Automation of ultrasound image acquisition and interpretation for dose guidance in prostate cancer radiotherapy workflows

Automation of ultrasound image acquisition and interpretation for dose guidance in prostate cancer radiotherapy workflows

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op dinsdag 4 juni 2019 om 13:30 uur

door

Saskia Maria Camps

geboren te Breda

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof.dr.ir. A.B. Smolders	(Eindhoven University of Technology)
1^e promotor:	prof.dr.ir. P.H.N. de With	(Eindhoven University of Technology)
2^e promotor:	prof.dr.ir. F. Verhaegen	(Department of radiation oncology - MAASTRO)
copromotor:	dr. D. Fontanarosa	(Queensland University of Technology)
leden:	prof.dr.ir. C.H. Slump	(University of Twente)
	prof.dr. H.P. Beerlage	(Amsterdam University Medical Center)
	prof.dr.ir. M. Mischi	(Eindhoven University of Technology)
	dr. E. Harris	(The Institute of Cancer Research, London, UK)
	dr.ir. R.G. Aarnink	(Philips Healthcare)

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

Summary

Automation of ultrasound image acquisition and interpretation for dose guidance in prostate cancer radiotherapy workflows

Radiotherapy (RT) is one of the possible curative treatment options for prostate cancer. It aims at irradiating tumor tissue, while at the same time sparing normal tissue as much as possible. Frequent imaging during the course of the RT treatment can be used to guide radiation dose and so increase delivery to the tumor tissue and decrease toxicity.

Ultrasound (US) imaging allows real-time volumetric imaging in the RT environment. In addition, this image modality facilitates relatively fast image acquisition, it does not deliver any potentially harmful radiation to the patient and it is cost-effective. Despite these advantages, the use of US imaging for guidance in prostate cancer RT workflows is presently not widespread. This can be partially attributed to the need for a trained operator who needs to interpret the US images during manual image acquisition in order to verify if the correct anatomical structures are visualized with sufficient image quality. In addition, electron density information from a CT scan is currently still necessary to calculate radiation dose. This implies that a CT scan has to be acquired to understand if re-planning of a RT treatment plan could be beneficial.

The aim of this PhD work was to optimize the US guided RT workflow of prostate cancer patients by automating steps of this workflow and to explore if US imaging can help to understand whether RT treatment plan adaptation is necessary. This could make the use of US imaging more appealing and allow prostate cancer patients to fully benefit from the unique characteristics of this imaging modality.

In the first part of this thesis, a brief introduction is given on US imaging and RT treatment in general (**Chapter 1**) and a more in-depth review is provided on the use of US imaging in external beam RT workflows of prostate cancer patients (**Chapter 2**).

The second part comprises the work related to automation of US image acquisition and interpretation. **Chapter 3** details a proof-of-concept study, which was performed on a male pelvic phantom. In this study, an algorithm is introduced that uses anatomical information derived from a CT scan, to automatically propose an US probe setup to the operator. This probe setup should allow visualization of the relevant anatomical structures. For evaluation purposes, several proposed probe setups were reproduced using a robotic arm and US volumes of the phantom were acquired. Quantitative evaluation and visual inspection of these volumes showed that the algorithm was able to propose probe setups that fulfill the clinical requirements. In addition, the experience was gained that the currently used robotic arm was not suitable for

US probe positioning on the body of a patient, due to the occurrence of unexpected arm movements.

In **Chapter 4**, the phantom work was used as a basis to create an algorithm that can propose a prostate cancer patient-specific transperineal US (TPUS) probe setup. This algorithm also performs the pre-processing by first identifying the perineal skin area of the patient, which is accessible for TPUS probe placement in clinical practice, using a simulation CT scan. Subsequently, it proposes several possible probe setups on this skin area and selects the optimal setup using anatomical structure delineations corresponding to the CT scan. In order to test the algorithm, patient-specific probe setups for three prostate cancer patients were proposed. These setups were then compared to the setups used by radiation oncologists during actual TPUS image acquisitions of these patients using a mechanical arm, instead of a robotic arm. This comparison revealed that the algorithm seems able to propose setups that allow visualization of 94% of the anatomical structures that were also visualized by the radiation oncologist. In addition, the algorithm-proposed setups allowed visualization of 100% of the anatomical structures that are of interest for US guided RT.

Chapter 5 introduces the first step towards fully automatic interpretation of US images. To allow for US image acquisition during a treatment fraction, the US probe would need to be positioned on the body of the patient prior to the treatment commencement using, for example, a mechanical arm, as the operator cannot be present in the room during radiation dose delivery. Changes in anatomical structures or small motion of the patient during the treatment delivery can compromise US image quality, which requires identification. This chapter details a prototype deep learning algorithm that can automatically assign a quality score to 2D TPUS images of the male pelvic region. This quality score gives information on the usability of the specific image in the US guided RT workflow. The promising results show that the algorithm can achieve an accuracy of 94% which is comparable with the three experts (accuracy range: 92% - 97%) who were consulted in this study.

In the third part of this thesis, the work related to obtaining CT electron density information without actually acquiring a CT scan is described. **Chapter 6** proposes a new concept workflow where deformable image registration between two transabdominal US (TAUS) volumes, acquired at simulation stage and prior to a treatment fraction, is performed. The deformation field so obtained is then applied to the simulation CT scan, resulting in the creation of a pseudo-CT scan. This workflow has been applied to image data of three prostate cancer patients, who had images acquired during several time points. Comparison of the created pseudo-CT scans and available ground truth CT scans showed improvements up to 11.2% in gamma failure for dose in the prostate and up to 20.5% improvement in gamma failure for Hounsfield units in anus and rectum. These results seem to

confirm the hypothesis that the pseudo-CT scan represents the anatomy of the patient at treatment stage better than the simulation CT scan does.

In **Chapter 6** the deformation field between the two US volumes was created using only deformable image registration, while varying the registration parameters up to a limited extend. These parameters, such as, the number of iterations can impact the performance of the registration algorithm significantly. For this reason, in **Chapter 7** a larger part of the registration parameter space were explored. In addition, six additional approaches to create the pseudo-CT scans were used. These approaches did not only include deformable registration, but also only rigid registration or a combination of both. This parameter space exploration and application of additional approaches was tested on image data from one prostate cancer patient, who had images available at five different time points. It was shown that at least four out of eight approaches resulted in more representative pseudo-CT scans with an average Dice similarity coefficient improvement of 40.1% (range: 4.3% - 126.5%) in comparison with the corresponding simulation CT scans.

Finally, in Chapter 8 the results presented in this thesis are thoroughly discussed and a vision is given for the future use of US imaging for guidance in prostate cancer RT workflows, as well as its use for guidance during other medical procedures.

Samenvatting

Automatisering van acquisitie en interpretatie van echobeelden voor dosisgestuurde radiotherapie behandeling van prostaatkanker

Radiotherapie (RT) is één van de behandelmethodes voor gelokaliseerde prostaatkanker die potentieel tot genezing kunnen leiden. Het doel van deze behandelmethode is om tumorweefsel te bestralen, terwijl tegelijkertijd normaal weefsel zo veel mogelijk wordt gespaard. Medische beelden die gedurende het RT-behandeltraject worden gemaakt, kunnen gebruikt worden om de stralingsdosis te sturen. Dit kan de stralingsafgifte aan de tumor verhogen en de radiotoxiciteit van het omliggende weefsel verlagen.

Echografie maakt het mogelijk om real-time volumetrische beelden te maken in een RT-omgeving. Daarnaast faciliteert deze beeldmodaliteit een relatief snelle beeldvorming, maakt het geen gebruik van straling die potentieel schadelijk is voor de patiënt en is het kosteneffectief. Ondanks deze voordelen is het gebruik van echografie voor sturing van stralingsdosis gedurende RT-procedures voor prostaatkanker niet gangbaar. Dit kan deels worden toegeschreven aan het feit dat een getrainde operator de echobeelden gedurende de handmatige acquisitie moet interpreteren om te verifiëren dat de juiste anatomische structuren in beeld zijn gebracht en dat de beeldkwaliteit voldoende is. Daarbij is, op dit moment, de elektronendichtheidsinformatie van een CT-scan nog steeds nodig om stralingsdosisberekeningen te kunnen uitvoeren. Dit impliceert dat een CT-scan gemaakt moet worden om te achterhalen of een patiënt baat zal hebben bij het herplannen van een RT-behandelplan.

Het doel van dit promotie-onderzoek was het optimaliseren van echogestuurde RT-procedures voor prostaatkankerpatiënten. Hiervoor zijn verschillende stappen in deze RT-procedures geautomatiseerd en is er onderzocht of echografie gebruikt kan worden om te beslissen of het herplannen van een RT-behandelplan nodig is. Beide aspecten zouden het gebruik van echografie aantrekkelijker kunnen maken en prostaatkankerpatiënten de mogelijkheid bieden om te profiteren van de unieke karakteristieken van deze beeldmodaliteit.

In het eerste deel van dit proefschrift wordt een introductie gegeven over echografie en RT-behandelingen in het algemeen (**Hoofdstuk 1**). In **Hoofdstuk 2** is een gedetailleerder overzicht te vinden van het huidige gebruik van echografie gedurende uitwendige RT-procedures van prostaatkankerpatiënten.

Het automatiseren van de acquisitie en interpretatie van echobeelden is beschreven in het tweede deel van dit proefschrift. In **Hoofdstuk 3** is een conceptstudie beschreven, die is uitgevoerd op een fantoom van de mannelijke pelvis. In deze studie wordt een algoritme geïntroduceerd dat gebruik maakt van de anatomische informatie verkregen van een CT-scan om automatisch een

geschikte echoprobe-opstelling aan de operator te suggereren. Deze probe-opstelling moet het visualiseren van de relevante anatomische structuren mogelijk maken. Ter evaluatie is een robotarm gebruikt om verschillende probe-opstellingen te reproduceren en zijn er vervolgens echovolumes van het fantoom gemaakt. Kwantitatieve evaluatie en visuele inspectie van deze volumes laten zien dat het algoritme in staat is om probe-opstellingen te suggereren die voldoen aan de klinische eisen. Daarbij kon worden geconcludeerd dat de gebruikte robotarm niet geschikt is om de echoprobe te positioneren op het lichaam van de patiënt, vanwege het feit dat de arm af en toe ongecontroleerd beweegt.

In **Hoofdstuk 4** is het fantoomwerk gebruikt als basis om een algoritme te ontwerpen dat een transperineale echoprobe-opstelling voor een specifieke prostaakankerpatiënt kan suggereren. Dit algoritme identificeert allereerst het perineale huidoppervlak van de patiënt dat toegankelijk is voor plaatsing van de transperineale echoprobe in een klinische setting. Vervolgens suggereert het algoritme verscheidene probe-opstellingen op dit huidoppervlak en selecteert het de optimale opstelling, waarbij gebruik wordt gemaakt van segmentaties van de anatomische structuren van de bijbehorende CT-scan. Er zijn patiëntspecifieke probe-opstellingen gesuggereerd voor drie prostaatkankerpatiënten om het algoritme te testen. Deze opstellingen zijn vervolgens vergeleken met de opstellingen die daadwerkelijk zijn gebruikt door radiotherapeut-oncologen gedurende het maken van transperineale echobeelden van deze patiënten. Hierbij is gebruik gemaakt van een mechanische arm in plaats van een robotische arm. De vergelijkingen laten zien dat het algoritme in staat is om een opstelling te suggereren die het potentieel mogelijk maakt om 94% van de anatomische structureren, die ook door de radiotherapeut-oncoloog in beeld zijn gebracht, te visualiseren. Daarnaast lijkt de opstelling die is gevonden door het algoritme visualisatie van 100% van de anatomische structuren, die relevant zijn voor echo gestuurde RT, mogelijk te maken.

De eerste stap in de richting van volledig geautomatiseerde interpretatie van echobeelden is omschreven in **Hoofdstuk 5**. Teneinde echobeelden gedurende de bestraling te maken, is het noodzakelijk dat de echoprobe voor de start van de behandeling op het lichaam van de patiënt wordt gepositioneerd met behulp van bijvoorbeeld een mechanische arm. Dit is vereist omdat de operator niet in de behandelkamer kan blijven gedurende de bestraling. Veranderingen van de anatomische structuren of kleine bewegingen van de patiënt gedurende de bestraling kunnen de echobeeldkwaliteit verslechteren. Het is belangrijk om deze kwaliteitsverslechtering tijdig te identificeren. In dit hoofdstuk wordt een prototype van een deep learning algoritme beschreven dat automatisch een kwaliteitsscore kan toeschrijven aan 2D echobeelden van de mannelijke pelvis. De hoogte van deze kwaliteitsscore geeft inzicht in de bruikbaarheid van het desbetreffende beeld in de echogestuurde RT procedure. De veelbelovende resultaten laten zien dat het algoritme een nauwkeurigheid kan behalen van 94%

welke vergelijkbaar is met de drie experts (nauwkeurigheidsbereik: 92% - 97%) die in deze studie geconsulteerd zijn.

Het derde deel van dit proefschrift omvat de studies waarin CT-elektronendichtheid, noodzakelijk voor dosisberekeningen, wordt verkregen zonder daadwerkelijk een CT-scan te maken. In **Hoofdstuk 6** wordt een workflow geïntroduceerd waarin elastische registratie van twee echobeelden, die gemaakt zijn gedurende de simulatiefase van de behandeling en voor een behandelfractie wordt uitgevoerd. Het deformatieveld dat op deze manier is verkregen wordt vervolgens toegepast op de simulatie CT-scan, waarbij een pseudo CT-scan wordt gecreëerd. Deze workflow is toegepast op data van drie prostaatkankerpatiënten waarvan beelden beschikbaar waren op verschillende tijdspunten. De vergelijking van de gecreëerde pseudo CT-scans en de gouden standaard CT-scan liet verbeteringen tot 11.2% in gammafalen voor dosis in de prostaat en verbeteringen tot 20.5% in gammafalen voor Hounsfield units in anus en rectum zien. Dit lijkt de hypothese, dat de pseudo CT-scan de anatomie van de patiënt gedurende de behandelfase beter representeert dan de simulatie CT-scan, te bevestigen.

In **Hoofdstuk 6** wordt gebruik gemaakt van elastische registratie om het deformatieveld tussen de twee echovolumes te creëren, waarbij er slechts gebruik wordt gemaakt van een gelimiteerd aantal parameters. Deze parameters, zoals het aantal iteraties, kunnen significante invloed hebben op de prestatie van het registratie-algoritme. Daarom is in **Hoofdstuk 7** een groter deel van de registratieparameter-ruimte verkend. Ook zijn er zes extra strategieën toegevoegd om de pseudo CT-scans mee te creëren. Hierbij werd niet alleen gebruik gemaakt van elastische registratie, maar ook van rigide registratie of een combinatie van beide. Het verkennen van de parameterruimte en het gebruik van extra strategieën is getest op beelddata van één prostaatkankerpatiënt, die beelden beschikbaar had op vijf verschillende tijdspunten. De resultaten lieten zien dat minimaal vier van de in totaal acht onderzochte strategieën tot representatievere pseudo CT-scans leiden met een gemiddelde Dice similarity coefficient verbetering van 40.1% (bereik: 4.3% - 126.5%) in vergelijking met de bijhorende simulatie CT-scans.

Als laatste worden in **Hoofdstuk 8** de resultaten die in dit proefschrift gepresenteerd zijn bediscussieerd en wordt er een visie gegeven voor zowel het toekomstig gebruik van echografie voor dosisgestuurde RT voor prostaatkankerbehandeling als voor het gebruik van echografie voor sturing gedurende andere medische procedures.

Contents

1

Introduction

1.1 Prostate cancer

Prostate cancer is the most frequently diagnosed cancer in men worldwide. The risk that a man develops prostate cancer over a lifetime is 1 in 6 [1]. In 2015, more than 1.6 million men worldwide were diagnosed with this type of cancer, which often arises in the gland cells of the prostate (adenocarcinoma) [2]. It tends to develop in elderly men and progress slowly. Therefore it is not unlikely that the cancer will never cause serious illness, as the patient may die of other causes before the cancer grows to problematic proportions. For this reason, one of the treatment options for localized prostate cancer is active surveillance. This approach involves close monitoring of the cancer by regular testing, for example, by using a Prostate-Specific Antigen (PSA) blood test. Maurice *et al.* [3] used data from the US National Cancer Base (2010-2011) to examine the use of active surveillance among 189,768 prostate cancer patients. Depending on the chosen guidelines, 10.7% - 39.8% of these patients were eligible for active surveillance. However, the majority of these patients, as well as the patients that were not eligible for active surveillance, received a different type of treatment. This could be, for example, a surgery during which the prostate together with surrounding tissues is resected from the body (radical prostatectomy). This is a highly invasive treatment option with potential side-effects as infections, damage to nearby organs and bleedings. A recent study has shown that treating prostate cancer with radiotherapy (RT) yields similar outcomes as with the radical prostatectomy, but without the occurrence of severe side-effects [4].

1.2 Radiotherapy treatment

The aim of RT treatment is to irradiate tumor tissue with ionizing radiation, while sparing the surrounding normal tissue as much as possible. The RT treatment plan is typically designed based on a computed tomography (CT) scan of the patient. During the design process, radiation dose prescriptions and delineations of the treatment target and the organs at risk are also taken into account. Subsequently, the radiation can be delivered to prostate cancer patients using two approaches.

In brachytherapy radioactive sources or seeds are implanted inside the treatment target. This allows for RT treatment from the inside out. In low-dose rate (LDR)

brachytherapy of the prostate, these radioactive sources are permanently implanted into the prostate tissue, while in high-dose rate (HDR) brachytherapy the sources are only temporarily placed into the prostate using implanted needles [5]. In external beam RT (EBRT) the treatment target is irradiated from the outside using ionizing radiation generated by a linear accelerator (Linac). The gantry of this Linac usually rotates around the patient, while collimators inside the gantry shape the radiation beam. Both intensity-modulated RT (IMRT) and volumetric-modulated arc therapy (VMAT) allow radiation delivery from several different angles, resulting in a more favorable dose distribution in comparison with the more traditional 3D conformal RT (3D-CRT) [6].

Which RT treatment option or combination of options is chosen for a patient is based on the risk classification of this patient. There are a number of guidelines available (e.g. [7]) that allow this classification of a prostate cancer patient in low, intermediate or high risk groups based on pretreatment factors including PSA level, biopsy Gleason score and clinical tumor staging. Grimm *et al.* [8] evaluated studies published from 2000 to 2011 in which biochemical (PSA) free progression of prostate cancer patients, who were treated with a variety of treatment options or a combination of these options. The findings of this evaluation suggest that brachytherapy approaches (LDR or HDR) provide a superior outcome for low-risk patients in comparison with other treatment options. For intermediate-risk patients, both brachytherapy alone or in combination with EBRT seems to be superior to the use of EBRT alone and to the use of other treatment options. Finally, for high-risk patients a combination of EBRT and brachytherapy appears superior to these localized treatment options alone.

In the remainder of this thesis, the focus will lie on EBRT treatment of prostate cancer and it will be referred to as RT treatment. In RT treatment, the radiation dose is often delivered to the patient in multiple treatment fractions, to allow normal tissue to recover in between the fractions. A daily dose of 1.8 to 2.0 Gy for 39-45 fractions [9] is used in standard RT, which implies that the patients will receive radiation for several weeks in a row.
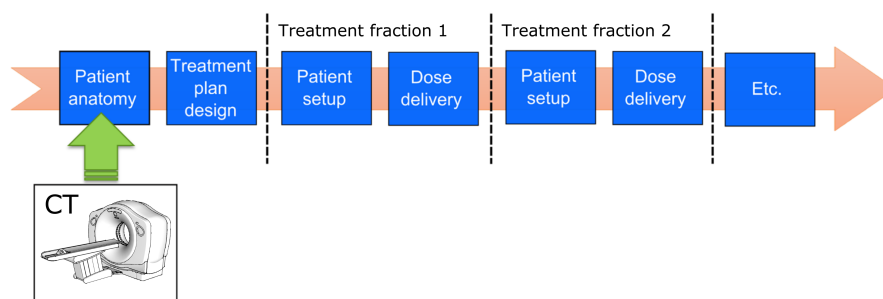
Figure 1.1: Typical RT workflow that includes CT scan acquisition for treatment plan design and delivery of the radiation dose during several treatment fractions.

A crucial aspect of a RT workflow is the patient positioning. Prior to each treatment fraction, the patient should be positioned the same way as he was during the CT scan acquisition, as this is the position that the treatment plan was prepared on. Currently, several solutions are clinically available to assist during the correct positioning of prostate cancer patients, such as skin markers on the body of the patient [10] and cone-beam CT imaging [11] with or without the use of implanted fiducial markers [12].

The available solutions might not be suited to fully identify soft tissue deformations, such as prostate shape and position changes that can occur due to a different bladder or rectal filling [13]. In addition, the solutions are primarily useful for patient positioning in between treatment fractions (inter-fraction), while also during a treatment fraction (intra-fraction) tissue distributions can change. Ballhausen *et al.* [14] have shown that the intra-fraction motion of the prostate can be considered a time-dependent random walk. According to the authors, the prostate tends to move away from the treatment isocenter and this drift increases over time. This seems to be confirmed by other studies [15], [16], including a more recent work of Ballhausen *et al.* [17]. In this work, the prostate remained within 1 mm of the setup position in 58% of the treatment time during IMRT (relatively long treatment time) and in 72% of the treatment time during VMAT (relatively short treatment time). The prostate remained within 2 mm of the setup position in 79% or 92% of the time, respectively.

If these changes would not be taken into account and the patient would be treated using the initially designed treatment plan, this could result in a suboptimal dose deposition in the target and extra undesired dose delivery to the surrounding organs at risk. According to the recent paper of Ballhausen *et al.* [17] future work is required to calculate the dosimetric consequences for the prostate and the organs at risk due to position and shape changes of these anatomical structures.

In the past, Li *et al.* [18] used the electromagnetic beacons of the Calypso system (Calypso Medical Technologies, Inc., Seattle, WA, USA) to show that the dose effects of intra-fraction prostate motion on individual treatment fractions, during which large motion occurs, are substantial. In the end, these effects become insignificant over the course of the whole treatment. However, with the beacons it is not possible to identify motion of the organs at risk and deformation of both prostate and the organs at risk. In addition, it has been suggested that using a hypo-fractionation scheme in the RT treatment could result in the same or better outcomes for the prostate cancer patients [19]. In such a scheme, a higher radiation dose per fraction is delivered in less treatment fractions. This would most probably lead to longer treatment times per fraction, which could result in more motion of the anatomical structures and so different dose consequences. This implies that, in order to understand the dosimetric consequences of the intra-fractional tissue distribution changes, another solution is desired, which

allows for soft tissue deformation identification, preferably using prolonged real-time imaging during the fraction.

1.3   Ultrasound imaging

Ultrasound (US) imaging is mostly known by the general public as the imaging modality that is used to image a fetus in the uterus of its mother. However, also for several other medical examinations of, for example, heart and liver, this image modality is a widely used diagnostic tool of choice. During US image acquisition the operator positions a probe equipped with piezoelectric elements on the body of the patient. These elements create high-frequency (1 - 18 MHz is clinically available) sound waves which are transmitted into the body. On their way through the body, the US waves encounter tissues which vary in physical properties, such as density. As acoustic impedance is defined as $Z = \rho \, c$ with ρ is the physical mass density and c is the speed of sound (SOS) in the tissue, also the acoustic impedance varies. This difference in acoustic impedance results in a partial reflection of the US waves when they encounter an interface between tissues. The echo times (or time of flight [TOF]) and amplitudes of these reflected waves are subsequently received and recorded by the US probe. Under the assumption of a fixed SOS (typically 1540 m/s [20]), it is then possible to convert the TOFs into depths. In the end, the echo amplitudes are plotted as a function of these depths, resulting in the reconstruction of the US image (see Fig. 1.2 for two examples). More in depth physics theory on US imaging can be found in literature (e.g. [21] and [22]).

(A)   (B)
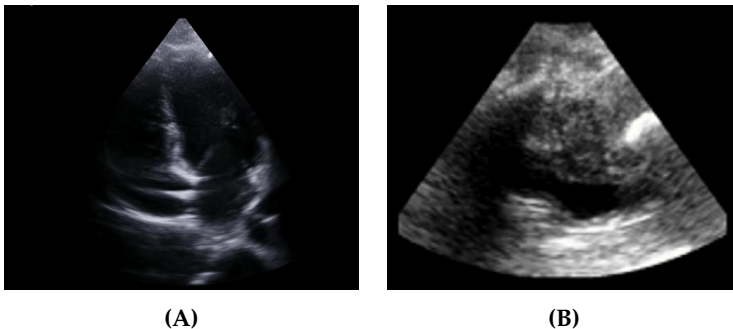
Figure 1.2: Examples of US images. (A) Apical four chamber view of the heart. (B) Transperineal US image of the male pelvic region.

US imaging has a unique combination of characteristics which makes it a potentially powerful imaging modality. It allows for real-time volumetric imaging without delivery of additional radiation dose to the patient. Therefore, frequent or prolonged image acquisition will not cause any harm to the patient. In addition, it allows for relatively fast and cheap image acquisition and most US imaging systems are positioned on an easily portable cart. Finally, US imaging

allows for acquisition of images with a good soft tissue contrast, potentially visualizing anatomical structures which cannot be identified on, for example, CT scans. However, also several challenges are associated with US imaging in general. For example, anatomical structures that are shielded by air or bone usually cannot be imaged with sufficient image quality and the imaging target needs to be within the US imaging depth range.

1.4 US guided RT and its challenges

Currently, US imaging is not solely used as a diagnostic tool anymore, but increasingly also as a tool for guidance during various medical procedures. One could think of using US guidance during, for example, catheter placement [23], but also during a RT workflow [24], [25]. Over the years, a small number of systems have been made commercially available that allowed intra- and/or inter-fraction motion monitoring of the prostate during a RT workflow using either transabdominal US (TAUS) or transperineal US (TPUS) imaging. To the best of my knowledge, the Clarity system (TAUS) and the Autoscan system (TPUS) from Elekta (Stockholm, Sweden) are currently still available on the market and are used mostly in recent publications.

In addition to the general challenges, which were described in the previous section, there are some other challenges that are less problematic during diagnostic US imaging, but become more serious when US imaging is used for guidance. For example, establishing a good acoustic coupling between the probe and the skin of the patient is crucial to allow the sound waves to penetrate into the body. US gel is often used as a medium to establish this coupling. However, US gel might be partially absorbed by the skin or spread out too much during prolonged use (which is more likely during guidance), causing a change in characteristics and so a potential decrease of coupling ability.

Also, a trained operator is needed during the US image acquisition to verify if the correct anatomical structures are visualized with sufficient image quality. This implies that either a trained operator needs to be present in addition to the clinician who performs the procedure, or the clinician has to undergo the training with a steep learning curve himself. For RT specifically, this implies that the technicians who usually are present during the dose delivery, would need to be trained how to acquire and use the US images during the course of the treatment.

Finally, US images cannot provide electron density information which is still necessary for radiation dose calculations and so for RT treatment plan adaptation. Nowadays, this density information can only be provided by a CT scan. This implies that the US images should be somehow registered to a CT scan, to be of real use during a RT workflow.

1.5 Problem statement and objective of this thesis

The use of US imaging for guidance during prostate cancer RT workflows is presently not widespread. This can be partially attributed to the need for a trained operator and the fact that electron density information from a CT scan is still necessary to understand if RT treatment plan re-planning could be beneficial. The aim of this PhD work was to optimize the US guided RT workflow of prostate cancer patients by automating steps of this workflow and to explore if US imaging can help to understand if RT treatment plan adaptation is necessary. This could make the use of US imaging more appealing and allow prostate cancer patients to fully benefit from the unique characteristics of this imaging modality.

In this thesis, three aspects of the US guided RT workflow of prostate cancer patients have been investigated:

US image acquisition
Can a prostate cancer patient-specific US probe setup be proposed prior to US image acquisition, which would allow visualization of the clinically required anatomical structures?

US image interpretation
Can a deep learning algorithm be trained to automatically assign a score to US images based on their quality or, in other words, on their usability during the US guided RT workflow?

Electron density information retrieval
Can US images be used to derive electron density information, necessary for radiation dose calculations at treatment stage, without the acquisition of an additional CT scan?

To investigate the aspect of US image acquisition, a phantom study and a patient study were performed (detailed in **Chapter 3** and **Chapter 4** of this thesis). In these studies, algorithms were developed and tested, which propose a patient-specific US probe setup based on the simulation CT scan of that specific patient. The second aspect, related to US image interpretation, is covered in **Chapter 5**. In the described study, a deep learning algorithm was implemented and trained to automatically assign a quality score to 2D TPUS images. For evaluation purposes, the scoring performance of the algorithm was then compared with the scoring performance of three experts. The final aspect, involving retrieval of electron density information, was investigated in two studies (see **Chapter 6** and **Chapter 7**). In both studies, (deformable) image registration of pairs of US volumes was used to create a pseudo-CT scan. In the end, an assessment showed if the pseudo-CT scan did represent the anatomy of the patient well. This was done to understand if the scan could be used to derive the necessary electron density information. A more detailed outline of this PhD thesis can be found in the next section.

1.6 Outline of the thesis

In this first chapter (**Chapter 1**) a brief introduction is given on prostate cancer, US imaging and RT workflows, as well as a problem statement and a thesis outline. In **Chapter 2** a more in-depth review on the use of US imaging in external beam RT workflows of prostate cancer patients is given. After the two introductory chapters, the second part of the thesis comprises the work that is related to automation of US image acquisition and interpretation.

Chapter 3 details a proof-of-concept study, which was performed on a male pelvic phantom. In this study, an algorithm is introduced that uses anatomical information derived from a CT scan to automatically propose an US probe setup to the operator. This probe setup should allow visualization of the relevant anatomical structures. For evaluation purposes, several proposed probe setups have been reproduced using a robotic arm and US volumes of the phantom have been acquired. Quantitative evaluation and visual inspection of these volumes showed that the algorithm was able to propose probe setups that fulfill the clinical requirements. In addition, the experience was gained that the currently used robotic arm was not suitable for US probe positioning on the body of a patient, due to the occurrence of unexpected arm movements.

In **Chapter 4**, the phantom work was used as a basis to create an algorithm that can propose a prostate cancer patient-specific TPUS probe setup. This algorithm also takes care of the pre-processing by identifying the perineal skin area of the patient, which is accessible for TPUS probe placement in clinical practice, using a simulation CT scan. Subsequently, it proposes several possible probe setups on this skin area and selects the optimal setup using anatomical structure delineations corresponding to the CT scan. In order to test the algorithm, patient-specific probe setups for three prostate cancer patients were proposed. These setups were then compared to the setups used by radiation oncologists during actual TPUS image acquisitions of these patients using a mechanical arm, instead of a robotic arm. This comparison revealed that the algorithm seems to be able to propose setups that could allow the visualization of all anatomical structures that are of interest for US guided RT.

Chapter 5 introduces the first step towards fully automatic interpretation of US images. To allow for US image acquisition during a treatment fraction, the US probe would need to be positioned on the body of the patient prior to the treatment commencement using e.g. a mechanical arm, as the operator cannot be present in the room during radiation dose delivery. Changes in anatomical structures or small motion of the patient during the treatment delivery can compromise US image quality, which requires identification. This chapter details a prototype deep learning algorithm that can automatically assign a quality score to 2D US images of the male pelvic region. This quality score gives information on the usability of the specific image in the US guided RT workflow. The obtained promising results

show that the algorithm can achieve an accuracy which is comparable with three experts who were consulted in this study.

In the third part of this thesis, the work related to obtaining CT electron density information, which is necessary for dose calculations, without actually acquiring a CT scan is described. **Chapter 6** introduces a new concept workflow where deformable image registration between two US volumes, acquired at simulation stage and prior to a treatment fraction, is performed. The deformation field so obtained is then applied to the simulation CT scan, resulting in the creation of a pseudo-CT scan. This workflow has been applied to image data of three prostate cancer patients who had images acquired during several time points. Comparison of the created pseudo-CT scans and available ground truth CT scans seemed to confirm the hypothesis that the pseudo-CT scan represents the patient's anatomy at treatment stage better, than the simulation CT scan does.

In **Chapter 6** the deformation field between the two US volumes was created using only deformable image registration, while exploring a limited number of parameters. These parameters, such as number of iterations can impact the performance of the registration algorithm significantly. For this reason, in **Chapter 7** a larger part of the registration parameter space has been explored. In addition, six additional approaches to create the pseudo-CT scans were used. These approaches did not only include deformable registration, but also only rigid registration or a combination of both. This parameter space exploration and application of additional approaches was tested on image data from one prostate cancer patient, who had images available at five different time points. It was shown that at least four out of eight approaches resulted in more representative pseudo-CT scans.

Finally, in **Chapter 8** the results presented in this thesis are thoroughly discussed and a vision is given for the future use of US imaging for guidance in prostate cancer RT workflows as well as its use for guidance during other medical procedures.

References

[1] Cancer-Research-UK. "Prostate cancer incidence statistics - Lifetime risk". In: *www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer* (2017).

[2] C. Fitzmaurice, C. Allen, R. Barber, L. Barregard, Z. Bhutta, H. Brenner, D. Dicker, O. Chimed-Orchir, R. Dandona, L. Dandona, et al. "Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study". In: *JAMA oncology* 3.4 (2017), pp. 524–548.

[3]   M. Maurice, R. Abouassaly, S. Kim, and H. Zhu. "Contemporary nationwide patterns of active surveillance use for prostate cancer". In: *JAMA internal medicine* 175.9 (2015), pp. 1569–1571.

[4]   F. Hamdy, J. Donovan, J. A. Lane, M. Mason, C. Metcalfe, P. Holding, M. Davis, T. Peters, E. Turner, and R. Martin. "10-year outcomes after monitoring, surgery, or radiotherapy for localized prostate cancer". In: *New England Journal of Medicine* 375.15 (2016), pp. 1415–1424.

[5]   M. Chao, P. Grimm, J. Yaxley, R. Jagavkar, M. Ng, and N. Lawrentschuk. "Brachytherapy: state-of-the-art radiotherapy in prostate cancer". In: *BJU international* 116 (2015), pp. 80–88.

[6]   D. Palma, E. Vollans, K. James, S. Nakano, V. Moiseenko, R. Shaffer, M. McKenzie, J. Morris, and K. Otto. "Volumetric modulated arc therapy for delivery of prostate radiotherapy: comparison with intensity-modulated radiotherapy and three-dimensional conformal radiotherapy". In: *International Journal of Radiation Oncology* Biology* Physics* 72.4 (2008), pp. 996–1001.

[7]   M. Zelefsky, Z. Fuks, M. Hunt, H. Lee, D. Lombardi, C. Ling, V. Reuter, E. Venkatraman, and S. Leibel. "High dose radiation delivered by intensity modulated conformal radiotherapy improves the outcome of localized prostate cancer". In: *The Journal of urology* 166.3 (2001), pp. 876–881.

[8]   P. Grimm, I. Billiet, D. Bostwick, A. Dicker, S. Frank, J. Immerzeel, M. Keyes, P. Kupelian, R. Lee, S. Machtens, et al. "Comparative analysis of prostate-specific antigen free survival outcomes for patients with low, intermediate and high risk prostate cancer treatment by radical therapy. Results from the Prostate Cancer Results Study Group". In: *BJU international* 109 (2012), pp. 22–29.

[9]   B. Vanneste, E. Van Limbergen, E. van Lin, J. van Roermund, and P. Lambin. "Prostate cancer radiation therapy: What do clinicians have to know?" In: *BioMed research international* 2016 (2016).

[10] G. Bentel. *Patient positioning and immobilization in radiation oncology*. McGraw-Hill, 1999.

[11] M. Oldham, D. Létourneau, L. Watt, G. Hugo, D. Yan, D. Lockman, L. Kim, P. Chen, A. Martinez, and J. Wong. "Cone-beam-CT guided radiation therapy: A model for on-line application". In: *Radiotherapy and oncology* 75.3 (2005).

[12] U. van der Heide, A. Kotte, H. Dehnad, P. Hofman, J. Lagenijk, and M. van Vulpen. "Analysis of fiducial marker-based position verification in the external beam radiotherapy of patients with prostate cancer". In: *Radiotherapy and Oncology* 82.1 (2007), pp. 38–45.

[13] J. Roeske, J. Forman, C. Mesina, T. He, C. Pelizzari, E. Fontenla, S. Vijayakumar, and G. Chen. "Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy". In: *International Journal of Radiation Oncology Biology Physics* 33.5 (1995), pp. 1321–1329.

[14] H. Ballhausen, M. Li, N. Hegemann, U. Ganswindt, and C. Belka. "Intrafraction motion of the prostate is a random walk". In: *Physics in medicine and biology* 60.2 (2014), p. 549.

[15] A. Richardson and P. Jacobs. "Intrafraction monitoring of prostate motion during radiotherapy using the Clarity® Autoscan Transperineal Ultrasound (TPUS) system". In: *Radiography* 23.4 (2017), pp. 310–313.

[16] M. Baker and C. Behrens. "Determining intrafractional prostate motion using four dimensional ultrasound system". In: *BMC cancer* 16.1 (2016), p. 484.

[17] H. Ballhausen, M. Li, U. Ganswindt, and C. Belka. "Shorter treatment times reduce the impact of intra-fractional motion". In: *Strahlentherapie und Onkologie* (2018), pp. 1–11.

[18] H. Li, I. Chetty, C. Enke, R. Foster, T. Willoughby, P. Kupellian, and T. Solberg. "Dosimetric consequences of intrafraction prostate motion". In: *International Journal of Radiation Oncology* Biology* Physics* 71.3 (2008), pp. 801–812.

[19] D. Dearnaley, I. Syndikus, H. Mossop, V. Khoo, A. Birtle, D. Bloomfield, J. Graham, P. Kirkbride, J. Logue, Z. Malik, et al. "Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHiP trial". In: *The Lancet Oncology* 17.8 (2016), pp. 1047–1060.

[20] P. Wells. *Biomedical ultrasonics*. Academic Pr, 1977.

[21] C. Hill, J. Bamber, and G. ter Haar. *Physical principles of medical ultrasonics*. 2004.

[22] J. Bushberg. *The essential physics of medical imaging*. Lippincott Williams & Wilkins, 2002.

[23] B. Saugel, T. Scheeren, and J. Teboul. "Ultrasound-guided central venous catheter placement: a structured review and recommendations for clinical practice". In: *Critical Care* 21.1 (2017), p. 225.

[24] D. Fontanarosa, S. Van der Meer, J. Bamber, E. Harris, T. O'Shea, and F. Verhaegen. "Review of ultrasound image guidance in external beam radiotherapy: I. Treatment planning and inter-fraction motion management". In: *Physics in Medicine & Biology* 60.3 (2015).

[25]   T. O'Shea, J. Bamber, D. Fontanarosa, S. van der Meer, F. Verhaegen, and E. Harris. "Review of ultrasound image guidance in external beam radiotherapy part II: intra-fraction motion management and novel applications". In: *Physics in Medicine & Biology* 61.8 (2016).

Chapter 2

Ultrasound in prostate cancer external beam radiotherapy

Abstract

External beam radiotherapy (EBRT) is one of the curative treatment options for prostate cancer patients. The aim of this treatment option is to irradiate tumor tissue, while sparing normal tissue as much as possible. Frequent imaging during the course of the treatment (image guided radiotherapy) allows for determination of the location and shape of the prostate (target) and of the organs at risk. This information is used to increase accuracy in radiation dose delivery resulting in better tumor control and lower toxicity.

Ultrasound imaging is harmless for the patient, it is cost-effective and it allows for real-time volumetric organ tracking. For these reasons, it is an ideal technique for image guidance during EBRT workflows. Review papers have been published in which the use of ultrasound imaging in EBRT workflows for different cancer sites (prostate, breast, etc.) was extensively covered. This new review paper aims at providing the readers with an update on the current status for prostate cancer ultrasound guided EBRT treatments.

2.1 Introduction

Prostate cancer is the most frequently diagnosed cancer in men worldwide. It accounted for 1.6 million new diagnoses and 366,000 deaths in 2015 [1]. In the next decades, the incidence of prostate cancer might increase due to the possible linkage of this cancer with risk factors associated with economic development (e.g. excess body weight and physical inactivity) [2] and the aging population [3].

One of the curative treatment modalities for prostate cancer is external beam radiotherapy (EBRT) [3]. The aim of this modality is to irradiate tumor tissue using ionizing radiation generated by an X-ray source (e.g. linear accelerator). At the same time, normal tissue must be spared as much as possible to avoid excessive toxicity. EBRT is one of the most common forms of RT treatment and therefore it is often denoted as just radiotherapy (RT) in literature (as will be done in the remainder of this chapter).

Prior research using kV radiography has shown [4], [5] that frequent imaging of the patient's anatomical structures of interest during the course of the prostate RT treatment (image guided RT, IGRT) can improve radiation targeting and tumor control. This improved targeting could allow reduction of safety margins, with consequently decreased toxicity. Next to kV radiographs also other imaging modalities have been used for IGRT, such as cone beam CT (CBCT) in combination with fiducial markers [6], magnetic resonance imaging (MRI) [7], implantation of electromagnetic transponders [8] and ultrasound (US) imaging [9].

In this review paper the focus solely lies on the use of US imaging during the IGRT workflow of prostate cancer patients. US imaging typically provides good soft tissue contrast and therefore it is a modality that allows contouring of structures such as the prostate [10]. It is also a real-time image modality, because the images are reconstructed and visualized directly during the acquisition. Some of the currently available US systems potentially even allow real-time volumetric imaging and soft-tissue tracking, using a matrix probe (e.g. X6-1 xMatrix array probe, center frequency: 3.2 MHz, Philips Healthcare, Bothell, WA, United States), or a mechanically swept probe (e.g. Clarity Autoscan probe, m4DC7-3/40, center frequency: 5 MHz, Sonix Series; Ultrasonix Medical Corporation, Richmond, BC, Canada).

Some of the limitations and challenges associated with US imaging include the inaccessibility of tissue shielded by bone or air, the proneness for imaging artifacts and the user dependency [11], due to its mostly manual operation. However, in comparison with other imaging modalities US is cost-effective and it does not deliver ionizing radiation to the patient. The combination of these characteristics with the real-time volumetric tracking ability makes US imaging a suitable image modality for inter- and intra-fraction organ motion monitoring during the course of a prostate RT treatment [12]. US imaging could then be used either as standalone system or possibly in combination with other imaging modalities.

In 2015 and 2016 two review articles [13], [14] were published in which the use of US for IGRT of different cancer sites (e.g. prostate, breast and liver) was extensively covered. The current review article updates this work for prostate cancer. After an introductory summary on US techniques and US systems that can potentially be used during the RT prostate cancer patient workflow, a comprehensive update on the latest developments in this field is presented.

2.2 EBRT workflow and US imaging

2.2.1 EBRT workflow

The typical RT workflow of prostate cancer patients consists of several steps, either belonging to the simulation stage (preparatory phase) or the treatment stage (radiation dose delivery phase) (Fig. 2.1). The first step involves the invasive implantation of fiducial markers in the prostate gland. These markers are considered a surrogate for the target and are currently used to monitor its motion between different treatment fractions using X-ray imaging.
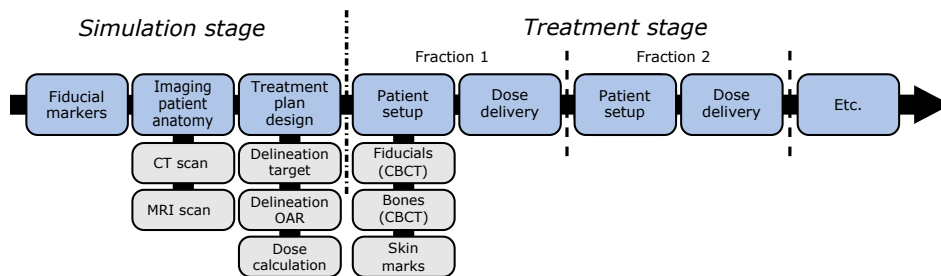
Figure 2.1: Typical RT workflow for prostate cancer patients. During the simulation stage, fiducial markers are implanted in the prostate, images of the patient's anatomy are acquired and a treatment plan is designed. Subsequently, the dose is delivered to the patient in several treatment fractions, while ensuring that the patient is setup as accurately as possible.

Subsequently, a computed tomography (CT) scan, and increasingly more often an MRI scan, are acquired. The CT scan provides electron-density information allowing for treatment plan preparation, based on prescribed radiation dose and delineations of the anatomical structures of interest (target and organs at risk [OARs]). In case also an MRI is acquired, it is registered with the CT scan based on the fiducial markers [15], which can be visualized with both imaging modalities. Then, the prostate (target) is delineated on the MRI instead of on the CT scan. As the volumes are registered, the delineation can be transferred to the CT scan and used during the treatment plan preparation. MRI-based delineation is preferred as MRI usually allows for a more accurate delineation of the prostate than the CT [16]–[18].

After finalizing the treatment plan design, the radiation dose will be delivered to the patient in multiple daily treatment fractions (up to 45) during 1-2 months [3].

The setup of the patient prior to each of these treatment fractions is an important step in the RT workflow. This procedure must be as accurate as possible to reproduce the setup at simulation stage, on which the treatment plan was designed.

Nowadays, setting up the patient is typically assisted by the use of skin marks on the patient's body [19], the previously mentioned fiducial markers [20] and CBCT [21]. However, even if the patient seems to be correctly aligned, internal soft-tissue deformations may still occur. The position and shape of the prostate can change, due to a different filling of the bladder and rectum [22]. To account for these deviations from the simulation CT, a safety margin is usually added to the treatment target [23]. Unfortunately, this leads to a larger volume being irradiated, potentially including larger portions of OARs.

Monitoring the position and shape of the prostate during the course of the RT treatment could potentially improve the accuracy of the radiation dose delivery and, in the end, potentially even allow for a margin reduction. In the ideal case, this prostate monitoring would not only include monitoring between different fractions (inter-fraction), but also during a treatment fraction (intra-fraction) [14]. As noted before, US imaging could be a suitable imaging modality for this purpose.

2.2.2   US imaging in RT workflow

US imaging makes use of a probe equipped with piezoelectric elements to create high-frequency sound waves and transmit these into the body. On their way through the body, these waves encounter interfaces between different tissues and scattering objects. Due to the differences in acoustic impedance between the tissues at each side of this interface, and between the scattering objects and the surrounding tissue, a part of the US waves is reflected, while the remaining waves keep penetrating deeper into the body. The reflected waves are received by the probe, processed and combined to generate an image.

As air reflects US waves very strongly, the presence of air between the probe and the body of the patient will prevent sufficient penetration of the waves into the body, which significantly degrades the image quality. It is therefore crucial to establish sufficient acoustic coupling between the probe and the body. For this purpose, a coupling medium, such as US gel or water is typically used.

Several US probes with different shapes and characteristics are commercially available for the different procedures possible with this technology. To image the prostate and OARs during the RT workflow, three US imaging techniques are presently used in clinical practice. These techniques and how they can potentially improve the accuracy of radiation dose delivery are described in the next sections. We refer to the literature (e.g. [24], [25]) for more general details on the physics theory and technology of US imaging.

A. Transrectal US imaging

Transrectal US (TRUS) imaging requires positioning of the probe through the anus inside the rectum (Fig. 2.2A) and is therefore a low invasive imaging procedure. As the prostate is located in close proximity of the rectum, TRUS allows imaging of the prostate with a good image quality (Fig. 2.3A) [26]. Challenges that can occur while making use of TRUS imaging are rectal filling, which can be removed using an enema [27] and the potential presence of air in the rectum, which results in a poor acoustic coupling between the probe and the body of the patient.

In the EBRT workflow, TRUS imaging is currently used to guide the fiducial marker placement during the simulation stage (Fig. 2.4) [28]. The invasive character of this US modality makes it less suitable for frequent imaging during the course of the treatment. In addition, the presence of the probe inside the rectum being potentially in the path of the radiation treatment beam (Fig. 2.2A), raises issues as well. For this reason, no research seems to have been conducted on the use of TRUS for inter- and intra-fraction organ motion monitoring during prostate EBRT.
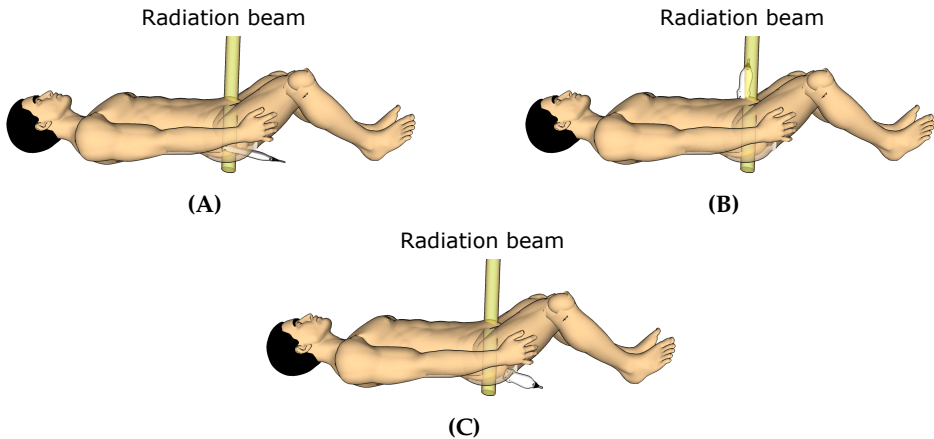
Figure 2.2: US probe setup using three US imaging techniques. (A) TRUS, (B) TAUS and (C) TPUS with the yellow beam indicating a possible location of a radiation beam during a treatment fraction.

B. Transabdominal US imaging

Transabdominal US (TAUS) imaging involves the positioning of the US probe on the abdomen (Fig. 2.2B) and it is therefore a non-invasive imaging modality. It is capable of measuring the same prostate volumes as TRUS imaging (considered the standard) [29] and it makes use of the acoustic window of the bladder for prostate visualization (Fig. 2.3B). For this reason, TAUS requires a reasonably full bladder, which can lead to discomfort for the patients. However, a filled bladder is

often requested during the RT treatment to prevent the whole bladder wall from being irradiated and to push the intestines away from the high dose regions.

During TAUS imaging, the probe is located relatively far from the prostate, which might influence the quality of the acquired images. Especially the acquisition of TAUS images of obese patients is a challenge [30]. Adipose tissue attenuates the US waves and increases the possibility for imaging artifacts, which can significantly degrade the image quality. Unfortunately, it is a challenge to predict the degree of adipose attenuation and the associated image quality degradation, due to the dependence on patient-specific characteristics, such as fat distribution [31].

The probe setup on the body of the patient during TAUS imaging, makes this imaging modality suitable for inter-fraction monitoring. However, it is more challenging to use TAUS imaging for intra-fraction monitoring (Fig. 2.4), as the probe is potentially located in the path of the radiation beam, especially for rotational therapy (Fig. 2.2B). Ways to overcome this challenge are currently not available in clinical practice, although they are being investigated. In Section 2.6 of this paper, the recent developments in this field will be discussed.

(A) (B) (C)
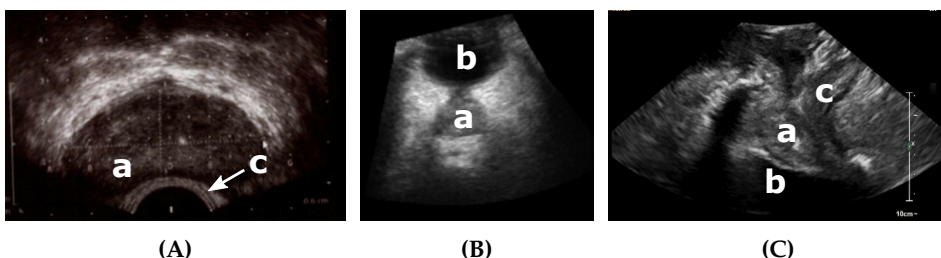
Figure 2.3: Three US techniques suitable for prostate and OARs imaging (A) TRUS, (B) TAUS and (C) TPUS with (a) prostate and with (b) bladder and (c) rectum which can partially be seen.

