

Time-of-Flight - A New Modality for Radiotherapy

Der Technischen Fakultät der
Universität Erlangen–Nürnberg

zur Erlangung des Grades

DOKTOR–INGENIEUR

vorgelegt von

Christian Schaller

Erlangen — 2011

Als Dissertation genehmigt von der
Technischen Fakultät der
Universität Erlangen-Nürnberg

Tag der Einreichung: 10.03.2011
Tag der Promotion: 14.09.2011
Dekan: Prof. Dr. Reinhard German
Berichterstatter: Prof. Dr. Joachim Hornegger
Prof. Dr. Hans-Peter Meinzer

Zusammenfassung

Diese Dissertation repräsentiert eine der ersten Arbeiten, welche sogenannte Time-of-Flight Kameras im medizinischen Umfeld anwendet. Mit Time-of-Flight Kameras ist es möglich 3-D Modelle mit einem einzigen Sensor in Echtzeit zu erfassen. Es werden im Folgenden verschiedene Systeme vorgestellt. Ziel dieser Systeme ist es Atembewegung in der Strahlentherapie zu kompensieren. Es gibt fünf wesentliche Beiträge in dieser Arbeit: Eine Online-Methode zur Verifikation der Korrelation einer internen Tumor Bewegung zu einem externen Atemsignal, die Anwendung einer neuen Technologie im Bereich der medizinischen Bildverarbeitung und die Vorstellung von drei neuen Systemen: einem System zur Messung der Atmung und zwei Systemen zur Patientenpositionierung: Der Algorithmus der eine externe mit einer internen Bewegungen korreliert ist ein bildbasiertes Synchronisierungsverfahren. Das Verfahren ordnet automatisch fluoroskopische Bilder den jeweilig korrespondierenden 4-D CT Aufnahmen zu. Es ist als Optimierungsprozess aufgebaut und findet die optimale Zuordnung zwischen beiden Sequenzen indem es, unter Berücksichtigung des zeitlichen Ablaufes, die Bildähnlichkeiten maximiert. Das System wird sowohl an synthetischen, als auch Patientendaten evaluiert. Im Mittel werden 93% der Bilder korret zugeordnet.

Das Time-of-Flight basierte Atemdetektionssystem ermöglicht es gleichzeitig verschiedene Atemsignale für unterschiedliche Regionen aufzunehmen. Das System wird mit Hilfe eines neuen mechanischen Körpermodells evaluiert. Tests konnten zeigen, dass das Signal des Systems und das des Körpermodells für Amplituden größer 5 mm mit mehr als 80% korreliert. Die Korrelation beider Systeme ist unabhängig (immer größer 80%) von der jeweiligen Atemfrequenz. Weiterhin werden die Atemkurven des Signales mit einem im klinischen Alltag verwendeten ANZAI Gurt verglichen. Auch hier konnte eine Korrelation von 88% gezeigt werden. Das erste der beiden Positionierungssysteme ermöglicht es ein C-Bogen ähnliches Gerät an einem Patienten auszurichten. Hierfür nimmt eine Time-of-Flight Kamera den kompletten Körper des Patienten auf und unterteilt ihn in verschiedene aussagekräftige Körperregionen, wie z.B. Kopf, Brust, Bauch, Beine, etc. Das System berechnet die 3-D Grenzen (Bounding Boxes) der Körperregionen und berechnet deren Schwerpunkt. Mit Hilfe dieser Information kann der C-Bogen sich selbst positionieren und die Körperregion scannen. Das System wird an einem Körperphantom evaluiert und es kann eine Genauigkeit von 1 cm gezeigt werden. Das zweite System beschäftigt sich mit der oberflächenbasierten Positionierung von Patienten in der Strahlentherapie. Das Verfahren nutzt einen speziell für Time-of-Flight Daten angepassten Iterative-Closest-Point Algorithmus. Das System erreicht eine Genauigkeit von $0.74 \text{ mm} \pm 0.37 \text{ mm}$ für Verschiebungen bis zu 10 mm in alle drei Raumrichtungen.

Abstract

In this work, one of the first approaches utilizing so-called Time-of-Flight cameras for medical applications is presented. Using Time-of-Flight cameras it is feasible to acquire a 3-D model in real-time with a single sensor. Several systems for managing motion within radiotherapy are presented.

There are five major contributions in this work: A method to verify internal tumor movement with an external respiratory signal on-line, the application of a novel technology to medical image processing and the introduction of three novel systems, one to measure respiratory motion and two other to position patients.

The algorithm to correlate external and internal motion is an image-based synchronization procedure that automatically labels pre-treatment fluoroscopic images with corresponding 4-D CT phases. It is designed as an optimization process and finds the optimal mapping between both sequences by maximizing the image similarity between the corresponding pairs while preserving a temporal coherency. It is both evaluated at synthetic and patient data and an average of 93% correctly labeled frames could be achieved.

The Time-of-Flight based respiratory motion system enables the simultaneously measurement of different regions. We evaluate the system using a novel body phantom. Tests showed, that the system signal and the ground truth signal of the phantom have a reliable correlation of more than 80% for amplitudes greater 5 mm. The correlation of both systems is independent (always more than 80%) of the respiratory frequency. Furthermore, the measured signals were compared with a well-established external gating system, the Anzai belt. These experiments were performed on human persons. We could show a correlation of about 88% of our system and the Anzai system.

The first positioning system is able to position a C-arm like device with respect to the patient. Therefore, a Time-of-Flight camera acquires the whole body of the patient and segments it into meaningful anatomical regions, like head, thorax, abdomen, legs. The system computes 3-D bounding boxes of the anatomical regions and computes the iso-center of the boxes. Using this information, the C-arm system can automatically position itself and perform a scan. The system is evaluated using a body phantom and an accuracy within the patient table accuracy of 1 cm could be shown. The second system deals with surface-based positioning of a patient with respect to a priorly acquired surface of the same patient. Such systems are necessary, e.g. in radiotherapy or multi-modal imaging. The method uses an Iterative-Closest-Point algorithm, tailored to Time-of-Flight cameras. It is evaluated using a body phantom and obtaines an overall accuracy of $0.74 \text{ mm} \pm 0.37 \text{ mm}$ for translations in all three room directions within 10 mm.

Acknowledgements

Dear Kristin thank you very much for your support during the last months while writing this thesis and since we know each other. I thank also my parents for making it possible to study computer science and their support during my whole life.

Having the opportunity of working on a PhD at the Lehrstuhl für Mustererkennung (LME) together with Prof. Dr.-Ing. Joachim Hornegger is a great chance. I would like to thank the whole research group at the LME and especially Achim for supporting me and my work during the last three and a half years. Special thanks goes to all who contributed to this work directly or indirectly: André, Christian, Kerstin, Peter and Simon. Furthermore, I thank Jochen for pointing me towards Time-of-Flight cameras and working together with him for several years on this topic. This includes also Michael S., who provided all of us a common framework to handle Time-of-Flight data. At Siemens Corporate Research, I especially thank Ali and Frank for giving me the chance to do research on respiratory motion related topics within radiotherapy. Working with both of you set the basis for this thesis.

During my time at the LME, I got various financial support. Thanks goes to the International Max Planck Research School for Optics and Imaging, softgate GmbH and the LME.

Last but not least, I would like to thank Michael B. and Simon for starting a company dealing with Time-of-Flight applications in the beginning of 2010 together with me: the Metrilus GmbH.

Contents

1	Introduction	1
1.1	Introduction	1
1.2	Contributions	4
1.3	Structure	5
2	Basic Principles of Radiotherapy	7
2.1	Treatment Planning	7
2.1.1	Planning CT	7
2.1.2	Clinical Target Volumes	8
2.1.3	Aim of Treatment Planning	10
2.2	Motion Management in Radiotherapy	11
2.2.1	Introduction	11
2.2.2	Gated Radiotherapy	12
2.3	Clinical Procedure	15
2.4	Summary	17
3	Motion Management	19
3.1	Introduction	19
3.2	Related Work	20
3.3	Materials and Methods	21
3.3.1	Problem Statement	21
3.3.2	Image Similarity Matrix	22
3.3.3	Respiratory Motion as a Markov Process	24
3.3.4	Mapping as a Solution of an Optimization Problem	24
3.4	Experiments and Results	27
3.5	Conclusion	30
4	Time-of-Flight Cameras	31
4.1	Introduction	31
4.2	Working Principle	34
4.3	Key Benefits of Time-of-Flight Cameras	37
4.4	Time-of-Flight Camera Calibration	40
4.4.1	Optical Calibration	40
4.4.2	Distance Calibration	43
4.5	Conclusion	47

5 Time-of-Flight based Patient Positioning	49
5.1 Introduction	49
5.2 Treatment Table Calibration	50
5.3 Patient Positioning for C-arm Systems	55
5.3.1 Introduction	55
5.3.2 Related Work	55
5.3.3 Materials and Methods	55
5.3.4 Experiments and Results	59
5.3.5 Summary	60
5.4 Patient Positioning for Radiotherapy	61
5.4.1 Introduction	61
5.4.2 Related Work	62
5.4.3 Materials and Methods	65
5.4.4 Experiments and Results	71
5.4.5 Summary	74
5.5 Conclusion	74
6 Time-of-Flight based Respiratory Motion Gating	77
6.1 Introduction	77
6.2 Related Work	78
6.2.1 Other Sensors	78
6.2.2 Optical Sensors	79
6.2.3 Discussion	80
6.3 Materials and Methods	81
6.3.1 Table Detection	83
6.3.2 Respiratory Motion Detection	85
6.4 Experiments and Results	88
6.4.1 Phantom Experiments	88
6.4.2 Patient Experiments	90
6.5 Conclusion	90
7 Summary and Outlook	93
7.1 Summary	93
7.2 Outlook	95
Bibliography	97

CHAPTER I

Introduction

1.1 Introduction

Lung and bronchus cancer are still the most fatal types of cancer, even though the cases of death have been declining in the last years. Regarding to the American Cancer Society (ACS), cancer within the respiratory system will be the reason for almost 29% of all cancer deaths in the United States in 2010. There are 222,520 new estimated cases diagnosed with an estimated number of 157,300 deaths. Although one can recognize a continuous downward movement of deaths caused by respiratory system cancer in the past 20 years (see Figure 1.1), it is important to improve early detection and treatment continuously. The one-year relative survival for lung cancer increased by 7% from 35% (1975-1979) to 42% (2002-2005). This is due to improvements in surgical techniques and combined therapies. However, the 5-year survival rate for all stages of lung cancer is only 16% [Amer 10].

There are several ways to treat lung cancer. The treatment options are determined by the type (small cell or non-small cell) and stage of cancer:

- surgery
- radiotherapy
- chemotherapy
- targeted therapies

Within this work, we will focus on radiotherapy and tools to improve this treatment option. The term radiotherapy describes the medical application of ionizing radiation to control malignant cells. It can be applied for various cases like curative, adjuvant or palliative treatment, and can also be combined with other methods like surgery or chemotherapy. Figure 1.3 shows a typical linear particle accelerator (LINAC). The device is used to deliver radiation to the patient. Using the patient table and the rotating gantry tumors can be irradiated from various directions.

A major goal in cancer treatment is increasing the local control of tumors via dose escalation. At the same time, it should be provided that the severity of normal tissue complication

1 Introduction

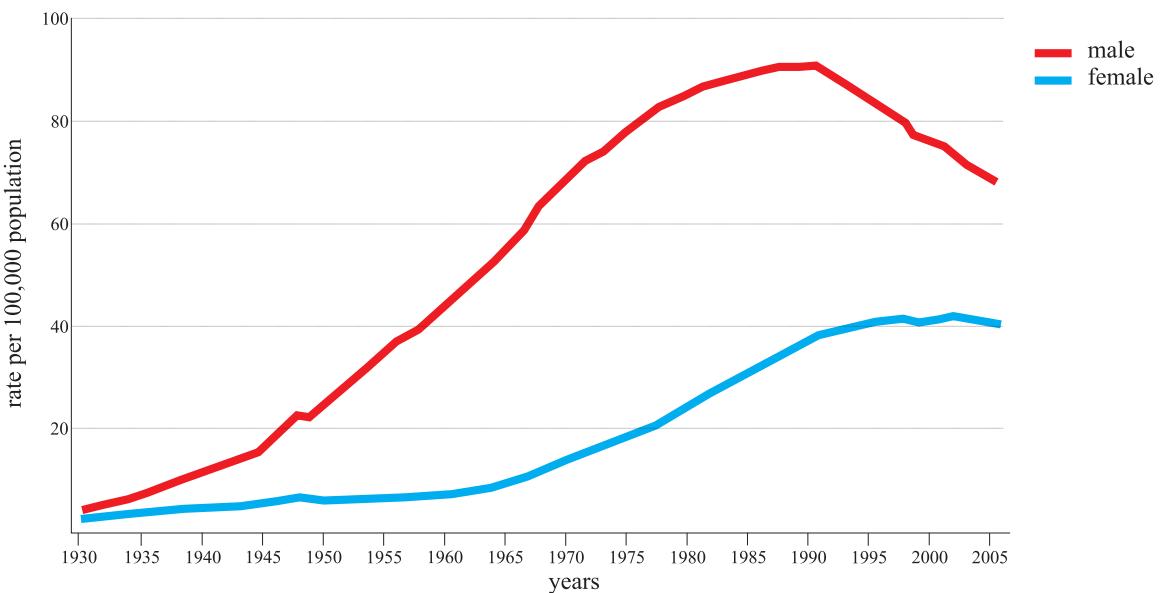


Figure 1.1: Age-adjusted cancer death rates (per 100,000, age adjusted to the 2000 US standard population) for lung and bronchus cancer, US, 1930-2006 [Amer 10].

is reduced. To achieve this goal, a precise localization of potential moving targets during the course of treatment is mandatory. This is exceedingly important for tumors in the thorax and abdomen area. Tumors located in these areas are affected by different sources of motion: the heart beat, breathing of the patient and general patient motion. To accomplish a successful treatment, several tasks within a radiotherapy workflow have to be accounted.

The four pillars in Figure 1.2 (showing a typical radiotherapy workflow) heavily depend on each other. To increase the local control of a tumor, its location first has to be determined. Therefore, a so-called planning CT (computed tomography) is taken. Commonly this is a usual CT scan acquired once a few days or even weeks prior to the particular treatment session. The dose applied to tumors which are located at non-critical positions in the body is usually applied in fractions. As a consequence various treatment sessions over a period of days / weeks are necessary. This process is called fractional treatment. Each of these treatment sessions refers to the planning CT. For critical tumor locations like in the brain, a single treatment session is preferred and the whole dose is applied once.

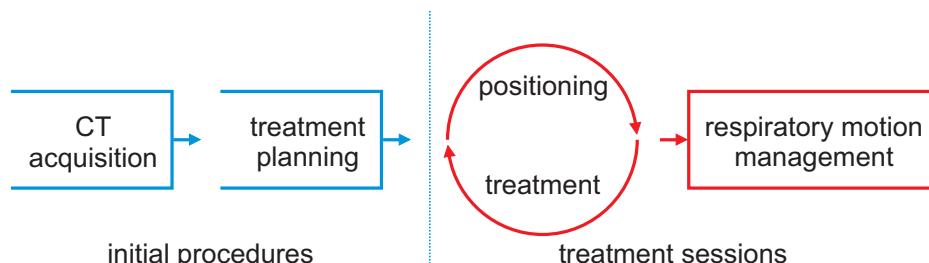


Figure 1.2: Schematic overview of a conventional radiotherapy workflow. The process usually starts with the acquisition of a CT, followed by the treatment planning. This initial procedure is usually done only once. Afterwards, the treatment sessions require patient positioning and also a motion management strategy, depending on the location of the tumor.

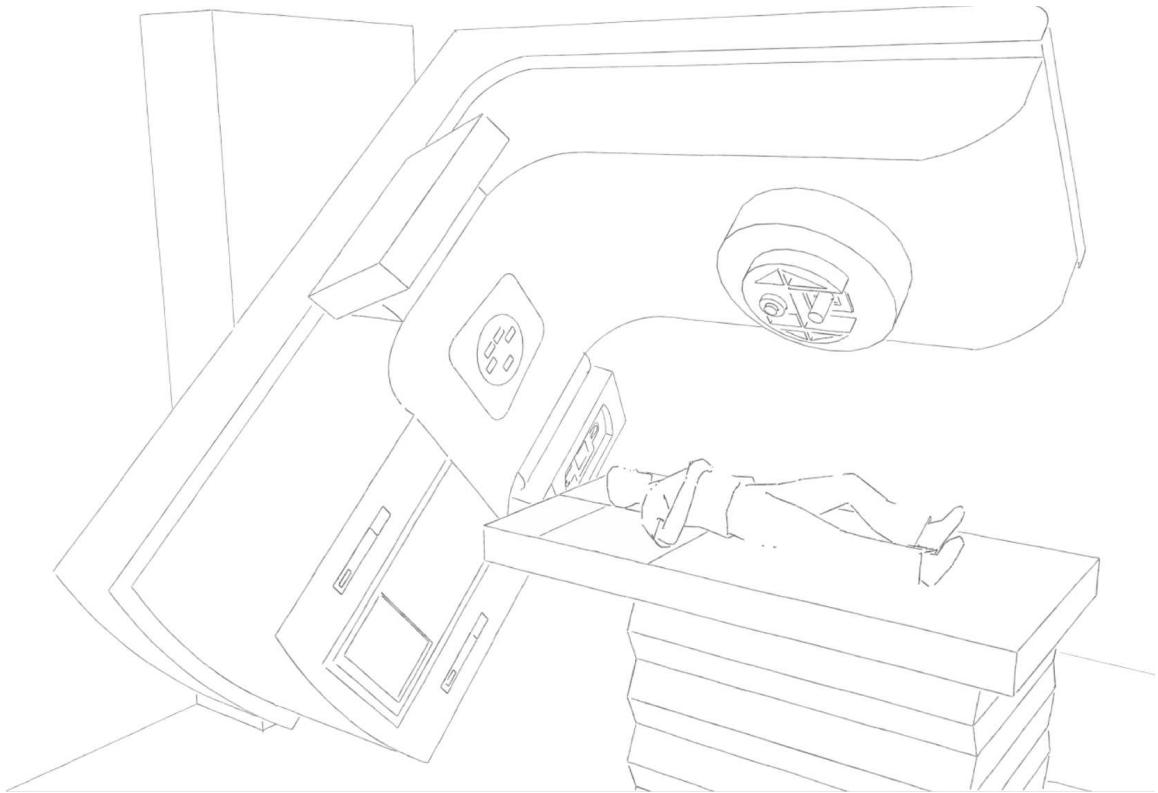


Figure 1.3: LINAC with rotating gantry and patient table. Source: based on linear particle accelerator model (Google Sketchup 3-D Warehouse ^a).

^a<http://sketchup.google.com/3dwarehouse/details?mid=dfafdfb263e16ef6abfe687359b7b6cf>

Both fractional and singular treatment sessions require a precise radiation plan. Based on the planning CT, the tumor is located and the radiation plan including dose distributions is precomputed. This is taken as a base for the exact delivery location of radiation in the patient's body.

Having the treatment plan set up, a problem arises: The patient cannot be treated on the CT scanner, as a LINAC is needed for radiation therapy. Hence, the patient has to be moved between the two modalities. This is essential for fractional treatment, where sometimes a long period of time lays between each session. The most accurate treatment plan is worthless, if one fails to properly align the patient according to the plan. For that reason the placement of the patient has to be done very accurately. By aligning the planning coordinate system with the patient coordinate system less harm for healthy tissue and maximum dose for the tumor can be expected. This results in a shorter and a less harmful radiation treatment.

If so-called moving targets exist within the scope of the treatment plan, a motion management strategy is mandatory. The problem of placing the dose precisely into the tumor and not to harm healthy tissue, is even more important, if the tumor is moving.

Within this work several strategies are presented to cope with these problems.

1.2 Contributions

The main focus of this work is on the treatment sessions and tools to handle the previously identified problems. This includes a system to position the patient each treatment session and a system to account for respiratory motion (see also Figure 1.2 right side). We also set a focus on cost effectiveness in this work and realize all the systems using a novel technology called Time-of-Flight (ToF) cameras. A single ToF camera serves as core component of each system. An algorithm to correlate an external respiratory signal and the internal movement of the tumor is presented as a scope for the ToF-based systems in advance. Among several minor contributions these major contributions arise from this work:

- **On-line verification of internal tumor movement with an external respiratory signal:** Well established methods in pattern recognition, like Markov processes are used to solve an important task in cancer treatment. Using the proposed method, it is possible to link the directly observed internal movement of the tumor prior to the radiation to an external surrogate respiratory signal. Hence, it is possible to focus the radiation beam more accurately on tumor tissue as just assuming some correlation between internal tumor motion and external respiratory signal.
- **Application of a novel technology to medical image processing:** Time-of-Flight cameras are a novel technology to acquire 3-D data in real time. Today these cameras are mainly used for research in the automotive- and consumer electronics industry. This work presents several systems based on this technology, which can seamlessly be integrated into existing radiotherapy workflows. The proposed systems provide a cost-effective alternative to existing technologies. Once, a ToF camera is included into the workflow, this technology is not limited to solve a single problem (like existing technologies), rather it can serve several tasks, like patient positioning, motion management, etc. This work represents one of the first approaches where ToF technology is implemented within a medical environment.
- **Improvements of this novel technologies to meet the demands of the given problem:** Off-the-shelf ToF cameras are not capable to solve challenging tasks, like patient positioning. An accuracy of about one millimeter is required to make the system reasonable for this task. Current cameras provide a per-pixel accuracy (in depth) within the range of one centimeter. Within this work, accurate calibration algorithms and pre-processing steps are introduced to improve the ToF camera data to meet the demands of patient positioning.
- **Introduction of a novel system to measure respiratory motion:** A ToF camera is also utilized to introduce a novel system to account for respiratory motion. The system is able to measure respiratory motion without touching the patient or the use of markers. Furthermore, the system can measure respiratory motion simultaneously on multiple positions in real-time. The combination of these features is novel and unique.
- **General improvements to the application of gated radiotherapy:** Today, the application of respiratory gating is limited due to intra-fraction deformations [Sieb 07b]. The combination of the above contributions leads towards a more robust application of gated radiotherapy. This results from a method to verify internal mo-

tion with an external respiratory signal (Chapter 3), which uses a novel positioning system (Chapter 5) and a novel system to account for respiratory motion (Chapter 6).

1.3 Structure

Chapter 2

The medical background for radiotherapy is given in this chapter. Basics about treatment planning and clinical target volumes are introduced. Furthermore, motion management in radiotherapy including gated radiotherapy is explained.

Chapter 3

This chapter introduces a general method to verify internal and external correlation of moving targets. It gives also the motivation of why the systems and algorithms described in the following chapters are of note for a general motion management system. An algorithm is introduced to combine the advantages of internal gating and to discard the disadvantages of external gating. This work is published in the following papers: [Scha 07, Kham 07]. Furthermore, a patent application was filed¹. Both papers provide a framework for motion management. The framework requires two core components: A system to position the patient prior to the treatment and a system to detect and measure the respiratory motion of the patient.

Chapter 4

The chapter deals with ToF cameras. It includes the working principle of the ToF camera. Furthermore it differentiates the technology from others, like stereo, structured light and laser scanners. Two essential calibration methods are presented in this chapter as well: an optical and a distance calibration.

Chapter 5

Two distinct systems using a Time-of-Flight camera are introduced in this chapter. The first system is able to position a C-arm like system (e.g. a LINAC) with respect to the patient. The second system can be used to align a patient e.g. for radiotherapy. This work is published within the following papers:

- Adelt et al.: Within the framework of a student thesis, pre-examinations were conducted. The feasibility of positioning patients using a Time-of-Flight camera was shown [Adel 08].
- Höller et al.: This work gives a resume of the development of medical applications using Time-of-Flight cameras [Holl 08].
- Ringbeck et al.: An overview paper about prospects of Time-of-Flight cameras for medical applications [Ring 09].

¹US. Patent. No. US7570738

- Schaller et al.: The paper introduces patient positioning using Time-of-Flight cameras [Scha 09a].
- Schaller et al.: C-arm positioning for interventional procedures. The paper was awarded with the "“poster of the day award”" at MICCAI 2009 [Scha 09b].
- Placht et al.: Improvements to the algorithm in [Scha 09a] and a detailed evaluation of the system. The evaluation was performed using the body phantom which was introduced within Chapter 4 [Plac 10b].

After a general introduction, a method is introduced to align the patient table coordinate system and the ToF camera coordinate system. The system for inverse C-arm positioning including its evaluation is presented afterwards. In the end of this chapter, a system for patient positioning in general is introduced including its evaluation.

Chapter 6

This chapter introduces a system enabling the detection of respiratory motion. The system does not use any marker or has to be in contact with the patient. This work is among the first approaches to utilize Time-of-Flight cameras for medical applications and is published in the following papers:

- Höller et al.: This work gives a resume of the development of medical applications using Time-of-Flight cameras [Holl 08].
- Müller et al.: Within the framework of a student thesis an extension to [Scha 08] is introduced. The paper includes the measurement of breathing volumes [Mull 09].
- Penne et al.: The system presented in [Scha 08] is examined with respect to noise aspects of Time-of-Flight cameras [Penn 08a, Penn 08b].
- Schaller et al.: The paper introduces respiratory motion gating using a Time-of-Flight camera. It includes a comparison to the well-established ANZAI belt system [Scha 08].
- Schuhmann et al.: Within the framework of a master thesis, pre-examinations were done. The feasibility of respiratory motion detection using a Time-of-Flight camera is shown [Schu 08].
- Ulrich et al.: Within the framework of a master thesis a plaster cast phantom was built and developed. The phantom provides a ground truth and is used to evaluate surface-based respiratory motion detection systems [Ulri 10].

Furthermore, a patent application was filed ².

Chapter 7

The chapter will conclude the work and will give an outlook to potential applications of the proposed system within radiotherapy and other modalities.

²German Patent No. DE102008033243

CHAPTER II

Basic Principles of Radiotherapy

Within this chapter basic principles of radiotherapy are introduced. Starting with treatment planning, the chapter introduces common clinical target volumes and the aim of treatment planning. Afterwards, motion management in radiotherapy is discussed and the concept of gated radiotherapy is presented. The chapter concludes with a recommendation for a clinical procedure for radiotherapy.

2.1 Treatment Planning

The purpose of treatment planning is to locate tumors in the human body for the actual process of radiotherapy. Therefore, physicians locate tumors and surrounding tissue in the planning CT scan. This information is used to define the dose which has to be delivered to the tumor. Based on the planning, several margins are set up. These margins are also called clinical target volumes and are used for each fractional treatment session.

2.1.1 Planning CT

The most recent technology to acquire planning CT data is called 4-D CT. Using this technology dynamic information can be captured. The mean tumor position and the range of the tumor motion can be determined by using this method. Figure 2.1 shows a schematic overview of a 4-D CT acquisition process. The patient table of the CT scanner is moving back and forth during the acquisition and multiple slices are acquired for each table position. Additionally, a surrogate signal capturing the respiratory motion is acquired. Each table position delivers several time synchronized images showing different breathing phases. The complete 4-D CT sequence consists of multiple 3-D CT datasets. To obtain these 3-D volumes, the images are sorted. The surrogate signal acts as a trigger signal for the sorting. After sorting the images, each 3-D volume represents a single breathing phase. However, there are also limitations for this method. Artifacts can occur, if the respiratory pattern varies during the acquisition process of the 4-D CT. This is caused if the patient

does not breath with constant speed and / or amplitude. A typical 4-D CT sequence can cover from about 8 to 25 phases.

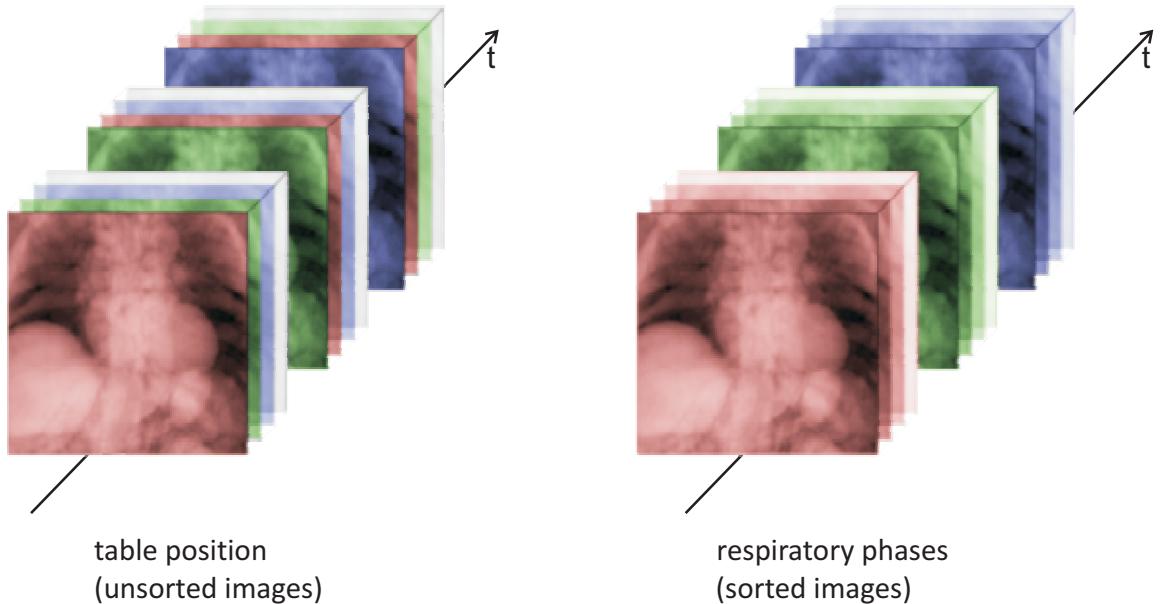


Figure 2.1: A schematic overview of the 4-D CT process. At each table position images for many respiratory phases are acquired (images of the same respiratory phase are indicated by colors). Afterwards, the time synchronized images are sorted and put together to a 3-D CT volume sequence.

2.1.2 Clinical Target Volumes

According to the ICRU Report 50 and its supplement 62 [Doug 94] the following recommendations emerge to address certain tumor- and treatment regions. These regions are defined within the planning phase and used as a basis for the radiation plan.

- **Gross Tumor Volume (GTV):**

The gross palpable or visible / demonstrable extent and location of the malignant growth. The GTV covers the primary tumor (GTV primary), lymph nodes (GTV nodal) and metastasis (GTV M).

- **Clinical Target Volume (CTV):**

Tissue volume that contains a GTV and / or sub-clinical microscopic malignant disease which has to be eliminated. This volume has to be related adequately in order to achieve the aim of therapy: cure or palliation. The CTV is part of the radiotherapy directions and covers single malignant cells, small cell clusters or non-detectable tumor cells.

- **Planning Target Volume (PTV):**

This region is a geometrical concept and defined to select an appropriate beam size and beam arrangements. It takes into consideration all possible geometrical variations and inaccuracies to ensure that the prescribed dose is absorbed in the CTV. The PTV includes the Internal Margin (IM) as well as the Setup Margin (SM). The region depends on intra- and inter-fractional movement (see Section 2.2), organ movement and patient movement. It particularly includes healthy tissue.

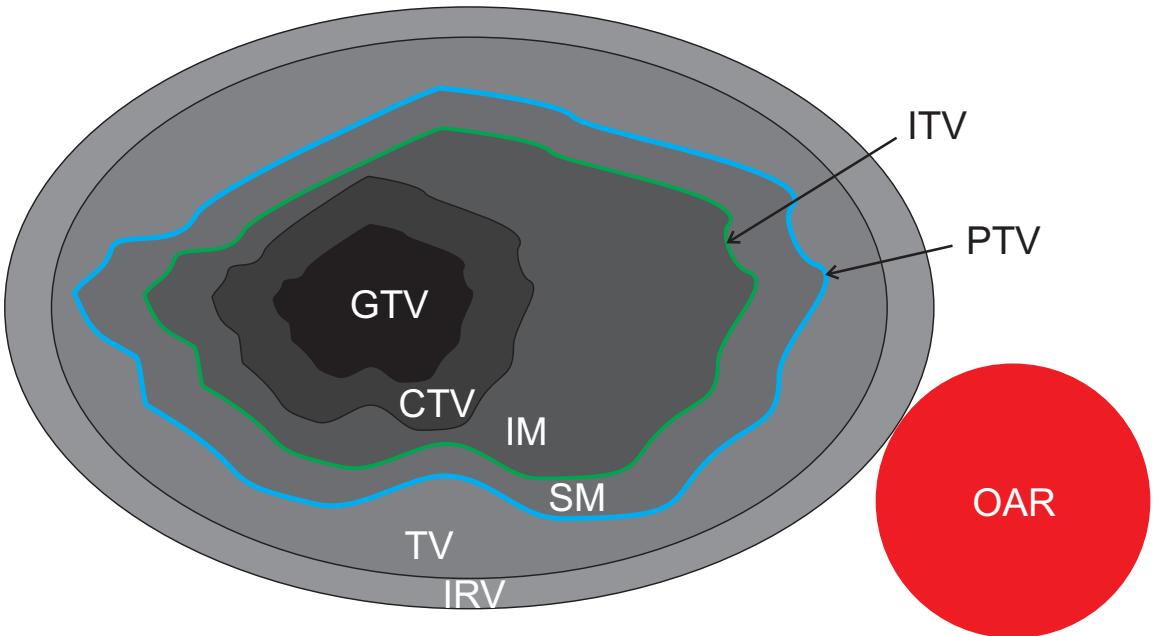


Figure 2.2: Clinical target volumes, according to the ICRU Reports 50 and its supplement 62 [Doug 94]. The ITV/PTV regions, which should be reduced to a minimum are indicated in red.

■ **Internal Margin (IM):**

Considers variations in size, shape, and position of the CTV in relation to anatomical reference points (e.g. filling of stomach, bladder, movements due to respiration, etc.). The underlying physiological processes causing these issues are usually very hard to control.

■ **Setup Margin (SM):**

Added to take into account all uncertainties in patient-beam positioning. This region is mostly depending on technical factors. It can be improved (shrinked) by accurate positioning and / or immobilization of the patient.

■ **Internal Target Volume (ITV):**

Defined as the sum of Clinical Target Volume (CTV) and Internal Margin (IM).

■ **Treated Volume (TV):**

Volume enclosed by an iso-dose surface, selected and specified by the radiation oncologist as being appropriate to achieve the purpose of treatment (e.g. tumor eradication, palliation). In an ideal case, the TV equals the PTV.

■ **Irradiated Volume (IRV):**

Volume which receives a dose that is considered significant in relation to normal tissue tolerance.

■ **Organs at risk (OAR):**

Normal tissues whose radiation sensitivity may significantly influence treatment planning and / or prescribed dose.

2.1.3 Aim of Treatment Planning

There exist a variety of software solutions to guide through the planning process. Figure 2.3 exemplarily shows a planning system from RaySearch and the MGH. The output of these software packages is used in every treatment session. Therefore, physicians have to take care that these plans are valid in each session. The patient has to be aligned exactly within the same coordinate system like he was placed during the planning CT scan acquisition. This is mandatory, as the LINAC is supposed to shoot on the iso-center of the tumor. If so-called moving targets are irradiated, this motion has also to be considered for in the treatment plan.

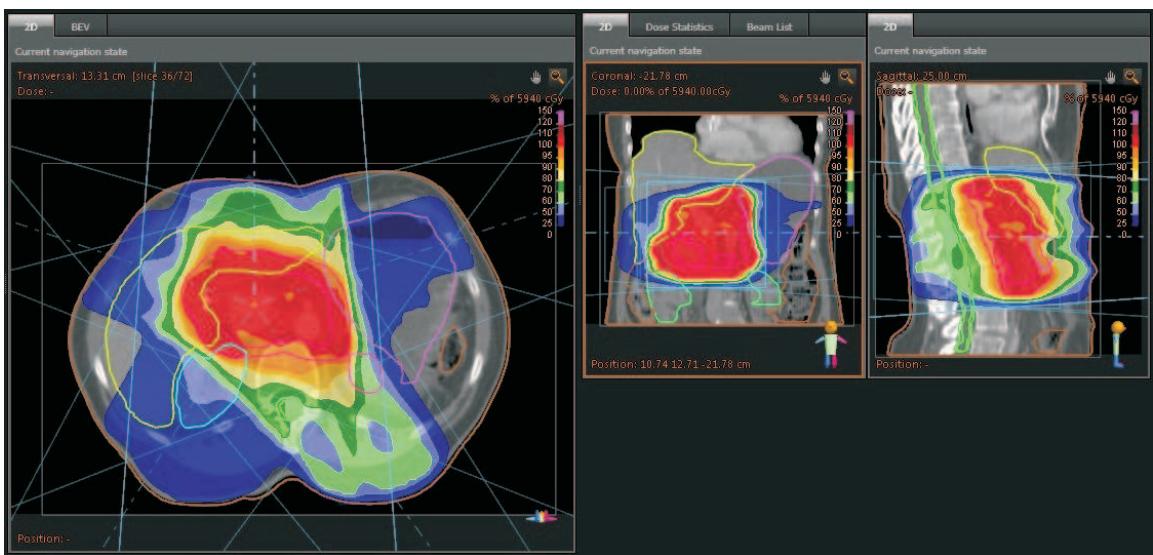


Figure 2.3: Screenshot of a treatment planning system from RaySearch and the Massachusetts General Hospital (MGH). The distribution of the applied radiation is displayed as a color-coded overlay. Source: <http://gray.mgh.harvard.edu/>

The main objective of the treatment plan is to reduce the gap between the CTV and PTV. It is obvious, that this gap should be as small as possible. Amongst others, the following components take influence on CTV-PTV margins definitions:

- **Inter- and intra-observer variations in GTV delineation:**

Due to different experience and education of physicians, target regions often are defined differently [Gira 02]. A selected GTV region is dependent on the physician and not unique.

- **Motion artifacts in the CT scan (during planning):**

As patient breathing and cardiac activity is present during the acquisition of a planning CT scan, these two components could cause artifacts. These artifacts lead to systematic errors during treatment because the defined region is based upon this data.

- **Respiratory motion and heartbeat (during delivery):**

As in the CT scan, respiration and heartbeat also occurs during treatment and causes tumor motion [Sepp 02].

■ **Daily variations of respiratory motion:**

Due to the fact that the planning sequence is usually acquired several days / weeks prior to the treatment, daily variations of respiratory motion have a quite significant effect on tumor position and movement [Sepp 02, Neic 06].

■ **Treatment-related anatomic changes including tumor size:**

As the goal of radiation therapy is a shrinkage of the tumor during treatment these changes are also effecting the shape of the regions and have to be taken into consideration for long-term treatment processes.

■ **Patient setup error:**

Because the region is defined using a prior acquisition, the patient has to be positioned exactly at the same position for the treatment. Possible positioning errors should also taken into consideration while planning the regions [Erri 03].

The proposed system and algorithms in this work will target issue numbers 2, 3, 4 and 6.

