie van 20 beelden/s, maakte dit van EVolution een haalbare oplossing voor het gegeven probleem. Als zodanig is aangenomen dat een groot deel van de aangeduide limitaties van interventionele geleiding gedurende therapie-afgifte is geadresseerd. Daarom verschuift **hoofdstuk 8** de focus naar begwegingsgerelateerde onnauwkeurigheden die ontstaan gedurende de herplanningsfase van kankertherapie. Beeldgestuurde EBRT is normaal gesproken gebaseerd op de registratie van de planning CT op de CBCT beelden die op verschillende momenten gedurende de behandeling worden opgenomen. In de kliniek wordt dit momenteel gedaan door hoge-Z markers te implanteren in de tumor en/of zijn directe omgeving. De CT en de CBCT beelden zijn vervolgens geregistreerd op basis van de locatie van de geplanteerde markers. Los van het feit dat deze implantatie ongemak oplevert voor de patint is deze methode alleen toepasbaar op rigide verplaatsing. Dit is duidelijk niet optimaal voor vervormbare organen. Bovendien zijn de geplanteerde markers slechts surrogaten voor de werkelijke tumor positie, wat de mogelijkheid openlaat voor een relatieve drift tussen marker en tumor. Een puur beeldgebaseerde registratie zou daarom wenselijk zijn. Een moeilijkheid hierbij is dat CBCT beelden vaak opgenomen worden met een klein aantal X-ray projecties en/of met lage stralingsintensiteit. Dit leidt vaak tot een laag contrast in de beelden, een lage signaal-ruisverhouding en streep-artefacten. Dit laatste in het bijzonder introduceert intensiteits-inconsistenties tussen CT en CBCT beelden, wat problematisch is voor een breed scala aan registratiemethoden. Vanwege zijn onderliggende principe, was het EVolution algoritme verondersteld robuust te zijn voor zulke effecten. **Hoofdstuk 8** onderzoekt deze veronderstelling door de methode te evalueren voor verschillende scenario's waarbij CT/CBCT data opgenomen is bij kankerpatinten met verschillende pathologin. De resultaten toonden aan dat de artefacten van CBCT beelden slechts

een sub-millimeter impact hebben op de nauwkeurigheid en precisie van het EVolution algoritme. Bovendien had de methode gemiddeld 60 s nodig om te convergeren voor $256 \times 256 \times 256$ beelden, waardoor zij toepasbaar is in online scenario's.
De Horn & Schunck en EVolution algortimen hebben een belangrijke rol gespeeld in de bewegingsestimatie en bewegingscompensatie oplossingen voorgesteld in dit proefschrift. Beide methoden hebben echter hun oorsprong in de optische beeldverwerking. Wanneer zij worden toegepast op medische beelden is er daarom geen garantie dat de geschatte deformaties anatomisch plausibel zijn. **Hoofdstuk 9** onderzoekt dit aspect door de methoden te analyseren vanuit het perspectief van twee bestaande contour-gebaseerde criteria en drie extra indicatoren van anatomische plausibiliteit. Twee nieuwe versies van de twee algoritmen zijn ook voor voorgesteld, die de gladheid-regularisatie door een beperking op de deformatie-samendrukbaarheid vervangt. Deze beperking is gerechtvaardigd door het feit dat de meerderheid van de biologische weefsels onsamendrukbaar is. De resultaten toonden aan dat hoewel alle vier methoden vergelijkbare contour-matching mogelijkheden hadden, de voorgestelde onsamendrukbare algoritmen deformaties gaven die beter overeenkomen met de materile eigenschappen van biologische zachte weefsels. Ondanks deze verbeteringen toont **hoofdstuk 9** aan dat zulke methoden selectief moeten worden ingezet, afhankelijk van de physische eigenschappen van de waargenomen anatomie.

Publications

Journal articles

C. Zachiu, B. Denis de Senneville, C. Moonen, and M. Ries. A Framework for the Correction of Slow Physiological Drifts During MR-Guided HIFU Therapies: Proof of Concept. *Med. Phys.*, 42(7): 4173 - 4148, 2015.

C. Zachiu, N. Papadakis, M. Ries, C. Moonen, and B. Denis de Senneville. An Improved Optical Flow Tracking Technique for Real-Time MR-Guided Beam Therapies in Moving Organs. *Phys. Med. Biol.*, 60(23): 9003 - 9029, 2015.

B. Denis de Senneville, **C. Zachiu**, M. Ries, and C. Moonen. EVolution: an edge-based variational method for non-rigid multi-modal image registration. *Phys Med Biol*, 61(20): 7377 - 7396, 2016.

C. Zachiu, M. Ries, C. Moonen, and B. Denis de Senneville. An Adaptive Non-Local-Means Filter for Real-Time MR-Thermometry. *IEEE Trans. Med. Imaging*, 36(4): 904 - 916, 2017.