In the past 20 years, three systems were commercially available that allowed inter-fraction monitoring of the prostate during the RT workflow by means of TAUS imaging: SonArray system (Varian Medical Systems, Palo Alto, CA, USA), B-mode Acquisition and Targeting (BAT) system (Best Nomos, Pittsburgh, PA, USA) and the Clarity system (Elekta, Stockholm, Sweden, formerly called Restitu and commercialized by Resonant Medical, Montreal, QC, Canada). To our knowledge, only the Clarity system is still available on the market and as there have been papers published on this system in the last years, it will be covered in this chapter.

The BAT system was only used in one study [32] since the publication of the previously mentioned review papers [13], [14]. In this study a comparison was made between the Clarity system and the BAT system, resulting in a good agreement between both. As the BAT system was extensively covered in the previous review papers, it will not be discussed further in this work.

In the RT workflow, a freehand sweep using a 2D TAUS probe (C5-2/60, center frequency: 3.5 MHz, Sonix Series; Ultrasonix Medical Corporation, Richmond, BC, Canada) can be acquired by the Clarity system during the simulation stage. Due to the use of a probe localization system, it is possible to reconstruct the sweeps such that a 3D TAUS volume is created. The same procedure is repeated prior to each treatment fraction. The requirement for manual sweep acquisition makes the Clarity system inherently sensitive to uncertainties associated with operator variability and probe pressure. These issues will be covered in more detail in Section 2.5.

Comparison of the US volumes acquired at treatment stage and the reference US volume acquired at simulation stage, allows the calculation and correction of inter-fractional prostate motion [12]. Besides the fact that the US probe is potentially located in the path of the radiation beam, the need of an operator performing the manual sweep for the 3D TAUS volume reconstruction, makes this system not suitable for intra-fraction monitoring.
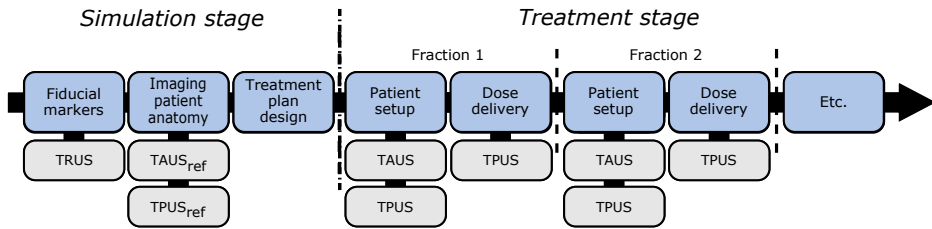
Figure 2.4: RT workflow of prostate cancer patients with US imaging implemented at several steps. The fiducial marker implantation is currently performed under TRUS guidance. The acquisition of the reference TAUS or TPUS images at simulation stage and also the acquisition of TAUS and TPUS prior to dose delivery, can provide valuable information for inter-fraction prostate motion correction. Finally, during dose delivery TPUS imaging could provide information on intra-fraction prostate motion.

C. Transperineal US imaging

Transperineal US (TPUS) imaging is a non-invasive imaging modality, as it involves the positioning of the US probe on the perineum of the patient (Fig. 2.2C). Also this imaging modality is capable of measuring the same prostate volumes as TRUS imaging [33]. TPUS imaging does not exploit the acoustic window of the bladder to obtain images of the prostate (Fig. 2.3C) and therefore it requires a less strict bladder filling protocol. A semi-filled bladder is still beneficial since it yields good imaging contrast distal to the prostate. In addition, as the distance between the prostate and the perineum is smaller, a relatively good image quality can potentially be achieved. However, just like with TAUS imaging, the body composition of the patient can affect the image quality. Finally, due to the fact that the probe setup does not interfere with the radiation beam

(Fig. 2.2C), TPUS imaging can potentially be used also for intra-fraction monitoring of the prostate (Fig. 2.4).

Currently there is only one commercial system available that enables the inter- and intra-fraction prostate motion monitoring during the RT workflow using TPUS imaging: Clarity Autoscan (Elekta, Stockholm, Sweden) [12]. This system is an extension of the Clarity system as described above. Like the Clarity system it employs a 2D probe (m4DC7-3/40, center frequency: 5 MHz, Sonix Series; Ultrasonix Medical Corporation, Richmond, BC, Canada). However, the Autoscan probe is mounted in a housing which also comprises a motorized control of the sweeping motion. This automation of the sweeping motion makes a manual sweep superfluous.

The Autoscan probe which can be localized in the room by a probe tracking system, is attached to a baseplate on the CT or on the linear accelerator (Linac) couch during the procedure (Fig. 2.5), allowing positioning and locking of the probe for TPUS imaging. The use of the baseplate and the automatically performed sweeping motion potentially reduce the operator dependence. The operator dependence will be covered in more detail in Section 2.5.
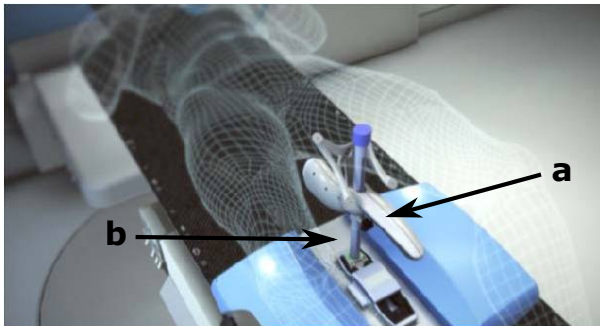
Figure 2.5: Clarity Autoscan system setup with (a) probe and (b) baseplate.

The Clarity Autoscan system follows the TPUS workflow, as represented in Fig. 2.4. First, a 3D TPUS volume is acquired at simulation stage. Then, prior to the dose delivery, a full sweep is acquired and reconstructed. Comparison of this full sweep with the image acquired at simulation allows the calculation of a required couch shift to account for inter-fraction prostate motion.

During the radiation dose delivery, continuous volumetric imaging using the US probe is performed. This allows position monitoring of the prostate in 3D. The therapist can interrupt the treatment and perform a couch correction, in case the motion in a certain Cartesian direction is exceeded for a certain amount of time. These motion direction and time thresholds can be set by the operator prior to the first treatment delivery [12].

2.3 Inter-fraction monitoring

2.3.1 Fiducial markers

As already introduced in Section 2.2, currently 3-4 fiducial markers are implanted inside the prostate prior to the start of the treatment. The most frequently used markers are made of gold and provide a surrogate for the prostate position. The markers are visible using kV imaging modalities (such as CBCT or 2D X-ray radiographs), but can also cause metal-induced image artifacts [20].

The implantation procedure is often performed under TRUS guidance and involves invasively positioning the markers in the prostate through the perineum or the rectum [34]. The procedure can be considered as well tolerated by the majority of patients [28], [35], but it is definitely not without risks. One study [36] even suggests that the risk associated with the implantation of the markers through the rectum is still underestimated. An overall rate of symptomatic infection with the fiducial marker implantation was reported to be 7.7% with one third requiring hospital admission.

The use of fiducial markers during the RT workflow is based on the assumption that the marker position inside the prostate will not change during the whole course of the treatment, from the simulation stage until the final treatment fraction. Changes in anatomy and physiology, however, can potentially cause or mimic marker migration [37]. Moreover, studies have shown that the presence of fiducial markers in the prostate can affect the dose deposition [38] and that imaging the fiducial markers using CBCT adds a non-negligible dose to the patient [39].

Therefore, inter-fraction motion monitoring should be ideally performed with a non-invasive image modality that does not require the presence of these fiducial markers inside the prostate. In this regard, US imaging is an excellent candidate. In the next section studies are discussed which used TAUS or TPUS imaging for inter-fraction motion monitoring of the prostate.

2.3.2 TAUS and TPUS imaging

In Table 2.1 the studies are reported that compared the use of TAUS (Clarity system) or TPUS (Clarity Autoscan system) with other imaging modalities for inter-fraction prostate monitoring. As the work of Tas *et al.* [40] only includes data from one prostate cancer patient, it is excluded from this table. The studies indicated with an asterisk (*) were included in the previously mentioned review papers [13], [14]. However, they have been added to this work to provide a complete overview.

The older studies primarily focused on TAUS imaging. In these studies, 2D techniques [11], [41], [42] and volumetric imaging techniques [43]–[46] were used for comparison with the TAUS imaging. One study [47] also compared the results of a surface imaging system (AlignRT, VisionRT, London, UK) with TAUS imaging. The four most recent studies [48]–[51] examined the use of TPUS

imaging in comparison with volumetric imaging only, such as CBCT and an additionally acquired planning CT.

All studies (TAUS and TPUS) reported the differences (using mean $\pm$ standard deviation (SD) or error notation including mean, systematic and random error [23]) between the US imaging technique and another image modality. The reported mean differences for the anterior-posterior (AP), left-right (LR) and superior-inferior (SI) directions were in 9 out of 11 studies in the absolute range of 0-3 mm. Some studies also reported the Bland-Altman 95% limits of agreement (LoA) [52] and/or the ranges of the measured differences. For the studies that did not report the LoA, the ranges are detailed in the final column of Table 2.1. The largest range difference was reported by Robinson *et al.* [44], ranging between 1.3 mm and 61.4 mm.

The Bland-Altman LoAs are detailed in the final column of Table 2.1 and were reported by 5 out of 12 studies. The LoA (bias $\pm$ 1.96·SD) are a measure for the inter-changeability of two methods or systems. If the limits are smaller than or equal to an a-priori defined tolerance, one method can be used inter-changeably with the other. The TPUS studies (min LoA: 3.2 mm; max LoA: 9.4 mm) tend to report slightly lower LoA values than the TAUS studies (min LoA: 5.3 mm; max LoA: 11.7 mm). Considering that the prostate safety margins currently used in clinical practice (using fiducial markers) range from 3 to 10 mm [53], neither TAUS nor TPUS could be considered inter-changeable with the imaging techniques they have been compared with. However, this does not automatically imply that the US techniques perform worse than the comparison technique, simply because there is no recognized ground truth. Therefore, potential inaccuracies in the imaging modality that the US is compared with can influence the results and associated conclusions.

The absence of ground truth is also reflected in conflicting conclusions regarding the potential performance of US imaging in the RT workflow. For example, Li *et al.* [45] concluded that it is feasible to use TAUS imaging for image guidance during the prostate RT workflow and that this image modality appears comparable to CBCT when used for the same purpose. On the other hand, Fargier-Voiron *et al.* [46] concluded that TAUS imaging cannot replace CBCT without increasing treatment margins. These conclusions seem to differ significantly, while the reported mean differences between the reference imaging modality and TAUS imaging are comparable.

In general, it seems that the studies investigating the use of TPUS imaging are more optimistic about the accuracy, inter-changeability and usability in comparison to the TAUS imaging studies. For example, Trivedi *et al.* [50], conclude that TPUS imaging provides excellent imaging of the prostate and comparable localization results. Also Li *et al.* [51] conclude that TPUS is a feasible image modality for IGRT and has a good accuracy.

Table 2.1: Studies comparing TAUS or TPUS with other imaging techniques used for IGRT. The first column details the first author and year of publication. The * indicates that the paper was also included in [13], [14]. The second column details the used US technique: TAUS (Clarity) or TPUS (Clarity Autoscan). The third column indicates the parameters varied in the study. The fourth column indicates the image modality used for comparison with FM (fiducial marker) and EPI (electronic portal imaging). The columns five and six detail the number of patients and the number of acquired scans, respectively. Most studies reported the difference using mean ± standard deviation (SD) (column seven), however, one study used the error notation [42]. The square brackets indicate absolute values. The final column details the Bland-Altman Limits of Agreement (LoA) or, if this was not reported, the range of measured differences denoted between [ ].

First author	US	Parameters	Compared	# pts	# US scans	Difference: mean ± SD [mm] AP	LR	SI	Bland-Altman / Range [mm] AP	LR	SI
Cury* (2006) [43]	TAUS	-	CT	10	30	-0.2 ± 1.6	0.2 ± 1.7	0.1 ± 1.4	-	-	-
Johnston* (2008) [41]	TAUS	Assisted segmentation	FM (EPI)	8	255	1.3 ± 6.6	0.9 ± 4.0	1.3 ± 5.1	-11.7 - 14.3	-7 - 8.8	-8.8 - 11.4
		Manual segmentation	FM (EPI)	7	181	2.1 ± 4.6	0.8 ± 3.5	0.2 ± 4.5	-6.8 - 11.1	-6.0 - 7.7	-8.7 - 9.0
Robinson* (2012) [44]	TAUS	1. All datasets	CT	17	CT: 136 US: 272		10.3 ± 7.9			[1.3 - 61.4]	
		2. Remove insufficient data	CT	-	US: 210		8.7 ± 4.9			[1.0 - 40.0]	
		3. Review (2) manufacturer	CT	-	US: 153		7.4 ± 3.1			[1.8 - 17.1]	
Van der Meer* (2013) [11]	TAUS	-	FM (EPI)	8	244	-2.3 ± 3.6	2.5 ± 4.0	0.6 ± 4.9	-9.3 - -4.7	-5.4 - 10.3	-8.9 - 10.2
Mayyas* (2013) [42]	TAUS	-	Bone (EPI)	27	1100	μ: -0.7 Σ: 2.4 σ: 3.4	μ: -0.5 Σ: 1.6 σ: 3.2	μ: -1.0 Σ: 2.4 σ: 3.6	μ = mean error Σ = systematic error [23] σ = random error		
Li (2015) [45]	TAUS	-	FM (CBCT)	6	78	0.0 ± 3.0	-0.2 ± 2.7	-1.9 ± 2.3	[-7.3 - 7.2]	[-5.6 - 6.9]	[-10.0 - -2.9]
Fargier-Voiron (2015) [46]	TAUS	Raw database	CBCT	25	284	2.8 ± 4.1	0.5 ± 3.3	-0.9 ± 4.2	-5.3 - 10.9	-5.9 - 6.9	-9.0 - 7.3
		Data corrected by mean US-CBCT difference of first 3 fractions	CBCT	25	284	-0.5 ± 3.9	0.3 ± 3.0	-1.0 ± 4.2	-8.1 - 7.1	-5.6 - 6.2	-7.7 - 7.3
Krengli (2016) [47]	TAUS	-	3D surface image	40	1318	-1.2 ± 4.9	-0.7 ± 5.0	-2.6 ± 6.4	[-25.8 - 18.0]	[-48.8 - 15.9]	[-22.5 - 22.1]
Richter (2016) [48]	TPUS	-	CBCT	10	150	[3.0 ± 2.4]	[2.7 ± 2.3]	[3.2 ± 2.7]	-7.1 - 8.2	-8 - 5	-9.4 - 6.5
Fargier-Voiron (2016) [49]	TPUS	With intra-fraction motion	CBCT	12	357	1.9 ± 3.2	0.2 ± 2.6	0.7 ± 2.6	-4.3 - 8.1	-4.4 - 5.8	4.9 - 5.3
		No intra-fraction motion	CBCT	12	357	2.8 ± 3.0	-0.1 ± 2.5	-0.3 ± 2.5	-3.2 - 8.8	-5 - 4.7	-5.1 - 4.6
Trivedi (2016) [50]	TPUS	-	FM (CT)	17	30	-0.06 ± 2.86	0.63 ± 3.27	-0.49 ± 3.49	[-4.55 - 7.52]	[-5.96 - 7.06]	[-6.70 - 6.78]
Li (2017) [51]	TPUS	-	FM (CBCT)	7	177	0.3 ± 1.7	0.0 ± 1.7	0.2 ± 2.0	[-4.2 - 5.5]	[-4.8 - 3.9]	[-4.5 - 5.7]

In conclusion, different opinions exist in the literature regarding the comparability between US (TAUS and TPUS) and other imaging modalities used for image guidance during the RT workflow. For this reason, more research is necessary before final conclusions can be drawn about the usability of US imaging in the prostate IGRT workflow. Also, it is very important that US imaging is standardized to reduce the operator dependency (see Section 2.5).

2.4 Intra-fraction monitoring

As discussed in the introduction section, the position and shape of the prostate can change, due to e.g. different bladder or rectum fillings. This phenomenon can occur not only between treatment fractions, but potentially also during a treatment slot. Intra-fractional prostate motion has been investigated in several studies using, e.g. the Calypso localization system (Calypso Medical Technologies, Inc., Seattle, WA, USA) (e.g. [8] and [54]). This system is based on the electromagnetic detection of beacon transponders which need to be implanted in the prostate. Calypso provides continuous, real-time localization of the prostate surrogates and it has been shown to have a sub-millimeter accuracy in a phantom [55].

These transponders need to be implanted in the prostate and, in addition, can cause image artifacts on MRI that could be used for treatment response assessment. In addition, an antenna which is necessary for the localization of the beacons is present in the path of the radiation beam. Finally, assumptions are needed to determine a relation between the position of the transponders and the shape and location of the prostate, making the Calypso system not a real volumetric tracking system.

As the Clarity Autoscan system (TPUS) does not involve implantation of transponders in the prostate, it allows for real volumetric tracking of the prostate. In addition, during the procedure no equipment is present in the beam path, which potentially makes it a more favorable solution for intra-fraction prostate motion tracking in comparison to the Calypso system. Abramowitz *et al.* [56] found a good agreement between the Clarity Autoscan system and the Calypso system, while examining the ability of both systems to track a prostate-like sphere in a phantom.

The accuracy and precision of the Clarity Autoscan system have been evaluated in a study using a male pelvic phantom [57]. In this study, a latency of 223 ± 45.2 milliseconds was reported between the motion of the phantom and the US tracking. In addition, a mean position error of 0.23 mm (LR) and 0.45 mm (SI) was reported. These positional and timing accuracies were found to be acceptable under the simulated treatment conditions examining, among others, the performance of the system while the radiation beam was on and while the image quality was degraded by the introduction of an air gap between the probe and the surface of the phantom. This was done to assess tracking performance under worse image quality conditions.

In the literature, three papers [58]–[60] and one abstract [61] are available in which intra-fraction prostate monitoring was clinically investigated using the Clarity Autoscan system. The authors of these publications reported different metrics. For example, Richardson *et al.* [60] reported the total frequency of intrafraction prostate displacements per direction for different thresholds, while Baker *et al.* [59] reported the percentage of fractions with displacements larger than 2 mm. These differences make it difficult to compare the results directly.

Ballhausen *et al.* [58] investigated data from 6 prostate cancer patients. This data was used to verify their hypothesis that the intra-fraction motion of the prostate can be modeled as a time-dependent 'random walk' [62]. It was shown that the prostate tends to move away from the treatment iso-center during a fraction and that this drift away from the iso-center increases over time. These findings imply that a shorter dose delivery time could be favorable. Such a reduction of the treatment time can be achieved by using e.g. volumetric modulated arc therapy (VMAT) or RapidArc Radiotherapy Technology (see [3] for more details on radiation techniques).

Table 2.2: Studies reporting on the use of TPUS imaging with the Clarity Autoscan system for intra-fraction prostate motion monitoring. The first column details the first author and publication year. The second column details the used system, while the third and fourth column indicate the number of patients and scans examined, respectively. The fifth column contains the examined time intervals in seconds, while the final column details some results and conclusions.

First author	System	# pts	# US scans	Time [sec]	Results and conclusions
Baker (2016) [59]	TPUS	10	51	120-150	– Largest displacement (2.8 mm) in posterior direction – Displacement insignificant during treatment time – Displacement increases over time
Richardson (2017) [60]	TPUS	20	526	385	– Posterior motion seems most common – 35% of patients displacement > 10 mm – Duration of displacement varies considerably between patients
Guillet (2017) [61]	TPUS	10	330	140 (+ 120 setup) 290 (+ 120 setup)	– Largest movement in AP direction – Dosimetric impact increases with treatment time duration

Baker *et al.* [59] investigated the prostate intra-fraction motion during a time interval corresponding to a beam-on time for RapidArc (120-150 seconds) (see Table 2.2). A tolerance of 2 mm was considered, as this value is perceived to be clinically irrelevant according to the British Ionization Radiation Medical Exposure Regulations 2000 (IRMER 2000). In the study, maximal intra-fractional displacements of -0.2 ± 1.1 mm (AP), -0.2 ± 0.8 mm (LR) and +0.2 ± 0.9 mm (SI) were found. The largest displacement of 2.8 mm was measured in the posterior direction.

Also, displacements larger than 2 mm were measured for 10% (AP), 2% (LR) and 4% (SI) of the examined fractions. The authors concluded that the displacement of the prostate is insignificant during the measured time interval. However, the conclusion was also drawn that the displacement increases over time, which is line with the findings of Ballhausen *et al.* [58].

Richardson *et al.* [60] instead used the Clarity Autoscan system to assess the intra-fraction prostate motion during intensity-modulated radiotherapy (IMRT) with static beams from different angles, which consequently has a longer treatment time (reported mean of 385 seconds). In this case, the authors considered three different thresholds; 3 mm (fine tolerance), 7 mm (future planning target volume) and 10 mm (current planning target volume). In addition to a technical overview, also the first clinical experiences of the physicians were captured in a letter [63] and article [64].

Also in this study, the motion of the prostate in the posterior direction seems to be the most common (Table 2.2). All patients experienced at least one displacement larger than 3 mm and 35% of the patients experienced one displacement larger than 10 mm. These higher rates of motion in comparison with [59] can potentially be explained by the fact that the evaluated time interval was much longer (385 seconds versus 120-150 seconds). In the study of Richardson *et al.* [60] also the duration of the intra-fraction prostate displacement was calculated as a proportion of the total treatment time. This duration varied considerably between patients. For example, for motion larger than 3 mm in the posterior direction, durations from 2% of the treatment time up to 92% of the treatment time were observed for individual patients.

Finally, also one abstract was published by Guillet *et al.* [61] in which the dosimetric impact of the intra-fraction motion was investigated and in which also some prostate movement results were reported. Also in this work, the largest movements were reported in the AP direction (Table 2.2), with 18% of the short treatment sessions (140 seconds) and 31% of the longer treatment sessions (290 seconds) displaying motions larger than 3 mm. In addition, in this work it was also shown that the dosimetric impact of the intra-fractional motion increases with the treatment time duration.

2.5 Operator dependece

Currently, the operator who acquires the US images in the clinic (not only in the RT environment) may need to (manually) place the US probe on the body of the patient, interpret the live images and then decide if the correct anatomical structures are visualized with sufficient image quality. This makes US imaging operator-dependent and this dependence may cause significant variability in the quality of the acquired US images and thus influence the ability to locate and track the prostate and OARs. Section 2.5.1 discusses the studies that investigated prostate displacement induced by probe pressure in both TAUS and TPUS. Inter- and intra-operator variability is detailed in Section 2.5.2.

2.5.1 Probe pressure effects

As introduced previously, the Clarity system requires the acquisition of a manual sweep along the abdomen of the patient using the TAUS probe prior to dose delivery. The acquired image can then be used for inter-fraction motion correction. Subsequently, the probe is removed from the body of the patient and the patient is irradiated. In case the prostate is displaced due to probe pressure, it might move to a different position when the probe is removed from the body. This displacement after the probe removal is not accounted for in the inter-fraction motion correction, which can lead to a suboptimal radiation dose delivery.

Table 2.3 details studies that investigated prostate displacement due to probe pressure. Two out of three TAUS studies used a relative method to assess the prostate displacement. For example, Van der Meer *et al.* [11] acquired images at no pressure (reference situation - probe touching the skin) and subsequently acquired images at low pressure, intermediate pressure and high pressure. To determine the displacement due to probe pressure the location of the prostate was compared to the reference situation.

Baker *et al.* [65] assessed the effect of TAUS probe positioning using TPUS imaging. In this work, a reference image was acquired using just a TPUS probe without the TAUS probe actually being in place on the body of the patient. The average displacement vector of the prostate found by Baker *et al.* [65] was significantly lower than the displacement found in the other studies (1.3 mm versus 2.5 mm and 3.0 mm). The studies concluded that even though the prostate displacements are small, a minimal pressure should be used in order to make the probe setup more reproducible.

The effect of probe pressure during TPUS imaging were reported in two studies. Mantel *et al.* [66] investigated the shift of the penile bulb after positioning the TPUS probe against the perineum. A superior shift of the penile bulb could bring it closer to the prostate and therefore closer to the high dose region. This could lead to an increase of dose delivered to the penile bulb, which has been correlated earlier (e.g. [67]) with the incidence of erectile dysfunction. The authors studied datasets from 10 patients and reported that the penile bulb had a significant median shift

of 6.2 mm in the superior direction. In addition, no relevant volume changes of the prostate and planning target volume due to probe pressure were observed and just minor motion of these structures were reported, mainly in the superior direction. No quantitative results on this prostate and planning target volume motion were reported in the paper.

Table 2.3: Studies reporting on prostate displacement induced by probe pressure. The first column details the first author and publication year. The $\triangle$ indicates that the specific study was mentioned in the previous review paper [13], but these specific results were not discussed. The second column details the used system, while the third column provides the imaging modality with which the prostate displacement was assessed. The fourth and fifth column specify the number of examined patients and the assessed scans, respectively. The prostate displacement in all directions are listed in column 6 with the [a] indicating results per 1-mm probe shift and in the final column the displacement vector can be found.

First author	System	Assessed with	# pts	# US scans	Prosate displacement mean $\pm$ SD [mm]			Displacement mean $\pm$ SD [mm]
					AP	LR	SI	
Van der Meer$^\triangle$ (2013) [11]	TAUS	Relative TAUS	13	376	0.7	-0.5	0.0	3.0
Fargier-Voiron$^\triangle$ (2014) [68]	TAUS	Relative TAUS	8	24	-	-	-	2.5 ± 1.2
Baker (2015) [65]	TAUS	TPUS	9	42	-0.1 ± 1.0	0.2 ± 0.7	-0.1 ± 0.8	1.3 ± 0.7
Li (2017) [69]	TPUS	Relative TPUS	10	16 series	0.07 ± 0.11[a]	0.04 ± 0.11[a]	0.42 ± 0.09[a]	2 - 4

In another study [69] the pressure applied by a TPUS probe was found to have a quantitatively similar impact on prostate displacement as the TAUS probe (Table 2.3). Since this conclusion contradicts the conclusion of Mantel *et al.* [66], it implies that more research is necessary to understand the impact of TPUS probe pressure on the displacement of the prostate and OARs. Li *et al.* [69] also detected a systematic intra-fraction drift of the prostate. They hypothesized that this drift was caused by the relaxation of the compressed tissue of the perineal area present between the prostate and the probe. As intra-fraction motion monitoring is possible using TPUS imaging, this drift can be monitored and, when needed, potentially compensated for.

With TPUS imaging the probe does not need to be removed prior to dose delivery. Therefore, no displacement of the prostate and organs at risk due to probe removal is expected. As long as the pressure is not so high that it produces a shift of the OARs into high dose regions (as reported e.g. for the penile bulb in the previous paragraph) and it is reproducible, the consequences of the pressure in the US

guided RT workflow should be minimal. For TAUS imaging, it was reported that it is difficult to reproduce the pressure [68], however, for TPUS imaging results on this issue are currently not available. If future studies prove that it is feasible to position the TPUS probe with a reproducible probe pressure, it would add another advantage to this imaging modality in comparison to TAUS imaging.

2.5.2 US image interpretation

The variation in US probe pressure applied by different operators may influence the displacement of the prostate and thus result in US image variation. However, also during interpretation of the images inter- and intra-operator variability can occur. This variability seems to be more present in operators with limited US imaging experience. For this reason, the importance of training has been emphasized by the American Association of Physicist in Medicine [70].

The inter- and intra-operator variability for different levels of expertise has been investigated in a few studies (Table 2.4). In these studies, the operators were asked to match a reference contour of the prostate to a newly acquired US image to determine the required setup shift during inter-fraction motion monitoring. Subsequently, differences in the performed matches were statistically examined.

The results reported by Fiandra *et al.* [71] show that the inter-user variability decreases with growing TAUS imaging experience. The same holds for the intra-user variability during TPUS imaging, as reported by Pang *et al.* [72]. The operators that matched the images in the study of van der Meer *et al.* [11] received thorough training and scanning instructions. These operators seem to perform similarly to the operators with more than one year experience of Fiandra *et al.* [71].

In Table 2.1, the results reported by Robinson *et al.* [44] regarding the differences in prostate localization between TAUS imaging and CT are listed. These results seem to confirm as well that more experience (clinical operator versus manufacturer representative) results in better agreement between the CT and TAUS based prostate locations.

In addition to providing training to the operators, making the system less prone to operator dependence could potentially reduce both inter- and intra-operator variability. In comparison with the Clarity system, the Clarity Autoscan system has already implemented several improvements to potentially reduce operator dependence. In particular, the mechanically swept probe could be attractive, since it minimizes the disadvantages of a manual sweep acquisition, such as the variance in probe pressure and sweeping motion. In addition, the probe is attached to a baseplate avoiding the need to hold it by hand and the operator is assisted to reproduce the earlier used probe pressure and setup by means of visual feedback.

Table 2.4: Studies reporting on the inter- and intra-operator variability of the Clarity system (TAUS) or the Clarity Autoscan system (TPUS). The first column details the first author and publication year. The * indicates that the detailed results were discussed in the review paper [13] as well, while the △ highlights that these specific results were not discussed, but the paper was included in the previous review. The second column specifies the used system, while the third and fourth column provide the number of compared operators and their experience, respectively. In the fifth column the number of examined patients is specified, while in the sixth column the number of matches made by the operators is detailed. The seventh column explains the metric that was used to quantify the intra-operator variability (column eight) and the inter-operator variability (column nine).

First author	System	# operators	Experience	# pts	# matches	Metric	Intra AP	Intra LR	Intra SI	Inter AP	Inter LR	Inter SI
Van der Meer* (2013) [11]	TAUS	2 (intra) 3 (inter)	-	13	817	SD (mm)	0.7	0.8	1.0	1.4	1.3	1.8
Fiandra△ (2014) [71]	TAUS	2	Expert (>5 years)	10	60	mean ± SD of operator Δ (mm)	-	-	-	-0.1 ± 1.4	-0.4 ± 1.2	0.1 ± 1.3
		5	> 1 year	10	150	root mean square error (mm)	-	-	-	2.1 ± 2.1	1.3 ± 1.7	1.7 ± 1.7
		4	< 1 year	10	120	with respect to expert (mm)	-	-	-	3.1 ± 2.7	2.7 ± 2.7	3.2 ± 3.2
Pang (2017) [72]	TPUS	7	All operators	10	70	Δ between operator and total group median (mm) in all directions	≤ 2 in 93.3% of the time			≤ 2 in 93.8% of the time		
		5	9 - 16 months	10	50		≤ 2 in 96.7% of the time			-		
		2	4 - 9 months	10	20		≤ 2 in 60% and 80% of the time			-		

Another approach to reduce operator dependence and potentially even allow less trained operators to acquire good-quality images was proposed by Camps *et al.* [73], [74]. In this work, the simulation CT scan of prostate cancer patients (currently almost always available for treatment planning purposes) was used to optimize the patient-specific US probe setup that would allow visualization of all the required anatomical structures with sufficient image quality. This helps to reduce the need for image interpretation during the acquisition and the operator variability in probe positioning.

2.6 Challenges

Some challenges associated with the use of US imaging in the RT workflow have already been described in the previous sections, such as the inter- and intra-operator variability and the displacement of anatomical structures due to probe pressure. In this section, a number of other challenges associated with the implementation of US imaging in the prostate RT workflow are discussed.

2.6.1 Intra-fraction US imaging

The presence of the US probe in the radiation beam during the treatment can potentially cause dose delivery errors, which might influence the treatment outcome for the patient. Three possible solutions have been proposed in the literature for this problem. One option is to design the treatment plan in such a way that the US probe is completely avoided during the treatment [75]. Second, the radiation can be delivered through the probe, but it requires that the possible dose deviations are taken into account during the treatment planning process, as investigated by e.g. by Bazalova-Carter *et al.* [76]. As a third solution, Schlosser *et al.* [77] designed a 4D radiolucent US probe with significantly less metal components close to the imaging field. This probe should produce a minimal interference with the radiation beam.

Martyn *et al.* [78] also investigated the effect of an US probe on the surface dose delivered to a phantom using a Monte Carlo study. In this study, a phantom was imaged using an Elekta Autoscan probe parallel to the radiation beam to mimic TAUS imaging, or perpendicular to the beam, to mimic TPUS imaging. It was shown that the presence of the probe in the TPUS configuration produces dose perturbations near the surface of the phantom, when there is overlap between the probe and the radiation field. However, the dose increase was of a similar order of magnitude as the one resulting from inter-fraction motion. In case no probe-field overlap occurred, the measured dosimetric effect was minimal. In the TAUS probe setup, instead, a dose increase near the surface of the phantom was measured and reported to be smaller than 5%.

Several studies (e.g. [75], [79]–[81]) also looked into the possibility of replacing a human operator handling the probe at the bedside with a robot. Schlosser *et al.* [75], for example, built a patient-safe robotic manipulator which

could be used to control the pitch and pressure of a TAUS probe. To safely control the robot remotely from outside the Linac room, a haptic device was added to the design. During the treatment delivery, the beam angles were restricted to prevent collision with the robotic hardware or the probe. The authors showed that the robotic system was able to image the prostate remotely. In addition, both the tracking ability of the US probe and the robot performance were not degraded during radiation beam operation. The use of such a robotic system could not only enable intra-fraction TAUS imaging, but also potentially allow for an easier probe pressure and position reproduction using both TAUS and TPUS imaging.

2.6.2 Speed of sound and refraction effects

Most clinical US systems work in pulse-echo mode, where the time of flight of the US pulses is used to infer the depth of the structures in the scanned tissues. This time of flight is calculated with the speed of sound (SOS) of the tissues traversed by the pulse. Different tissues have a different SOS. For example, adipose tissue typically has an SOS around 1450 m/s, while for connective tissue it is around 1600 m/s [82].

However, the US systems usually assume a fixed average SOS value of 1540 m/s for all human soft tissues [83]. This assumption may produce wrong quantitative estimates of organ boundary positions up to several millimeters. Fontanarosa *et al.* published multiple studies ([84]–[87]) in which CT scans were used to create SOS maps for correcting these aberrations. These corrections are essential to restore quantitative comparability with the reference simulation CT scan.

The usability of US imaging in the RT workflow does not only rely on the acquisition and interpretation of the US images. Also, the precision of the calibration procedure of the localization system and associated with that, the precision that can be achieved while localizing the US probe in absolute coordinates in the simulation or treatment room, is of importance. How well the US probe is localized influences the co-registration between e.g. the simulation CT scan and the reference US image, or two US images acquired at different time points.

The phantoms used in a calibration procedure are typically made of homogeneous tissue equivalents to avoid the SOS effects. In addition, refractions inside the phantom should not affect the calibration procedure. However, in the work of Ballhausen *et al.* [88] it has been shown that the calibration of a 3D US system can be affected by refraction of the sound waves at the phantom surface. Especially when the probe was tilted during the calibration procedure this could result in a position difference of more than 0.5 mm.

Van der Meer *et al.* [89] simulated five different scenarios mimicking the errors that could occur when using the Clarity system for TAUS image guidance. These errors could be due to e.g. the above-mentioned inaccurate calibration, but also due to laser offsets or patient motion between the simulation CT and simulation US

image acquisition. It has been shown that it is important to take SOS aberrations into account and to assess the matching of US and CT images. In case these images do not match, a manual correction could be performed, potentially introducing operator variability. In such a case, the authors recommend to rescan the patient to avoid problems during the dose delivery procedure.

Summarizing, it is important to take SOS aberrations into account while registering US images to another image modality. In addition, caution should be used while performing calibration and image acquisition, to avoid image matching issues.

2.6.3 Hypo-fractionation and adaptive radiotherapy

In current clinical practice, it is common to deliver the radiation dose to prostate cancer patients in multiple treatment fractions (even up to 45). It has been suggested that hypo-fractionation could result in the same or better outcomes for the prostate patients [90]. In a hypo-fractionation scheme, a higher dose per fraction is delivered to the patient in less treatment fractions. The treatment is then delivered over a shorter amount of time and with a total lower dose.

As the dose delivered per treatment fraction is higher and there are fewer fractions to potentially perform corrections or compensate for errors originating from the previous fractions, it is even more crucial to deliver the radiation correctly. Ricardi *et al.* [91] used the Clarity system in the treatment of intermediate risk prostate cancer patients treated with a hypo-fractionated schedule. It was shown that the hypo-fractionated schedule under US guidance was a safe and effective treatment approach with consistent biochemical control and a mild toxicity profile.

Patient immobilization during the treatment fraction is also an important aspect of the RT workflow. For this reason, a wide range of immobilization devices is available on the market, ranging from a simple leg immobilizer (Civco Medical Solutions, IA, USA) to vacuum cushions (e.g. Vac-Lok, Civco Medical Solutions, IA, USA) that can adapt to the body composition of the patient. Pang *et al.* [92] investigated the inter-fraction setup differences, patient satisfaction and radiation therapist satisfaction regarding two immobilization devices: the traditionally used leg immobilizer and the Clarity Autoscan immobilization device. The results showed that the setup errors were smaller with the Clarity device and the patients were satisfied with the new device. The radiation therapist, though, had some issues with the weight and bulkiness of the new device.

Adaptive RT (ART) aims at reducing or compensating for the effects of patient-specific treatment variation measured during the course of a radiotherapy treatment [93], [94] by adaptively modifying the treatment plan of the patient. This approach could be used to further improve the accuracy of radiation dose delivery. However, in current clinical practice, typically CT scans provide the electron-density information necessary for treatment planning and dose

calculation. So, in case re-planning proves necessary, one or multiple additional CT scans during the course of the treatment must be acquired. This does not only result in extra radiation dose delivery to the patient, but also high costs are associated with the rather complex CT acquisition procedure.

Van der Meer *et al.* [95] and Camps *et al.* [96] have investigated the feasibility of creating pseudo-CT scans of the pelvic region, based on combinations of rigid and deformable image registrations of TAUS images. These TAUS images, acquired at simulation stage and during treatment stage, were used to create a deformation field that represented the changes that occurred in tissue distribution between these time points. The subsequent application of this deformation field on the simulation CT, resulted in the creation of a pseudo-CT scan.

It was shown that this pseudo-CT scan represents the anatomy of the patient at treatment stage better than the simulation CT. These results are promising and may, lead to the ability to re-plan based on a pseudo-CT scan, instead of on a regular CT scan.

2.7 Conclusion

In this work, the recent relevant studies regarding the use of US imaging for guidance during the prostate EBRT workflow have been discussed. Several US based guidance systems have been introduced to the market in the last 15 years with varying success. TPUS imaging seems to overcome some of the issues associated with the limitations of TAUS imaging during intra-fraction organ motion monitoring, such as displacement of the organs due to probe pressure and the interference with the radiation beam.

The studies that investigated TPUS imaging show promising results and, for this reason, we recommend the use of TPUS imaging during the US guided EBRT workflow of prostate cancer patients. However, there are still several challenges to be addressed, which are associated with inter- and intra-operator variability during the acquisition of the images and the interpretation of these images. In addition, technical aspects of the US image modality, such as SOS aberrations and refractions should be investigated further to understand if these cause issues while using TPUS imaging for both inter- and intra-fraction monitoring.

If a decrease in user variability and an increase of usability of the US guided EBRT systems can be achieved, this would potentially make the use of this approach more appealing to physicians and medical experts, in the end, resulting in smaller margins with less toxicities for prostate cancer patients undergoing EBRT.

References

[1]   C. Fitzmaurice, C. Allen, R. Barber, L. Barregard, Z. Bhutta, H. Brenner, D. Dicker, O. Chimed-Orchir, R. Dandona, L. Dandona, et al. "Global, regional, and national cancer incidence, mortality, years of life lost, years

lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study". In: *JAMA oncology* 3.4 (2017), pp. 524–548.

[2] L. Torre, R. Siegel, E. Ward, and A. Jemal. "Global cancer incidence and mortality rates and trends—an update". In: *Cancer Epidemiology and Prevention Biomarkers* 25.1 (2016), pp. 16–27.

[3] B. Vanneste, E. Van Limbergen, E. van Lin, J. van Roermund, and P. Lambin. "Prostate cancer radiation therapy: What do clinicians have to know?" In: *BioMed research international* 2016 (2016).

[4] M. Zelefsky, M. Kollmeier, B. Cox, A. Fidaleo, D. Sperling, X. Pei, B. Carver, J. Coleman, M. Lovelock, and M. Hunt. "Improved clinical outcomes with high-dose image guided radiotherapy compared with non-IGRT for the treatment of clinically localized prostate cancer". In: *International Journal of Radiation Oncology* Biology* Physics* 84.1 (2012), pp. 125–129.

[5] J. Sveistrup, P. af Rosenschöld, J. Deasy, J. Oh, T. Pommer, P. Petersen, and S. Engelholm. "Improvement in toxicity in high risk prostate cancer patients treated with image-guided intensity-modulated radiotherapy compared to 3D conformal radiotherapy without daily image guidance". In: *Radiation Oncology* 9.1 (2014), p. 44.

[6] U. van der Heide, A. Kotte, H. Dehnad, P. Hofman, J. Lagenijk, and M. van Vulpen. "Analysis of fiducial marker-based position verification in the external beam radiotherapy of patients with prostate cancer". In: *Radiotherapy and Oncology* 82.1 (2007), pp. 38–45.

[7] J. Lagendijk, B. Raaymakers, and M. Van Vulpen. "The magnetic resonance imaging–linac system". In: *Seminars in radiation oncology*. Vol. 24. 3. 2014, pp. 207–209.

[8] R. Foster, T. Solberg, H. Li, A. Kerkhoff, C. Enke, T. Willoughby, and P. Kupelian. "Comparison of transabdominal ultrasound and electromagnetic transponders for prostate localization". In: *Journal of applied clinical medical physics* 11.1 (2010), pp. 57–67.

[9] A. Fung, K. Ayyangar, D. Djajaputra, R. Nehru, and C. Enke. "Ultrasound-based guidance of intensity-modulated radiation therapy". In: *Medical Dosimetry* 31.1 (2006), pp. 20–29.

[10] D. Shen, Y. Zhan, and C. Davatzikos. "Segmentation of prostate boundaries from ultrasound images using statistical shape model". In: *IEEE transactions on medical imaging* 22.4 (2003), pp. 539–551.

[11] S. van der Meer, E. Bloemen-van Gurp, J. Hermans, R. Voncken, D. Heuvelmans, C. Gubbels, D. Fontanarosa, P. Visser, L. Lutgens,

F. van Gils, and F. Verhaegen. "Critical assessment of intramodality 3D ultrasound imaging for prostate IGRT compared to fiducial markers". In: *Medical physics* 40.7 (2013).

[12] M. Lachaine and T. Falco. "Intrafractional prostate motion management with the Clarity Autoscan system". In: *Med. Phys. Int.* 1 (2013).

[13] D. Fontanarosa, S. Van der Meer, J. Bamber, E. Harris, T. O'Shea, and F. Verhaegen. "Review of ultrasound image guidance in external beam radiotherapy: I. Treatment planning and inter-fraction motion management". In: *Physics in Medicine & Biology* 60.3 (2015).

[14] T. O'Shea, J. Bamber, D. Fontanarosa, S. van der Meer, F. Verhaegen, and E. Harris. "Review of ultrasound image guidance in external beam radiotherapy part II: intra-fraction motion management and novel applications". In: *Physics in Medicine & Biology* 61.8 (2016).

[15] C. Parker, A. Damyanovich, T. Haycocks, M. Haider, A. Bayley, and C. Catton. "Magnetic resonance imaging in the radiation treatment planning of localized prostate cancer using intra-prostatic fiducial markers for computed tomography co-registration". In: *Radiotherapy and Oncology* 66.2 (2003), pp. 217–224.

[16] M. Milosevic, S. Voruganti, R. Blend, H. Alasti, P. Warde, M. McLean, P. Catton, C. Catton, and M. Gospodarowicz. "Magnetic resonance imaging (MRI) for localization of the prostatic apex: comparison to computed tomography (CT) and urethrography". In: *Radiotherapy and oncology* 47.3 (1998), pp. 277–284.

[17] C. Rasch, I. Barillot, P. Remeijer, A. Touw, M. van Herk, and J. Lebesque. "Definition of the prostate in CT and MRI: a multi-observer study". In: *International Journal of Radiation Oncology* Biology* Physics* 43.1 (1999), pp. 57–66.

[18] P. Horsley, N. Aherne, G. Edwards, L. Benjamin, S. Wilcox, C. McLachlan, H. Assareh, R. Welshman, M. McKay, and T. Shakespeare. "Planning magnetic resonance imaging for prostate cancer intensity-modulated radiation therapy: Impact on target volumes, radiotherapy dose and androgen deprivation administration". In: *Asia-Pacific Journal of Clinical Oncology* 11.1 (2015), pp. 15–21.

[19] G. Bentel. *Patient positioning and immobilization in radiation oncology*. McGraw-Hill, 1999.

[20] A. O'Neill, S. Jain, A. Hounsell, and J. O'Sullivan. "Fiducial marker guided prostate radiotherapy: a review". In: *The British journal of radiology* 89.1068 (2016).

[21] M. Oldham, D. Létourneau, L. Watt, G. Hugo, D. Yan, D. Lockman, L. Kim, P. Chen, A. Martinez, and J. Wong. "Cone-beam-CT guided radiation therapy: A model for on-line application". In: *Radiotherapy and oncology* 75.3 (2005).

[22] K. Langen and D. Jones. "Organ motion and its management". In: *International Journal of Radiation Oncology* Biology* Physics* 50.1 (2001), pp. 265–278.

[23] M. Van Herk. "Errors and margins in radiotherapy". In: *Seminars in radiation oncology*. Vol. 14. 1. Elsevier, 2004, pp. 52–64.

[24] C. Hill, J. Bamber, and G. ter Haar. *Physical principles of medical ultrasonics*. 2004.

[25] J. Bushberg. *The essential physics of medical imaging*. Lippincott Williams & Wilkins, 2002.

[26] R. Aarnink, H. Beerlage, J. De La Rosette, F. Debruyne, and H. Wijkstra. "Transrectal ultrasound of the prostate: innovations and future applications". In: *The Journal of urology* 159.5 (1998), pp. 1568–1579.

[27] S. Gill, J. Li, J. Thomas, M. Bressel, K. Thursky, C. Styles, K. Tai, G. Duchesne, and F. Foroudi. "Patient-reported complications from fiducial marker implantation for prostate image-guided radiotherapy". In: *The British journal of radiology* 85.1015 (2012), pp. 1011–1017.

[28] J. Langenhuijsen, E. van Lin, L. Kiemeney, L. van der Vight, G. McColl, A. Visser, and J. Witjes. "Ultrasound-guided transrectal implantation of gold markers for prostate localization during external beam radiotherapy: complication rate and risk factors". In: *International Journal of Radiation Oncology* Biology* Physics* 69.3 (2007), pp. 671–676.

[29] J. Huang Foen Chung, S. De Vries, R. Raaijmakers, R. Postma, J. Bosch, and R. van Mastrigt. "Prostate volume ultrasonography: the influence of transabdominal versus transrectal approach, device type and operator". In: *European urology* 46.3 (2004), pp. 352–356.

[30] R. Uppot, D. Sahani, P. Hahn, M. Kalra, S. Saini, and P. Mueller. "Effect of obesity on image quality: fifteen-year longitudinal study for evaluation of dictated radiology reports". In: *Radiology* 240.2 (2006), pp. 435–439.

[31] R. Uppot. "Impact of obesity on radiology". In: *Radiologic clinics of North America* 45.2 (2007), pp. 231–246.

[32] B. Salter, M. Szegedi, C. Boehm, V. Sarkar, P. Rassiah-Szegedi, B. Wang, H. Zhao, J. Huang, L. Huang, K. Kokeny, and J. Tward. "Comparison of 2 transabdominal ultrasound image guidance techniques for prostate

and prostatic fossa radiation therapy". In: *Practical Radiation Oncology* 7.2 (2017).

[33] K. Griffiths, L. Ly, B. Jin, L. Chan, and D. Handelsman. "Transperineal ultrasound for measurement of prostate volume: validation against transrectal ultrasound". In: *The Journal of urology* 178.4 (2007), pp. 1375–1380.

[34] K. Shinohara and M. Roach. "Technique for implantation of fiducial markers in the prostate". In: *Urology* 71.2 (2008), pp. 196–200.

[35] S. Igdem, H. Akpinar, G. Alco, F. Agacayak, S. Turkan, and S. Okkan. "Implantation of fiducial markers for image guidance in prostate radiotherapy: patient-reported toxicity". In: *The British journal of radiology* 82.983 (2009), pp. 941–945.

[36] J. Loh, K. Baker, S. Sridharan, P. Greer, C. Wratten, A. Capp, S. Gallagher, and J. Martin. "Infections after fiducial marker implantation for prostate radiotherapy: are we underestimating the risks?" In: *Radiation Oncology* 10.1 (2015), p. 38.

[37] P. Kupelian, T. Willoughby, S. Meeks, A. Forbes, T. Wagner, M. Maach, and K. Langen. "Intraprostatic fiducials for localization of the prostate gland: monitoring intermarker distances during radiation therapy to test for marker stability". In: *International Journal of Radiation Oncology* Biology* Physics* 62.5 (2005), pp. 1291–1296.

[38] J. Chow and G. Grigorov. "Dose measurements near a non-radioactive gold seed using radiographic film". In: *Physics in medicine and biology* 50.18 (2005).

[39] J. Perks, J. Lehmann, A. Chen, C. Yang, R. Stern, and J. Purdy. "Comparison of peripheral dose from image-guided radiation therapy (IGRT) using kV cone beam CT to intensity-modulated radiation therapy (IMRT)". In: *Radiotherapy and oncology* 89.3 (2008), pp. 304–310.

[40] B. Tas, I. F. Durmus, and S. T. Ozturk. "Image Guided Radiotherapy (IGRT) comparison between cone beam CT and ultrasound system for prostate cancer". In: *Universal Journal of Physics and Application* 10.4 (2016), pp. 110–114.

[41] H. Johnston, M. Hilts, W. Beckham, and E. Berthelet. "3D ultrasound for prostate localization in radiation therapy: a comparison with implanted fiducial markers". In: *Medical physics* 35.6 (2008), pp. 2403–2413.

[42] E. Mayyas, I. Chetty, M. Chetvertkov, N. Wen, T. Neicu, T. Nurushev, L. Ren, M. Lu, H. Stricker, D. Pradhan, B. Movsas, and M. Elshaikh. "Evaluation of multiple image-based modalities for image-guided

radiation therapy (IGRT) of prostate carcinoma: A prospective study". In: *Medical physics* 40.4 (2013).

[43] F. Cury, G. Shenouda, L. Souhami, M. Duclos, S. Faria, M. David, F. Verhaegen, R. Corns, and T. Falco. "Ultrasound-based image guided radiotherapy for prostate cancer-comparison of cross-modality and intramodality methods for daily localization during external beam radiotherapy". In: *International Journal of Radiation Oncology* Biology* Physics* 66.5 (2006), pp. 1562–1567.

[44] D. Robinson, D. Liu, S. Steciw, C. Field, H. Daly, E. Saibishkumar, G. Fallone, M. Parliament, and J. Amanie. "An evaluation of the Clarity 3D ultrasound system for prostate localization". In: *Journal of applied clinical medical physics* 13.4 (2012), pp. 100–112.