2.2 Motion Management in Radiotherapy

2.2.1 Introduction

Receiving radiation therapy in the thorax and abdomen area is a major health risk for patients. Usually, a radiation therapy for a specific patient is split into several treatment sessions. The overall dose is separated into small fractions. The dose spreading over time allows a better recovery of non tumor cells affected by the radiation process. Nevertheless, it is still not optimal for the patient that healthy tissue catches radiation as well. Due to practical issues and cost-effective reasons, the planning of the treatment sessions is only done once in advance of the whole treatment. Based on this plan the tumor is irradiated in each session. Therefore, two specific problems arise within fractional treatment:

- Intra-fraction motion (during the actual session) is mainly caused by respiratory motion and the heartbeat and highly appears at tumors in or near the lung.
- Inter-fraction motion (between different fractions) is caused by various factors like gravity, gain / loss of weight, full / empty bladder, etc. Therefore, tissue, organ, and tumor positions can vary during sessions.

Both problems induce tumor motion which has to be handled by a motion management system. Furthermore, there is a clinical evidence of a survival advantage for higher dose levels [Okun 95, Pere 86, Mach 05]. Without any knowledge about the motion of a moving tumor, the surrounding healthy tissue would have to be irradiated as well. Hence, healthy tissue would be harmed even more, if a successful treatment based on higher dose levels was implemented. Therefore it is very significant to reduce the radiation delivered to healthy tissue to a minimum. This is only possible, if either the direct tumor motion or the main source of the tumor motion is known: the respiratory motion. Using the state-of-the-art technology there are three major approaches for dealing with tumor motion in radiation treatment.

- Breath-hold: The patient is asked to hold the breath or to do a forced shallow breathing. This results in less motion for the tumor.
- Tracking: The patient can breath freely and no compliance is expected. The radiation beam or table is moved so that the tumor always remains at a fixed point within the treatment field [Schw 00].
- Gating: The radiation beam is gated in a way that the tumor is only treated at the prescribed or planned position [Kubo 96]. The patient can also breath freely using this method.

The major drawback of the first method are the high requirements on the patient. Most patients cannot fully cooperate as their state of health is crucial. Furthermore, the implementation of the strategy is rather subjective and it depends on the skill of the therapist and support of the patient. Although, tracking has the advantage of a higher duty cycle (ratio of beam on time to the total treatment time) and a shorter delivery time, it is technically challenging to implement [Rang 05]. In contrast, respiratory gating is more practical and has been adopted in clinical practice by a number of cancer centers [Keal 06]. For both methods gating and tracking, a surrogate signal representing the respiratory cycle is required. It is assumed, that the surrogate signal and the tumor motion correlate. Within this work, we will focus on gating applications only. [Sieb 07b] refers to the limitations of respiratory gating due to intra-fraction deformations of organs during treatment. Therefore, for accurate gating, methods that only use valid (existing) correspondences for gating and which are able to compensation patient motion are required. A solution to cope with this problem is presented in this work.

2.2.2 Gated Radiotherapy

Respiratory gated radiation therapy was developed in Japan in the late 1980s and early 1990s [Ohar 89]. The general idea of gated radiotherapy is to reduce the incidence and severity of normal tissue complications and to increase local control through dose escalation. Therefore, the main purpose of gated radiotherapy is the reduction of the ITV. In an idealized gated treatment, the tumor position should be directly detected and the delivery of radiation is only allowed when the tumor is at the right position. However, the direct detection of the tumor mass in real-time during the treatment often is difficult or even impossible. The system needs to know the source of the tumor motion: the respiration of the patient. The respiratory signal acts as a surrogate signal. Therefore, gated radiotherapy relies on the just mentioned correlation between tumor motion and the patient respiration.

Only if the breathing cycle is within a certain state, an irradiation of the tumor is performed. It is assumed that the tumor is roughly at the same location within the body each time when the patient is in a specific breathing cycle state. When the tumor is moving within this specified part of the breathing cycle, it is moving within the so called "gate" or "gating window" (see Figure 2.4). According to this term the name of gated radiotherapy is derived.

There are two different types of gating:

- Displacement Gating:

The relative position between two extrema of the breathing motion (inhalation and

exhalation) is measured. The beam is activated whenever the respiration signal is within a range of relative positions.

- Phase Gating:

The respiratory cycle is divided into several breathing phases, like begin of inhale, maximum inhale or end of exhale. These phases are also often referred to percentage levels, like 10% inhale or 10% exhale. Whenever the patient's respiration is within certain phases, the beam is turned on.

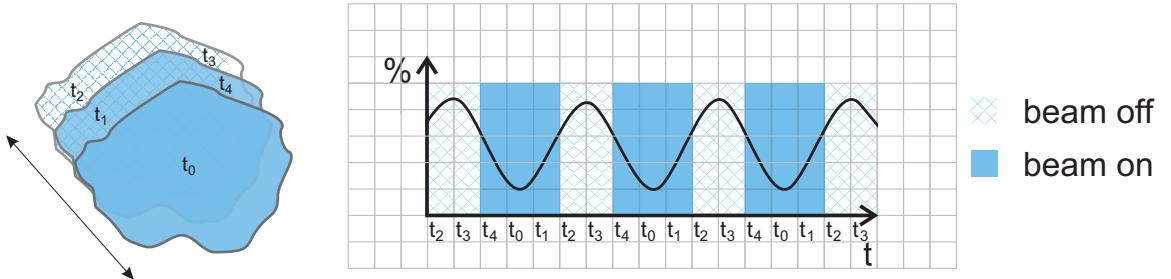


Figure 2.4: Phase gating. Left: Moving tumor at different points in time. Right: Corresponding breathing cycle with a gate applied on t_0 , t_1 and t_4 .

Apparently, the duration of a treatment session is extended if gated radiation therapy is applied. The beam is not able to deliver X-rays continuously. However, this leads to the primary objective of gated radiation therapy. Treating the tumor just at specific positions opens a potential for a significant decrease of the ITV / PTV margin. Therefore tumors can be treated with higher dose levels. According to this, gated radiotherapy enables the possibility of exposing less healthy tissue to high dose levels and therefore the patient has a chance to achieve a better long-term survival rate.

Some tumor motion still occurs within the gate, this motion is also called residual motion. However, there is always a trade-off between duty cycle and residual motion. Increasing the duty cycle results in a larger residual motion of the tumor and vice versa. Usually both aspects are already taken into consideration during the treatment planning. According to physicians there is usually a larger duty cycle achieved during exhale than during inhale¹. Therefore most patients are treated at full exhale if there is a larger duty cycle required. For a more accurate gating and a stable residual motion patients are gated at full inhale, because this breathing phase is more reproducible than full exhale.

Basically, there are two methods to acquire the surrogate respiratory signal (see Figure 2.5):

- External Gating:

A correlation between an external signal and the tumor movement is assumed. It might be quite inaccurate as there often is only a poor correlation between the external surrogate and the real tumor position. The big advantage of this method is that it is non-invasive and has no side effects for the patient.

- Internal Gating:

A marker is implanted near the tumor. This requires additional dose, as image acquisition is necessary to locate the implanted marker properly. The additional dose can

¹Based on discussions with various physicians

be more than what is clinically acceptable for patients with many treatment fractions or a long treatment time of a single fraction. For children it is extraordinary important to avoid unnecessary dose. Furthermore internal gating is difficult for thoracic tumors or even not possible because fiducial markers cannot be inserted and there is a risk of pneumothorax.

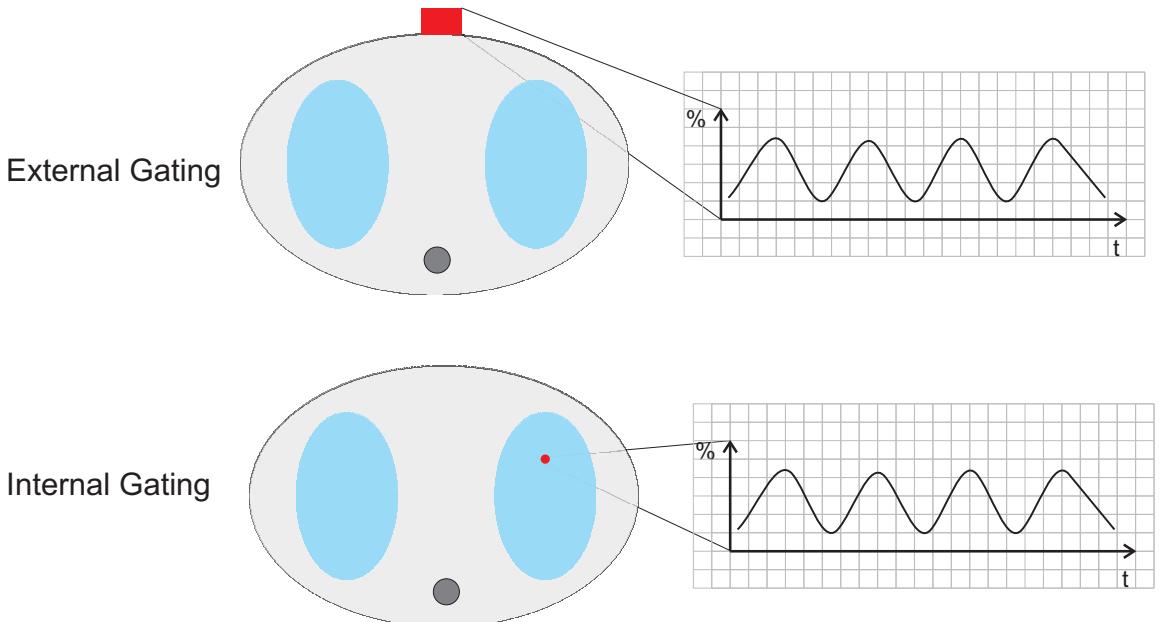


Figure 2.5: External Gating: An external surrogate (red box) is used to measure e.g. chest or abdominal movement. A correlation between the tumor movement and the surrogate signal is assumed. Internal Gating: A fiducial marker (red circle) is implanted near the tumor. The information gained through tracking the fiducial marker is used as a surrogate respiratory signal.

Both methods are based on the observation of a surrogate, where the correlation between an internal surrogate is considered to be more accurate than a correlation between an external surrogate and the real tumor position. Considering these two methods, there are various problems and each method has its own advantages and disadvantages, respectively.

A novel external gating system will be presented within this work (see Chapter 6). Therefore, we will give a short overview of external gating systems. A first feasibility study with a Varian 2100C accelerator, as well as an evaluation of different external surrogates signals were made at the University of California at Davis [Kubo 96]. Later, in a joint venture with Varian Medical Systems, a gated radiotherapy system using a video camera to track infrared reflective markers on a patients' abdomen was developed [Ford 02]. This system was commercialized later by Varian and is one of the most widely discussed external respiration systems today. A similar system is developed by BrainLAB. This system is also able to determine the internal anatomy position using X-ray imaging during treatment. Siemens Healthcare e.g. offers a gating interface with its linear accelerator using a belt measuring pressure. The system is developed by Anzai Medical. VisionRT recently introduced their systems GateCT and GateRT. Both systems use a vision based system based on 3-D surface reconstruction.

As all these products rely on an external surrogate and no intervention is needed, this method can be applied to almost all patients. However, sometimes breathing training is needed to improve the likelihood of the patient completing the simulation session. Usually an external gating system like the Varian Real-time Position Management (RPM) system observes an external fiducial marker. Related to the RPM system, this marker is an infrared reflective plastic box placed on the patient's anterior abdominal surface to maximize the AP respiratory-induced motion. This surrogate is monitored by a video camera mounted on the treatment room wall. The system allows both displacement and phase gating. This filter checks the regularity of the breathing waveform and disables the beam immediately when the breathing waveform becomes irregular. Such an irregular breathing waveform occurs if the patient is coughing or moving. The beam is enabled again after establishing regular breathing again. The VisionRT system provides similar functionality. However, this system does not need any marker.

As stated before, Chapter 3 will present a solution to join the advantages of internal and external gating. It is based on an arbitrary external gating system and abandons the disadvantages of both techniques. The proposed algorithm computes an actual image-based mapping right before the treatment starts. The mapping is based upon a reference breathing cycle and the current breathing of the patient. During the treatment the breathing pattern is observed by the external surrogate which can be compared to the computed mapping. Therefore no surgery is required and the patient is only exposed to a short additional dose which is just used to acquire a few breathing cycles. Furthermore the correlation is computed prior to the treatment and is not only based upon external surrogates, but on patient images where breathing phases and tumor position are known.

2.3 Clinical Procedure

Looking at the following procedure, a brief overview of the clinical use-case of gated radiotherapy is given. The Task Group 76 of the American Association of Physicists in Medicine recommend for patients in whom respiratory motion may be a concern the application of a decision process [AAPM 06].

In topic two the task group recommends that either a motion range in any direction greater than 5 mm or significant normal tissue sparing is a criterion to apply motion management techniques. The Massachusetts General Hospital (MGH) introduced a treatment procedure which depicts the clinical process of radiation therapy [Jian 05]. The procedure describes an image guided respiration gated (IGRG) treatment procedure for gated liver and lung radiotherapy. It suggests seven steps for patient respiration monitoring and gating. The Varian RPM system and additionally an Integrated Radiology Information System (IRIS) for image guidance is suggested. As the usage of only an external surrogate is quite uncertain, the MGH recommends to use 4-D CT scanning, gated radiographic setup and cine EPID (Electronic Portal Imaging Device) as well as patient breath coaching together with the RPM system. Furthermore, the procedure requires implanted fiducial markers for liver tumors and clear anatomic features near the target for lung cancer.

- Patient selection and preparation:

There are several issues to be considered before a patient is selected to receive gated

radiotherapy. The MGH procedure determines suitable patients within three steps. Since a breath coaching technique is used during treatment, it is important that patients are willing and able to follow such instructions. For liver tumor treatment sessions, the patient should already have radio-opaque markers implanted inside or near the tumor and for lung tumors clear anatomic features near the tumor. The last criterion is the range of intra-fraction motion. To gain most advantages in using gated radiotherapy, this range should be quite large.

- Breath training and motion assessment:

Due to breathing control, physicians are able to control the patient's breathing and therefore are furthermore able to reproduce breathing patterns throughout the whole treatment course. Therefore, a breath training session of one hour is scheduled on the simulator, where fluoroscopic images are taken and a initial gating window is determined.

- 4-D CT simulation:

Before the treatment a 4-D CT simulation is scheduled, where both a free breathing and a coached helical 10 phase CT scan is acquired. Using these scans, the physicians specify a gating window, considering the balance between residual motion and duty cycle.

- Treatment planning:

Physicians contour the GTV and/or CTV in each of the ten 4-D CT datasets. These contours are used to define a composite target volume that includes the residual motion. As the end of exhale phase is used as planning CT, the composite target volume is fused to the 4-D CT data set within this phase. Critical structures are contoured at the end of exhale CT data set and a margin is added to the CTV to obtain the PTV. Additionally, a backup plan with a larger margin for non-gated treatment is also created using the free breath CT scan.

- Image guided patient setup:

As in conventional treatment the patient is initially set up using laser alignment to skin tattoos (see Chapter 5). Furthermore, the RPM system is applied to be able to monitor and coach the patient's breathing. A pair of gated AP and lateral IRIS radiographs are taken at the end of exhale phase after the patient is properly coached. These gated radiographs are matched with digital reconstructed radiographs (DRR) to detect patient shifts.

- Gated treatment delivery:

After the patient is positioned, the patient is treated under breath control using the RPM system. Additionally, EPID images are taken in cine mode during the delivery of each field. These images are used for treatment verification within the next step.

- Treatment verification and assessment:

Within the last step, the recorded EPID images are analyzed retrospectively. To verify the gated treatment, the residual marker motion in the gating window is measured. Therefore, the treatment can be modified by reducing the gate, if the residual motion is significantly larger than what was estimated during the simulation [Berb 05].

2.4 Summary

Within this chapter, a brief introduction about the medical background of radiotherapy was presented. Basic principles, like treatment planning and clinical target volumes were introduced. Furthermore, the importance of shrinking the ITV using gated radiotherapy was discussed. Last a recommendation of the American Association of Physicists in Medicine for radiotherapy was presented.

CHAPTER III

Motion Management

This chapter introduces a novel method to verify a correlation between a reference 4-D CT respiratory cycle and live fluoroscopic images. After a short introduction, the algorithm is introduced. Afterwards an evaluation is presented and a conclusion is given.

3.1 Introduction

Recent technological advances in radiotherapy enable delivering a highly conformal radiation dose distribution to a morphologically complex target volume. Main sources of delivery errors include the inter- and intra-fraction target and organ motions. To face this problem, gating can be applied. External gating methods are non-invasive and the external surrogate signal can be generated from various sources. These sources include strain gauges around the patient's thorax, optical markers on the patient skin, 3-D surface acquisitions or airbag systems measuring the air flow. For these reasons, we will focus on external gating.

Having a real-time external signal, however, it is not equivalent to knowing the position of the tumor at all times. Exploiting external signal blindly to turn the radiation beam on and off can cause significant errors. In particular, the correlation between the tumor motion and the surrogate signal may change from the planning phase to the treatment sessions. The major drawback of this approach is the variability of the correlation between the external signal amplitude and the internal target position. The main challenge of the gated treatment is to adapt the gating intervals based on the up-to-date breathing pattern of the patient. This section focuses on an image-based method, which maps the up-to-date images of the patient to the various phases captured during the planning phase. The mapping function is then used to determine an optimal gating interval for the breathing surrogate for a specific treatment session.

3.2 Related Work

Respiration induced tumor motion has been studied indirectly by movements of organs, radio-opaque markers (implanted in the vicinity of the tumor) [Shir 03], and surrogate organ/structure [Keal 01]. Real-time imaging technologies can be used to either observe tumor motion or indirectly perceive the tumor motion [Onim 05]. It has been shown that the tumor motion is within a range of 2–6 cm [Keal 06].

Organ/source	Respiratory signal	N patients (measurem.)	Correlation range	Phase shift	Source
Diaphragm SI fluoroscopy	Abdominal displacement	5(60)	0.82-0.95	Not observed	[Veda 03]
Tumor and diaphragm fluoroscopy	Abdominal displacement	43	0.41-0.94	Short delays observed	[Ahn 04]
Tumor, SI fluoroscopy	Spirometry & abdominal displacement	11(23)	0.36-0.99	-0.65-0.5s	[Hois 04]
Tumor, 3-D biplane radiography	Abdominal displacement	26	Respiratory waveform cycle and AP tumor motion	Principally within 0-0.3s existence of >1.0s	[Tsun 04]
Lung vessels, cine MRI	Abdominal displacement	4	SI 0.87 ± 0.23 AP 0.44 ± 0.27	-	[Koch 04]
Lung tumor respiration correlated CT	Abdominal displacement	9 where tumor SI motion >5 mm	0.74-0.98	<1.0s 4 pts <0.5s 5 pts	[Mage 04]
Lung tumor respiration correlated CT	Diaphragm displacement	12	0.73-0.96	<1.0s 4 pts <0.5s 5 pts	[Mage 04]

Table 3.1: Correlation of tumor/organ motion with the respiratory signal; 3-D: three - dimensional, AP: anterior - posterior, CT: computed tomography, MRI: magnetic resonance imaging, pts: patients, s: second(s), SI: superior - inferior [AAPM 06]

The variation of the motion extent as well as the breathing pattern makes the prediction of the tumor position very difficult. Table 3.1 shows various studies within the AAPM task group report 76 [AAPM 06] where the correlation between a surrogate respiratory signal and organ / tumor movement is shown. This correlation is used to manage moving targets in radiotherapy, e.g. with gated radiotherapy.

3.3 Materials and Methods

To have a better understanding of the tumor motion, time resolved volumetric CT images (4-D CT) are acquired in the planning phase (see Figure 3.1 left). A surrogate signal is usually used to segment the projections prior to reconstruction into a set of phases (usually 8-12) [Riet 06]. After reconstruction, various phases of the 4-D CT are represented by either an amplitude or phase from the surrogate signal. The 4-D CT scans are processed and the gross target volume (GTV) of the tumor is determined for each phase. Furthermore, based on the tumor location, residual motion, and duty cycle, a gating phase interval is selected. This interval is delimited by two phases of the 4-D CT. The union of the GTVs from the phases within the gating interval forms a new volume, the internal target volume (ITV). Based on the prescribed dose and shape of the ITV, a treatment plan is devised. Furthermore, amplitude or phase of the surrogate signal at the two end phases of the gating phase interval marks the planned gating surrogate interval.

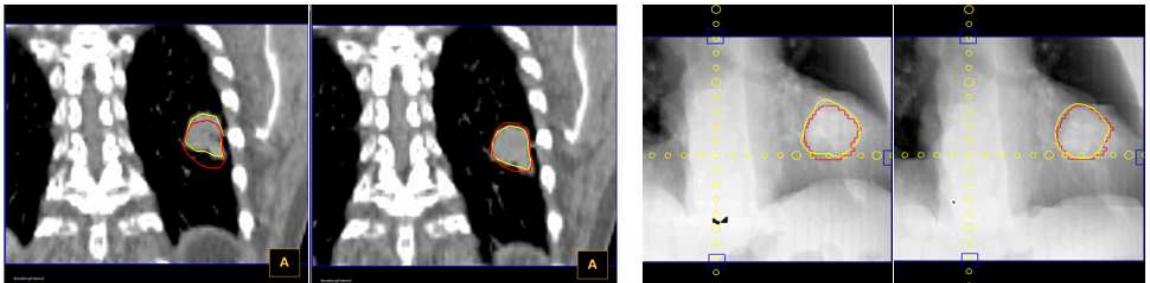


Figure 3.1: Left: Depicts the coronal slices of planning 4-D CT data sets with visible contour lines. Right: Demonstrates two frames of a fluoroscopic acquisition overlaid with the outlines of the projected structures from the planning 4-D dataset [Kham 07].

At the day of treatment, the initial patient setup is performed. In order to reduce the positioning bias due to respiratory motion, we select a single phase out of the 4-D CT that roughly resembles the same breathing phase apparent from the acquired 2-D setup x-ray or portal images. Either an image guided method like [Kham 06] (using the selected image) or the system presented in Chapter 5 can be used to correct for the residual positioning error. After the setup verification process, the patient is at the optimal treatment position. Since the surrogate signal correlation of the target position varies day by day, the gating intervals set during the planning phase might not be valid anymore. It is desirable to check and possibly to adapt the gating intervals using an up-to-date respiratory pattern seen in a pre-treatment fluoroscopic acquisition.

3.3.1 Problem Statement

The image based verification process starts with the acquisition of a pre-treatment image sequence. Mega voltage or kilo voltage fluoroscopic acquisition synchronized with a breathing surrogate is considered (see Fig. 3.1 right). We assume that the pre-treatment image sequence $I_j^{(pt)}$ with $j \in \{0, 1, \dots, K - 1\}$ has K frames taken at intervals of δt covering at least several respiratory cycles (both j and K are non negative integers). Furthermore,

a corresponding synchronized surrogate signal $s_j^{(pt)}$ again with $j \in \{0, 1, \dots, K - 1\}$ is acquired. We also assume that we have a planning 4-D CT acquisition which includes L phases of breathing. We generate a set of DRRs [Russ 03] $I_i^{(p)}$, where $i \in \{0, 1, \dots, L - 1\}$ using each phase of the planning 4-D CT with the exact geometry of the pre-treatment imaging system and the known patient position from the initial patient setup step (both i and L are non negative integers, pt stands for "pre-treatment" and p for "planning").

As mentioned in the previous section, the gating phase interval is delimited by two phases i_{min} and i_{max} both within $\{0, 1, \dots, L - 1\}$. In order to map the gating phase interval which is set at the planning step onto the surrogate signal acquired prior to treatment, we need to establish a mapping function $\mathcal{L} : \{0, 1, \dots, K - 1\} \mapsto \{0, 1, \dots, L\}$:

$$\mathcal{L}_j = \begin{cases} i \in \{0, 1, \dots, L - 1\} & \text{if } i \text{ is a correspondence for the frame } j \\ L & \text{otherwise (no correspondence)} \end{cases} \quad (3.1)$$

The mapping function needs to be determined based on the image information from the two time series $I_i^{(p)}$, where $i \in \{0, 1, \dots, L - 1\}$ and $I_j^{(pt)}$, where $j \in \{0, 1, \dots, K - 1\}$. Once the mapping function is estimated, the treatment day gating surrogate interval can be estimated as follows:

$$s_{min}^{(pt)} = \min_{j|\mathcal{L}_j \in \{i_{min}, i_{max}\}} s_j^{(pt)} \quad (3.2)$$

For computing $s_{max}^{(pt)}$, we switch the min to max in equation (3.2). The treatment day gating surrogate interval $\{s_{min}^{(pt)}, s_{max}^{(pt)}\}$ can also be used to estimate a new duty cycle. The surrogate signal amplitudes of $s_j^{(pt)}$ for j satisfying the condition $\mathcal{L}_j \in \{i_{min}, i_{max}\}$ can also be further analyzed to detect outliers, mainly to increase the duty cycle as much as possible.

Figure 3.2 shows a schematic overview of an empty similarity matrix with the DRRs and the fluoroscopic images arranged.

3.3.2 Image Similarity Matrix

The first step in establishing an optimal mapping function is to generate a similarity image matrix between the two image sequences of $I_i^{(p)}$ where $i \in \{0, 1, \dots, L - 1\}$ and $I_j^{(pt)}$ where $j \in \{0, 1, \dots, K - 1\}$. We generate a matrix \mathbf{M} of size $(L + 1) \times K$, where $\mathbf{M}_{i,j}$ is the normalized mutual information (NMI) value between the two images of $I_i^{(p)}$ and $I_j^{(pt)}$. The normalized mutual information is defined as follows, where $p_A(a)$ is the probability that A has value a and H is the Shannon entropy:

$$H(A) = - \sum_a p_A(a) \log p_A(a) \quad (3.3)$$

$$H(A, B) = - \sum_{a,b} p_{AB}(a, b) \log p_{AB}(a, b) \quad (3.4)$$

$$\text{NMI}(A, B) = \frac{H(A) + H(B)}{H(A, B)} \quad (3.5)$$

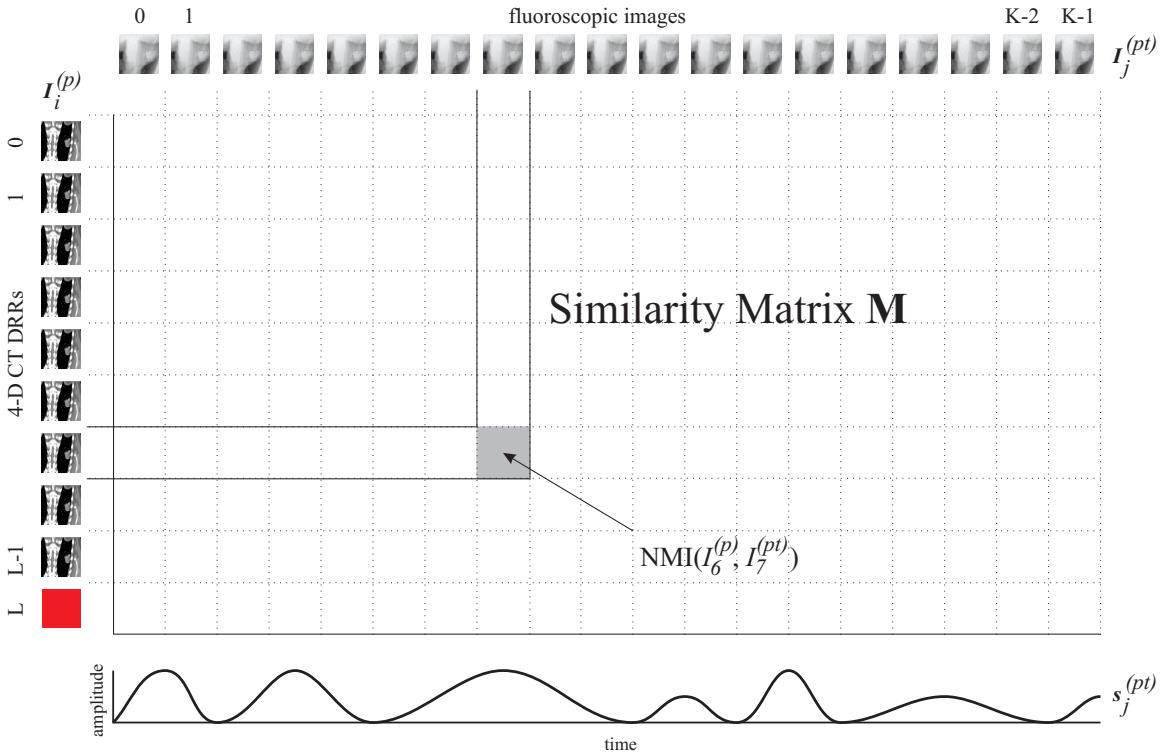


Figure 3.2: Schematic overview of the similarity matrix based on 4-D CT DRRs $I_i^{(p)}$ and a sequence of fluoroscopic images $I_j^{(pt)}$ with a corresponding respiratory signal $s_j^{(pt)}$. The last row (red box) indicates an additional row for "no-match" situation.

An additional row ($L + 1$) is added for "no-match" situations (see Figure 3.2). As the quality of the similarity measure might decrease despite an increasing mutual information (MI) value, NMI is used [Stud 99]. \mathbf{M} is furthermore normalized between 0 and 1.

Each column of the matrix \mathbf{M} can be thought of as matching likelihood profile between a frame of the fluoroscopic sequence and various phases of the planning DRRs. A clear peak in that profile could indicate a potential match. However, in case there is no match, one would expect to have consistently low value across the whole column. In order to deal with these cases properly, we have considered the last row, which is in fact representing the likelihood of having a "no-match" situation. For each element in the last row of the matrix \mathbf{M} , we consider a constant value that is the median of the maximums for various columns minus the range of the similarity values of the corresponding column. In practice this computation showed good results.

Finding maximum values for each column of the matrix \mathbf{M} could be a potential solution for the mapping function \mathcal{L} . However, the problem of this approach is that no temporal consistency is considered, since all columns are processed independently. For example, images from certain phases of inhale and exhale could look quite similar if they are observed independently and not considered in the context of a sequence. Not considering temporal aspects of the process also makes it difficult to deal with cases, without match. Furthermore, robustness to noise and/or to variations of image intensities from the two sequences could be easily compromised.

3.3.3 Respiratory Motion as a Markov Process

In order to enforce temporal consistency in finding the mapping function \mathcal{L} from the image similarity matrix \mathbf{M} , we first need to find a reasonable model for the breathing process [Wu 04]. We assume that the breathing can be modeled as a Markov process.

We consider S_u , where $u \in \{0, 1, \dots, L\}$ to be the states for a Markov process. Since the pre-treatment acquisition is not triggered, it can start at any breathing phase. Therefore the initial probabilities π_u of the various states are constant (i.e. $\pi := \pi_u = \frac{1}{L+1}$). We consider that the observations are identical to the process states. The state transition probability matrix \mathbf{A} , which is of size $(L + 1) \times (L + 1)$ can be estimated from a set of observation sequences using the expectation maximization algorithm or defined by a function with the following properties:

- \mathbf{A} is a circulant matrix [Gray 06]
- The first row of \mathbf{A} is defined by a function, having a negative quadratic form, e.g.:

$$\mathbf{A}_{0,x} = -\frac{L}{2}x^2 + \frac{1}{2L}, \text{ where } x = 0, 1, \dots, L - 1 \quad (3.6)$$

- The following rows are shifted by one:

$$\mathbf{A}_{1,x} = -\frac{L}{2}x^2 + \frac{1}{2L}, \text{ where } x = L - 1, 0, 1, \dots, L - 2 \quad (3.7)$$

$$\mathbf{A}_{2,x} = -\frac{L}{2}x^2 + \frac{1}{2L}, \text{ where } x = L - 2, L - 1, 0, 1, \dots, L - 3 \quad (3.8)$$

$$\dots = \dots \quad (3.9)$$

$$\mathbf{A}_{L-1,x} = -\frac{L}{2}x^2 + \frac{1}{2L}, \text{ where } x = 1, 2, \dots, L - 1, 0 \quad (3.10)$$

- Each row is normalized that the sum of all values is one.
- The last row and column of \mathbf{A} (the additional state) holds a constant value for the probability of the transition to the additional state and is not part of the circulation.

The probability of a given sequence having been produced by a specific Markov process can be computed using the initial state probability π and the transition matrix \mathbf{A} using a forward-backward propagation algorithm [Rabi 89]. An example of an heuristically determined transition matrix for the breathing cycle is depicted in Figure 3.3.

3.3.4 Mapping as a Solution of an Optimization Problem

The mapping function \mathcal{L} is in fact a constellation of states S_u for $u \in \{0, 1, \dots, L\}$ from the planning phase. The likelihood of a certain mapping due to the imaging cues can be computed using the image similarity matrix \mathbf{M} . Furthermore, the probability of a certain mapping can be computed using π and \mathbf{A} . Therefore, we setup an optimization to maximize the posteriori probability that is expanded to the product of the likelihood and prior probability using Bayesian rule. The optimization formulation can then be written as:

$$\begin{aligned} \tilde{\mathcal{L}} &= \arg \max_{\mathcal{L}} \left(\prod_{j=0}^K M_{\mathcal{L}_j, j} \right) \cdot \left(\pi \prod_{j=1}^{K-1} \mathbf{A}_{\mathcal{L}_j, \mathcal{L}_{j+1}} \right) \\ &\approx \text{likelihood} \qquad \qquad \qquad \text{prior} \end{aligned} \quad (3.11)$$

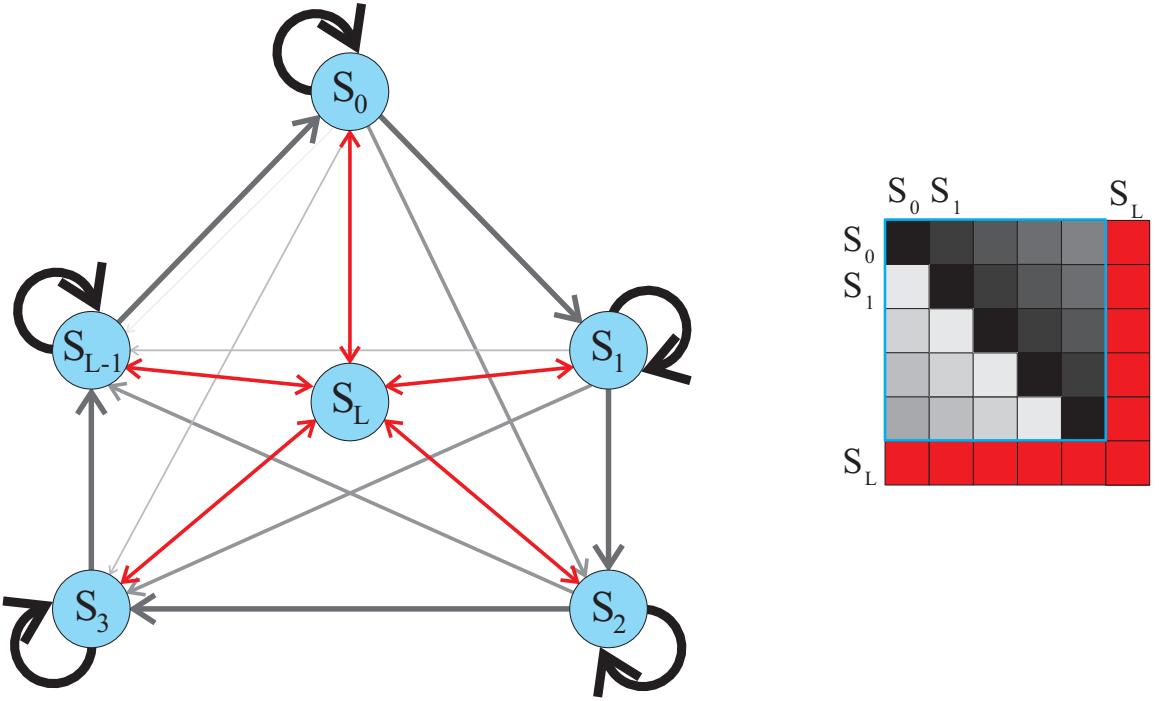


Figure 3.3: Left: Depicts the state transition network for $L = 5$. Thicker lines denote higher transition probability, whereas red lines denote the transition to an undefined state. Right: Example for a state transition matrix \mathbf{A} . Darker entries show higher probabilities to stay in this state.

where $\tilde{\mathcal{L}}$ is the solution. Equation (3.11) depicts a combinatorial optimization problem, since the number of possibilities for state constellations are discrete and finite. The evaluation of the cost function in this problem is computationally inexpensive. Since the problem can be looked at as overlapping sub-problems, we use the Viterbi algorithm [Vite 03] to find the optimal solution. Similar to Dijkstra shortest path algorithm, we can define and store the maximum cost of getting at the time index j in a $(L + 1) \times K$ matrix \mathbf{Q} :

$$\mathbf{Q}_{i,j} = \begin{cases} \infty & j < 0 \\ \mathbf{M}_{i,j} \cdot \pi_i & j = 0 \\ \mathbf{M}_{i,j} + \max_n (\mathbf{Q}_{i,j-1} \cdot \mathbf{A}_{i,n}) & \text{otherwise, with } n \in \{0, 1, \dots, L\} \end{cases} \quad (3.12)$$

Once the matrix \mathbf{Q} in the equation (3.12) is computed recursively, the solution of the optimization problem in the equation (3.11) is:

$$\tilde{\mathcal{L}}_j = \arg \max_i \mathbf{Q}_{i,j} \quad \text{for } j \in \{0, 1, \dots, K\} \quad (3.13)$$

Estimated mapping function $\tilde{\mathcal{L}}$ can be used in equation (3.2) to compute the updated gating surrogate interval. Algorithm 1 shows the algorithm in pseudo-code. The important part of the algorithm has a complexity of $\mathcal{O}(K \times L^2)$.

A similar method, not implying periodicity can be seen in [Rohk 08].

Algorithm 1 Verification

```

1: Input:
     $\mathbf{I}_j^{(pt)}, j \in \{0, 1, \dots, K - 1\}$  {fluoroscopic images}
     $\mathbf{s}_j^{(pt)}, j \in \{0, 1, \dots, K - 1\}$  {respiratory signal}
     $\mathbf{I}_i^{(p)}, i \in \{0, 1, \dots, L - 1\}$  {DRRs}
     $\mathbf{A}^{(L+1) \times (L+1)}$  {precomputed transition matrix}
2: for  $j = 0..K - 1$  do
3:   for  $i = 0..L - 1$  do
4:      $\mathbf{M}_{i,j} = \text{NMI}(\mathbf{I}_i^{(p)}, \mathbf{I}_j^{(pt)})$ 
5:   end for
6: end for
7: for  $j = 0..K - 1$  do
8:    $\mathbf{M}_{L,j} = \text{median of all column maxima minus the range of } j\text{th column values}$ 
9: end for
10: for  $i = 0..L$  do
11:    $\mathbf{Q}_{i,0} = \mathbf{M}_{i,0} \cdot \pi_i$  {Initialize first column}
12: end for
13: for  $j = 1..K - 1$  do
14:   for  $i = 0..L$  do
15:      $\mathbf{Q}_{i,j} = \mathbf{M}_{i,j} + \max_n(\mathbf{Q}_{i,j-1} \cdot \mathbf{A}_{i,n})$ 
16:   end for
17: end for
18: for  $j = 0..K - 1$  do
19:    $\mathcal{L}_j = \arg \max_i \mathbf{Q}_{i,j}$ 
20: end for
21: Output: Mapping function  $\mathcal{L}$ .