C. Zachiu, M.G. Ries, P. Ramaekers, J.L. Guey, C.T.W. Moonen, and B. Denis de Senneville. Real-Time Non-Rigid Target Tracking for Ultrasound-Guided Clinical Interventions. *Phys. Med. Biol.*, 62(20): 8154-8177, 2017.

C. Zachiu, B. Denis de Senneville, I.D. Dmitriev, C.T.W. Moonen, and M. Ries. A Framework for Continuous Target Tracking During MR-Guided High Intensity Focused Ultrasound Thermal Ablations in the Abdomen. *J Ther Ultrasound*, 5: 27, 2017.

C. Zachiu, B. Denis de Senneville, R.H.N. Tijssen, A.N.T.J. Kotte, A.C. Houweling, L.G.W. Kerkmeijer, J.J.W. Lagendijk, C.T.W. Moonen, and M. Ries. Non-rigid CT/CBCT to CBCT registration for online external beam radiotherapy guidance. *Phys. Med. Biol.*, 63(1): 015027, 2017.

C. Zachiu, B. Denis de Senneville, C. T. W. Moonen, B. W. Raaymakers and M. G. Ries. Anatomically plausible models and quality assurance criteria for online mono- and multi-modal medical image registration. *Phys. Med. Biol.*, 2018. [In press]

Conference proceedings

C. Zachiu, B. Denis de Senneville, S. P. M. Crijns, B. W. Raaymakers, C. T. W. Moonen, and M. G. Ries. A Framework for Slow Physiological Mo-

tion Compensation During HIFU Interventions in the Liver: Proof of Concept. *International Symposium on Focused Ultrasound*, 2014, Bethesda, Maryland, USA.

C. Zachiu, B. Denis de Senneville, S. P. M. Crijns, B. W. Raaymakers, C. T. W. Moonen, and M. G. Ries. A Framework for Slow Physiological Motion Compensation During HIFU Interventions in the Liver: Proof of Concept. *Symposium of the Focused Ultrasound Foundation*, 2014, Rockville, Maryland, USA.

C. Zachiu, B. Denis de Senneville, S. P. M. Crijns, B. W. Raaymakers, C. T. W. Moonen, and M. G. Ries. A correction framework for slow physiological drifts during MR-guided HIFU interventions in the abdomen: Proof of Concept. *Focused Ultrasound Therapy Symposium*, 2016, London, United Kindgdom.

C. Zachiu, N. Papadakis, M. G. Ries, C. T. W. Moonen, and B. Denis de Senneville. An Improved Real - Time Optical Flow Tracking Technique for MR - Guided Beam Therapies in Moving Organs. *International Symposium on Biolmedical Imaging*, 2016, Prague, Czech Repuplic.

C. Zachiu, N. Papadakis, M. G. Ries, C. T. W. Moonen, and B. Denis de Senneville. An improved tracking technique for real-time MR-guided beam therapies in moving organs. *International Society of Magnetic Resonance In Medicine*, 2016, Singapore.

J. Tan, **C. Zachiu**, P. Ramaekers, C. T. W. Moonen, and M. G. Ries. Optical Flow Based Motion Tracking in the Abdomen: Ultrasound vs MRI; A direct comparison. *International Symposium on Focused Ultrasound*, 2016, Bethesda, Maryland, USA.

C. Zachiu, M. G. Ries, C. T. W. Moonen, and B. Denis de Senneville. An adaptive non-local denoising method for real-time MR-thermometry. *International Society of Magnetic Resonance in Medicine*, 2017, Hawaii, USA.

Acknowledgements

Working towards my PhD has been a journey which not only lead to my growth as a scientist, but first of all as a person. At the end of this road, I would like to sincerely thank all the people who have aided me in my journey in one way or another.

First of all I would like to thank my promoters Prof. Dr. Chrit Moonen and Prof. Dr. Bas Raaymakers for having me on-board in the advancement of this fascinating technology: completely non-invasive cancer treatment. This is something I would not have thought possible if I hadn't joined your group here at the UMC Utrecht. Chrit, your ever-positive and calm attitude have been an inspiration. It taught me that regardless of how difficult a problem or a situation is, approaching it in such a mindstate, can make them 10 times easier. I really hope that the work we did in the past few years and the work that we will do in the years to come, will finally lead to a widespread clinical adoption of image-guided HIFU. My thanks also goes to you Bas, for your constant support ever since my first employment interview, until this point and beyond. I sometimes wonder how in your over-booked schedule, while getting the first MR-Linac into clinics, you somehow always found time to discuss about my project and ideas and greeted me with an open door. I hope that our collaboration will continue for many years to come and that together we will manage to bring the MR-Linac to its full potential.

Baudouin, there are no words to thank you for the guidance and support that you provided me ever since we met in Bordeaux. Thank you for introducing me to this fascinating field of image-guided non-invasive therapies and for the patience and diligence in training me from scratch, always with a smile on your face. Your amazing creativity and calm and open-minded nature have been an inspiration to me. I truly hope that regardless on where life's paths may take us, there will always be time for a quick scientific or friendly chat. Chapeau!