[45] M. Li, H. Ballhausen, N. Hegemann, U. Ganswindt, F. Manapov, S. Tritschler, A. Roosen, C. Gratzke, M. Reiner, and C. Belka. "A comparative assessment of prostate positioning guided by three-dimensional ultrasound and cone beam CT". In: *Radiation Oncology* 10.1 (2015), p. 82.

[46] M. Fargier-Voiron, B. Presles, P. Pommier, A. Munoz, S. Rit, D. Sarrut, and M. Biston. "Ultrasound versus Cone-beam CT image-guided radiotherapy for prostate and post-prostatectomy pretreatment localization". In: *Physica Medica* 31.8 (2015), pp. 997–1004.

[47] M. Krengli, G. Loi, C. Pisani, D. Beldì, G. Apicella, V. Amisano, and M. Brambilla. "Three-dimensional surface and ultrasound imaging for daily IGRT of prostate cancer". In: *Radiation Oncology* 11.1 (2016), p. 159.

[48] A. Richter, B. Polat, I. Lawrenz, S. Weick, O. Sauer, M. Flentje, and F. Mantel. "Initial results for patient setup verification using transperineal ultrasound and cone beam CT in external beam radiation therapy of prostate cancer". In: *Radiation Oncology* 11.1 (2016), p. 147.

[49] M. Fargier-Voiron, B. Presles, P. Pommier, A. Munoz, S. Rit, D. Sarrut, and M. Biston. "Evaluation of a new transperineal ultrasound probe for inter-fraction image-guidance for definitive and post-operative prostate cancer radiotherapy". In: *Physica Medica* 32.3 (2016), pp. 499–505.

[50] A. Trivedi, T. Ashikaga, D. Hard, J. Archambault, M. Lachaine, D. Cooper, and H. Wallace. "Development of 3-dimensional transperineal ultrasound for image guided radiation therapy of the prostate: Early evaluations of feasibility and use for inter-and intrafractional prostate localization". In: *Practical radiation oncology* 7.1 (2017).

[51] M. Li, H. Ballhausen, N. Hegemann, M. Reiner, S. Tritschler, C. Gratzke, F. Manapov, S. Corradini, U. Ganswindt, and C. Belka. "Comparison of

prostate positioning guided by three-dimensional transperineal ultrasound and cone beam CT". In: *Strahlentherapie und Onkologie* 193.3 (2017), pp. 221–228.

[52] J. Bland and D. Altman. "Statistical methods for assessing agreement between two methods of clinical measurement". In: *The lancet* 327.8476 (1986), pp. 307–310.

[53] G. Meijer, J. De Klerk, K. Bzdusek, H. Van Den Berg, R. Janssen, M. Kaus, P. Rodrigus, and P. van der Toorn. "What CTV-to-PTV margins should be applied for prostate irradiation? Four-dimensional quantitative assessment using model-based deformable image registration techniques". In: *International Journal of Radiation Oncology* Biology* Physics* 72.5 (2008), pp. 1416–1425.

[54] P. Kupelian, T. Willoughby, A. Mahadevan, T. Djemil, G. Weinstein, S. Jani, C. Enke, T. Solberg, N. Flores, D. Liu, D. Beyer, and L. Levine. "Multi-institutional clinical experience with the Calypso System in localization and continuous, real-time monitoring of the prostate gland during external radiotherapy". In: *International Journal of Radiation Oncology* Biology* Physics* 67.4 (2007), pp. 1088–1098.

[55] J. Balter, J. Wright, L. Newell, B. Friemel, S. Dimmer, Y. Cheng, J. Wong, E. Vertatschitsch, and T. Mate. "Accuracy of a wireless localization system for radiotherapy". In: *International Journal of Radiation Oncology* Biology* Physics* 61.3 (2005), pp. 933–937.

[56] M. Abramowitz, E. Bossart, R. Flook, X. Wu, R. Brooks, M. Lachaine, F. Lathuiliere, and A. Pollack. "Noninvasive real-time prostate tracking using a transperineal ultrasound approach". In: *International Journal of Radiation Oncology* Biology* Physics* 84.3 (2012).

[57] A. Yu, M. Najafi, D. Hristov, and T. Phillips. "Intrafractional tracking accuracy of a transperineal ultrasound image guidance system for prostate radiotherapy". In: *Technology in Cancer Research & Treatment* (2017).

[58] H. Ballhausen, M. Li, N. Hegemann, U. Ganswindt, and C. Belka. "Intrafraction motion of the prostate is a random walk". In: *Physics in medicine and biology* 60.2 (2014), p. 549.

[59] M. Baker and C. Behrens. "Determining intrafractional prostate motion using four dimensional ultrasound system". In: *BMC cancer* 16.1 (2016), p. 484.

[60] A. Richardson and P. Jacobs. "Intrafraction monitoring of prostate motion during radiotherapy using the Clarity® Autoscan Transperineal Ultrasound (TPUS) system". In: *Radiography* 23.4 (2017), pp. 310–313.

[61]　L. Guillet, M. Fargier-Voiron, D. Sarrut, and M. Biston. "Evaluation of intrafraction motions with a transperineal ultrasound imaging system: Dosimetric impact for prostate cancer". In: *Physica Medica: European Journal of Medical Physics* 31 (2015).

[62]　H. Ballhausen, M. Reiner, S. Kantz, C. Belka, and M. Söhn. "The random walk model of intrafraction movement". In: *Physics in medicine and biology* 58.7 (2013), p. 2413.

[63]　S. Hilman, R. Smith, S. Masson, H. Coomber, A. Bahl, A. Challapalli, and P. Jacobs. "Implementation of a daily transperineal ultrasound system as image-guided radiotherapy for prostate cancer". In: *Clinical Oncology* 29.1 (2017).

[64]　S. Hilman and P. Jacobs. "Image-guided radiotherapy for prostate cancer using transperineal ultrasound". In: *RAD magazine* (2017), pp. 29–30.

[65]　M. Baker and C. Behrens. "Prostate displacement during transabdominal ultrasound image-guided radiotherapy assessed by real-time four-dimensional transperineal monitoring". In: *Acta Oncologica* 54.9 (2015), pp. 1508–1514.

[66]　F. Mantel, A. Richter, C. Groh, I. Lawrenz, S. Weick, B. Polat, M. Guckenberger, and M. Flentje. "Changes in penile bulb dose when using the Clarity transperineal ultrasound probe: A planning study". In: *Practical radiation oncology* 6.6 (2016).

[67]　S. Mangar, M. Sydes, H. Tucker, J. Coffey, S. Sohaib, S. Gianolini, S. Webb, V. Khoo, D. Dearnaley, and M. R. T. M. Group. "Evaluating the relationship between erectile dysfunction and dose received by the penile bulb: using data from a randomised controlled trial of conformal radiotherapy in prostate cancer (MRC RT01, ISRCTN47772397)". In: *Radiotherapy and Oncology* 80.3 (2006), pp. 355–362.

[68]　M. Fargier-Voiron, B. Presles, P. Pommier, S. Rit, A. Munoz, H. Liebgott, D. Sarrut, and M. Biston. "Impact of probe pressure variability on prostate localization for ultrasound-based image-guided radiotherapy". In: *Radiotherapy and Oncology* 111.1 (2014), pp. 132–137.

[69]　M. Li, N. Hegemann, F. Manapov, A. Kolberg, P. Thum, U. Ganswindt, C. Belka, and H. Ballhausen. "Prefraction displacement and intrafraction drift of the prostate due to perineal ultrasound probe pressure". In: *Strahlentherapie und Onkologie* (2017), pp. 1–7.

[70]　J. Molloy, G. Chan, A. Markovic, S. McNeeley, D. Pfeiffer, B. Salter, and W. Tome. "Quality assurance of US-guided external beam radiotherapy for prostate cancer: Report of AAPM Task Group 154". In: *Medical physics* 38.2 (2011), pp. 857–871.

[71]   C. Fiandra, A. Guarneri, F. Muñoz, F. Moretto, A. Filippi, M. Levis, R. Ragona, and U. Ricardi. "Impact of the observers' experience on daily prostate localization accuracy in ultrasound-based IGRT with the Clarity platform". In: *Journal of Applied Clinical Medical Physics* 15.4 (2014), pp. 168–173.

[72]   E. Pang, K. Knight, M. Baird, and J. Tuan. "Inter-and intra-observer variation of patient setup shifts derived using the 4D TPUS Clarity system for prostate radiotherapy". In: *Biomedical Physics & Engineering Express* 3.2 (2017).

[73]   S. Camps, F. Verhaegen, G. P. Fonseca, P. de With, and D. Fontanarosa. "Automatic transperineal ultrasound probe positioning based on CT scan for image guided radiotherapy". In: *Proc. SPIE medical imaging*. Vol. 10135. Orlando, USA, 2017.

[74]   S. Camps, F. Verhaegen, P. de With, and D. Fontanarosa. "CT scan based prostate cancer patient-specific transperineal ultrasound probe setups for image guided radiotherapy". In: *Proc. IEEE International Ultrasonics Symposium*. 2017.

[75]   J. Schlosser, K. Salisbury, and D. Hristov. "Telerobotic system concept for real-time soft-tissue imaging during radiotherapy beam delivery". In: *Medical physics* 37.12 (2010), pp. 6357–6367.

[76]   M. Bazalova-Carter, J. Schlosser, J. Chen, and D. Hristov. "Monte Carlo modeling of ultrasound probes for image guided radiotherapy". In: *Medical physics* 42.10 (2015), pp. 5745–5756.

[77]   J. Schlosser and D. Hristov. "Radiolucent 4D ultrasound imaging: System design and application to radiotherapy guidance". In: *IEEE transactions on medical imaging* 35.10 (2016), pp. 2292–2300.

[78]   M. Martyn, T. O'Shea, E. Harris, J. Bamber, S. Gilroy, and M. Foley. "A Monte Carlo study of the effect of an ultrasound transducer on surface dose during intrafraction motion imaging for external beam radiation therapy". In: *Medical Physics* 44.10 (2017), pp. 5020–5033.

[79]   M. Bell, H. Şen, I. Iordachita, P. Kazanzides, and J. Wong. "In vivo reproducibility of robotic probe placement for a novel ultrasound-guided radiation therapy system". In: *Journal of Medical Imaging* 1.2 (2014).

[80]   H. Şen, M. Bell, Y. Zhang, K. Ding, J. Wong, I. Iordachita, and P. Kazanzides. "System integration and preliminary in-vivo experiments of a robot for ultrasound guidance and monitoring during radiotherapy". In: *Proceedings of International Conference on Advanced Robotics*. 2015, p. 53.

[81]   S. Gerlach, I. Kuhlemann, P. Jauer, R. Bruder, F. Ernst, C. Fürweger, and A. Schlaefer. "Robotic ultrasound-guided SBRT of the prostate: feasibility

with respect to plan quality". In: *International journal of computer assisted radiology and surgery* 12.1 (2017), pp. 149–159.

[82] T. Mast. "Empirical relationships between acoustic parameters in human soft tissues". In: *Acoustics Research Letters Online* 1.2 (2000), pp. 37–42.

[83] P. Wells. *Biomedical ultrasonics*. Academic Pr, 1977.

[84] D. Fontanarosa, S. Meer, E. Harris, and F. Verhaegen. "A CT based correction method for speed of sound aberration for ultrasound based image guided radiotherapy". In: *Medical physics* 38.5 (2011), pp. 2665–2673.

[85] D. Fontanarosa, S. Meer, E. Bloemen-van Gurp, G. Stroian, and F. Verhaegen. "Magnitude of speed of sound aberration corrections for ultrasound image guided radiotherapy for prostate and other anatomical sites". In: *Medical physics* 39.8 (2012), pp. 5286–5292.

[86] D. Fontanarosa, S. van der Meer, and F. Verhaegen. "On the significance of density-induced speed of sound variations on US-guided radiotherapy". In: *Medical physics* 39.10 (2012), pp. 6316–6323.

[87] D. Fontanarosa, S. Pesente, F. Pascoli, D. Ermacora, I. Rumeileh, and F. Verhaegen. "A speed of sound aberration correction algorithm for curvilinear ultrasound transducers in ultrasound-based image-guided radiotherapy". In: *Physics in medicine and biology* 58.5 (2013), p. 1341.

[88] H. Ballhausen, B. Ballhausen, M. Lachaine, M. Li, K. Parodi, C. Belka, and M. Reiner. "Surface refraction of sound waves affects calibration of three-dimensional ultrasound". In: *Radiation Oncology* 10.1 (2015), p. 119.

[89] S. van der Meer, E. Seravalli, D. Fontanarosa, E. Bloemen-van Gurp, and F. Verhaegen. "Consequences of intermodality registration errors for intramodality 3D ultrasound IGRT". In: *Technology in cancer research & treatment* 15.4 (2016), pp. 632–638.

[90] D. Dearnaley, I. Syndikus, H. Mossop, V. Khoo, A. Birtle, D. Bloomfield, J. Graham, P. Kirkbride, J. Logue, Z. Malik, et al. "Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHiP trial". In: *The Lancet Oncology* 17.8 (2016), pp. 1047–1060.

[91] U. Ricardi, P. Franco, F. Munoz, M. Levis, C. Fiandra, A. Guarneri, F. Moretto, S. Bartoncini, F. Arcadipane, S. Badellino, C. Piva, E. Trino, A. Ruggiere, A. Filippi, and R. Ragona. "Three-dimensional ultrasound-based image-guided hypofractionated radiotherapy for intermediate-risk prostate cancer: results of a consecutive case series". In: *Cancer investigation* 33.2 (2015), pp. 23–28.

[92] E. Pang, K. Knight, M. Baird, J. Loh, A. San Boo, and J. Tuan. "A comparison of interfraction setup error, patient comfort, and therapist acceptance for 2 different prostate radiation therapy immobilization devices". In: *Advances in Radiation Oncology* (2017).

[93] D. Yan, F. Vicini, J. Wong, and A. Martinez. "Adaptive radiation therapy". In: *Physics in medicine and biology* 42.1 (1997), p. 123.

[94] M. Ghilezan, D. Yan, and A. Martinez. "Adaptive radiation therapy for prostate cancer". In: *Seminars in radiation oncology*. Vol. 20. 2. Elsevier. 2010, pp. 130–137.

[95] S. van der Meer*, S. Camps*, W. van Elmpt, M. Podesta, P. Sanches, B. Vanneste, D. Fontanarosa, and F. Verhaegen. "Simulation of pseudo-CT images based on deformable image registration of ultrasound images: A proof of concept for transabdominal ultrasound imaging of the prostate during radiotherapy". In: *Medical Physics* 43.4 (2016), pp. 1913–1920.

[96] S. Camps, S. van der Meer, F. Verhaegen, and D. Fontanarosa. "Various approaches for pseudo-CT scan creation based on ultrasound to ultrasound deformable image registration between different treatment time points for radiotherapy treatment plan adaptation in prostate cancer patients". In: *Biomedical Physics & Engineering Express* 2.3 (2016).

Chapter 3

Probe setup calculation for pelvic phantom imaging

Abstract

Image interpretation is crucial during ultrasound image acquisition. A skilled operator is typically needed to verify if the correct anatomical structures are all visualized and with sufficient quality. The need for this operator is one of the major reasons why presently ultrasound is not widely used in radiotherapy workflows.

To solve this issue, we introduce an algorithm that uses anatomical information derived from a CT scan to automatically provide the operator with a patient-specific ultrasound probe setup. The first application we investigated, for its relevance to radiotherapy, is 4D transperineal ultrasound image acquisition for prostate cancer patients.

As initial test, the algorithm was applied on a CIRS multi-modality pelvic phantom. Probe setups were calculated in order to allow visualization of the prostate and adjacent edges of bladder and rectum, as clinically required. Five of the proposed setups were reproduced using a precision robotic arm and ultrasound volumes were acquired. A gel-filled probe cover was used to ensure proper acoustic coupling, while taking into account possible tilted positions of the probe with respect to the flat phantom surface.

Visual inspection of the acquired volumes revealed that clinical requirements were fulfilled. Preliminary quantitative evaluation was also performed. The mean absolute distance (MAD) was calculated between actual anatomical structure positions and positions predicted by the CT-based algorithm. This resulted in a MAD of 2.8 ± 0.4 mm for prostate, 2.5 ± 0.6 mm for bladder and 2.8 ± 0.6 mm for rectum. These results show that no large systemic errors due to e.g. probe misplacement were introduced.

3.1 Introduction

Radiotherapy (RT) is one of the available treatment modalities for various types of cancer. The aim of RT is to deliver a high radiation dose to the tumor tissue, while sparing normal tissue as much as possible. Prior to treatment delivery (at the simulation stage), usually a computed tomography (CT) scan of the patient is acquired, which is then used for treatment plan preparation. The treatment typically involves multiple irradiation fractions, prior to each of which the patient is positioned as accurately as possible to reproduce the position at simulation. A positioning accuracy of a few millimeters is desirable.

Even when the patient seems correctly aligned externally, internal tissue distributions may have changed. For example, in case of prostate cancer, a different filling of the bladder or the rectum can cause a shift of the prostate [1]. If the initial treatment plan is still delivered, this could result in a suboptimal dose deposition in the tumor and/or in extra unwanted irradiation of the normal tissue, possibly producing excessive toxicity.

It is now recognized (see, for example, Zelefsky *et al.* [2] or Yorke *et al.* [3]) that frequent imaging of the patients' anatomical structures of interest during the course of the RT treatment (image guided RT, IGRT) can improve the biochemical tumor control and decrease toxicity. It was shown by Zelefsky *et al.* [2] that correction of the inter-fraction motion of the prostate using kilovoltage imaging of implanted fiducial markers, significantly improved the outcome of the patients. In comparison with kilovoltage imaging of fiducial markers, ultrasound (US) imaging has many advantages, such as: no extra radiation dose delivery, non-invasiveness and a superior soft tissue contrast. In addition, it is presently the only volumetric imaging modality that allows for real-time anatomical structure tracking. For this reason, US imaging could be used not only for inter-fraction motion corrections, but also for intra-fraction monitoring [4], [5].

However, in current clinical practice, image interpretation during US image acquisition in the RT workflow is still crucial. Typically, a skilled operator needs to verify if the correct anatomical structures are visualized and if the image quality is sufficient. This makes the process heavily dependent on this operator. Introducing a certain degree of automation into the US-guided RT workflow can decrease operator dependence and improve the usability for unskilled operators. In addition, the patients could fully benefit from the unique characteristics of US imaging, potentially resulting in higher treatment quality and therefore in better outcomes.

In this proof-of-concept study, we introduce an algorithm that uses anatomical information derived from a CT scan to automatically provide the operator with a patient-specific US probe setup. These proposed probe setups should allow good anatomical structure visualization based on clinical requirements. As far as we know, this was never reported in literature before. We present the first

results achieved using a pelvic phantom, in order to mimic the acquisition of 4D transperineal US images of prostate cancers patients.

3.2 Materials and methods

In Section 3.2.1 both the acquisition of the phantom images as well as the required preprocessing steps are detailed. After these preprocessing steps, suitable US probe setups are proposed by the algorithm. More insight into the workflow of this algorithm is given in Section 3.2.2. The procedure followed to reproduce the proposed probe setups in the experimental setting is detailed in Section 3.2.3. Finally, in Section 3.2.4 the evaluation methods are described.

3.2.1 Image acquisition and preprocessing

Since, we focused on prostate cancer patients as first application, a multi-modality male pelvic phantom (CIRS, model 048, Norfolk, USA) was used. A CT scan of this phantom was acquired (Fig. 3.1A) using a SOMATOM Sensation Open CT scanner (Syngo CT 2006A, Siemens Healthcare GMBH, Germany; voxel size: 0.5 mm x 0.5 mm x 1 mm). The first preprocessing step involved resampling of the voxel size to 1 mm x 1 mm x 1 mm in order to make further processing steps easier. Subsequently, the anatomical structures were segmented using the Eclipse treatment planning system (Varian Medical System, Palo Alto, CA, USA).

| (A) | (B) | (C) | (D) |
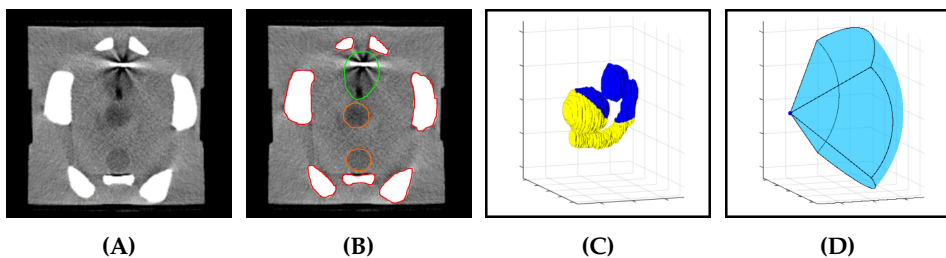
Figure 3.1: Image acquisition and preprocessing workflow for the pelvic phantom. (A) CT scan of the phantom with voxels resampled to 1 mm x 1 mm x 1 mm. (B) CT scan with superimposed the target in green, the organs at risk in orange and the blocking structures in red. (C) Required structures in the FOV in blue, optional structures in yellow. (D) 3D FOV model of the used matrix probe with the blue dot representing the apex of the FOV.

The scanning area of the cubic pelvic phantom was a 16 cm x 16 cm square on one of the faces of the phantom cube. On this scanning area the US probe should be positioned. The anatomical structures located inside the phantom were classified: the prostate was the target, bladder and rectum were considered organs at risk and the bones were identified as US blocking structures (Fig. 3.1B). Clinical requirements provided by an experienced radiation oncologist demanded that the whole prostate be visualized, as well as the adjacent edges of bladder and rectum. To ensure that the required anatomical structures were not too close to the edges of

the field of view (FOV), a margin of 2 mm in all directions was added. The required structures with added margin will be referred to as "expanded structures" from here on. The parts of the bladder and rectum that were not required in the FOV were classified as optional (Fig. 3.1C). Due to the addition of the margin, there was partial overlap between the expanded structures and the optional structures.

The final preprocessing step was the construction of a 3D model of the FOV of the matrix probe, which we used for image acquisition (X6-1 matrix probe in combination with an EpiQ7 US system, Philips Medical Systems, Andover, MA, USA), (Fig. 3.1D).

3.2.2 Probe setup calculation

In Fig. 3.2A the setup of the pelvic phantom with the corresponding coordinate system is shown. The orange surface represents the available scanning area on which the US probe can be positioned. The setup of a virtual patient in the same coordinate system is detailed in Fig. 3.2B. For the patient, the area on which the US probe could be positioned is the perineum.

In this study only four degrees of freedom were allowed to position the US probe on the phantom: translation along Z (up and down), translation along X (left and right) and rotation around these same axes. The position on the Y-axis was considered fixed, as in a clinical situation the probe is positioned against the patient's skin. The rotation around the Y-axis should be evaluated for clinical relevance and may be included in future research.

(A)   **(B)**
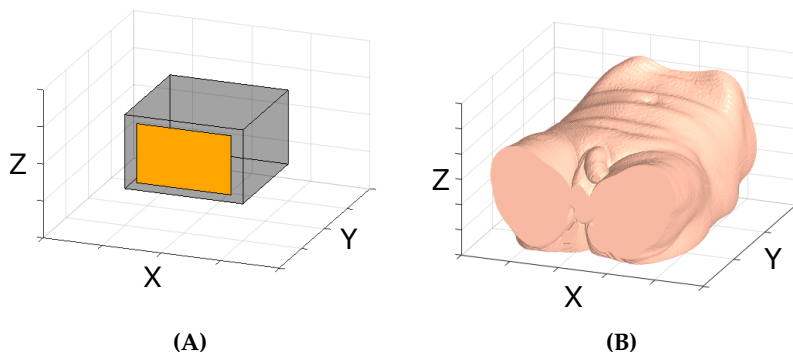
Figure 3.2: Phantom and patient setup in the same coordinate system. (A) Sketch of the phantom configuration with the orange rectangle representing the scanning area. (B) Possible setup of the patient in supine position.

As previously mentioned, the probe setup proposed by the algorithm should allow visualization of the whole prostate as well as the adjacent edges of bladder and rectum. In Fig. 3.3A these required structures are visualized, together with the 2 mm margin added in the preprocessing steps. The red dot located inside the

required structures represents the centroid of the combination of the expanded structures.

In order to propose suitable probe setups to the operator, the algorithm first determined an initial setup. As the setup of the US probe is directly connected to the position of its FOV, the constructed FOV model (Fig. 3.3B) was used to determine this initial setup via three steps.

In the first step, the FOV model was translated along the Y-axis until the Y-coordinate of the origin was at the same Y-coordinate value as the available scanning area of the phantom. The assumption was made that the gel layer needed for proper US scanning was infinitely thin, so the probe could be positioned directly on the surface. At this stage, the algorithm assumed that there was no pressure applied while positioning the probe and consequently that the probe position along the Y-axis was fixed.
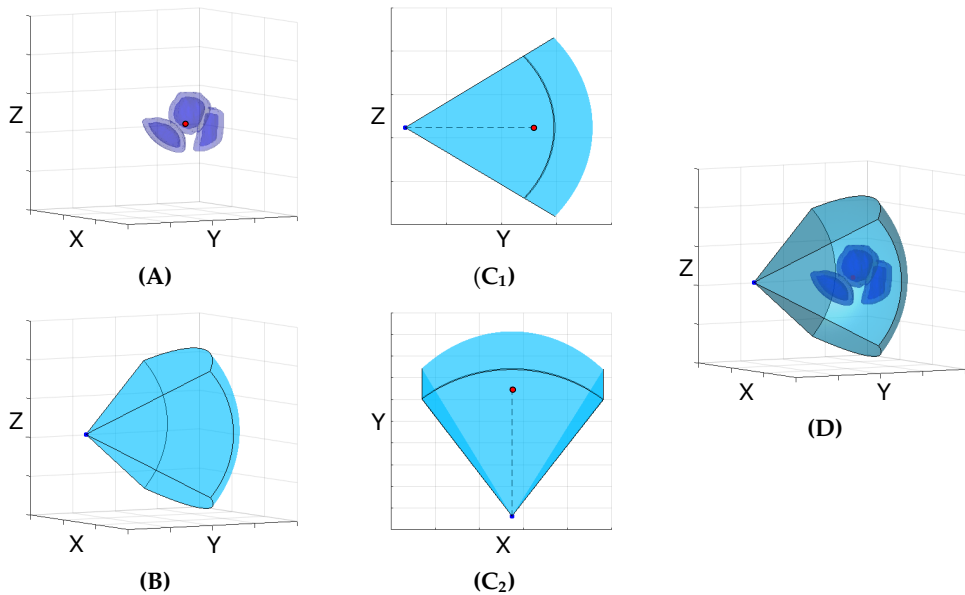
Figure 3.3: Workflow for initial probe setup determination. (A) The darker blue voxels indicate the required structures, while the lighter blue indicates the added margin of 2 mm. The red dot corresponds to the centroid of the expanded structures. (B) 3D FOV model of the X6-1 probe. (C_1) Alignment of the apex of the FOV (blue dot) with the centroid of the expanded structures (red) in the YZ-plane. (C_2) Alignment of the apex and the centroid in the XY-plane. (D) Final result of the initial probe setup determination.

In the second step, the FOV model was translated along the Z-axis until the Z-value of the position of the apex of the FOV (blue dot in Fig. 3.3C_1) was aligned with the corresponding coordinate of the centroid of the expanded structures (red dot in Fig. 3.3C_1) on the YZ-plane. The same procedure was repeated in the third

step, translating the FOV along the X-axis until alignment of the X-coordinates was established on the XY-plane (Fig. 3.3C$_2$).

In Fig. 3.3D the results of this initial setup determination are displayed. Since the centroid of the expanded structures was positioned in the FOV, this procedure was such that the initial FOV position would most likely also result in at least a part of the expanded structures being automatically included in the FOV. In this specific case, the clinical requirements were completely met, as already all expanded structures were in the FOV.

The following iterative steps were optional and based on the outcome of the initial process. If the expanded structures could not be visualized using the initial setup, a procedure was performed to enforce the inclusion of these structures in the FOV. If another procedure was followed instead, because the structures were already included, a further optimization of the position of the FOV aiming at including a larger fraction of the optional structures was performed, but never at the expense of the expanded structures.

The flow chart of the procedure that involves several iterations is plotted in Fig. 3.4. All the voxels of the expanded structures located outside the FOV determined with the initial setup were marked and their centroid was determined. Moreover, also the centroid of the FOV itself was calculated (Column 3.4A).

Subsequently, the flow was split in two arms: procedures on the XY- or on the YZ-plane. For each plane, the angle ϕ between two vectors (marked I and II in Fig. 3.4B) was calculated. These calculated angles will be referred to as XY_ϕ and YZ_ϕ from here on. If the absolute value of XY_ϕ and/or YZ_ϕ was larger than 37.5 degrees (half of the FOV viewing angle of 75 degrees), the FOV position and orientation were adjusted. The position of the FOV could be adjusted by translating the US probe along the X- or Z-axis using steps of 1 mm. The sign of the previously calculated angles XY_ϕ and YZ_ϕ determined if this translation occurred in positive or negative direction. A positive XY_ϕ resulted in translation along the negative X-axis, while a negative XY_ϕ resulted in translation along the positive X-axis. On the other hand, a positive YZ_ϕ resulted in translation along the positive Z-axis, while a negative YZ_ϕ resulted in translation along the negative Z-axis (Column 3.4C). If only translation was taken into account, two different new probe setups were proposed, resulting from translation along X-axis or Z-axis.

As mentioned before, also the orientation of the FOV could be adjusted. This could be done by performing rotation around the X- or Z-axis in clockwise or counterclockwise direction. Also in this case, the sign of the angles XY_ϕ and YZ_ϕ determined in which direction the rotation was performed. The procedure to determine this direction was similar to the procedure as used in the translation and it is detailed in Column 3.4C. If only rotation was taken into account, two additional new probe setups were proposed, resulting from rotation around X- or Z-axis. The rotation angles depended on the absolute values of the calculated angles XY_ϕ and YZ_ϕ and the viewing angle of the FOV (75 degrees).

In addition, the angle of rotation could not exceed a certain set maximum. This maximum was set to 5 degrees in order to ensure that the acoustic coupling between the probe and the scanning area of the phantom would not be lost. More research in the future is necessary to verify the exact magnitude of this maximum angle. In order to calculate the angle of rotation supposing that YZ_ϕ is 39 degrees, the difference between the absolute angle (39 degrees) and half of the viewing angle (37.5 degrees) was calculated, resulting in 1.5 degrees. As this rotation angle is smaller than the set maximum of 5 degrees, the FOV should be rotated over 1.5 degrees in clockwise direction as indicated by the last box of Column 3.4C.

Also a combination of translation and rotation is among the possibilities, as can be seen in Column 3.4D of the flow chart. This results in eight possible probe setups in total. Subsequently, one or multiple probe setups must be chosen to feed back into the flow chart to continue the iterative process. For this reason, the cost of each of the eight proposed probe setups was calculated. This cost was defined as the number of expaned voxels still located outside of the FOV. The setups could then be ranked based on their cost.
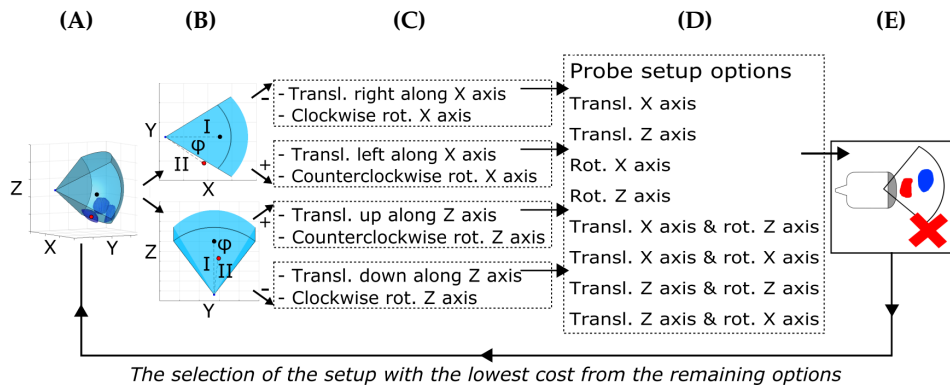
Figure 3.4: Flow chart of the steps necessary to iteratively determine suitable US probe setups. (Column A) Initial setup of the FOV with the red dot representing the centroid of the expanded structures that are located outside of this FOV (outside structures) and the black dot representing the centroid of the FOV itself. (Column B) Calculation of the angles (ϕ) between two vectors (FOV apex - FOV centroid [I] and FOV apex - outside structures centroid [II]) in both YZ and XY-plane. (Column C) Possible actions for the FOV depending on the sign and magnitude of the angles XY_ϕ and YZ_ϕ. (Column D) All possible probe setup options. (Column E) Check to ensure that no blocking occurs by US wave blocking structures. Finally, the best setup (with the lowest cost) among the probe setup options is selected and used as initial setup. This workflow is repeated until the solution converges.

Some anatomical structures (in particular bones, which have an acoustic impedance significantly different from soft tissues) can reflect a substantial percentage of the US wave intensity. Therefore, a check for all eight possible probe setups was performed to ensure that no blockage occurred (Column 3.4E). In case blockage occurred, the corresponding probe setup was removed from the list with possible setups.

Finally, among the list with possible setups, the setup with the lowest cost was selected. This setup was used as the new initial probe setup and the flow was repeated. This was done until no better probe setup could be found and the solution converged. In case two setups had the exact same cost, both solutions were used as initial probe setup and basically the flow was extended with two extra arms.

In case all expanded structures were already inside the FOV using the initial setup, the procedure is slightly different. Instead of determining the centroid of the expanded structures outside of the FOV, the centroid of the optional structures outside the FOV needs to be determined. This means that Column 3.4A of the flow chart is different. The Columns 3.4B - 3.4E are then identical and an additional column must be added at the end. Before the best setup is chosen, first it needs to be determined if the expanded structures are still fully in the FOV with the newly proposed probe setups. If this is not the case, the corresponding probe setups are removed from the list with possibilities and the best setup is chosen subsequently. Also in this case, this procedure was repeated until the solution converged.

3.2.3   Probe setup reproduction

After calculation of suitable probe setups, these configurations need to be reproduced in the experimental setting. For this purpose, the US probe was attached to a robotic arm (R17 5-axis, ST Robotics, Cambridge, UK) and the robot (Fig. 3.5A) was used to position the probe according to the calculated setups.

It was important to ensure sufficient airless coupling between the US probe and the scanning area of the phantom. For this purpose a lubricated condom without a reservoir tip (World's Best, Denmark) was partially filled with US gel. This gel-filled cover was then placed on top of the US probe (Fig. 3.5B and Fig. 3.5C). As the gel-filled cover was constructed of flexible material, it allowed deformation when the probe was pressed against the scanning area of the phantom. However, the gel could not escape from the cover, so sufficient coupling was ensured at all times.

Subsequently, a calibration procedure for the robotic arm was performed to ensure alignment of the origin and coordinate systems between the algorithm and the robot. In this way, the translations and rotations provided by the algorithm could be directly input to the software of the robot.
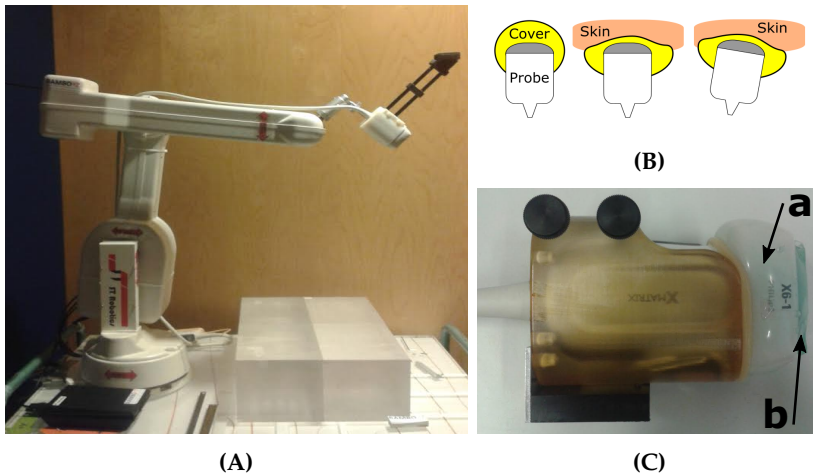
Figure 3.5: Materials necessary during the reproduction of the US probe setup in the experimental setting. (A) Robotic arm which was used to reproduce the calculated US probe setups in the room. (B) Schematic representation of the gel-filled probe cover which was placed on top of the US probe. (C) The X6-1 matrix probe used for image acquisition with (a) indicating the condom placed on top of the US probe and (b) pointing towards the blue US gel with which this condom was partially filled to ensure sufficient acoustic coupling.

3.2.4 Evaluation

In order to evaluate the whole workflow, five of the algorithm-proposed probe setups were reproduced using the robotic arm and US volumes of the phantom were acquired. These setups did not necessarily allow visualization of most of the optional anatomical structures. The best probe setups were located very close together (1 mm or 1 degree apart), and therefore they were hard to distinguish. For this reason, the five setups were randomly chosen from the list with proposed probe setups, provided that the individual setups were at least 2 mm or 2 degrees apart. In addition, three of the picked setups only involved translation of the probe, while the other two setups also included the rotation parameters.

All the US volumes acquired with the proposed US probe setups were then resampled to 1 mm x 1 mm x 1 mm and the prostate, bladder and rectum were segmented by a trained operator using Eclipse. A visual inspection of the acquired US volumes was performed to determine if the prostate as well as adjacent edges of the bladder and rectum were clearly visualized. In addition, the mean absolute distance (MAD) [6] was calculated between the algorithm-predicted positions of the different anatomical structures and the actual positions in the US volumes. The MAD calculation was performed on sub-volumes, consisting of fifteen consecutive slices on which the structures were easily recognizable.

3.3 Results

Fig. 3.6A displays a slice of one of the acquired 3D US phantom volumes with the corresponding segmentations of the anatomical structures, superimposed in green. Visual inspection reveals that this volume fulfills all clinical requirements. Inspection of the other acquired volumes resulted in the same conclusion.

(A)   **(B)**
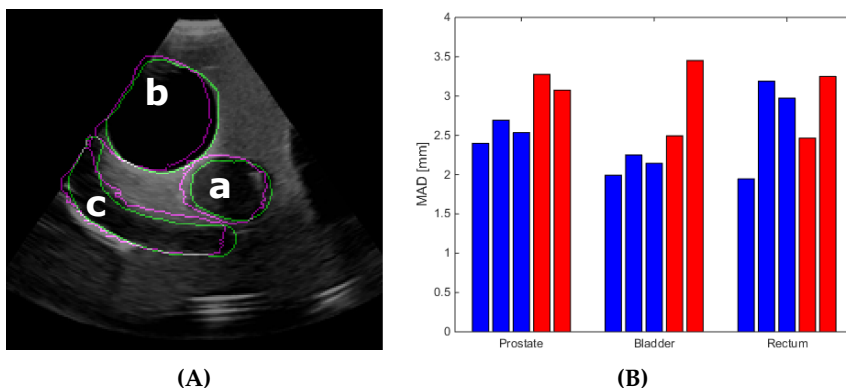
Figure 3.6: (A) Slice of an acquired 3D US volume with the corresponding segmentations of the prostate (a), bladder (b) and rectum (c), superimposed in green. The magenta lines indicate the positions of the same anatomical structures based on the FOV model and the CT scan of the pelvic phantom. (B) Resulting MAD for each anatomical structure displayed per acquired US volume. During the acquisition of the volumes represented in blue, only translation of the US probe was performed, while during acquisition of the red volumes also rotation was included.

The magenta lines in Fig. 3.6A indicate the positions and orientations of the anatomical structures, based on the FOV model and the CT scan of the pelvic phantom, i.e. the algorithm-proposed contours. As described before, the MAD between the green and magenta contours was calculated and these results are displayed in Fig. 3.6B. Each bar in this figure corresponds to one of the picked setups and therefore to one of the acquired US volumes.

The blue bars represent the setups with only translation and the red bars also include rotation of the US probe. The average results over all the acquired volumes resulted in an MAD of 2.8 ± 0.4 mm for the prostate, 2.59 ± 0.6 mm for the bladder and 2.8 ± 0.6 mm for the rectum. The overall MAD average of the setups in which only translation was performed is 2.5 ± 0.4 mm, while the average of the setups in which also rotation was included is 3.0 ± 0.4 mm.

3.4 Discussion

The aim of the proposed algorithm was to provide the operator with a patient-specific US probe setup that would allow visualization of the required structures. In this paper, the probe setup on which the algorithm converged could be a local

minimum of the cost function. However, the intention of the algorithm was to find possible probe setups, not necessarily find the optimal probe setup for the phantom. As the whole scanning area of the phantom (16 cm x 16 cm) was used as a possible location for the probe, brute force calculation was not an option. On the other hand, as the available space for positioning a US probe on a patient will probably be smaller, not much effort was put into the optimization of the process yet. In the end, the aim of the algorithm will be to find an optimal solution for the patient. So in case more optimization turns out to be necessary to achieve this goal, the probe setup selection process might be (partly) changed and an algorithm could be designed.

At this point in time, four degrees of freedom for the US probe were allowed: both translation and rotation along/around the X- and Z-axis. Translation along the Y-axis was not possible, due to the fact that the US probe should be positioned against the skin of the patient. On the other hand, the rotation around the Y-axis should be evaluated for its clinical relevance. The expectation is that the rotation around the Y-axis will not significantly improve the amount of structures that could be visualized. In addition, as the space between the legs of the patient is very limited, rotation around this axis might be impossible in clinical practice. In any case, more research is necessary to confirm this.

In this study, the MAD between the anatomical structure contours as predicted by the algorithm and the actual position of these structures was used as quantitative measure. The MAD is a typical local metric in comparison with, for example, the Dice similarity coefficient [7], [8], which allows global evaluation. Both the location of the algorithm-predicted contours as well as the actual contours (and therefore also the MAD) depend on several factors. The overall MAD measured in this study was 2.7 mm, which corresponds to about 3 voxels. This can be considered a good result, especially taking into account that the MAD depends on the segmentation accuracy.

Another factor that must be taken into account is the reproduction of the calculated probe-setup. As discussed before, five of the proposed setups were reproduced in the room using a robotic arm. The accuracy of this arm in all directions is not reported by the manufacturer, however, the sag with a partly extended arm at a nominal payload is reported [9] to be 1.0 mm. So potentially, the robotic arm did not position the US probe exactly at the position it was instructed to. In addition, not only translations of the US probe were required, but also rotations. If the axes of the probe were not exactly aligned with the axes of rotation of the robotic arm, rotating the probe around these axes might also have induced a small translation. Finally, the gripper of the robotic arm was not intended to hold an US probe. The US probe was fixed to the arm as rigidly as possible, but some small movements with respect to the robotic arm might still have been possible. Hence all the points mentioned above might have affected the exact positioning of the US probe on the phantom, in this way affecting the actual position contours and so the MAD.

To obtain a good overview of the accuracy of the robotic arm and therefore also of the accuracy of the probe setup reproduction at this point in time, an extensive study on the robotic arm itself should be performed. However, as the used robotic arm does not move around smoothly and it tends to sometimes respond unexpectedly, it is not completely suitable to be used for US probe positioning on a patient. So for the reproduction of the setup on a patient another solution should be found. A possible solution could be the use of a high-precision surgery robot or the use of a mechanical arm with encoders on the joints instead of electrical components. These encoders should then be set by the operator. Finally, also options like wearable US probes that can be glued to the skin of the patient instead of held by a robotic arm, could be among the possibilities. More research is necessary to explore the possibilities and define the best solution for this particular purpose.

3.5 Conclusion

We introduced an algorithm that uses anatomical information derived from a simulation CT scan to automatically provide the operator with a patient-specific US probe setup for US based IGRT. Visual inspection and quantitative evaluation show that this algorithm is able to propose probe setups that fulfill all clinical requirements.

The average MAD obtained for all organs is 2.7 mm. As mentioned before, two potential error sources for this remaining MAD error were identified: segmentation mistakes and probe setup errors. In future steps, it is important to investigate and preferably quantify the typical magnitude of the errors made while segmenting US or CT volumes. In addition, another solution for the reproduction of the setup should be found, because the robotic arm used in this study is not completely suitable for US probe positioning on a patient.

References

[1] J. Roeske, J. Forman, C. Mesina, T. He, C. Pelizzari, E. Fontenla, S. Vijayakumar, and G. Chen. "Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy". In: *International Journal of Radiation Oncology Biology Physics* 33.5 (1995), pp. 1321–1329.

[2] M. Zelefsky, M. Kollmeier, B. Cox, A. Fidaleo, D. Sperling, X. Pei, B. Carver, J. Coleman, M. Lovelock, and M. Hunt. "Improved clinical outcomes with high-dose image guided radiotherapy compared with non-IGRT for the treatment of clinically localized prostate cancer". In: *International Journal of Radiation Oncology* Biology* Physics* 84.1 (2012), pp. 125–129.

[3] E. Yorke, P. Keall, and F. Verhaegen. "Anniversary paper: role of medical physicists and the AAPM in improving geometric aspects of treatment accuracy and precision". In: *Medical physics* 35.3 (2008), pp. 828–839.

[4] D. Fontanarosa, S. Van der Meer, J. Bamber, E. Harris, T. O'Shea, and F. Verhaegen. "Review of ultrasound image guidance in external beam radiotherapy: I. Treatment planning and inter-fraction motion management". In: *Physics in Medicine & Biology* 60.3 (2015).

[5] T. O'Shea, J. Bamber, D. Fontanarosa, S. van der Meer, F. Verhaegen, and E. Harris. "Review of ultrasound image guidance in external beam radiotherapy part II: intra-fraction motion management and novel applications". In: *Physics in Medicine & Biology* 61.8 (2016).

[6] G. Gerig, M. Jomier, and M. Chakos. "Valmet: A new validation tool for assessing and improving 3D object segmentation". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2001, pp. 516–523.

[7] T. Sørensen. "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons". In: *Biol. Skr.* 5 (1948), pp. 1–34.

[8] L. Dice. "Measures of the amount of ecologic association between species". In: *Ecology* 26.3 (1945), pp. 297–302.

[9] *ST Robotics - R17 5/6-axis robot arm*. URL: http://www.strobotics.com/articulated-robot-arm.htm.

4 Patient-specific transperineal ultrasound probe setups

Abstract

The use of ultrasound imaging is not widespread in prostate cancer radiotherapy workflows, despite several advantages (e.g. allowing real-time volumetric organ tracking). This can be partially attributed to the need for a trained operator during acquisition and interpretation of the images. We introduce and evaluate an algorithm that can propose a patient-specific transperineal ultrasound probe setup, based on a CT scan and anatomical structure delineations. The use of this setup during the simulation and treatment stage could improve usability of ultrasound imaging for relatively untrained operators (radiotherapists with less than one year experience with ultrasound).

The internal perineum boundaries of three prostate cancer patients were identified based on bone masks extracted from their CT scans. After projection of these boundaries to the skin and exclusion of specific areas, this resulted in a skin area accessible for transperineal ultrasound probe placement in clinical practice. Several possible probe setups on this area were proposed by the algorithm and the optimal setup was automatically selected. In the end, this optimal setup was evaluated based on a comparison with a corresponding transperineal ultrasound volume acquired by a radiation oncologist.

The algorithm-proposed setups allowed visualization of 100% of the clinically required anatomical structures, including the whole prostate and seminal vesicles, as well as the adjacent edges of the bladder and rectum. In addition, these setups allowed visualization of 94% of the anatomical structures, which were also visualized by the physician during the acquisition of an actual ultrasound volume.

Provided that the ultrasound probe setup proposed by the algorithm, is properly reproduced on the patient, it allows visualization of all clinically required structures for image guided radiotherapy purposes. Future work should validate these results on a patient population and optimize the workflow to enable a relatively untrained operator to perform the procedure.

4.1 Introduction

Radiotherapy (RT) is one of the curative treatment options for prostate cancer. It aims at irradiating tumor tissue while sparing organs at risk (e.g. rectum and bladder) as much as possible. A treatment plan is usually prepared based on a computed tomography (CT) scan of the pelvic region, after which different radiation techniques [1] (e.g. volumetric modulated arc therapy [VMAT]) can be used to deliver the prescribed radiation dose to the patient in multiple treatment fractions.

Correct patient positioning prior to each of these fractions is crucial to ensure correct dose delivery. Several solutions are available to assist in this inter-fraction positioning procedure, such as skin marks [2], implanted fiducial markers [3] and cone-beam CT [4] (CBCT). However, these solutions might not fully identify soft tissue deformations, such as prostate shape and position changes due to a different bladder or rectal filling [5]. In addition, during the treatment fraction (intra-fraction) tissue distributions might also change [6]. In both cases, delivery of the initial treatment plan could result in a suboptimal dose deposition in the tumor and additional undesired dose delivery to the organs at risk.

Frequent soft tissue imaging during the course of the treatment (image guided RT, IGRT [7]) could potentially be used to guide dose and increase delivery to the tumor tissue and decrease toxicity. If images would be acquired prior to and during each treatment fraction, patients will have to undergo imaging up to 40 times during a time span of one to two months [8]. Therefore, the use of an image modality that is non-invasive and that also does not involve ionizing radiation is preferred. However, if an image modality provides images with significantly higher quality or images that are better suited for the purpose, this might outweigh the discomfort of an invasive procedure and the potential risk of radiation use. In addition, the image modality of choice should enable real-time volumetric tracking of anatomical structures in the RT environment as well.

Magnetic Resonance (MR) imaging is one of the promising imaging modalities for IGRT. However, to include MR imaging in the RT workflow, the linear accelerators currently available in the hospitals would have to be replaced by a new MRI-Linac system [9] which is a costly procedure. An image modality that does not require the replacement of the available linear accelerators is ultrasound (US) imaging. This image modality was already suggested by Fontanarosa *et al.* [10] and O'Shea *et al.* [11] to be suitable for IGRT, as it also meets the previously described requirements and it is the only clinically available modality that allows real-time volumetric imaging in RT.

Despite the many advantages, the use of US imaging is presently not widespread in RT workflows. This can be attributed to the need for a trained operator who needs to interpret the images during the acquisition, in order to verify if the correct anatomical structures are visualized with sufficient image quality. Other challenges that are associated with US imaging include that a good acoustic

coupling between the probe and patient's skin needs to be established and that structures shielded by air or bone are inaccessible [10], [11]. In addition, the imaging target needs to be within the US imaging depth range and a high inter-operator variability [12], [13] was occasionally reported.

In a previously published proof-of-concept study [14], we introduced the preliminary version of an algorithm that uses anatomical information derived from a CT scan to provide the operator automatically with a US probe setup. The algorithm was evaluated using a pelvic phantom, in order to mimic the acquisition of 4D transperineal US (TPUS) images of prostate cancer patients.

In this study, the phantom algorithm was used as a basis to create an algorithm that can propose a prostate cancer patient-specific TPUS probe setup, based on a simulation CT scan and corresponding delineations of anatomical structures. Subsequently, the proposed setup was compared with the setup used by a radiation oncologist (RTO) during an actual TPUS image acquisition, to assess if the same anatomical structures could be visualized. In addition, an assessment was performed to understand if the clinically required structures could be visualized.

The algorithm-proposed setup is intended to be used during both the simulation and treatment stage. During the simulation stage, the US volume would need to be acquired after the CT scan, as the anatomical information from the CT scan is necessary to calculate the setup. The probe should then be positioned on the body of the patient according to this calculated setup, after which a reference US volume would be acquired.

Then, prior to each treatment fraction, the US probe should again be positioned according to the calculated setup. This would allow the acquisition of a US volume that could be used for inter-fraction motion compensation, by comparing it to the reference US volume acquired at simulation stage. Leaving the probe in place during the radiation treatment, would finally also allow intra-fraction monitoring of the anatomical structures of interest.