```

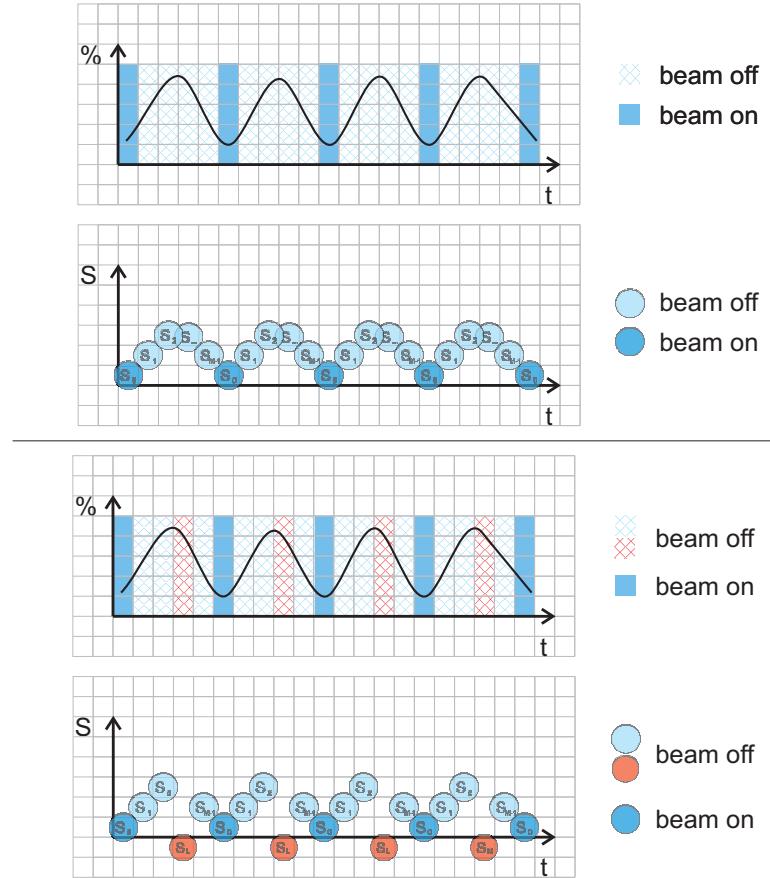


Figure 3.4: Left: Breathing cycle, represented by different states, where all states are known. Right: An unknown state is included, which causes the LINAC to stop irradiating.

3.4 Experiments and Results

To evaluate the performance of the proposed method, we define several error measures. The first error measurement e_1 is the percentage of wrongly labeled fluoroscopic frames. The second error measure e_2 is the percentage of wrongly labeled "no match". Although the wrongly labeled "no match" does not have any effect on the computed gating surrogate interval, it has an adverse impact on the duty cycle. For the first series of tests, we used a breathing phantom that has a moving target within a hallow cavity. The speed and the range of the target motion within the phantom can be adjusted. A ten phase 4-D CT of the breathing phantom was acquired using Siemens Sensation 64 CT Scanner (Siemens Health Care, Forchheim, Germany). Simulator fluoroscopic sequences of the phantom were taken from the phantom with three different motion ranges and speeds using Siemens Mevasim S Simulator (Siemens Health Care, Concord CA, USA). For the second and third series, the 4-D CT of patients with synthetic and real fluoroscopic sequences were used. We acquired two 4-D CT scans of the same patient at two different points in time. Based on one 4-D CT we generated DRRs from the 4-D CT phases and puzzled them together to form synthetic fluoroscopic sequences. In total three sequences were generated. For the first one, we concatenated the original 4-D CT phases multiple times. For the second and third ones, we varied the sequence order and generated two distinct synthetic breathing patterns.

3 Motion Management

Finally, we acquired two 4-D CT scans of the two different patients with corresponding simulator acquisitions. In all cases, correct correspondences were selected manually. We used split-screen and a blending display method to identify the correct correspondences for each fluoroscopic frame out of various 4-D CT phases. Sample images and the similarity matrix with the overlaid mapping function are depicted in Figure 3.5. The results of the two error measures are shown in Table 3.2. The processing time depends on the number of fluoroscopic sequence frames and it is roughly one second per frame. All the tests were performed on an Intel Centrino Duo CPU with 2.0 GHz and 2 GB of RAM and a NVIDIA Quadro FX 2500 display adapter.

Measure	$e_1(\%)$	$e_2(\%)$
Phantom Original (i.e., same as 4-D CT)	0	0
Phantom Slow Breathing	8	0
Phantom Fast Breathing	5	0
Phantom Deep Breathing	8	15
Phantom Shallow Breathing	8	15
Patient 1 with Syn. Fluoro	4	10
Patient 1 with Syn. Fluoro Modified Pattern 1	7	10
Patient 1 with Syn. Fluoro Modified Pattern 2	8	10
Patient 2 with Fluoro	15	15
Patient 3 with Fluoro	14	10
Average	8	8.5

Table 3.2: Results of the two error measures for various datasets.

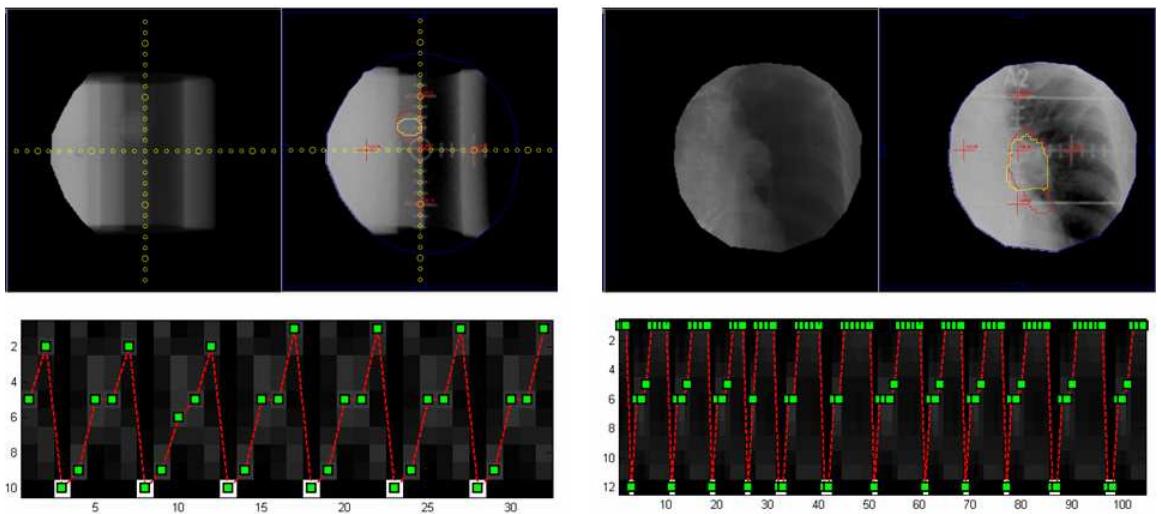


Figure 3.5: In the top row of the figures (left) and (right) a 4-D CT frame with matching fluoroscopic frame with overlaid contours for phantom and patient datasets is depicted, respectively. The bottom row shows the corresponding image similarity matrix for each case [Kham 07].

Within a second evaluation series, we used the NCAT phantom [Sega 03] to generate multiple series of a breathing thorax including a tumor. Overall nine sequences were gener-

ated: normal breathing at normal speed (NN), normal breathing at slow speed (NS), normal breathing at fast speed (NF), shallow breathing at normal speed (SN), shallow breathing at slow speed (SS), shallow breathing at fast speed (SF), deep breathing at normal speed (DN), deep breathing at slow speed (DS) and deep breathing at fast speed (DF).

Each sequence consists of a 4-D CT and corresponding DRRs representing fluoroscopic images. Using these nine sequences random sequences concatenating different sequences were created. These sequences were evaluated by looking at four parameters:

- Beam to healthy tissue volume: The percentage of irradiated healthy tissue (see equation (3.17)).
- Beam to tumor volume: The percentage of the beam hitting the tumor (see equation (3.18)).
- Irradiated tumor volume: The percentage of the tumor volume which is irradiated (see equation (3.19)).
- Missed tumor volume: The percentage of the tumor's volume not to be irradiated (see equation (3.20)).

These measurements are defined as follows. Let T be the voxel set of points corresponding to the tumor and B the voxel set of points irradiated by the beam. We than can define three sets:

$$lrS_0 := T \setminus B \quad \text{non-irradiated (missed) tumor volume} \quad (3.14)$$

$$S_1 := T \cap B \quad \text{irradiated tumor volume} \quad (3.15)$$

$$S_2 := B \setminus T \quad \text{healthy tissue} \quad (3.16)$$

These voxel sets can now be used to compute the four measurements:

$$\text{beamToHealthyTissue} = \frac{|S_1 \cup S_2|}{|S_2|} \quad (3.17)$$

$$\text{beamToTumor} = \frac{|S_1 \cup S_2|}{|S_0 \cup S_1|} \quad (3.18)$$

$$\text{irradiatedTumorVolume} = |S_1| \quad (3.19)$$

$$\text{missedTumorVolume} = |S_0| \quad (3.20)$$

Figure 3.6 shows an example sequence. The sequence consists of six different patterns, starting with normal speed and amplitude twice it is followed by a deep and slow sequence and a normal one again and ends with two shallow and fast sequences. The algorithm correctly recognizes that there are no corresponding 4-D CT volumes available for the deep sequence. However, it does some misclassification for the shallow sequences and labels the 4-D CT volumes slightly wrong (see around frame 120 and 150). Table 3.3 shows the computed mean, median, and variance for the sequence.

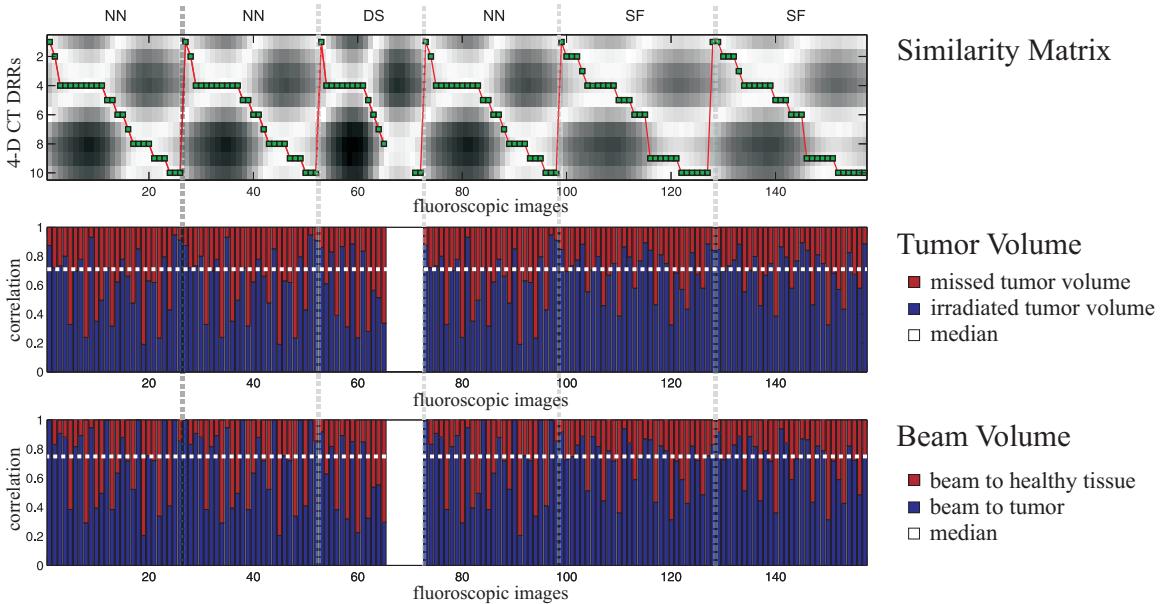


Figure 3.6: Similarity matrix for a specific sequence (NN-NN-DS-NN-SF-SF). Between frame 65 and 73 there is no information available as there are no corresponding 4-D CT volumes. The green dots in the upper image indicate corresponding frames, where the red line shows the computed path.

Measure	Mean [%]	Median [%]	Variance [%]
Irradiated tumor volume	0.6448	0.7000	0.0440
Missed tumor volume	0.3552	0.3000	0.0440
Beam to tumor volume	0.6926	0.7534	0.0533
Beam to healthy tissue volume	0.3074	0.2466	0.0533

Table 3.3: Results for the sample sequence: NN-NN-DS-NN-SF-SF

3.5 Conclusion

We have developed an image based mapping / synchronization procedure that automatically labels pre-treatment fluoroscopic image frames with the corresponding phase from 4-D CT. The mapping procedure is formulated as an optimization process. It finds an optimal mapping maximizing the image similarity between the corresponding pairs while preserving a temporal coherency to an established Markov model for breathing. The mapping procedure also detects the frames with no corresponding phase from the planning 4-D CT. The fluoroscopic image based verification addresses the problem of establishing the correlation between surrogate signal and internal target prior to treatment. By virtue of having the labels generated from a mapping process, we can adaptively change the surrogate gating interval to the up-to-date breathing pattern of the patient. Results on phantom, synthetic, and patient data show in average 93 percent of correctly labeled frames.

Based on synthetic data we could furthermore show, that at different speed and amplitude with respect to a reference 4-D CT scan, an average irradiated tumor volume of 64% and only a harm of 30% of healthy tissue within the beam can be realized.

CHAPTER IV

Time-of-Flight Cameras

This chapter introduces Time-of-Flight cameras. The basic working principle of the camera is presented. To improve the data acquired by the ToF camera both an optical- and a distance calibration is introduced. Furthermore, the key benefits of ToF cameras are discussed.

4.1 Introduction

Time-of-Flight cameras provide a direct way to acquire 3-D surface information of objects in real-time [Xu 98]. Conventional imaging sensors consist of multiple photo diodes. In cameras these diodes are arranged within a matrix and provide an image of, e.g. color or gray values. In opposite to normal cameras, a Photon Mixing Device (PMD) sensor simultaneously acquires a distance value for each pixel in addition to the common intensity (gray) value (see Figure 4.1, 4.2). The theoretical principle behind the technology relies on the ToF principle. This means that a PMD camera has at least one illumination source to actively illuminate the scene. Typical ToF cameras use intensity modulated infrared light, which is not visible to human eyes to illuminate the scene. Other light sources than infrared light are possible, too. The PMD sensor captures the reflected light and evaluates the distance information on the pixel. This is done by correlating the emitted signal with the received signal. Concluding, a PMD sensor is composed by a matrix of distance sensors. Despite these functional improvements (compared to conventional imaging sensors) the sensor itself is still a standard CMOS sensor. Therefore imaging and 3-D measurement capabilities can be placed next to system-relevant electronics like analog-digital converters, etc. All "intelligence" of the sensor is included on the chip, meaning, meaning that the distance is computed per pixel. Therefore PMD pixels are also called "smart pixels". Some ToF cameras are additionally equipped with a special pixel-integrated circuit. This circuit guarantees the independence of sunlight influence by the Suppression of Background Illumination (SBI).



Figure 4.1: Examples of ToF camera models. Left side: CamCube from PMD Technologies GmbH, Germany (<http://www.pmdtec.com>). Right side: SR4000 from MESA Imaging AG, Switzerland (<http://www.swissranger.ch>)

There are various PMD camera models covering a broad field of applications. The most recent camera developed by PMD Technologies is called CamCube 3.0. It is an evaluation platform with a modular concept and is optimized for flexibility. The sensor matrix has one of the highest ToF camera resolution with 200×200 pixels to date. It uses USB 2.0 to transfer the data to the computer. Framerates of up to 40 frames per second are possible. Depending on the modulation frequency of the emitted light, the camera has a non-ambiguous range. For example at a modulation frequency of 40 MHz this range is between 0 and 3.75 meters, as the "length" of the emitted wave is 7.5 m and the light has to travel from the camera to the object and from the object back to the camera. The theoretical computation of the non-ambiguous range is presented in the next section.

Attached to the camera there is a removable CS-mount lense and two separate light sources. This modular design enables the camera to be adjusted to specific demands. One can change the aperture or improve the light power and modulation frequency of the light source to cover higher distances. There is also a great potential to improve the accuracy of the camera by tuning the light source. In general, the light source is an important factor for the data quality. Factors, like the homogeneity of the illumination or the quality of the modulation frequency have a direct impact on the data. It is also possible to detach the light source completely from the camera and to put it to another position. An example for such a scenario is the automotive industry. The illumination unit is next to the car light, whereas the camera is behind the windshield of the car.

Current and future fields of application of a ToF camera are for example:

- **Security:** ToF cameras can be used to observe rooms / places. Compared to simple motion sensors a ToF camera can easily distinguish between various objects crossing the scene.
- **Robot control:** A real-time 3-D scenery is very important for robots (e.g. for collision detection, etc) [Prus 08].
- **Virtual- and augmented-reality applications:** Having 3-D information of an object, it is much easier to match it with virtual 3-D objects [Koch 09].

- **"Grab in the box":** This is a classical problem. Having various objects in a box (e.g. at an assembly line), a robot has to grab the right tool. 3-D information can be used to locate and classify the objects.
- **Human-machine interaction:** 3-D gesture interaction, with hands or the whole body, are a very recent topic in consumer electronics today. ToF cameras provide important 3-D information to interact with computers or machines [Koll 08, Sout 08].
- **Automotive applications:** ToF sensors provide new possibilities for applications in the automotive industry, e.g. pre-crash detection or collision warning [Hsu 06]
- **Medical applications:** Recently, various applications for medical purposes have arisen. Among respiratory motion and patient positioning, there are also applications within 3-D endoscopy or orthopedics [Scha 08, Scha 09a, Penn 07].

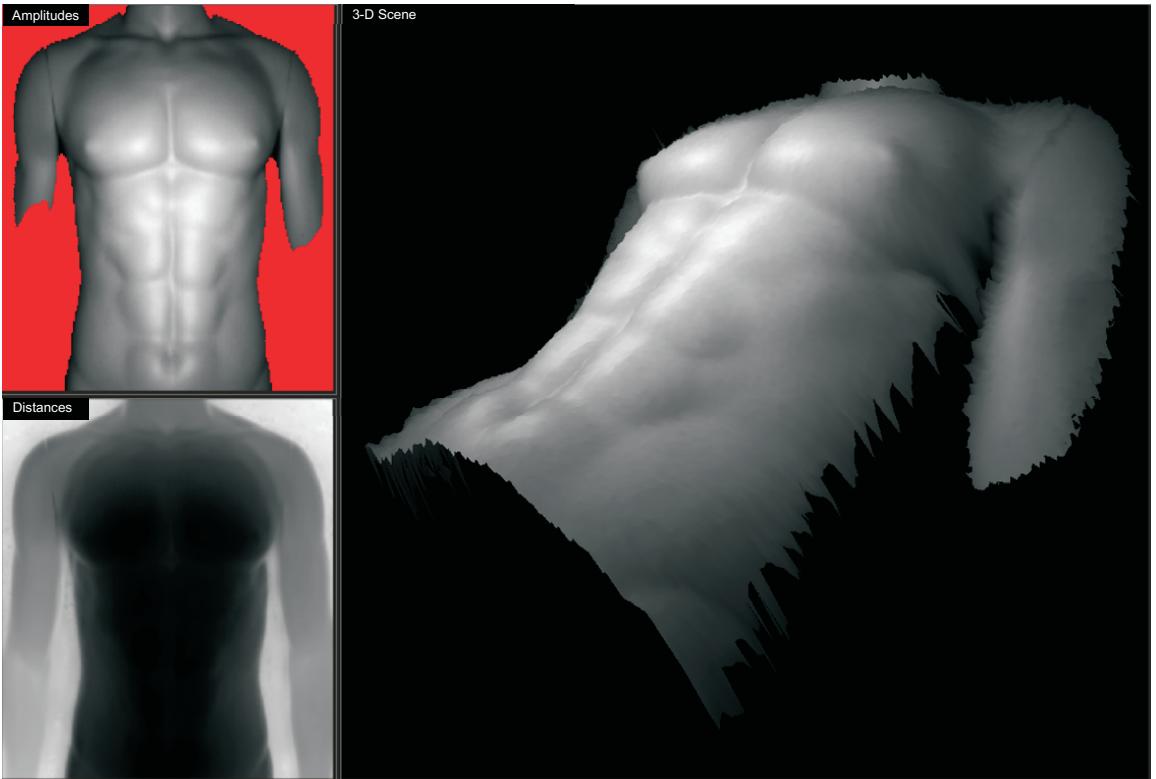


Figure 4.2: ToF camera data. Upper left: amplitude image (red pixels indicate pixels not considered in the 3-D reconstruction). Lower left: distance image. Right: 3-D reconstruction of data.

More recently, applications like gesture recognition [Koll 08, Sout 08] or automotive passenger classification [Deva 07] are utilizing ToF cameras. ToF cameras have several advantages over other 3-D surface acquisition techniques like stereo vision or structured light techniques. The most promising advantage of ToF is that it is on its way to become a component of mass markets like consumer electronics and the automotive industry. Currently a high-end ToF camera is available for about USD 7,000 (see Figure 4.1). However, in the near future, a target price of a few hundred dollars for ToF cameras can be expected. ToF cameras also render calibration steps which are mandatory for stereo based systems, for most applications unnecessary. This benefit is based on the monocular all-solid-state architecture of ToF cameras. However, in this thesis, we recalibrate the whole ToF camera.

This is necessary, as the intended purpose and therewith the accuracy and precision of current ToF cameras is not within the scope of patient positioning. To achieve best results, we override the manufacturer calibration. The compact architecture also enables a high portability of the system and a variety of integration prospects in existing systems. Knowing the lens properties, it is furthermore possible to compute precise metric 3-D information for each pixel.

Currently there are several supplier for Time-of-Flight and depth cameras:

- **3DV (Israel)**: Former supplier for consumer electronics ToF cameras, acquired by Microsoft in the beginning of 2009. The technology is currently not available any more.
- **Canesta (USA)**: Canesta (USA): Own ToF sensor, acquired by Microsoft in the end of 2010. Still acts as chip supplier for various companies.
- **Fotonic (Sweden)**: Builds own cameras for industrial applications using sensors of Canesta.
- **MESA Imaging (Swiss)**: Own ToF sensor and camera, for industrial applications.
- **Optex (Japan)**: Builds cameras based on Canesta sensors for consumer electronic applications.
- **Optrima (Belgium)**: Own ToF sensor and camera for consumer electronics and industrial applications.
- **Panasonic (Japan)**: Consumer electronics camera based on second MESA Imaging chip generation.
- **PMD Technologies (Germany)**: Own ToF sensor and camera prototypes with modular concept for consumer electronics, industrial and automotive applications
- **Prime Sense (Israel)**: sensor supplier for Microsoft XBOX 360 Kinect (not a ToF Sensor)

4.2 Working Principle

ToF cameras provide a real-time $2\frac{1}{2}$ -D¹ representation of an object. The object, e.g. a patient is actively illuminated with an incoherent light signal. This signal is intensity modulated by a cosine-shaped signal of frequency f_{mod} . Usually the emitted light is part of the non-visible area of the spectrum in the near infrared spectral range (e.g. 780 nm).

Non-ambiguous Range

Due to the measurement principle, ToF cameras have a non-ambiguous range of d_{nb} . It simply defines a range, where distances can be computed uniquely. The range is dependent on the modulation frequency of the camera, as this frequency defines the wave length of the emitted signal. To compute distances, the camera evaluates the phase shift between a reference (emitted) signal and the received signal. Figure 4.3 shows the relation of a

¹Only a part of the surface can be seen by the ToF camera.

distance d to a phase shift ϕ_d . Furthermore, this relation can also be seen in equation 4.4, where

- c [m/s] denotes the speed of light,
- d_n [m] the distance the light travels,
- t [s] the time the light needs to travel the distance d_n ,
- f_{mod} [MHz] the modulation frequency,
- d_{nb} [m] the distance from a light source to an object (non-ambiguous range),
- ϕ_d [rad] the phase shift.

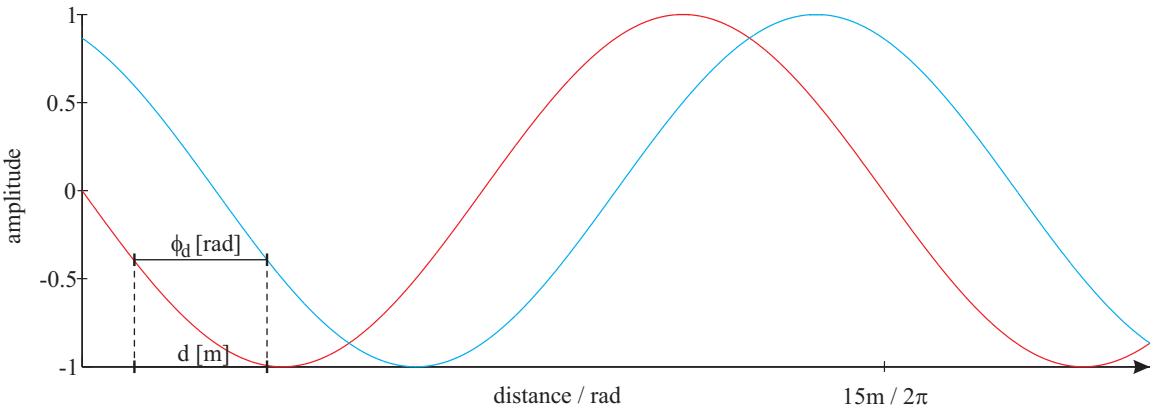


Figure 4.3: Relation of a phase shift to the distance at a fixed modulation frequency of $f_{mod} = 20$ MHz. A single wave is of length 2π or 15 m. Given a phase shift ϕ_d , the phase shift is proportional to the distance d .

$$c = \frac{d_n}{t}, \text{ where } t = \frac{1}{f_{mod}} \quad (4.1)$$

$$c = d_n \cdot f_{mod} \rightarrow d_n = \frac{c}{f_{mod}} \quad (4.2)$$

$$d_{nb} = \frac{c}{2f_{mod}} \quad (4.3)$$

$$d = \frac{c}{2f_{mod}} \cdot \frac{\phi_d}{2\pi} \quad (4.4)$$

Determination of the Phase Shift

Traveling with constant speed of light c , the light signal is reflected by the surface of the object and gets back to the camera after a time of flight τ_d .

In addition to depth values (**D**), ToF cameras provide intensity values, representing the amount of light sent back from a specific point. Therefore, this image is also called amplitude image (**A**). Some ToF cameras also provide a gray value image (**I**), representing the overall amount of light captured by the sensor.

By estimating the phase shift ϕ_d between both, the emitted $g(t)$ (with phasing ϕ_T) and reflected/received light signal $s_d(t) \sim g(t - \tau_d)$ (with phasing ϕ_R), the distance d can

be computed (see equations 4.6). Assuming constant speed of light c , the distance d is proportional to the phase shift ϕ_d of the emitted and reflected signal (see equation 4.8). Figure 4.4 illustrates the ToF principle for various distances.

$$\bar{\omega} = 2\pi f_{\text{mod}} \quad (4.5)$$

$$\phi_d = \phi_R - \phi_T = \bar{\omega}\tau_d \quad (4.6)$$

$$d = \frac{c \cdot \phi_d}{4\pi \cdot f_{\text{mod}}} \quad (4.7)$$

$$= \frac{c}{2f_{\text{mod}}} \cdot \frac{\phi_d}{2\pi} \quad (4.8)$$

The phase shift ϕ_d is measured by sampling the received signal $s_d(t)$ at N equidistant measurement points. This can be realized by mixing / correlating the signal in a so-called charging swing ("Ladungsschaukel") with an electrical reference signal. This reference signal is initially in phase with the sent signal $g_k(t)$ and is stepwise increased by the phase shift $\bar{\omega}\tau_k$ (see equation 4.9). According to [Lang 00] the phase shift is computed as follows:

$$s_d(t) = 1 + a \cdot \cos(\bar{\omega}t - \phi_d) \quad (4.9)$$

$$g_k(t) = g(t + \tau_k) \quad (4.10)$$

$$= \cos(\bar{\omega}t + \bar{\omega}\tau_k), \text{ where} \quad (4.11)$$

$$\bar{\omega}\tau_k = \frac{2\pi}{N} \cdot k \text{ with } k = 0, 1, \dots, N-1 \quad (4.12)$$

Typically $N = 4$ is applied which leads to an iterative phase shift for the reference signal $g_k(t)$ by 90° with $\bar{\omega}\tau_0 = 0^\circ$ and $\bar{\omega}\tau_3 = 270^\circ$ while mixing the reference signal with the received and transformed signal $s_d(t)$.

By the correlation function $c(\tau_k) \sim s_d(t) \otimes g_k(t)$ the correlation can be computed as stated in equation 4.13 - 4.20 (where a : signal amplitude, b : background illumination influence).

$$c(\tau_k) \sim s_d(t) \otimes g_k(t) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} s_d(t) \cdot g_k(t) dt \quad (4.13)$$

$$= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} (1 + a \cdot \cos(\bar{\omega}t - \phi_d)) \cdot \cos(\bar{\omega}t + \bar{\omega}\tau_k) dt \quad (4.14)$$

$$= \frac{a}{2} \cdot \cos(\phi_d + \bar{\omega}\tau_k) \quad (4.15)$$

$$= \frac{a}{2} \cdot \cos(\phi_d + \bar{\omega}\tau_k) \quad (4.16)$$

$$c_0 = c(\tau_0) = b + \frac{a}{2} \cdot \cos(\phi_d) \quad (4.17)$$

$$c_1 = c(\tau_1) = b - \frac{a}{2} \cdot \sin(\phi_d) \quad (4.18)$$

$$c_2 = c(\tau_2) = b - \frac{a}{2} \cdot \cos(\phi_d) \quad (4.19)$$

$$c_3 = c(\tau_3) = b + \frac{a}{2} \cdot \sin(\phi_d) \quad (4.20)$$

Accordingly, a pair of phase shift dependent voltage differences Δc_{31} and Δc_{02} can be computed (see equation 4.21 - 4.22). These differences depend on ϕ_d (see equation 4.23)

$$\Delta c_{31} = c_3 - c_1 = a \cdot \sin(\phi_d) \quad (4.21)$$

$$\Delta c_{02} = c_0 - c_2 = a \cdot \cos(\phi_d) \quad (4.22)$$

$$\frac{\Delta c_{31}}{\Delta c_{02}} = \tan(\phi_d) \quad (4.23)$$

Because no knowledge is given about the unit circle quadrant, the common arctan function is not sufficient to compute ϕ_d non-uniquely. For the correct mapping of the values between 0 and 2π an offset of 2π has to be added to negative ϕ_d values in equation 4.24.

$$\phi_d = \text{atan2}(\Delta c_{31}, \Delta c_{02}) \quad (4.24)$$

The amplitude a depends on Δc_{31} and Δc_{02} as well (see equation 4.25) and the offset b can be computed with the voltages $c_0 \dots c_3$ (see equation 4.26).

$$a = \frac{\sqrt{\Delta C_{31}^2 + \Delta C_{02}^2}}{\sqrt{2}} = \frac{\sqrt{\sin(a \cdot \phi_d)^2 + \cos(a \cdot \phi_d)^2}}{\sqrt{2}} \quad (4.25)$$

$$b = \frac{c_0 + c_1 + c_2 + c_3}{4} \quad (4.26)$$

The distances can be computed according to equation 4.8. Based on the periodicity of the cosine-shaped modulation signal, only smaller distances than $\frac{c}{2f_{\text{mod}}}$ can be computed. Today available ToF cameras operate at a modulation frequency of for example about 20 MHz. Thus, the upper limit for observable distances is approx. 7.5 m which is sufficient for all the proposed systems within this thesis. Detailed information about the working principle of ToF cameras can be found in Xu et al. [Xu 98, Lang 00].

4.3 Key Benefits of Time-of-Flight Cameras

To emphasize technical differences between laser range, structured light and stereo camera based techniques versus ToF sensors these issues will be briefly discussed:

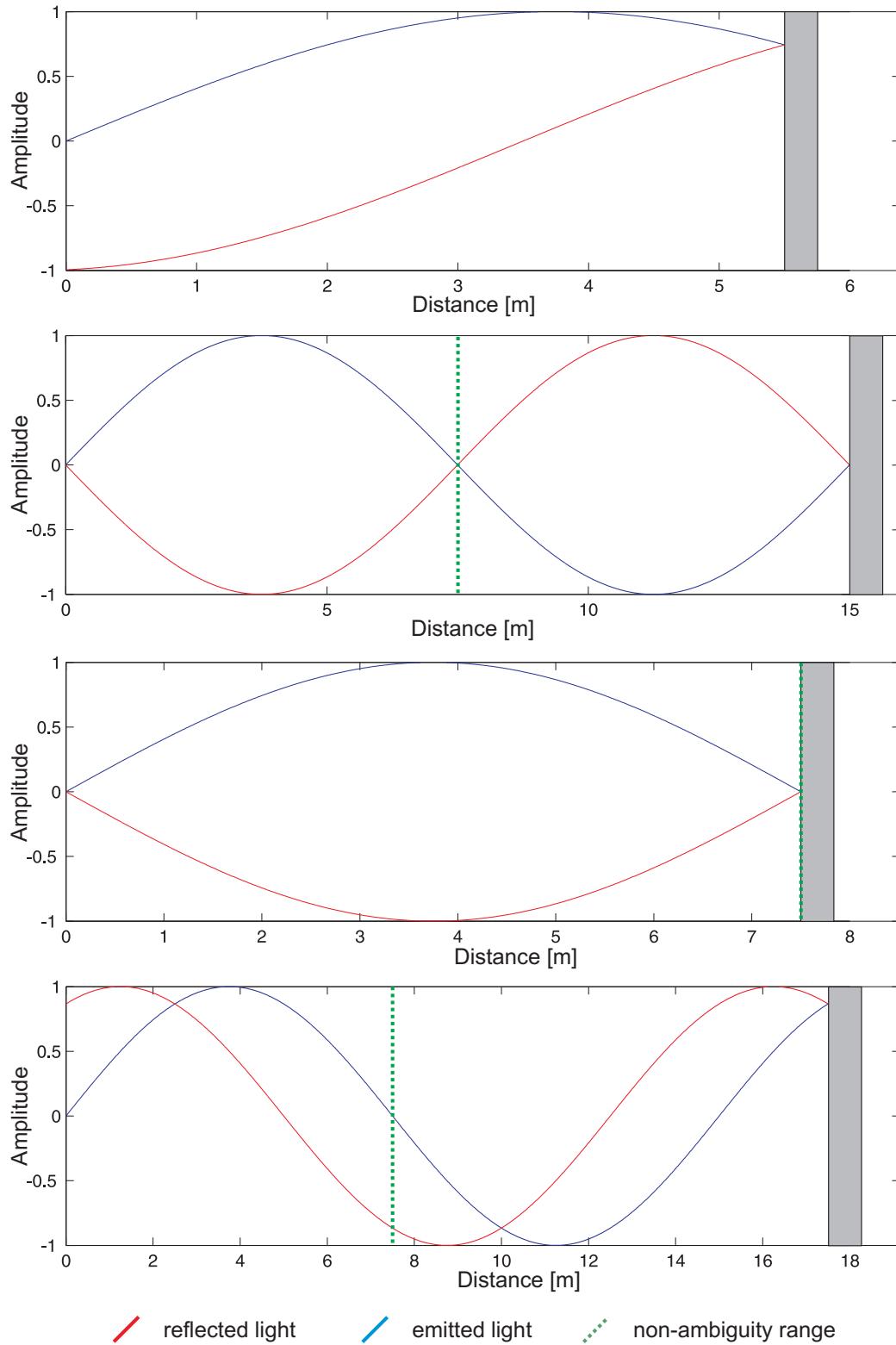


Figure 4.4: ToF principle for various distances at a modulation frequency of 20 MHz. Dashed green line indicates border of first non-ambiguity range for this modulation frequency. If the object is beyond this line, the distance cannot non-ambiguously be determined.

- **Laser range vs. ToF:** Real-time operation using a laser range scanner is usually not possible. Common laser range scanners combine a laser and a mechanical mirror to scan the environment. Other systems move a line across the object. There are also combinations using a laser line and a color camera to detect the line and extract 3-D information. Using a laser, eye-safety issues have to be considered. However, laser range scanner achieve a very high accuracy in the measured data. A major drawback for many applications are the high costs of high precision laser range systems.
- **Structured light vs. ToF:** It is often difficult to achieve real-time capability with structured light techniques, especially in setups that require a sequential light scan through the scene. Furthermore, structured light obtains direct measurements only at scene points illuminated by the light pattern. Depth values at non-illuminated points also have to be derived via interpolation or surface fitting like for stereo based systems.

Again, the accuracy depends on the setup and is depending on parameters like baseline, focus, light pattern, etc. For structured light systems it is also mandatory to calibrate the system in order to map the observed light pattern to 3-D point values. Besides calibrating the camera itself, the relative geometry of the light rays with respect to the camera has to be known. The projection unit has to be mounted very accurately and is also highly application dependent, what is leading to high costs.

- **Stereo vs. ToF:** The main difference between both systems is how depth information is achieved. Most stereo methods extract depth values at pixels where either corresponding features or texture information is available. To obtain dense depth information interpolation between these pixels is required. In contrast, due to their high lateral resolution ToF cameras innately have very dense depth information at constant resolution and high frame rates without the need of interpolation.

Both systems are able to operate in real-time, but the accuracy for stereo systems usually drops when trying to achieve real-time. Accuracy in stereo systems depends on the setup, where typical parameters are for example baseline, focal length, etc. ToF cameras are monocular all-solid-state cameras, hence no hardware setup depended parameters have influence on the accuracy. However, the main camera-specific parameters concerning accuracy are the amount of emitted light as well as the quality of the modulation signal and the transmission properties of the LEDs.

An inevitable step for stereo setups is calibration. Unless calibrated, a stereo system will not be able to convert disparity values to true depth estimates. Furthermore, many correspondence algorithms assume known epipolar geometry which is typically derived from calibration. Lastly, regular re-calibration and very robust stereo mounting are an issue for long-term use of stereo systems [Huss 08].

Table 4.1 shows a summary of the described differences of all 3-D surface acquisition technologies.

Method	Calibration	Accuracy	Speed (Motion)	Density	Costs
Laser range	●	●	○	●	○
Structured light	○	●	○	○	○
Stereo	○	○	○	○ ⁽¹⁾	○
Time-of-Flight	●	○	●	○	●

Table 4.1: Overview of advantages and disadvantages of different 3-D surface acquisition technologies. ⁽¹⁾ There is a strong dependency on the observed object. (Full circle represents the best.)