Mario, working with you during these years was truly a privilege. Just like with Baudouin, mere words are not enough to thank you for the guidance you've provided, always encouraging me to go one step beyond my abilities. This lead to my development as a scientist far beyond any level I imagined before undertaking this PhD. Your critical attitude towards any new idea: "What's the angle?", "What's the bigger picture?", "Let's take a step back" and a few others, will stick with me for the remainder of my scientific career as after all, I truly believe that this always paves the way for high quality science.

I would like to express my sincere gratitude to the thesis assessment committee members: Prof. Dr. Paul J. van Diest, Prof. Dr. Chris H. J. Terhaard, Prof. Dr. Inne H. M. Borel Rinkes, Prof. Dr. Jurgen J. Futterer and Prof. Dr. Wiro J. Niessen for evaluating this thesis.

I would also like to give a special thanks to you, Ivan, for helping me bring my computer programming skills to the next level. Your direct and highly educated manner of criticizing the shape, output and performance of any piece of code

has definitively changed my view on how proper programming should look like. At the same time, I will also always appreciate our insightful discussions during the coffee breaks, about the past, present and future of image processing.
My thanks also goes to the other PhD's, post-docs and staff members with whom I had the pleasure of working in the past few years, in no particular order: Markus (paranymph), Yulia, Xiaoxu, Charles, Pascal, Martijn, Pim (paranymph), Marijolein, Gerald, Marc, Jean-Michel, Charis, Maxence, Matteo and any others that I may have forgotten. The lunches, coffee breaks and OIO events and outings that we had together, made this journey a lot more enjoyable. I wish you all the best!
Un mic paragraf in română pentru a le mulţumi părinţilor mei Cornel şi Ana pentru toată grija şi sprijinul pe care mi l-au oferit ca eu să ajung mai departe decât mi-aş fi imaginat vreodată posibil. Pe parcursul anilor, a fost un drum cu multe greutăţi şi de-a lungul căruia aţi făcut nenumărate sacrificii ca eu să fiu unde sunt astăzi. Vreau ca această carte să vă fie în mare parte dedicată vouă şi în acelaşi timp să fie o amintire că sacrificiile pe care le-aţi făcut nu au fost in zadar. Cu capul plecat in faţa voastră: Vă mulţumesc!
Cătă şi Iuli, nu vă faceţi griji, nu v-am uitat. Probabil e greu de crezut, dar nu aş fi ajuns să scriu o teză de doctorat fără sprijinul şi încurajările voastre. Orice început e greu, iar în ceea ce priveste cariera mea ştiinţifică, aici e momentul cheie în care aţi contribuit amândoi. Cătă, n-o să uit niciodată zecile de desene pe care le făceai până reuşeam să elucidăm vreo problemă de fizică sau de matematică, ce-mi punea dificultăţi majore când eram doar un învăţăcel. Cât despre tine Iuli, rolul de profesoară pe care ţi l-ai însuşit voluntar (fie că am vrut sau nu) m-a ajutat mai mult decât îţi imaginezi să mă pun pe picioare în primii ani de şcoala. Fratele vostru mai mic vă mulţumeşte nu numai pentru asta, ci şi pentru tot timpul pentru care mi-aţi fost alături. Cinste vouă!
My beloved Tianyu, many people do not believe in destiny, but you came into my life when I needed the most. Your unconditional love and support has helped me get through the most difficult moments of my training as a PhD. Thank you for your patience and for teaching me to greet each day with a smile. 我愛你！我總是會。

Curriculum vitae

Cornel Zachiu was born on September 4th, 1988 in the city of Sebeş, Romania. After completing his studies in mathematics and informatics at the "Lucian Blaga" high-school of Sebeş in 2007, he enrolled at the Technical University of Cluj - Napoca and obtained his diploma in Electronics and Telecommunications Engineering in 2011. He then pursued the courses of the "Signal and Image Processing" cotutelle masters programme, obtaining in 2013 his double diploma from both the Technical University of Cluj - Napoca and the University of Bordeaux, with the master thesis entitled "Real-Time Image Based Deformation Estimation of Abdominal Organs". In the same year he started pursuing a PhD within the HIFU group at the University Medical Center Utrecht, The Netherlands, under the supervision of Dr. Mario G. Ries, Dr. Baudouin Denis de Senneville, Prof. Dr. Chrit T. W. Moonen and Prof. Dr. Bas W. Raaymakers. The results of this research, focusing on real-time and online image-based motion compensation during non-invasive cancer therapies in moving organs, are the topic of this thesis.
Cornel is currently hired as an applied mathematician and computer scientist in the Department of Radiotherapy and University Medical Center Utrecht, where he continues to pursue his vision of a fully automated motion compensation solution encompassing the entire work-flow of a non-invasive cancer treatment.