4.2 Materials and methods

4.2.1 Patient image acquisition

CT scans of the pelvic region of three patients (Table 4.1) with localized prostate cancer were acquired using a SOMATOM Sensation Open CT scanner (Syngo CT 2006A, Siemens, Erlangen, Germany; voxels: 1 mm x 1 mm x 3 mm). During the scan, the patients were positioned in a supine position with the skin markers aligned to the isocenter lasers of the scanner (Fig. 4.1). The knees and feet of the patients were supported by a Kneefix^{TM} 2 and a Feetfix^{TM} 2 (Combifix^{TM} 2, CIVCO Medical Solutions, Coralville, IA, USA).

Table 4.1: Summary of the available patient data in this study. Three RTOs acquired the US volumes (see second column), while only one performed all the registrations (third column). In column four all the available delineations for each patient are detailed.

Patient	RTO acquisition	RTO registration	Delineations
1	A	C	Prostate, Seminal vesicles, Bladder, Rectum, Anal canal
2	B	C	Prostate, Bladder, Rectum, Anal canal
3	C	C	Prostate, Seminal vesicles, Bladder, Rectum, Anal canal

Right after the CT scan acquisition (within about two minutes), a 3D TPUS volume of each patient was acquired using an X6-1 xMatrix array probe (Philips Healthcare, Bothell, WA, USA) in combination with an EpiQ7 US system (Philips Medical Systems, Andover, MA, USA).

The US probe was attached to a custom-made mechanical arm (item f in Fig. 4.1) and it was positioned by a RTO at the desired location and orientation, based on the live US images provided on the EpiQ7 monitor. Subsequently, the probe setup was fixed using a rotary knob on the arm and the TPUS volume was acquired. During the whole procedure, the patients were instructed to restrict motion as much as possible.
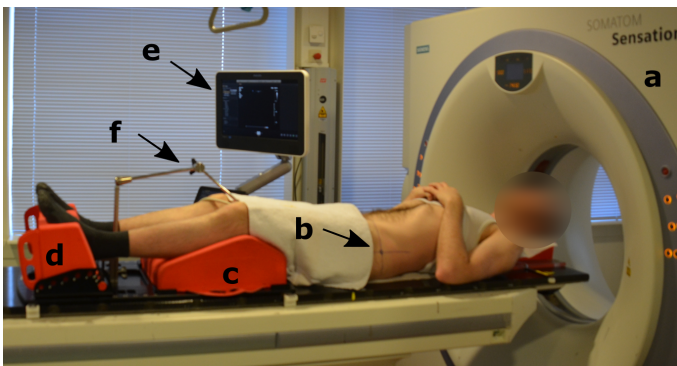
Figure 4.1: Volunteer mimicking the patient setup for CT and US volume acquisition: (a) CT scanner (b) Skin markers on the volunteer's body for isocenter laser alignment. (c) Kneefix^{TM}2 (d) Feetfix^{TM}2 (e) EpiQ7 US system monitor (f) Mechanical arm with rotary knob for US probe fixation.

It has been reported that application of probe pressure during US volume acquisition can result in a displacement of the anatomical structures [15]. In this work, US volumes were only acquired at the simulation stage, so potentially induced displacements would not be reproduced at treatment stage. For this

reason, the CT scan was acquired without the probe in place. In addition, the interference with the current clinical workflow was minimized by acquiring the US volumes after the CT scan.

4.2.2 Image pre-processing

For each patient, the TPUS volume was registered to the corresponding simulation CT scan by one RTO. The Eclipse treatment planning system (Varian Medical Systems, Palo Alto, CA, USA) offers this registration procedure, which comprises an initial manual registration, potentially followed by an automatic point-match registration based on the fiducial markers in the prostate.

Subsequently, the CT table was removed from the scan, using a similar approach as proposed by Zhu *et al.* [16] and the body outline of the patient was identified using a thresholding approach (threshold: -250 HU), as proposed by Huang *et al.* [17] This allowed the removal of image artifacts present outside the patient's body.

Finally, the voxels of the CT scan, corresponding structure contours (Table 4.1) and the US volume were resampled to 1 mm x 1 mm x 1 mm using MeVisLab (Version 2.8, MeVis Medical Solutions AG, Bremen, Germany) to make the subsequent calculations easier. After resampling, the image dimensions of all datasets were (500 x 500 x 507/582/456) voxels. However, not all slices in the cranial-caudal direction contained useful anatomical information. As the processing of these slices would only prolong the calculation time without any additional benefit, the data sets were, after visual inspection, cut to (500 x 500 x 275/350/225) voxels.

4.2.3 Bone mask extraction

Bones and soft tissue have a significantly different acoustic impedance, causing bones to reflect a substantial part of the US wave intensity producing shadows in the US images. For this reason, it was crucial to take the pelvic bones into account during probe-setup calculation. A 3D binary bone mask was extracted from the patient's CT scan using a segmentation tool, which was developed in MATLAB (Version 9.0.0 (R2016a), The Mathworks Inc. Natick, MA, USA). Voxels corresponding to bone tissue have higher Hounsfield Unit (HU) values assigned than soft tissue. Therefore, thresholding (see Appenidx A.1) was used to outline the bones, resulting in a bone mask as shown in Fig. 4.2A.

4.2.4 Localization of the internal perineum boundaries

The perineal area on the skin surface of the patient was identified to determine possible setups of the US probe. According to literature, the definition of the internal perineum is based on several points on the skeleton of the patient. The posterior boundary of the perineum is defined by the tip of the coccyx, the anterior boundary of perineum by the inferior margin of the pubic symphysis and the two lateral boundaries by the ischial tuberosities [18].

A. Localization of the tip of the coccyx

The localization of the tip of the coccyx was performed in three automatic steps. In the first step, the binary bone mask (Fig. 4.2A), was projected on the XZ-plane (Fig. 4.2B) and subsequently, the space between the femurs (Fig. 4.2B) could be identified. After identifying the interface between the top of the inter-femur space and the pelvic bone, local maxima on this interface were found (Fig. 4.2B). Finally, the two outer local maxima were used to remove the femurs and other parts of the skeleton that certainly did not belong to the spine (Fig. 4.2C).
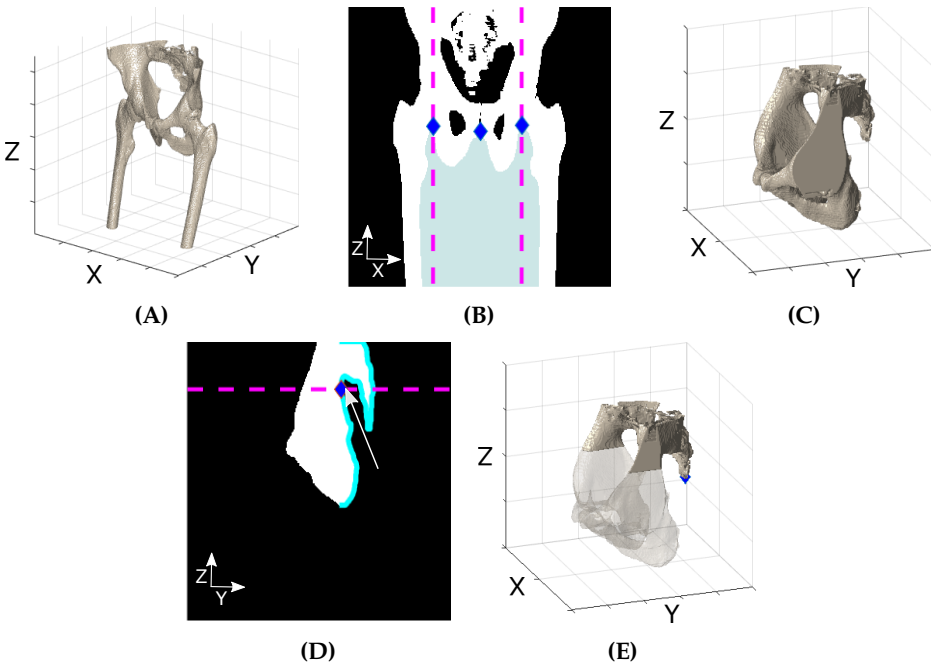
Figure 4.2: Workflow for localization of the tip of the coccyx based on image data of Patient 1. (A) Binary bone mask extracted from the patient's CT scan. (B) Projection of the mask in the XZ-plane with the light blue area highlighting to the inter-femur space, and the blue diamonds indicating the local maxima of the interface between this space and the bone mask. The magenta, dashed lines indicate the cut-off lines. (C) Bone mask without femurs resulting after cut application as displayed in (B). (D) Projection of the remaining bone mask in the YZ-plane with the cyan line referring to the perimeter and the blue diamond (pointed out by the arrow) indicating the global minimum. The magenta, dashed line indicates the cut-off line. (E) Bone mask with a transparent pelvic bone and a non-transparent spine with the blue diamond indicating the tip of the coccyx.

In the second step, the remaining mask was projected onto the YZ-plane and the perimeter on the coccyx side of the mask was automatically constructed (Fig. 4.2D). The global minimum in the negative Y-direction (Fig. 4.2D) was used as a custom cut-off point between the pelvic bone and the spine. Finally, the

connected components were again labeled and an area based threshold value *T1* was automatically calculated using the basic global thresholding method of Gonzalez and Woods [19]. Connected components with an area $\geq$ *T1* were assigned to the pelvic bone, while connected components with an area $<$ *T1* were assigned to the spine, see Fig. 4.2E for the resulting bone mask.

The third and final step, involved the actual localization of the tip of the coccyx. This tip was found by identifying the spine voxel with the smallest Z coordinate (blue diamond in Fig. 4.2E).

B. Localization of the pubic symphysis

The pelvic bone (Fig. 4.2C) was the starting point for the pubic symphysis localization performed in three automatic steps. First, the pelvic bone was projected on the XY-plane and the obturator foramina were identified (Fig. 4.3A). As the pubic symphysis lies between the two obturator foramen, the other areas could be excluded from further evaluation.

(A)   **(B)**   **(C)**

(D)
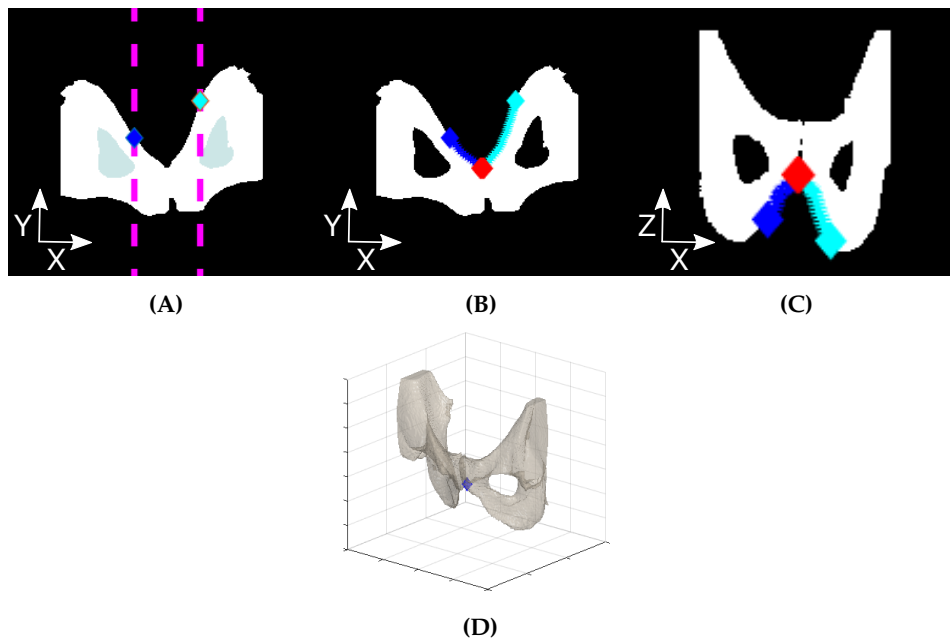
Figure 4.3: Workflow for pubic symphysis localization based on image data of Patient 1. (A) Projection of the pelvic bone mask in XY-plane with the dashed magenta lines representing the edges of the region of interest, the light blue areas being the obturator foramina and the blue and cyan diamond being the voxels with maximum Y-coordinate in the region between the dashed lines. (B) Projection of the mask in XY-plane with the blue and cyan search paths resulting in overlapping red diamonds. (C) Projection of the mask in XZ-plane with the blue and cyan search paths resulting in overlapping red diamonds. (D) 3D representation of the pelvic bone with the location of the pubic symphysis represented by the blue diamond.

C. Localization of the ischial tuberosities

For the localization of the ischial tuberosities, the pelvic bone (Fig. 4.2C) was used in combination with the established cutting lines (dashed magenta lines in Fig. 4.3A). This allowed the division of the pelvic bone in three parts, see Fig. 4.4A. Next, the total pelvic bone mask was projected on the YZ-plane and the pixels with the lowest Z-value were found (Fig. 4.4B).

Subsequently, in both the left and right part of the pelvic bone, the voxels with a Z-value as close as possible, were identified. Finally, from these voxels the most posterior voxels were identified as the perineum boundary points on the ischial tuberosities (blue diamonds in Fig. 4.4C).

(A)   **(B)**   **(C)**
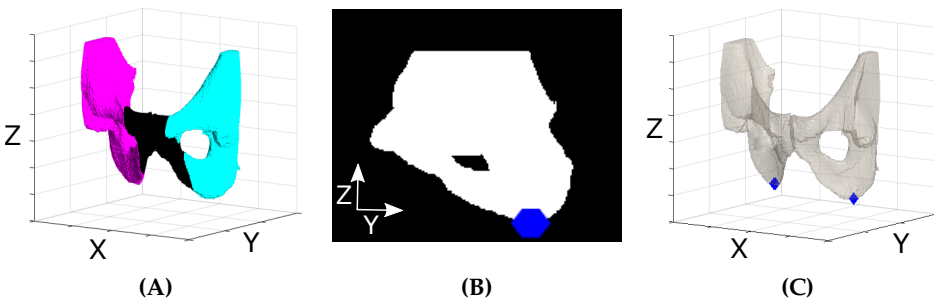
Figure 4.4: Workflow for ischial tuberosities localization based on image data of Patient 1. (A) Binary pelvic bone mask with cyan representing the left part, black representing the middle part and the magenta representing the right part. (B) Pelvic bone projection in the YZ-plane with the blue diamonds indicating the voxels with the lowest Z-coordinate. (C) Pelvic bone mask with the blue diamonds representing the ischial tuberosities.

D. Total perineal skin area

Even though the perineum boundary points were not located in one plane (Fig. 4.5A), they were connected by straight lines resulting in a diamond-shaped perineal area while looking from the caudal direction (Fig. 4.5B). Subsequently, all points in this diamond-shaped area were projected, using parallel lines along the Z-axis, onto the skin between the legs resulting in a situation as displayed in Fig. 4.5C.

4.2.5 Identification of accessible perineal skin

In Fig. 4.5C it can be seen that part of the skin area that was identified as perineal skin was obstructed by the legs, or was covered by part of the scrotum, or too close to the anus. Therefore, first, the blue perineum voxels belonging to the legs instead of to the perineal skin area (Fig. 4.6A) were removed from the perineum projection using a binary leg mask, created using the workflow described in Appendix A.2. Then, the actually accessible skin area on which the probe could be positioned

in clinical practice was determined using two automatic post-processing steps described in the next two sections.
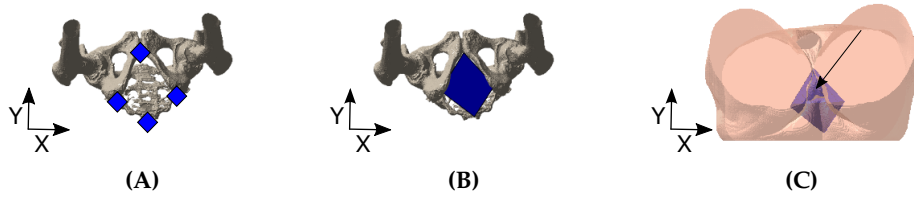
Figure 4.5: Projection of the internal perineum to the skin based on image data of Patient 1. (A) Binary bone mask with the blue diamonds representing the boundary points of the perineum. (B) Binary bone mask with the surface between these diamonds represented by the large blue diamond. (C) Body outline of the patient with the perineal area projected to the skin. The arrow indicates the scrotum located inside the perineal skin area.

A. Scrotum identification and exclusion

It is typical for male patients in the supine position that the scrotum partially covers the defined perineal skin area. To make the skin underneath the scrotum accessible for TPUS probe placement, the patient is usually asked to support the scrotum with his hands. So to define an accessible perineal skin area, the skin area underneath the scrotum was estimated.
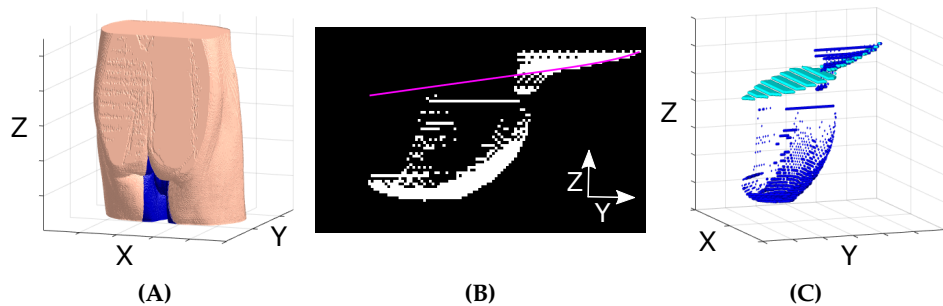
Figure 4.6: Identification of the skin area underneath the scrotum based on image data of Patient 1. (A) Projection of the perineal skin area in blue on the body outline of the patient. (B) Projection of the remaining perineal points in the YZ-plane with the magenta line being a fitted two-term exponential. (C) Representation of the perineal skin area after the leg removal in blue and the perineal skin area after the identification of the skin underneath the scrotum in cyan.

The perineum points remaining after leg exclusion were projected on the YZ-plane (Fig. 4.6B) showing clearly where the perineum/buttocks of the patient are located and where the scrotum starts. The assumption was made that the skin underneath the scrotum follows roughly the same curve as the skin in front of it. For this reason, a two-term exponential was fitted to the first 20-40 data points of this

skin (magenta line in Fig. 4.6B). Subsequently, all the perineal skin points located underneath this curve, thus with a lower Z-coordinate, were mapped to this curve. During this mapping procedure, the X- and Y-coordinate of each perineal skin point were kept constant. The Z-coordinate was changed into a value as proposed by the curve for the corresponding Y-coordinate resulting in an estimated perineal skin (Fig. 4.6C).

B.   Anus exclusion

In clinical practice it is not desirable to position the US probe on top of the anus. For this reason, the exact position of the anus needed to be identified and used to exclude an area from the accessible perineal skin area. For this purpose, the anal canal delineation (Table 4.1) was used, see Fig. 4.7A. The assumption was made that the most caudal axial slice of this canal was the anus. By assessing the boundaries of the anus, an accessible perineum edge could be identified (Fig. 4.7B). Finally, the usable perineal skin area was restricted using this edge (Fig. 4.7C).

(A)   **(B)**   **(C)**
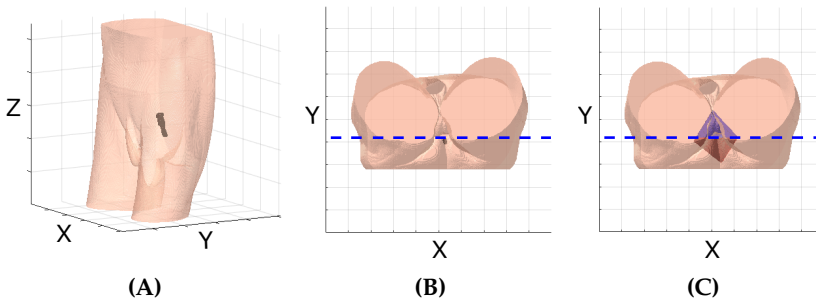
Figure 4.7: Anus exclusion from usable perineal skin area based on image data of Patient 1. (A) Body outline of the patient with the anal canal shown in black. (B) Body outline of the patient in the XY-plane with the anus shown in black and the blue dashed line representing the cut-off line. (C) Body outline in the XY-plane with the red part of the diamond representing the perineal skin area that is not accessible and the blue part indicating the accessible skin area.

C.   Final identification of accessible perineal area

Combining Section 4.2.5.A and Section 4.2.5.B resulted in a perineal skin area that was accessible in clinical practice for TPUS probe positioning (Fig. 4.8A). During US image acquisition, operators tend to apply pressure on the probe to improve the acoustic coupling, which results in significant perineal skin displacement.

To mimic this displacement, the perineal skin area was moved 35 mm parallel to the longitudinal axis of the patient in the cranial direction (Fig. 4.8B). This displacement was estimated under the assumption that the scanning head of the probe was flat and the used US gel layer was infinitely thin. CT scans of prostate cancer patients who had a TPUS probe in place during the scan acquisition were used for this estimation. These scans were available in our research group.
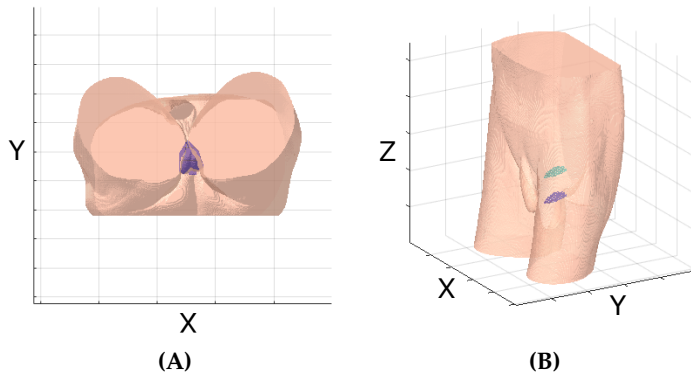
Figure 4.8: Identification of the final accessible perineal area based on image data of Patient 1. (A) Resulting accessible perineal skin area (blue). (B) Original position of the accessible skin area in blue and the shifted skin area to compensate for probe pressure in cyan.

4.2.6 Identify probe setups

In this section the workflow used to determine possible US probe setups will be described. First the possible positions of the US probe on the XY-plane projection of the perineal skin area were identified (Fig. 4.9A). To this end, the virtual probe was positioned on the edges of the accessible perineal skin (see item a in Fig. 4.9A) and translated along the X- and Y-axes using steps of 1 mm until the whole perineum was covered, resulting in N_{XY} possible probe setups.
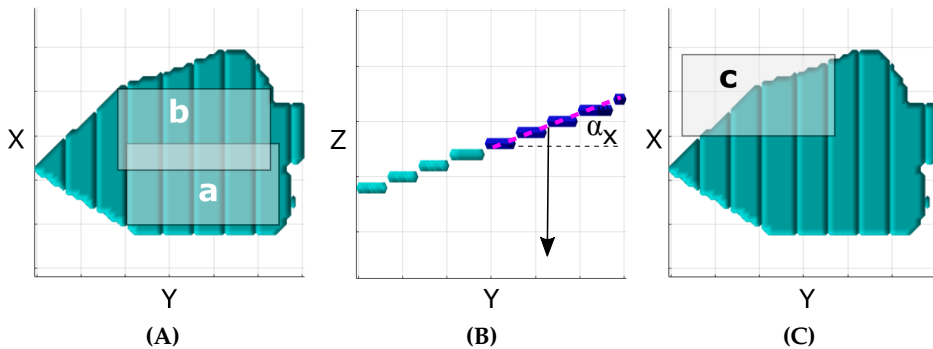
Figure 4.9: Workflow to determine the possible US probe setups based on image data of Patient 1. (A) Projection of the perineal skin area on the XY-plane with a and b indicating possible probe setups. (B) Projection of the skin area on YZ-plane with a possible probe position in Y-direction highlighted in blue. The dashed magenta line corresponds to the first-order polynomial fit. The center of the fit determines the translation in Z-direction (arrow) and α_X identifies the angle of rotation around the X-axis. (C) Projection of the skin area on XY-plane with c indicating a possible setup with insufficient coverage with the perineal skin area.

As the US probe needs to be positioned on the patient's skin, the Z-position of the probe depends on the shape of the perineum. In Fig. 4.9B, a projection of the perineum on the YZ-plane is shown in cyan, with the highest possible probe position along the Y-direction highlighted in blue. A 2D first-order polynomial was fitted to this projection of the probe position (dashed magenta line in Fig. 4.9B), in order to mimic the placement of a flat probe. Subsequently, the required translation along the Z-direction could be determined, based on the Z-coordinate of the center of this polynomial fit (Fig. 4.9B).

As can be seen in Fig. 4.9B, the dashed magenta line representing the probe is positioned under an angle. This implies that in addition to a translation in Z-direction, also a rotation around the X-axis was required, in order to position the virtual probe against the perineum. Calculating the inverse tangent of the slope of the first-order polynomial fit provided the angle α_x over which the probe should be rotated around the X-axis (Fig. 4.9B). Combining translation in Z-direction with rotation around the X-axis resulted in N_Z possible setups and $N_{Setups} = N_{XY} \cdot N_Z$ setups in total.

For all N_{Setups} also the orientation of the probe could be varied by adding rotations around the different axes. For the X-axis, these rotations varied from -15 to +15 degrees in steps of 3 degrees (11 options), while for the Y- and Z-axes these rotations varied from -3 to +3 degrees in steps of 1 degree (7 options). In the end, this resulted in $N_{Total} = N_{XY} \cdot N_Z \cdot 11 \cdot 7 \cdot 7$ setups in total.

By covering the whole perineal skin area with virtual probe setups, also a setup as indicated by item c in Fig. 4.9C could be identified. In this case, a large part of the probe is located outside of the perineal skin area. Positioning the probe according to this setup would involve enforcing skin from, for example, the legs aside. At this point in time, the decision was made to exclude the setups that did not have a 100% coverage with the perineal skin area (in XY-plane) from the possibilities. After exclusion of these setups, N_{Final} possible probe setups could be examined in the next steps.

4.2.7 Pre-processing anatomical structures

As noted before, the aim of this study was to propose a probe setup that would allow the visualization of the same structures that the physician was able to visualize. In this process, the CT contours (Table 4.1) created during the treatment plan design (Fig. 4.10A - 4.10B) could be used, as the assumption was made that the structures did not move between the acquisition of the CT scan and the US volume.

The structures visualized by the physician during the actual acquisition of the US volume were considered the preferred structures (blue in Fig. 4.10C). So, preferably it should be possible to visualize these structures using the setup proposed by the algorithm. The remaining structures were considered extra (yellow in Fig. 4.10C) and the more of these structures could be visualized, the

better. However, this could not be achieved at the expense of the preferred structures.
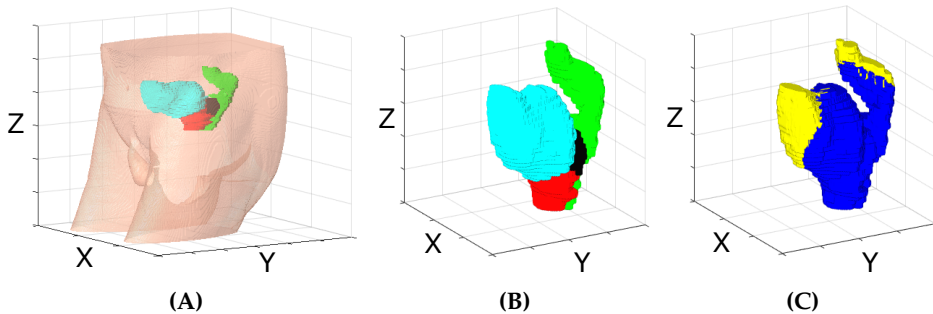
Figure 4.10: Pre-processing of the anatomical structure based on image data of Patient 1 (A) 3D representation of anatomical structures with prostate (red), seminal vesicles (black), rectum (green) and bladder (cyan). (B) Anatomical structures without the body outline. (C) The preferred structures are blue, while the extra structures are yellow.

4.2.8 Ranking of probe setups

The probe setups were ranked based on the presence of the different anatomical structures in the field of view (FOV) of the US probe. During this ranking, we assumed that the probe would be setup on the body of the patient according to the calculated setup and that all acquired US images would be of sufficient quality.

First, a 3D model of the FOV of the US probe was created. A customized pre-set for TPUS imaging of the prostate was available on the EpiQ7 system, which was used for image acquisition during this study. The variables present in this pre-set, such as penetration depth (120 mm), azimuthal angle (70 degrees) and elevation angle (65 degrees), were used as input for the model construction. Since the position of the FOV with respect to the active surface of the probe was known, the location of the FOV inside the body of the patient could be determined for each setup.

Subsequently, the fraction of the preferred and extra structures that would be located inside the FOV of the probe, was calculated based on the contours of these structures for each probe setup. As the number of structure voxels in the FOV should be maximized, the setup with the largest preferred structure fraction was considered optimal. However, if two setups had an identical preferred structure fraction, the setup with the larger fraction of extra structures was prioritized. This provided the initial ranking of the probe setups.

Then, the top 25 setups were selected and possible bone shadowing was checked using the binary bone mask (Section 4.2.3). First, the bones located outside the FOV were removed. Subsequently, lines were drawn from the origin of the FOV towards all boundary voxels of the preferred structures. If the drawn lines intersected with a bone, the corresponding setup potentially could cause

shadowing effects. In this work, preventing shadowing effects while visualizing the extra structures was not part of the requirements. The final result was a ranked list of possible probe setups, which potentially would allow imaging of all preferred structures.

4.2.9 Evaluation of probe setups

As noted before, co-registered US and CT volumes of the patients were available. This data was used to determine if the algorithm-proposed optimal probe setup was similar to the setup used by the physician. To this end, several comparisons were made.

First, the overlap of both US volumes was quantified using the Dice Similarity Coefficient (DSC) [20], [21]. In this case, these volume samples did not contain any anatomical information, but were purely based on the shape and location of the US sectors (see e.g. the cyan and yellow sectors in Fig. 4.11). A DSC=1 indicates perfect overlap, whereas a value of DSC=0 means no overlap at all. Subsequently, the visualization of the different anatomical structures was quantified by calculating the fraction of the structures that was in the FOV of the US probe in the proposed setup and in the setup actually used by the physician.

Clinical requirements provided by an experienced RTO demanded that at least the whole prostate and seminal vesicles, as well as the adjacent edges of the bladder and rectum should be visualized, in order to make the US volume usable for IGRT during prostate cancer treatments. For this reason, two additional sets of anatomical structure voxels were defined. The required structures included all voxels of the prostate and seminal vesicles. In addition, the voxels of the bladder and rectum that were not further than 20 mm away from the borders of the prostate, were included as well. All the remaining structure voxels were considered optional. The preferred structures introduced earlier included all required structures and a substantial part of the optional structures (Table 4.2). Finally, also the fractions of the required and optional structures present in the FOVs were calculated.

4.3 Results

On average 46,088 possible probe setups (range: 14,736 - 86,401) were proposed for the examined patients. The top 25 setups (after ranking) underwent a check for bone shadowing. For Patients 1 and 2 for none of the top 25 setups bone blockage occurred, while for Patient 3 bone blockage also occurred in the TPUS volume acquired by the physician. The identification of the accessible perineal skin area, the ranking and evaluation of the possible probe setups took on average 24 hours on a standard PC (i5 CPU, 2.5 GHz, 4 GB RAM).

During the manual registration of the US and CT volumes performed by one RTO, three out of four fiducial markers were distinguishable in the US volumes of Patients 1 and 2, and two out of four in the volume of Patient 3.

The overlap between the US volume as acquired by the physician and the algorithm-proposed volume was, on average, DSC = 0.76 (Table 4.2). However, the algorithm-proposed setups only allowed visualization of up to 98% of the structure voxels that the physician was able to visualize (preferred structures). The not-visualized voxels were located further than 20 mm away from the prostate edges.

The first setup of each patient was considered the optimal setup and therefore assessed further. In Fig. 4.11, the center slices of the CT scan of Patient 1 in sagittal and coronal direction are shown with the relevant anatomical structures in white. In addition, the outlines of the US sector scanned by the physician (cyan) and as proposed by the algorithm (yellow), are superimposed. As can be seen, the algorithm-proposed setup seems to have a slightly different sector size than the setup used by the physician.

(A)   **(B)**
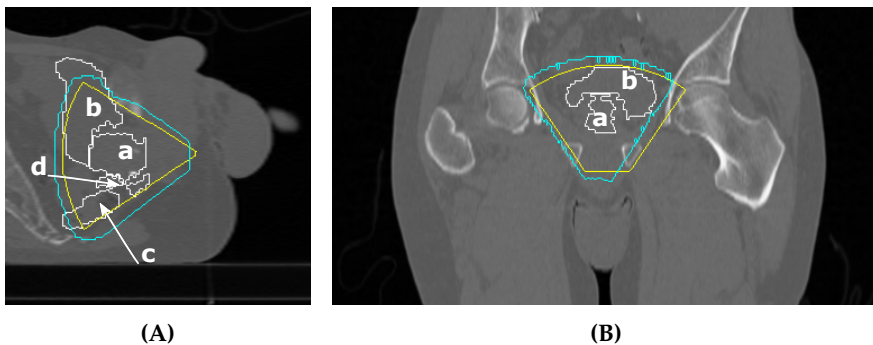
Figure 4.11: Center slices of the CT scan of Patient 1 in both sagittal (A) and coronal direction (B). The US volume as setup by the physician is superimposed in cyan, while the US volume as proposed by the algorithm is superimposed in yellow. In addition, the contours of the relevant anatomical structures are superimposed in white: (a) prostate, (b) bladder, (c) rectum and (d) seminal vesicles.

For both the physician-proposed setup and the algorithm-proposed setup, the visualization using these setups was quantified. These results are detailed in Table 4.2 and in this section the average results over all patients are reported. The algorithm-proposed setup visualized 100% of the required structures and 94% (range: 88% - 98%) of the preferred structures. In addition, the algorithm-proposed setup allowed visualization of 65% of the bladder in comparison to 70% achieved by the physician. The physician was able to visualize 61% of the rectum, while the algorithm-proposed setup was able to visualize 59%. The lower visualization abilities of the bladder and the rectum by the algorithm-proposed

setup automatically results in a lower ability to visualize the optional structures (51% versus 56%).

Table 4.2: Results of the probe setup evaluation. In the first column all the relevant (combinations of) anatomical structures and the DSC are detailed. In the remaining columns, for each patient the visualization of the structures using the physician's probe setup is expressed in percentages of the whole volume of the corresponding structure(s) in the left column. The visualization using the algorithm-proposed optimal setup is detailed in the right column. The asterisk (*) indicates that the fraction of the preferred structures was 100% by definition.

	Patient 1		Patient 2		Patient 3	
	Physician [%]	Algorithm [%]	Phys. [%]	Alg. [%]	Phys. [%]	Alg. [%]
Required structures	100	100	100	100	100	100
Preferred structures	100*	88	100*	98	100*	96
Optional structures	76	57	19	22	72	73
Prostate	100	100	100	100	100	100
Seminal vesicles	100	100	-	-	100	100
Bladder	84	73	26	28	100	95
Rectum	83	65	41	45	58	66
DSC	0.8		0.8		0.7	

4.4 Discussion

In this work, an algorithm was introduced, that proposed a patient-specific TPUS probe setup for imaging of the pelvic region of a prostate cancer patient. The input for this algorithm was a simulation CT scan and contours of several, manually delineated, anatomical structures. The subsequent steps that were performed to determine the optimal setup were fully automatic. As far as we know, the use of a 3D image modality (in this case a CT scan) to determine a US probe setup was never reported earlier by other research groups.

The evaluation of the algorithm-proposed setup was based on manually co-registered CT and US data. During this registration procedure, two or three out of four fiducial markers in the prostate could be identified. Unfortunately, not for all patients these markers are distinguishable in the US volumes. In addition, due to the lack of ground truth, it was not possible to verify how accurate this registration was. In the ideal case, this evaluation should be repeated in the future using datasets which are, for example, co-registered in absolute coordinates using localization hardware, such as a tracking system (e.g. Polaris, NDI, Waterloo, Ontario, Canada).

The binary bone mask used to identify the internal perineum boundaries was extracted from the patient's CT scan using primarily a thresholding approach, which gave a reasonable outline of the bones. Though the algorithm tended to overestimate the edges of the bones in the lower pelvic area. However, overestimation was in this case preferred over underestimation, as the bones potentially might block the US waves. In the upper part of the skeleton, the thresholding algorithm had more problems identifying the bone edges, as the corresponding HU were less distinct with respect to the surrounding structures. This led to some underestimation and has influenced the localization of the tip of the coccyx primarily in the Z-direction. See Appendix A.3 for a figure to support the explanation.

However, the coccyx is one of the vertices of the diamond-shaped perineum and it is located behind the anus, see Fig. 4.7C. After excluding the area behind the anus from the accessible skin area, the perineum is more triangular-shaped. The basis of this triangle seems to be approximately located on the straight line between the ischial tuberosities. Therefore, identifying both ischial tuberosity points and the point on the pubic symphysis could potentially be sufficient to identify the accessible perineal skin. The whole algorithm should be tested on a patient population to evaluate a larger range of anatomical variability, while paying special attention to this hypothesis.

A limitation of this work is that only the potential bone shadowing of the preferred structures was taken into account, even though also shadowing of the extra structures could affect the ranking of the setups. Another limitation is the need for manually delineated structure contours, such as the prostate and the anal canal. In the current situation, the US volume acquisition would have to be postponed until these delineations are available. However, automatic interpretation of the CT images using machine learning strategies could potentially allow the automatic delineation of these structures (e.g. [22]).

As noted before, during the acquisition of the simulation CT scan the scrotum was partially covering the perineal skin. Therefore, the skin area underneath the scrotum had to be estimated to add that part to the skin accessible for TPUS probe placement. In the ideal case, the scrotum would be already lifted during the CT scan acquisition, in order to mimic the procedure necessary for probe placement. However, lifting the scrotum would interfere with the current clinical workflow. For the short term, more research is recommended regarding the ability to distinguish the scrotum skin from the perineal skin on the CT scan.

For this work, the perineal skin displacement was preliminarily estimated based on CT scans with the US probe in place, which were acquired for a different study. However, for these patients no scan was acquired without the probe. A possible experimental setup to allow displacement calculation instead of estimation could be the acquisition of two MR scans of the patient. These scans include one scan without a probe and one scan with an MR-compatible probe (unfortunately the

X6-1 matrix probe was not MR-compatible), to be able to quantitatively assess the displacement of the perineal skin due to the probe placement.

In this study, all setups that did not have a 100% coverage with the underlying perineal skin area were excluded from further evaluation. However, potentially even better setups can be found if some skin displacement is allowed. The discomfort experienced by the patient, while positioning the US probe involving pushing some skin aside from, e.g., the legs, could be examined in the same experimental setup.

Finally, the identification of the accessible perineal skin area and the subsequent ranking and evaluation of the possible probe setups took on average about 24 hours on a standard PC. This long calculation time makes the algorithm not yet suitable for implementation in the simulation stage of the RT workflow. Ideally the US volume would be acquired right after the CT scan, so a 24-hour waiting time to calculate the optimal probe setup is not acceptable. However, the hypothesis is that the use of an optimization method to identify the optimal setup instead of evaluating the usability of all setups using a full-search approach could dramatically decrease the calculation time. For this reason, optimization approaches will be explored in the future.

4.5 Conclusion

The algorithm introduced in this work was able to propose a patient-specific probe setup for each examined patient, which allowed visualization of all anatomical structures that are of interest for US guided RT.

References

[1] B. Vanneste, E. Van Limbergen, E. van Lin, J. van Roermund, and P. Lambin. "Prostate cancer radiation therapy: What do clinicians have to know?" In: *BioMed research international* 2016 (2016).

[2] G. Bentel. *Patient positioning and immobilization in radiation oncology*. McGraw-Hill, 1999.

[3] U. van der Heide, A. Kotte, H. Dehnad, P. Hofman, J. Lagenijk, and M. van Vulpen. "Analysis of fiducial marker-based position verification in the external beam radiotherapy of patients with prostate cancer". In: *Radiotherapy and Oncology* 82.1 (2007), pp. 38–45.

[4] M. Oldham, D. Létourneau, L. Watt, G. Hugo, D. Yan, D. Lockman, L. Kim, P. Chen, A. Martinez, and J. Wong. "Cone-beam-CT guided radiation therapy: A model for on-line application". In: *Radiotherapy and oncology* 75.3 (2005).

[5]   J. Roeske, J. Forman, C. Mesina, T. He, C. Pelizzari, E. Fontenla, S. Vijayakumar, and G. Chen. "Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy". In: *International Journal of Radiation Oncology Biology Physics* 33.5 (1995), pp. 1321–1329.

[6]   H. Ballhausen, M. Li, N. Hegemann, U. Ganswindt, and C. Belka. "Intra-fraction motion of the prostate is a random walk". In: *Physics in medicine and biology* 60.2 (2014), p. 549.

[7]   L. Xing, B. Thorndyke, E. Schreibmann, Y. Yang, T. Li, G. Kim, G. Luxton, and A. Koong. "Overview of image-guided radiation therapy". In: *Medical Dosimetry* 31.2 (2006), pp. 91–112.

[8]   D. Brenner and E. Hall. "Fractionation and protraction for radiotherapy of prostate carcinoma". In: *International Journal of Radiation Oncology Biology Physics* 43.5 (1999), pp. 1095–1101.

[9]   J. Lagendijk, A. Raaymakers B.W.and Raaijmakers, J. Overweg, K. Brown, E. Kerkhof, R. van der Put, B. Hårdemark, M. van Vulpen, and U. van der Heide. "MRI/linac integration". In: *Radiotherapy and Oncology* 86.1 (2008), pp. 25–29.

[10]   D. Fontanarosa, S. Van der Meer, J. Bamber, E. Harris, T. O'Shea, and F. Verhaegen. "Review of ultrasound image guidance in external beam radiotherapy: I. Treatment planning and inter-fraction motion management". In: *Physics in Medicine & Biology* 60.3 (2015).

[11]   T. O'Shea, J. Bamber, D. Fontanarosa, S. van der Meer, F. Verhaegen, and E. Harris. "Review of ultrasound image guidance in external beam radiotherapy part II: intra-fraction motion management and novel applications". In: *Physics in Medicine & Biology* 61.8 (2016).

[12]   C. Enke, K. Ayyangar, C. Saw, W. Zhen, R. Thompson, and N. Raman. "Inter-observer variation in prostate localization utilizing BAT". In: *International Journal of Radiation Oncology• Biology• Physics* 54.2 (2002), p. 269.

[13]   K. Langen, J. Pouliot, C. Anezinos, M. Aubin, A. Gottschalk, I. Hsu, D. Lowther, Y. Liu, K. Shinohara, L. Verhey, V. Weinberg, and M. Roach III. "Evaluation of ultrasound-based prostate localization for image-guided radiotherapy". In: *International Journal of Radiation Oncology*Biology*Physics* 57.3 (2003), pp. 635–644.

[14]   S. Camps, F. Verhaegen, G. P. Fonseca, P. de With, and D. Fontanarosa. "Automatic transperineal ultrasound probe positioning based on CT scan for image guided radiotherapy". In: *Proc. SPIE medical imaging*. Vol. 10135. Orlando, USA, 2017.

[15] M. Li, N. Hegemann, F. Manapov, A. Kolberg, P. Thum, U. Ganswindt, C. Belka, and H. Ballhausen. "Prefraction displacement and intrafraction drift of the prostate due to perineal ultrasound probe pressure". In: *Strahlentherapie und Onkologie* (2017), pp. 1–7.

[16] Y. Zhu, S. Cochoff, and R. Sukalac. "Automatic Patient Table Removal in CT Images". In: *Journal of Digital Imaging* 25.4 (2012), pp. 480–485.

[17] S. Huang, J. Jia, R. Cao, G. Li, M. Cheng, and Y. Wu. "Automatic segmentation of the body and the spinal canal in CT images based on a priori information". In: *2011 5th International Conference on Bioinformatics and Biomedical Engineering*. May 2011, pp. 1–4.

[18] V. Singh. *Textbook of anatomy abdomen and lower limb*. Vol. 2. Elsevier Health Sciences, 2014.

[19] R. Gonzalez and R. Woods. "Digital image processing". In: *Pearson Education* (2002).

[20] L. Dice. "Measures of the amount of ecologic association between species". In: *Ecology* 26.3 (1945), pp. 297–302.

[21] T. Sørensen. "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons". In: *Biol. Skr.* 5 (1948), pp. 1–34.

[22] W. Li, S. Liao, Q. Feng, W. Chen, and D. Shen. "Learning image context for segmentation of the prostate in CT-guided radiotherapy". In: *Physics in medicine and biology* 57.5 (2012), p. 1283.

Chapter 5

US image quality assessment using deep learning

Abstract

Ultrasound guidance is not widespread in prostate cancer radiotherapy workflows. This can be partially attributed to the need for image interpretation by a trained operator during ultrasound image acquisition. In this work, a one-class regressor, based on DenseNet and Gaussian processes, was implemented to automatically assess the quality of transperineal ultrasound images of the male pelvic region. The implemented deep learning approach achieved a scoring accuracy of 94%, a specificity of 95% and a sensitivity of 92% with respect to the majority vote of three experts, which was comparable with the results of these experts. This is the first step towards a fully automatic workflow, which could potentially remove the need for ultrasound image interpretation and make real-time volumetric organ tracking in the RT environment using ultrasound more appealing.

5.1 Introduction

One of the curative treatment modalities for prostate cancer is radiotherapy (RT). This modality aims to irradiate tumor tissue in the prostate (sometimes including the seminal vesicles), while sparing the surrounding organs at risk (e.g. bladder and rectum) as much as possible. The radiation dose is typically delivered to the patient in multiple treatment fractions, in accordance with a treatment plan designed based on a computed tomography (CT) scan.

It has been shown that shape and position of the prostate might differ between treatment fractions (inter-fraction), due to changes in bladder and/or rectal filling [1]. Also during a treatment fraction (intra-fraction), the tissue distributions might change [2]. If the original treatment plan was delivered on the changed tissue configuration, this could result in a suboptimal dose deposition in the tumor and the organs at risk could receive extra undesired dose [3].

For this reason, several solutions have been proposed to identify the position and shape differences of the anatomical structures during the treatment course with respect to the treatment plan. This information can be used to potentially improve the precision of dose delivery. Most of the proposed solutions require frequent imaging during the course of the RT treatment (image guided RT, IGRT) with or without implanted fiducial markers [4] using X-ray, magnetic resonance imaging (MRI) [5], [6], or ultrasound (US) imaging [7]–[9].

In this work, we focused on the use of US imaging for intra-fraction guidance during RT. US imaging allows real-time volumetric organ tracking in the RT environment and, in addition, it is cost-effective and harmless for the patient. Despite these advantages, its use is not yet widespread. This can be partly attributed to the need for a trained operator during manual image acquisition to verify if the correct anatomical structures are visualized with sufficient quality.

To allow for intra-fraction monitoring of anatomical structures, the operator needs to position the US probe prior to treatment fraction commencement. As the operator cannot stay in the treatment room during radiation delivery, the probe would need to be fixed using either a mechanical or a robotic arm. During the treatment fraction, small motion of the patient or changes in anatomical structures can compromise image quality due to, for example, a loss of acoustic coupling and/or a sudden appearance of shadowing artifacts. The operator would therefore need to be present in the control room to promptly identify this quality loss and, if necessary, take appropriate action.

The aim of this study was to develop a prototype deep learning algorithm to automatically score 2D US images out of a 3D volume of the male pelvic region based on their quality or, in other words, on their usability during the US guided RT (USgRT) workflow. In particular, we developed a novel one-class regressor, based on DenseNet [10] and Gaussian processes (GPs) [11]. This is the first step towards a fully automated workflow that would eventually remove the need for

a trained operator and therefore potentially make the use of US imaging more appealing for hospitals.

Machine learning has been used earlier in the assessment of US image quality, primarily in the obstetrics field (e.g. [12], [13]). In these studies, the assessment was based on initial segmentations, or on the presence of specific anatomical structures in the image. In the work of Schwaab *et al.* [14] the quality of 3D US images of the breast was automatically assessed; however, this work made use of handcrafted features, such as, the total 2D physical area of the breast. Instead, we aimed to perform the quality assessment using solely automatically learned deep features from the image without relying on any initial segmentations or specific anatomical structure detection.

5.2 Materials and methods

5.2.1 Image data acquisition

In this work, datasets from three different studies conducted after local institutional review board or medical ethics committee approval at the MAASTRO Clinic (Maastricht, the Netherlands) were combined (Table 5.1) The 36 male subjects were either healthy volunteers or patients with localized prostate cancer and all signed an informed consent.

Table 5.1: Summary of the available datasets in this study in total comprising 11,148 TPUS volumes from 36 male subjects.

Study	Subject type	# subjects	Age mean [range]	Total # volumes
Study 1	Volunteers	6	35 (range: 26 - 52)	840
Study 2	Patients	21	74 (range: 58 - 85)	1,269
Study 3	Volunteers	9	51 (range: 31 - 73)	9,039
Total	-	36	-	11,148

For each subject, several 3D and 4D transperineal US (TPUS) volumes of the pelvic region were acquired using an X6-1 xMatrix array probe (Philips Healthcare, Bothell, WA, USA) and an EpiQ7 US system (Philips Medical Systems, Andover, MA, USA). The datasets show a significant variability in image characteristics due to, for example:

1. different volume dimensions and voxel sizes, due to a requirement to achieve an acceptable frame rate in the 4D sequences;

2. varying body composition, age and medical history of the subjects;

3. possible anatomical structure displacements, which were artificially introduced by instructing the subjects to consciously contract muscles in the pelvic area or to cough;

4. the exact settings on the US system such as imaging depth and focus, which could vary between the different studies and between the different subjects (this also affected the voxel sizes and dimensions of the volumes);

5. the involvement of four radiation oncologists in the acquisition of the volumes, each of them with their own approach to TPUS image acquisition.

5.2.2 Initial image data pre-processing

Three initial pre-processing steps were necessary to prepare the datasets for deep learning algorithm processing. These steps were all performed using MATLAB (Version 9.3.0 (R2017b), The Mathworks Inc. Natick, MA, USA). First, the volumes were resampled to identical voxel sizes, which allowed easy volume comparisons and batch processing of the data in the next steps. Second, the TPUS volumes were sliced to 2D images along the sagittal direction, as this was the direction with the highest resolution. Visual inspection of these images then revealed that the anatomical structures of interest were most often located at the center of each volume. For this reason, only the central 16 sagittal 2D images from each volume were selected for further processing, which also reduced the total computational cost. Then all 2D images were symmetrically padded with black pixels to ensure that all images had the same dimensions as the largest 2D image (namely 216x180 pixels) in the entire dataset. Finally, a fixed region of interest was defined by automatically cropping the images, while preserving the crucial information of all anatomical structures. This resulted in 178,368 2D TPUS images overall composed of 116x100 pixels, originating from 11,148 TPUS volumes.

5.2.3 2D US image classification

The crucial anatomical structures for prostate RT treatments are: prostate, seminal vesicles, bladder and rectum. Prostate is the target of the treatment and should therefore be always completely visible on an acquired US volume. In the ideal case, also the edges of the bladder and rectum adjacent to the prostate should be visible to potentially spare these organs at risk from excessive radiation exposure. Since it was not possible to identify the seminal vesicles with sufficient certainty on the acquired US volumes, these were not evaluated in this study.

Based on the above-mentioned criteria, three image categories were defined which are detailed in Table 5.2. An example of each category is displayed in Fig. 5.1. Category 1 involves images that have insufficient quality to be used clinically for USgRT. The quality of Category 2 and 3 images was considered sufficient as the target was visualized and could potentially be tracked.

Table 5.2: Definition of three image criteria used to classify 2D TPUS images based on their quality.