4.4 Time-of-Flight Camera Calibration

The manufacturer calibration of the ToF camera is overrided within this section. We apply both an optical and a distance calibration to the ToF camera. Further details about ToF camera calibration can also be found in [Fuch 08, Plac 10a].

4.4.1 Optical Calibration

Optical calibration is a well known problem in computer vision. Therefore a lot of methods already exist for conventional color cameras. For details see e.g. [Tsai 87] [Zhan 00] [Heik 97]. These methods can also be applied to ToF cameras. The optical calibration of the ToF camera has a huge influence on the quality of the 3-D coordinates. For the camera calibration the so-called pinhole model is used, where:

- s represents the cosine of the angle between x - and y -axis on the sensor
- $\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$ are the coordinates of a 3-D point in the world coordinate system
- $\begin{pmatrix} u \\ v \end{pmatrix}$ are the coordinates of the projection point in pixels
- A holds the intrinsic camera parameters, $R|t$ is a rotation matrix including translation
- c_x and c_y is the principle point (usually at the image center)
- f_x and f_y are the focal lengths (in pixel-related units)
- r_{ij} , where $i, j \in \{1, 2, 3\}$ are rotational and t_i translation parameters to describe the camera motion around a static scene or vice versa.

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \mathbf{A}(\mathbf{R}|\mathbf{t}) \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (4.27)$$

$$= \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (4.28)$$

Without lens distortion $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ can be computed using:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{R} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \mathbf{t} \quad (4.29)$$

$$x' = \frac{x}{z} \quad (4.30)$$

$$y' = \frac{y}{z} \quad (4.31)$$

$$u = f_x x' + c_x \quad (4.32)$$

$$v = f_y y' + c_y \quad (4.33)$$

$$(4.34)$$

ToF cameras usually have lenses with some distortion, mostly radial and slight tangential distortion. Therefore, the above model is extended as stated in the following equation, where

- k_i , where $i \in \{1, 2, 3\}$ are radial distortion coefficients
- p_j , where $j \in \{1, 2\}$ are tangential distortion coefficients

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{R} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \mathbf{t} \quad (4.35)$$

$$x' = \frac{x}{z} \quad (4.36)$$

$$y' = \frac{y}{z} \quad (4.37)$$

$$r^2 = x'^2 + y'^2 \quad (4.38)$$

$$x'' = x' (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 x' y' + p_2 (r^2 + 2x'^2) \quad (4.39)$$

$$y'' = y' (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1 (r^2 + 2y'^2) + 2p_2 x' y' \quad (4.40)$$

$$u = f_x x'' + c_x \quad (4.41)$$

$$v = f_y y'' + c_y \quad (4.42)$$

OpenCV² provides an implementation of the just introduced model. The optical calibration is performed using a planar checkerboard pattern. Using OpenCV, the nine degrees of freedom ($f_x, f_y, c_x, c_y, k_1, k_2, k_3, p_1, p_2$) are determined. To meet the special demands of ToF cameras, some adoptions of the calibration routine are applied:

- ToF cameras have a resolution of 200×200 pixels. It turned out that the OpenCV methods are performing better on higher resolution images. Thus, an upsampling of the amplitude image using bi-linear interpolation in each image dimension by the factor of 3 is performed before detecting the checkerboard.
- About 100 amplitude images with different orientations of a 8×12 checkerboard pattern are acquired. The inner checkerboard corners are detected by OpenCV (see Figure 4.5).
- Using the 2-D point correspondences of the corners, OpenCV computes the above stated nine parameters.

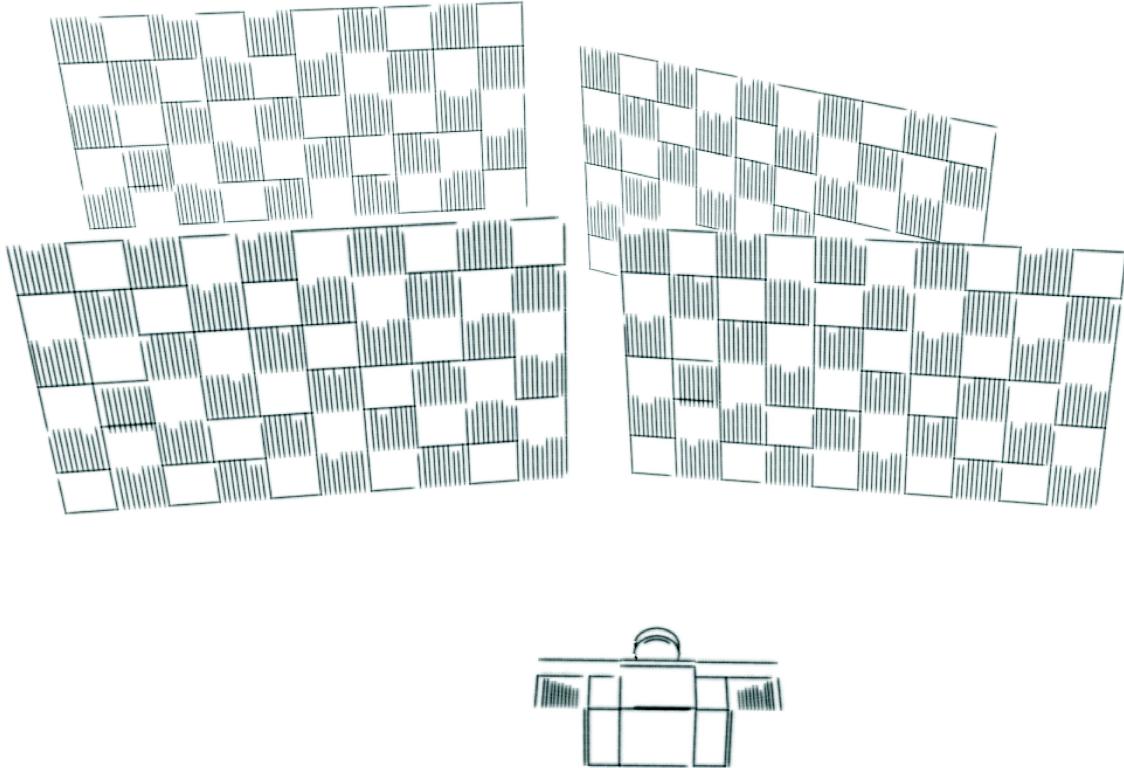


Figure 4.5: Setup for optical calibration. A checkerboard is placed in front of the camera at different distances and orientations.

Determination of 3-D Coordinates from Intrinsics and Distances

For a fast 3-D coordinate computation we precompute unit vectors for each camera ray. Therefore the computation of the 3-D points is reduced to a scaling of the vector with the

²<http://sourceforge.net/projects/opencvlibrary>

related distance value. By inverting equations (4.39) - (4.42) for each integer (u, v) the unit vectors can be computed. After having determined (x'', y'') the vector $(x'', y'', 1)$ is normalized to length one and stored in a pixel-dependent lookup table.

4.4.2 Distance Calibration

Distance values acquired by current ToF cameras are interfered with a systematic distance offset. This intensity-related distance error depends on:

- Integration time
- Intensity
- Distance
- Pixel coordinates

This section introduces a method to compensate the systematic errors by applying a calibration based on the test setup illustrated in Figure 4.6. The calibration is based on acquiring $u \in \{0, \dots, 7\}$ planar boards with different intensity at $v \in \{0, \dots, 2\}$ different distances.

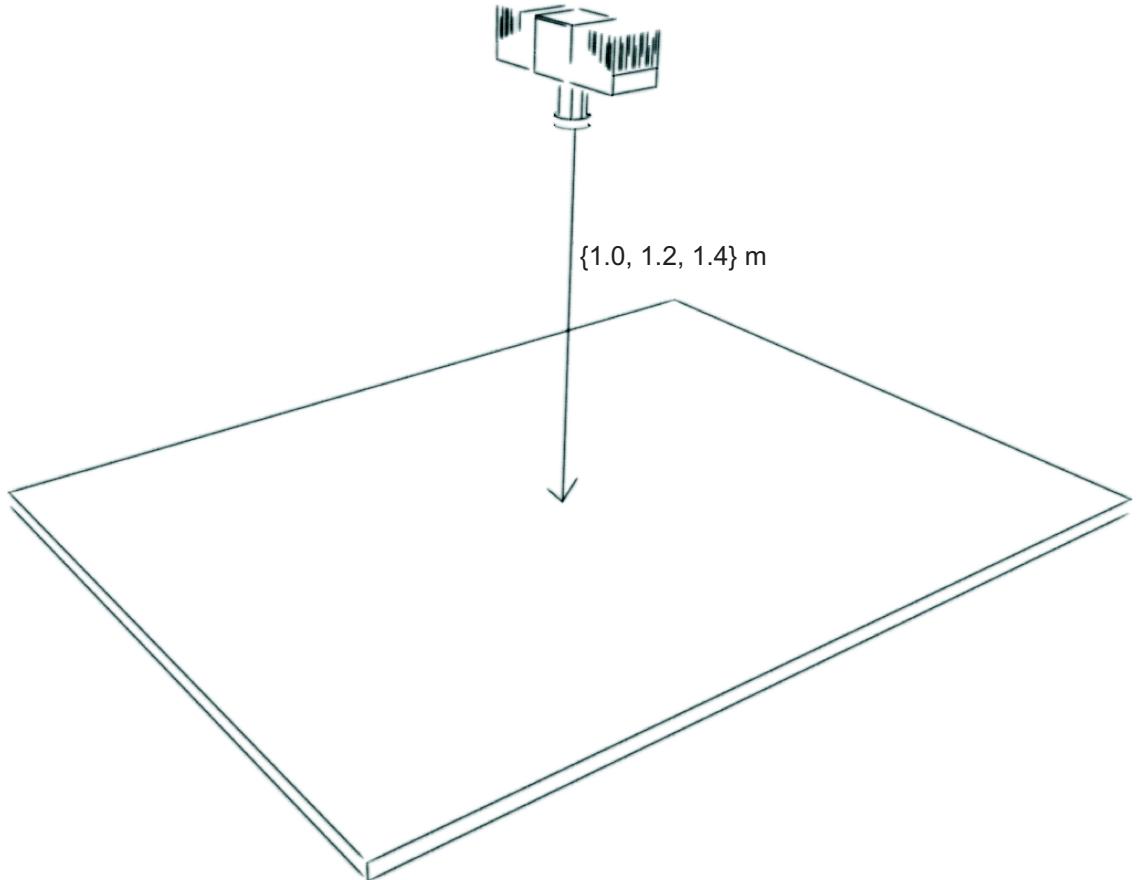


Figure 4.6: Setup for distance calibration. Boards with different reflectivity (gray values) are placed at different distances (1.0m, 1.2m, 1.4m) in front of the camera. The acquired distances are used to perform a distance calibration.

Integration Time

For the systems introduced in this thesis we can assume that the distance between the ToF camera and the patient is constant within a certain range. Therefore we can set the integration time of the ToF camera to a fixed value and reduce the dependency of the systematic error to three dimensions.

Intensity, Distance and Pixel Coordinates

We use eight different planar boards covered with paper in different gray-intensities. A standard laser printer is used to generate the gray shaded paper. Starting at 0% white (black), we process to 100% white in eight steps uniformly. Each board has a uniform gray tone and is positioned orthogonally at three different positions (1000 mm, 1200 mm and 1400 mm) to the viewing direction of the camera. The field of view of the camera has to cover the whole board. In order to eliminate temporal noise, all acquisitions are averaged over 1000 frames. Thereby, we sample the intensity domain at eight positions and the distance domain at three positions at each pixel coordinate.

Knowing that the acquired data represents a plane, we approximate each plane $E_{u,v}$ ($e_{u,v}^{(0)}x + e_{u,v}^{(1)}y + e_{u,v}^{(2)}z + e_{u,v}^{(3)} = 0$) by solving the following overdetermined system of equations for each board:

$$\begin{pmatrix} \hat{x}_{u,v}^{(0)} & \hat{y}_{u,v}^{(0)} & \hat{z}_{u,v}^{(0)} & 1 \\ \hat{x}_{u,v}^{(1)} & \hat{y}_{u,v}^{(1)} & \hat{z}_{u,v}^{(1)} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \hat{x}_{u,v}^{(n)} & \hat{y}_{u,v}^{(n)} & \hat{z}_{u,v}^{(n)} & 1 \end{pmatrix} \mathbf{e}_{(u,v)} = \mathbf{0}, \text{ where } n = (K - 1) \cdot (L - 1) \quad (4.43)$$

The set $\mathbf{I}_{u,v}$, $\mathbf{D}_{u,v}$ include the $K \times L$ intensities, distances acquired by a ToF camera for the measurements corresponding to $E_{u,v}$, where $k \in \{0, 1, \dots, (K - 1) \cdot (L - 1)\}$, $v \in \{0, \dots, 2\}$, $u \in \{0, \dots, 7\}$:

$$\mathbf{I}_{u,v} = \{o_{u,v,k} | o_{u,v,k} \in \mathbb{R}\} \quad (4.44)$$

$$\mathbf{D}_{u,v} = \{d_{u,v,k} | d_{u,v,k} \in \mathbb{R}\} \quad (4.45)$$

$$\mathbf{P}_{u,v} = \begin{pmatrix} x_{u,v,k} \\ y_{u,v,k} \\ z_{u,v,k} \end{pmatrix} \text{ corresponding 3-D coordinate} \quad (4.46)$$

An offset value depending on the intensity and the pixel index is computed for each acquired pixel:

$$\kappa_{u,v,k} = d_{u,v,k} \left(1 - \frac{e_{u,v}^{(3)}}{z_{u,v,k}}\right) \quad (4.47)$$

The formula can be verified by applying the theorem on intersecting lines (see Figure 4.7), where κ is the correction offset:

$$\frac{z_{u,v,k}}{e_{u,v}^{(3)}} = \frac{d_{u,v,k}}{d_{u,v,k} - \kappa} \quad (4.48)$$

$$z_{u,v,k}(d_{u,v,k} - \kappa) = d_{u,v,k}e_{u,v}^{(3)} \quad (4.49)$$

$$-z_{u,v,k}\kappa + z_{u,v,k}d_{u,v,k} = d_{u,v,k}e_{u,v}^{(3)} \quad (4.50)$$

$$z_{u,v,k}\kappa = z_{u,v,k}d_{u,v,k} - d_{u,v,k}e_{u,v}^{(3)} \quad (4.51)$$

$$\kappa = \frac{z_{u,v,k}d_{u,v,k}}{z_{u,v,k}} - \frac{d_{u,v,k}e_{u,v}^{(3)}}{z_{u,v,k}} \quad (4.52)$$

$$\kappa = d_{u,v,k} - \frac{d_{u,v,k}e_{u,v}^{(3)}}{z_{u,v,k}} \quad (4.53)$$

$$= d_{u,v,k}\left(1 - \frac{e_{u,v}^{(3)}}{z_{u,v,k}}\right) \quad (4.54)$$

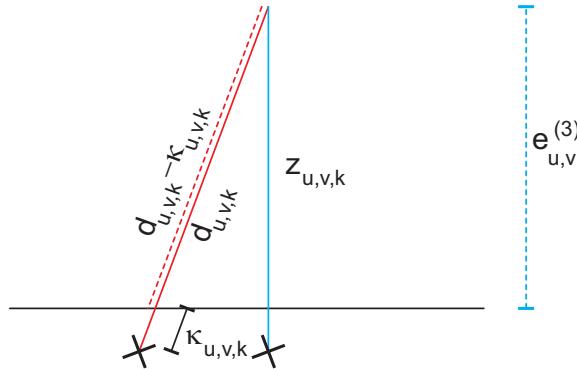


Figure 4.7: Computation of κ by the theorem on intersecting lines.

Based on these values, three look-up tables $\text{LT}_v \in \mathbb{R}$ are computed. One for each acquired plane $E_{u,v}$. The first dimension of the look-up table is the intensity domain. It is sampled equidistantly between the minimum intensity value and the maximum intensity value of all intensities. k sets the second dimension of the look-up table. As there are also intensities between the acquired samples, a spline based interpolation is applied to other values. We use Kochanek-Bartels Splines [Koch 84] to interpolate the systematic error $\kappa_{u,v,k}$ between the eight acquired intensities. The following parameter settings are used for the splines: bias = 0.0, tension = 0.0, continuity = 0.0:

Figure 4.8 shows the linear interpolation in the distance domain, after finding the nearest intensity values $\kappa_{\text{low}}, \kappa_{\text{high}}$ in both distance values, the interpolated offset value κ_{ip} is computed.

Gain of Distance Correction

To measure the quantitative quality G_c of the calibration, a flat board is acquired at distances lying between the samples used for the calibration. The N measured 3-D plane coordinates $\mathbf{u}_i \in \mathbb{R}^3$ are approximated by a plane P_c . After that, the quality criterion G_c is

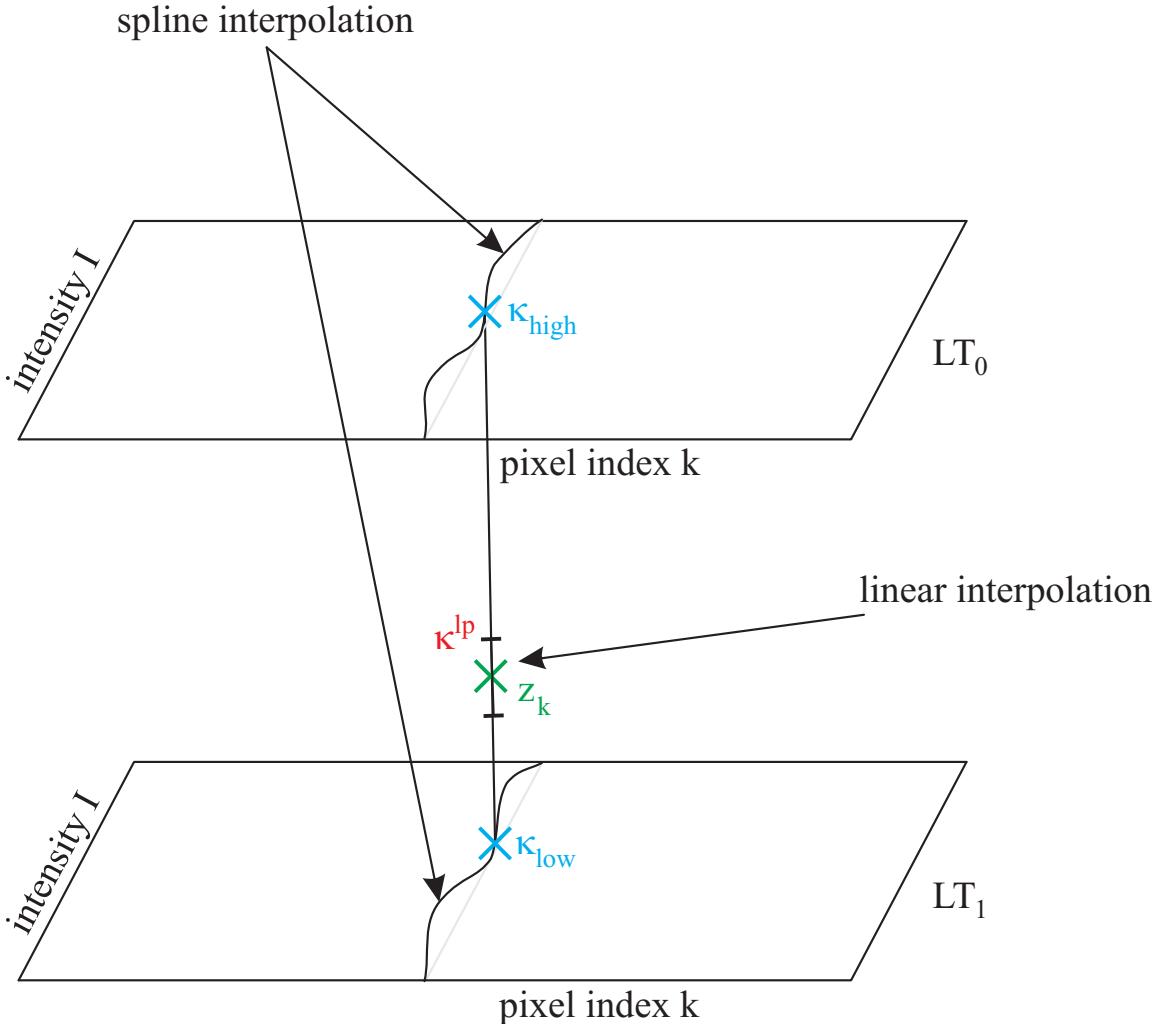


Figure 4.8: Computation of distance correction for a measured z_k value. The two planes represent the look-up table samples, with the pixel position in one direction and the intensity samples in the other direction, for different positions (e.g. 1.0 m and 1.2 m). κ^{high} and κ^{low} are computed using a spline interpolation. The distance offset $\kappa^{(ip)}$ is computed using a linear interpolation.

computed as the mean of absolute distances (MAD) between the measured points and the approximation plane:

$$G_c = \frac{1}{N} \sum_{i=1}^N \text{distance}(P_c, \mathbf{u}_i)^2 \quad (4.55)$$

The quality function G_c is evaluated for the 85.7% white board at three distances between the 1200 and the 1400 distance calibration sample (see Figure 4.9).

Using the quality criterion G_c , the distance correction increases the (relative) accuracy of the distance data by approximately 400%. Figure 4.10 shows the effect of the distance calibration on realistic data. It shows the rigid plaster cast phantom introduced in the next section with no correction (left) and enabled correction (right).

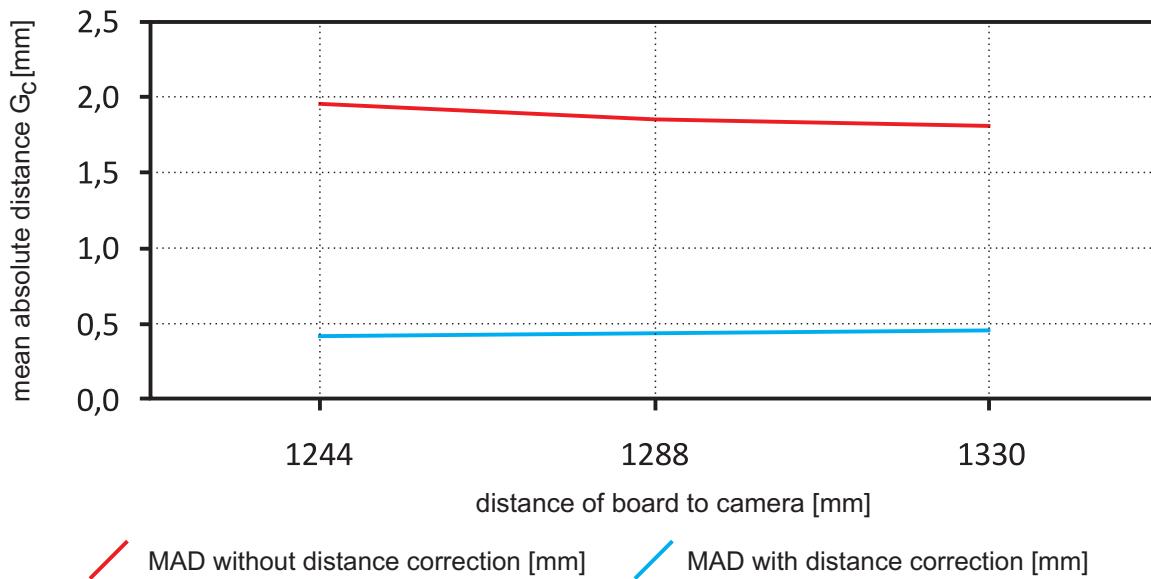


Figure 4.9: Gain of distance correction for three different samples.

4.5 Conclusion

Within this chapter an overview of the Time-of-Flight technology was given. The basic principle of the technology was presented and important key advantages of ToF cameras over other technologies were discussed. To improve data acquired by ToF cameras, two calibration methods were presented. An optical calibration, using a well-established camera calibration model and a distance calibration. Using both calibrations, we could show an improvement of the data by approximatley 400%.

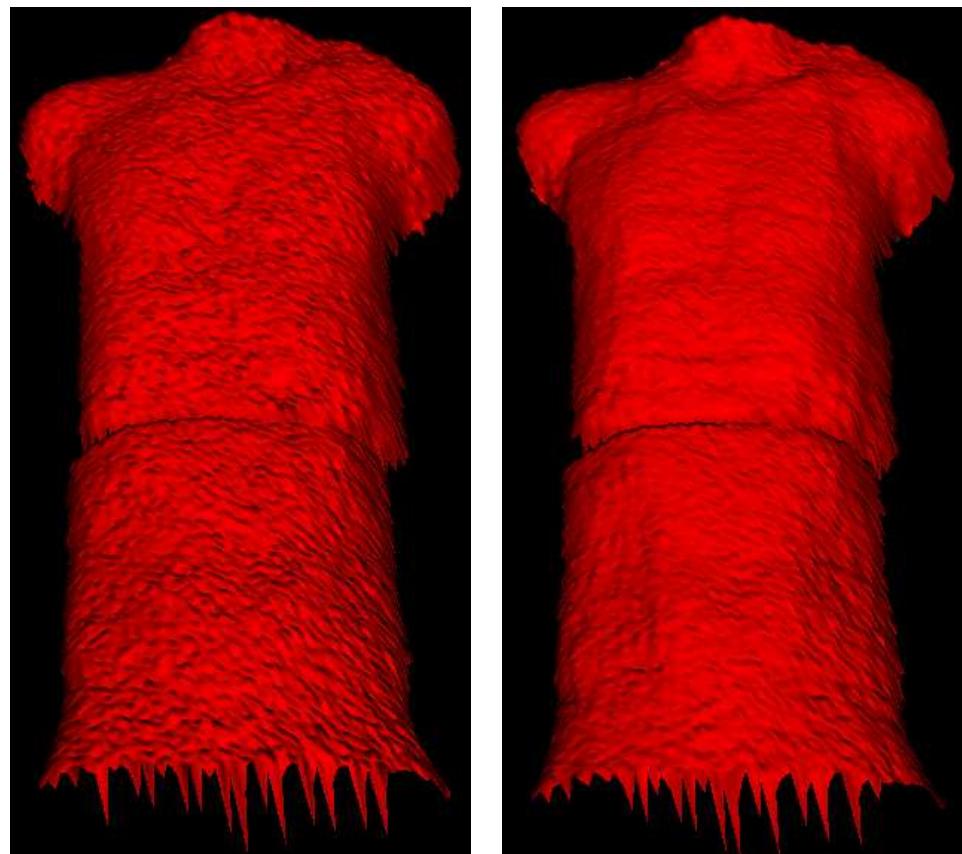


Figure 4.10: Left: surface without distance correction. Right: surface after correction.

CHAPTER V

Time-of-Flight based Patient Positioning

The chapter presents two positioning systems. In advance, a calibration method to align both, the LINAC and the ToF coordinate system is presented. The first method positions a C-arm like device with respect to the patient. The second system introduces a method to position a patient, e.g. for a radiotherapy treatment. Each section has a short state-of-the art overview, a method part and an evaluation part.

5.1 Introduction

Workflow optimization is an important task during clinical procedures. The combination of increased life expectancy and advancements in the field of medicine have resulted in a constantly increasing number of patients. This in-turn creates a heavier workload for hospitals and healthcare systems in general. Thus, it is becoming progressively more important to develop efficient healthcare procedures without any decline of the quality of patient care. In the future either healthcare costs will increase dramatically or smart cost-effective solutions for optimizing current workflows have to be found.

With that last goal in mind we examined the image acquisition workflow and treatment workflow respectively and analyzed it in terms of bottlenecks. In this chapter we introduce two systems to shorten the probably most important and for sure most time consuming part of these procedures: the initial patient positioning step. Image acquisition and treatment procedures have constantly been optimized within the past years. Today, the scanning time for CT and C-arm systems is basically negligible. Tumors can be irradiated with high dose within radiation therapy.

What is still very time consuming within the workflow of image acquisition procedures and very critical within radiotherapy is the patient-dependent setup procedure. Each patient is physiologically comparatively unique. There is also a diverse number of clinical procedures. As a result, the patient setup process cannot be easily generalized. Nevertheless, the correct positioning of the patient is a crucial parameter for the quality of image acquisition or radiation treatment. It is important that patient positioning is performed ac-

curately, which can make it a very time consuming task. For retrospective hybrid-image acquisitions it is also very important to acquire images at the correct positions. The fusion of different imaging modalities is much more easier when the relation is known between both images. This does not only save time but also enables better diagnostic results.

The Time-of-Flight camera technology will be used to propose methods to align patients for various image acquisition devices. This will include modern C-arm systems, CT scanners and C-arm like systems (like radiotherapy devices). Furthermore it can be used for PET- and SPECT acquisitions as well. In general the proposed systems also work for magnetic resonance tomography. However, because of the high magnetic fields, the ToF camera itself has to be modified first.

The proposed systems are tailored to the target systems and have some specific requirements. Basically there are two different ways of positioning. We separate the options into conventional positioning and so-called *inverse positioning*. The term conventional positioning expresses the patient alignment with respect to the imaging system. By means of *inverse positioning* we position the whole system with respect to the patient. *Inverse positioning* is used to position for example C-arm systems. All proposed systems have a one-time calibration step in common which is introduced in the next section.

5.2 Treatment Table Calibration

For conventional positioning as well as *inverse positioning* the positioning system coordinate system has to be calibrated with the patient table coordinate system. Therefore, the table coordinate system has to be determined. Patient tables are usually able to move the patient in all three space dimensions and to rotate the patient around the isocenter of the *z*-axis. There are other systems, which also allow rotations around the *x*-axis and the *y*-axis. Placht et al. [Plac 10a] introduced a system to calibrate both coordinate systems automatically. The system works with basic table systems, providing a three axes translation and a rotational part around the *z*-axis.

In Figure 5.1 the setup used to detect the table coordinate system is illustrated. A standard black and white checkerboard is placed on the patient table. It has to be visible by the ToF camera in various positions. Please note that we also use the backside of the checkerboard. It should be uniformly white and diffusively reflecting. The checkerboard detection is mainly based on the amplitude image of the ToF camera. To determine the coordinate system, the distance image is used as well.

The overall process can be separated in seven steps. Each step will be explained in detail next:

- **2-D corner detection:** The translational components of the table are reset to (0, 0, 0) and the rotation angle is reset to 0.0°. The process is starting with translating the table stepwise along the *x*- and *y*-direction of the table and rotating the table around the isocentric *z*-axis. As already used for the optical calibration of the Time-of-Flight cameras, the OpenCV framework (`cvFindChessboardCorners` and `cvFindCornerSubPix`) is used again to determine the inner checkerboard corners in 2-D. We then assign the 2-D corner coordinates to the three classes *x*-translation, *y*-translation

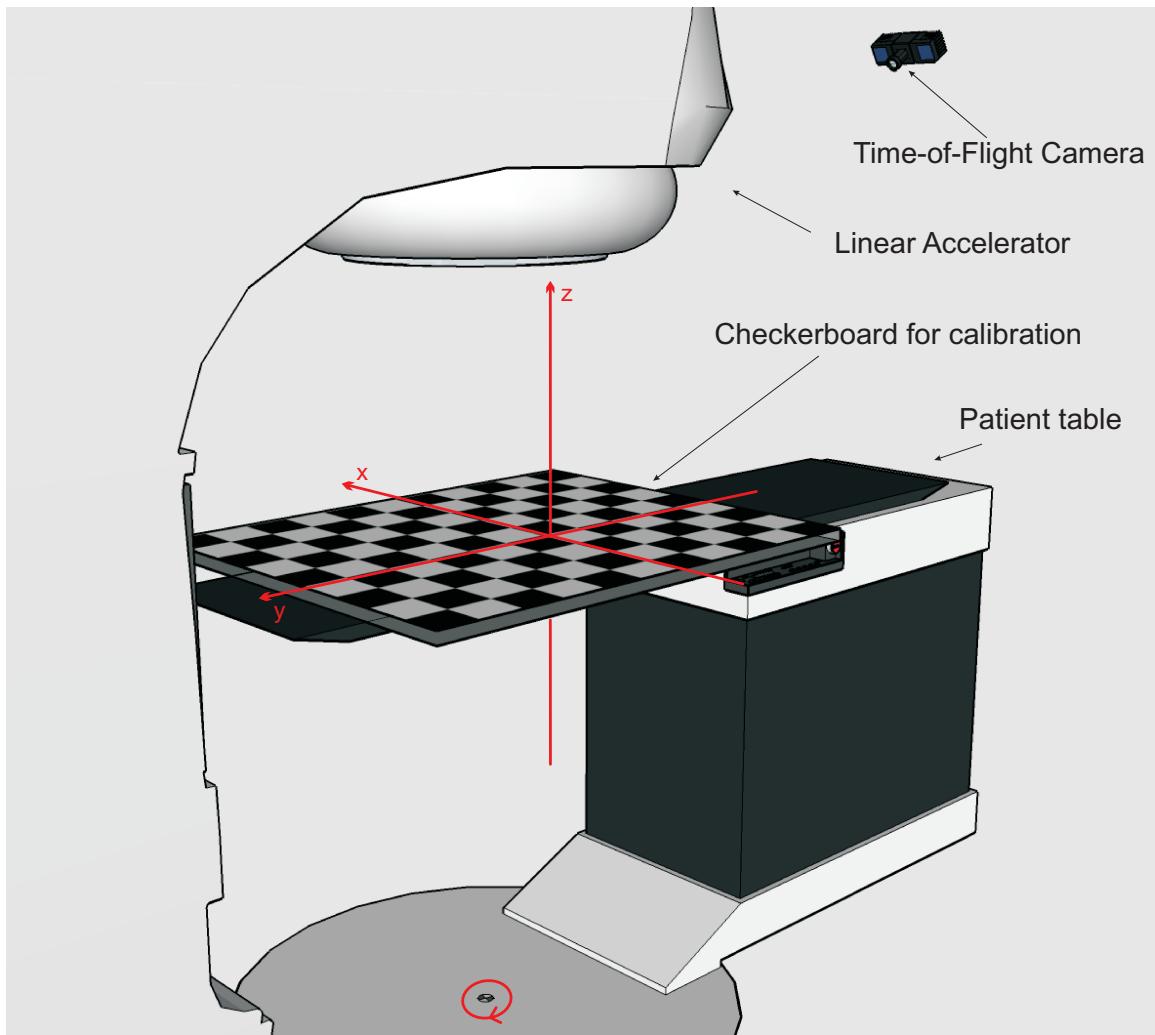


Figure 5.1: Setup to detect the treatment couch coordinate system. The x , y , z coordinate system and the rotational part around the z -axis is shown in red. The checkerboard has to be visible by the ToF camera all time. This calibration is a one-time calibration.

and isocentric rotation. Pairs are built in relation to corresponding corners in the real checkerboard.

- **3-D plane acquisition:** To determine a plane in 3-D, the checkerboard has to be turned over. This is required due to the so-called intensity related distance error of current ToF cameras, which can not be completely eliminated by the distance calibration (see Chapter 4). Temporal noise is suppressed by adding the distance image of the calibration pattern over 3000 frames. We then can use the distance image of the white board to approximate a plane P_w in 3-D.
- **3-D corner detection:** To determine the 3-D corner coordinates from the pixel position and the plane P_w we take again the optical calibration model used by OpenCV under consideration. We already introduced this model in Chapter 4. The equations (4.39, 4.42) of the model have to be inverted, so that the 3-D rays corresponding to the corner pixel positions can be identified. By intersecting these rays with the plane P_w , we know the 3-D coordinates of all checkerboard corners.

- **Computation of x - and y -axis:** We define the $\mathbf{x}_i^{(j)}$ and $\mathbf{y}_i^{(j)}$ to be the i -th 3-D corner of the $W \times H$ checkerboard pattern acquired in the j -th translation image along the x/y -direction. The N_t translation indices j correspond to translations in ascending order (from negative to positive translations). By averaging over all difference vectors (5.1, 5.2), the unnormalized directions $\mathbf{d}_x^{(u)}$ and $\mathbf{d}_y^{(u)}$ of the x - and y -axis are determined.

$$\mathbf{d}_x^{(u)} = \sum_{i=1}^{WH} \sum_{j=1}^{N_t} \sum_{k=j+1}^{N_t} (\mathbf{x}_i^{(k)} - \mathbf{x}_i^{(j)}) \quad (5.1)$$

$$\mathbf{d}_y^{(u)} = \sum_{i=1}^{WH} \sum_{j=1}^{N_t} \sum_{k=j+1}^{N_t} (\mathbf{y}_i^{(k)} - \mathbf{y}_i^{(j)}) \quad (5.2)$$

Finally, $\mathbf{d}_x^{(u)}$ and $\mathbf{d}_y^{(u)}$ are normalized to the unit vectors \mathbf{d}_x and \mathbf{d}_y .

- **Computation of z -axis direction:** The desired coordinate system should be Cartesian. Therefore $\mathbf{d}_z = \mathbf{d}_x \times \mathbf{d}_y$. Figure 5.2 shows the coordinate system determination.

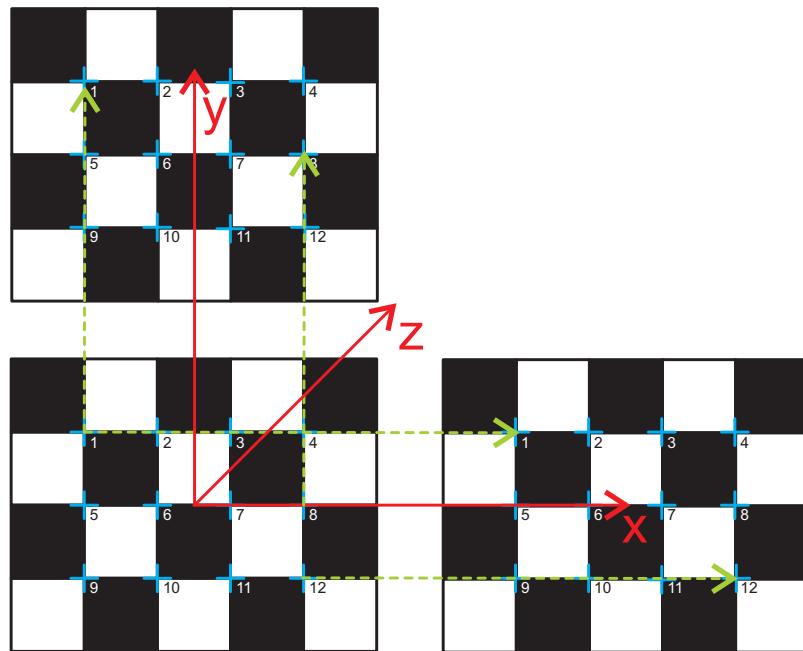


Figure 5.2: Determination of x , y and z axes of the coordinate system. The corresponding vectors (green) are computed and averaged (red).