Category	Critera
Category 1	Prostate could not be identified
Category 2	Prostate alone or in combination with either a part of the bladder or the rectum could be identified
Category 3	Prostate could be identified, as well as a part of the bladder and the rectum

In order to provide the deep learning algorithm with labeled training, validation and test samples, a subset of the available 2D TPUS images was manually and independently scored by four members of our research team, as the experts involved in this study had very limited time available. The central 16 2D images (see Section 5.2.2) of each volume were presented to each team member. They could then scroll through the images of each volume and assign a score between 1 and 3, corresponding to Categories 1 to 3, respectively, to each image.

Some of the 2D images were horizontally flipped, due to the fact that the probe was sometimes held upside down. This resulted in a flipped anatomical structure configuration. During the scoring process, the orientation of these images was manually corrected, to ensure that the bladder was located on the left side and the rectum on the right. The team members were instructed to only assign a score to an image if they were highly confident, so it was also possible to leave images unscored. Following this procedure, 1000 randomly selected volumes were scored by each team member.

(A)   (B)   (C)
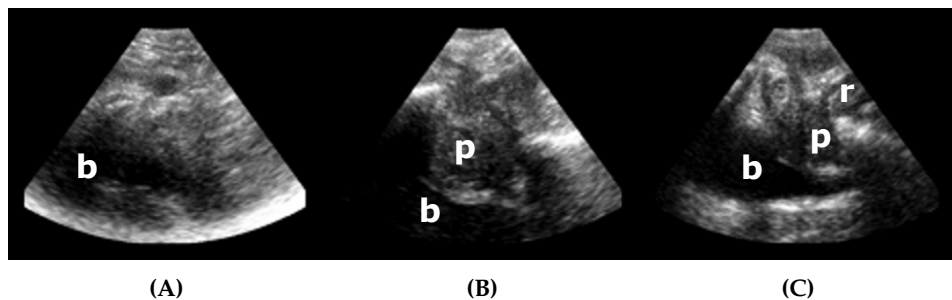
Figure 5.1: Example 2D TPUS image of each quality category. (A) Category 1 with only bladder (b) identifiable. (B) Category 2 with bladder (b) and prostate (p). (C) Category 3 with bladder (b), prostate (p) and rectum (r).

The images that received a consistent score from at least three out of four team members were included in a database (*Database_NonBinary*) with the majority vote of the scores given by the team members assumed to be their ground-truth annotations. Subsequently, the scores of each team member were binarized, with binary score 1 = 0 (poor quality) and binary score 2 or 3 = 1 (good quality). Then, the same procedure of including images in the database that at least three out of four team members scored consistently was followed, resulting in a binary database (*Database_Binary*).

The research team evaluated overall 16,000 2D TPUS images distributed over 1,000 volumes. In total 13,463 of these images (from 34 out of 36 subjects) received a consistent score from three out of four team members and were therefore included in *Database_Binary*. Subsequently, the data were split into training (60% = 20 subjects), validation (20% = 7 subjects) and test (20% = 7 subjects) sets. For each subject, the number of classified images varied. In addition, not for all subjects images of all categories (Category 1-3) were available. For this reason, the split was not just done randomly, but performed using an optimization approach based on simulated annealing [15].

As mentioned in the previous section, the quality of Category 2 and 3 images was considered sufficient for use in clinical practice. Therefore, *Database_Binary* was used to train and test the algorithm; however, the subject split into training, validation and test sets was performed based on *Database_NonBinary*. This was done to ensure a balance between good (Category 2) and very good (Category 3) images in the positive binary group.

First, the data were split into a test and train set by randomly assigning the subjects to one of the groups, while not exceeding the defined sizes of each group. Subsequently, in each iteration, a random subject from the test set was swapped with a random subject from the training set. The aim was to obtain similar ratios between the number of images of a certain category (1-3) in each group (test or train) with respect to the total size of that group. So, for example, if 20% of the training images were from Category 1, also about 20% of the test images should be from Category 1. In total 1000 iterations were executed, in which more weight was assigned to the ratios of Category 2 and 3 images. The ratios of the Category 2 and 3 images were more important, due to the fact that a one-class approach was implemented. This is explained further in Section 5.2.5. The same process was repeated to extract the validation set from the training set. In the end, this resulted in a distribution of poor-quality images (binary score 0) and good-quality images (binary score 1) over the train, validation and test set, as shown in Fig. 5.2A. In Fig. 5.2B the distributions of the binary score 0 and 1 images per subject and per group are detailed. In the remainder of this chapter, this subject distribution will be referred to as *Distr0*.
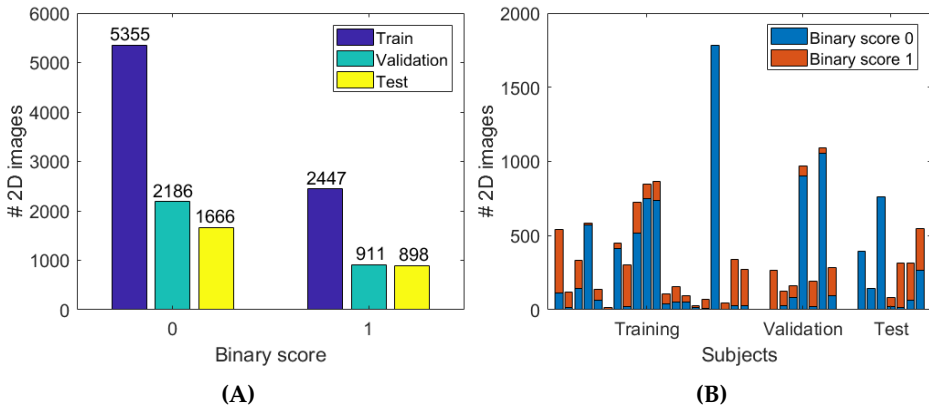
Figure 5.2: Distribution of the image data in *Distr0* per binary score (A) and per subject (B) in the training, validation and test set.

To allow for cross-validation of the hyper-parameters of the deep learning algorithm, nine additional subject distributions were created (*Distr1 - Distr9*). These distributions were also created using the approach based on simulated annealing, as described above. However, since the hyper-parameters were optimized based only on the validation set of *Distr0*, it was not necessary to perform the second step in which the training set is again split into a training set and a validation set. In addition, the distributions were chosen in such a way that each of the 34 subjects appeared at least once in the test set of a distribution. Fig. 5.3 shows the number of test and training images in each distribution (including *Distr0*) and the subjects indicated by their numbers appearing in the test set are detailed in Table 5.3.
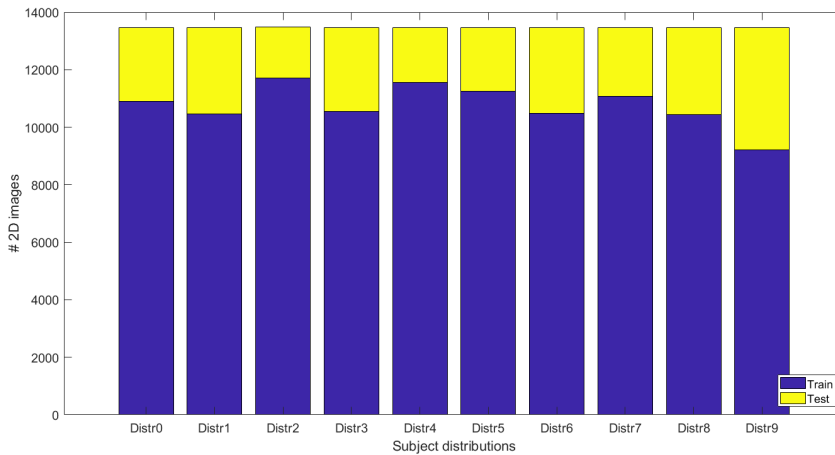
Figure 5.3: Number of training (purple) and test (yellow) images per subject distribution.

Table 5.3: Subjects in the test set of each distribution.

	Test subject numbers						
Distr0	4	8	12	19	22	32	33
Distr1	7	9	13	15	18	25	31
Distr2	5	13	17	24	25	26	32
Distr3	7	13	14	22	23	28	33
Distr4	1	2	5	18	20	24	30
Distr5	3	4	5	12	20	29	34
Distr6	2	15	16	17	26	28	33
Distr7	5	11	21	22	26	29	31
Distr8	4	5	12	13	17	22	27
Distr9	10	12	14	17	19	23	29

5.2.4 Quality score validation

Quality score validation was performed by an accredited medical sonographer and by two of the radiation oncologists involved in the acquisition of the images. These experts were each presented with the same 300 2D TPUS images, which were randomly selected from the test set of *Distr0*, and asked to score these images between 1 and 3. The inter-expert agreement, the test data agreement and the performance of the algorithm were then compared to the majority vote of the experts using Fleiss' kappa [16], accuracy, sensitivity and specificity metrics.

5.2.5 Deep learning algorithm selection

Several different aspects of the image acquisition procedure as well as the subject body composition affect US image quality. Quality deterioration can be caused by many factors including insufficient acoustic coupling between the US probe and the skin, bones causing shadowing artifacts on critical anatomical structures and insufficient penetration due to (fat) tissue distributions. This makes describing features for classification challenging.

For this reason, we approached this problem as a one-class classification (OCC) problem. This approach involves the definition of a single class that should contain all images with "good" (according to clinical requirements) quality, while considering the images with "poor" (according to clinical requirements) quality as outliers. One-class support vector machines (OCSVM) can construct a hyper-sphere with a minimum radius, which contains all positive data points in the multi-dimensional feature space [17]. However, even though this technique is widely used, it does not perform well on noisy data [18].

In this work, the use of Gaussian processes (GPs) instead of conventional SVM was explored for OCC of US image quality. In line with Kemmler *et al.* [11], GPs were used for regression acting as a one-class classifier. In contrast to SVMs, GPs deliver probabilistic predictions and are able to automatically learn regularization

and kernel parameters as well as feature importance. However, GPs lack characterization power for complex data [19]. For this reason, a combination of two techniques was considered: a convolutional neural network (CNN) was used as an autonomous feature descriptor. Then its output was supplied to the GP for OCC.

5.2.6 Architecture and implementation

In this work, DenseNet [10] was used for feature description. This CNN provides a robust architecture which reduces the chance to over-fit and for vanishing gradients, while giving state-of-the-art results on fundamental datasets, like ImageNet [10]. A wide-sense variant with 2 dense blocks, 18 layers per block and a growth rate k of 12 (see Table 5.4) was used and no bottleneck layers were included. Prior to the first dense block, a convolutional operation with a 7x7 filter was performed, followed by a max pooling operation. Finally, the last fully connected layer was removed and replaced by a GP regressor (see Fig. 5.4).
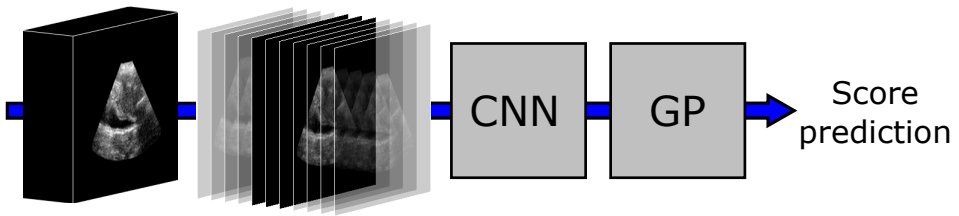
Figure 5.4: Deep learning approach. Selection of 2D images from the center of a TPUS volume, followed by processing using a CNN and a GP, resulting in a quality score prediction.

This regressor was implemented using GPflow [20]. A major advantage of GPflow is that it supports sparse GPs [21], which reduces computation time and memory usage (one of the main drawbacks of GPs). The regressor used a radial basis kernel function (RBF) with an initial variance of 0.1 to fit the data (see Table 5.4). The number of points used during the GP calculations was 150, which was 75% of the outputs from the CNN. As the GPflow library is built on TensorFlow [22], the DenseNet was also implemented in TensorFlow to make end-to-end training possible.

Prior to providing the deep learning algorithm with the image datasets, two final processing steps needed to be performed on the data. First, all pixel values were normalized by setting the total mean to zero and the standard deviation to one, to ensure that the training backpropagation algorithm of the CNN would work efficiently. Second, the training data was randomly permuted and then split in mini-batches to ensure subject balance in the mini-batches.

Table 5.4: Algorithm parameters per implementation step with the asterisks indicating the optimized hyper-parameters.

		CNN + GP (one class)
	Parameter	**Value**
DenseNet	Number of blocks*	2
	Number of layers*	18
	Growth rate k*	12
	Outputs*	200
GPflow	Model	Sparse GP Regression (SGPR)
	Kernel	Radial Basis Function (RBF)
	Initial kernel variance*	0.1
	Inducing points*	150
Training	Batch size*	200
	Epochs*	75
	Optimizer	Adam
	Learning rate*	1e-8
	Drop-out rate*	0.05

All training and testing was performed on a Linux Cluster with an NVIDIA Tesla K40 GPU (NVIDIA, Santa Clara, CA, USA). During training (200 images per batch), the one-class classifier algorithm was only provided with images with good quality (binary score 1). The optimization was done using the Adam optimizer [23] and a learning rate of 1e-8 (see Table 5.4). After the deep learning hyper-parameters were optimized (indicated with an asterisk in Table 5.4) using the validation set of *Distr0*, the training and validation sets were combined into the final training set of *Distr0*.

5.2.7 Comparison with other deep learning algorithms

In addition to the one-class approach in which a CNN was combined with GPs, two additional deep learning approaches were implemented for comparison purposes. The first approach also consists of a CNN in combination with GPs, but instead of only training on the positive data (one-class), the network was trained on both the negative and positive classes. This was possible in this study as sufficient negative class data was available. The parameters used in this implementation are detailed in Table B.1 of Appendix B and again the asterisks indicate the optimized parameters based on the validation set of *Distr0*.

The second deep learning approach consisted of a DenseNet implementation with a softmax classifier attached to it, as described in the paper by Huang *et al.* [10]. With this approach a binary classification was performed, again with the hyper-parameters optimized using the validation set of *Distr0* (see Table B.1 of Appendix B).

5.3 Results

In Table 5.5 the cross-validation results are reported as per the implemented deep learning approach. These results are based on the full test sets and not just on the expert validated test subset. As the training of the CNN + Softmax based on *Distr9* with the corresponding hyper-parameters ran out of GPU memory, only the results of *Distr0 - Distr8* were averaged. The CNN + Softmax approach had the worst accuracy and sensitivity in comparison with the CNN + GP approaches. Both CNN + GP approaches performed comparably and the hyper-parameters seem to be able to generalize.

Table 5.5: Cross-validation results per deep learning approach reporting the mean and σ of the accuracy, sensitivity and specificity calculated over *Distr0 - Distr8*.

	Accuracy [mean $\pm \sigma$]	Sensitivity [mean $\pm \sigma$]	Specificity [mean $\pm \sigma$]
CNN + GP (One class)	92 $\pm$ 1.7 %	89 $\pm$ 5.8 %	93 $\pm$ 2.6 %
CNN + GP (Two class)	92 $\pm$ 1.5 %	90 $\pm$ 6.1 %	93 $\pm$ 2.2 %
CNN + Softmax	90 $\pm$ 1.3 %	79 $\pm$ 7.6 %	95 $\pm$ 2.5 %

The Fleiss' kappa among the three experts, calculated based on the subset of 300 images randomly picked from the test set of *Distr0*, was equal to 0.80. The kappa among the three experts and the test subset was equal to 0.79. The accuracy, sensitivity and specificity results with respect to the majority vote of the experts are detailed in Table 5.6. The accuracy of the test subset with respect to the majority vote was 91%, while the accuracy from the experts ranged within 92% - 97%. The test subset had the lowest sensitivity (80% compared to 90% - 99%), but a specificity of 99%.

Table 5.6: Accuracy, sensitivity and specificity (1-false positive rate) results for the algorithms, test subset and three experts calculated with respect to the majority vote of the three experts.

	GP One class	GP Two class	Softmax	Test subset	Expert 1	Expert 2	Expert 3
Accuracy	94%	93%	92%	91%	96%	97%	92%
Sensitivity	92%	96%	87%	80%	99%	95%	90%
Specificity	95%	91%	96%	99%	93%	99%	94%

All algorithms achieved an accuracy, which was equal to or higher than the accuracy of the worst expert (Expert 3) with respect to the majority votes of the experts. The CNN + GP approaches achieved a better accuracy than the CNN + Softmax, while the one-class approach achieved a better accuracy and specificity in comparison with the two-class approach.

In Fig. 5.5 the receiver operating characteristics (ROC) curves of the one-class CNN + GP approach is plotted, again with respect to the majority vote of the experts.
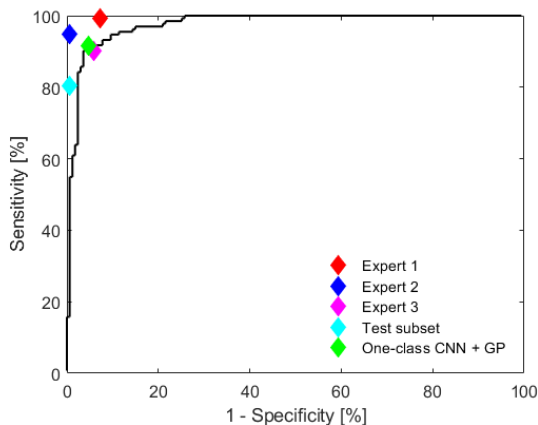
Figure 5.5: ROC-curve of the one-class CNN + GP algorithm with respect to the majority vote of the experts, where the red, blue and cyan diamonds are indicating the performance of the experts, the cyan diamond gives the performance of the test subset and the green diamond gives the performance of the algorithm.

The green diamond indicates the highest accuracy of the algorithm (94%), which corresponded to a sensitivity of 92% and a specificity of 95% (see Table 5.6). The red, blue and magenta diamonds indicate the performance of the experts, while the cyan diamond corresponds to the test subset. The Fleiss' kappa for the experts and the algorithm was equal to 0.82.

5.4 Discussion

In this work, a one-class deep learning approach was proposed that could be used to automatically assess the quality of TPUS images of the male pelvic region. For comparison purposes, two additional approaches were also implemented. The CNN + Softmax was not able to train on *Distr9* as it ran out of memory. This can potentially be explained by the size of the test set of *Distr9* and by the network depth of the CNN + Softmax network in comparison with the network depth of the CNN + GP approaches. In addition, both the cross-validation results and the validation by experts showed that better accuracy and sensitivity results was achieved using the CNN + GP approaches. It has to be noted that during the optimization of the hyper-parameters, the hyper-parameter space has only been explored up to a certain extend. However, the current results seem to justify the use of GPs instead of softmax for classification.

The cross-validation results show a comparable performance of both the one-class and binary classification using CNN + GP, while the expert validation shows a better performance of the one-class approach (accuracy and specificity). So, the one-class regressor seems to be able to identify the not-usable 2D TPUS images well, even though it was only trained on the usable images (i.e. images belonging to Categories 2 and 3, which consists of 32% of the training images). This is a very promising result in cases where there is a lack of available not-usable images or where it is difficult to capture the whole range of appearances of not-usable images for training purposes.

All the algorithms were trained based on a subset of a larger database and the labels used for training were generated by a small research team. The team members were asked to only assign a score when they were highly confident of their results. In addition, only images to which at least three out of four team members assigned a consistent score were included in the database. This was done to partly eliminate the inter-user variability from the database. In the ideal case, the labels would be generated by experts, but this was not feasible due to time constraints. However, the kappa values (0.80 vs. 0.79) showed a good agreement between the scores of the team-members and the experts. This agreement also shows that, even if the four team members were not fully independent, the resulting database was in agreement with the experts. The accuracy of the test subset was 91%, which is lower than the accuracy of the experts (range: 92% - 97%), but still comparable.

The initial aim was to achieve an accuracy equal to the performance of the worst performing expert (92%). In Fig. 5.5 it can be observed that the algorithm is able to achieve a sensitivity and specificity that are comparable with the experts, which resulted in an overall accuracy of 94%. Calculating the Fleiss' kappa of the experts and the algorithm resulted in 0.82, which seems to imply that there is almost perfect agreement between the experts and the algorithm (according to the interpretation of Fleiss' kappa from Landis and Koch [24]). The current performance evaluation was performed with a subset of the test set, due to limited availability of the experts. In future research, this subset will be expanded and the algorithm parameters will be optimized further in order to achieve a 96% accuracy goal, which is the performance of the second to best expert.

The scores assigned to the 2D images were binarized, as currently the quality of Category 1 was considered insufficient for use in clinical practice, while the quality of Categories 2 and 3 were considered sufficient. In Category 2 images, none or just one of the organs at risk (bladder and rectum) is visualized. As the organs at risk should be spared from radiation as much as possible, in the future not only the position of the prostate should be monitored, but also the position of these organs. This would introduce the need to also make a distinction between Category 2 and Category 3 images. In addition, a single poor-quality 2D image does not necessarily imply that the whole volume is not able to provide useful clinical

information. Therefore, the next steps should move towards the interpretation of a whole volume, for example, using recurrent neural networks which can take into account inter-slice context (e.g. [25]).

The potential of the database that was available in this work has not been fully exploited, as only 16,000 2D images of the 178,368 images were examined by the team, resulting in 13,463 images with labels. Potentially, the performance of the algorithm can be improved by using more images for training. Also, the orientation of the images that had a flipped anatomical structure configuration were manually corrected during the scoring process. However, during the actual image acquisition the probe might be held upside down as well, so the algorithm should be robust for any image orientation changes. This robustness will also be examined in future research.

The focus in this work was on the use of US imaging during a prostate cancer RT workflows. However, a similar approach could be adapted for use in other medical procedures in which US imaging may be beneficial for anatomical localization, but where it is not yet feasible and/or desirable to have a trained operator present at the time. These procedures could be, for example, USgRT workflows of other cancer sites (e.g. liver, bladder of cervical cancer) or US guided surgeries.

5.5 Conclusion

The purpose of this work was to propose a deep learning approach that could automatically assess the quality of TPUS images of the male pelvic region. This could potentially remove the need for quality interpretation by a trained operator. The performance of the implemented one-class GP regressor was compared with three experts and the results showed that the algorithm is comparable with these experts in a binary scoring scenario. Future work will involve exploring the non-binary scoring scenario, including adding additional data into the database and assessing the overall quality of the TPUS volume instead of judging individual 2D images.

References

[1] J. Roeske, J. Forman, C. Mesina, T. He, C. Pelizzari, E. Fontenla, S. Vijayakumar, and G. Chen. "Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy". In: *International Journal of Radiation Oncology Biology Physics* 33.5 (1995), pp. 1321–1329.

[2] H. Ballhausen, M. Li, N. Hegemann, U. Ganswindt, and C. Belka. "Intra-fraction motion of the prostate is a random walk". In: *Physics in medicine and biology* 60.2 (2014), p. 549.

[3]   D. Fraser, Y. Chen, E. Poon, F. Cury, T. Falco, and F. Verhaegen. "Dosimetric consequences of misalignment and realignment in prostate 3DCRT using intramodality ultrasound image guidance". In: *Medical physics* 37.6 (2010), pp. 2787–2795.

[4]   U. van der Heide, A. Kotte, H. Dehnad, P. Hofman, J. Lagenijk, and M. van Vulpen. "Analysis of fiducial marker-based position verification in the external beam radiotherapy of patients with prostate cancer". In: *Radiotherapy and Oncology* 82.1 (2007), pp. 38–45.

[5]   A. McPartlin, X. Li, L. Kershaw, U. Heide, L. Kerkmeijer, C. Lawton, U. Mahmood, F. Pos, N. van As, M. van Herk, D. Vesprini, J. van der Voort van Zyp, A. Tree, and A. Choudhury. "MRI-guided prostate adaptive radiotherapy–A systematic review". In: *Radiotherapy and Oncology* 119.3 (2016), pp. 371–380.

[6]   J. Lagendijk, A. Raaymakers B.W.and Raaijmakers, J. Overweg, K. Brown, E. Kerkhof, R. van der Put, B. Hårdemark, M. van Vulpen, and U. van der Heide. "MRI/linac integration". In: *Radiotherapy and Oncology* 86.1 (2008), pp. 25–29.

[7]   D. Fontanarosa, S. Van der Meer, J. Bamber, E. Harris, T. O'Shea, and F. Verhaegen. "Review of ultrasound image guidance in external beam radiotherapy: I. Treatment planning and inter-fraction motion management". In: *Physics in Medicine & Biology* 60.3 (2015).

[8]   T. O'Shea, J. Bamber, D. Fontanarosa, S. van der Meer, F. Verhaegen, and E. Harris. "Review of ultrasound image guidance in external beam radiotherapy part II: intra-fraction motion management and novel applications". In: *Physics in Medicine & Biology* 61.8 (2016).

[9]   S. Camps, D. Fontanarosa, P. de With, F. Verhaegen, and B. Vanneste. "The use of ultrasound imaging in the external beam radiotherapy workflow of prostate cancer patients". In: *BioMed Research International* 2018 (2018).

[10]   G. Huang, Z. Liu, K. Weinberger, and L. van der Maaten. "Densely connected convolutional networks". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Vol. 1. 2. 2017, p. 3.

[11]   M. Kemmler, E. Rodner, E. Wacker, and J. Denzler. "One-class classification with Gaussian processes". In: *Pattern Recognition* 46.12 (2013), pp. 3507–3518.

[12]   L. Zhang, N. Dudley, T. Lambrou, N. Allinson, and X. Ye. "Automatic image quality assessment and measurement of fetal head in two-dimensional ultrasound image". In: *Journal of Medical Imaging* 4.2 (2017).

[13] B. Rahmatullah, I. Sarris, A. Papageorghiou, and J. Noble. "Quality control of fetal ultrasound images: Detection of abdomen anatomical landmarks using adaboost". In: *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE. 2011, pp. 6–9.

[14] J. Schwaab, Y. Diez, A. Oliver, R. Martí, J. van Zelst, A. Gubern-Mérida, J. Mourri A.B.and Gregori, and M. Günther. "Automated quality assessment in three-dimensional breast ultrasound images". In: *Journal of Medical Imaging* 3.2 (2016).

[15] S. Kirkpatrick, C. Gelatt, and M. Vecchi. "Optimization by simulated annealing". In: *Science* 220.4598 (1983), pp. 671–680.

[16] J. Fleiss. "Measuring nominal scale agreement among many raters". In: *Psychological Bulletin* 76.5 (1971), pp. 378–382.

[17] S. Khan and M. Madden. "One-class classification: taxonomy of study and review of techniques". In: *The Knowledge Engineering Review* 29.3 (2014), pp. 345–374.

[18] Z. Ghahramania. "A Tutorial on Gaussian Processes (or why I don't use SVMs)". In: *Mlss2011.Comp.Nus.Edu.Sg* (2011).

[19] J. Bradshaw, A. Matthews, and Z. Ghahramani. "Adversarial examples, uncertainty, and transfer testing robustness in Gaussian process hybrid deep networks". In: *arXiv preprint arXiv:1707.02476* (2017), pp. 1–33.

[20] A. G. G. Matthews, M. van der Wilk, T. Nickson, K. Fujii, A. Boukouvalas, P. León-Villagrá, Z. Ghahramani, and J. Hensman. "GPflow: A Gaussian process library using TensorFlow". In: *Journal of Machine Learning Research* 18.40 (2017), pp. 1–6.

[21] M. Titsias. "Variational learning of inducing variables in sparse Gaussian processes". In: *Aistats* 5 (2009), pp. 567–574.

[22] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al. "TensorFlow: A system for large-scale machine learning". In: *OSDI*. Vol. 16. 2016, pp. 265–283.

[23] D. Kingma and J. Ba. "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980* (2014), pp. 1–15.

[24] J. Landis and G. Koch. "The measurement of observer agreement for categorical data". In: *Biometrics* (1977), pp. 159–174.

[25] J. Chen, L. Yang, Y. Zhang, M. Alber, and D. Chen. "Combining fully convolutional and recurrent neural networks for 3d biomedical image segmentation". In: *Advances in Neural Information Processing Systems*. 2016, pp. 3036–3044.

6 Proof-of-concept study for pseudo-CT scan creation

Abstract

Imaging of patient anatomy during treatment is a necessity for position verification, and for adaptive radiotherapy based on daily dose recalculation. Ultrasound (US) image guided radiotherapy systems are currently available to collect US images at the simulation stage (US_{sim}), co-registered with the simulation computed tomography (CT), and during all treatment fractions. We hypothesize that a deformation field derived from US-based deformable image registration can be used to create a daily pseudo-CT image (CT_{ps}) that is more representative of the patient's geometry during treatment than the CT acquired at simulation stage (CT_{sim}).

The three prostate patients, considered to evaluate this hypothesis, had co-registered CT and US scans on various days. In particular, two patients had two US-CT datasets each and the third one had five US-CT datasets. Deformation fields were computed between pairs of US images of the same patient and then applied to the corresponding CT_{sim} scan to yield a new deformed CT_{ps} scan. The original treatment plans were used to recalculate dose distributions in the simulation, deformed and ground truth CT (CT_{GT}) images to compare Dice similarity coefficients, maximum absolute distance, and mean absolute distance on CT delineations and gamma index (γ) evaluations on both the Hounsfield units (HU) and the dose.

In the majority, deformation did improve the results for all three evaluation methods. The change in gamma failure for dose (γ_{Dose}, 3%, 3 millimeter) ranged from an improvement of 11.2% in the prostate volume to a deterioration of 1.3% in the prostate and bladder. The change in gamma failure for the CT images (γ_{CT}, 50 HU, 3 millimeter) ranged from an improvement of 20.5% in the anus and rectum to a deterioration of 3.2% in the prostate.

This new technique may generate CT_{ps} images that are more representative of the actual patient anatomy than the CT_{sim} scan.

6.1 Introduction

Image guidance has become an essential part of radiotherapy (RT) treatment to allow for safe delivery of radiation doses. Image guided RT (IGRT) is often performed for several or all treatment fractions to position the patient correctly. Beyond the aim of image guidance, the availability of daily imaging also allows for the possibility of adaptive RT (ART) [1], [2]. The goal of ART is to improve RT treatment by systematically monitoring dose discrepancies and incorporating them to re-optimize the treatment plan. Normally, only the planning computed tomography (CT) image, acquired at simulation stage, is available for the dose calculation, but both inter- and intra-fraction patient anatomy motion and changes (like tumor shrinkage, nodal volume changes and weight loss) may alter the dose distribution [3]–[6]. In ART, the anatomy from the planning CT is updated by the anatomy from the daily imaging, acquired during the IGRT workflow to monitor dose distribution and if necessary adapt the treatment plan.

CT scanners are usually not available in the treatment room. Instead, cone-beam computed tomography (CBCT) can be used for dose calculations either directly [7]–[10], or indirectly with deformable image registration (DIR) [11], [12], even though they offer a lower image quality when compared to CT scanners. In some studies, using the CBCT directly for dose calculations, the inaccuracies in the Hounsfield units (HU) are large enough to result in clinically relevant dose errors [13]–[15].

In this chapter a workflow is introduced to produce pseudo-CT images based on deformable registration of ultrasound (US) volumes. A 3D US IGRT system can acquire volumetric, high-contrast soft-tissue images non-invasively on a daily basis without using ionizing radiation (Fig. 6.1). Subsequently, deformable registration of these volumes can reveal changes in tissue distribution that occurred over time.
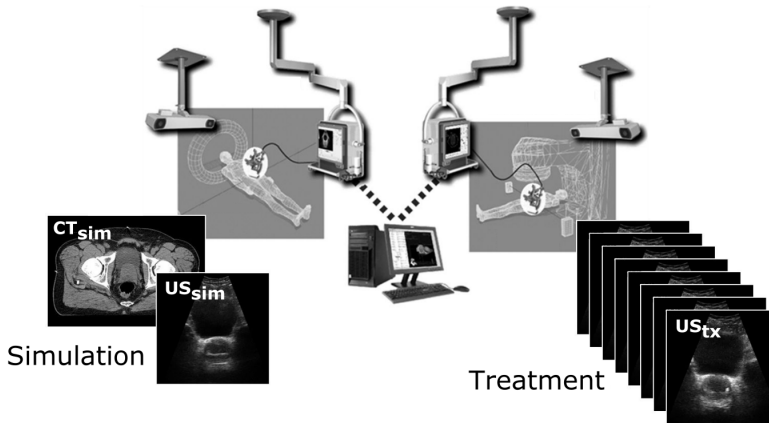
Figure 6.1: Workflow of acquisition of CT_{sim}, US_{sim} and US_{tx} images (Clarity US system; Elekta). (Figure adapted from Elekta with their permission)

Relatively few papers on US to US deformable registration can be found in the literature and as far as we could find, there are presently no papers involving deformable registration of pelvic or abdominal US volumes in RT. In other medical fields, however, some publications are available. For example, Shekhar *et al.* [16] proposed a non-rigid method based on mutual information to register cardiac US images in different phases throughout the complete cardiac cycle.

A similar workflow as proposed in this study was presented for brain surgery applications by Pennec *et al.* [17]. In this study, pre-operative magnetic resonance (MR) images and US images were acquired. Subsequently, intra-operative US images were used to create pseudo-MR images of the brain. This resulted in acceptable representations of the brain anatomy during surgery.

As these results were promising, we used a similar approach to create pseudo-CT images (CT_{ps}). We hypothesize that a pseudo-CT image can be created based on CT_{sim} using a deformation field calculated between US_{sim} and US_{tx}. We expect that the CT_{ps} so created gives a better representation of the patient's anatomy during treatment delivery than the planning CT_{sim}.

6.2 Materials and methods

6.2.1 The concept

In the proposed workflow (Fig. 6.2) for CT_{ps} image creation, DIR has to be performed to calculate a deformation field between US_{sim} and US_{tx}. Subsequently, this deformation field has to be applied to CT_{sim} which results in the creation of CT_{ps}.
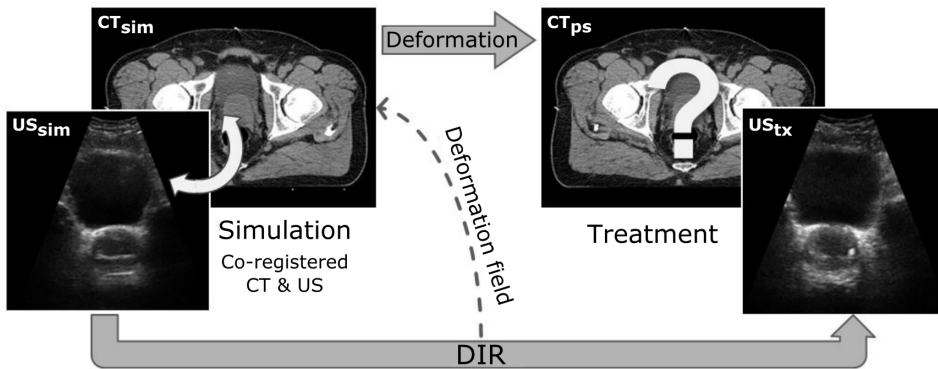
Figure 6.2: Deformable image registration (DIR) is computed between the two US images (US_{sim} & US_{tx}) and then applied to CT_{sim} resulting in a new pseudo CT_{ps}. The question is whether this CT_{ps} is indeed representative for the patient anatomy during treatment.

6.2.2 Patient scans

Clinical examples with multiple co-registered US-CT combinations at the simulation stage (instead of the treatment stage), were used to validate the concept. In this study three prostate cancer patients from a previous study [18] were used. Due to clinical reasons, these patients underwent additional US and CT imaging next to US_{sim} and CT_{sim} acquisition. In the normal clinical workflow, these extra CT and US images are not acquired. The extra CT scans were used as ground truth scans (CT_{GT}) to which the derived CT_{ps} scans could be compared in this proof of concept study. In Table 6.1, the method used to calculate and evaluate the result from the deformations is described.

Table 6.1: Overview of the US-CT combinations used to calculate and evaluate US-based deformable image registration. Patient 3 sets A-D differ in the second US-CT pair that was used to test the method.

Patient	Set	Used as US_{sim}	Used as US_{tx}	Used as CT_{sim}	Used as CT_{tx}
1		US_1	US_2	CT_1	CT_2
2		US_1	US_2	CT_1	CT_2
3	A	US_1	US_2	CT_1	CT_2
3	B	US_1	US_3	CT_1	CT_3
3	C	US_1	US_4	CT_1	CT_4
3	D	US_1	US_5	CT_1	CT_5

The co-registered CT-US images were acquired at two time points for Patients 1 and 2 (3 and 1 weeks apart, respectively). Acquisitions for Patient 3 were made for five time points where the first two were 2 weeks apart and the following three time points were one week apart.

All co-registered US-CT combinations were acquired in the CT room with the patient's external skin markers positioned along the room lasers. The 3D US scans (Clarity system; Elekta, Stockholm, Sweden, voxels: 1 x 1 mm^2 x 3 mm slice thickness; US probe type C5-2/60, center frequency 3.5 MHz, Sonix Series; Ultrasonix Medical Corporation, Richmond, BC, Canada) were performed trans-abdominally immediately before or after the CT-scan. The number of voxels of the US images varied between [512, 512, 90] and [512, 512, 131]. For each patient, the images were resampled to match the dimensions of the first acquired US volume (US_{sim}).

The CT scans were acquired using a SOMATOM Sensation Open (Syngo CT 2006A, Siemens, Germany; voxels: 1 x 1 mm^2 x 3 mm slice thickness). Both scans were performed in the same supine patient position, stabilized with knee fix and foot support (Combifix, Civco Medical Solutions, Kalona, Iowa, USA), resulting in a correct automatic fusion of the US and CT images [19].

In all US images, the prostate was delineated. All CT images had delineations of the body contour, prostate, seminal vesiculae (SV, except for Patient 3), anus, rectum and bladder (except for Patient 1).

6.2.3 Deformation

For each US-CT combination (as detailed in Table 6.1) deformation fields were calculated using a DIR algorithm (B-spline method from ElastiX; Utrecht, the Netherlands) [20], [21]. Prior to the deformation field calculation, all volumes were resampled to the same image dimensions per patient. In addition, segmentation of the CT_{sim} images resulted in a binary mask of the bones and the region of interest (ROI) was defined as the overlapping parts of the US images (ROI: $US_{sim} \cap US_{tx}$). All these preprocessing steps were performed with the Matlab (MathWorks Inc., Natick, MA) software.

During the acquisition of the different US-CT combinations, the patients were in the same position with the body markers aligned to the lasers. For this reason, no rigid transformation was performed prior to the deformable registration, in particular to prevent erroneous full body shifts based on internal shifts of the prostate [22].

As mentioned before, the deformable registration was performed using the ElastiX software. This software package requires three inputs: fixed image (US_{tx}), moving image (US_{sim}) and a parameter file. The parameter file contains all the parameters that determine the characteristics of the registration. In Appendix C.1 an example of such a parameter file is detailed.

In this study the deformable registration was performed either on the overlapping parts of the US images or on binary masks of the delineated prostate volumes only. In total five different parameter sets (Parameters A-E in Table 6.2) were defined for this purpose using the file in Appendix C.1 as a basis.

Table 6.2: Five different parameter sets (A-E) were used during the deformable registration. This registration could be based on the whole US volume or on the binary mask of the delineated prostate volume only (reported in the columns: Fixed image and Moving image). In addition, both the metric (normalized-correlation [NC] or mean-squares [MS]) and iterations were varied among the different sets.

Parameter set	Fixed image	Moving image	Metric	Iterations
A	US_{tx}	US_{sim}	NC	10
B	US_{tx}	US_{sim}	NC	50
C	US_{tx}	US_{sim}	NC	100
D	Prostate mask US_{tx}	Prostate mask US_{sim}	MS	100
E	Prostate mask US_{tx}	Prostate mask US_{sim}	MS	300

The deformation field calculations were based on the overlapping parts of the US images, but were propagated further through the image (Fig. 6.3). Also bones

were sometimes present in these overlapping parts. As bones are in principle rigid structures, they are not expected to undergo deformations. Therefore, the binary bone mask defined during preprocessing was input in the rigidity penalty [23] of ElastiX to prevent bones from deforming.
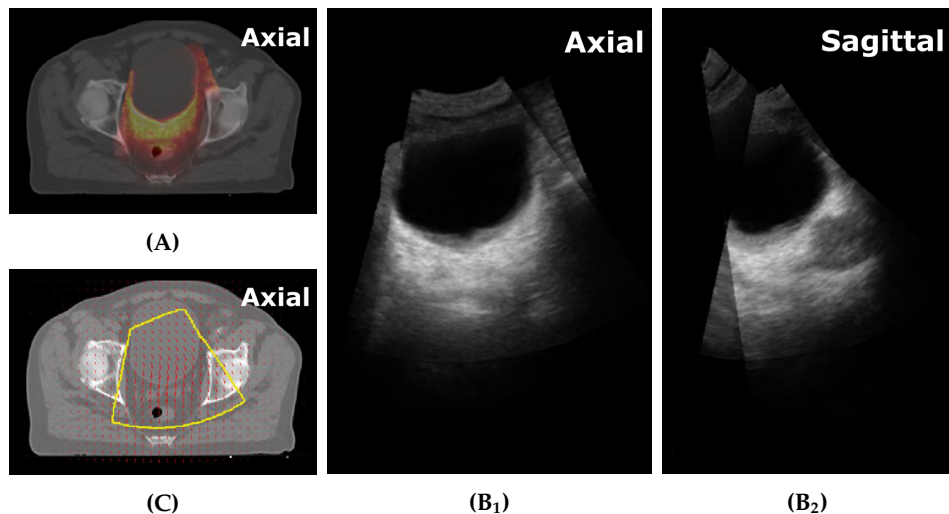
Figure 6.3: Example of overlap between CT (grey) and US (color) (A), and between two US images (B) of Patient 1. US based DIR can only be performed on the area where both CT and US information (of both US_{sim} and US_{tx}) is available. In this example only the prostate and its surrounding tissue, e.g. a part of the bladder, are present in both US images. In panel (C) only the overlapping area of both US images (yellow contour) contains information where the deformation field (2D representation with red arrows) is based on. The field propagates further beyond this border.

6.2.4 Evaluation of the deformation

The created CT_{ps} and the deformed CT delineations were then compared to the ground truth, i.e. the corresponding CT_{GT} and its delineations. The contours were evaluated using the Dice similarity coefficient [DSC = $(2|X \cap Y|)/(|X| + |Y|)$]. A DSC ratio of 1 indicates complete overlap, while 0 indicates no overlap.

In addition the prostate contours were also evaluated using both the maximum absolute distance (MAX) and mean absolute distance (MAD) [24]. The MAX defines the largest difference between two contours, e.g., prostate contour A and prostate contour B. For each point a on prostate contour A, the minimal distance to all points on prostate contour B was calculated. The same was repeated for each point b on prostate contour B with respect to prostate contour A. This resulted in a set of minimal distances and the maximum of this set is referred to as MAX. Calculating the mean of this set gave the MAD.

The CT_{sim} and CT_{ps} images were compared to CT_{GT} using a gamma (γ) index evaluation [25], [26]. The γ index is commonly used for dose evaluations. Prior to the index calculation, two acceptance criteria need to be set: voxel-by-voxel numerical dose difference and distance-to-agreement (DTA: distance between a voxel of one volume and the nearest voxel in the other volume that has the same dose). The resulting index gives information on a voxel scale, while taking the voxels in the vicinity into account as well.

In this case not only dose was evaluated with the γ index, but also HU (γ_{CT}). The γ values were calculated using an in-house developed method [27], [28] using Matlab and C++. The used method allows the sign of the γ value to indicate whether an overdose ($\gamma > 0$) or underdose ($\gamma < 0$) is found for each voxel [28]. In this case, because we evaluate HU, a $\gamma > 0$ means that the HU is relatively higher than the reference and $\gamma < 0$ means that the HU is relatively lower. A value $|\gamma| > 1$ in a voxel indicates that the voxel fails to meet the acceptance criteria, in this case a 50 HU voxel intensity difference and a 3 mm distance to agreement. (The 50 HU is a conservative measure based on that for typical radiotherapy beams; to produce a 1% error in dosimetry would require errors of over 8% in bone electron density [29] and hence HU. The 3 mm distance to agreement is a commonly used criterion in dosimetry [26]).

The percentages of the volume with a $|\gamma_{CT}| > 1$ within the contours 'intersection body contours', 'prostate', 'anus & rectum', and 'bladder' were reported. The percentages of gamma failure and DSC evaluations are reported using the contours of the CT_{GT}, except for the 'intersection body contours' which is the overlapping part of the body contours of both CT_{sim} and CT_{GT}.

6.2.5 Dose calculation and evaluation

Dose distributions were obtained by recalculating the original treatment plans (5-beam IMRT plans; XiO CMS 4.51, Elekta, Stockholm, Sweden) designed on the planning CT_{sim}, on the CT_{sim}, CT_{ps} and CT_{GT} scans. For this, an in-house developed software package was used, based on Monte Carlo simulation using the XVMC code [30], [31]. Dose distributions on the CT_{sim} and CT_{ps} images were compared to the dose on CT_{GT} using a γ evaluation (γ_{Dose}) [25], with acceptance criteria of 3% dose difference and 3 mm distance to agreement. Again the percentage of the volume with a $|\gamma_{Dose}| > 1$ within the contours 'intersection body contours', 'prostate', 'anus & rectum', and 'bladder' was reported.

6.3 Results

In most cases, deformation did improve the results according to all evaluation methods, although these improvements were in some cases very small or even negligible. Only for Patient 1 there was a large improvement (more than 10% decrease in the volume with $|\gamma_{Dose}| > 1$) in the dose of the prostate when the

intensity-based normalized-correlation metric with 100 iterations (parameter set C) was used (Table 6.3).

In Fig. 6.4, an example is given for Patient 1 using parameter set C. In the second column the overlap of the prostate and anus & rectum contours is shown. DSC increased by 0.3 when the deformations were used. The third and fourth column show the γ_{CT} and γ_{Dose} values. In the overlapping body contours, the percentage of γ_{CT} failure decreased by 1.7% in volume. For the prostate and anus & rectum contours, there was a γ_{CT} failure decrease of 9% and 8.4%, respectively. For the dose, the volume percentage of γ_{Dose} failure decreased by 11.2% in volume for the prostate. Yet the percentage of γ_{Dose} failure decreased by only 0.6% and 0.0% for the overlapping body contours and anus & rectum, respectively.

| | CTs | Overlap of prostate and anus/rectum | γ_{CT}(50HU, 3 mm) | γ_{Dose}(3%, 3 mm) | γ |
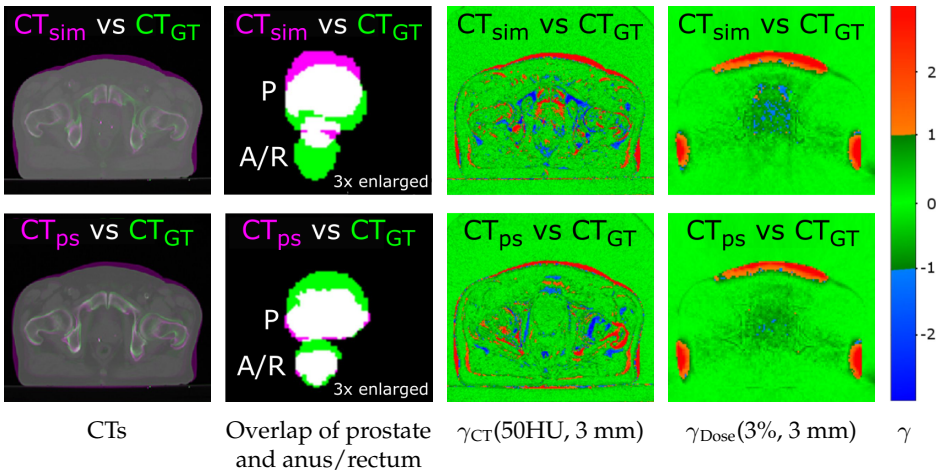
Figure 6.4: Results for Patient 1 (parameter set C). In the first colums the CT_{sim} and CT_{ps} (in pink) are compared to CT_{GT} (in green). In the second column, the contours of prostate (P) and anus/rectum (A/R) are compared. When the images are greyscale (Column 1) or white (Column 2) there is overlap between the compared images. The third and fourth columns show the γ_{CT} (Column 3) and γ_{Dose} (Column 4). In green, the γ values are between -1 and 1. In red and blue are the voxels in which the γ failed to meet the criteria of (50HU, 3 mm) for the CT values and (3%, 3 mm) for the dose. For Column 4, the areas where there is an underdosages compared to CT_{GT} ($\gamma_{Dose} < $ -1) are shown in blue. In red there is an overdosage compared to CT_{GT} ($\gamma_{Dose} > 1$).

All available results for Patient 1 are summarized in Fig. 6.5. Fig. 6.5A shows that the DSC improved for all parameter sets. For prostate the best results were obtained with parameter set E, for anus & rectum set C performed best. Both the MAD and MAX where smaller compared to the reference situation (Fig. 6.5B). Fig. 6.5C and 6.5D detail results on gamma failure, based on CT values and dose, respectively. In case of CT-based evaluations the best results were achieved using parameter set B for prostate and anus & rectum, and using parameter set D for

the body contours. For the dose-based evaluations, parameter set C gave the best results in all cases. The analyses were repeated for all available patient data and the overview of the results is detailed in Appendix C.2.

Evaluation of all patient cases (Table 6.3 and Fig. C1.A,E,I,M,Q,U in Appendix C.2) shows that the DSC of the prostate increased the most for the two contour-based parameter sets (D and E). Parameter set E with 300 iterations did not succeed in the deformation of Patient 2 because there was a too small overlapping volume. Therefore not enough voxels could be mapped and the registration failed to find a solution. Only for Patient 3a none of the parameter sets gave an improvement for any of the contours. Overall, the maximum changes in DSC for the intensity-based normalized-correlation parameter sets were a decrease of -0.5 or an improvement of +0.3. For the contour-based parameter sets, these values were -0.3 and +0.4.
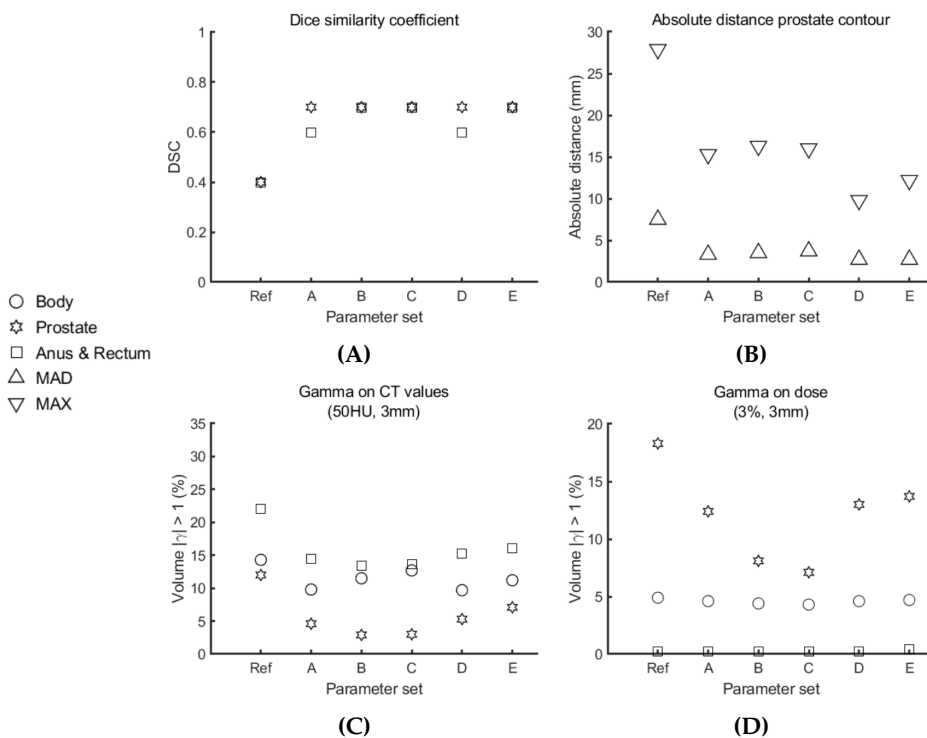
Figure 6.5: Results for all five parameter sets used on Patient 1. The circle represents the body contours, the star the prostate contours and the square a combination of anus and rectum. (A) DSC. (B) Absolute distance for the prostate contours (MAX and MAD). (C) Volume percentage of the gamma failure ($|\gamma_{CT}(50HU, 3 \text{ mm})| > 1$. (D) Percentage of a gamma failure ($|\gamma_{Dose}(3\%, 3 \text{ mm})| > 1$).