- **Computation of the coordinate system origin:** To compute the coordinate system origin, we define an arbitrary 2-D coordinate system inside the plane P_w . The coordinate system origin is \mathbf{o} , its abscissa \mathbf{v}_1 and its ordinate \mathbf{v}_2 ($\mathbf{o}, \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3$). All checkerboard 3-D coordinates $\mathbf{r}_i^{(j)} \in \mathbb{R}^3$ belonging to the rotation class are transformed into the new coordinate system. Again, i is the index for the current chessboard corner and j represents the index for the number of rotations N_r . The

transformed 2-D points are denoted $\mathbf{p}_i^{(j)} \in \mathbb{R}^2$. As rotations around the origin of a coordinate system do not change the distance of the rotated point to the origin, the detection of the 2-D rotation center $\mathbf{c} = (c_1, c_2)^T$ inside the plane P_w can be formulated as an optimization problem:

$$f_{i,j,k} = \left(\left\| \mathbf{p}_i^{(j)} - \mathbf{c} \right\|_2^2 - \left\| \mathbf{p}_i^{(k)} - \mathbf{c} \right\|_2^2 \right)^2 = \quad (5.3)$$

$$\left[(\mathbf{p}_i^{(j)} - \mathbf{c})^T (\mathbf{p}_i^{(j)} - \mathbf{c}) - (\mathbf{p}_i^{(k)} - \mathbf{c})^T (\mathbf{p}_i^{(k)} - \mathbf{c}) \right]^2 = \quad (5.4)$$

$$\left[\mathbf{p}_i^{(j)T} \mathbf{p}_i^{(j)} - 2\mathbf{p}_i^{(j)T} \mathbf{c} + \mathbf{c}^T \mathbf{c} - \mathbf{p}_i^{(k)T} \mathbf{p}_i^{(k)} + 2\mathbf{p}_i^{(k)T} \mathbf{c} - \mathbf{c}^T \mathbf{c} \right]^2 = \quad (5.5)$$

$$\left[\mathbf{p}_i^{(j)T} \mathbf{p}_i^{(j)} - \mathbf{p}_i^{(k)T} \mathbf{p}_i^{(k)} - 2(\mathbf{p}_i^{(j)T} - \mathbf{p}_i^{(k)T}) \mathbf{c} \right]^2 \quad (5.6)$$

$$\hat{\mathbf{c}} = \underset{\mathbf{c}}{\operatorname{argmin}} \sum_{i=1}^{WH} \sum_{j=1}^{N_r} \sum_{k=j+1}^{N_r} f_{i,j,k} \quad (5.7)$$

For the solution of the optimization problem partial derivatives have to be computed. Since all non-linear terms of c_1, c_2 disappear after derivation, a linear solver $\mathbf{M}\hat{\mathbf{c}} = \mathbf{b}$ can be implemented:

$$\frac{df_{i,j,k}}{dc} = 2 \left[\mathbf{p}_i^{(j)T} \mathbf{p}_i^{(j)} - \mathbf{p}_i^{(k)T} \mathbf{p}_i^{(k)} - 2(\mathbf{p}_i^{(j)T} - \mathbf{p}_i^{(k)T}) \mathbf{c} \right] . \quad (5.8)$$

$$(-2(\mathbf{p}_i^{(j)} - \mathbf{p}_i^{(k)})) \quad (5.9)$$

$$= 4 \left[\mathbf{p}_i^{(j)T} \mathbf{p}_i^{(j)} - \mathbf{p}_i^{(k)T} \mathbf{p}_i^{(k)} - 2(\mathbf{p}_i^{(j)T} - \mathbf{p}_i^{(k)T}) \mathbf{c} \right] . \quad (5.10)$$

$$(-(\mathbf{p}_i^{(j)} - \mathbf{p}_i^{(k)})) \quad (5.11)$$

$$= 4 \underbrace{\left(\left\| \mathbf{p}_i^{(k)} \right\|_2^2 - \left\| \mathbf{p}_i^{(j)} \right\|_2^2 \right) (\mathbf{p}_i^{(k)} - \mathbf{p}_i^{(j)})}_{b_{i,j,k}} + \quad (5.12)$$

$$2 \underbrace{(\mathbf{p}_i^{(k)} - \mathbf{p}_i^{(j)}) (\mathbf{p}_i^{(k)} - \mathbf{p}_i^{(j)})^T}_{\mathbf{M}_{i,j,k}} \mathbf{c} \quad (5.13)$$

$$\stackrel{!}{=} 0 \quad (5.14)$$

$$\mathbf{M} = \sum_{i=1}^{WH} \sum_{j=1}^{N_r} \sum_{k=j+1}^{N_r} 2 (\mathbf{p}_i^{(k)} - \mathbf{p}_i^{(j)}) (\mathbf{p}_i^{(k)} - \mathbf{p}_i^{(j)})^T \quad (5.15)$$

$$\mathbf{b} = \sum_{i=1}^{WH} \sum_{j=1}^{N_r} \sum_{k=j+1}^{N_r} \left(\left\| \mathbf{p}_i^{(k)} \right\|_2^2 - \left\| \mathbf{p}_i^{(j)} \right\|_2^2 \right) (\mathbf{p}_i^{(k)} - \mathbf{p}_i^{(j)}) \quad (5.16)$$

Using SVD, the estimated 2-D rotation center $\hat{\mathbf{c}}$ can be computed. By applying a back-transformation to the original coordinate system, the real 3-D rotation center \mathbf{c}_{3D} is computed:

$$\mathbf{c}_{3D} = \mathbf{o} + c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 \quad (5.17)$$

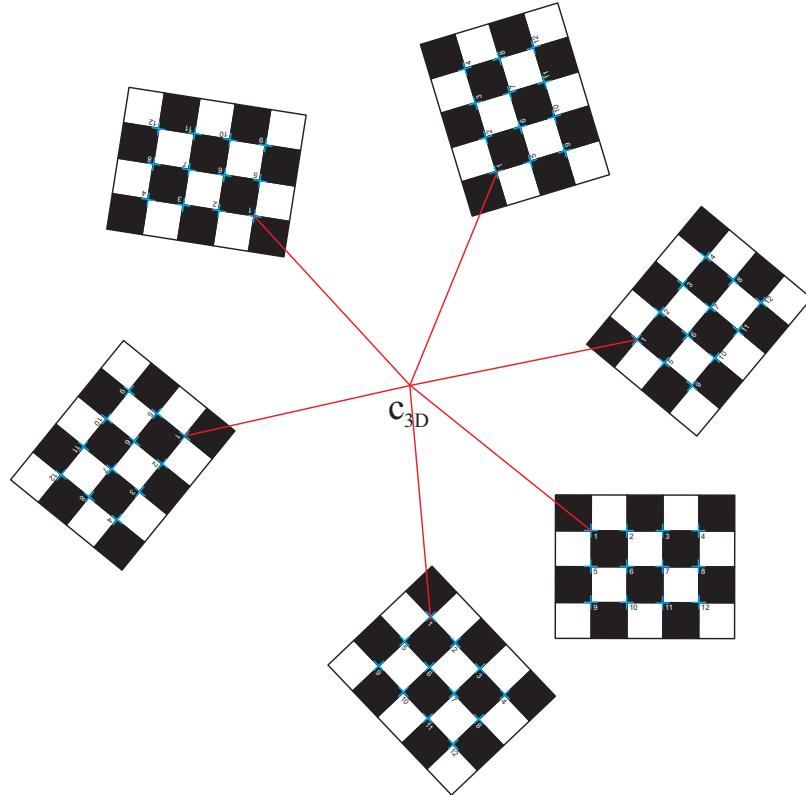


Figure 5.3: Determination of rotation center, by optimizing c_{3D} .

Figure 5.3 shows the coordinate origin computation.

- **Definition of a transformation from camera to table coordinates:** The treatment couch coordinate system is completely given by the origin c_{3D} and the axes d_x , d_y and d_z . The transformation from camera coordinates p_c to table coordinates p_t is given by:

$$p_t = (d_x, d_y, d_z)^T (p_c - c_{3D}) \quad (5.18)$$

This calibration step links the result of the ToF based patient registration to the table control. It is now possible to directly take the registration results as an input for moving the patient table to the correct position.

5.3 Patient Positioning for C-arm Systems

5.3.1 Introduction

Especially in interventional procedures correct positioning of the image acquisition device is mandatory. Todays C-arm systems are already very intuitive to use. Strobel et al. [Reis 09] describe the setup procedure and the associated time consuming steps. Several fluoroscopic images have to be taken before the required body part is within the isocenter of the C-arm. These steps are both very time consuming and apply additional radiation dose to the patient.

To automate and speed-up the setup procedure for interventional image acquisition procedures we suggest a system based on a single ToF sensor. The proposed system partitions the whole body into several 3-D bounding boxes. It operates marker-lessly and does not rely on any other assumptions. The C-arm system receives exact metric coordinates from the proposed system which then can be used in adjusting the C-arm's position with respect to the patient and for data acquisition automatically. We call this *inverse positioning*, as the system is positioned with respect to the target and not the target with respect to the system. Furthermore the C-arm exactly knows the bounds of the object to scan and an instantaneous collision detection can be provided. Last but not least, the proposed solution is very cost-effective as only one ToF camera is used for positioning and collision detection.

5.3.2 Related Work

The importance of a solution to these problems is also reflected by prior work. Crimson et al. [Grim 98] introduced a system which supports optical tracking of patient and instrument locations using surface data. Navab et al. [Nava 06] proposed a system for intraoperative positioning and repositioning of mobile C-arms using a camera-augmented mobile C-arm. This system speeds up the whole procedure and also reduces the radiation of the patient as it decreases the number of images which have to be acquired for patient positioning. Using an optical camera and x-ray/optical markers the system supports guidance for C-arm repositioning. C-arm systems also suffer from their narrow field-of-view. Wang et al. [Wang 08] suggest a stitching algorithm for scanning long bones using a C-arm system. They, also use an additional optical camera to augment different views. Ladikos et al. [Ladi 08] provide a method for collision detection for a C-arm environment. An array of optical cameras is used to generate a 3-D representation of the operation room. All of these applications are either based on markers and/or use multiple cameras/systems. Our proposed system is independent of markers and only needs a single sensor.

5.3.3 Materials and Methods

In this section we will give an overview of how a ToF sensor enables inverse real-time positioning. We will introduce a generalized, very fast, and fairly efficient algorithm to solve this task. The proposed algorithm for body part detection consists of two main parts which require preprocessing and segmentation of the ToF sensor data. The first part describes an effective and robust two-stage classification procedure. At the end of this first subtask, the

3-D body surface is subdivided into multiple anatomically meaningful regions of interest. The second part computes bounding boxes for each of these regions and their corresponding isocenter. Furthermore a fairly efficient calibration method is introduced to align the computed isocenter with the C-arm isocenter.

The ToF sensor is rigidly mounted above the patient table and the whole patient is within the field of view of the ToF sensor. We denote \mathbf{P} as a set of 3-D points, e.g. the $(K \times L)$ 3-D points of interest acquired by a ToF sensor.

$$\mathbf{P} = [\mathbf{p}_{i,j}] , i \in \{0, 1, \dots, K - 1\}, j \in \{0, 1, \dots, L - 1\} \quad (5.19)$$

Typically ToF data is affected by noise. To reduce this noise we apply both a bilateral filter [Pari 09] and a temporal averaging filter. The averaging filter returns the average 3-D point cloud using data from the last n accumulated 3-D point clouds.

To reduce the amount of data and to identify 3-D points belonging to the body we detect the patient table and compute a virtual plane \mathbf{T} . For this task a Hough-Transform like method based on surface normals described in Chapter 6 can be used. One can also use the just described calibration routine to determine the patient table plane. Knowing the virtual table plane we can discard all 3-D points which do not belong to the patient. As a result for further computations, only relevant 3-D points $\hat{\mathbf{P}} \subseteq \mathbf{P}$ (\subseteq denotes a subset of points) belonging to the body have to be considered.

To speed-up the classification process, we apply a normalization on the remaining 3-D points $\hat{\mathbf{P}}$. Therefore a Karhunen-Loeve-Transformation (Principle Component Analysis, PCA) is performed. As a result the origin of the coordinate system is placed in the centroid μ of the point cloud $\hat{\mathbf{P}}$ and the axes are aligned with the axial, sagittal, and coronal plane of the patient. The main advantage we achieve is parallelism of the bounding boxes to the axes. This results in a more computationally efficient run-time for the algorithm. The three principal axes computed by the transformation are denoted as ρ_0, ρ_1, ρ_2 (see Fig. 5.4). We have prior knowledge about the shape of the human body. Therefore, we can assume that ρ_0 corresponds to the height of the patient, ρ_1 to the width, and ρ_2 to the depth.

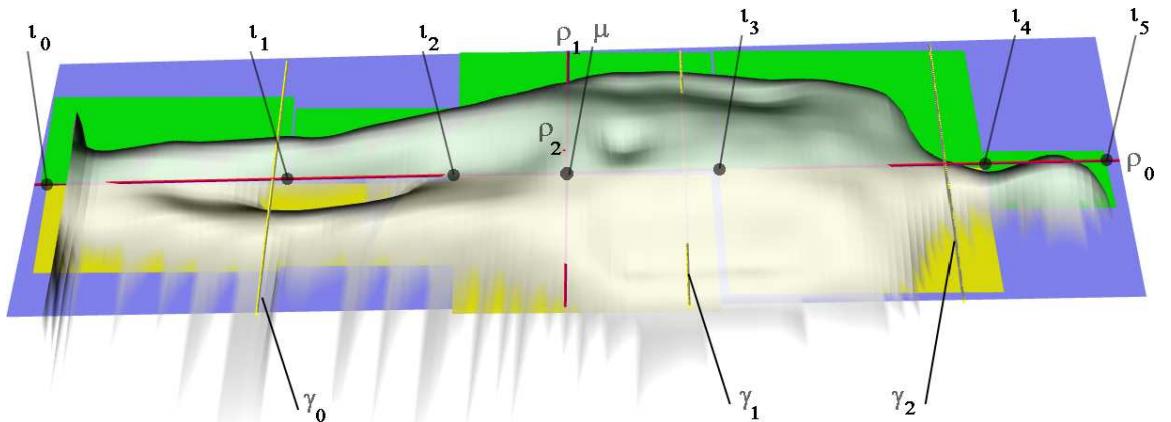


Figure 5.4: Overview: The red lines indicate the three principal axes ρ_0, ρ_1, ρ_2 . The three golden sections $\gamma_0, \gamma_1, \gamma_2$ are shown as yellow lines. Furthermore all intersection points i_k and the centroid μ are also depicted.

One can then compute six intersections (ι_k , where $k \in \{0, 1, \dots, 5\}$) along the first principle axis (see Fig. 5.4). The outer most intersection points ι_0 and ι_5 are defined by the outermost 3-D points on ρ_0 .

Again, we utilize prior knowledge about the shape of the human body. For a coarse first stage initialization of the classification, we compute three golden intersections γ_0 , γ_1 , γ_2 along ρ_0 (see Fig. 5.4). The golden section γ_2 roughly indicates the position of the neck. γ_1 is located near the diaphragm and γ_0 at the knees of the person. These points constrain the search space for the exact position of the remaining intersection points.

To refine these coarse initial values we introduce a second stage in our classification. Two histograms H_0 and H_1 along ρ_0 are computed. The vertical axis of H_0 bins the silhouette of the body along the positive ρ_1 axis, while the vertical axis of H_1 bins it along the positive ρ_2 direction. The horizontal axis of H_0 ranges from $\gamma_2 - d_0$ to $\gamma_2 + d_0$. The horizontal axis of H_1 ranges from $\mu - d_1$ to $\mu + d_1$. d_0 and d_1 are empirically determined. Based on these histograms, ι_4 and ι_5 are the minima of H_0 and H_1 accordingly. ι_3 is then defined as the mid-point between ι_4 and ι_5 , while ι_1 is set to γ_0 (see Figure 5.5).

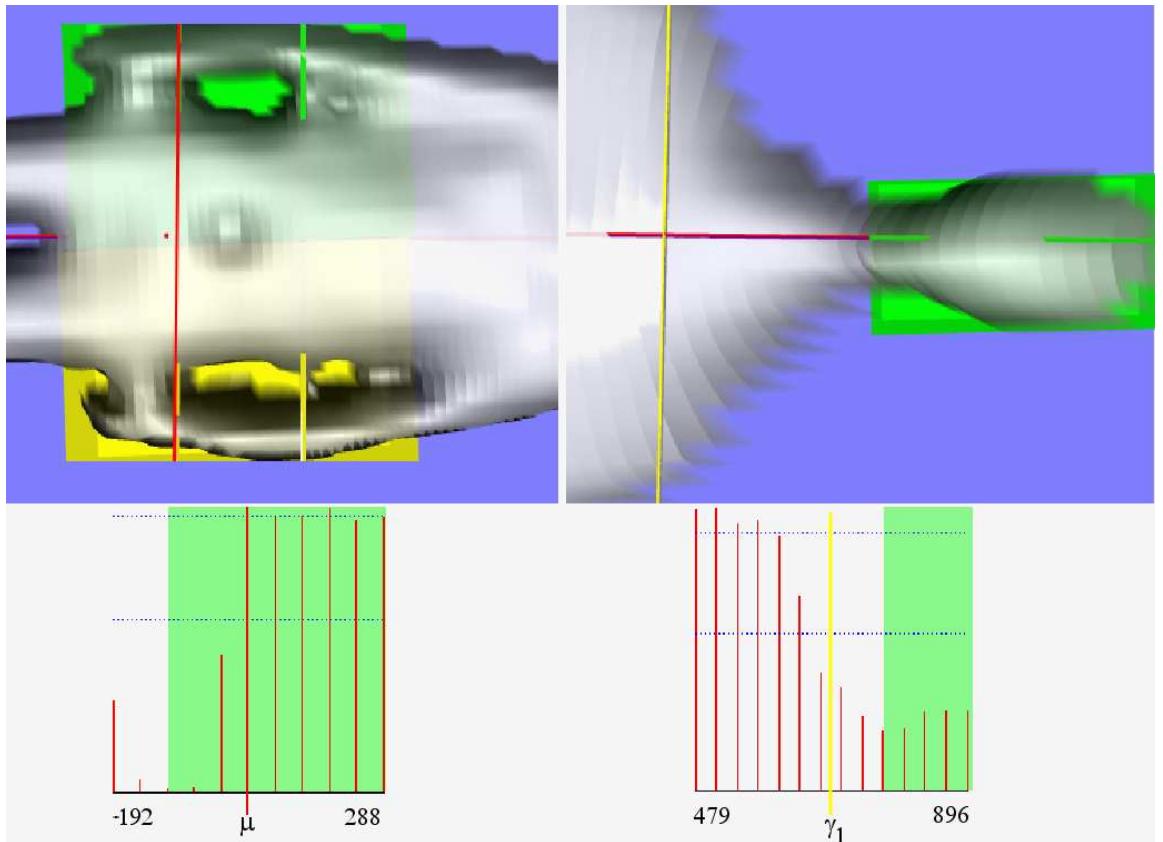


Figure 5.5: Refinement of the bounding boxes. Left: Detection of the pelvic region with the corresponding histogram below. μ is the origin of the coordinate system. The search interval d_1 has a length of 30 cm in each direction with a binning size of 4.8 cm. Right: Detection of the neck with the corresponding histogram below. γ_1 is the upper golden section intersection. The search interval d_0 has a length of 10 cm with a binning size of 3.2 cm.

5 Time-of-Flight based Patient Positioning

We can then use these intersections as an input for the second part of the algorithm. This part computes bounding boxes using ι_0, \dots, ι_5 . With the exception of the head, left and right bounding boxes are computed for each body segment (e.g. left and right abdomen), where ρ_0 acts as a delimiter. In addition to the boundary values, the isocenter and the volume is computed for each box. Figure 5.6 shows the full body part segmentation.

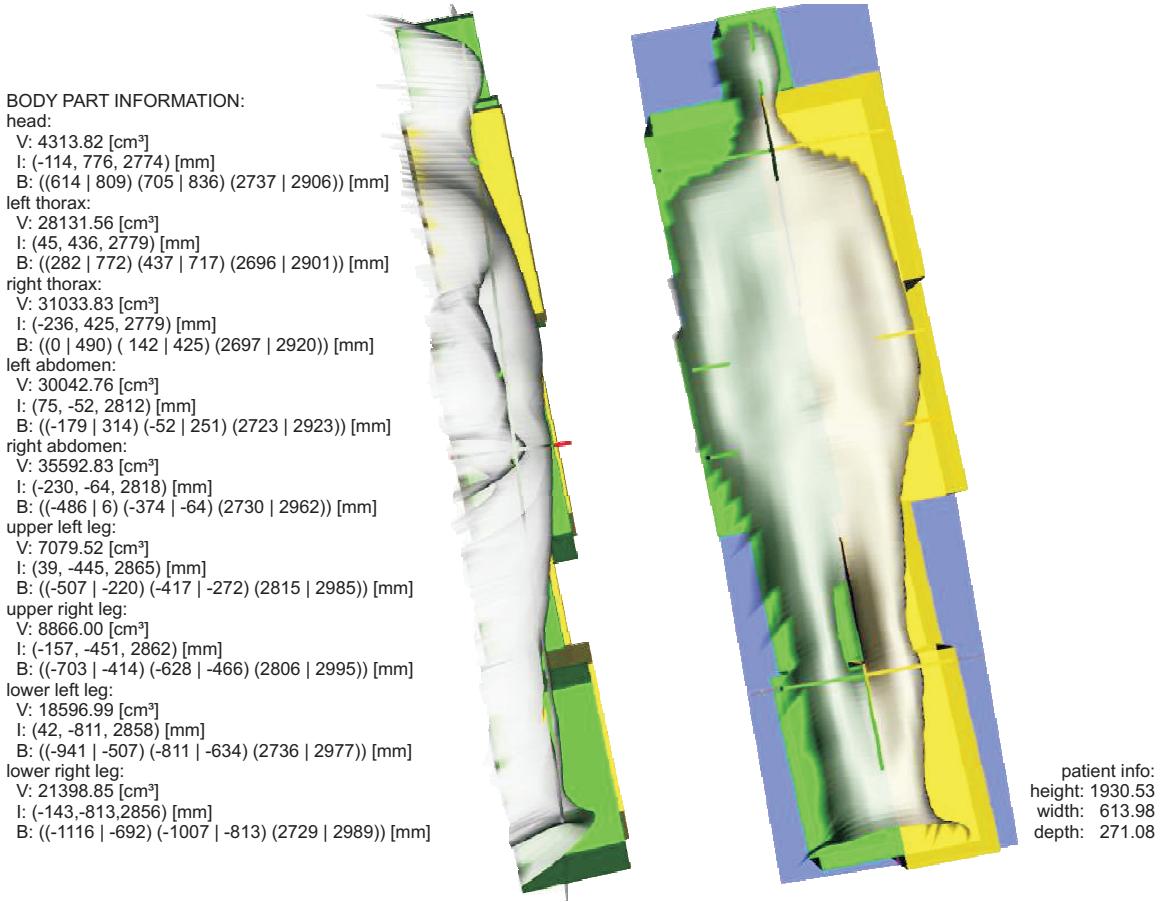


Figure 5.6: Full body acquisition using a ToF sensor including bounding boxes (side view and top view). On the left side information about the different bounding boxes is shown. V shows the volume of the bounding box in cm^3 , I the isocenter of the bounding box in mm and B the bounds of the actual bounding box in mm. On the lower right side, some basic information about the dimensions of the patient is displayed.

Before we are able to position a target, the before described table coordinate system calibration has to be applied. An alternative for C-arm systems also gives the following fairly efficient calibration step.

The following has to be done only once for the whole system. A coin is placed on a box on the patient table (see Fig. 5.7). The isocenter I_{Carm} of the C-arm is manually aligned with the center of the coin. We use two fluoroscopic images, one from 0° and one from 90° to do this. After determining I_{Carm} , the corresponding 3-D coordinate (isocenter) of the center of the coin I_{ToF} in the ToF coordinate system has to be identified. These two points, I_{Carm} and I_{ToF} can be considered as the origins of each of the corresponding coordinate

systems. To position a bounding box isocenter, the 3-D coordinate of the isocenter is shifted to I_{ToF} .

Algorithm 2 shows the whole algorithm in pseudo-code.

Algorithm 2 Inverse patient positioning algorithm

- 1: **Input:**
 $\mathbf{P} = [\mathbf{p}_{i,j}], i \in \{0, 1, \dots, K - 1\}, j \in \{0, 1, \dots, L - 1\}$ {3-D points}
 \mathbf{T} {patient table plane}
 ρ_0, ρ_1, ρ_2 {principle axis}
 $\gamma_0, \gamma_1, \gamma_2$ {golden intersections}
 $\iota_0 \iota_1 \iota_2 \iota_3 \iota_4 \iota_5$ {bounding box borders}
 $\hat{\mathbf{P}}$ {segmented points}
 - 2: Apply bilateral / temporal filtering to \mathbf{P}
 - 3: $\hat{\mathbf{P}} \leftarrow$ Clip \mathbf{P} against \mathbf{T} .
 - 4: $\rho_0, \rho_1, \rho_2 \leftarrow$ Apply a PCA to $\hat{\mathbf{P}}$.
 - 5: $\gamma_0, \gamma_1, \gamma_2 \leftarrow$ Compute three golden intersections along ρ_0 .
 - 6: $\mathbf{H}_0 \leftarrow$ Compute a histogram along ρ_0 binning the silhouette of ρ_1 ranging from $\gamma_2 - d_0$ to $\gamma_2 + d_0$ (see Figure 5.5).
 - 7: $\mathbf{H}_1 \leftarrow$ Compute a histogram along ρ_0 binning the silhouette of ρ_2 ranging from $\mu - d_1$ to $\mu + d_1$ (see Figure 5.5).
 - 8: $\iota_4 \leftarrow \min(\mathbf{H}_0)$ {minimum value of H_0 }
 - 9: $\iota_5 \leftarrow \min(\mathbf{H}_1)$ {minimum value of H_1 }
 - 10: $\iota_3 \leftarrow \frac{\iota_4 + \iota_5}{2}$
 - 11: $\iota_1 \leftarrow \mu$
 - 12: Compute bounding boxes using $\iota_0, \iota_1, \iota_2, \iota_3, \iota_4, \iota_5$ as bounds.
 - 13: **Output:** multiple bounding boxes representing anatomical body regions.
-

5.3.4 Experiments and Results

For the evaluation we used a SR-3000 ToF sensor from MESA Imaging AG, Switzerland and a C-arm system of the Artis zee family from Siemens AG, Healthcare Sector, Germany. The evaluation was performed on a standard Pentium Dual Core 2GHz processor. We rigidly mounted the ToF sensor on the ceiling above the patient table so that the patient can be fully visible. The ToF sensor has a resolution of 176x144 pixels with a field of view of 47.5° and 39.6° respectively.

We first examined the reproducibility of the isocenter computation. Without loss of generality, this was done for the head isocenter. For the other isocenters this could be done in a similar manner. The head isocenter of three people was computed 70 times per person. We took the mean x - y - z - coordinate as our reference isocenter. We again computed 140 head isocenters for each of these persons while staying in the same position. The mean squared errors between those measurements and the reference in x - y - z - direction (corresponding to ρ_0, ρ_1, ρ_2) are: $x: 1.73 \pm 1.11$ mm, $y: 1.87 \pm 1.31$ mm, $z: 2.91 \pm 2.62$ mm. Furthermore we computed the inter-subject standard deviation on these datasets: $x: 3.16$ mm $y: 2.85$ mm $z: 4.42$ mm.

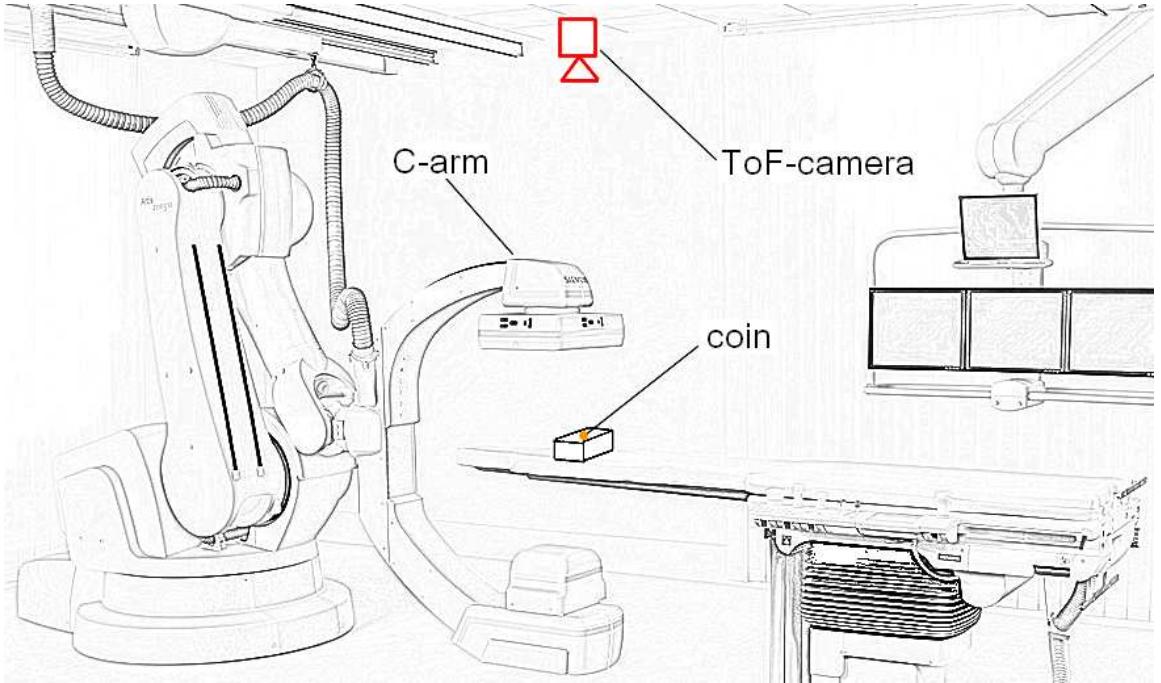


Figure 5.7: Coin calibration: A coin is placed on a box on the patient table. The coin is aligned in the isocenter of the C-arm. Afterwards, the corresponding 3-D coordinate of the coin position can be determined in the ToF point cloud. This calibration is a fast and fairly efficient alternative to the more complex calibration presented in the beginning of this chapter.

To evaluate our approach within a clinical setup, a body phantom was placed in a C-arm environment. We were again interested in the isocenter of the phantom head. The phantom was shifted in all three space dimensions multiple times. After computing the isocenter I_h of the head bounding box with respect to the previously computed I_{ToF} we could directly compute the translation to align I_h with I_{Carm} . A gold standard was defined by manually aligning the isocenter of the head using the same technique as for the previously described coin calibration (see Fig. 5.8). This is also the way it is done in hospitals today. We compared the table position of the gold standard alignment and the automatic alignment. In all cases we were able to automatically position the head in the x and the y direction within the accuracy of the patient table. According to the manufacturer the accuracy is 1 cm. In the z-direction, we had an error of ± 1 cm (see Fig. 5.8) because the z-coordinate of the isocenter heavily depends on the quality of the table segmentation. On a standard CPU (2.0 GHz dual-core) our algorithm has an execution time of 65 ms. Including data acquisition, preprocessing and segmentation, the overall execution time is about 143 ms.

5.3.5 Summary

Automation of time consuming steps is a key competency of future clinical procedures and workflow aware hospitals. We have shown that our system for fully automatic patient setup is able to identify anatomical regions for image acquisition in real-time. This information can be used to align the isocenter of anatomical targets with the isocenter of a

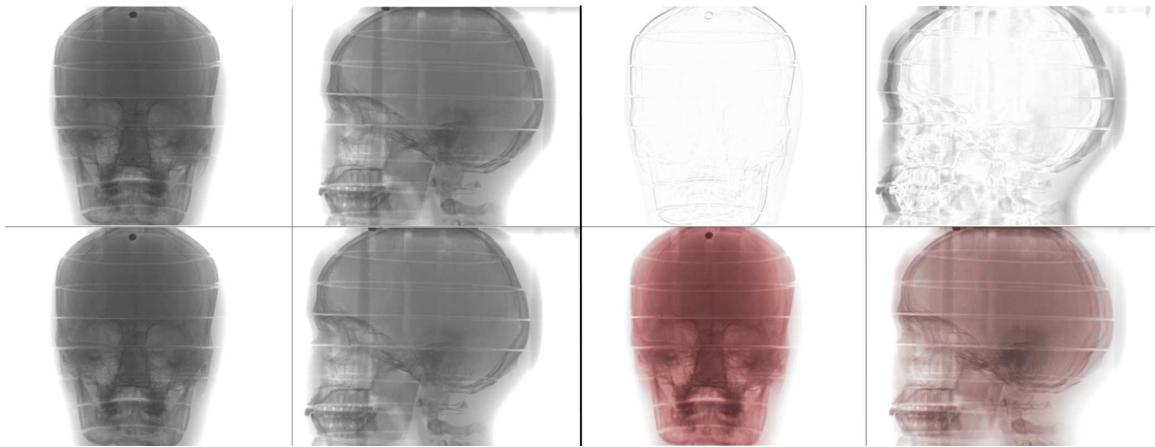


Figure 5.8: Phantom evaluation: Two upper left images: 0° and 90° projections of manually aligned head (gold standard). Two lower left images: 0° and 90° projections of automatically aligned head using the proposed method. Two upper right images: 0° and 90° projection difference images of manually and automatically aligned head. Two lower right images: 0° and 90° projections of manually and automatically aligned head (automatically aligned projections are superimposed in red).

C-arm system fully automatically. For this task, a 3-D surface acquired by a ToF sensor is analyzed. A two-stage classification process is introduced to identify body parts and to compute the corresponding isocenters. The extracted information can be directly used to control any image acquisition devices. This results in an *inverse positioning*, where the image acquisition device can be positioned according to the patient. Furthermore, several patient dependent metric information is provided by the system. This information can be used for e.g., initial positioning, collision detection, dose estimation or respiratory motion correction (see Chapter 6).

5.4 Patient Positioning for Radiotherapy

5.4.1 Introduction

Recent advances in computer technology and medical imaging devices open new chances for medical applications such as interventional procedures and therapy. Today, various imaging technologies are used to plan and execute such procedures, enabling new therapeutic possibilities. In order to benefit from the complementary information from these new imaging modes, a registration of either mono-modal or multi-modal image data is often required.

To cope with this problem, different solutions are available, e.g. image-based or surface-based registration [Gier 08b]. In radiotherapy, image-based registration plays a key role in the treatment planning as well as in the delivery phase. The introduction of Image Guided RadioTherapy (IGRT) has gained great attention, promising to result in a more effective and safer radiotherapy treatment [Olse 07]. On the other hand, due to the use of x-Ray based imaging for correct patient positioning and monitoring, additional dose is

applied to the patient. Simultaneously, care has also been focused on the optimization of the radiotherapy workflow so as to decrease time slots allocated to treat patients, potentially increase the throughput of the radiotherapy device and reduce the burden of the treatments. These two drivers – the need for IGRT and workflow optimization – have motivated the investigation of non-invasive surface based positioning systems to:

- properly align and monitor patients in a quick way;
- reduce the imaging dose in all cases where only surface-based positioning or monitoring is considered sufficiently.

Most of the available surface-based positioning systems are based on complex technologies, like laser range scanners [Cash 03] or active stereo camera systems [Scho 07b]. This leads to high system costs which may prevent the large scale use of these solutions. To tackle this drawback, the purpose of this section is to introduce a surface-based registration framework using a Time-of-Flight (ToF) camera.

5.4.2 Related Work

There are many different systems using different attempts to ensure the patient can be treated appropriately. Starting with manual systems offering worst precision, continuing with very accurate image-based systems and finally surface based systems are presented.

Laser Alignment

Today most commonly used and wide spread positioning systems in clinical workflows are based on wall mounted laser systems. These kind of systems rely on skin tattoos which are manually aligned with laser pointers. The laser pointers are projected on the patient skin and the treatment table has to be manually aligned. As these systems are installed in almost every hospital, people are quite familiar in dealing with such a system as it is well established. Also the costs for these systems are relatively low. The major drawback of a system based on laser pointers is the lack of accuracy. The system is very inaccurate and a mean positioning error of 4-8 mm [Gier 08a] is quite common. Another disadvantage is the setup time to manually align each patient. A third inconvenience factor is that the skin tattoos have to remain the complete treatment duration of several days to some weeks on the patients' skin.

Fiducial Markers

Other quite common systems are based on external fiducial or index markers. Such systems are using about 5-10 reflective markers as surface landmarks. One or more cameras are used to detect and track these markers. These markers can also be used to track respiratory motion. An example for such a system is the NOVALIS system from BrainLAB, Germany¹. One major drawback of this method is the setup procedure. First of all it is very time consuming and second it is also vulnerable to setup errors, e.g. markers can move, etc.

¹<http://www.brainlab.com>



Figure 5.9: The BrainLAB NOVALIS system. An infrared camera is used to track reflective marker on the patients' skin. Source: <http://www.brainlab.com>.

Image-guided Radiotherapy

Using x-ray images to verify the position of the patient is also known as IGRT. There are various systems using one or more flat panel detectors. If more than one detector is used, a 3-D registration can be done by a CT scan. Usually a very good accuracy can be achieved by these systems as the registration relies on internal structures. But there is additional irradiation necessary to acquire the images used for registration. Clinical studies showed a target registration error of about 2 mm [Gier 08a]. For additional motion management these systems mostly rely on implanted fiducial marker, which is not always possible. The BrainLAB ExacTrac x-Ray 6D system (which is a part of the Novalis system) is an example for such a system. It is an IGRT system including a dual x-ray system that allows for automatic couch adjustment and an additional infrared tracking system. The major drawback of such systems is the additional dose for the patient and the eventually necessary surgery to implant the marker.