Table 6.3: Five evaluation methods were used to evaluate the delineated prostate contours. The first and second column detail the patient and the used evaluation method. Both gamma index values show the volume percentage of gamma failure, (γ_{CT}(50HU, 3mm) > 1) and (γ_{Dose}(3%, 3mm) > 1) respectively. In the third column the reference situation (comparison beween CT_{sim} and CT_{GT}) can be found. In the final five columns the results for each of the parameter sets (A-E) are detailed. The bold numbers indicate which parameter sets resulted in the same result or in an improvement with respect to the reference.

Patient	Metric		Ref	A	B	C	D	E
1	DSC		0.4	**0.7**	**0.7**	**0.7**	**0.7**	**0.7**
	MAD	[mm]	7.5	**3.3**	**3.5**	**3.7**	**2.7**	**2.7**
	MAX	[mm]	27.9	**15.3**	**16.3**	**16.0**	**9.8**	**12.2**
	γ_{CT}	[%]	12.0	**4.6**	**2.9**	**3.0**	**5.3**	**7.1**
	γ_{Dose}	[%]	18.3	**12.4**	**8.1**	**7.1**	**13.0**	**13.7**
2	DSC		0.5	**0.5**	**0.5**	**0.5**	**0.6**	-
	MAD	[mm]	5.1	5.5	5.5	5.7	**3.9**	-
	MAX	[mm]	16.0	18.1	21.6	23.2	**13.9**	-
	γ_{CT}	[%]	11.5	14.6	12.8	12.3	**10.2**	-
	γ_{Dose}	[%]	1.6	2.2	2.8	2.9	**1.4**	-
3a	DSC		0.8	0.6	0.4	0.4	0.6	0.6
	MAD	[mm]	1.6	4.5	6.1	6.5	3.8	3.8
	MAX	[mm]	5.8	14.3	16.3	18.0	10.7	10.7
	γ_{CT}	[%]	6.9	7.4	7.8	8.5	7.5	7.3
	γ_{Dose}	[%]	2.3	**2.0**	**1.6**	**1.9**	**2.1**	**1.9**
3b	DSC		0.8	0.6	0.6	0.6	**0.8**	**0.8**
	MAD	[mm]	2.3	4.3	4.2	4.4	**2.1**	**2.3**
	MAX	[mm]	8.4	16.4	18.9	19.1	**5.9**	**6.1**
	γ_{CT}	[%]	6.6	**6.6**	7.0	6.8	**5.0**	**5.1**
	γ_{Dose}	[%]	3.8	**3.5**	**2.7**	**3.2**	**1.8**	**2.2**
3c	DSC		0.6	0.5	0.5	0.5	**0.7**	**0.7**
	MAD	[mm]	4.4	**3.8**	4.4	4.9	**2.5**	**2.2**
	MAX	[mm]	12.4	14.0	**12.4**	13.0	**7.1**	**6.2**
	γ_{CT}	[%]	9.9	**7.1**	**7.7**	**7.2**	**7.9**	**6.8**
	γ_{Dose}	[%]	4.1	**3.3**	**3.5**	4.2	**3.4**	**3.8**
3d	DSC		0.4	**0.5**	**0.6**	**0.7**	**0.8**	**0.8**
	MAD	[mm]	6.6	**5.9**	**4.5**	**4.0**	**2.1**	**1.9**
	MAX	[mm]	20.1	22.0	25.7	27.5	**10.4**	**9.7**
	γ_{CT}	[%]	11.8	**6.7**	**4.4**	**4.0**	**10.1**	**9.3**
	γ_{Dose}	[%]	10.3	**9.9**	**9.5**	**8.7**	**6.9**	**6.5**

For the changes in CT HU values, the percentage of the volume with a $|\gamma_{CT}$(50HU, 3mm)$| > 1$ for prostate is shown in Table 6.3 and for the other contours in Appendix C (Table C.1 and Fig. C1.C,G,K,O,S,W). A maximum

improvement was seen of 20.5% (14.6% for contour-based) and the poorest results gave an increase of 3.2% (2.2% for contour-based) in the volume with $|\gamma_{CT}(50HU, 3\,mm)| > 1$.

Looking at the prostate results as shown in Table 6.3, in case an improvement was achieved, the contour parameter set (D, 100 iterations) seemed to give an improvement in most cases, yet it was not always the best one. The results for the other contours (body, anus & rectum and bladder), that can be found in Table C.1 in Appendix C.2 confirm this as well.

6.4 Discussion

We have evaluated the impact of applying US-derived tissue deformations to approximate CT images to the real anatomical organ position of prostate patients during radiation therapy. As noted earlier, a similar workflow was presented by Pennec *et al.* [17] for brain surgery applications. However, in that study pseudo-MR images of the brain were created. To our knowledge, this is the first time a similar method is used for RT applications.

In this study, Patients 1 and 3d would have benefited most from the deformations (> 3% volume decrease for the volume with a $|\gamma_{Dose}|>1$). In addition, the difference in dose between CT_{sim} and CT_{GT} was there also the largest (> 10% volume with a $|\gamma_{Dose}|>1$). For the other patient cases, the improvements were not clinically relevant.

Ideally, one should be able to evaluate beforehand which patients would benefit from applying the deformations. The only metric that is available prior to DIR and could be suitable is the DSC of the prostate contours on US_{sim} and US_{tx}. A statistical evaluation was performed to find a possible correlation between these DSC and the effect on the dose deposition on the prostate ($|\gamma_{Dose}|>1$). Unfortunately such a correlation was not found, possibly due to the limited number of patients. However, there seems to be a trend that the patients with the largest geometric changes benefit most from deformations, but a future study with a larger image database will be necessary to validate the predictive power of this DSC parameter to obtain a clearer indication when it is worthwhile to perform DIR.

Besides a larger database to perform statistics, such a database could be used to find an optimal metric and parameter set for the DIR. For this proof-of-principle study, two deformation metrics were used and only the number of iterations varied. Optimization of the metrics and parameter set may improve the results. In the current study, the results of the evaluation methods were not always in agreement. Even between the CT and dose values there were some differences due to the cumulative effect of the dose along the beam path. The differences between change in γ_{CT} and γ_{Dose} are caused by the fact that the dose in the organs is not only dependent on the local HU, but also on the HU along the beam path.

The best evaluation method is dependent on the purpose; the evaluation of the best parameter set should therefore always be assessed with the correct evaluation method. In case of ART, this could be γ_{Dose}(3%, 3 mm).

A limitation of an US-based deformation field is that the volume of the CT on which one can directly calculate the deformation field is limited to the volume of the US data available (Fig. 6.3). The deformation field propagates further, but this is not based on image data and is therefore maybe less reliable. For Patient 2, a small overlap of US volumes resulted in a failure in parameter set E. Standardization of scanning, such that at least the complete prostate is visible, the US volume overlap is maximal, and US images with larger fields of view may improve the results. Transperineal scanning with a larger image sector or perhaps even fusion of multiple US scans from different directions, can extend the field of view.

However, the US image will never completely overlap the CT image, therefore part of the deformation field will still be based on only an extrapolated deformation field. For an ideal exact extrapolation, it may be crucial to take into account the mechanical properties of tissues and organs, such as skin, bones and bladder, which are positioned outside of the overlapping US images. In this work, some deformation field propagation outside of the overlapping US volumes is already inherently taken into account, due to the use of the so-called multi-resolution approach during the deformable registration. In this approach the registration starts with images that have a lower complexity. For example, images that were smoothed and possibly down-sampled. During the registration, a B-spline control point grid is overlaid on the fixed image. This grid is always rectangular. Control points that are outside of the region of interest (overlapping parts of the US volumes) are in principle not affected. However, due to the multi-resolution approach the control-point spacing is larger at lower resolutions than at higher resolutions. For this reason, a larger area around the region of interest is affected at lower resolutions, which typically produces deformations outside of the region of interest.

Another reason why it is important to have standardization of the US scanning is that, just like with the IGRT usage of the US images, it is important to have reproducible US images. In particular the probe pressure [18], [32] and speed-of-sound aberration [33] along the imaging beam should be comparable. One cannot distinguish between the US imaging dependent changes caused by non-standardized procedures or real anatomy changes. Therefore it is best to prevent them, or correct [34]–[37] for them prior to the DIR procedure. For our specific cases, preliminary inspection revealed that these corrections were not necessary.

Validation of the DIR methods in general is also still necessary to reliably perform DIR for ART. Different deformation algorithms lead to different results, therefore more research is necessary.

6.5 Conclusion

It was possible to generate a pseudo-CT (CT_{ps}) with the use of DIR based on US imaging, which was more representative of CT_{GT} than CT_{sim}. For the patients with the smaller prostate change over time, the procedure did not improve the dose calculations significantly. The largest improvements were seen for patients with the largest anatomical changes. More research with a larger image database is necessary to find an optimal deformation metric and parameter set. With a larger database, it might be possible to find a predictive measure and associated criteria to decide whether DIR is worthwhile for individual patients.

References

[1] D. Yan, E. Ziaja, D. Jaffray, J. Wong, D. Brabbins, F. Vicini, and A. Martinez. "The use of adaptive radiation therapy to reduce setup error: a prospective clinical study". In: *International Journal of Radiation Oncology* Biology* Physics* 41.3 (1998), pp. 715–720.

[2] Q. Wu, T. Li, Q. Wu, and F. Yin. "Adaptive radiation therapy: technical components and clinical applications". In: *The Cancer Journal* 17.3 (2011), pp. 182–189.

[3] E. Hansen, M. Bucci, J. Quivey, V. Weinberg, and P. Xia. "Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer". In: *International Journal of Radiation Oncology* Biology* Physics* 64.2 (2006), pp. 355–362.

[4] D. Fraser, Y. Chen, E. Poon, F. Cury, T. Falco, and F. Verhaegen. "Dosimetric consequences of misalignment and realignment in prostate 3DCRT using intramodality ultrasound image guidance". In: *Medical physics* 37.6 (2010), pp. 2787–2795.

[5] D. Fraser, P. Wong, K. Sultanem, and F. Verhaegen. "Dosimetric evolution of the breast electron boost target using 3D ultrasound imaging". In: *Radiotherapy and Oncology* 96.2 (2010), pp. 185–191.

[6] S. van Kranen, A. Mencarelli, S. van Beek, C. Rasch, M. van Herk, and J. Sonke. "Adaptive radiotherapy with an average anatomy model: evaluation and quantification of residual deformations in head and neck cancer patients". In: *Radiotherapy and Oncology* 109.3 (2013), pp. 463–468.

[7] Y. Yang, E. Schreibmann, T. Li, C. Wang, and L. Xing. "Evaluation of on-board kV cone beam CT (CBCT)-based dose calculation". In: *Physics in Medicine & Biology* 52.3 (2007), p. 685.

[8] S. F. Petit, W. J. van Elmpt, S. M. Nijsten, P. Lambin, and A. L. Dekker. "Calibration of megavoltage cone-beam CT for radiotherapy dose

calculations: correction of cupping artifacts and conversion of CT numbers to electron density". In: *Medical physics* 35.3 (2008), pp. 849–865.

[9] 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". In: *Radiation Oncology* 3.1 (2008), p. 42.

[10] W. van Elmpt, S. Petit, D. De Ruysscher, P. Lambin, and A. Dekker. "3D dose delivery verification using repeated cone-beam imaging and EPID dosimetry for stereotactic body radiotherapy of non-small cell lung cancer". In: *Radiotherapy and oncology* 94.2 (2010), pp. 188–194.

[11] M. Moteabbed, G. Sharp, Y. Wang, A. Trofimov, J. Efstathiou, and H. Lu. "Validation of a deformable image registration technique for cone beam CT-based dose verification". In: *Medical physics* 42.1 (2015), pp. 196–205.

[12] C. Veiga, J. McClelland, S. Moinuddin, A. Lourenço, K. Ricketts, J. Annkah, M. Modat, S. Ourselin, D. D'souza, and G. Royle. "Toward adaptive radiotherapy for head and neck patients: Feasibility study on using CT-to-CBCT deformable registration for "dose of the day" calculations". In: *Medical physics* 41.3 (2014).

[13] J. Hatton, B. McCurdy, and P. Greer. "Cone beam computerized tomography: the effect of calibration of the Hounsfield unit number to electron density on dose calculation accuracy for adaptive radiation therapy". In: *Physics in Medicine & Biology* 54.15 (2009).

[14] S. Yoo and F. Yin. "Dosimetric feasibility of cone-beam CT-based treatment planning compared to CT-based treatment planning". In: *International Journal of Radiation Oncology* Biology* Physics* 66.5 (2006), pp. 1553–1561.

[15] C. Houser, A. Nawaz, J. Galvin, and Y. Xiao. "TH-D-ValA-04: Quantitative evaluation of cone beam CT data used for treatment planning". In: *Medical Physics* 33.6 (2006), pp. 2285–2286.

[16] R. Shekhar and V. Zagrodsky. "Mutual information-based rigid and nonrigid registration of ultrasound volumes". In: *IEEE transactions on medical imaging* 21.1 (2002), pp. 9–22.

[17] X. Pennec, P. Cachier, and N. Ayache. "Tracking brain deformations in time sequences of 3D US images". In: *Pattern Recognition Letters* 24.4-5 (2003), pp. 801–813.

[18] S. van der Meer, E. Bloemen-van Gurp, J. Hermans, R. Voncken, D. Heuvelmans, C. Gubbels, D. Fontanarosa, P. Visser, L. Lutgens, F. van Gils, and F. Verhaegen. "Critical assessment of intramodality 3D ultrasound imaging for prostate IGRT compared to fiducial markers". In: *Medical physics* 40.7 (2013).

[19] S. van der Meer, E. Seravalli, D. Fontanarosa, E. Bloemen-van Gurp, and F. Verhaegen. "Consequences of intermodality registration errors for intramodality 3D ultrasound IGRT". In: *Technology in cancer research & treatment* 15.4 (2016), pp. 632–638.

[20] S. Klein, M. Staring, K. Murphy, M. Viergever, and J. Pluim. "Elastix: a toolbox for intensity-based medical image registration". In: *IEEE transactions on medical imaging* 29.1 (2010), pp. 196–205.

[21] D. Shamonin, E. Bron, B. Lelieveldt, M. Smits, S. Klein, and M. Staring. "Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease". In: *Frontiers in neuroinformatics* 7 (2014), p. 50.

[22] M. Ghilezan, D. Jaffray, J. Siewerdsen, M. Van Herk, A. Shetty, M. Sharpe, S. Jafri, F. Vicini, R. Matter, D. Brabbins, and A. Martinez. "Prostate gland motion assessed with cine-magnetic resonance imaging (cine-MRI)". In: *International Journal of Radiation Oncology* Biology* Physics* 62.2 (2005), pp. 406–417.

[23] M. Staring, S. Klein, and J. Pluim. "A rigidity penalty term for nonrigid registration". In: *Medical physics* 34.11 (2007), pp. 4098–4108.

[24] G. Gerig, M. Jomier, and M. Chakos. "Valmet: A new validation tool for assessing and improving 3D object segmentation". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2001, pp. 516–523.

[25] D. Low, W. Harms, S. Mutic, and J. Purdy. "A technique for the quantitative evaluation of dose distributions". In: *Medical physics* 25.5 (1998), pp. 656–661.

[26] D. Low and J. Dempsey. "Evaluation of the gamma dose distribution comparison method". In: *Medical physics* 30.9 (2003), pp. 2455–2464.

[27] L. Persoon, M. Podesta, W. van Elmpt, S. Nijsten, and F. Verhaegen. "A fast three-dimensional gamma evaluation using a GPU utilizing texture memory for on-the-fly interpolations". In: *Medical physics* 38.7 (2011), pp. 4032–4035.

[28] M. Podesta, L. Persoon, and F. Verhaegen. "A novel time dependent gamma evaluation function for dynamic 2D and 3D dose distributions". In: *Physics in Medicine & Biology* 59.20 (2014), p. 5973.

[29] S. Thomas. "Relative electron density calibration of CT scanners for radiotherapy treatment planning." In: *The British Journal of Radiology* 72.860 (1999), pp. 781–786.

[30]   M. Fippel. "Fast Monte Carlo dose calculation for photon beams based on the VMC electron algorithm". In: *Medical physics* 26.8 (1999), pp. 1466–1475.

[31]   W. Van Elmpt, S. Nijsten, R. Schiffeleers, A. Dekker, B. Mijnheer, P. Lambin, and A. Minken. "A Monte Carlo based three-dimensional dose reconstruction method derived from portal dose images". In: *Medical physics* 33.7 (2006), pp. 2426–2434.

[32]   M. Baker and C. Behrens. "Prostate displacement during transabdominal ultrasound image-guided radiotherapy assessed by real-time four-dimensional transperineal monitoring". In: *Acta Oncologica* 54.9 (2015), pp. 1508–1514.

[33]   B. Salter, B. Wang, M. Szegedi, P. Rassiah-Szegedi, D. Shrieve, R. Cheng, and M. Fuss. "Evaluation of alignment error due to a speed artifact in stereotactic ultrasound image guidance". In: *Physics in Medicine & Biology* 53.23 (2008).

[34]   D. Fontanarosa, S. Meer, E. Harris, and F. Verhaegen. "A CT based correction method for speed of sound aberration for ultrasound based image guided radiotherapy". In: *Medical physics* 38.5 (2011), pp. 2665–2673.

[35]   D. Fontanarosa, S. Meer, E. Bloemen-van Gurp, G. Stroian, and F. Verhaegen. "Magnitude of speed of sound aberration corrections for ultrasound image guided radiotherapy for prostate and other anatomical sites". In: *Medical physics* 39.8 (2012), pp. 5286–5292.

[36]   D. Fontanarosa, S. van der Meer, and F. Verhaegen. "On the significance of density-induced speed of sound variations on US-guided radiotherapy". In: *Medical physics* 39.10 (2012), pp. 6316–6323.

[37]   D. Fontanarosa, S. Pesente, F. Pascoli, D. Ermacora, I. Rumeileh, and F. Verhaegen. "A speed of sound aberration correction algorithm for curvilinear ultrasound transducers in ultrasound-based image-guided radiotherapy". In: *Physics in medicine and biology* 58.5 (2013), p. 1341.

7 Various approaches for pseudo-CT scan creation

Abstract

The purpose of this study was to evaluate eight possible approaches to create pseudo-CT images for radiotherapy (RT) treatment re-planning. These re-planning CT scans would normally require a separate CT scan session. If important changes occur in the patient's anatomy between simulation (SIM) and treatment (TX) stages, 3D ultrasound (US) images acquired at the two stages, available in US guided RT workflows, can be used to produce a deformation field. Proof of concept research showed that the application of this deformation field to the SIM CT image yields a pseudo-CT which can be more representative of the patient at TX than SIM CT. Co-registered CT and US volumes acquired at five different time points during the RT course of a prostate cancer patient were combined into data pairs, providing ground truth CT images (CT_{tx}). Eight different methods were explored to create the deformation field that was used to produce the pseudo-CT scan. Anatomical structure comparison and γ index calculations were used to compare the similarity of the pseudo-CT volumes and the reference TX CT volumes.

In five out of ten data pairs, all the eight approaches resulted in the creation of a pseudo-CT equal or more similar to the TX CT than the SIM CT within the region of interest, with an average improvement of 54.1% (range: 5.1% - 126.5%) in Dice similarity coefficient (DSC) and 32.3% (range: 0.3% - 52.6%) in the γ index. For the remaining data pairs, four up to seven approaches resulted in an improvement in both DSC (range: 4.3% - 54%) and γ index (range: 0.8% - 41.3%).

In conclusion, at least four out of eight explored approaches resulted in more representative pseudo-CT images in all the data pairs. In particular, the approaches in which an initial rigid alignment was combined with deformable registration performed best.

7.1 Introduction

Radiotherapy (RT) is one of the treatment modalities for various types of cancer. The goal of RT is to deliver a high radiation dose over several treatment fractions to tumor tissue, while at the same time sparing normal tissue as much as possible. Prior to treatment commencement (simulation stage, SIM) typically an initial computerized tomography scan (simulation CT, CT_{sim}) is acquired, which provides electron density information, necessary for radiation dose calculations. Subsequently, a treatment plan is prepared where regions of interest (ROI, typically target and organs at risk, OARs) are outlined and the optimal dose distribution is calculated.

An important aspect of the RT workflow is patient setup. Prior to each treatment (TX) fraction, the patient's position must correspond to the position at SIM, since the treatment plan that is going to be delivered with highly conformal beams was prepared for this position. Even when the patient appears correctly aligned externally, internal anatomic changes may have occurred. For example, in the case of prostate cancer, different filling of bladder or rectum can cause a shift of the target [1]. Delivery of the initial treatment plan might then result in a suboptimal dose deposition in the tumor and/or in extra unwanted irradiation of normal tissues, possibly producing excessive toxicity.

This implies that, ideally, in case changes have occurred, the treatment plan should be updated prior to (or even during) each treatment fraction (Adaptive Radiation Therapy [2], ART). As CT images are currently required to calculate dose, this would involve acquisition of multiple additional CT images during the treatment course, to enable re-planning. This is not part of the standard clinical routine, though (although some attempts were made in the past [3]), due to the long times, complex procedures, extra dose depositions and high costs involved.

In the proof of concept study by Van der Meer *et al.* [4] a workflow was introduced to produce pseudo-CT image datasets prior to each treatment fraction. In that workflow, based on an ultrasound (US) guidance for RT (USgRT) [5] approach, the US scan taken before treatment fraction n (at time t_n) was used to reveal the geometric changes that occurred between SIM and t_n and to create a pseudo-CT. Van der Meer *et al.* [4] showed that it is possible to create a pseudo-CT scan using deformable image registration which was more representative of the patient's tissue distribution at TX than CT_{sim}.

No other papers that involve deformable registration of US volumes to create pseudo-CT images for RT purposes were found in the literature. However, Pennec *et al.* [6] used deformable registration of preoperative and intraoperative US images of the brain, to create pseudo-MR images for brain surgery applications. This resulted in an acceptable representation of the brain anatomy during surgery.

In this study, the workflow proposed by Van der Meer *et al.* [4], that included two approaches to perform registration of the US volumes between SIM and TX and create a deformation field, was extended by introducing six additional approaches. The workflow including in total eight approaches was applied on image data from one patient who underwent RT for localized prostate cancer. This resulted in the creation of multiple pseudo-CT scans. Subsequently, the similarity between these scans and available ground truth images (CT_{tx}) was assessed. This gave more insight into which of the explored approaches was the most suitable for this particular purpose.

7.2 Materials and methods

7.2.1 Workflow

Van der Meer *et al.* [4] proposed a workflow (Fig. 7.1) in which US images were used to create pseudo-CT images using deformable image registration. We extended this workflow and explored six additional approaches to create a deformation field between US_{sim} and US_{tx}. Deformable registration based on either the total US volume or the prostate contour extracted from this volume was not only used as standalone procedure (as done by Van der Meer *et al.* [4]), but also combined with an initial alignment (translation) which can be based on the total US volume or just the prostate contour as well. Also the execution of the initial alignment as standalone procedure is possible. All explored approaches are detailed in Table 7.1.

When the deformation field so created is then applied on CT_{sim}, this results in the creation of pseudo-CT images (CT_{ps}). In the present work, the CT_{ps} volume is compared to ground truth images (CT_{tx}, in normal clinical workflow these images are not acquired) to assess if the pseudo-CT images represent CT_{tx} more accurately than the CT_{sim} volume. This evaluation, where only the similarity between CT_{ps} and CT_{tx} is assessed, might also give more insight into which approach is the most suitable for the deformation field creation.

Table 7.1: Eight possible approaches to create pseudo-CT images that were explored in this study. The approaches marked with * were used by Van der Meer *et al.* [4] as well.

Procedure	Based on	Followed by	Based on
Translation	US volume		
Translation	Prostate contour US		
Deformable registration*	US volume		
Deformable registration*	Prostate contour US		
Translation	US volume	Deformable registration	US volume
Translation	US volume	Deformable registration	Prostate contour US
Translation	Prostate contour US	Deformable registration	US volume
Translation	Prostate contour US	Deformable registration	Prostate contour US
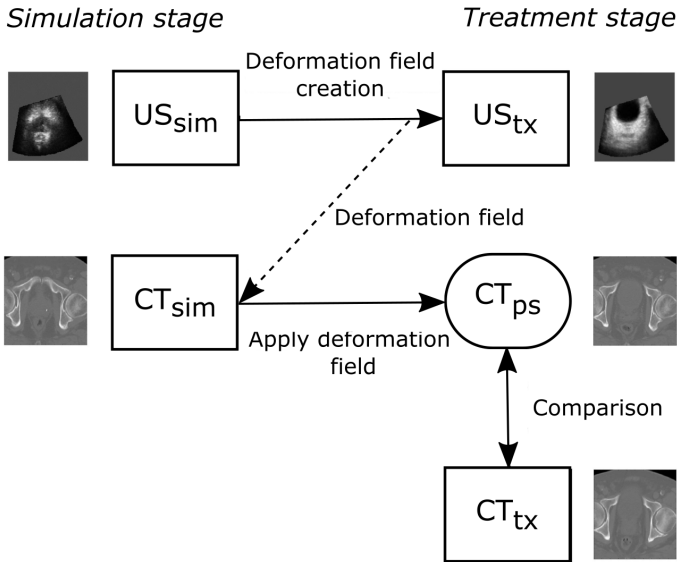
Figure 7.1: Schematic representation of the workflow proposed by Van der Meer *et al.* [4]: US volumes are acquired at both simulation (US$_{sim}$) and treatment stage (US$_{tx}$). A CT volume is normally only acquired at simulation stage (CT$_{sim}$), but for this study four additional CT scans were acquired that could be used as ground truth (CT$_{tx}$). Subsequently, eight approaches (Table 7.1) are investigated to create a deformation field. This deformation field is then applied on the CT$_{sim}$ volume resulting in an approximation of CT image data at treatment stage (CT$_{ps}$). For evaluation purposes CT$_{ps}$ is then compared to CT$_{tx}$.

7.2.2 Data description

Five co-registered CT and transabdominal 3D US (TAUS) image sets were available at different time points for one patient who underwent RT for localized prostate cancer. The time between the imaging sessions was two weeks and subsequent weekly follow-up occurred (Fig. 7.2). All the data were acquired using a Clarity 3D US system (Elekta, Stockholm, Sweden) in a CT-room with the external skin markers of the patient aligned with the room lasers.
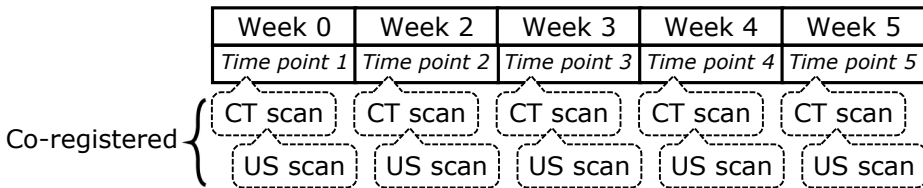
Figure 7.2: Image acquisition time line of one patient who underwent RT for localized prostate cancer. At each time point, a co-registered CT-US scan is acquired. The US scan is acquired right before or after the CT scan acquisition. .

For US volume acquisition a convex transducer was used (type C5-2/60, center frequency 3.5 MHz, Sonix Series, Ultrasonix Medical Corporation, Richmond, BC, Canada). The CT images were acquired with a SOMATOM Sensation Open CT scanner (Syngo CT 2006A, Siemens, Erlangen, Germany; voxels: 1x1x3 mm). On each dataset (US and CT), the prostate was manually segmented by an experienced user. In addition, on the CT volumes anus and rectum were manually segmented as well. No patient setup correction was performed after acquisition of the volumes.

To execute the workflow multiple times and increase statistical significance, cross-combinations of the five co-registered CT and TAUS volumes were made. This resulted in 20 possible pairs, in which one dataset was designated ground truth and the other simulation. For each pair, the overlap between the prostate contours on CT_{sim} and CT_{tx} was expressed with a Dice similarity coefficient (DSC, Section 7.2.6). We decided to focus on the datasets which had an initial prostate overlap ≤ 0.75 DSC, as these might benefit the most from the proposed workflow. This narrowed the number of pairs down to 10 (Table 7.2). From now on, these combinations of ground truth and simulation datasets are referred to as 'data pairs'.

Table 7.2: Image data acquired at different time points (1-5) were combined into data pairs, designating one dataset (co-registered CT and US volumes) as simulation and the other as ground truth. Only data pairs with an initial overlap (DSC) between the prostate contour on CT_{sim} and the contour on CT_{tx} (detailed in Table 7.4) which was smaller than or equal to 0.75 were used in this study.

Data pair	Simulation	Ground truth
T1 - T4	Time point 1	Time point 4
T1 - T5	Time point 1	Time point 5
T2 - T4	Time point 2	Time point 4
T2 - T5	Time point 2	Time point 5
T3 - T5	Time point 3	Time point 5
T4 - T1	Time point 1	Time point 1
T4 - T2	Time point 1	Time point 2
T5 - T1	Time point 5	Time point 1
T5 - T2	Time point 5	Time point 2
T5 - T3	Time point 5	Time point 3

Prior to initial alignment (described in the following section), all the volumes were resampled to the same image dimensions (512x512x128 voxels) using the software REGGUI (version 2.5, Université Catholique de Louvain, Belgium) [7] to make them compatible with pre-processing, registration and evaluation requirements, which arise from the eight different approaches.

7.2.3 Initial alignment

A preliminary visual inspection of the US and CT volumes revealed that not only non-rigid deformation occurred between the time points, but also translation of the whole body, primarily in the lateral direction. This can be explained by the fact that the table of the used CT scanner could not move in this direction. So to ensure correct setup, the patient had to move in lateral direction with respect to the table and this apparently introduced errors. The software ElastiX (Version 4.7, Image Sciences Institute, University Medical Center Utrecht, the Netherlands) [7]–[9] was used to initially align the images applying translation alone. Hereby, a multi-resolution approach could be used [10].

The translation was either based on the total US volumes or on the manually outlined binary prostate contours on the US images alone. In the latter case, the assumption made was that the translation of the prostate is representative for the translation of the whole body. As noted before, primarily lateral translation of the CT image was detected. For the datasets used in this study, trial and error showed that excluding translation in the other directions gave the best results.

ElastiX requires three inputs to perform the registration: a fixed image (in our case the prostate contour US_{tx} or the whole US_{tx} volume), a moving image (the prostate contour US_{sim} or the whole US_{sim} volume), and a parameter file. The parameters in this file determine the registration characteristics, e.g. the similarity metric used. We used a general parameter set that was based on the example translation file provided by ElastiX for the initial alignment (Appendix D.1). In this parameter set, three resolutions were used in the multi-resolution approach. For four data pairs the translation based on the prostate contour only failed. The contour was translated excessively outside the field of view of the image. No excessive translation occurred when, for these cases, less than three resolutions were used in the multi-resolution approach.

At the end, the TransformiX [7]–[9] software, provided with the ElastiX package, was used to apply the deformation field on all SIM datasets. If no translation of the whole body occurred, the initial alignment is potentially useless and time could be saved by not performing it.

7.2.4 Pre-processing

To prepare the volumes for the deformable registration process, a few initial steps were required. In particular, since each voxel must have a spatially corresponding voxel to map to, to make sure that a valid metric value can be calculated, only the overlapping parts of the US volumes were used. In addition, it was important to ensure that the volumes had sufficient overlap, otherwise not enough voxels would be mapped and the registration would fail. In Fig. 7.3 the complete procedure is shown: first, threshold based segmentation of the original volumes (Fig. 7.3A) produced binary masks (Fig. 7.3B) in which potential holes were filled performing a morphological closing with a 4x4 kernel (Fig. 7.3C). Next,

multiplication of these masks resulted in a binary mask of the overlapping regions of the US volumes (Fig. 7.3D). This mask was used as a ROI during the registration procedure, which is described in the next section. Finally, multiplication of this mask by the original US volumes (Fig. 7.3A · Fig. 7.3D) extracted the regions that contain information in both US_{sim} and US_{tx} volumes (Fig. 7.3E).

The CT volumes used for prostate cancer patients usually encompass the whole abdomen, while the US volumes only encompass (parts of) the pelvic region. Due to the fact that the Clarity system outputs the co-registered volumes in CT dimensions, some US volume regions only contain black background voxels. To speed up calculations, all the slices in the US volumes not containing information were removed and the US volumes were cropped to remove surrounding background voxels (Fig. 7.3F). From now on, all mentioned image datasets refer to the cropped images with removed slices, unless stated otherwise.
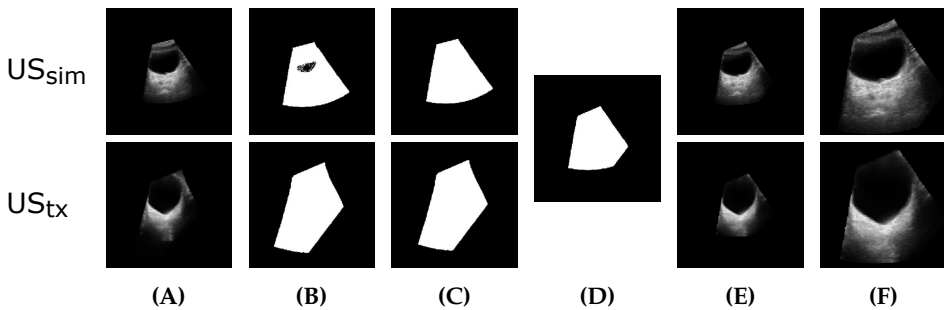
Figure 7.3: Overview of the pre-processing procedure. (A) Spatially corresponding 2D slice of the initial US volumes. (B) Binary masks of the 2D slices obtained with thresholding. (C) Potential holes are closed using morphological closing. (D) Overlapping parts of both US masks obtained by multiplication of the masks in (C). (E) Overlapping parts of the original volumes extracted, (A)·(D). (F) Background pixels are removed to speed up calculations.

Another important observation is that not all types of tissue present in the US volumes deform non-rigidly. For example, bones are rigid structures and therefore are primarily expected to translate or rotate. To asses if the bones deformed non-rigidly during the registration procedure the exact position of the bones at SIM has to be known. For this reason, bones were extracted from CT_{sim} performing semi-automatic segmentation based on thresholding resulting in a binary mask of the bones ($Bone_{sim}$).

Finally, for all CT_{sim} and CT_{tx} volumes, the corresponding contours of prostate, rectum and anus were combined into one volume per dataset containing binary masks of all three structures ($CT_{sim\_contours}$ and $CT_{tx\_contours}$).

7.2.5 Deformable registration and deformation field application

The implemented GUI and ElastiX were also used to perform 3D multi-resolution deformable registration of the US volumes. The parameters needed during this registration are detailed in Section 7.2.7.

Two different strategies to execute the deformable registration were explored. In the first strategy, the voxels of the total US_{sim} volume (moving image) were spatially mapped onto the voxels of US_{tx} (fixed image). In the second strategy, just the binary masks of the prostate contours on the US_{sim} and US_{tx} volumes were used as moving and fixed images, respectively. Both strategies could be executed preceded by one of the initial alignment procedures (Section 7.2.3), or as standalone procedure as well.

During the registration, a B-spline control point grid is overlaid on the fixed image. This grid is always rectangular and control points that are outside the ROI (binary mask of overlapping parts of the US volumes, Section 7.2.4), in principle, are not affected. However, since a multi-resolution schedule is used, the control point spacing is larger at the lower resolutions than at the higher resolutions. This means that a larger area around the ROI is affected at lower resolutions, which can lead to deformations outside the ROI. In subsequent resolutions these deformations are preserved. In this way, to some extent, deformation field propagation outside of the US volumes is inherently taken into account.

After creation of the deformation field, the final step of the registration workflow was its application to CT_{sim} to create the CT_{ps} volume. The deformation field was applied on the whole CT_{sim} volume, specifying which slices should be affected. Prior to evaluation, the resulting CT_{ps} was cropped to match the US dimensions again. The same procedure was repeated for $CT_{sim\_contours}$.

7.2.6 Evaluation

To assess if the CT_{ps} volume represents the CT_{tx} volume better than the CT_{sim} volume, the similarities between CT_{ps} and CT_{tx} and between CT_{sim} and CT_{tx} should be investigated. This evaluation can be based on e.g. position of anatomical structures or correspondence of image intensity values (Hounsfield units [HU]).

First, the determinant of the Jacobian of the deformation fields was calculated to determine if folds in the deformation field were present [11]. Therefore, all the CT_{ps} volumes corresponding to determinant values smaller than zero were excluded from further evaluation.

The overlap of the same anatomical structure contours on different volumes can provide information on the similarity of those volumes. For this reason, the overlap of $CT_{sim\_contours}$ and $CT_{tx\_contours}$, as well as the overlap of $CT_{ps\_contours}$ and $CT_{tx\_contours}$ were calculated using the Dice Similarity Coefficient (DSC) [12], [13]:

$$DSC(X, Y) := \frac{2|X \cap Y|}{|X| + |Y|} \tag{7.1}$$

where X and Y are the binary structure contours and the operator |.| includes all the voxels that have a value equal to 1. A DSC of 1 indicates perfect overlap, while a value of 0 means no overlap at all. For evaluation purposes, also the overlap between the prostate contours on the US volumes as well as the overlap between the prostate contours on US_{tx} and CT_{tx} were expressed in DSC. It is important to mention that the DSC is always calculated with respect to the overlapping parts of the US volumes, not the total available volume.

In addition to the DSC, the gamma (γ) index [14] was proposed as a second metric for evaluation. This index is typically based on two acceptance criteria: distance-to-agreement [DTA] and dosimetric difference. The γ index is commonly used to evaluate discrepancies between measured and calculated dose on a voxel scale, while taking into account the neighboring voxels as well. An index value $\gamma > 1$ indicates that the corresponding voxels in both volumes exceed the maximum acceptable dosimetric difference/distance-to-agreement according to the criteria set.

In this work, we propose to use the same concept to assess the similarity between whole CT volumes at a voxel level. Therefore, the γ index was based on Hounsfield units (HU) instead of dose. An evaluation algorithm [15] developed in-house was used. As negative HU cannot be input in the software, the CT image data were first normalized by increasing all voxel values with the absolute value of the minimum negative HU present in the dataset.

Prior to γ index calculations, acceptance criteria need to be set. According to a survey performed by Nelms *et al.* [16], in most institutions that perform these calculations a 3 mm DTA criterion is used. In our case, the DTA was also set to 3 mm.

In this study, the focus was on prostate cancer patients and for this reason the second acceptance criterion (HU difference) was based on in vivo prostate HU as determined by D'Souza *et al.* [17]. They used several human subjects to determine that in case a voxel has an intensity value within 45 ± 17 HU it can be considered a prostate voxel. To translate this value to a HU difference criterion, it was first normalized using the same approach as described before. The minimum absolute HU value was added and this resulted in 1069 ± 17 HU, which implies $1069 \pm 2\%$ HU. For this reason, the γ index calculations were performed using a 3 mm DTA and 2% HU difference criterion.

The last evaluation method involves the assessment of bone edge deformation. The deformation field created with the different approaches was applied on the previously defined $Bone_{sim}$ (binary mask of bones extracted from CT_{sim}) resulting in a deformed binary bone mask ($Bone_{def}$). If the bone edges only underwent translation and/or rotation, it should be possible to map $Bone_{def}$ back to $Bone_{sim}$ just using rigid registration.
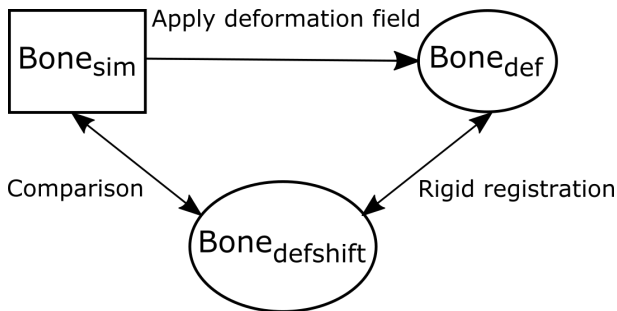
Figure 7.4: Workflow of bone rigidity assessment. First the created deformation field was applied on $Bone_{sim}$ resulting in $Bone_{def}$. If bone edges did not deform during registration, it should be possible to rigidly map $Bone_{def}$ back to $Bone_{sim}$. The rigid registration of $Bone_{def}$ results in $Bone_{defshift}$, which is, in the end, compared to $Bone_{sim}$.

The parameter file used for this registration is detailed in Appendix D.2. The rigid registration results in $Bone_{defshift}$ and subsequently the overlap between $Bone_{defshift}$ and $Bone_{sim}$ are expressed with a DSC (Fig. 7.4). This was a measure for how much the bone edges deformed during the deformable registration procedure. A DSC value of 1 indicates perfect overlap and therefore implies that no non-rigid deformation of the bone edges occurred.

7.2.7 Parameter set selection

Most of the parameters that are required to perform the deformable registration are directly related to the task that needs to be performed. For example, a specific value needs to be assigned to the 'Transform' parameter to perform a B-spline transformation. On the other hand, some parameters, such as the maximum number of iterations, directly influence the performance of the registration. Unfortunately, it is hard to determine a priori which settings will produce good results. For this reason, multiple parameter sets for each data pair were tested per approach. Table 7.3 details the four parameters that could not be determined reliably a priori. These parameters are varied in the parameter set selection procedure, exploring the values listed in Column 2 of Table 7.3. During this process, the other parameters were kept fixed.

Table 7.3: Parameters which were varied during the parameter set selection, including the values that were explored during this selection procedure.

Parameters	Explored values
Metric	Sum of squared differences, mutual information or cross validation
Number of resolutions	1, 2, ... or 5
Maximum number of iterations	5, 15, ... or 305
Final grid spacing (mm)	10, 15, ... or 35

In the parameter set selection procedure, for each data pair at least 50 different parameter sets were used per approach that involved deformable registration. Each set was used to perform the deformable registration, followed by calculation of the overlap of the prostate contours on US_{deform} and US_{tx} and calculation of the determinant of the Jacobian. If this determinant was smaller than zero, the result was excluded from further evaluation and an additional parameter set was randomly selected and tested.

After performing the registration per approach at least 50 times, the parameter set that resulted in the largest overlap of the prostate contours on US_{deform} and US_{tx} for that specific data pair was selected and the corresponding deformation field was used to continue the workflow with. These parameter sets can be found in Appendix D.3.

7.3 Results

Each of the eight approaches resulted in the creation of a deformation field that was subsequently applied on $CT_{sim\_contours}$. The resulting $CT_{ps\_contours}$ were then compared with $CT_{tx\_contours}$ using the DSC. These results are listed in Table 7.4. In the second column, the reference situation is detailed, defined as the overlap of $CT_{sim\_contours}$ and $CT_{tx\_contours}$ expressed in DSC. For each data pair, the result of the approach(es) leading to the highest improvement is displayed bold.

Next to the overlap of the contours, also the deformation of bones was assessed. Following the procedure described in Section 7.2.6, the rigidity of the bone edges is expressed in DSC per approach for each data pair. The last row of Table 7.4 contains the average bone rigidity DSC per approach. As noted before, a DSC of 1 implies that no non-rigid deformation of the bone edges occurred during the registration procedure.

After the deformation field application, some approaches (e.g. deformable registration only) based on the total US volumes for some data pairs failed to result in an improvement of the overlap. For each data pair, one of the approaches led to an improvement of at least 4.3%. The best improvements were as high as 126.5%. In three out of ten data pairs the contour based translation followed by the contour based non-rigid deformation performed best. In all the other data pairs the difference between this approach and the best performing approach was not more than 0.05 DSC. In addition, the approaches that gave the best results in structure overlap, mostly fall in the 0.95 - 1 DSC range in terms of bone rigidity.

To evaluate the similarity of the CT volumes quantitatively, the deformation fields were applied to the total volumes. The calculation of the deformation field and the subsequent application performed on a standard PC (i5 CPU, 2.6 GHz, 4 GB RAM) took on average 98 seconds (range: 14 - 196 seconds). All calculation times per approach and data pair are detailed in Appendix D.4.

Table 7.4: For each studied data pair (column 1), the overlap of CT$_{contours}$ was calculated and expressed in DSC. The reference is defined as the overlap between CT$_{sim\_contours}$ and CT$_{tx\_contours}$. In the remaining columns the overlap between CT$_{tx\_contours}$ and CT$_{ps\_contours}$, resulting from one of the performed approaches (prostate contour [C] and/or US volume [US] based) is listed. Next to the DSC value, also the relative improvement with respect to the reference situation was calculated and listed. The bold numbers indicate which approach(es) result in the highest improvement with respect to the reference. In the final row, the bone rigidity results are displayed as the average DSC per approach.

Data pairs	Ref	Translation [US]		Translation [C]		Deform [US]		Deform [C]		Translation [US] Deform [US]		Translation [US] Deform [C]		Translation [C] Deform [US]		Translation [C] Deform [C]	
	DSC	DSC	+%	DSC	+%	DSC	+%	DSC	+%	DSC	+%	DSC	+%	DSC	+%	DSC	+%
T1-T4	0.50	0.70	40.0	0.71	42.0	0.63	26.0	0.69	38.0	0.68	36.0	0.73	46.0	0.74	48.0	**0.75**	50.0
T1-T5	0.34	0.67	97.1	0.68	100.0	0.59	73.5	0.70	105.9	0.64	88.2	**0.77**	126.5	0.61	79.4	**0.77**	126.5
T2-T4	0.69	**0.81**	17.4	0.78	13.0	0.60	-13.0	0.78	13.0	0.74	7.2	0.77	11.6	0.72	4.3	0.76	10.1
T2-T5	0.50	0.45	-10.0	0.68	36.0	0.22	-56.0	0.55	10.0	0.35	-30.0	**0.71**	42.0	0.55	10.0	0.70	40.0
T3-T5	0.59	**0.76**	28.8	0.75	27.1	0.73	23.7	0.71	20.3	**0.76**	28.8	0.74	25.4	0.75	27.1	0.74	25.4
T4-T1	0.50	0.70	40.0	0.71	42.0	0.46	-8.0	**0.77**	54.0	0.76	42.0	0.75	50.0	0.68	36.0	**0.77**	54.0
T4-T2	0.69	**0.79**	14.5	**0.79**	14.5	0.64	-7.3	0.75	8.7	0.77	11.6	0.78	13.0	0.77	11.6	0.78	13.0
T5-T1	0.34	0.68	100.0	0.47	38.2	0.57	67.7	0.66	94.1	0.41	20.6	**0.75**	120.6	0.61	79.4	0.74	117.7
T5-T2	0.50	0.36	-28.0	**0.70**	40.0	0.46	-8.0	0.67	34.0	0.46	-8.0	0.53	6.0	0.57	14.0	0.67	34.0
T5-T3	0.59	0.74	25.4	**0.76**	28.8	0.75	27.1	0.62	5.1	0.74	25.4	0.73	23.7	0.75	32.2	0.75	27.1
Bone	-	1.00 ± 0.00		1.00 ± 0.00		0.82 ± 0.12		0.92 ± 0.05		0.84 ± 0.10		0.95 ± 0.05		0.86 ± 0.07		0.95 ± 0.04	

Table 7.5: For each data pair, the number of voxels with $\gamma > 1$ in the corresponding overlapping parts of the US volume was calculated for each approach (prostate contour [C] anor US volume [US] based) and expressed in percentages. The values in the reference column were based on CT$_{sim}$ and CT$_{tx}$. Next to the relative number of non-matching voxels, also the relative improvement with respect to the reference situation was calculated. The bold numbers indicate which approach(es) result in the highest improvement with respect to the reference. In the final row, the bone rigidity results are displayed as the average DSC per approach.

Data pairs	Ref	Translation [US]		Translation [C]		Deform [US]		Deform [C]		Translation [US] Deform [US]		Translation [US] Deform [C]		Translation [C] Deform [US]		Translation [C] Deform [C]	
	Fail [%]	Fail [%]	Δ [%]	Fail [%]	Δ [%]	Fail [%]	Δ [%]	Fail [%]	Δ [%]	Fail [%]	Δ [%]	Fail [%]	Δ [%]	Fail [%]	Δ [%]	Fail [%]	Δ [%]
T1-T4	34.4	22.8	33.7	21.4	37.8	28.2	18.0	25.9	24.7	26.0	24.4	19.3	43.9	23.4	32.0	**17.8**	48.3
T1-T5	35.2	21.2	39.8	21.6	38.6	23.2	34.1	22.0	37.5	21.2	39.8	**16.7**	52.6	22.1	37.2	17.9	49.1
T2-T4	34.7	**25.2**	27.4	26.8	22.8	36.8	-6.1	28.4	18.2	29.5	15.0	27.3	21.3	30.2	13.0	27.0	22.2
T2-T5	38.6	40.0	-3.6	26.8	30.6	45.4	-17.6	38.3	0.8	41.8	-8.3	27.5	28.8	28.5	26.2	**25.3**	34.5
T3-T5	34.8	**20.3**	41.7	20.6	40.8	29.6	14.9	30.5	12.4	21.4	38.5	20.4	41.4	23.5	32.5	20.7	40.5
T4-T1	43.3	33.4	22.9	31.0	28.4	40.9	5.5	29.8	31.2	33.6	22.4	30.5	29.6	30.7	29.1	**29.0**	33.0
T4-T2	34.7	25.2	27.4	**23.5**	32.3	35.8	-3.2	33.1	4.6	23.6	32.0	25.7	25.9	24.8	28.5	25.6	26.2
T5-T1	39.2	24.6	37.2	36.3	7.4	26.5	32.4	33.6	14.3	32.8	16.3	**24.0**	38.8	28.8	26.5	25.9	33.9
T5-T2	34.4	38.8	-12.8	20.4	40.7	38.1	-10.8	26.1	24.1	37.6	-9.3	35.0	-1.7	31.7	7.8	**20.2**	41.3
T5-T3	38.1	22.9	39.9	**22.6**	40.7	31.2	18.1	38.0	0.3	29.6	22.3	22.9	39.9	26.7	29.9	23.4	38.6
Bone	-	1.00 ± 0.00		1.00 ± 0.00		0.82 ± 0.12		0.92 ± 0.05		0.84 ± 0.10		0.95 ± 0.05		0.86 ± 0.07		0.95 ± 0.04	

After application of the deformation field, the corresponding γ indices were calculated. As an example, the outcome for data pair T1 - T4 is shown in Fig. 7.5A and Fig. 7.5B. As noted before, yellow voxels do not match within the previously defined acceptance criteria. Comparison of CT_{ps} with CT_{tx} clearly results in less yellow voxels, than if CT_{sim} is compared to CT_{tx}.

(A) **(B)**
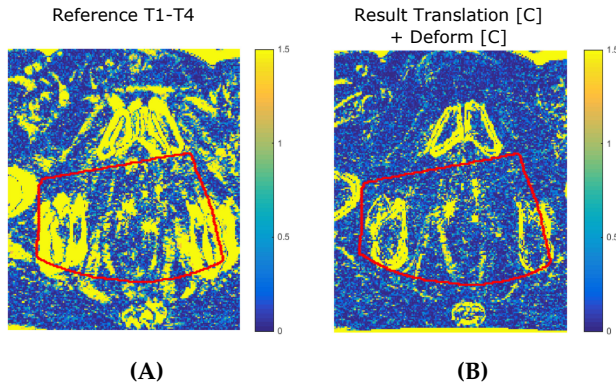
Figure 7.5: (A) γ index between CT_{sim} and CT_{tx} (reference) of data pair T1 - T4. The red contour indicates the overlapping parts of the co-registered US volumes. A green-yellow voxel indicates $\gamma > 1$, so these voxels of the compared volumes do not match within the acceptance criteria. (B) γ index between CT_{ps} and CT_{tx}.

For each data pair, the number of voxels with $\gamma > 1$ with respect to the total US volume (red contours in Fig. 7.5A and 7.5B) expressed in percentages, as well as the relative improvement (decrease of the relative number of yellow voxels) with respect to the reference, were calculated and listed in Table 7.5.

7.4 Discussion

This study has shown that it is possible with at least four out of the eight proposed approaches to create a pseudo-CT volume for each pair of co-registered CT-US volumes which, according to the used evaluation criteria, represents the patient anatomy at TX better than at CT_{sim}. All the approaches that failed to result in an improvement required the use of total US volumes (and not only the prostate contours) for either translation and/or deformable registration. In addition, the approaches failed for all data pairs in which data from Time point 2 were involved. This suggests that the data at Time point 2 might contain artifacts that influence the registration procedures. Unfortunately, visual inspection of the data did not reveal any clear issues.

A combination of initial alignment followed by deformable registration produced the best results. When this initial alignment was based on the prostate contour only, the assumption was made that the translation of the prostate was representative

for the translation of the whole body. This can be justified by the assumption that the motion of the total body could be considered one order of magnitude larger than the motion of the prostate. Possible (relatively small) translation of the prostate with respect to the body could be corrected by subsequent deformable registration. Performing the translation in a case where actually no translation of the whole body occurred might introduce the same type of errors, which can be fixed with the deformable registration afterwards. Future research on a larger patient population will show if initial alignment is still necessary and if including only lateral direction is sufficient.

Even better results might be achieved when different US scanning techniques are used for US image acquisition, as this may improve the quality of the global information available from the US volumes. For example, transperineal US for prostate imaging has a larger field of view (FOV) than TAUS and usually produces images with a better contrast. This US scanning technique is already included in the Autoscan system of Elekta (Stockholm, Sweden). We hypothesize that the performance of the approaches involving total US volumes might be better when these transperineal images are used.

The evaluation of the similarity of the CT_{ps} and CT_{tx} volumes based on anatomical structures and HU proposed in this work may provide a valuable insight on which CT_{ps} volumes are acceptable. The ultimate evaluation, though, is to perform dose calculations to assess what the use of CT_{ps} instead of CT_{sim} images would mean in terms of radiation dose deposition.

In this study, the γ index was used to perform the whole CT evaluation. Depending on the acceptance criteria set, the index seemed to be able to quantify differences in soft tissue. Unfortunately, bone similarity could not be assessed with the same acceptance criteria. Neighboring bone voxels can initially differ more than the set 2% in HU. So even when there is a misalignment of less than one voxel position, this might already be not accepted. This phenomenon could also explain the discrepancy between the differences in improvement between the DSC and γ index, which sometimes resulted in different recommendations for the best approach. In the ideal case, it would be possible to assign different acceptance criteria to bone regions and soft tissue.

Next to the similarity of the CT volumes, also bone deformation was assessed. For most data pairs (80%) the best performing approach had a bone rigidity between 0.95 and 1 DSC. ElastiX has a rigidity penalty [18] parameter available to constrain certain regions (such as bones) to only undergo rigid influence of the registration of the surrounding tissue. The bone rigidity results show that the bone edges did not undergo deformation. Constraining bone regions could also result in a worse overlap of anatomical structures. For these reasons, no constraints were taken into account during the registration procedure in this study. If in future studies more deformation of the bones will be detected, dose calculations can reveal what the

effect is on the dose deposition and if the use of the rigidity penalty would be beneficial.

In general, the limited FOV of US with respect to other imaging modalities such as CT, is one of the issues connected with registration based on US volumes. Only information inside the US volume is, in principle, used to create the deformation field during the registration. However, due to the multi-resolution approach some deformation field is available outside of the US volume as well (described in Section 7.2.5). During the registration procedure no correction for this was performed because, eventually, extrapolation of the deformation field should to some extent be performed to deform the whole CT volume. This is important because changes in tissue distribution may occur outside the US volume as well. The earlier mentioned transperineal US has a larger field of view and might therefore be suitable to create a larger deformation field. In addition, a deformation field extrapolation procedure could be implemented, taking into account the mechanical properties of the tissue and organs, such as skin, bones and bladder, which are positioned outside the US volume.

7.5 Conclusion

The purpose of this study was to investigate the differences between created pseudo-CT images and reference CT images, using various approaches for deformation field calculation. In addition, evaluations were performed to assess how well these images represent the patient at TX. Similarity analysis of CT_{ps} and CT_{tx} showed that, for five out of eight data pairs, all approaches resulted in an improvement of the similarity with respect to the CT_{sim}. For the remaining pairs four up to seven approaches resulted in an improvement. On average, the approaches that involved translation followed by deformable registration gave the best results. The promising results warrant more research on different US methods and parameter set optimization techniques in a larger patient group. In addition, the research could be extended to other cancer sites, such as liver.

References

[1] J. Roeske, J. Forman, C. Mesina, T. He, C. Pelizzari, E. Fontenla, S. Vijayakumar, and G. Chen. "Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy". In: *International Journal of Radiation Oncology Biology Physics* 33.5 (1995), pp. 1321–1329.

[2] D. Yan, E. Ziaja, D. Jaffray, J. Wong, D. Brabbins, F. Vicini, and A. Martinez. "The use of adaptive radiation therapy to reduce setup error: a prospective clinical study". In: *International Journal of Radiation Oncology* Biology* Physics* 41.3 (1998), pp. 715–720.

[3] M. Uematsu, T. Fukui, A. Shioda, H. Tokumitsu, K. Takai, T. Kojima, Y. Asai, and S. Kusano. "A dual computed tomography linear accelerator unit for stereotactic radiation therapy: a new approach without cranially fixed stereotactic frames". In: *International Journal of Radiation Oncology•Biology•Physics* 35.3 (1996), pp. 587–592.

[4] S. van der Meer*, S. Camps*, W. van Elmpt, M. Podesta, P. Sanches, B. Vanneste, D. Fontanarosa, and F. Verhaegen. "Simulation of pseudo-CT images based on deformable image registration of ultrasound images: A proof of concept for transabdominal ultrasound imaging of the prostate during radiotherapy". In: *Medical Physics* 43.4 (2016), pp. 1913–1920.

[5] D. Fontanarosa, S. Van der Meer, J. Bamber, E. Harris, T. O'Shea, and F. Verhaegen. "Review of ultrasound image guidance in external beam radiotherapy: I. Treatment planning and inter-fraction motion management". In: *Physics in Medicine & Biology* 60.3 (2015).

[6] X. Pennec, P. Cachier, and N. Ayache. "Tracking brain deformations in time sequences of 3D US images". In: *Pattern Recognition Letters* 24.4-5 (2003), pp. 801–813.

[7] G. Janssens. "Registration models for tracking organs and tumors in highly deformable anatomies: applications to radiotherapy". PhD thesis. Universite catholique de Louvain, Sept. 2010.

[8] S. Klein, M. Staring, K. Murphy, M. Viergever, and J. Pluim. "Elastix: a toolbox for intensity-based medical image registration". In: *IEEE transactions on medical imaging* 29.1 (2010), pp. 196–205.

[9] D. Shamonin, E. Bron, B. Lelieveldt, M. Smits, S. Klein, and M. Staring. "Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease". In: *Frontiers in neuroinformatics* 7 (2014), p. 50.

[10] S. Klein and M. Staring. *Elastix 4.7 - Manual*. 62nd ed. 2014.

[11] S. Chun and J. Fessler. "A simple regularizer for B-spline nonrigid image registration that encourages local invertibility". In: *IEEE Journal of selected topics in Signal Processing* 3.1 (2009), pp. 159–169.

[12] T. Sørensen. "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons". In: *Biol. Skr.* 5 (1948), pp. 1–34.

[13] L. Dice. "Measures of the amount of ecologic association between species". In: *Ecology* 26.3 (1945), pp. 297–302.

[14]   D. Low, W. Harms, S. Mutic, and J. Purdy. "A technique for the quantitative evaluation of dose distributions". In: *Medical physics* 25.5 (1998), pp. 656–661.

[15]   L. Persoon, M. Podesta, W. van Elmpt, S. Nijsten, and F. Verhaegen. "A fast three-dimensional gamma evaluation using a GPU utilizing texture memory for on-the-fly interpolations". In: *Medical physics* 38.7 (2011), pp. 4032–4035.

[16]   B. Nelms and J. Simon. "A survey on planar IMRT QA analysis". In: *Journal of applied clinical medical physics* 8.3 (2007), pp. 76–90.

[17]   W. D'Souza, E. Madsen, O. Unal, K. Vigen, G. Frank, and B. Thomadsen. "Tissue mimicking materials for a multi-imaging modality prostate phantom". In: *Medical physics* 28.4 (2001), pp. 688–700.

[18]   M. Staring, S. Klein, and J. Pluim. "A rigidity penalty term for nonrigid registration". In: *Medical physics* 34.11 (2007), pp. 4098–4108.

8 General discussion

For this PhD work, I have investigated how the ultrasound (US) guided radiotherapy (RT) workflow for prostate cancer patients can be improved. In particular, there were two aspects of the workflow that in my opinion had to be optimized, because they were the major causes why this guidance technique is currently not widespread in RT: the process is operator dependent and it does not provide enough information to adapt or re-plan the treatment in case errors are detected. These aspects are covered in the three different parts of this thesis and include proof-of-concept and feasibility studies, as well as more mature studies. The first part (**Chapter 3** and **Chapter 4**) describes the proposition of an approach to automatically provide an operator with a patient-specific US probe setup. In the second part (**Chapter 5**), an automatic quality assessment procedure for 2D US images using deep learning is detailed. In the third part (**Chapter 6** and **Chapter 7**), the creation of pseudo-CT scans that can provide crucial electron density information for adaptive dose re-calculations is described.

In all three parts of this thesis promising results have been reported. Nevertheless, a lot of work still needs to be done to make US image acquisition, interpretation and utilization integrated in the RT workflow as well as some other available imaging modalities. In the next sections, the parts of this thesis are thoroughly discussed while pointing out contributions, limitations and recommending future work. In addition, a future perspective is given on the use of US guidance in RT workflows in general and in prostate cancer RT workflows specifically. Finally, a brief future outlook on US guidance during medical interventions other than RT is given.

8.1 Overview and discussion of thesis chapters

Automatic US probe setup

In current clinical practice, an operator needs to interpret the live US images, while positioning the US probe on the body of the patient, in order to understand if the correct anatomical structures are visualized with sufficient quality. This requires a significant amount of training with a steep learning curve. If the operator was provided with an (initial) patient-specific US probe setup, this could remove part of the need for training and potentially decrease operator

dependence. The phantom study and preliminary patient study described in **Chapter 3** and **Chapter 4** of this thesis, respectively, detailed a possible approach to automatically provide the operator with this patient-specific probe setup.

In the proof-of-concept study, several possible setups were calculated for a phantom, based on an acquired CT scan and reproduced using a robotic arm. The visual inspection of the subsequently acquired US volumes implied that this algorithm is able to propose probe setups that fulfill all clinical requirements. Therefore the decision was made to move onto patient data. In the patient study, the probe setups were automatically calculated based on simulation CT scans of the individual patients. These patient-specific probe setups were retrospectively compared with the setups used by a radiation oncologist during the actual image acquisition. The results obtained for three prostate cancer patients showed that it seems possible to propose a patient-specific probe setup that could allow visualization of the required anatomical structures. This patient study was a proof-of-concept and is an initial step on the path towards fully automatic proposition of patient-specific US probe setups for prostate cancer patients.

In order to draw final conclusions on the ability of the algorithm to propose a suitable setup, a prospective analysis should be performed next. In that case, a probe setup should be determined, after which the operator should replicate this setup on the body of the patient, acquire an US image and evaluate if indeed the correct anatomical structures are visualized. To be able to do this, the development of a probe fixation system including a system that can localize the probe in the room is crucial. This was not in the scope of this PhD work. Nevertheless, in the phantom study a first attempt was made to use a robotic arm for this purpose, but due to unexpected motion of the arm this could not be used in the patient study as well. For probe positioning on the body of the patients, a mechanical arm was used instead. However, also this prototype arm should be developed further, in order to allow the operator to reproduce the proposed probe setup precisely on the body of the patient.

After the confirmation that the algorithm can indeed propose suitable probe setups, the focus should be on how this principle can be effectively implemented in the RT workflow. The approach described in **Chapter 3** and **Chapter 4** requires positioning of the US probe after the acquisition of the CT scan. In the ideal case, the CT scan and US volume would be perfectly registered. Unfortunately, positioning the probe after the CT scan would most probably result in motion and so in a sub-optimal registration of the US and CT images. In future work, it is important to quantify this motion and to understand how this could influence the radiation dose deposition. If the motion would lead to unacceptable dose deviations, the proposed approach described in this thesis could be modified in several ways.

A possible solution could be the use of a scout-CT scan instead of the simulation CT scan. A scout-CT scan usually has less slices than a regular CT scan and it is

acquired in some medical centers before the simulation CT scan to check for air in the rectum. Potentially, the anatomical information captured in this scout-CT scan is sufficient to propose a suitable probe setup. Another solution could be the use of CT scans and probably also MR scans that have been acquired for diagnosis. This could then yield an initial probe setup that subsequently should be fine-tuned when the probe is positioned on the body of the patient. The live US images that are available as soon as the probe is in place, can provide anatomical information for this fine-tuning. In order to allow a relatively untrained operator to perform this fine-tuning, automatic interpretation of these live images is key. In **Chapter 5** already a first step towards this automatic interpretation of US images is made and in the next section of this discussion, this chapter is discussed in further detail.

The algorithm introduced in **Chapter 4** made assumptions including, for example, how bony structures would restrict the visualization of subsequent anatomical structures. In addition, the algorithm followed a relatively straightforward approach to propose patient-specific probe setups. Nevertheless, promising results were obtained. One could wonder if a more precise calculation of the probe-setup would, in the end, be worth the effort. Unless the time between the acquisition of the CT scan that is used for this setup calculation and the US volume acquisition is very short, anatomical structures might move and therefore the calculated probe setup might not be optimal anymore. Potentially, aiming for a setup that positions the anatomical structures of interest in the center of the field of view of the probe is a better approach than aiming for the optimal setup. This could allow for some displacements before the structures would escape the field of view. This approach would only require a small modification in the final assessment procedure of the possible probe setups in the current algorithm and should therefore be straightforward to realize. In case a more precise probe-setup calculation turns out to be necessary, several options could be investigated, including simulation of US images with (e.g. [1]) or without the use of deep learning strategies (e.g. [2] and [3]) and the use of 3D surface mapping strategies (e.g. [4]) to retrieve the perineal skin area on which the probe can be positioned in clinical practice.

Automatic image quality assessment

As noted in the previous section, a significant amount of training is required to be able to interpret US images. This does not only include interpretation during the setup of a US probe, but also during the subsequent continuous US imaging. In order to use these US images for radiation dose guidance, specific information needs to be extracted from them by an operator. This information could be, for example, the location of the treatment target or the location of an organ at risk. Ultimately, the operator would only be presented with the already extracted information, for example, in the form of anatomical structure segmentations, and not with the underlying image data. In this way, operator dependence can probably be reduced and the usability of US imaging in RT workflows improved.

Before this ultimate goal can be achieved, there are several challenges that need to be faced first, one of them being quality assessment of US images. Small motion of the patient or of his anatomical structures during the RT treatment might compromise the US image quality. This could result in US images that cannot be used for reliable dose guidance anymore. Currently, this quality assessment needs to be done by a trained operator, who might not be available for the whole treatment fraction of every patient.

In **Chapter 5** of this thesis, an exploratory study is described which was performed to understand if it is feasible to train a deep learning algorithm so it can automatically assign a quality score to 2D TPUS images. To the best of my knowledge, just a few publications were available on US image quality assessment in general when this study was started, which mostly assessed fetal images (e.g. [5] and [6]) or made use of hand crafted features (e.g. [7]). Therefore, a feasibility study was justified. The results achieved in this study, show that the used deep learning algorithm can achieve a higher accuracy (94%) than the worst performing expert (92%). This does not only prove feasibility of quality assessment by means of deep learning, but it also warrants further improvements of the currently implemented approach and exploration of other deep learning algorithms.

In order to make the described approach more suitable for usage during the current RT workflow of prostate cancer patients, there are two main aspects that need attention. First, the quality of a whole 3D TPUS volume is more relevant than the quality of a single 2D image, as the US volume is more suitable to be used for tracking of the prostate and the surrounding organs at risk. However, the assessment of a volume instead of a 2D image brings along several challenges. In this study, even though the algorithm was executed on a GPU, it sometimes ran out of memory. So, in order to process an even larger database consisting of volumes, more calculation power would be required and/or the focus should lie on developing a more memory efficient deep learning algorithm. Also, the current database consisted of 13,463 2D images, which were picked from 16,000 evaluated images, which were again part of a much larger database of 178,368 images (resulting from 11,148 volumes). If the volumes had been used instead of the images, the largest database would have been only 11,148 volumes. It is less labor intensive and time consuming to acquire sufficient 2D images with a 3D probe to train and evaluate the algorithm, than it is to acquire sufficient 3D volumes for the same purpose. Finally, the experts would also need to assign a quality score to a whole volume instead of to single images, and establishing the criteria for the different quality scores would, most probably, be more challenging in three dimensions than it is for a single slice of the volume.

Second, currently only images acquired by four radiation oncologists with an X6-1 xMatrix probe (Philips Healthcare, Bothell, WA, USA) were included in the database. However, the deep learning algorithm will probably make mistakes

when it is presented with a type of images that it has never seen before. For this reason, it would be desired to expand the database with TPUS volumes acquired by a broader range of clinicians, in order to make the algorithm more robust.

The use of deep learning could be of great potential for US image interpretation and the promising results reported in **Chapter 5** may only be just the beginning. In addition to the binary image quality assessment, there are several other aspects of US image interpretation in which deep learning could potentially be of help. One could think of, for example, automatic segmentation (e.g. [8] and [9]), automatic assessment in which direction the US probe should be moved to achieve a better visualization of anatomical structures of interest and the automatic tracking of relevant anatomical structures [10]. In addition, deep learning could also be beneficial in US beamforming (e.g. [11]–[13]) allowing (partial) elimination of image artifacts. However, with the currently used supervised learning method, which involves learning from ground truth labels, this would require a lot of input from the experts in the form of, among others, quality scores and manual delineations of anatomical structures, so these can be used to train the algorithm. In addition, the experts should somehow come to an agreement on what would be the correct quality score or where the contour should be drawn. In particular for US imaging, achieving such an agreement is not trivial.

For this reason, the developments in weakly supervised and unsupervised deep learning should be followed closely, as this might be part of the solution. In this type of deep learning, the algorithm is not provided with a pixel level annotated image for training, but just with course object locations (e.g. bounding boxes) or an indication if a certain anatomical structure is present in the image [14] (weakly supervised) or the algorithm is not provided with any kind of ground truth label (unsupervised). These approaches seem worth investigating further in the medical domain, as interesting results have been shown already for, for example, automatic lesion detection on brain MR scans, multiple organ detection on 4D patient data and breast density segmentation [15]–[17].

Pseudo-CT scan creation

After the acquisition and interpretation of the US images, the next step is the usage of these images in the RT workflow. It has been extensively covered in several parts of this thesis that the aim of RT treatment is to irradiate tumor tissue, while sparing normal tissue as much as possible. For this reason, final dose distributions are the most relevant metric to measure efficacy of US guidance during the course of the treatment. In order to calculate these distributions, electron density information (and so the CT scans that can provide this information) are of interest. As frequent CT imaging during the course of the treatment is not part of the standard clinical routine, we propose an approach in which US images are used to determine the required electron density information.

In **Chapter 6** and **Chapter 7** of this thesis an approach is described to create pseudo-CT scans using (deformable) image registration of transabdominal US (TAUS) volumes. The presented results show that several registration approaches can result in the creation of a pseudo-CT scan that anatomically represents the patient at treatment stage better than the simulation CT scan does. In addition, some preliminary dosimetric evaluations have been performed. However, it is important to investigate further what could be the exact benefit of such a pseudo-CT scan creation approach for adaptive dose recalculation. Especially, because, for example, Fraser *et al.* [18] have reported that the amount of under-dosage or over-dosage of an anatomical structure is not directly proportional to the magnitude of the occurred shifts of that anatomical structure. Therefore, the relatively limited dosimetric evaluations could be considered limitation of this study and should definitely be taken care of in the future, preferably looking at hypo-fractionation [19] schemes as well.

The US image data used for the pseudo-CT scan creation was acquired with the transabdominal probe of a Clarity system (Elekta, Stockholm, Sweden). The usage of this transabdominal probe for prostate imaging for RT guidance is not optimal due to the limited field of view, the requirement for a full bladder, possible probe pressure effects (e.g. [20] and [21]) and the presence of the probe in the radiation beam path. For this reason, it would be interesting to repeat the same work using the newer Clarity Autoscan system (Elekta, Stockholm, Sweden) which involves US imaging with a TPUS probe which does not, or to a lesser extent, suffer from these disadvantages. In order to compare created pseudo-CT scans with ground truth CT scans, acquisition of several ground truth scans from the same patient over the course of the treatment would be required. Multiple CT scans expose the patient to additional radiation and are therefore not part of the standard clinical routine. Usually an additional CT scan is only acquired when there are clear clinical implications that the treatment plan requires re-planning, such as severe weight loss or gain. Therefore, it could be challenging to get a study approved by a medical ethics committee that purposely involves the acquisition of multiple CT scans. However, as the currently achieved results are promising, this might convince the committee that the potential benefits outweigh the radiation exposure risk.

The performance of the algorithms used in this thesis for registration of the US volumes is still highly dependent on the used parameters, such as number of iterations and the similarity metric. Unfortunately, it is not yet possible to determine a priori which parameter set would yield good results. Also, the optimal parameter set might differ for individual image data sets. Nowadays, the use of deep learning approaches for several types of image processing problems is expanding very fast. Potentially, a way to eliminate the influence of these parameters can be found in this area of research. Several papers have been published on performing image registration by means of deep learning, such as [22], [23] and [24]. Unfortunately, no work seems to be published yet on the

registration of US volumes. Due to the peculiar characteristics of US imaging this is a harder problem to tackle than, for example, registration of MR images. However, as the deep learning field keeps expanding, this could rapidly change in the near future.

Two other deep learning approaches that could possibly be of use for the creation of pseudo-CT scans are the use of style transfer or generative adversarial networks (GANs). A style transfer approach renders the content of an image in a different style. For example, a picture of houses along a canal can be rendered in the style of a painting of Vincent van Gogh, as shown in [25] and [26]. The same kind of principle could potentially be applied to US images in order to render them in the style of a CT scan. However, no publications in this direction seem to be available in literature and currently, the only evaluation metric seems visual judgment if the style "looks like" the reference style [26]. Therefore, it seems difficult to judge the accuracy of the algorithm and so the potential usability in medical imaging. The use of GANs ([1], [27], [28]), however, could be very promising.

Created pseudo-CT scans that represent the anatomy of the patient well, can provide more insight into if treatment re-planning could be beneficial for that specific patient. In addition, these scans could potentially even be used for the actual re-planning as well. To design a treatment plan for a prostate cancer patient, electron density information of a substantial part of the abdomen is usually required, in order to calculate the optimal radiation beam paths. The field of view of an US probe is currently not big enough to cover a sufficient part of the abdomen and so to create a pseudo-CT scan that covers a substantial part of the abdomen. However, in the current registration implementation, already some deformation is estimated outside of the US volume. Potentially this could somehow be extended using assumptions on the deformable characteristics of the surrounding tissues. Another solution could be the use of image compounding as has been done by, for example, [29] and [30]. This would involve the acquisition of several US volumes, for example, from different angles and compounding them, resulting in an US image with a bigger field of view and also an improved signal to noise ratio.

8.2 Future perspective

Developments in immunotherapy treatment for several types of cancer have shown very promising results, even for patients with metastatic cancer (e.g. [31]–[33]). Intuitively this type of treatment, during which the patient's own immune system is assisted and encouraged to fight the cancer by itself, seems to make more sense than attempting to destroy the individual cancer cells using chemotherapy or RT treatment. Therefore, I am wondering if the use of radiation for ablative cancer treatment, as we currently know it, will still be considered a viable treatment option in 5, 10, or 20 years from now.

RT has always been seen as an effective, but local cancer treatment option [34]. Surprisingly, over 20 case reports have been published in which tumor regression is reported outside of the regions that have been treated with RT [35]. The underlying mechanisms of this so-called abscopal effects have not been fully understood yet. However, there is a suspicion that this effect might contribute to the enhanced immune response [35] when the patient is treated with a combination of RT and immunotherapy, which is reported in several publications (e.g. [36] and [37]).

As the human immune system is extremely complex, it might take decades and a lot of research to understand how RT and immunotherapy can be combined most effectively and what kind of role (US) image guidance can play in this. In the meantime, men will keep receiving a prostate cancer diagnosis, undergo RT treatment and so they can potentially benefit from US image guidance already. Therefore, in the remainder of this discussion, an outlook in the closer future on US guidance during RT in general and for prostate cancer patients specifically is given. Finally, US guidance can also be of help during other medical interventions than RT treatment. For this reason, also a brief outlook on US guidance for other medical interventions is given.

US guidance in RT in general

The use of US imaging for dose guidance in RT workflows is not only limited to prostate RT workflows, but has been applied to other RT workflows, such as liver and breast, as well. However, no matter which anatomical structure is imaged using US imaging, most of the challenges associated with this image modality, such as the need for a trained operator, the need for image interpretation during the acquisition and inaccessibility of anatomical structures shielded by bone, will still occur. Therefore, also for these other RT workflows including, for example, cervix [38], breast [39] and liver [40], research, such as anatomical structure tracking evaluation, is conducted to allow for US guidance during the course of the treatment. Joining forces and trying to come up with general solutions that could fit all these fields of research could speed up the developments and so make the use of US imaging more appealing.

Currently, there seems to be a focus on the potential use of an MR guided RT system, the MR-Linac ([41] and [42]). In this system a linear accelerator is integrated into an MR scanner, which could allow for MR guidance during the radiation dose delivery. Most clinicians are more familiar with MR scans than with US images and the workflow of this MR-Linac is most probably less operator dependent than an US guided workflow. These two factors could explain the growing interest for the MR-Linac [43]. However, in comparison with US imaging this system is way more expensive, not only including the cost of the system itself, but also the unavoidable replacement of the currently available Linacs. In addition, this magnetic field might influence the direction of the radiation beam. For example, Perik *et al.* [44] have shown that the magnetic fringe

field of an MR-Linac can influence the beam steering of the regular Linacs in adjacent bunkers. Even though this could be corrected for, this raises questions about the influence of the field on the beam of the MR-Linac itself and if there are any other (negative) influences that should be thoroughly investigated before the MR-Linac is introduced in the clinics at a larger scale.

Moreover, the number of proton therapy centers in the world is currently rapidly growing. This type of external beam RT makes use of a beam of protons instead of an X-ray beam. In contrast with more traditional X-ray irradiation, a single proton beam has a maximal dose at a user-defined depth and no exit dose [45]. These characteristics make it possible to reduce to radiation dose to normal tissue substantially, while preserving dose to the target. This results in a more favorable dose distribution in comparison with conformal photon therapy. However, with a maximal dose at a user-defined depth, the need for (image) guidance is crucial. The influence of the magnetic field of a MR scanner on the proton beam and the absence of an exit dose necessary for electronic portal imaging make these imaging modalities not suitable for usage during a guided proton therapy workflow. As US imaging does not suffer from these problems, it can therefore be a good candidate for proton therapy guidance.

US guidance for prostate RT

The motion of the prostate, both inter- and intra-fraction is relatively small in comparison with anatomical structures that are affected by breathing motion, such as liver and pancreas. Therefore one could reason that real-time US guidance is more relevant for structures that are affected by breathing motion. However, the dosimetric consequences of motion of the prostate and the organs at risk during the RT workflow have not been fully understood yet. Especially, in the hypo-fractionation schemes [19], during which more radiation dose is delivered in each (longer) treatment fraction, these anatomical structure motions could potentially result in significant dose deviations.

Also, with 1.6 million men being diagnosed with prostate cancer worldwide every year, the prostate cancer patient population is significantly larger than the number of patients diagnosed with, for example, liver (about 850,000 diagnoses) and pancreas (about 400,000 diagnoses) cancer [46]. This larger patient population could potentially make it easier to recruit patients for studies. For this reason, some proof-of-concept studies, that look into generic challenges of US imaging, such as image quality assessment, could be performed based on prostate patient data and then further fine-tuned for the smaller patient populations of other cancer sites.

Finally, as mentioned in the previous section, the use of proton therapy is growing, also for the treatment of prostate cancer (e.g. [47]). Due to the unsuitability of other image modalities in a proton environment, also for prostate proton therapy US guidance could be the image guidance approach of choice.

Summarizing, even though the anatomical structure motion in prostate RT is relatively small, further development of US guided RT workflows for prostate cancer patients is still justified, due to still unknown dose consequences, the potential in hypo-fractioned schemes and proton therapy, and the ability to transfer research on prostate to other cancer sites.

US guidance outside RT

There seems to be a growing realization in the medical community that US imaging can provide valuable (volumetric) information that can be used for navigation and guidance during a wide range of medical procedures. This image modality can provide a cost effective and harmless solution for this guidance during procedures in which guidance was absent before or which was cumbersome, expensive and/or potentially harmful for the patient. Procedures during which US guidance can be of use range from epidural needle placement [48] and breast conserving surgery [49] to lung biopsies [50].

Especially with the improving 3D, 4D and elastography capabilities which are currently available in the clinic, the range of applications might grow and the usage of US imaging for guidance might increase. Also, the developments towards faster image acquisition, higher frame rates, smaller US probes, stick-on probes, flatter probes and many others, can contribute to an increase in the popularity of US imaging.

However, as also noted in an earlier section, no matter which anatomical structure is imaged using US imaging, most of the common challenges (e.g. need for image interpretation, operator dependence and establishment of acoustic coupling) associated with this image modality will occur (up to a certain extent). Therefore, part of the research presented in this thesis, such as the automatic image quality assessment of US images, could not only be of use for US guidance in prostate cancer RT workflows, but also for other guidance procedures and potentially even for diagnostic US imaging as well.

In conclusion, in my opinion the use of US imaging, not only for guidance during (prostate) cancer RT workflows, but also for guidance during other medical procedures will keep increasing in the upcoming years. This increased usage can be boosted even further by tackling the generic and most predominant challenges associated with US imaging. The work presented in this thesis has contributed to the tackling of these challenges for US guidance in prostate RT workflows specifically and partially even for US guidance in general.

References

[1] Y. Hu, E. Gibson, L. Lee, W. Xie, D. Baratt, T. Vercauteren, and A. Noble. "Freehand ultrasound image simulation with spatially-conditioned generative adversarial networks". In: *Molecular Imaging, reconstruction*

and analysis of moving body organs, and stroke imaging and treatment. Springer, 2017, pp. 105–115.

[2] B. Burger, S. Bettinghausen, M. Radle, and J. Hesser. "Real-time GPU-based ultrasound simulation using deformable mesh models". In: *IEEE transactions on medical imaging* 32.3 (2013), pp. 609–618.

[3] M. Salehi, S. Ahmadi, R. Prevost, N. Navab, and W. Wein. "Patient-specific 3D ultrasound simulation based on convolutional Ray-tracing and appearance optimization". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2015, pp. 510–518.

[4] M. Krengli, S. Gaiano, E. Mones, A. Ballarè, D. Beldì, C. Bolchini, and G. Loi. "Reproducibility of patient setup by surface image registration system in conformal radiotherapy of prostate cancer". In: *Radiation Oncology* 4.1 (2009), p. 9.

[5] L. Zhang, N. Dudley, T. Lambrou, N. Allinson, and X. Ye. "Automatic image quality assessment and measurement of fetal head in two-dimensional ultrasound image". In: *Journal of Medical Imaging* 4.2 (2017).

[6] B. Rahmatullah, I. Sarris, A. Papageorghiou, and J. Noble. "Quality control of fetal ultrasound images: Detection of abdomen anatomical landmarks using adaboost". In: *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE. 2011, pp. 6–9.

[7] J. Schwaab, Y. Diez, A. Oliver, R. Martí, J. van Zelst, A. Gubern-Mérida, J. Mourri A.B.and Gregori, and M. Günther. "Automated quality assessment in three-dimensional breast ultrasound images". In: *Journal of Medical Imaging* 3.2 (2016).

[8] O. Ronneberger, P. Fischer, and T. Brox. "U-net: Convolutional networks for biomedical image segmentation". In: *International Conference on Medical image computing and computer-assisted intervention*. Springer. 2015, pp. 234–241.

[9] F. Milletari, S. Ahmadi, C. Kroll, A. Plate, V. Rozanski, J. Maiostre, J. Levin, O. Dietrich, B. Ertl-Wagner, K. Bötzel, et al. "Hough-CNN: deep learning for segmentation of deep brain regions in MRI and ultrasound". In: *Computer Vision and Image Understanding* 164 (2017), pp. 92–102.

[10] G. Carneiro and J. Nascimento. "Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data". In: *IEEE transactions on pattern analysis and machine intelligence* 99.1 (2013), p. 1.

[11]   A. Nair, T. Tran, A. Reiter, and M. Bell. "A deep learning based alternative to beamforming ultrasound images". In: *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2018, pp. 3359–3363.

[12]   A. Luchies and B. Byram. "Deep neural networks for ultrasound beamforming". In: *IEEE transactions on medical imaging* 37.9 (2018), pp. 2010–2021.

[13]   S. Vedula, O. Senouf, A. Bronstein, O. Michailovich, and M. Zibulevsky. "Towards CT-quality ultrasound imaging using deep learning". In: *arXiv preprint arXiv:1710.06304* (2017).

[14]   L. Papandreou G.and Chen, K. Murphy, and A. Yuille. "Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation". In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 1742–1750.

[15]   M. Kallenberg, K. Petersen, M. Nielsen, A. Ng, P. Diao, C. Igel, C. Vachon, K. Holland, R. Winkel, N. Karssemeijer, et al. "Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring". In: *IEEE transactions on medical imaging* 35.5 (2016), pp. 1322–1331.

[16]   X. Chen and E. Konukoglu. "Unsupervised detection of lesions in brain MRI using constrained adversarial auto-encoders". In: *arXiv preprint arXiv:1806.04972* (2018).

[17]   H. Shin, M. Orton, D. Collins, S. Doran, and M. Leach. "Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data". In: *IEEE transactions on pattern analysis and machine intelligence* (2012), p. 1.

[18]   D. Fraser, Y. Chen, E. Poon, F. Cury, T. Falco, and F. Verhaegen. "Dosimetric consequences of misalignment and realignment in prostate 3DCRT using intramodality ultrasound image guidance". In: *Medical physics* 37.6 (2010), pp. 2787–2795.

[19]   D. Dearnaley, I. Syndikus, H. Mossop, V. Khoo, A. Birtle, D. Bloomfield, J. Graham, P. Kirkbride, J. Logue, Z. Malik, et al. "Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHiP trial". In: *The Lancet Oncology* 17.8 (2016), pp. 1047–1060.

[20]   S. van der Meer, E. Bloemen-van Gurp, J. Hermans, R. Voncken, D. Heuvelmans, C. Gubbels, D. Fontanarosa, P. Visser, L. Lutgens, F. van Gils, and F. Verhaegen. "Critical assessment of intramodality 3D

ultrasound imaging for prostate IGRT compared to fiducial markers". In: *Medical physics* 40.7 (2013).

[21] M. Fargier-Voiron, B. Presles, P. Pommier, S. Rit, A. Munoz, H. Liebgott, D. Sarrut, and M. Biston. "Impact of probe pressure variability on prostate localization for ultrasound-based image-guided radiotherapy". In: *Radiotherapy and Oncology* 111.1 (2014), pp. 132–137.

[22] X. Yang, R. Kwitt, M. Styner, and M. Niethammer. "Quicksilver: Fast predictive image registration–a deep learning approach". In: *NeuroImage* 158 (2017), pp. 378–396.

[23] G. Wu, M. Kim, Q. Wang, Y. Gao, S. Liao, and D. Shen. "Unsupervised deep feature learning for deformable registration of MR brain images". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2013, pp. 649–656.

[24] G. Balakrishnan, A. Zhao, M. Sabuncu, J. Guttag, and A. Dalca. "An unsupervised learning model for deformable medical image registration". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 9252–9260.

[25] L. Gatys, A. Ecker, and M. Bethge. "A neural algorithm of artistic style". In: *arXiv preprint arXiv:1508.06576* (2015).

[26] L. Gatys, A. Ecker, and M. Bethge. "Image style transfer using convolutional neural networks". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, pp. 2414–2423.

[27] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. "Generative adversarial nets". In: *Advances in neural information processing systems*. 2014, pp. 2672–2680.

[28] S. Kazeminia, C. Baur, A. Kuijper, B. van Ginneken, N. Navab, S. Albarqouni, and A. Mukhopadhyay. "GANs for medical image analysis". In: *arXiv preprint arXiv:1809.06222* (2018).

[29] J. Krücker, C. Meyer, G. LeCarpentier, J. Fowlkes, and P. Carson. "3D spatial compounding of ultrasound images using image-based nonrigid registration". In: *Ultrasound in medicine & biology* 26.9 (2000), pp. 1475–1488.

[30] S. Mason. "Ultrasound-guided radiotherapy for cervical cancer". PhD thesis. The institute of Cancer Research and the Royal Marsden NHS Fonundation Trust, Oct. 2018.

[31] P. Kantoff, C. Higano, N. Shore, E. Berger, E. Small, D. Penson, C. Redfern, A. Ferrari, R. Dreicer, R. Sims, et al. "Sipuleucel-T immunotherapy for

castration-resistant prostate cancer". In: *New England Journal of Medicine* 363.5 (2010), pp. 411–422.

[32] F. Hodi, S. O'day, D. McDermott, R. Weber, J. Sosman, J. Haanen, R. Gonzalez, C. Robert, D. Schadendorf, J. Hassel, et al. "Improved survival with ipilimumab in patients with metastatic melanoma". In: *New England Journal of Medicine* 363.8 (2010), pp. 711–723.

[33] M. Reck, D. Rodríguez-Abreu, A. Robinson, R. Hui, T. Csőszi, A. Fülöp, M. Gottfried, N. Peled, A. Tafreshi, S. Cuffe, et al. "Pembrolizumab versus chemotherapy for PD-L1–positive non–small-cell lung cancer". In: *New England Journal of Medicine* 375.19 (2016), pp. 1823–1833.

[34] E. Van Limbergen, D. De Ruysscher, V. Olivo Pimentel, D. Marcus, M. Berbee, A. Hoeben, N. Rekers, J. Theys, A. Yaromina, L. Dubois, et al. "Combining radiotherapy with immunotherapy: the past, the present and the future". In: *The British journal of radiology* 90.1076 (2017), p. 20170157.

[35] K. Reynders, T. Illidge, S. Siva, J. Chang, and D. de Ruysscher. "The abscopal effect of local radiotherapy: using immunotherapy to make a rare event clinically relevant". In: *Cancer treatment reviews* 41.6 (2015), pp. 503–510.

[36] G. Herter-Sprie, S. Koyama, H. Korideck, J. Hai, J. Deng, Y. Li, K. Buczkowski, A. Grant, S. Ullas, K. Rhee, et al. "Synergy of radiotherapy and PD-1 blockade in Kras-mutant lung cancer". In: *JCI insight* 1.9 (2016).

[37] K. Koller, H. Mackley, J. Liu, H. Wagner, G. Talamo, T. Schell, C. Pameijer, R. Neves, B. Anderson, K. Kokolus, et al. "Improved survival and complete response rates in patients with advanced melanoma treated with concurrent ipilimumab and radiotherapy versus ipilimumab alone". In: *Cancer biology & therapy* 18.1 (2017), pp. 36–42.

[38] S. Mason, T. O'Shea, I. White, S. Lalondrelle, K. Downey, M. Baker, C. Behrens, J. Bamber, and E. Harris. "Towards ultrasound-guided adaptive radiotherapy for cervical cancer: Evaluation of Elekta's semiautomated uterine segmentation method on 3D ultrasound images". In: *Medical Physics* 44.7 (), pp. 3630–3638.

[39] A. Landry, T. Berrang, I. Gagne, C. Popescu, T. Mitchell, H. Vey, L. Sand, S. Soh, J. Wark, I. Olivotto, et al. "Investigation of variability in image acquisition and contouring during 3D ultrasound guidance for partial breast irradiation". In: *Radiation Oncology* 9.1 (2014), p. 35.

[40] E. Bloemen-van Gurp, S. van der Meer, J. Hendry, J. Buijsen, P. Visser, D. Fontanarosa, M. Lachaine, G. Lammering, and F. Verhaegen. "Active breathing control in combination with ultrasound imaging: a feasibility

study of image guidance in stereotactic body radiation therapy of liver lesions". In: *International Journal of Radiation Oncology* Biology* Physics* 85.4 (2013), pp. 1096–1102.

[41] J. Lagendijk, A. Raaymakers B.W.and Raaijmakers, J. Overweg, K. Brown, E. Kerkhof, R. van der Put, B. Hårdemark, M. van Vulpen, and U. van der Heide. "MRI/linac integration". In: *Radiotherapy and Oncology* 86.1 (2008), pp. 25–29.

[42] J. Lagendijk, B. Raaymakers, and M. Van Vulpen. "The magnetic resonance imaging–linac system". In: *Seminars in radiation oncology*. Vol. 24. 3. 2014, pp. 207–209.

[43] A. McPartlin, X. Li, L. Kershaw, U. Heide, L. Kerkmeijer, C. Lawton, U. Mahmood, F. Pos, N. van As, M. van Herk, D. Vesprini, J. van der Voort van Zyp, A. Tree, and A. Choudhury. "MRI-guided prostate adaptive radiotherapy–A systematic review". In: *Radiotherapy and Oncology* 119.3 (2016), pp. 371–380.

[44] T. Perik, J. Kaas, and F. Wittkämper. "The impact of a 1.5 T MRI linac fringe field on neighbouring linear accelerators". In: *Physics and Imaging in Radiation Oncology* 4 (2017), pp. 12–16.

[45] J. Slater, C. Rossi, L. Yonemoto, D. Bush, B. Jabola, R. Levy, R. Grove, W. Preston, and J. Slater. "Proton therapy for prostate cancer: the initial Loma Linda University experience". In: *International Journal of Radiation Oncology* Biology* Physics* 59.2 (2004), pp. 348–352.

[46] C. Fitzmaurice, C. Allen, R. Barber, L. Barregard, Z. Bhutta, H. Brenner, D. Dicker, O. Chimed-Orchir, R. Dandona, L. Dandona, et al. "Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study". In: *JAMA oncology* 3.4 (2017), pp. 524–548.

[47] B. Hoppe, J. Michalski, N. Mendenhall, C. Morris, R. Henderson, R. Nichols, W. Mendenhall, C. Williams, M. Regan, J. Chipman, et al. "Comparative effectiveness study of patient-reported outcomes after proton therapy or intensity-modulated radiotherapy for prostate cancer". In: *Cancer* 120.7 (2014), pp. 1076–1082.

[48] P. Beigi, P. Malenfant, A. Rasoulian, R. Rohling, A. Dube, and V. Gunka. "Three-dimensional ultrasound-guided real-time midline epidural needle placement with epiguide: a prospective feasibility study". In: *Ultrasound in medicine & biology* 43.1 (2017), pp. 375–379.

[49] J. Volders, M. Haloua, N. Krekel, V. Negenborn, R. Kolk, A. Lopes Cardozo, A. Bosch, L. de Widt-Levert, H. van der Veen,

H. Rijna, et al. "Intraoperative ultrasound guidance in breast-conserving surgery shows superiority in oncological outcome, long-term cosmetic and patient-reported outcomes: final outcomes of a randomized controlled trial (COBALT)". In: *European Journal of Surgical Oncology (EJSO)* 43.4 (2017), pp. 649–657.

[50] M. Lee, M. Lubner, J. Hinshaw, and P. Pickhardt. "Ultrasound guidance versus CT guidance for peripheral lung biopsy: performance according to lesion size and pleural contact". In: *American Journal of Roentgenology* 210.3 (2018), W110–W117.

Appendices

A Patient-specific transperineal ultrasound probe setup

A.1 Bone mask extraction

Voxels in the CT volume that correspond to bone tissue, tend to be brighter and so tend to have a higher Hounsfield Unit (HU) than the surrounding soft tissue. This principle was used to perform initial thresholding on the whole CT volume using a threshold value of 200 HU (empirically chosen).

In Fig. A.1A an axial projection of the simulation CT scan is shown while the corresponding coarse outline of the bones resulting from the thresholding procedure is displayed in Fig. A.1B. Since it was necessary to obtain a bone mask with a connected boundary, post-processing steps were required, which were performed per axial slice.

(A)   (B)   (C)   (D)   (E)
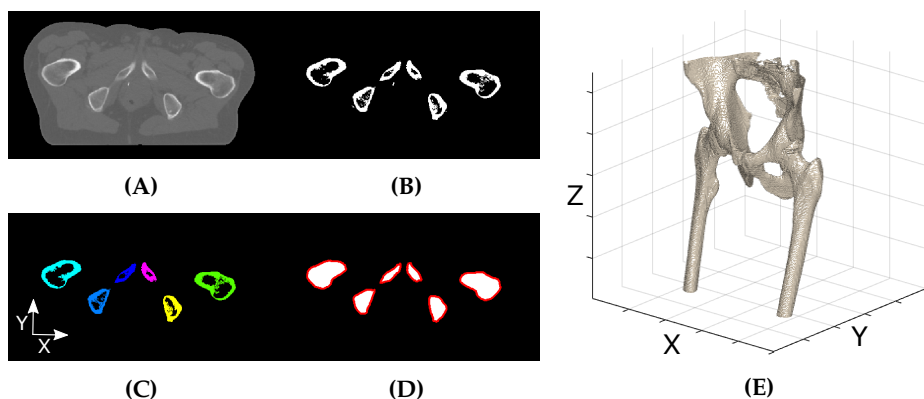
Figure A.1: Workflow to extract a 3D binary bone mask from a CT volume. (A) One axial slice of an acquired CT volume. (B) Bone mask after initial thresholding. (C) Each color corresponds to one connected component identified in the 2D slice. (D) The red line represents the boundary that envelopes the pixels of the different connected components. This boundary was then used to create a solid 2D binary bone mask. (E) Bone mask resulting from repeating the procedure on all axial slices displayed using a 3D interpolation process.

After bridging single, unconnected pixels and removing connected components that were smaller than a specific threshold of 25 pixels (empirically chosen), each remaining connected component (Fig. A.1C) was assessed separately.

Subsequently, a compact boundary per component was retrieved that enveloped the pixels of this component (red boundaries in Fig. A.1D). This was done to ensure that a mask with a connected boundary could also be retrieved, even when the original boundary of the bones was not closed.

Finally, the new boundary was filled with white pixels, resulting in a 2D binary mask of the bones of the corresponding axial slice (Fig. A.1D). The above procedure was repeated for all axial slices of the volume, resulting in a 3D bone mask, as displayed in Fig. A.1E.

A.2 Leg mask extraction

In order to identify the part of the legs that were covering the perineal skin area, each 2D axial slice of the CT volume of the patient was investigated separately. In this study, five different 2D slice categories were identified, of which visual examples are displayed in Fig. A.2. Category A included all 2D slices on which two unconnected legs were present (Fig. A.2A). In Category B these two legs were also present, however, in addition, one or multiple parts of the scrotum were visualized as well (Fig. A.2B). All these body parts were still unconnected. Category C included all slices on which two legs were present with one or multiple parts of the scrotum connected to one of the legs (Fig. A.2C). In Category D the scrotum was connected to both legs (Fig. A.2D), while in Category E neither legs nor scrotum were visualized (Fig. A.2E). Of course, also combinations of the categories existed, such as two legs with part of the scrotum connected to one of the legs and part of the scrotum unconnected. However, these could be handled by combining the different steps described below.

(A) **(B)** **(C)**

(D) **(E)**
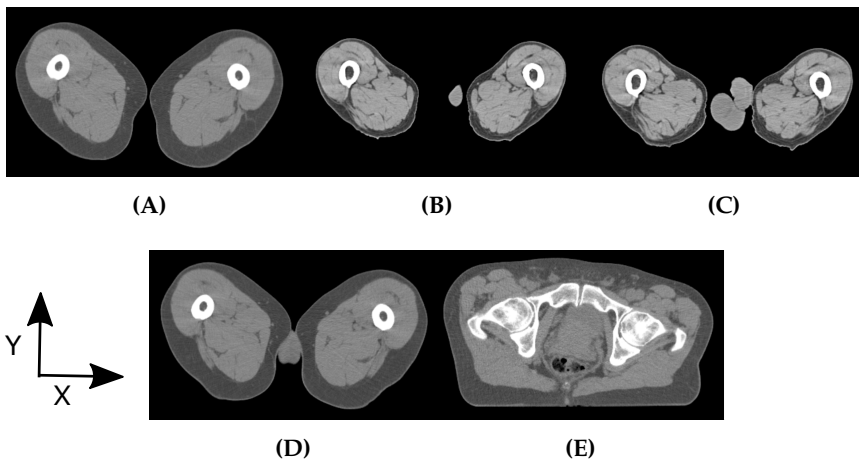
Figure A.2: Five defined 2D slice categories. (A) Two unconnected legs. (B) Two unconnected legs with one or multiple unconnected parts of the scrotum. (C) Two legs with one or multiple parts of the scrotum connected to one of the legs. (D) The scrotum connected to both of the legs. (E) Legs and scrotum are absent in the slice.

In order to distinguish Category A from C, the perimeters of the two connected components on each slice were determined and plotted starting from the most caudal slice. A clear increase in the length of the perimeter could be recognized when the slices transitioned from Category A to Category C. Before this transition point, all the slices with two components were classified as Category A, while after this point the slices were classified as Category C.

The distinction between the Categories D and E was performed based on the shape of the bone present in the slice. If the femur bone significantly started to change shape, the slice was classified as not containing legs or scrotum anymore. Hence also in this case, a transition point could be recognized, marking the transition from Category D to E.

After the category division, a binary leg mask of the patient was created. As Category A only shows the legs of the patient, these slices were directly added to the mask. The slices of Category B were converted into Category A by removing the smallest components (based on their area) from the slices, until only two components were left. These were then also added to the mask.