AlignRT

The AlignRT system consists of two camera pods. Each pod holds an active stereo vision setup, including two cameras: a speckle projector and a texture camera. Using these parts, the system can reconstruct the 3-D surface of the patient using triangulation. This reconstruction results in approximately 20.000 3-D points. Involuntary patient movement can be detected automatically. The major drawback of the system is the high purchase price of about 150,000 USD. Furthermore the calibration of the camera pods is very sensitive. It

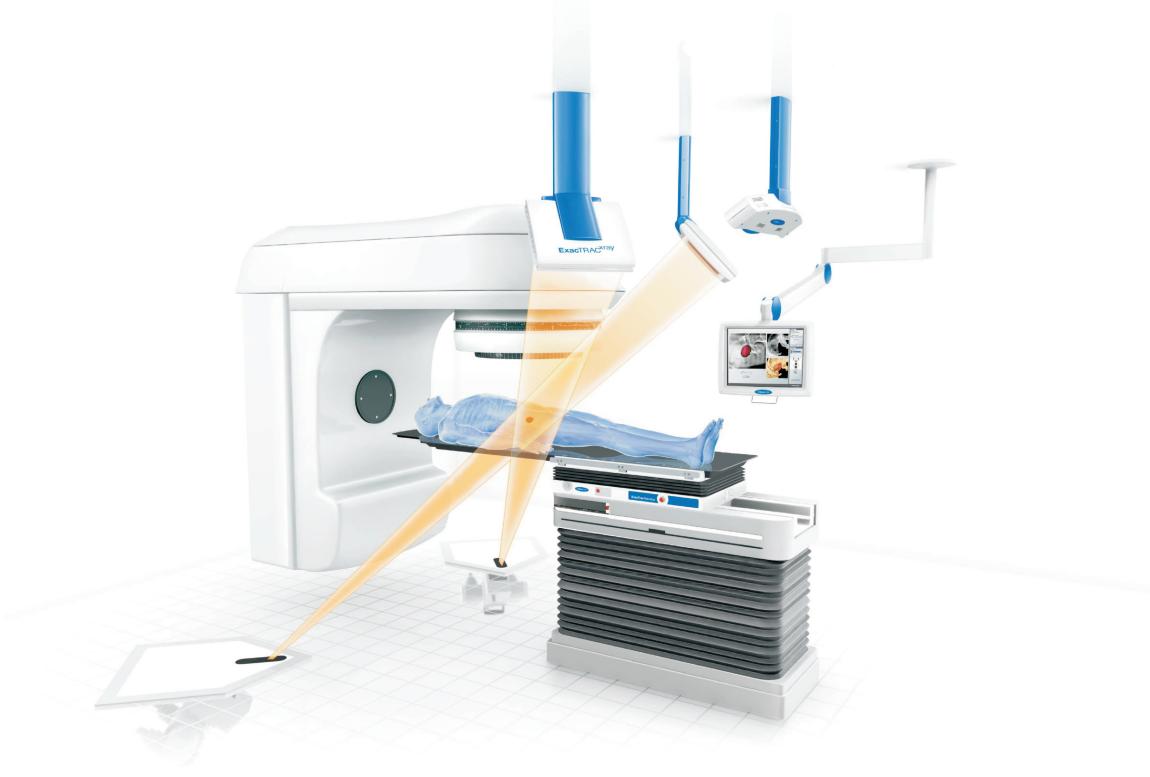


Figure 5.10: The BrainLAB ExacTrac x-Ray 6D system. Showing the x-ray sources and the corresponding flat panel detectors as well as the LINAC and the patient lying on the couch. Source: <http://www.brainlab.com>.

has to be calibrated on a daily basis [Scho 07a]. Figure 5.11 shows the Align RT system. AlignRT is distributed by VisionRT Ltd., London, UK².

Discussion

Surface based solutions are usually quite expensive because special hardware, like lasers or very accurate stereo vision cameras are needed. Furthermore these systems have to be calibrated on a dialy and weekly base. To address these issues we present a method to determine patient translation and rotation in 3-D by the acquisition of the patient's body surface during the planning CT and again right before or during each treatment session with the help of a ToF camera. By using a ToF camera, off-the-shelf hardware is used to solve the task of patient positioning. Therefore a solution for a possible system can be realized quite cost effectivly.

Table 5.1 shows a comparison between the introduced methods.

²<http://www.visionrt.com>



Figure 5.11: VisionRT System. The system uses an active illumination to project a speckle pattern on the patient. Using this pattern, the 3-D surface can be reconstructed. Based on this reconstruction, the system can compute a respiratory signal. Source: <http://www.visionrt.com>.

5.4.3 Materials and Methods

We basically start with similar requirements as we already defined in the last section. Figure 5.12 shows a schematic overview of the registration pipeline.

Again, the ToF sensor is rigidly mounted above the patient table. It is also possible that only parts of the patient are visible this time. In order to register a reference dataset with a moving dataset, we denote \mathbf{P} the $K \times L$ 3-D points of interest of the moving dataset acquired by a ToF sensor. \mathbf{Q} the $K \times L$ denotes the 3-D points of interest of the reference dataset acquired by a ToF sensor:

5 Time-of-Flight based Patient Positioning

Method	Accuracy	Time	Convenience	Dose	Costs
Wall laser	○	○	○	●	●
External Fiducial	●	○	○	●	○
IGRT	●	○	○	○	○
AlignRT	○	●	○	●	○
Time-of-Flight	●	●	●	●	○

Table 5.1: Overview of advantages and disadvantages of different patient positioning systems. (Full circle represents the best. In the case of IGRT one has to count the additional process of surgery as well.)

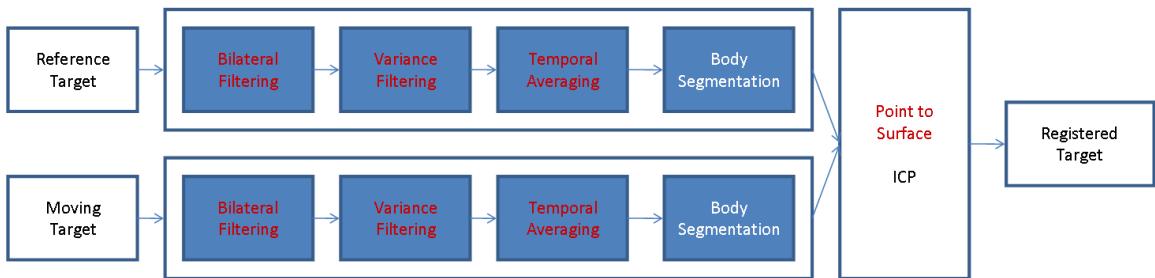


Figure 5.12: Processing pipeline: Each target, reference, and moving target is processed the same way. After applying a bilateral filter [Toma 98, Pari 09], a variance filter is applied. This filter is necessary to eliminate so-called "flying pixel". Before the 3-D body points are segmented from the background, a temporal filtering is applied as well. Using a point-to-plane ICP both targets are registered in the end. Note: The "gap" in the phantom is due to its mechanical parts.

$$\mathbf{P} = [\mathbf{p}_{i,j}], i \in \{0, 1, \dots, K-1\}, j \in \{0, 1, \dots, L-1\} \quad (5.20)$$

$$\mathbf{Q} = [\mathbf{q}_{i,j}], i \in \{0, 1, \dots, K-1\}, j \in \{0, 1, \dots, L-1\} \quad (5.21)$$

After the body segmentation (see last section) \mathbf{P} and \mathbf{Q} only contain 3-D points corresponding to the patient. Using these two point sets a surface registration of both can be computed. However, for registering ToF data in a robust manner, several issues have to be considered:

- **Noise:** ToF data suffers from high noise. Therefore the data has to be filtered. This can be done by e.g. using a Gaussian filter. However, in order to preserve important surface information, which can be used by the registration algorithm, we use bilateral filtering [Toma 98, Pari 09]. Furthermore, a temporal filter is applied, which averages the data over a certain amount of frames.
- **Flying pixels:** Another prominent problem regarding surface registration are so-called flying pixels. These occur whenever a distance discontinuity is given between two objects in the field of view of the ToF camera. As these pixels are often "jumping" between two distances, they can have a huge effect on the registration result. We

face this problem by applying a variance filter. For each pixel, we compute the variance based on the last frames. By applying a threshold, high variances (the "flying pixels") can be eliminated.

- **Grid effect:** Due to the ToF sensor matrix, the computed 3-D coordinates exactly match a grid structure in x and y direction. A standard ICP algorithm using an Euclidean point-to-point distance measurement prone to suffer from a "snap-to-grid" effect. For a distance of 1.0 m the grid spacing is already about 3.5 mm (CamCube). Instead of a point-to-point metric, we use a point-to-plane metric.

Figure 5.13 shows the original ToF data on the left side. After applying the just mentioned pre-processing filters, the data is improved significantly, as the right side of the figure shows. Within the next section an optimized version of the ICP algorithm is introduced.

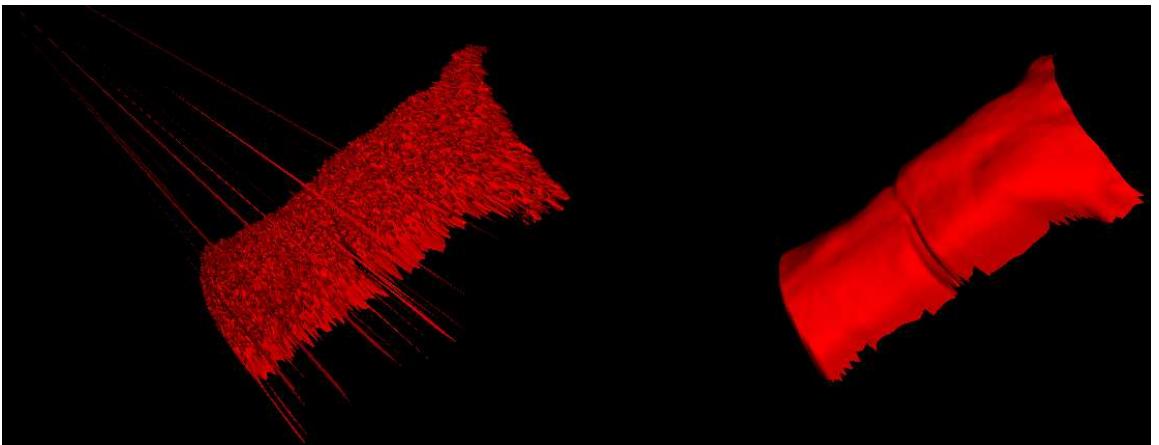


Figure 5.13: Comparison of preprocessing steps: Original ToF data on the left side (with flying pixels), pre-processed data on the right side

Iterative Closest Point Algorithm

Iterative Closest Point (ICP) algorithms can be considered as a family of algorithms. They vary in the choice of a specific distance measurement and an adequate optimization method. In general, ICP algorithms can be formulated as a minimization problem:

$$\sum_i \sum_j \|R\mathbf{p}_{i,j} + \mathbf{t} - \mathbf{q}_{i,j}\|_2^2 \rightarrow \min_{R,t} \quad (5.22)$$

Given two surfaces P and Q , the algorithm has to find the rigid transformation (R, t) ($t \in \mathbb{R}^3$) which aligns both surfaces to each other as close as possible. This minimal distance is expressed in some surface-to-surface distance $d(RP + t, Q)$. [Chen 92] introduced the algorithm about 20 years ago. It still can be considered as state-of-the-art for registering two point clouds where the matching is unknown. Algorithm 3 shows the classical version of the algorithm.

Looking at line 6-8 in Algorithm 3 one recognizes that a good registration result depends on the initial alignment of the two datasets. If the datasets are too far away from each other, the algorithm most likely will converge into a local optimum. For the specific task of

Algorithm 3 Classical ICP algorithm used e.g. in the VTK (www.vtk.org) toolkit. A quaternion based estimator is used.

```

1: Input:
    Two sets of 3-D points  $X$  and  $Y$ , where  $\mathbf{x}_i \in X, \mathbf{y}_j \in Y$  and  $i \in \{0, 1, \dots, |X|\}, j \in \{0, 1, \dots, |Y|\}$ 
2:  $c \leftarrow$  Array of size  $|X|$ 
3:  $\mathbf{R}_{\text{out}} \leftarrow \text{identity}(3)$  {rotation matrix}
4:  $\mathbf{t}_{\text{out}} \leftarrow \text{zeros}(3)$  {translation vector}
5: repeat
6:   for  $i = 1..|X|$  do
7:      $c(i) \leftarrow \underset{j}{\operatorname{argmin}} \|\mathbf{x}_i - \mathbf{y}_j\|_2^2$ 
8:   end for
9:    $(\mathbf{R}, \mathbf{t}) = \underset{(\mathbf{R}, \mathbf{t})}{\operatorname{argmin}} \sum_i \|(\mathbf{R}\mathbf{x}_i + \mathbf{t}) - \mathbf{y}_{c(i)}\|_2^2$ 
10:  for  $i = 1..|X|$  do
11:     $\mathbf{x}_i \leftarrow \mathbf{R}\mathbf{x}_i + \mathbf{t}$ 
12:  end for
13:   $\mathbf{R}_{\text{out}} \leftarrow \mathbf{R}\mathbf{R}_{\text{out}}$ 
14:   $\mathbf{t}_{\text{out}} \leftarrow \mathbf{R}\mathbf{t}_{\text{out}} + \mathbf{t}$ 
15: until  $\mathbf{R}_{\text{out}}$  and  $\mathbf{t}_{\text{out}}$  are stable
16: Output: Transformation  $(\mathbf{R}_{\text{out}}, \mathbf{t}_{\text{out}})$  and transformed point set  $X$ 
```

patient registration using ToF cameras this problem can be solved easily. Due to the depth information of the ToF camera, the patient can be segmented easily. The remaining 3-D points in both point clouds P and Q only belong to the patient. Computing the centroid of both point clouds and matching the centroid of Q to the centroid of P gives a good initialization in most cases. However, this is not valid if e.g. Q covers only a small subset of P .

To cope the problem of registering ToF datasets, we applied a lot of modifications to the original ICP algorithm. Those modifications affect both the runtime and the accuracy of the algorithm. [Rusi 01] subdivided the algorithm into five major parts. According to this subdivision, we will modify the algorithm to match the special demands for patient positioning:

- **Selection of points:** The lateral resolution of ToF cameras depends on the object to camera distance. Therefore, we choose to take all available points to maximize the surface information.
- **Point matching:** To guarantee a fast runtime, we precompute closest points on a dense voxel grid with a voxel spacing of Δv . The corresponding pseudo-code is given in algorithm 4 and algorithm 5. During runtime, this method has a complexity of $\mathcal{O}(1)$ as there is only one memory lookup required. We set Δv to 1 mm, which leads to a memory requirement of 288 MB for a region of interest of 600 mm \times 400 mm \times 300 mm.

- **Weighting of pairs:** To avoid discontinuities in the iterative optimization process a linear weighting w_{cor} for two corresponding points \mathbf{p} and \mathbf{q} is used depending on the point distance:

$$w_{\text{cor}} = \max \left(1 - \frac{\|\mathbf{p} - \mathbf{q}\|_2}{d_{\text{cor}}^{\max}}, 0 \right) \quad (5.23)$$

A value of about 3 mm is taken for d_{cor}^{\max} .

- **Rejecting pairs:** Again, this part is important to avoid discontinuities in the optimization process. d_{cor}^{\max} is already serving as an absolute threshold for the elimination of bad correspondences. Furthermore, a rejection of pairs containing mesh boundary points is added to avoid false correspondences in not completely overlapping surfaces.
- **Error metric and minimization:** We compute the sum of squared differences (SSD) from each source point to the tangent plane around the corresponding point. This metric is called point-to-plane metric (see Figure 5.14). Using this error metric, less iterations are required as the metric allows two surfaces to slide over each other. It furthermore avoids the snap-to-grid effect. This effect can be observed when a point-to-point error metric is chosen. The accuracy of the whole system is then limited to the spacing of the pixels in x and y direction. In general, there is no closed form solution available for the point-to-plane optimization. However, if the residual rotation angles are small, the problem can be linearized ($\sin \theta = \theta$, $\cos \theta = 1$).

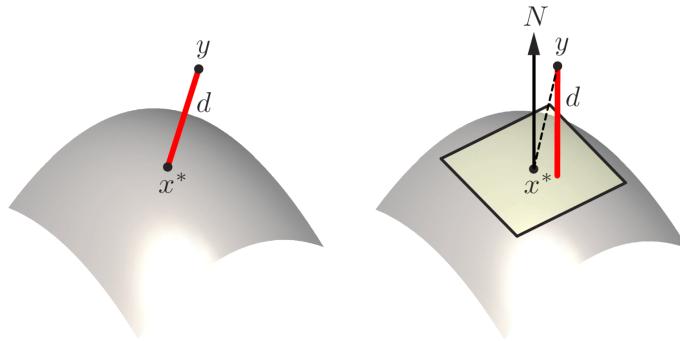


Figure 5.14: Comparison of distance metrics (Source: [Bron 08]).

The following equation describes the point-to-plane metric SSD optimizer used for the computation of a rigid body transform:

$$\begin{aligned} (\mathbf{R}, \mathbf{t}) &= \underset{(\mathbf{R}, \mathbf{t})}{\operatorname{argmin}} \sum_{\mathbf{b}_j \in B} w_{\text{cor}}^j \left((\mathbf{R}\mathbf{b}_j + \mathbf{t} - \mathbf{a}_{c(j)})^T \mathbf{n}_{c(j)}^{(A)} \right)^2 \\ &= \underset{(\mathbf{R}, \mathbf{t})}{\operatorname{argmin}} \sum_{\mathbf{b}_j \in B} w_{\text{cor}}^j \end{aligned} \quad (5.24)$$

$$\left((\mathbf{R}\mathbf{b}_j + \mathbf{t})^T \mathbf{Q}_{c(j)} (\mathbf{R}\mathbf{b}_j + \mathbf{t}) + (\mathbf{r}_{c(j)})^T (\mathbf{R}\mathbf{b}_j + \mathbf{t}) \right) \quad (5.25)$$

- $\mathbf{n}_{c(j)}^{(A)}$ represents the normal in point $\mathbf{a}_{c(j)}$ of the static dataset A. The normals are approximated with an area weighted averaging over all (maximum 8) adjacent triangles.

Algorithm 4 Pre-computation of closest points on a voxel grid

```

1: Input:
    Target dataset  $A$ , voxel spacing  $\Delta v$ , tolerance margin  $d_{\text{cor}}^{\max}$ 
2:  $x_{\min}, x_{\max}, y_{\min}, y_{\max}, z_{\min}, z_{\max} \leftarrow$  Bounding box of  $A$  extended by tolerance margin
    $d_{\text{cor}}^{\max}$  in all three dimensions
3:  $n_x \leftarrow \lceil \frac{x_{\max}-x_{\min}}{\Delta v} \rceil; n_y \leftarrow \lceil \frac{y_{\max}-y_{\min}}{\Delta v} \rceil; n_z \leftarrow \lceil \frac{z_{\max}-z_{\min}}{\Delta v} \rceil$ 
4:  $V \leftarrow$  integer array of dimension  $n_z \times n_y \times n_x$ 
5:  $K \leftarrow$  kd-tree for all points  $\mathbf{a}_m$  in dataset  $A$ 
   {kd-tree (short for k-dimensional tree) is a space-partitioning data structure. In this
   case, a  $k = 3$  dimensional space is used. This enables a fast method to search corre-
   sponding points.}
6: for  $i = 0..(n_z - 1)$  do
7:    $z \leftarrow z_{\min} + i\Delta v$ 
8:   for  $j = 0..(n_y - 1)$  do
9:      $y \leftarrow y_{\min} + j\Delta v$ 
10:    for  $k = 0..(n_x - 1)$  do
11:       $x \leftarrow x_{\min} + k\Delta v$ 
12:       $m \leftarrow \underset{m}{\operatorname{argmin}} \left\| (x, y, z)^T - \mathbf{a}_m \right\|_2$  {Use kd-tree  $K$ }
13:       $V(i, j, k) \leftarrow m$ 
14:    end for
15:  end for
16: end for
17: Output: Index array  $V$ 

```

Algorithm 5 Determination of a closest point during live mode

```

1: Input:
    Lower bounds  $x_{\min}, y_{\min}, z_{\min}$ ; Index array  $V$ ; Landmark point  $\mathbf{b}_l$  of dynamic dataset
     $B$ ; voxel spacing  $\Delta v$ 
2:  $i \leftarrow \text{round} \left( \frac{b_l^{(3)} - z_{\min}}{\Delta v} \right)$ 
3:  $i \leftarrow \min(\max(i, 0), n_z - 1)$  {adapts to array boundaries if required}
4:  $j \leftarrow \text{round} \left( \frac{b_l^{(2)} - y_{\min}}{\Delta v} \right)$ 
5:  $j \leftarrow \min(\max(j, 0), n_y - 1)$ 
6:  $k \leftarrow \text{round} \left( \frac{b_l^{(1)} - x_{\min}}{\Delta v} \right)$ 
7:  $k \leftarrow \min(\max(k, 0), n_x - 1)$ 
8:  $m \leftarrow V(i, j, k)$ 
9: Output: Index  $m$  of closest point in dataset  $A$  for vertice  $\mathbf{b}_l$ 

```

- $w_{\text{cor}}^{(j)}$ is the weighting factor for the point pair $(\mathbf{a}_{c(j)}, \mathbf{b}_j)$ according to Equation (5.23).
- $\mathbf{Q}_{c(j)}$ and $\mathbf{r}_{c(j)}$ are variables defined as following:

$$\mathbf{Q}_{c(j)} = \mathbf{n}_{c(j)}^{(\mathbf{A})} \left(\mathbf{n}_{c(j)}^{(\mathbf{A})} \right)^T \quad (5.26)$$

$$\mathbf{r}_{c(j)} = -2\mathbf{Q}_{c(j)}\mathbf{a}_{c(j)} \quad (5.27)$$

The objective function is solved by first applying a linearization to the trigonometric functions of the rotation angles ϕ_x , ϕ_y and ϕ_z in \mathbf{R} :

$$\mathbf{R} = \begin{pmatrix} 1 & \phi_z & -\phi_y \\ -\phi_z & 1 & \phi_x \\ \phi_y & -\phi_x & 1 \end{pmatrix} \quad (5.28)$$

Alternatively, the rotation matrix can also be linearized for a scenario with four degrees of freedom (DoF) by setting ϕ_x and ϕ_y to zero. The objective function of Equation (5.25) is optimized by computing the partial derivatives with respect to the searched parameters and setting them to zero. Please note, that this is only valid for small rotations:

$$\underbrace{\sum_{\mathbf{b}_j \in B} \mathbf{M}_j}_{\mathbf{M}} \begin{pmatrix} \phi_x \\ \phi_y \\ \phi_z \\ t_1 \\ t_2 \\ t_3 \end{pmatrix} \stackrel{!}{=} \underbrace{\sum_{\mathbf{b}_j \in B} \mathbf{c}_j}_{\mathbf{c}} \quad \text{or} \quad \mathbf{M}_{4\text{DoF}} \begin{pmatrix} \phi_z \\ t_1 \\ t_2 \\ t_3 \end{pmatrix} \stackrel{!}{=} \mathbf{c}_{4\text{DoF}} \quad (5.29)$$

\mathbf{M} is a symmetric, 6×6 measurement matrix containing information about the point set B and the normals of A . The 6-DoF measurement vector \mathbf{c} is additionally containing information about the vertices of A . Equation (5.29) is solved by computing the pseudo-inverse with singular value decomposition.

The customized ICP algorithm used for the patient positioning framework is detailedly described for the 6-DoF scenario in Algorithm 6. Since the transformation $(\mathbf{R}_{\text{out}}, \mathbf{t}_{\text{out}})$ is already computed in table coordinates, the parameters can be directly used to reposition the patient.

5.4.4 Experiments and Results

Body Phantom

To evaluate the proposed systems, we introduce a unique phantom to simulate human respiratory motion. The phantom can be used to evaluate surface based respiratory gating systems. These systems have special requirements on phantoms. The shape of the phantom should be human-like and also the motion of the phantom should correspond to human respiratory motion. Existing phantoms for example the Quasar phantom from Modus Medical Devices Inc., do not meet these requirements. They do not provide a human-like body geometry and -motion. These phantoms are also not able to simulate various breathing

Algorithm 6 Customized ICP algorithm for 6 DoF

```

1: Input:
    R and t as result of pre-registration or of the last frame, point sets A and B, Normals
     $n_k^{(A)}$  of A, Maximum number of iterations  $i_{\max}$ 
2:  $R_{\text{out}} \leftarrow R$ 
3:  $t_{\text{out}} \leftarrow t$ 
4: for  $i = 1..i_{\max}$  do
5:   for all  $b_j \in B$  do
6:      $b_j^{\text{trafo}} \leftarrow R_{\text{out}} b_j + t_{\text{out}}$ 
7:   end for
8:    $M \leftarrow \text{zeros}(6, 6)$ 
9:    $c \leftarrow \text{zeros}(6)$ 
10:  for all  $b_j^{\text{trafo}} \in B$  not lying on the mesh margin do
11:     $k \leftarrow \underset{k}{\operatorname{argmin}} \|b_j^{\text{trafo}} - a_k\|_2$  {Lookup in pre-computed voxel cube}
12:    if  $a_k$  lying on mesh margin then
13:      continue {Skip correspondences containing margin points}
14:    end if
15:     $w_{\text{cor}}^{(j)} \leftarrow$  Weight for point pair  $(b_j^{\text{trafo}}, a_k)$  according to Equation (5.23)
16:    if  $w_{\text{cor}}^{(j)} = 0$  then
17:      continue {Skip correspondences exceeding the distance threshold}
18:    end if
19:     $M_j \leftarrow$  Measurement matrix using  $b_j^{\text{trafo}}$ ,  $n_k^{(A)}$  and  $w_{\text{cor}}^{(j)}$ 
20:     $c_j \leftarrow$  Measurement vector using  $a_k$ ,  $b_j^{\text{trafo}}$ ,  $n_k^{(A)}$  and  $w_{\text{cor}}^{(j)}$ 
21:     $M \leftarrow M + M_j$ 
22:     $c \leftarrow c + c_j$ 
23:  end for
24:  solve Equation (5.29) for  $(\phi_x, \phi_y, \phi_z, t)^T$  using SVD
25:   $R \leftarrow$  Rotation matrix corresponding to yaw, pitch and roll (DIN 9300)
26:   $R_{\text{out}} \leftarrow RR_{\text{out}}$ 
27:   $t_{\text{out}} \leftarrow Rt_{\text{out}} + t$ 
28:  if  $\phi_x, \phi_y, \phi_z < 0.001^\circ$  and  $\|t\|_2 < 0.1$  mm then
29:    break {Transformation is stable.}
30:  end if
31: end for
32: Output: Transformation  $(R_{\text{out}}, t_{\text{out}})$  for table correction

```

amplitudes and velocities. Due to its human-like surface, the phantom can also be used to verify patient positioning systems.

A plaster cast of a male person is used as a basis for the phantom (see Figure 5.16). Mechanical support to simulate various human respiratory motion is added to the phantom chassis. To control the mechanical components an 8-bit AVR Atmega32 micro controller (μ C) with a clock frequency of 16MHz and 32kByte flash memory is used. It manages the communication with a PC via USB and controls the movement of the phantom's abdomen and thorax.

The abdominal motion is generated by a DC³-motor with a rotary encoder as control feedback. The thorax is lifted and lowered with three servomotors (Figure 5.15). The intra-abdominal motion of the organs in the abdominal area was simplified as a lift and lowering of the abdominal surface. The ribs are connected to the spine in the neck area, which results in a rotatory movement of the thorax. It is therefore simulated as a rotatory movement by the servo-motors.

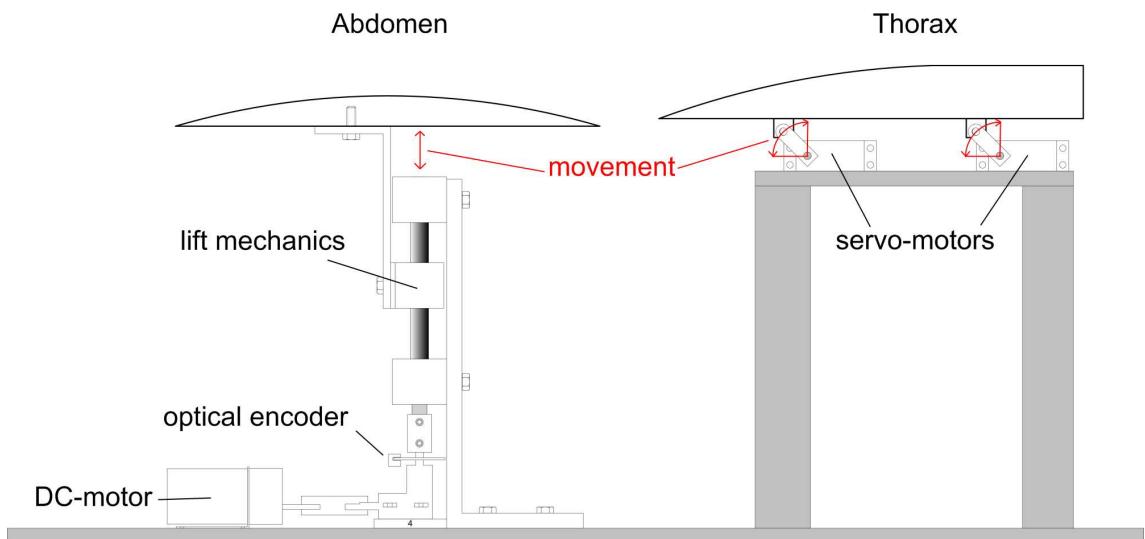


Figure 5.15: Schematic overview of the mechanical motion of the thorax and the abdomen (sideview).

The reference signal is a cosine-shaped signal which simulates human respiratory motion. The phantom can be parameterized using a software tool. Various frequencies and amplitudes for both thorax and abdomen can be set. It is also possible to simulate the connection of both respiratory systems using a couple factor. Using these components various respiratory motion movements can be simulated. The corresponding reference motion signals can be used as ground truth for correlation computations and evaluation. In general, the system is also capable to simulate irregular motion like coughing.

The introduced phantom will be used to evaluate the proposed patient positioning system and the respiratory motion gating system (see Chapter 6).

Figure 5.17 shows a schematic overview of the test setup. It is similar to the setup Schoeffel et al. [Scho 07a] are using. The movement of a Siemens ONCOR patient table

³direct current

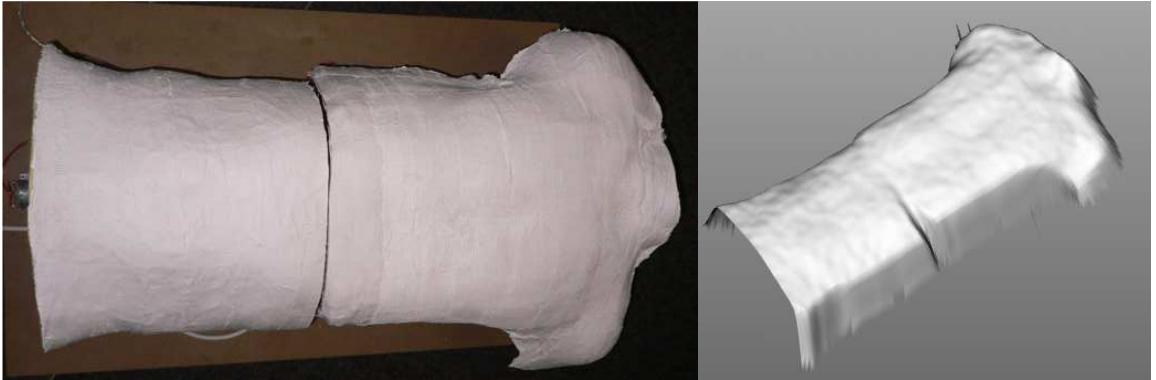


Figure 5.16: Plaster cast phantom used to simulate respiratory motion. Left: surface of the plaster cast. Right: ToF acquisition of the phantom.

is observed by three gauges - one for each axis. The gauges used for this evaluation had a total measurement range of 10 mm and an accuracy of $10 \mu\text{m}$. The ToF camera was perpendicularly aligned to the table in a distance of about 1.5 m which is a suitable distance for a practical setup. For the evaluation the patient table was moved arbitrarily in one direction with the other two directions locked. The evaluated relative table translations were in a range between 0.5 mm and 9.5 mm which is again similar to the experiments described in Schoeffel et al. [Scho 07a]. Rotations as well as translations in three dimensions were not considered within these experiments at the same time. The translation provided as output of the ICP algorithm was compared to the ground truth translation given by the gauges.

Table 5.2 shows the RMS error of the Euclidean distance (without any rotation) between ground truth and computed result. Table 5.3 shows the RMS error of the rotational error (including Euclidean distance error) between ground truth and computed result. The statistical indicators mean, median, and standard deviation are thereby determined upon a sequence of 50 consecutive frames of the non-moving phantom. Figure 5.18 illustrates the surface distance between the reference and the transformed source dataset.

5.4.5 Summary

We have developed a ToF based system for automatic surface registration in medical applications with a particular focus on radiotherapy. However, the framework is also suitable for other modalities, as it is purely surface based. We evaluated the presented framework on a typical patient positioning application. Overall an accuracy of $0.74 \text{ mm} \pm 0.37 \text{ mm}$ can be achieved at a distance of approximately 1.50 m. The proposed non-optimized algorithm is capable to process with 1 fps on a standard dual-core 2.0 GHz CPU.

5.5 Conclusion

It could be shown that ToF cameras are able to perform various positioning tasks in medical imaging. The achieved accuracy meets the required demands about 1 mm for patient positioning and about 1 cm for inverse C-arm positioning. Combined with image-based

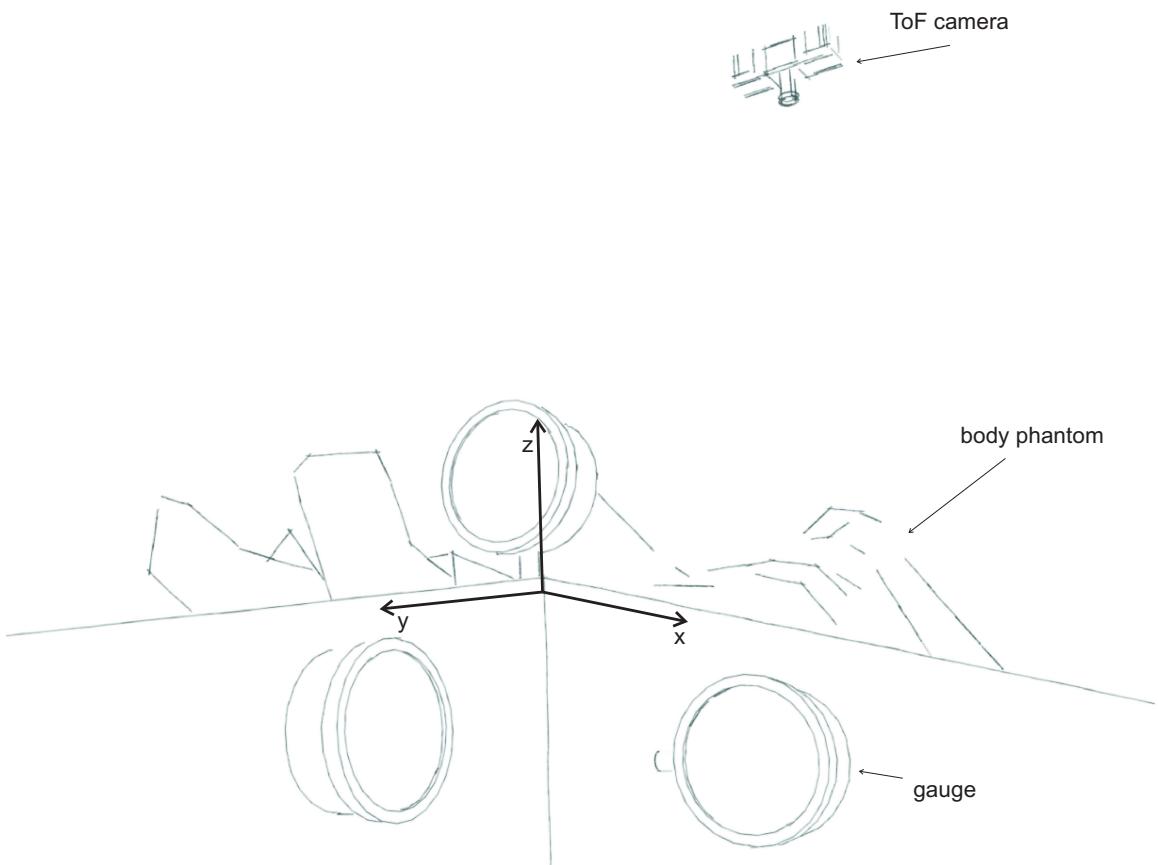


Figure 5.17: Schematic overview of the test setup. Three gauges with an accuracy of $\frac{1}{100}$ mm and a range of 10 mm are attached to the patient table. One gauge for each table direction.

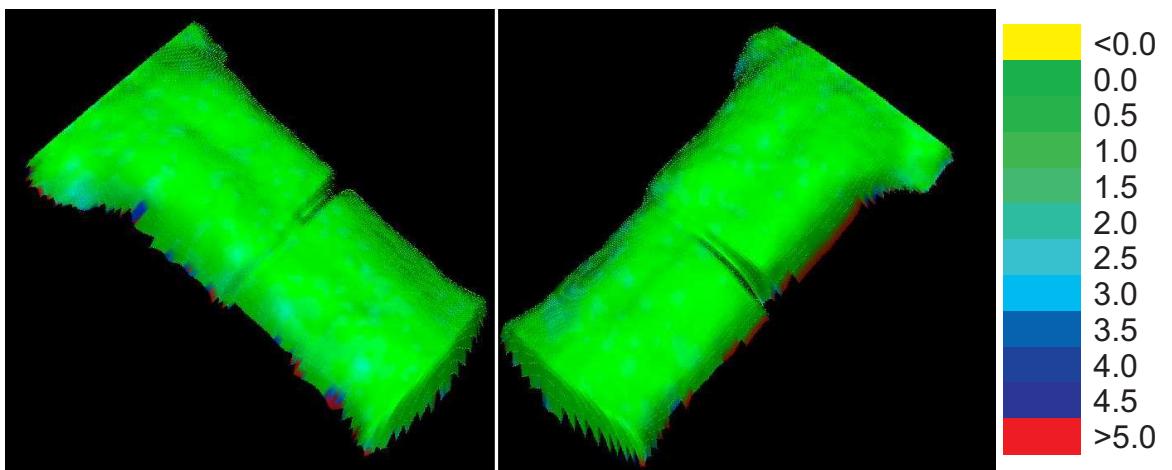


Figure 5.18: Exemplary registration result: Color coded distance between transformed source and target dataset (range: 0 mm - 5 mm).

system, the surface positioning results provided by a ToF-based system can be used to quickly align 3-D datasets.

5 Time-of-Flight based Patient Positioning

ground truth / [dir]	min⁽¹⁾	max⁽¹⁾	mean⁽¹⁾	median⁽¹⁾	std⁽¹⁾
0.50 mm/[x]	0.17	0.39	0.30	0.30	0.17
0.95 mm/[x]	0.28	0.56	0.42	0.40	0.24
2.50 mm/[x]	0.26	0.41	0.35	0.36	0.14
5.35 mm/[x]	0.20	0.63	0.47	0.49	0.32
6.45 mm/[x]	0.37	1.20	0.77	0.62	0.66
7.15 mm/[x]	0.30	0.67	0.51	0.51	0.30
9.50 mm/[x]	1.64	1.88	1.75	1.75	0.49
<hr/>					
0.50 mm/[y]	0.00	0.42	0.22	0.17	0.24
5.15 mm/[y]	1.03	1.43	1.26	1.20	0.56
7.15 mm/[y]	0.30	0.93	0.54	0.42	0.49
9.15 mm/[y]	0.81	1.46	1.20	1.17	0.66
<hr/>					
0.95 mm/[z]	0.10	0.55	0.39	0.42	0.32
2.95 mm/[z]	0.45	0.66	0.54	0.53	0.24
6.35 mm/[z]	1.08	1.24	1.16	1.16	0.32
7.60 mm/[z]	1.14	1.37	1.23	1.22	0.37
<hr/>					
average	0.54	0.92	0.74	0.71	0.37

Table 5.2: RMS of the Euclidean error of the computed translation with respect to the ground truth for lateral (x), longitudinal (y) and vertical (z) displacements. ⁽¹⁾ all measurements are in mm.

ground truth	min trans / rot⁽¹⁾	max trans / rot⁽¹⁾	mean trans / rot⁽¹⁾	median trans / rot⁽¹⁾	std trans / rot⁽¹⁾
-3°	0.65 / 0.11	0.78 / 0.14	0.72 / 0.13	0.72 / 0.13	0.03 / 0.00
-6°	1.36 / 0.15	1.63 / 0.13	1.47 / 0.14	1.47 / 0.14	0.06 / 0.00
-9°	1.73 / 0.13	2.06 / 0.17	1.96 / 0.15	2.00 / 0.15	0.09 / 0.01
-12°	3.05 / 0.33	3.31 / 0.35	3.18 / 0.35	3.18 / 0.35	0.07 / 0.00
<hr/>					
+3°	0.63 / 0.00	0.90 / 0.03	0.77 / 0.01	0.78 / 0.01	0.07 / 0.01
+6°	1.28 / 0.04	1.60 / 0.06	1.42 / 0.05	1.41 / 0.05	0.08 / 0.00
+9°	1.92 / 0.11	2.11 / 0.13	2.04 / 0.12	2.05 / 0.12	0.05 / 0.01
+12°	2.45 / 0.06	2.75 / 0.09	2.58 / 0.07	2.60 / 0.07	0.09 / 0.01
<hr/>					
average	1.63 / 0.12	1.89 / 0.13	1.76 / 0.13	1.64 / 0.13	0.07 / 0.01

Table 5.3: RMS of the Euclidean error of the computed rotations with respect to the ground truth. ⁽¹⁾ translation measurements are in mm, rotation measurements in °.

CHAPTER VI

Time-of-Flight based Respiratory Motion Gating

The chapter introduces a system to measure respiratory motion using a ToF camera. First, a state-of-the-art overview is given, followed by a method part. The chapter concludes with an evaluation of the system.

6.1 Introduction

There have been various approaches to account for respiratory motion in the past. The measured signal is mostly used for gating. We already introduced this term in Chapter 2. In general, there are two ways of gating, internal and external. Therefore, there are also two ways to detect respiratory motion. It can be measured internally or externally.

Internal gating requires the implantation of a fiducial marker near the tumor. This marker can be tracked using additional x-ray imaging, while radiotherapy is applied. External gating always requires a surrogate signal for the respiratory motion. Most of the external gating methods are based upon tracking a discrete point set e.g. with a stereo camera system. Manufacturers like BrainLAB and Varian are using infrared cameras to track markers on the chest of a patient. VisionRT provides a marker-less solution to account respiratory motion using an active stereo camera system. There are also systems utilizing sensors, like pressure sensors (e.g. AZ-733V, ANZAI Medical Co.) to measure respiratory motion.

Within this chapter, a system using Time-of-Flight cameras to account for respiratory motion gating is presented. The proposed system is able to detect independent respiratory signals for different anatomical regions of the body simultaneously and in real-time. This enables the system to deal with changing breathing patterns, e.g. the patient switches from breast breathing to abdominal breathing during an image acquisition or radiotherapy treatment. Hence it is also possible to determine the instantaneous type of breathing. This information can be used to adapt reconstruction algorithms or treatment protocols. Using a ToF camera to measure respiratory motion, no additional markers or complex calibration steps are necessary to acquire respiratory signals. Potential applications for the proposed method are manifold and can be found within 4-D CT- [Bied 07] and 4-D MRI [Sieb 07a]

reconstruction, PET- [Cook 07] and SPECT-motion compensation [Ue 06] and gated radiotherapy [Sepp 07, Iona 07].

In the next section an overview of state of the art technologies will be given. Afterwards, detailed instructions how to use Time-of-Flight cameras for respiratory motion gating are presented. Concluding, the evaluation of the proposed system with phantom and patient data is presented.

6.2 Related Work

Current commercially available technologies to account for respiratory motion in radiotherapy can be divided into two groups. The first group combines systems using external sensors for measuring respiration. Systems relying on optical technologies form the second group. Please note, that there is also the possibility to extract respiratory motion from image data using image processing. We will not go into detail for these methods within this work.

Regarding patient throughput and a more convenient therapy course for patients, systems belonging to the second group are more favorable. The main drawback of such systems is that some of them still need the attachment of manual markers. Furthermore, such systems are very expensive. Two representatives of each group will be introduced next.

6.2.1 Other Sensors

AZ-733V Anzai Belt

This commonly established system is using a sensor, either a load cell (pressure sensor) or a laser sensor, to monitor respiratory motion. Furthermore, it is possible to use the system in combination with several medical image acquisition devices such as common linear accelerators. The entire system consists of three different parts: elastic belt with a load cell, sensor port and a wave deck. In addition, a portable computer is included within the wave deck and provides a software application to monitor and display the patient's respiration. Radiation oncologists can define a respiratory gate and are able to use the information provided from the belt system for treatment. In a typical setup, the load cell is fixed on the patient's chest with the elastic belt (see Fig. 6.1). Connected to the sensor port, the load cell is able to measure the patients respiration and to pass it to the application. The AZ-733V is distributed by Anzai Medical (Anzai Medical, Tokyo, Japan¹).

The belt is well established and has a rather large installed base today. However, it always needs a time consuming setup period, which can also cause errors while attaching the belt to the patient. Another drawback of the system is that it is not possible to measure respiratory motion on different points. For instance, if the patient changes his respiration from abdominal breathing to chest breathing, the correlation between the measured signal and the movement within the body might change.

¹<http://www.anzai-med.co.jp>

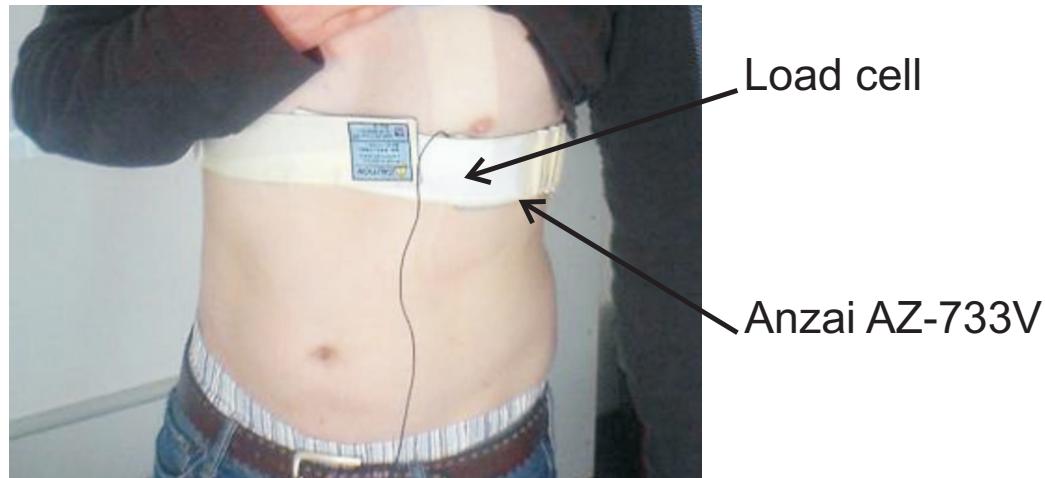


Figure 6.1: AZ-733V Anzai Belt: The belt holds a pressure sensor on the patients' body, which is used to measure the expansion and contraction of the chest.

CPX Spirometer

Another system utilizing sensors to measure respiratory motion is a spirometer. Spirometers can be used to measure expiration and inspiration of patients. One example for such a device is the preVent Pneumotach system, which measures the flow from and to the lungs. This flow is converted into a pressure signal. The acquired data is also processed by a computer which acts as an interface between the system and the attached device. The acquired data is sent to a computer for illustration and further proceedings [Zhan 03]. This method provides good results but still uses devices which bother the patient during the procedure [Hois 04]. In Figure 6.2 a typical test setup with a spirometer is illustrated. The system is very easy to use and no specific setup is required. The convenience for the patient is disputable, as the patient might have to wear the device for a long time during radiotherapy. The system is distributed by Medgraphics, St.Paul, USA².

6.2.2 Optical Sensors

Real-time Position Management System (RPM)

Usually an external gating system like the Varian Real-time Position Management (RPM) system is observing an external fiducial marker (see Fig. 6.3). Related to the RPM system, this marker is an infrared reflective plastic box placed on the patient's anterior abdominal surface to maximize the AP respiratory-induced motion. This surrogate is monitored by a digital video camera mounted on the treatment room wall. The system allows both displacement and phase gating. To secure the treatment the system is using a periodicity filter. This filter checks the regularity of the breathing waveform and disables the beam immediately when the breathing waveform becomes irregular. Such an irregular breathing waveform occurs if the patient is coughing or moving. The beam is enabled again after

²<http://www.medgraphics.net>



Figure 6.2: Test Person with the CPX spirometer and the preVent Pneumotach. Source: [Zhan 03].

establishing regular breathing again. The system is distributed by Varian Medical Systems Inc., California, USA³.

This also well-established system has been available for several years. In general it has the same disadvantage as the ANZAI system, as it can only measure respiratory motion at one distinct point on the body. Setup errors also can cause errors in the measurement.

GateRT

In Chapter 5 the AlignRT system from VisionRT is already introduced. Using the same technology, the company offers also GateCT / GateRT, which additionally offers markerless surface tracking for 4-D CT reconstruction. According to [Tart 06] the system can distinguish between thorax and abdominal breathing with a framerate of approx. 5 frames per second.

6.2.3 Discussion

Today most of the described systems are limited to focus on a single point or a limited area of the upper part of the body. This results in an acquisition of a 1-D respiratory signal. However, there is only one system (GateRT/GateCT, VisionRT) which is competitive to the system proposed in this chapter.

³<http://www.varian.com>

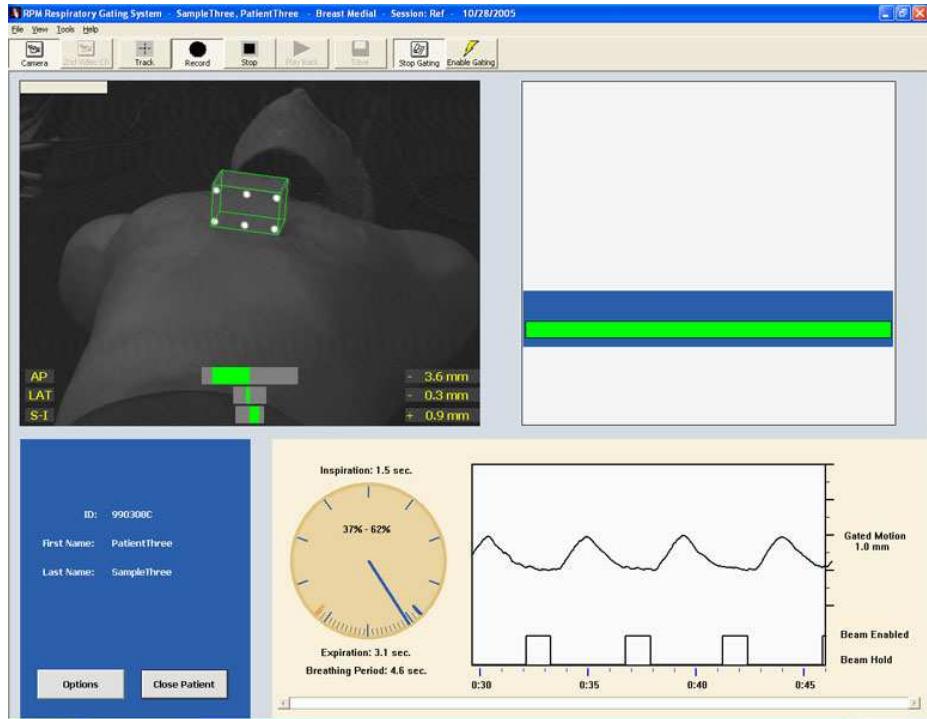


Figure 6.3: Screenshot of the software of the RPM system. The green box represents the detected box, which provides the respiratory curve. Based on this curve, the gating intervals can be set. Source: <http://www.varian.com>.

Table 6.1 shows a comparison between the introduced systems. For respiratory motion gating, each system has particular disadvantages. Only the ToF based system shows no significant disadvantages for respiratory motion gating. Therefore, this technology suites optimal for the given problem.

The overall process of human respiration is complex and dependent on many different factors. The human body is able to utilize both thorax muscles and the diaphragm for breathing. This yields to an almost infinite number of possible combinations of these two systems to keep the respiratory cycle alive. Therefore a multi-dimensional respiratory signal accounting for different breathing patterns could potentially improve solutions for problems dealing with respiratory motion, like motion artifacts or tumor motion due to respiration. A signal with only two dimensions (thorax and abdomen signal) e.g. can handle a switch between respiration techniques.

6.3 Materials and Methods

In this section we will give an overview of how a ToF sensor enables multi-dimensional respiratory tracking. Indices in the following equations are considered to be integer values. Furthermore we assume, the ToF camera is rigidly mounted above the patient table.

The set P includes the $K \times L$ 3-D points of interest acquired by a ToF camera, where $i \in \{0, 1, \dots, K-1\}$, $j \in \{0, 1, \dots, L-1\}$. The corresponding amplitude values are represented by the set A :

6 Time-of-Flight based Respiratory Motion Gating

System	Time	Complexity	Convenience	Usability	Costs
RPM	●	●	●	●	●
GateCT/GateRT	●	○	●	●	○
Anzai	●	●	●	●	●
CPX	●	●	○	●	●
Time-of-Flight	●	●	●	●	●

Table 6.1: Overview of advantages and disadvantages of different respiratory motion gating systems. (Full circle is best. In the case of complexity, a full circle means, that the system is less complex.)

$$\mathbf{P} = [\mathbf{p}_{i,j}], i \in \{0, 1, \dots, K - 1\}, j \in \{0, 1, \dots, L - 1\} \quad (6.1)$$

$$\mathbf{A} = [\mathbf{a}_{i,j}], i \in \{0, 1, \dots, K - 1\}, j \in \{0, 1, \dots, L - 1\} \quad (6.2)$$

To measure respiratory motion precisely, several steps are required (see Figure 6.4).

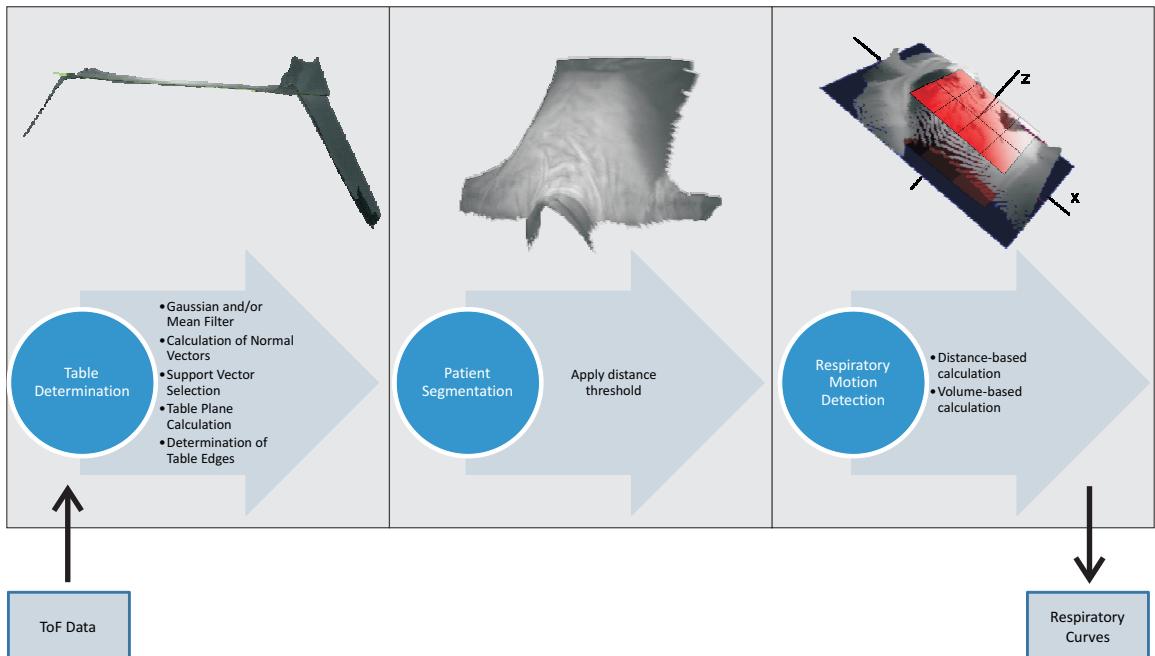


Figure 6.4: Schematic overview of ToF-based respiratory motion detection. The algorithm is divided into three parts. First an independent detection of the patient table plane is performed. This plane is then used to segment the patient from the background in 3-D. To compute a respiratory signal two methods are available. The first one computes the distance of a plane to the camera, whereas the second measures the volume between the table plane and the thorax and abdomen, respectively.

Chapter 5 introduces a treatment table calibration. For the proposed stand-alone system it is not necessary to know the relation of the table coordinate system and the camera coordinate system. This information is not required to measure respiratory motion. Therefore, we can take advantage of this and simplify the table detection. The following algorithm does not need any additional tools like a calibration pattern. Therefore it works fully automatically. The algorithm detects a plane representing the table where the patient is lying on. This plane then is used to segment the 3-D points belonging to the patient from the background of the acquired scene. In order to reduce noise it is beneficial to filter the ToF data by applying a mean filter with a kernel of 5 pixels before applying the following algorithms.

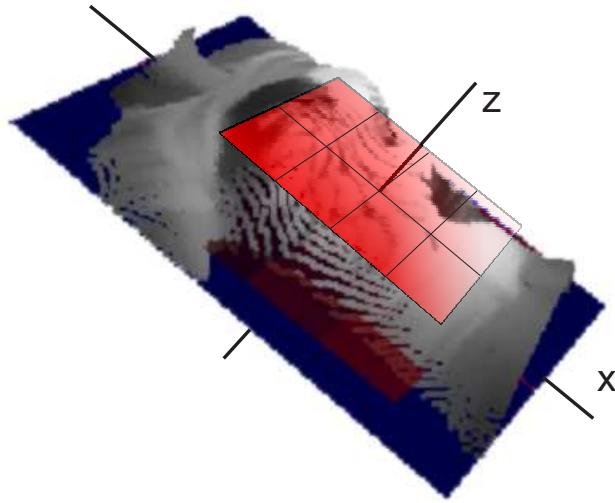


Figure 6.5: Segmented 3-D point cloud including two axis with superimposed regions of interest (red). The partitioning of \mathbf{I} into n parts uses the x-axis for the first division in two disjunct parts and equally subdivides the rest parallel to the y-axis.

6.3.1 Table Detection

The basic idea of detecting the treatment table is to fit a plane through all 3-D points belonging to the table. The following algorithm will describe a common method to fit distinct planes into a set of 3-D points. The most dominant plane of the algorithms' output is the table plane.

Unit Normal Vector Computation

Given \mathbf{P} we compute a set \mathbf{N} of unit normals using $\mathbf{h}_{i,j} = \mathbf{p}_{i+1,j} - \mathbf{p}_{i-1,j}$ and $\mathbf{v}_{i,j} = \mathbf{p}_{i,j+1} - \mathbf{p}_{i,j-1}$, where $i \in \{1, 2, \dots, K-2\}$, $j \in \{1, 2, \dots, L-2\}$:

$$\mathbf{N} = \{\mathbf{n}_{i,j} | \mathbf{n}_{i,j} = \frac{\mathbf{h}_{i,j} \times \mathbf{v}_{i,j}}{\|\mathbf{h}_{i,j} \times \mathbf{v}_{i,j}\|_2}\}. \quad (6.3)$$

Due to the normalization each normal vector $\mathbf{n}_{i,j}$ has only two degrees of freedom. The 3-D surface coordinates are given in a left-handed Euclidean coordinate system, whose

origin coincides with the optical center of the ToF camera and the camera looks into the direction of the z -axis. Therefore, the two degrees of freedom are the angle α_x between the x -coordinate and the z -axis and the angle α_y between the y -coordinate and the z -axis.

Histogram Computation

Using these two angles we compute a 2-D $K_g \times K_g$ histogram \mathbf{H} with a bin-size of δ_B degrees, where $g \in \{0, 1, \dots, \lceil \frac{2\pi}{\delta_B} \rceil\}$ and δ_B is the range size of similar normal vectors we want to group. The histogram is initialized with zeros. If the angles α_x and α_y associated with the unit normal vector are in the range of the bin, each bin $h_{k,m} \in \mathbf{H}$ is incremented by the associated intensity value $a_{i,j}$ of the input image. The amplitude intensity value acts as a weighting factor here. Thus, 3-D coordinates of points with similar normals are associated with each bin. Examples of such histograms are given in Figure 6.6 and 6.7.

Using a threshold t_H we can select corresponding cluster peaks out of the 2-D histogram (see below). After this step, l planes E_u with $u \in \{0, 1, \dots, l-1\}$ are identified and a certain number n_u of 3-D points are associated with each plane E_u , where $e_{u,v}$ represents the v -th vertex of the cluster E_u .

$$E_u = \{e_{u,v} | e_{u,v} \in \mathbf{R}^3\} \quad (6.4)$$

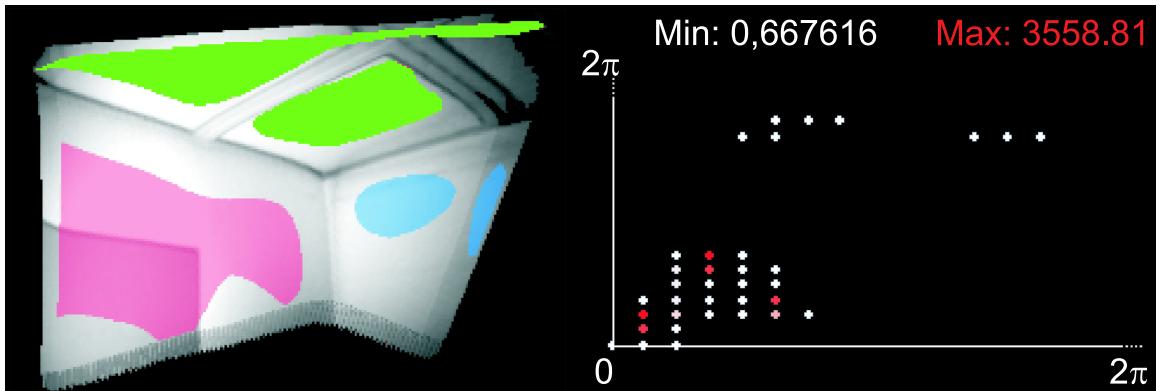


Figure 6.6: Example of the histogram computation: 3-D reconstruction of three pair-wise perpendicular walls and the significant part of the computed histogram representing the angles α_x and α_y of the surface normals. Bright red spots in the histogram indicate a high number of normals with the same orientation. The three planar walls yield three clearly observable peaks in the histogram. Note: Parts of the colored best fitting planes are hidden due to noise of the ToF signal and spatial averaging (near edge regions).

Determination of Best Fitting Planes

We are now able to use this 2-D histogram information to compute different 3-D boundaries of an object surface patch determined by similar unit normal vectors. In order to group these points, we average 3-D point coordinates of E_u associated with a histogram bin $h_{k,m}$, where $k, m \in \{0, 1, \dots, \lceil \frac{2\pi}{\delta_B} \rceil\}$. To get stable planes, we are interested in big surfaces patches, therefore only histogram values greater than t_H are averaged.

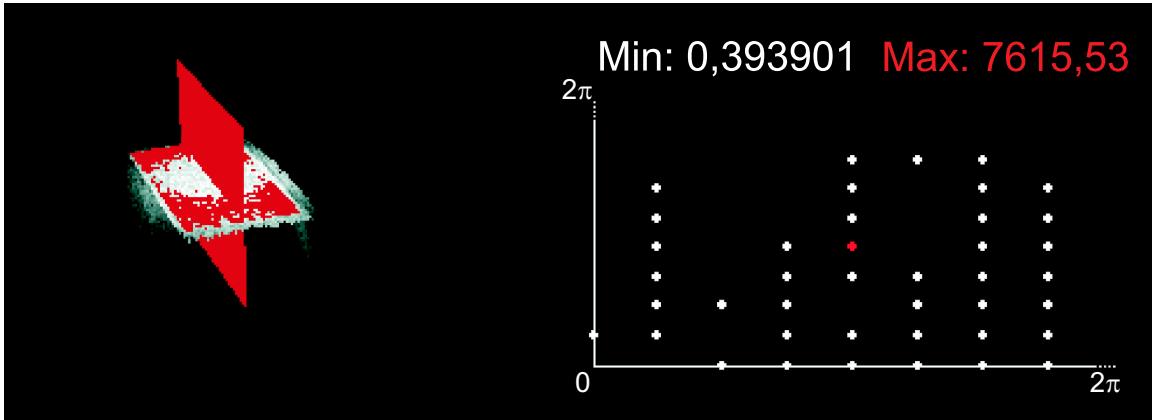


Figure 6.7: Example of the histogram computation: 3-D reconstruction of a table rear side (red) and the significant part of the computed histogram representing the angles of surface normals to x-axis and the y-axis. The red point in the histogram indicates the obviously dominant surface of the table rear side.

An arbitrary 2-D point is used as a seed point for a flood-fill algorithm in order to identify neighbouring points belonging to a certain plane. The points have to be connected in the ToF image matrix and have the same histogram bin. Such an algorithm is suitable for this type of application due to the relatively small lateral resolution of available ToF cameras. Therefore, only a small amount of 3-D points has to be processed by the flood-fill algorithm. By applying such an algorithm it is ensured that only a set of pixels next to each other in the sensor matrix are grouped into a plane which is not necessarily given if only 3-D points associated with a certain histogram bin are used.

The obtained 3-D point-set of connected 3-D points with similar normal vectors is denoted \mathbf{E}'_u . We use the normal vector \mathbf{n}_u determined by the corresponding histogram bin $h_{k,m}$ as a reference normal vector. Furthermore, the 3-D points of \mathbf{E}'_u which are grouped into a plane are used in computing a best fitting plane \hat{E}_u :

$$\hat{E}_u : a_u x + b_u y + c_u z + d = 0. \quad (6.5)$$

Each 3-D point $\mathbf{e}'_{u,v} \in \mathbf{E}'_u$ is then projected on \hat{E}_u along the normal vector \mathbf{n}'_u of the computed plane. The resulting points are denoted $\mathbf{e}_{u,v}^{(p)} \in \mathbb{R}^3$ and the set of these points is denoted $\mathbf{E}_u^{(p)}$.

For a typical setup to detect a table usually the most prominent cluster represents the table plane. Using the introduced algorithm it is possible to identify the patient table with the absence of a patient. Once the best fitting plane corresponding to the patient table is known, a patient lying on the table can be segmented. Algorithm 7 shows the algorithm in pseudo-code.

6.3.2 Respiratory Motion Detection

The just computed table plane is used to segment the patient from the background. The segmented points, which only belong to the patient are denoted $\hat{\mathbf{P}}$. We want to compute a certain region of interest $\mathbf{I} \subseteq \hat{\mathbf{P}}$ and partition this region into n sub-regions

Algorithm 7 Table detection algorithm

```

1: Input:
2:    $\mathbf{P} = [\mathbf{p}_{i,j}], i \in \{0, 1, \dots, K - 1\}, j \in \{0, 1, \dots, L - 1\}$  {3-D points}
3:    $\mathbf{A} = [\mathbf{a}_{i,j}], i \in \{0, 1, \dots, K - 1\}, j \in \{0, 1, \dots, L - 1\}$  {corresponding amplitudes}
4:    $t_H$  {threshold}
5:    $k$  {number of planes}
6:    $\mathbf{N} \leftarrow$  Array of size  $K \times L$  {normals}
7:    $\mathbf{H} \leftarrow$  Array of size  $K_h \times L_h, h \in \{0, 1, \dots, \lceil \frac{2\pi}{\delta_B} \rceil\}$ ,  $\delta_B \leftarrow$  range size {histogram}
8:   for  $i = 0..K - 1$  do
9:     for  $j = 0..L - 1$  do
10:     $\mathbf{N}_{i,j} = \text{ComputeNormal}(\mathbf{P}_{i,j})$ 
11:    Compute  $\alpha_x$  and  $\alpha_y$ 
12:    Increment histogram  $\mathbf{H}_{\alpha_x, \alpha_y}$  by  $A_{i,j}$ 
13:   end for
14: end for
15: select  $k$  peaks in  $\mathbf{H}$ 
16: fit  $k$  planes to corresponding points in  $\mathbf{P}$ 
17: Output:  $k$  planes

```

$\mathbf{I} = \{\mathbf{S}_1 \cup \mathbf{S}_2 \cup \dots \cup \mathbf{S}_n\}$ and $\mathbf{S}_q \cap \mathbf{S}_w = \emptyset$, where $q \in \{0, 1, \dots, n-1\}, w \in \{0, 1, \dots, n-1\}$. Using these n regions of interest we can compute anatomical adaptive respiratory signals of the patient's chest and abdomen (see Fig. 6.5). We assume that the coarse orientation of the patient is known and apply a Karhunen-Loëve-Transformation to all remaining 3-D points $\hat{\mathbf{P}}$. Resulting, the origin of the coordinate system is placed in the center of gravity of the point cloud $\hat{\mathbf{P}}$ and the axis are aligned with the axial, sagittal and coronal plane of the patient. By applying this transformation in advance an accurate and stable computation of the volume signal can be assured. For distance based respiratory motion detection this transformation is not required. Instead the algorithm presented in Section 5.3 can e.g. be used to automatically identify thorax and abdomen.

To actually measure the respiratory motion of the patient, two methods are presented. The first one calculates the distance between the chest and / or abdomen of the patient and the camera position. Whereas the second method computes the volume between the patient table and the abdomen or chest of the patient.

Distance based Respiratory Motion Detection

Regions of interest for the thorax and abdomen area can be computed either automatically (see Chapter 5) or selected by hand. For a coarse determination of the thorax and abdomen region the knowledge about the orientation of the patient can be used. Usually, only the upper body of the patient is visible to the camera. As an initial segmentation the point cloud of the upper body is divided into two equal parts, either manually or automatically. Experiments showed that dividing the upper body into two equal parts is valid for most datasets. This results in two regions e.g. $\mathbf{S}_{\text{thorax}}$ and $\mathbf{S}_{\text{abdomen}}$. For a finer partition these regions can also be subdivided. A best fitting plane is computed for each region. We can

now directly use the distance d_{thorax} and d_{abdomen} of each plane to get a respiratory signal for each plane.

By observing d over time for each plane we can derive multiple respiratory signals (see Fig. 6.11).

Volume based Respiratory Motion Detection

To calculate accurate volume signals for each sub-region S_k ($k \in \{0, 1, \dots, n - 1\}$) each point $p_{i,j} \in S_k$ has to be clipped against the 2-D border lines of S_k . This algorithm is necessary to compute the exact volume of each sub-region as the borders of each region intersect the 3-D points arbitrary. We are using the Cohen-Sutherland clipping algorithm to perform this task [Fole 08]. By using this approach the partition is always consistent and independent from the patient's position. To calculate the volume of S_k we take a triangle of three neighboring points $p_{i,j}, p_{i,j+1}, p_{i+1,j} \in S_k$ and the distance between the segmentation plane and the mean z -coordinate of these three points to calculate the partial volume. If one or more triangle points are not present in S_k points computed by the clipping algorithm are used. All particular triangle volumes are summed up to obtain the whole volume value for a specific subregion S_k . By plotting the volume values for each subregion S_k over time n respiratory graphs can be generated.

Determination between Thorax- and Abdominal Breathing

We now can use, either the computed distance or volume signals to distinguish between chest and abdominal breathing. Therefore the signal energy F_k of each subregion S_k is calculated over a certain timespan T , where $A_k(t)$ denotes the offset (mean) corrected volume values of subregion S_k .

$$F_k = \sum_{t=-T}^0 |A_k(t)|^2 \quad (6.6)$$

After computing the energy for every subregion an energy coefficient α is calculated. Therefore the signal values of the subregions assigned to the chest area F_{chest} and the signal energy of the subregions assigned to the abdominal area $F_{\text{abdominal}}$ are summed up. The relation α of the energies is achieved by dividing the abdominal energy by the overall computed energy F_{all} . For the classification an appropriate border λ has to be set:

$$\alpha = \frac{F_{\text{abdominal}}}{F_{\text{all}}} \quad (6.7)$$

$$\text{Abdominal Respiration : } \alpha > \lambda \mid \text{Chest Respiration : } \alpha \leq \lambda \quad (6.8)$$

Experiments showed, that test persons are predominantly breathing with their abdomen. Therefore λ has to be chosen at about 0.7 for a good distinction between both respirations.

6.4 Experiments and Results

6.4.1 Phantom Experiments

We used a CamCube ToF camera from PMD Technologies GmbH and observed the movements of the phantom described in Section 5.4.4 within a distance of about 1 m. The test setup is illustrated in Figure 6.8. The ToF camera provides a 3-D representation of the shape and motion of the phantom. For the evaluation we used the just described methods, where two independent respiratory signals for thorax and abdomen can be acquired.

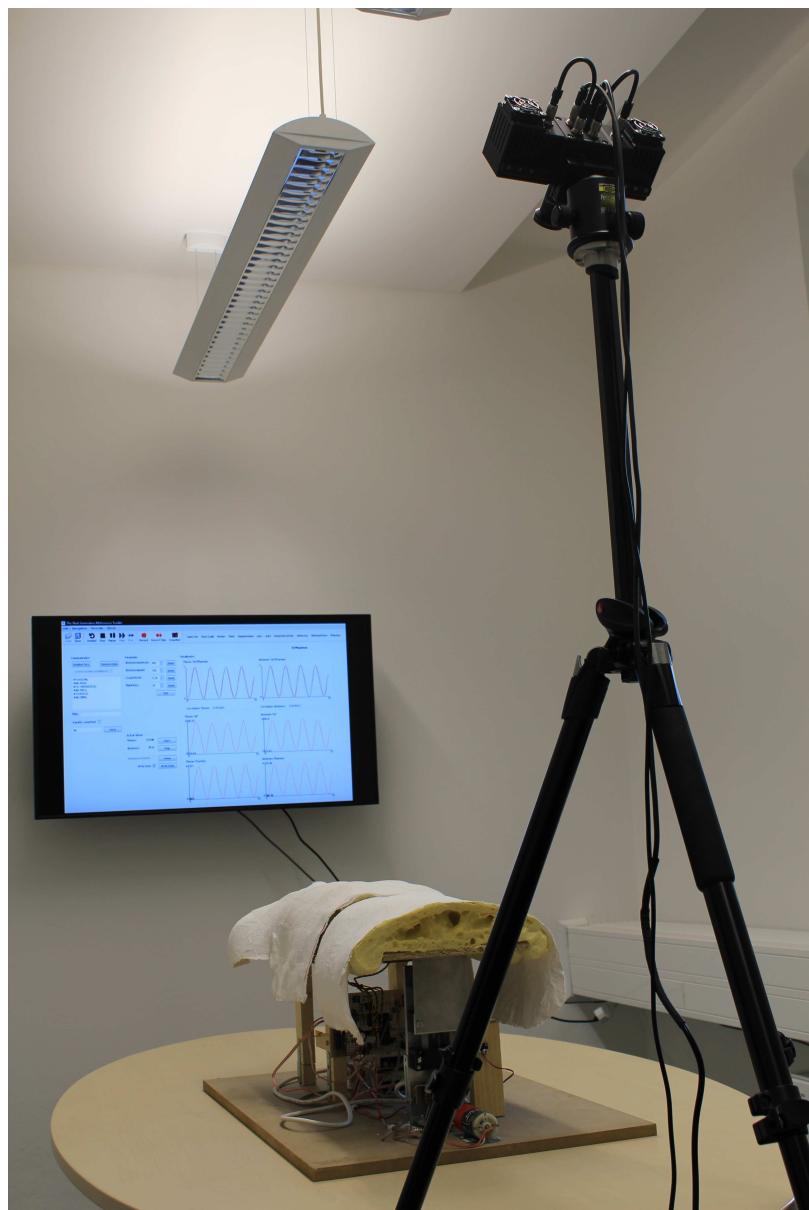


Figure 6.8: Overview of the test setup for the evaluation of the respiratory gating system. The system can be used to evaluate surface-based motion detection systems.

To investigate different respiratory signals the phantom breathing frequency was varied between $3\text{--}25 \text{ min}^{-1}$, the phantom abdomen and thorax amplitude from 1.3–18 mm. The resting respiratory frequency of an adult is $12\text{--}15 \text{ min}^{-1}$, during stress, e.g. during radiotherapy the respiratory frequency can increase. On the opposite a sedated patient can have a reduced respiratory activity.

After a gain-offset correction of the measured and reference signal the correlation factor according to Pearson was computed. For synchronization purposes, the ToF signal was shifted two timestamps to match the phantom signal. This phase shift is the result of the non-deterministic thread handling in the current operating system. To acquire the two respiratory signals the USB and the serial port of the PC have to be polled simultaneously.

Several measurements were performed. The correlation factors of the ToF- and phantom signal varied from 0.36 up to 0.99 depending on amplitude and frequency. Figure 6.4.1 illustrates a respiratory signal with a high correlation, whereas Figure 6.4.1 shows a lower correlation.

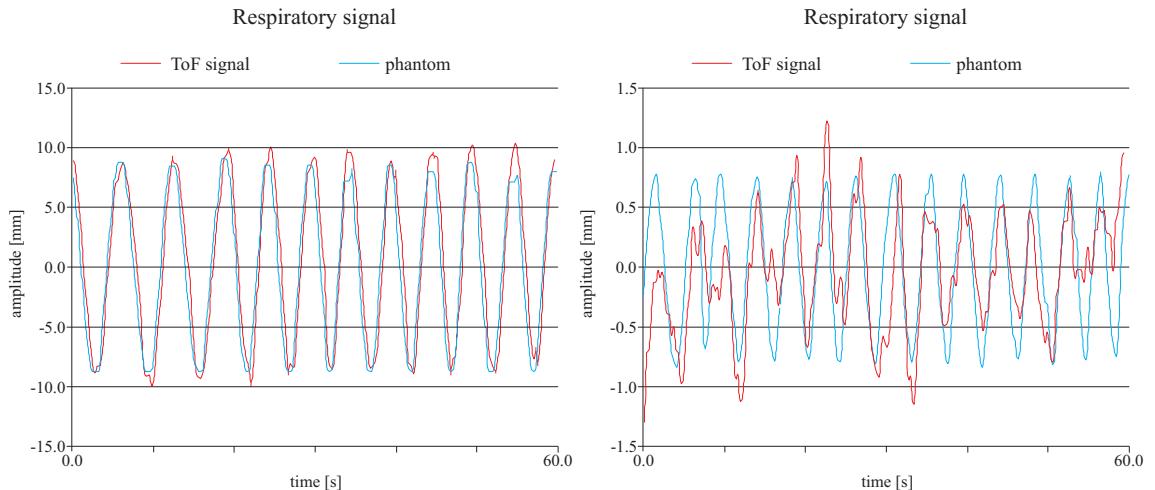


Figure 6.9: Evaluation of the thorax movement in ToF camera direction. Left: Respiratory frequency of 11 min^{-1} , amplitude of 17.8 mm, correlation factor of 0.94. Right: Respiratory frequency of 14 min^{-1} , amplitude of 1.5 mm, correlation factor of 0.65.