Prior to adding the legs of Category C to the legs mask the scrotum needed to be removed. To this end, the leg to which the scrotum was connected was identified, as well as the maximum voxel in Y-direction and minimum voxel in X-direction (see magenta diamonds in Fig. A.3A).
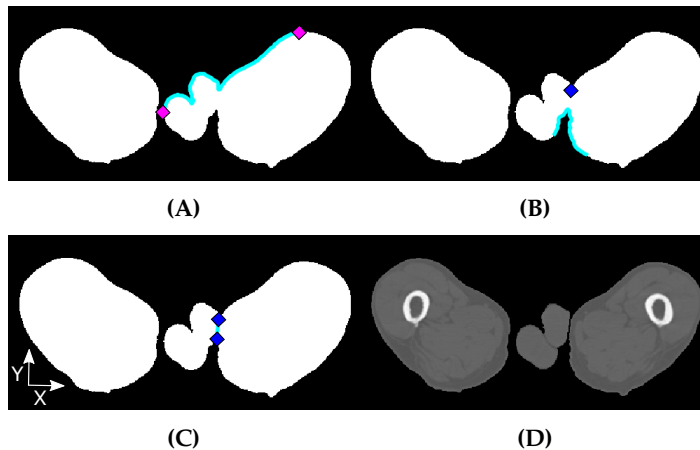
Figure A.3: Workflow to remove the scrotum from 2D slices of Category C. (A) Binary mask of a Category C slice with the magenta diamonds representing maximum voxel in Y-direction and minimum voxel in X-direction. The cyan line shows the perimeter of the mask between these points. (B) Binary mask with the blue diamond representing the last local minimum of the anterior perimeter and the cyan line representing the posterior perimeter. (C) Binary bone mask with the blue diamonds being the local minimum of the anterior perimeter and the global maximum of the posterior perimeter connected by a cyan straight line. (D) Corresponding CT slice after the straight line of (C) was used as cutting line.

Subsequently, the last local minimum of the perimeter between these points (cyan line in Fig. A.3A) was found and represented by the blue diamond in Fig. A.3B. The X-coordinate of this local minimum was then used to select the perimeter on the posterior side using a set margin. In this case, the margin was 20 voxels and the resulting perimeter selection is displayed with the cyan line in Fig. A.3B. Next, the global maximum of this selected perimeter was connected by a straight line to the earlier found local minimum of the anterior side, see Fig. A.3C. Finally, this straight line was used as a cutting line to disconnect the scrotum from the leg (Fig. A.3D), allowing the slice to be processed the same way as slices of Category B.

To remove the scrotum from the Category D slices, a similar procedure was followed. After identification of the left and right legs, the voxels with the maximum and minimum Y-coordinates for both legs were determined (magenta diamonds in Fig. A.4A). Subsequently, the local minima and maxima on the perimeters between these points (cyan lines in Fig. A.4A and Fig. A.4B) were identified (blue diamonds in Fig. A.4B). Based on some set constraints and the neighboring slices, the correct local minima could then be selected and connected by straight lines (cyan lines in Fig. A.4C). Finally, the original CT slice was then cut along those lines, which allowed processing of the slice the same way as a Category B slice (Fig. A.4D).
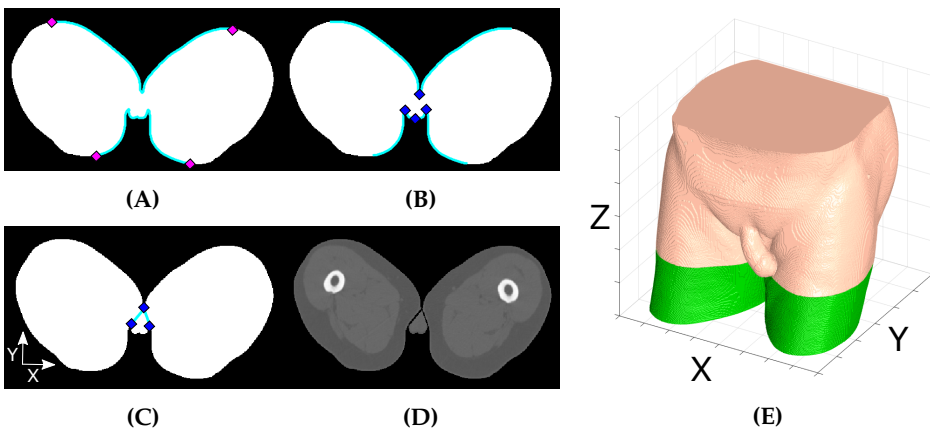
Figure A.4: Workflow to remove the scrotum from 2D slices of Category D. (A) Binary mask of a Category D slice with the magenta diamonds representing the voxels with maximum and minimum Y-coordinates and the cyan lines indicating the perimeters between these points. (B) Binary mask with the same perimeters and the local minima and maxima of these perimeters represented by the blue diamonds. (C) Selected minima and maxima being connected by straight cyan lines. (D) Corresponding CT slice after the straight lines of (C) were used as cutting line. (E) The body outline of the patient with the identified bone mask displayed in green.

After processing all the slices which contained legs (Category A-D), a binary leg mask was constructed which could be used later on to restrict the accessible perineum area (Fig. A.4E).

A.3 Bone mask underestimation
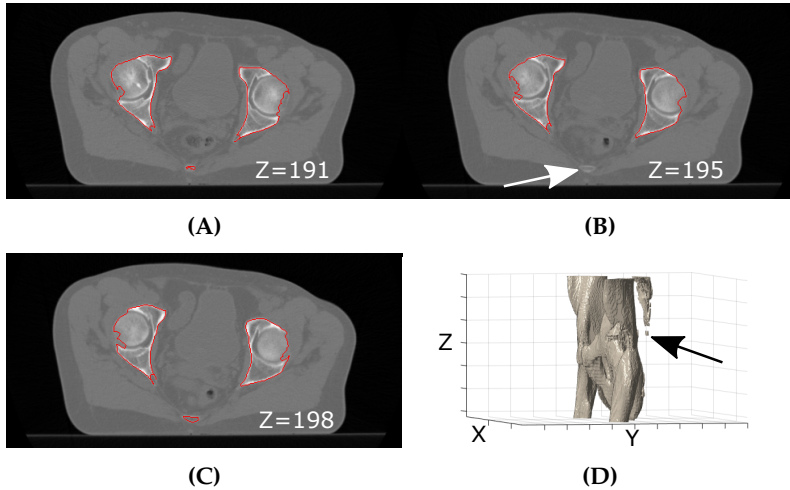
Figure A.5: Example of underestimation during bone segmentation based on image data of Patient 1. (A) CT slice at Z=191 with superimposed in red the detected bone edges. (B) CT slice at Z=195 with a missing spine segmentation at the arrow. (C) CT slice at Z=198 with the spine found again. (D) 3D representation of the bone mask. The arrow indicates a piece of spine disconnected from the mask due to underestimation in slices Z=195-197. If this piece was smaller than 50 voxels, it was removed before coccyx identification.

B US image quality assessment using deep learning

B.1 Parameters of additional deep learning approaches

Table B.1: Parameters of the additional deep learning algorithms per implementation step with the asterisks indicating the optimized hyper-parameters.

		CNN + GP (two class)	CNN + Softmax
	Parameter	**Value**	**Value**
DenseNet	Number of blocks*	2	3
	Number of layers*	18	18
	Growth rate k*	12	12
	Outputs*	200	-
GPflow	Model	SGPR	-
	Kernel	RBF	-
	Initial kernel variance*	0.2	-
	Inducing points*	150	-
Training	Batch size*	200	50
	Epochs*	75	75
	Optimizer	Adam	Adam
	Learning rate*	1e-8	1e-6
	Drop-out rate*	0.05	0.05

C Proof-of-concept study for pseudo-CT scan creation

C.1 Parameter file

Intensity based deformation	Changes for the contour based deformation
(FixedInternalImagePixelType "float")	
(MovingInternalImagePixelType "float")	
(UseDirectionCosines "false")	
(Registration "MultiMetricMultiResolutionRegistration")	
(FixedImagePyramid "FixedRecursiveImagePyramid")	
(MovingImagePyramid "MovingRecursiveImagePyramid")	
(Interpolator "BSplineInterpolator")	
(Optimizer "AdaptiveStochasticGradientDescent")	
(ResampleInterpolator "FinalBSplineInterpolator")	
(Resampler "DefaultResampler")	
(Transform "BSplineTransform")	
(Metric "AdvancedNormalizedCorrelation" "TransformRigidityPenalty")	(Metric "AdvancedMeanSquares" "TransformRigidityPenalty")
(NumberOfResolutions 3)	
(HowToCombineTransforms "Compose")	
(MaximumNumberOfIterations 100)	(MaximumNumberOfIterations 300)
(10, 50 or 100)	*(100 or 300)*
(Metric0Weight 1.0) (Metric1Weight 10.0)	
(LinearityConditionWeight 100.0)	
(OrthonormalityConditionWeight 1.0)	
(PropernessConditionWeight 2.0)	
(UseLinearityCondition "true")	
(UseOrthonormalityCondition "true")	
(UsePropernessCondition "true")	
(CalculateLinearityCondition "true")	
(CalculateOrthonormalityCondition "true")	
(CalculatePropernessCondition "true")	
(DilateRigidityImages "false" "false" "false")	
(DilationRadiusMultiplier 1.0 1.0 1.0)	
(UseFixedRigidityImage "false")	
(UseMovingRigidityImage "true")	
(MovingRigidityImageName "filepath and -name of the mask")	
(FinalGridSpacingInPhysicalUnits 20)	
(NumberOfSpatialSamples 2500)	
(ImageSampler "Random")	
(BSplineInterpolationOrder 3)	
(FinalBSplineInterpolationOrder 1)	(FinalBSplineInterpolationOrder 0)
(DefaultPixelValue 0)	
(ResultImagePixelType "short")	
(ResultImageFormat "mhd")	

C.2   Results for bladder, anus & rectum and body contours

Table C.1: Gamma index evaluations based on HU and dose as well as DSC calculations were performed for body, anus & rectum and bladder. The first three columns detail the patient, the evaluated contour and the used evaluation metric (gamma failure based on HU or dose expressed in volume % and DSC). The values in the Ref (reference) column were based on evaluation of CT_{sim} and CT_{tx}. The final columns detail the evaluation of the created CT_{ps} using the specific parameter sets A-E and CT_{tx}. The bold numbers indicate which parameter sets resulted in the same result or in an improvement with respect to the reference.

Patient	Contour	Metric		Ref	A	B	C	D	E
1	Body	γ_{CT}	[%]	14.3	**9.8**	**11.5**	**12.7**	**9.7**	**11.2**
		γ_{Dose}	[%]	4.9	**4.6**	**4.4**	**4.3**	**4.6**	**4.7**
	Anus & Rectum	DSC		0.4	**0.6**	**0.7**	**0.7**	**0.6**	**0.7**
		γ_{CT}	[%]	22.0	**14.5**	**13.4**	**13.6**	**15.2**	**16.1**
		γ_{Dose}	[%]	0.2	**0.2**	**0.2**	**0.2**	**0.2**	0.4
2	Body	γ_{CT}	[%]	23.1	**21.7**	**19.8**	**19.9**	**19.9**	-
		γ_{Dose}	[%]	2.8	**2.8**	**2.8**	**2.8**	**2.6**	-
	Anus & Rectum	DSC		0.7	0.6	**0.7**	**0.7**	**0.7**	-
		γ_{CT}	[%]	28.9	**25.4**	**27.3**	**28.8**	**28.1**	-
		γ_{Dose}	[%]	0.8	**0.2**	**0.5**	**0.7**	0.9	-
	Bladder	DSC		0.6	**0.6**	**0.6**	**0.6**	0.7	-
		γ_{CT}	[%]	10.4	**7.7**	**4.7**	**3.7**	**5.8**	-
		γ_{Dose}	[%]	0.0	0.3	0.7	1.0	**0.0**	-
3a	Body	γ_{CT}	[%]	8.5	**7.4**	**6.8**	**7.2**	**8.0**	**8.0**
		γ_{Dose}	[%]	0.6	**0.6**	**0.6**	**0.6**	**0.6**	**0.6**
	Anus & Rectum	DSC		0.7	0.4	0.2	0.2	0.4	0.4
		γ_{CT}	[%]	22.7	**20.3**	**22.5**	23.9	**21.1**	**21.5**
		γ_{Dose}	[%]	4.2	4.6	4.3	4.4	4.4	4.7
	Bladder	DSC		0.8	**0.8**	**0.8**	**0.8**	**0.8**	**0.8**
		γ_{CT}	[%]	4.2	**3.0**	**2.7**	**3.2**	6.5	6.4
		γ_{Dose}	[%]	0.0	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
3b	Body	γ_{CT}	[%]	10.6	**9.8**	**8.2**	**7.9**	**5.7**	**6.0**
		γ_{Dose}	[%]	0.9	**0.8**	**0.8**	**0.8**	**0.5**	**0.5**
	Anus & Rectum	DSC		0.6	**0.6**	**0.6**	**0.6**	**0.7**	**0.7**
		γ_{CT}	[%]	11.7	**7.0**	**6.7**	**7.2**	**11.2**	12.9
		γ_{Dose}	[%]	0.0	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
	Bladder	DSC		0.8	**0.8**	**0.8**	**0.8**	**0.8**	**0.8**
		γ_{CT}	[%]	11.1	**8.0**	**7.0**	**6.6**	**7.4**	**7.7**
		γ_{Dose}	[%]	0.0	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**

Patient	Contour	Metric		Ref	A	B	C	D	E
3c	Body	γ_{CT}	[%]	18.1	**16.5**	**15.0**	**13.5**	**12.8**	**12.7**
		γ_{Dose}	[%]	1.5	1.6	1.8	1.7	**1.1**	**1.1**
	Anus & Rectum	DSC		0.5	**0.5**	**0.6**	**0.6**	**0.6**	**0.6**
		γ_{CT}	[%]	19.0	**14.2**	**11.5**	**9.3**	**12.9**	**12.4**
		γ_{Dose}	[%]	0.0	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
	Bladder	DSC		0.7	**0.7**	0.6	0.6	**0.8**	**0.8**
		γ_{CT}	[%]	7.6	**5.9**	**6.6**	**6.5**	**1.2**	**1.5**
		γ_{Dose}	[%]	0.0	**0.0**	1.3	1.0	**0.0**	**0.0**
3d	Body	γ_{CT}	[%]	23.8	**21.9**	**19.7**	**18.5**	**18.3**	**17.8**
		γ_{Dose}	[%]	2.0	**2.0**	**1.9**	**1.8**	**1.5**	**1.5**
	Anus & Rectum	DSC		0.3	**0.3**	**0.4**	**0.4**	**0.5**	**0.5**
		γ_{CT}	[%]	32.8	**24.2**	**14.9**	**12.3**	**18.3**	**18.2**
		γ_{Dose}	[%]	0.5	**0.2**	0.6	**0.4**	**0.1**	**0.2**
	Bladder	DSC		0.7	**0.7**	**0.8**	**0.8**	**0.8**	**0.8**
		γ_{CT}	[%]	20.8	**17.1**	**12.2**	**9.6**	**11.5**	**11.6**
		γ_{Dose}	[%]	0.3	**0.2**	**0.1**	**0.0**	**0.0**	**0.0**

Figure C.1: Results for all five different parameter sets for all patient data. In the first column, the results of the Dice similarity coefficient are plotted. The second column details both MAD and MAX of the prostate contours. In the third column, the gamma evaluation on the CT values is shown. The percentage of a gamma failure ($|\gamma_{CT}(50HU, 3\,mm)| > 1$) is plotted. In the fourth column, the gamma evaluation on the dose. Again the percentage of a gamma failure ($|\gamma_{Dose}(3\%, 3\,mm)| > 1$) is plotted.

Figure C.1: Results for all five different parameter sets for all patient data. In the first column, the results of the Dice similarity coefficient are plotted. The second column details both MAD and MAX of the prostate contours. In the third column, the gamma evaluation on the CT values is shown. The percentage of a gamma failure ($|\gamma_{CT}(50HU, 3\,mm)| > 1$) is plotted. The last column shows the gamma evaluation on the dose. Again the percentage of a gamma failure ($|\gamma_{Dose}(3\%, 3\,mm)| > 1$) is plotted.
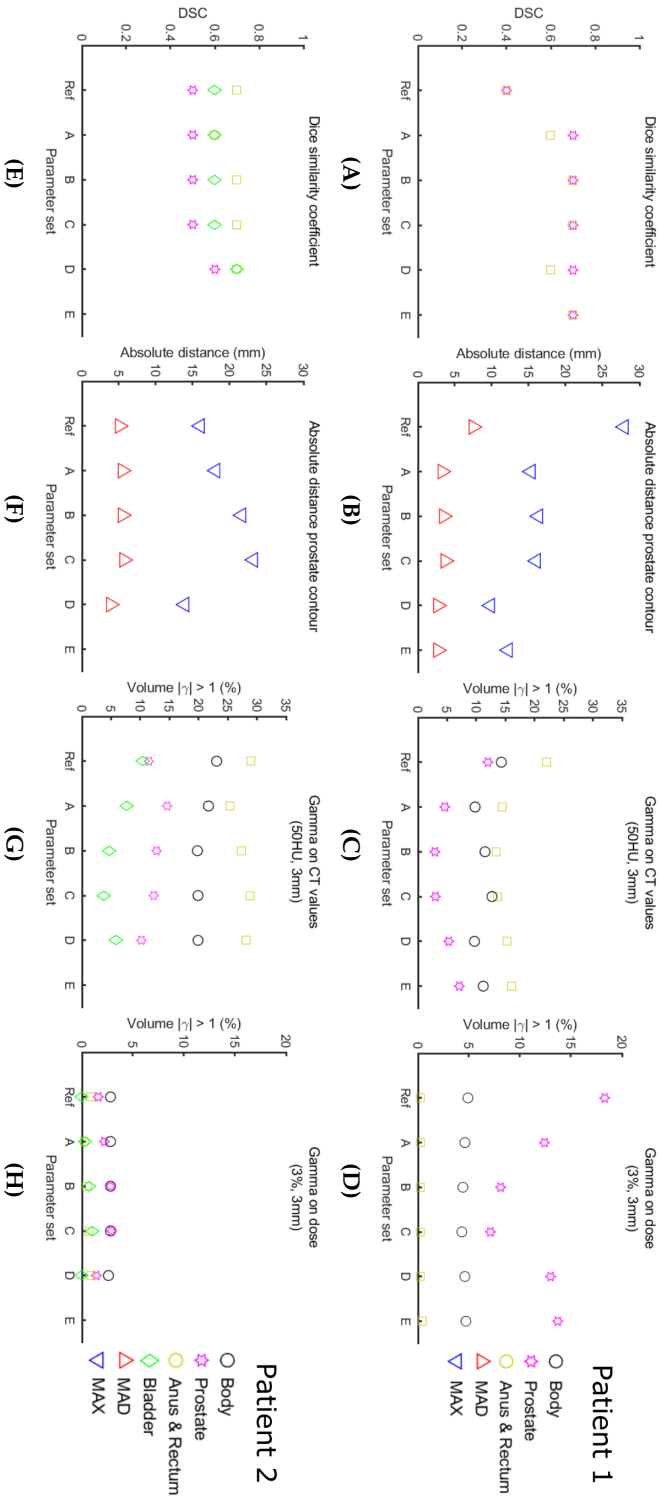
Figure C.1: Results for all five different parameter sets for all patient data. In the first column, the results of the Dice similarity coefficient are plotted. The second column details both MAD and MAX of the prostate contours. In the third column, the gamma evaluation on the CT values is shown. The percentage of a gamma failure ($|\gamma_{CT}(50HU, 3\,mm)| > 1$) is plotted. The last column shows the gamma evaluation on the dose. Again the percentage of a gamma failure ($|\gamma_{Dose}(3\%, 3\,mm)| > 1$) is plotted.
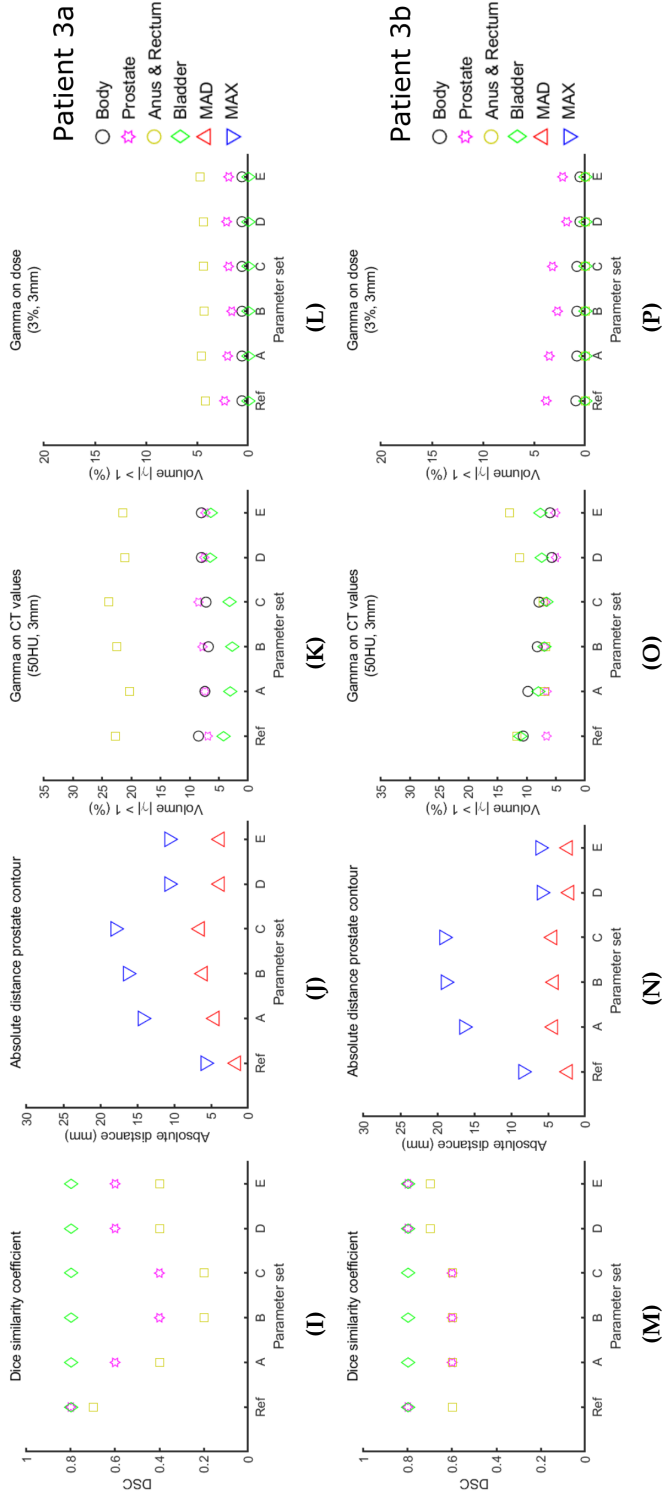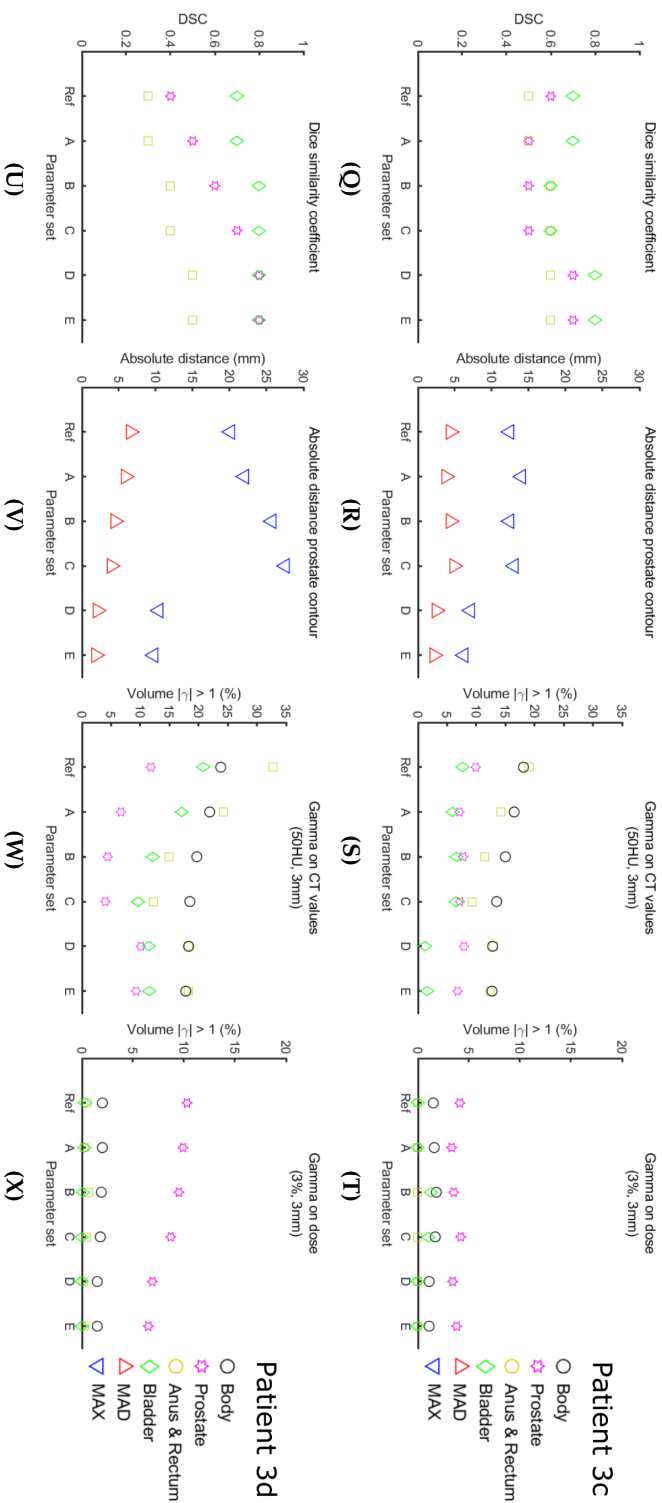
D Various approaches for pseudo-CT scan creation

D.1 Parameter set for initial alignment

(FixedInternalImagePixelType "float")
(MovingInternalImagePixelType "float")
(UseDirectionCosines "true")

(Registration "MultiResolutionRegistration")
(Interpolator "BSplineInterpolator")
(ResampleInterpolator "FinalBSplineInterpolator")
(Resampler "DefaultResampler")
(FixedImagePyramid "FixedRecursiveImagePyramid")
(MovingImagePyramid "MovingRecursiveImagePyramid")

(Optimizer "AdaptiveStochasticGradientDescent")
(Transform "TranslationTransform")
(Metric "AdvancedMeanSquares")

(AutomaticScalesEstimation "true")
(AutomaticTransformInitialization "false")
(ErodeMask "false")
(HowToCombineTransforms "Compose")

% NumberOfResolutions was set to 2 or 1 when excessive translation occurred
(NumberOfResolutions 3)

(MaximumNumberOfIterations 750)
(NumberOfSpatialSamples 2500)
(NewSamplesEveryIteration "true")
(ImageSampler "Random")

(FinalGridSpacingInPhysicalUnits 30)
(BSplineInterpolationOrder 1)
(FinalBSplineInterpolationOrder 0)
(MovingImageDerivativeScales 1 0 0)
(DefaultPixelValue 0)

(WriteResultImage "true")
(ResultImagePixelType "short")
(ResultImageFormat "mhd")

D.2 Parameter set for bone rigidity evaluation

(FixedInternalImagePixelType "float")
(MovingInternalImagePixelType "float")
(UseDirectionCosines "true")

(Registration "MultiResolutionRegistration")
(Interpolator "BSplineInterpolator")
(ResampleInterpolator "FinalBSplineInterpolator")
(Resampler "DefaultResampler")

(FixedImagePyramid "FixedRecursiveImagePyramid")
(MovingImagePyramid "MovingRecursiveImagePyramid")

(Optimizer "AdaptiveStochasticGradientDescent")
(Transform "EulerTransform")
(Metric "AdvancedMattesMutualInformation")

(AutomaticScalesEstimation "true")
(AutomaticTransformInitialization "true")
(HowToCombineTransforms "Compose")
(NumberOfHistogramBins 32)

(NumberOfResolutions 3)
(MaximumNumberOfIterations 250)
(NumberOfSpatialSamples 2500)

(NewSamplesEveryIteration "true")
(ImageSampler "Random")

(BSplineInterpolationOrder 3)
(FinalBSplineInterpolationOrder 0)
(DefaultPixelValue 0)

(WriteResultImage "true")
(ResultImagePixelType "short")
(ResultImageFormat "mhd")

D.3   Parameter sets for deformable registration

(FixedInternalImagePixelType "float")
(MovingInternalImagePixelType "float")
(UseDirectionCosines "false")

(Interpolator "BSplineInterpolator")
(ResampleInterpolator "FinalBSplineInterpolator")
(Resampler "DefaultResampler")

(FixedImagePyramid "FixedRecursiveImagePyramid")
(MovingImagePyramid "MovingRecursiveImagePyramid")

(Optimizer "AdaptiveStochasticGradientDescent")
(Transform "BSplineTransform")
(HowToCombineTransforms "Compose")

(Registration "MultiResolutionRegistration")
(Metric **METRIC**)
(NumberOfResolutions **RESOLUTIONS**)

(FinalGridSpacingInPhysicalUnits **SPACING**)
(MaximumNumberOfIterations **ITERATIONS**)
(NumberOfHistogramBins 32)

(NumberOfSpatialSamples 2500)
(NewSamplesEveryIteration "true")
(ImageSampler "Random")
(BSplineInterpolationOrder 3)

% The FinalBSplineInterpolationOrder is set to 0 when the deformable registration is based on contours only, otherwise it is set to 1.
(FinalBSplineInterpolationOrder 0)

(DefaultPixelValue 0)
(WriteResultImage "true")
(ResultImagePixelType "short")
(ResultImageFormat "mhd")

METRIC
SSD = (Metric "AdvancedMeanSquares")
NCC = (Metric "AdvancedNormalizedCorrelation")
MI = (Metric "AdvancedMattesMutualInformation")

Approach	Pair	Metric	Resolutions	Spacing	Iterations
Deformable registration	T1 - T4	SSD	3	15	295
Prostate contour	T1 - T5	SSD	4	10	265
	T2 - T4	SSD	3	35	275
	T2 - T5	NCC	2	10	205
	T3 - T5	NCC	4	5	175
	T4 - T1	NCC	5	10	285
	T4 - T2	NCC	3	10	195
	T5 - T1	SSD	4	5	185
	T5 - T2	NCC	4	15	285
	T5 - T3	NCC	1	10	165
Deformable registration	T1 - T4	MI	3	20	75
US volume	T1 - T5	NCC	3	10	125
	T2 - T4	NCC	1	25	105
	T2 - T5	NCC	4	25	255
	T3 - T5	SSD	2	30	105
	T4 - T1	SSD	1	20	75
	T4 - T2	SSD	4	15	5
	T5 - T1	SSD	3	30	45
	T5 - T2	NCC	5	20	265
	T5 - T3	NCC	4	35	65
Translation	T1 - T4	NCC	4	10	295
Prostate contour	T1 - T5	SSD	3	10	205
	T2 - T4	SSD	2	5	285
Deformable registration	T2 - T5	NCC	4	15	225
Prostate contour	T3 - T5	SSD	3	5	195
	T4 - T1	NCC	5	10	275
	T4 - T2	SSD	4	15	225
	T5 - T1	SSD	3	15	285
	T5 - T2	NCC	2	10	235
	T5 - T3	SSD	3	10	275
Translation	T1 - T4	MI	2	30	45
Prostate contour	T1 - T5	NCC	5	15	235
	T2 - T4	NCC	1	15	35
Deformable registration	T2 - T5	MI	1	20	75
US volume	T3 - T5	SSD	1	30	115
	T4 - T1	NCC	1	35	45
	T4 - T2	MI	1	25	15
	T5 - T1	NCC	2	20	195
	T5 - T2	NCC	2	20	225
	T5 - T3	MI	2	20	25

Approach	Pair	Metric	Resolutions	Spacing	Iterations
Translation	T1 - T4	SSD	4	10	265
US volume	T1 - T5	NCC	3	10	255
	T2 - T4	NCC	4	15	225
Deformable registration	T2 - T5	NCC	5	10	245
Prostate contour	T3 - T5	NCC	1	5	285
	T4 - T1	NCC	3	10	175
	T4 - T2	NCC	1	20	285
	T5 - T1	NCC	3	5	275
	T5 - T2	NCC	3	10	255
	T5 - T3	SSD	4	10	195
Translation	T1 - T4	MI	2	25	65
US volume	T1 - T5	NCC	3	10	255
	T2 - T4	MI	3	30	15
Deformable registration	T2 - T5	NCC	2	15	275
US volume	T3 - T5	SSD	1	20	45
	T4 - T1	SSD	4	15	5
	T4 - T2	MI	2	35	15
	T5 - T1	SSD	2	15	105
	T5 - T2	NCC	4	20	245
	T5 - T3	MI	3	15	115

D.4 Calculation times

Data pairs	Translation [US]	Translation [C]	Deform [US]	Deform [C]	Translation [US] Deform [US]	Translation [US] Deform [C]	Translation [C] Deform [US]	Translation [C] Deform [C]
	Calculation Time [s]	Calculation Time [s]	Calculation Time [s]	Calculation Time [s]	Calculation Time [s]	Calculation Time [s]	Calculation Time [s]	Calculation Time [s]
T1-T4	125	149	25	30	150	195	164	189
T1-T5	81	155	26	41	121	123	172	178
T2-T4	79	67	15	39	98	118	83	99
T2-T5	141	57	28	32	164	186	73	106
T3-T5	146	110	23	62	160	196	126	176
T4-T1	72	133	16	61	89	109	148	183
T4-T2	75	51	14	38	91	96	63	81
T5-T1	71	44	14	45	93	121	65	70
T5-T2	90	119	53	50	113	117	141	143
T5-T3	124	118	19	18	144	152	133	134

Publications and patents

Journal papers

[J1] S. Camps, T. Houben, C. Edwards, M. Antico, M. Dunnhofer, E. Martens, J. Baeza, B. Vanneste, E. van Limbergen, P. de With, F. Verhaegen, G. Carneiro, and D. Fontanarosa. "Automatic quality assessment of transperineal ultrasound images of the male pelvic region using deep learning". In: *Ultrasound in Medicine and Biology (submitted)* (2018).

[J2] S. Camps, F. Verhaegen, B. Vanneste, P. de With, and D. Fontanarosa. "Automated patient-specific transperineal ultrasound probe setups for prostate cancer patients undergoing radiotherapy". In: *Medical Physics* 45.7 (2018), pp. 3185–3195.

[J3] S. Camps, D. Fontanarosa, P. de With, F. Verhaegen, and B. Vanneste. "The use of ultrasound imaging in the external beam radiotherapy workflow of prostate cancer patients". In: *BioMed Research International* 2018 (2018).

[J4] S. Camps, S. van der Meer, F. Verhaegen, and D. Fontanarosa. "Various approaches for pseudo-CT scan creation based on ultrasound to ultrasound deformable image registration between different treatment time points for radiotherapy treatment plan adaptation in prostate cancer patients". In: *Biomedical Physics & Engineering Express* 2.3 (2016).

[J5] S. van der Meer*, S. Camps*, W. van Elmpt, M. Podesta, P. Sanches, B. Vanneste, D. Fontanarosa, and F. Verhaegen. "Simulation of pseudo-CT images based on deformable image registration of ultrasound images: A proof of concept for transabdominal ultrasound imaging of the prostate during radiotherapy". In: *Medical Physics* 43.4 (2016), pp. 1913–1920.

International conference contributions

[IC1] S. Camps, T. Houben, C. Edwards, M. Antico, M. Dunnhofer, E. Martens, J. Baeza, B. Vanneste, E. van Limbergen, P. de With, F. Verhaegen, G. Carneiro, and D. Fontanarosa. "Automatic deep learning based quality assessment of transperineal ultrasound guided prostate radiotherapy". In: *Australian Society of Medical Imaging and Radiation Therapy's (ASMIRT) conference*. Adelaide, Australia, 2019.

[IC2] S. Camps, T. Houben, C. Edwards, M. Antico, M. Dunnhofer, E. Martens, J. Baeza, B. Vanneste, E. van Limbergen, P. de With, F. Verhaegen, G. Carneiro, and D. Fontanarosa. "Quality assessment of transperineal ultrasound images of the male pelvic region using deep learning". In: *Proc. IEEE International Ultrasonics Symposium*. Kobe, Japan, 2018.

[IC3] S. Camps*, T. Houben*, C. Edwards, M. Antico, M. Dunnhofer, E. Martens, J. Baeza, B. Vanneste, E. van Limbergen, P. de With, F. Verhaegen, G. Carneiro, and D. Fontanarosa. "One-class Gaussian process regressor for quality assessment of transperineal ultrasound images". In: *Medical imaging with deep learning conference*. Amsterdam, the Netherlands, 2019.

[IC4] S. Camps, F. Verhaegen, P. de With, and D. Fontanarosa. "CT based prostate cancer patient-specific transperineal ultrasound probe setups for image guided radiotherapy". In: *Proc. IEEE International Ultrasonics Symposium*. Washington DC, USA, 2017.

[IC5] S. Camps, F. Verhaegen, P. de With, and D. Fontanarosa. "CT scan based patient-specific transperineal ultrasound probe setups for image guided radiotherapy". In: *American Association of Physicists in Medicine (AAPM) conference*. Denver, USA, 2017.

[IC6] S. Camps, S. van der Meer, F. Verhaegen, and D. Fontanarosa. "Pseudo-CT scan creation using registration of transabdominal ultrasound volumes of a prostate cancer patient". In: *American Association of Physicists in Medicine (AAPM) conference*. Denver, USA, 2017.

[IC7] S. Camps, F. Verhaegen, P. de With, and D. Fontanarosa. "Patient-specific transperineal ultrasound probe setups for image guided radiotherapy". In: *ESTRO conference*. Vienne, Austria, 2017.

[IC8] S. Camps, F. Verhaegen, G. P. Fonseca, P. de With, and D. Fontanarosa. "Automatic transperineal ultrasound probe positioning based on CT scan for image guided radiotherapy". In: *Proc. SPIE medical imaging*. Orlando, USA, 2017.

[IC9] S. Camps, F. Verhaegen, P. de With, and D. Fontanarosa. "Geometry based transperineal ultrasound probe positioning for image guided radiotherapy". In: *American Association of Physicists in Medicine (AAPM) conference*. Washington DC, USA, 2016.

[IC10] S. Camps, S. van der Meer, F. Verhaegen, and D. Fontanarosa. "Pseudo-CT scan creation based on ultrasound to ultrasound deformable image registration in prostate cancer patients for radiotherapy treatment adaptation". In: *Ultrasound guidance in radiotherapy conference*. London, UK, 2015.

Regional conference contributions

[RC1] S. Camps, F. Verhaegen, P. de With, and D. Fontanarosa. "Geometry based transperineal ultrasound probe positioning for image guided radiotherapy". In: *IEEE Student Branch Eindhoven*. Eindhoven, the Netherlands, 2016.

Published and/or filed patents

[P1] A. Elevelt, S. Box, S. Camps, J. Geraats, A. van Noort, T. Rijken, and J. Kustra. "A system to maintain ultrasound probe setup and ensure constant pressure (force) during hands free procedures". Patent. Filed in 2018.

[P2] D. Fontanarosa and S. Camps. "Image guided treatment delivery". Patent. Published in 2017.

Acronyms

3D-CRT   3D conformal radiotherapy

AP   Anterior-Posterior

ART   Adaptive radiotherapy

BAT   B-mode acquisition and targeting

CBCT   Cone beam computed tomography

CNN   Convolutional neural network

CT   Computed tomography

DIR   Deformable image registration

DSC   Dice similarity coefficient

DTA   Distance-to-agreement

EPI   Electronic portal imaging

EBRT   External beam radiotherapy

FOV   Field of view

FM   Fiducial marker

GP   Gaussian process

HDR   High-dose rate

HU   Hounsfield units

IGRT   Image guided radiotherapy

IMRT   Intensity-modulated radiotherapy

LoA   Limits of agreement

LDR   Low-dose rate

Linac   Linear accelerator

LR   Left-Right

MAD   Mean absolute distance

MAX   Maximum absolute distance

MR   Magnetic resonance

MRI   Magnetic resonance imaging

OAR   Organs at risk

OCC   One-class classification

OCSVM	One-class support vector machine
PSA	Prostate-specific antigen
RBF	Radial basis kernel function
ROC	Receiver operating characteristics
ROI	Region of interest
RT	Radiotherapy
RTO	Radiation oncologist
SD	Standard deviation
SGPR	Sparse Gaussian process regression
SI	Superior-Inferior
SIM	Simulation
SOS	Speed of sound
SV	Seminal vesiculae
SVM	Support vector machine
TAUS	Transabdominal ultrasound
TOF	Time of flight
TPUS	Transperineal ultrasound
TRUS	Transrectal ultrasound
TX	Treatment
US	Ultrasound
USgRT	Ultrasound guided radiotherapy
VMAT	Volumetric-modulated arc therapy

Acknowledgements

The seed of my PhD was planted in the summer of 2014. Back then, my internship in Boston was almost over and I 'met' several Philips employees in a Skype call to discuss possible master thesis projects. Among them, there was a nice sounding Italian who was really passionate about ultrasound imaging. I ended up choosing his project proposal and here we are: not only a master thesis, but also a PhD thesis and uncountable amazing experiences later.

Davide, I cannot thank you enough for dragging me into this adventure. Thanks to you I realized that pursuing a PhD might be the right next step for me and I enjoyed (almost) every minute of it. You gave me a lot of freedom and responsibility, while always being there when I needed you. Even when you left Philips and started working at the other end of the globe, you kept supervising me with a lot of dedication. Many supervisors would probably have done differently and, if this was not enough, you also gave me the opportunity to spend a couple months with your research group in Australia. I will always remember our numerous sessions during which I had to practice my presentations over and over (and over and over) again to make sure that I would correctly explain everything from A to Z. Even though these sessions have caused some sleepless nights, they are one of the many things that have contributed to me becoming a better researcher. Time will learn if we will collaborate again, but so far it has been a blast!

Naast Davide heeft ook professor Frank Verhaegen mijn interesse voor ultrasoundgestuurde radiotherapie aangewakkerd. Frank, ik leerde je kennen toen ik in het kader van mijn masterthesis de reguliere ultrasound meetings bij MAASTRO ging bijwonen. Op het begin moest ik wennen aan jouw recht-door-zee aanpak, maar naarmate de tijd vorderde leerde ik deze aanpak steeds meer te waarderen. Door de jaren heen kon ik bij jou terecht voor gedegen inhoudelijk advies met een klinische inslag. Ik weet zeker dat jouw visie op mijn rebuttals menig reviewer heeft weten te overtuigen. Ondanks dat ik officieel geen onderdeel uitmaakte van het MAASTRO team heb je mij altijd proberen te betrekken bij je andere PhD studenten en dit heb ik als heel fijn ervaren.

Natuurlijk wil ik ook heel graag mijn eerste promotor professor Peter de With bedanken. Peter, aangezien ik al langer met Philips en MAASTRO samenwerkte was het in het begin even zoeken naar een goede begeleidingsvorm. Deze vonden we uiteindelijk in reguliere werkoverleggen waarin je al mijn vragen beantwoordde en mijn papers tot in detail reviewde. Ik wil je bedanken dat je mijn koppige zelfstandigheid niet als een probleem hebt ervaren. Als ik regelmatig contact bleef houden was niks te gek; ontelbare conferenties bezoeken, werken vanuit San Diego en zelfs een aantal maanden in Australië aan de slag. Ik realiseer me heel goed dat dit absoluut niet vanzelfsprekend is. Al die jaren sprak

je waardering uit voor mijn inzet en de kwaliteit van mijn werk en dat heeft mijn zelfvertrouwen en groei als onderzoeker goed gedaan.

I want to thank the members of my PhD committee, Prof. Cees Slump, Dr. Emma Harris, Prof. Harrie Beerlage, Prof. Massimo Mischi and Dr. Rene Aarnink for their time and effort devoted to reviewing this thesis and providing me with comments.

Antonie, Jack, Peter, Rick, Sjef, Aaldert, Jacek, Bart, Ton and all other Philips colleagues I have worked with over the years, thanks to you my PhD journey was extra special. In addition to my academic experiences at the university, you gave me insight into the world of conducting research in industry. I also want to thank Akash, Marco, Meghna, Shyam, Maria, Jacopo, Stephan, Julia and many other Philips interns and PhD students that have made the lunches interesting with the many discussions and walks around the pond at the HTC.

Matilde, Esther, José, Gabriel and Pedro, thanks for the numerous afternoons and evenings that we have spent in the CT room of MAASTRO. Together we have acquired ultrasound images of volunteers and phantoms, made sure that the robot would not collide with the CT scanner, assessed localization systems and so much more. Without your help and dedication it would have been really hard to acquire all the data that has been used in this thesis. Ik wil ook graag de radiotherapeut-oncologen van MAASTRO, Ben, Evert en Ludy, bedanken voor het scannen van patiënten en vrijwilligers, voor het scoren van ultrasound beelden en voor jullie klinische input.

Wat ik toch hele dagen deed daar bij Philips en of daar ook wel eens een schroevendraaier aan te pas kwam? Lieve pap, ondanks veelvuldige uitleg bleef mijn PhD werk al die jaren een mysterieus en abstract gebeuren. Dit weerhielt je er alleen absoluut niet van om aan iedereen die het maar wilde horen zo goed mogelijk over mijn project te vertellen. Door de jaren heen zette jouw kijk op de gang van zaken, bijvoorbeeld omtrent betalingen voor open access journals, mij vaak weer met beide benen op de (niet academische) grond. En als je aan de eettafel in Rijswijk weer eens informeerde hoe het ervoor stond, stelde je regelmatig de vraag die elke PhD student koude rillingen bezorgt: *Lig je nog op schema?* Nu kan ik volmondig zeggen *Ja ik lig nog op schema* en zonder jouw goede adviezen, opbeurende telefoontjes als het even tegenzat, nakijksessies van stukken tekst en nog zo veel meer was dat misschien wel heel anders geweest. . .

Lieve mam, rozumiałaś mój projekt dużo lepiej od papy. Nawet o tyle lepiej, że gdy czasem utknęłam w kodowaniu, to razem dyskutowałyśmy nad możliwym rozwiązaniem. Aby mi dobrze pomóc, zadawałaś pytania aż wszystkie szczegóły stały się jasne. Właśnie dzięki tych wyjaśnień, przychodziły mi do głowy nowe pomysły i rozwiązania problemów. Zawsze pamiętałaś też o moich wysłanych artykułach i z zainteresowaniem pytałaś czy już zostały zaakceptowane przez czasopismo naukowe lub na konferencję. Wcale ci się nie podobało, że tak często latałam po całym świecie, żeby uczestniczyć w tych konferencjach. Mimo tego życzyłaś mi powodzenia i z ciekawością słuchałaś moich opowieści po powrocie

do domu. Dziękuję za wszystkie porady, niezliczoną ilość rozmów telefonicznych, herbat u was przy stole gdy jeszcze do późna pracowałam i o wiele więcej!

Lieve jongens, Jantje en Maxie, als we met z'n tweeën of drieën waren kwam mijn onderzoek niet zo vaak ter sprake, behalve misschien wat verbaasde uitroepen als ik álweer naar de andere kant van de wereld vloog. Eigenlijk was het juist wel fijn om af en toe eens over wat anders te babbelen en dit heeft jullie zeker niet minder betrokken gemaakt. Maxie, jij bent zelfs mijn paranimf tijdens mijn verdediging. Toen ik je vroeg of je dat wilde doen had je geen idee wat dit inhield, maar toen Google je leerde dat dit zoiets is als een getuigen bij een huwelijk en je zelf concludeerde dat je zoals in de rechtbank naast mij moest blijven staan, ging je meteen akkoord.

Mar! Als tweede paranimf mag jij natuurlijk absoluut niet ontbreken in dit dankwoord. Als ik in Nederland was waren onze wekelijkse etentjes echt iets om naar uit te kijken. In de eerste jaren nog redelijk vaak buiten de deur, de laatste tijd lekker bij jou thuis met een uitdagend bordspelletje als toetje. Tijdens deze etentjes bespraken we van alles en nog wat en natuurlijk was mijn PhD daar een vast onderdeel van. Ik denk dat naast mijn begeleiding jij het beste op de hoogte was van wat ik allemaal uitspookte op onderzoeksgebied. Je had mij immers een aantal keer uit de brand geholpen door mijn papers te reviewen. Ik vind het super leuk dat je mijn paranimf wil zijn en dat er nog veel etentjes mogen komen!

In de afgelopen jaren zijn mijn vrienden en familie ieder op hun eigen manier betrokken geweest bij mijn onderzoek. Alle interesse en adviezen hebben mij zeker geholpen en ik beloof dat ik jullie vanaf nu niet meer zal vervelen met verhalen over prostaten ;)

Lieve Jeroen, ik ben trots op het afronden van mijn promotie-onderzoek, maar er is één ding waar ik nog trotser op ben. Wij hebben het het lange-afstands-relatie-monster verslagen! Onze omgeving verklaarde ons voor gek toen ik deze PhD positie overwoog en later zelfs accepteerde, terwijl jij al een aantal maanden van het zonnige San Diego je thuis had gemaakt. Hoe moest dat toch zo ver van elkaar vandaan en hadden we wel een plan? Er was geen plan en met vallen en opstaan leerden wij hoe dat moest, 9108 kilometer bij elkaar vandaan wonen. Skype werd onze beste vriend en een ellelange vliegreis niks anders dan een wandelingetje naar de bakker om de hoek. San Diego was mijn tweede thuis; literatuur bijlezen aan de rand van het zwembad, samen mijn papers reviewen onder het genot van een biertje, weekenden er op uit in Anza Borrego om de gedachten te verzetten, midden in de nacht opstaan om een call bij te wonen met mijn Philips collega's, mijn onderzoek bediscussiëren tijdens wielrentochten en nog zo veel meer. Hoe gek het ook klinkt, ik zou het zo weer over doen (al was een jaartje korter beter geweest). Het heeft ons de kans gegeven om samen, maar toch onafhankelijk, te doen wat we graag wilden en tegelijkertijd heeft het voor een ijzersterke basis gezorgd voor alle avonturen die nog komen gaan. Ons eerste gezamelijk avontuur in Zwitserland is inmiddels begonnen en ik heb er nog lang geen genoeg van!

Curriculum vitae

Saskia Camps was born on the 26th of November 1990 in Breda. She grew up in Rijswijk and went to the Maerlant Lyceum in The Hague for her secondary education. After graduating high school in 2009, Saskia moved to Eindhoven to study Biomedical Engineering at the Eindhoven University of Technology. During her bachelor studies she followed not only biomedical courses, but also courses at the Mathematics department where her interest for medical image analysis kept growing. This resulted in a bachelor thesis project in which she collaborated with a neurosurgeon to perform semi-automatic volume analyses of brain tumor volumes. Next to her studies, Saskia did many extracurricular activities, such as, organizing concert trips abroad for her former youth orchestra, tutoring high school students in mathematics and organizing a large student sports event.

Saskia obtained her Bachelor's degree in 2013, after which she continued her education at the Biomedical Image Analysis research group at the Eindhoven University of Technology. She spent six months at the Surgical Planning Laboratory of Brigham and Women's Hospital in Boston (USA) where she worked as an intern on touchless image data accessibility in intra-operative settings. After coming back to the Netherlands, she started writing her Master thesis focusing on ultrasound guided radiotherapy in collaboration with Philips Research (Eindhoven, the Netherlands) and the MAASTRO Clinic (Maastricht, the Netherlands).

Directly after receiving her Master's degree in 2015, Saskia started working as a PhD candidate in the IMPULS 2 project in a collaboration between Philips Research, the Electrical Engineering Department of the Eindhoven University of Technology and the MAASTRO Clinic. In her PhD research she kept focusing on ultrasound guided radiotherapy under the supervision of prof. Peter de With and prof. Frank Verhaegen. In the third year of her PhD, Saskia spent four months with the research group of dr. Davide Fontanarosa at the Queensland University of Technology (Brisbane, Australia) to explore deep learning approaches for automatic ultrasound image interpretation.