Smaller correlation values, having small peak-to-peak amplitudes, can be explained by the noise of the ToF signal in z -direction (figure 6.4.1). No preprocessing to smooth the ToF data was applied. Therefore, the z -component is noisy. It can be reduced with a mean filter, which can increase the correlation and therewith the signal equality of the real respiratory motion and the measured motion. An adequate correlation without filtering can be achieved with a peak-to-peak amplitude of 5 mm and above.

The second parameter of respiratory motion is the frequency. Figure 6.4.1 shows that the correlation factor can be considered independent of the respiratory frequency in the range of $3\text{--}25 \text{ min}^{-1}$. Concluding, the ToF camera can acquire a stable respiratory signal independent from the respiratory frequency with at least 25 breathings per minute.

Figure 6.4.1 shows the limitations of respiratory gating with a Time-of-Flight camera. Here a breathing amplitude of 1.5 mm peak-to-peak with a respiratory frequency of 14 min^{-1} was simulated. Thereby a correlation factor of 0.65 was achieved.

6 Time-of-Flight based Respiratory Motion Gating

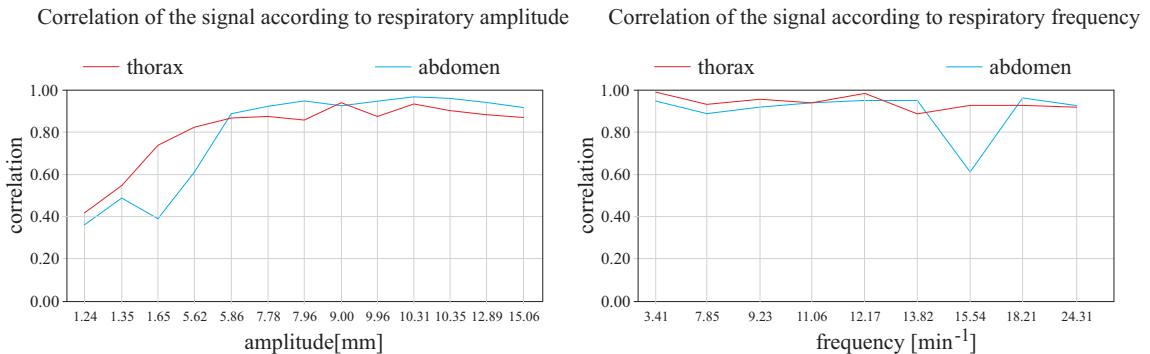


Figure 6.10: Correlation factor of the ToF signal and the phantom signal, Left: amplitude, right: frequency

The ToF system returned a stable respiratory frequency and a well fitting amplitude signal. With suitable signal post-processing and filtering of the raw data the signal similarity and thereby the correlation might be furthermore increased.

6.4.2 Patient Experiments

To furthermore evaluate the proposed methods we measured the systematic error of the system by computing the root mean square RMS error of the signal while looking at a rigid phantom. We observe a table RMS of about 0.1 mm. We also compared the ToF based signal in both thorax and abdomen with a respiratory signal acquired by the external gating system AZ-733V, ANZAI Medical Co. Therefore, we acquired both signals synchronized and computed the correlation coefficient. Based on 13 different persons, we computed mean, median, and variance of the computed correlation coefficients for two different setups (see Table 6.2). We placed the ANZAI belt at the abdomen of the test person for the first test scenario. The test person was required to do abdominal breathing first and breast breathing afterwards. In the second setup, we modified the position of the ANZAI belt and put it near the breast of the test person, which was also required to perform both breathing types successively. As we acquire two signals, it is also possible to measure the ratio between abdominal breathing and breast breathing by comparing the amplitude of both signals for the last 20 seconds (see Table 6.2).

6.5 Conclusion

We could show, that the deployment of a ToF camera system within respiratory tracking applications is comparable to existing systems. Furthermore, using a ToF sensor it is possible to acquire multiple respiratory signals for different anatomical regions in real-time. As an example we showed the simultaneous acquisition of two respiratory signals, one for abdominal breathing and one for breast breathing. This example also showed, that it is not sufficient to measure respiratory motion at only one point on the body. If the patient changes his respiration, the correlation drops down. Using the proposed system, a change in respiration can be detected and the measuring area can be switched easily. In general, the generic method is also feasible to partition the body into more regions of interest, if

			Average	Median	Variance	Ratio ⁽¹⁾
	Breathing: ANZAI: TOF:	Abdomen Abdomen Abdomen	0.91	0.94	0.01	79%
	Breathing: ANZAI: TOF:	Thorax Thorax Thorax	0.85	0.86	0.01	53%
	Breathing: ANZAI: TOF:	Thorax Abdomen Thorax	0.62	0.76	0.17	54%
	Breathing: ANZAI: TOF:	Abdomen Thorax Abdomen	0.67	0.75	0.06	79%

Table 6.2: Experimental results showing average, median, variance of the correlation ANZAI to ToF signal for 13 test persons; ⁽¹⁾ average abdominal breathing percentage

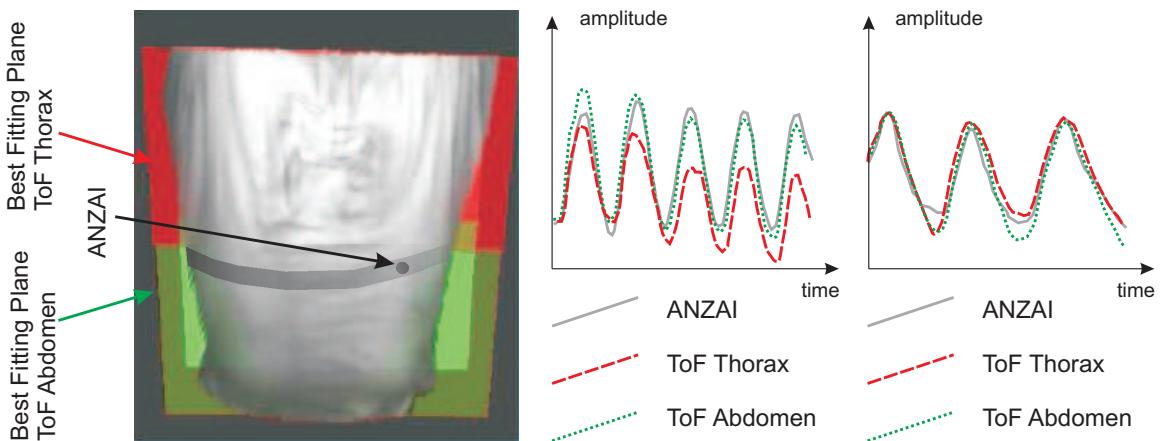


Figure 6.11: On the left: 3-D model of the automatically segmented upper part of the body of a patient. The colored planes are best fitting planes used to determine two independent respiratory signals (red: chest; green: abdomen). On the right: Derived 1-D signals. Vertical axis: time in seconds. Horizontal axis: motion amplitude of the planes representing the respiratory motion during breathing of chest (red) and abdomen (green) as well as the ANZAI belt signal (gray).

necessary. The computation time on a standard single core computer (Pentium M, 2GHz) is below 50 ms per frame, therefore we are able to sample respiratory signals with a frequency of about 20 Hz.

CHAPTER VII

Summary and Outlook

7.1 Summary

Within this work, a common motion management framework for radiotherapy is introduced. The whole framework consists of three independent systems. The first system targets gating and introduces a method to combine the actual internal movement of tumors with an external surrogate respiratory signal. The second system provides two independent surface-based methods to position either a C-arm like device with respect to the patient or the patient with respect to a reference surface. Using the last system it is possible to detect respiratory motion for gating applications without using marker or the need of contact to the patient.

The first major contributions of this work is an on-line verification of internal tumor movement with an external respiratory signal. Markov processes, a well established method in pattern recognition are used to verify the correlation of internal tumor motion and an external respiratory signal. The method enables a more accurate treatment of tumor tissue and gives more control over the overall process. The second contribution is the application of a novel technology to medical image processing. Originated in the automotive- and consumer electronics industry, a novel technology called Time-of-Flight cameras is introduced for medical problems within this work. The third contribution is the improvements of this novel technologies to meet the demands of the given problem. An optical and a distance calibration for ToF camera is presented in this work. Applying both calibrations, the ToF technology is feasible for solve tasks like patient positioning, where an accuracy of about 1 mm is mandatory. Therefore, a ToF camera can be seamlessly integrated in a radiotherapy workflow to position patients prior the treatment. The fourth contribution is the introduction of a novel system to measure respiratory motion. The system enables contact- and markerless measurement of a multidimensional respiratory signal. The sum of the above methods leads towards a more robust application of gated radiotherapy. The provided methods enable a more reliable way to control the overall process of radiotherapy.

The core technology of the framework is a novel camera sensor called Time-of-Flight camera. Nowadays these cameras are about to push several markets to a new dimension. Originally driven by the automotive industry, several systems for the consumer market have arisen in 2010 and will arise within the next years. The implementation of this technology in a medical environment within this thesis is among the first works in this field. However, there is a great potential in this technology and other research groups are starting to work with ToF cameras recently.

Clinical background and a general motivation for this work was given in Chapter 2. The chapter gives an overview of a radiotherapy workflow and gives clinical definitions of target volumes to irradiate tumors. It also holds an introduction to gated radiotherapy and points out on several problems within radiotherapy.

Within Chapter 3 an algorithm to combine the advantages of internal and external gating was presented. The method is an image-based mapping / synchronization procedure that automatically labels pre-treatment fluoroscopic images with corresponding 4-D CT phases. It is designed as an optimization process and finds the optimal mapping between both sequences by maximizing the image similarity between the corresponding pairs while preserving a temporal coherency. This mapping can be used by a external surrogate signal to trigger the gates for gated radiotherapy. Unlike only using an external respiratory motion detection system, without correlating the signal with true internal movement this method produces more reliable results. It was both evaluated at synthetic and patient data and an average of 93% correctly labeled frames could be achieved. For various sequences differing in speed and amplitude of the respiratory signal, in average the irradiated tumor volume was 64%, whereas only 30% of the beam targets healthy tissue.

Chapter 4 introduced the underlying technology of ToF cameras. Furthermore methods for an optical and a distance calibration of these cameras is presented. This calibration is mandatory to obtain the required accuracy for tasks like patient positioning. A mechanical phantom which is capable to simulate respiratory motion for thorax and abdomen is presented in this chapter as well.

Chapter 5 introduced two independent systems to solve positioning tasks. The first system is able to position a C-arm like device with respect to the patient. Therefore, a ToF camera acquires the whole body of the patient and segments it into meaningful anatomical regions, like head, thorax, abdomen, legs. The system computes 3-D bounding boxes of the anatomical regions and computes the iso-center of the boxes. Using this information, the C-arm system can automatically position itself and perform a scan. The system was evaluated using a body phantom and an accuracy within the patient table accuracy of 1 cm could be shown. A second system deals with surface-based positioning of a patient with respect to a priorly acquired surface of the same patient. Such systems are necessary, e.g. in radiotherapy or multi-modal imaging. The method uses an ICP algorithm, tailored to ToF cameras. We evaluated the system using the body phantom introduced in Chapter 4 and obtained an overall accuracy of $0.74 \text{ mm} \pm 0.37 \text{ mm}$ for translations in all three room directions within 10 mm.

In Chapter 6 a novel external system to measure respiratory motion was introduced. The system uses a ToF camera to measure the motion of thorax and abdomen of the patient independently. It is purely vision-based and does not need any additional information, like marker or other sensors. The system can be used as a stand-alone system. It is capable to

automatically detect the patient table and to segment the patient from the background. The user can manually select or automatically compute several regions on the upper part of the patient's body, where he want to measure respiratory motion. Due to the high framerate of about 20 Hz, several signals for different positions can be measured simultaneously. We evaluated the system using a novel body phantom introduced in Chapter 4. The tests showed, that the system signal and the ground truth signal of the phantom have a reliable correlation of more than 80% for amplitudes greater 5 mm. The correlation of both systems is independent (always more than 80%) of the respiratory frequency. Furthermore, the measured signals were compared with a well-established external gating system, the Anzai belt. These experiments were performed on human persons. We could show a correlation of about 88% of our system and the Anzai system.

7.2 Outlook

Time-of-Flight cameras provide a novel technology to further improve medical engineering. In the future, there will be more applications than the just introduced to benefit from this technology in various fields of medicine. Currently, more and more universities and research labs are about to work with ToF cameras. There is a great potential e.g. in 3-D endoscopy [Holl 10]. Once, the cameras are available as mass-products for the automotive and / or consumer electronic market, the prices will drop to a few hundred USD.

Radiotherapy could also benefit from the above presented systems. This work showed the basic principles to use a ToF camera to improve a radiotherapy workflow. Open tasks are still, the multimodal fusion of CT and ToF data or the feature based pre-registration, to furthermore improve registration results for large rotations or translations. It might be also worth to use modern graphic cards using OpenCL and CUDA to furthermore improve the speed of the presented methods.

Based on this work, a research project called motion management in the scope of "Bayern 2020" started about one year ago. The project is a joint research project of Siemens AG (Health Care), softgate GmbH and the pattern recognition lab of the University Erlangen-Nuremberg. With a runtime of three years, the project has the aim to develop prototypes close to a product using ToF cameras for a radiotherapy workflow.

Bibliography

- [AAPM 06] AAPM Task Group 76. “The Management of Respiratory Motion in Radiation Oncology”. Tech. Rep., American Association of Physicists in Medicine, College Park, MD 20740-3846, 2006.
- [Adel 08] A. Adelt, C. Schaller, J. Penne, and J. Hornegger. “Patient positioning using 3-D surface registration”. In: R. Bauernschmitt, Y. Chaplygin, H. Feußner, Y. Gulyaev, J. Hornegger, E. Mayr, N. Navab, S. Schookin, and S. Umnyashkin, Eds., *Proceedings of the 4th Russian-Bavarian Conference on Biomedical Engineering*, pp. 202–207, Moscow, Russia, 2008.
- [Ahn 04] S. Ahn, B. Yi, Y. Suh, J. Kim, S. Lee, S. Shin, and E. Choi. “A feasibility study on the prediction of tumor location in the lung from skin motion”. *British Journal of Radiology*, Vol. 77, No. 919, pp. 588–596, 2004.
- [Amer 10] American Cancer Society. “Cancer Facts and Figures 2010”. Tech. Rep., American Cancer Society, Atlanta, 2010.
- [Berb 05] R. Berbeco, T. Neicu, E. Rietzel, G. Chen, and S. Jiang. “A technique for respiratory-gated radiotherapy treatment verification with an EPID in cine mode”. *Physics in Medicine and Biology*, Vol. 50, No. 16, pp. 3669–79, 2005.
- [Bied 07] J. Biederer, J. Dinkel, H. Bolte, T. Welzel, B. Hoffmann, C. Thierfelder, U. Mende, J. Debus, M. Heller, and H.-U. Kauczor. “Respiratory-gated helical computed tomography of lung: reproducibility of small volumes in an ex vivo model”. *International Journal of Radiation Oncology, Biology, Physics*, Vol. 69, No. 5, pp. 1642–1649, 2007.
- [Bron 08] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. *Numerical Geometry of Non-Rigid Shapes (Monographs in Computer Science)*. Springer, 1 Ed., October 2008.
- [Cash 03] D. M. Cash, T. K. Sinha, R. L. Galloway, M. I. Miga, W. C. Chapman, and B. M. Terawaki, H. amd Dawant. “Incorporation of a laser range scanner into image-guided liver surgery: Surface acquisition, registration, and tracking”. *Medical Physics*, Vol. 30, No. 7, pp. 1671–1682, July 2003.
- [Chen 92] Y. Chen and G. Medioni. “Object modelling by registration of multiple range images”. *Image and Vision Computing*, Vol. 10, No. 3, pp. 145–155, 1992.
- [Cook 07] R. A. H. Cook, G. Carnes, T.-Y. Lee, and R. G. Wells. “Respiration-averaged CT for attenuation correction in canine cardiac PET/CT”. *Journal of Nuclear*

Bibliography

- Medicine*, Vol. 48, No. 5, pp. 811–818, 2007.
- [Deva 07] P. Devarakota, M. Castillo-Franco, R. Ginhoux, B. Mirbach, and B. Ottersen. “Occupant Classification Using Range Images”. *IEEE Transactions on Vehicular Technology*, No. 4, July 2007.
- [Doug 94] Douglas Jones Reviewer. “ICRU Report 50—Prescribing, Recording and Reporting Photon Beam Therapy”. *Medical Physics*, Vol. 21, No. 6, pp. 833–834, 1994.
- [Erri 03] S. C. Erridge, Y. Seppenwoolde, S. H. Muller, M. van Herk, K. de Jaeger, J. S. A. Belderbos, L. J. Boersma, and J. V. Lebesque. “Portal imaging to assess set-up errors, tumor motion and tumor shrinkage during conformal radiotherapy of non-small cell lung cancer”. *Radiotherapy and Oncology*, Vol. 66, No. 1, pp. 75–85, 2003.
- [Fole 08] J. D. Foley. *Computer graphics: Principles and practice. Addison-Wesley systems programming series*, Addison-Wesley, Boston, 2. ed. in c, reprinted with corr., july 1997, 24. print. Ed., 2008.
- [Ford 02] E. Ford, G. Mageras, E. Yorke, K. Rosenzweig, R. Wagman, and C. Ling. “Evaluation of respiratory movement during gated radiotherapy using film and electronic portal imaging”. *International Journal Radiation Oncology Biology Physics*, Vol. 52, No. 2, pp. 522–31, 2002.
- [Fuch 08] S. Fuchs and G. Hirzinger. “Extrinsic and depth calibration of ToF-cameras”. In: *CVPR08*, pp. 1–6, 2008.
- [Gier 08a] D. P. Gierga, M. Riboldi, J. C. Turcotte, G. C. Sharp, S. B. Jiang, K. P. Doppke, A. Katz, A. G. Taghian, and G. T. Y. Chen. “Target Registration Error with Three-Dimensional Surface Imaging in Setup of Image-Guided Partial Breast Irradiation”. *International Journal of Radiation Oncology Biology Physics*, Vol. 70, No. 4, pp. 1239–1246, 2008.
- [Gier 08b] D. P. Gierga, M. Riboldi, J. C. Turcotte, G. C. Sharp, S. B. Jiang, A. G. Taghian, and G. T. Y. Chen. “Comparison of target registration errors for multiple image-guided techniques in accelerated partial breast irradiation.”. *International Journal Radiation Oncology Biology Physics*, Vol. 70, No. 4, pp. 1239–46, 2008.
- [Gira 02] P. Giraud, S. Elles, S. Helfre, Y. de Rycke, V. Servois, M. F. Carette, C. Alzieu, P. Y. Bondiau, B. Dubray, E. Touboul, M. Housset, J. C. Rosenwald, and J. M. Cosset. “Conformal radiotherapy for lung cancer: different delineation of the gross tumor volume (GTV) by radiologists and radiation oncologists”. *Radiotherapy and Oncology*, Vol. 62, No. 1, pp. 27–36, 2002.
- [Gray 06] R. Gray. *Toeplitz and Circulant Matrices: A review*. Now Publishers Inc, 2006.
- [Grim 98] W. Grimson, M. Leventon, G. Ettinger, A. Chabrierie, F. Ozlen, S. Nakajima, H. Atsumi, R. Kikinis, and B. P. “Clinical Experience with a High Precision Image-Guided Neurosurgery System”. In: William M. Wells III and Alan C. F. Colchester and Scott L. Delp, Eds., *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98, First International Con-*

- ference, Cambridge, MA, USA, October 11-13, 1998, Proceedings, pp. 63–73, Springer, 1998.
- [Heik 97] J. Heikkila and O. Silven. “A Four-step Camera Calibration Procedure with Implicit Image Correction”. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, Vol. 0, pp. 1106–1112, 1997.
- [Hois 04] J. Hoisak, K. Sixel, R. Tirona, P. Cheung, and J. Pignol. “Correlation of lung tumor motion with external surrogate indicators of respiration”. *International Journal of Radiation Oncology Biology Physics*, Vol. 60, No. 4, pp. 1298–1306, 2004.
- [Holl 08] K. Höller, C. Schaller, D. Tacke, F. Höpfl, and J. Hornegger. “Contributions of Time-of-Flight cameras for biomedical applications”. In: D. G. für Biomedizinische Technik, Ed., *Innovationen bei der Erfassung und Analyse bioelektrischer und biomagnetischer Signale*, pp. 180–183, Braunschweig und Berlin, 2008.
- [Holl 10] K. Höller. *Novel Techniques for Spatial Orientation in Natural Orifice Translumenal Endoscopic Surgery (NOTES)*. PhD thesis, University Erlangen, Germany, 2010.
- [Hsu 06] S. Hsu, S. Acharya, A. Rafii, and R. New. “Performance of a Time-of-Flight Range Camera for Intelligent Vehicle Safety Applications”. *Advanced Microsystems for Automotive Applications 2006*, pp. 205–219, 2006.
- [Huss 08] S. Hussmann, B. Hagebeuker, and T. Ringbeck. “A performance review of 3D TOF vision systems in comparison to stereo vision systems”. In: A. Bhatti, Ed., *Stereo Vision*, pp. 103–120, I-Tech Education and Publishing, Vienna, Austria, 2008.
- [Iona 07] D. Ionascu, S. B. Jiang, S. Nishioka, H. Shirato, and R. I. Berbeco. “Internal-external correlation investigations of respiratory induced motion of lung tumors”. *Medical Physics*, Vol. 34, No. 10, pp. 3893–3903, 2007.
- [Jian 05] S. Jiang, R. Berbeco, J. Wolfgang, G. Sharp, K. Doppke, T. Neicu, Y. Chen, P. Busse, and G. Chen. “Image-Guided Respiration-Gated Treatment”. *Medical Physics*, Vol. 32, No. 6, p. 2160, 2005.
- [Keal 01] P. J. Keall, V. R. Kini, S. S. Vedam, and R. Mohan. “Motion adaptive x-ray therapy: a feasibility study”. *Physics in Medicine and Biology*, Vol. 46, No. 1, pp. 1–10, 2001.
- [Keal 06] P. J. Keall, G. S. Mageras, J. M. Balter, R. S. Emery, K. M. Forster, S. B. Jiang, J. M. Kapatoes, D. A. Low, M. J. Murphy, B. R. Murray, C. R. Ramsey, M. B. van Herk, S. S. Vedam, J. W. Wong, and E. Yorke. “The management of respiratory motion in radiation oncology report of AAPM Task Group 76”. *Medical Physics*, Vol. 33, No. 10, pp. 3874–3900, 2006.
- [Kham 06] A. Khamene, P. Bloch, W. Wein, M. Svatos, and F. Sauer. “Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy”. *Medical Image Analysis*, Vol. 10, No. 1, pp. 96–112, 2006.
- [Kham 07] A. Khamene, C. Schaller, J. Hornegger, J. C. Celi, B. Ofstad, E. Rietzel, X. A. Li, A. Tai, and J. Bayouth. “A Novel Image Based Verification Method for Respiratory Motion Management in Radiation Therapy”. In: D. Metaxas,

Bibliography

- B. C. Vemuri, A. Shashua, and H. Shum, Eds., *Eleventh IEEE International Conference on Computer Vision*, DVD Proceedings, 2007.
- [Koch 04] N. Koch, H. Liu, G. Starkschall, M. Jacobson, K. Forster, Z. Liao, R. Komaki, and C. Stevens. “Evaluation of internal lung motion for respiratory-gated radiotherapy using MRI: Part II-margin reduction of internal target volume”. *International Journal Radiation Oncology Biology Physics*, Vol. 60, No. 5, pp. 1473–1483, 2004.
- [Koch 09] R. Koch, I. Schiller, B. Bartczak, F. Kellner, and K. Koeser. “MixIn3D: 3D Mixed Reality with ToF-Camera”. In: *Dynamic 3D Imaging DAGM 2009 Workshop, Dyn3D, LNCS 5742*, pp. 126–141, Jena, Germany, September 2009.
- [Koch 84] D. H. U. Kochanek and R. H. Bartels. “Interpolating splines with local tension, continuity, and bias control”. *SIGGRAPH: Computer Graphics*, Vol. 18, No. 3, pp. 33–41, 1984.
- [Koll 08] E. Kollarz, J. Penne, J. Hornegger, and A. Barke. “Gesture recognition with a Time-Of-Flight camera”. *International Journal of Intelligent Systems Technologies and Applications*, Vol. 5, No. 3/4, pp. 334–343, 2008.
- [Kubo 96] H. D. Kubo and B. C. Hill. “Respiration gated radiotherapy treatment: a technical study”. *Physics in Medicine and Biology*, Vol. 41, No. 1, pp. 83–91, 1996.
- [Ladi 08] A. Ladikos, S. Benhimane, and N. Navab. “Real-Time 3D Reconstruction for Collision Avoidance in Interventional Environments”. In: *MICCAI ’08: Proceedings of the 11th International Conference on Medical Image Computing and Computer-Assisted Intervention, Part II*, pp. 526–534, Springer-Verlag, Berlin, Heidelberg, 2008.
- [Lang 00] R. Lange. *3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-Technology*. PhD thesis, University Siegen, Germany, 2000.
- [Mach 05] M. Machtay. “Higher BED is associated with improved local-regional control and survival for NSCLC treated with chemoradiotherapy: An RTOG analysis”. *International Journal Radiation Oncology Biology Physics*, Vol. 63, No. 2, p. 66, 2005.
- [Mage 04] G. S. Mageras, A. Pevsner, E. Yorke, K. Rosenzweig, E. Ford, A. Hertanto, S. Larson, D. Lovelock, Y. Erdi, S. Nehmeh, J. Humm, and C. Ling. “Measurement of lung tumor motion using respiration-correlated CT”. *International Journal Radiation Oncology Biology Physics*, Vol. 60, No. 3, pp. 933–941, 2004.
- [Mull 09] K. Müller, C. Schaller, J. Penne, and J. Hornegger. “Surface-based Respiratory Motion Classification and Verification”. In: H.-P. Meinzer, T. M. Deserno, H. Handels, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin 2009*, pp. 257–261, Berlin, 2009.
- [Nava 06] N. Navab, S. Wiesner, S. Benhimane, E. Euler, and S. Heining. “Visual Servoing for Intraoperative Positioning and Repositioning of Mobile C-arms”. In: R. Larsen, M. Nielsen, and J. Sporrings, Eds., *Medical Image Comput-*

- ing and Computer-Assisted Intervention - MICCAI 2006, 9th International Conference, Copenhagen, Denmark, October 1-6, 2006, Proceedings, Part I*, pp. 551–560, Springer, 2006.
- [Neic 06] T. Neicu, R. Berbeco, J. Wolfgang, and S. B. Jiang. “Synchronized moving aperture radiation therapy (SMART): improvement of breathing pattern reproducibility using respiratory coaching”. *Physics in Medicine and Biology*, Vol. 51, No. 3, pp. 617–636, 2006.
- [Ohar 89] K. Ohara, T. Okumura, M. Akisada, T. Inada, T. Mori, H. Yokota, and M. J. Calaguas. “Irradiation synchronized with respiration gate”. *International Journal Radiation Oncology Biology Physics*, Vol. 17, No. 4, pp. 853–857, 1989.
- [Okun 95] P. Okunieff, D. Morgan, A. Niemierko, and H. Suit. “Radiation dose-response of human tumors”. *International Journal Radiation Oncology Biology Physics*, Vol. 32, No. 4, pp. 1227–1237, 1995.
- [Olse 07] D. R. Olsen and D. I. Thwaites. “Now you see it... Imaging in radiotherapy treatment planning and delivery”. *Radiotherapy and oncology: journal of the European Society for Therapeutic Radiology and Oncology*, Vol. 85, No. 2, pp. 173–175, November 2007.
- [Onim 05] R. Onimaru, H. Shirato, M. Fujino, K. Suzuki, K. Yamazaki, M. Nishimura, H. Dosaka-Akita, and K. Miyasaka. “The effect of tumor location and respiratory function on tumor movement estimated by real-time tracking radiotherapy (RTRT) system”. *International Journal of Radiation Oncology, Biology, Physics*, Vol. 63, No. 1, pp. 164–169, 2005.
- [Pari 09] S. Paris and F. Durand. “A Fast Approximation of the Bilateral Filter Using a Signal Processing Approach”. *International Journal on Computer Vision*, Vol. 81, No. 1, pp. 24–52, 2009.
- [Penn 07] J. Penne, K. Höller, S. Krüger, and H. Feußner. “NOTES 3D: Endoscopes learn to see 3-D; Basic Algorithms for a Novel Endoscope”. In: *VISAPP 2007, Second international Conference on Computer Vision Theory and Applications*, pp. 134–139, 2007.
- [Penn 08a] J. Penne, C. Schaller, and J. Hornegger. “Robust Real-Time 3D Respiratory Motion Detection Using Time-of-Flight Cameras”. In: H. Lemke, K. Inamura, K. Doi, M. Vannier, and A. Farman, Eds., *Computer Assisted Radiology and Surgery 2008 - Proceedings of the 22nd International Congress and Exhibition*, Heidelberg, 2008.
- [Penn 08b] J. Penne, C. Schaller, J. Hornegger, and T. Kuwert. “Robust Real-Time 3D Respiratory Motion Detection Using Time-of-Flight Cameras”. *Computer Assisted Radiology and Surgery 2008*, Vol. 3, No. 5, pp. 427–431, 2008.
- [Pere 86] C. Perez, M. Bauer, S. Edelstein, B. Gillespie, and R. Birch. “Impact of tumor control on survival in carcinoma of the lung treated with irradiation”. *International Journal Radiation Oncology Biology Physics*, Vol. 12, No. 4, pp. 539–547, 1986.
- [Plac 10a] S. Placht. *Development and Design of a Feature Based Surface Registration System for Patient Positioning Using Time-of-Flight Cameras*. Master’s

Bibliography

- thesis, University of Erlangen-Nuremberg, Germany, 2010.
- [Plac 10b] S. Placht, C. Schaller, M. Balda, A. Adelt, C. Ulrich, and J. Hornegger. “Improvement and Evaluation of a Time-of-Flight based patient positioning system”. In: H.-P. Meinzer, T. M. Deserno, H. Handels, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin*, pp. 177–181, Berlin, 2010.
- [Prus 08] A. Prusak, O. Melnychuk, H. Roth, I. Schiller, and R. Koch. “Pose estimation and map building with a Time-Of-Flight camera for robot navigation”. *International Journal of Intelligent Systems Technologies and Applications*, Vol. 5, No. 3/4, pp. 355–364, 2008.
- [Rabi 89] L. R. Rabiner. “A tutorial on hidden markov models and selected applications in speech recognition”. In: *Proceedings of the IEEE*, pp. 257–286, 1989.
- [Rang 05] D. Rangaraj and L. Papiez. “Synchronized delivery of DMLC intensity modulated radiation therapy for stationary and moving targets”. *Medical Physics*, Vol. 32, No. 6, pp. 1802–1817, 2005.
- [Reis 09] M. F. Reiser, C. R. Becker, G. Glazer, and K. Nikolaou. *Multislice CT. Medical Radiology*, Springer Berlin Heidelberg, Berlin, Heidelberg, 3. Ed., 2009.
- [Riet 06] E. Rietzel and G. T. Y. Chen. “Improving retrospective sorting of 4D computed tomography data”. *Medical Physics*, Vol. 33, No. 2, pp. 377–379, 2006.
- [Ring 09] T. Ringbeck, M. Profitlich, and C. Schaller. “Kameras für die dritte Dimension”. *Optik & Photonik*, Vol. Oktober, No. 3, pp. 30–33, 2009.
- [Rohk 08] C. Rohkohl, M. Prümmer, G. Lauritsch, and J. Hornegger. “Cardiac C-arm CT: image-based gating”. In: J. Hsieh and E. Samei, Eds., *Medical Imaging 2008: Physics of Medical Imaging. Proceedings of the SPIE*, 2008.
- [Rusi 01] S. Rusinkiewicz and M. Levoy. “Efficient Variants of the ICP Algorithm”. In: *Proceedings of the Third International Conference on 3D Digital Imaging and Modeling*, pp. 145–152, 2001.
- [Russ 03] D. A. Russakoff, T. Rohlfing, C. R. Maurer, and Jr. “Fast Intensity-based 2D-3D Image Registration of Clinical Data Using Light Fields”. In: *Proc. 9th IEEE International Conference on Computer Vision (ICCV)*, pp. 416–423, 2003.
- [Scha 07] C. Schaller, A. Khamene, and J. Hornegger. “Image Based Time Series Synchronization for Periodically Moving Targets”. In: J. Hornegger, E. W. Mayr, S. Schookin, H. Feußner, N. Navab, Y. V. Gulyaev, K. Höller, and V. Ganzha, Eds., *3rd Russian-Bavarian Conference on Biomedical Engineering*, pp. 106–110, Erlangen, 2007.
- [Scha 08] C. Schaller, J. Penne, and J. Hornegger. “Time-of-Flight Sensor for Respiratory Motion Gating”. *Medical Physics*, Vol. 35, No. 7, pp. 3090–3093, 2008.
- [Scha 09a] C. Schaller, A. Adelt, J. Penne, and J. Hornegger. “Time-of-flight sensor for patient positioning”. In: E. Samei and J. Hsieh, Eds., *Proceedings of SPIE*, 2009.
- [Scha 09b] C. Schaller, C. Rohkohl, J. Penne, M. Stürmer, and J. Hornegger. “Inverse C-arm Positioning for Interventional Procedures Using Real-Time Body Part

- Detection”. In: G.-Z. Y. et al., Ed., *MICCAI 2009, Part I, LNCS 5761*, Berlin - Heidelberg, 2009.
- [Scho 07a] P. Schöffel, W. Harms, G. Sroka-Perez, W. Schlegel, and C. Karger. “Accuracy of a commercial optical 3D surface imaging system for realignment of patients for radiotherapy of the thorax”. *Physics in Medicine and Biology*, Vol. 53, No. 13, pp. 3949–3963, 2007.
- [Scho 07b] P. J. Schöffel, W. Harms, G. Sroka-Perez, W. Schlegel, and C. P. Karger. “Accuracy of a commercial optical 3D surface imaging system for realignment of patients for radiotherapy of the thorax”. *Physics in Medicine and Biology*, Vol. 52, No. 13, pp. 3949–3963, 2007.
- [Schu 08] P. Schuhmann, J. Penne, C. Schaller, J. Zeintl, J. Hornegger, and T. Kuwert. “Optical Tracking of Respiratory Motion Using a Time of Flight Camera and its Applicability in Emission Tomography”. In: D. G. für Nuklearmedizin, Ed., *NuklearMedizin 2008*, 2008.
- [Schw 00] A. Schweikard, G. Glosser, M. Bodduluri, M. J. Murphy, and J. R. Adler. “Robotic motion compensation for respiratory movement during radiosurgery”. *Computer Aided Surgery*, Vol. 5, No. 4, pp. 263–277, 2000.
- [Sega 03] W. P. Segars, B. M. W. Tsui, E. C. Frey, and E. K. Fishman. “Extension of the 4D NCAT phantom to dynamic X-ray CT simulation”. *IEEE Nuclear Science Symposium Conference Record*, Vol. 5, pp. 3195–3199, 2003.
- [Sepp 02] Y. Seppenwoolde, H. Shirato, K. Kitamura, S. Shimizu, M. van Herk, J. V. Lebesque, and K. Miyasaka. “Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy”. *International Journal of Radiation Oncology, Biology, Physics*, Vol. 53, No. 4, pp. 822–834, 2002.
- [Sepp 07] Y. Seppenwoolde, R. I. Berbeco, S. Nishioka, H. Shirato, and B. Heijmen. “Accuracy of tumor motion compensation algorithm from a robotic respiratory tracking system: a simulation study”. *Medical Physics*, Vol. 34, No. 7, pp. 2774–2784, 2007.
- [Shir 03] H. Shirato, T. Harada, T. Harabayashi, K. Hida, H. Endo, K. Kitamura, R. Onimaru, K. Yamazaki, N. Kurauchi, T. Shimizu, N. Shinohara, M. Matsushita, H. Dosaka-Akita, and K. Miyasaka. “Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers for precise setup and real-time tumor tracking in radiotherapy”. *International journal of Radiation Oncology, Biology, Physics*, Vol. 56, No. 1, pp. 240–247, 2003.
- [Sieb 07a] M. von Siebenthal, G. Székely, U. Gamper, P. Boesiger, A. Lomax, and P. Cattin. “4D MR imaging of respiratory organ motion and its variability”. *Physics in Medicine and Biology*, Vol. 52, No. 6, pp. 1547–1564, 2007.
- [Sieb 07b] M. von Siebenthal, G. Szekely, A. J. Lomax, and P. C. Cattin. “Systematic errors in respiratory gating due to intrafraction deformations of the liver.”. *Medical Physics*, Vol. 34, No. 9, pp. 3620–9, 2007.
- [Sout 08] S. Soutschek, J. Penne, J. Hornegger, and J. Kornhuber. “3-D Gesture-Based Scene Navigation in Medical Imaging Applications Using Time-Of-Flight

Bibliography

- Cameras”. In: *IEEE Computer Society Conference on Computer Vision, Omnipress*, p. DVD Proceedings, 2008.
- [Stud 99] C. Studholme, D. L. G. Hill, and D. J. Hawkes. “An overlap invariant entropy measure of 3D medical image alignment”. *Pattern Recognition*, Vol. 32, No. 1, pp. 71–86, 1999.
- [Tart 06] S. Tarte, J. McClelland, H. S., J. Blackall, D. Landau, and D. Hawkes. “A Non-Contact Method for the Acquisition of Breathing Signals that Enable Distinction Between Abdominal and Thoracic Breathing”. *Radiotherapy and Oncology*, Vol. 81(1), p. S209, October, 2006.
- [Toma 98] C. Tomasi and R. Manduchi. “Bilateral Filtering for Gray and Color Images”. In: *ICCV ’98: Proceedings of the Sixth International Conference on Computer Vision*, p. 839, IEEE Computer Society, Washington, DC, USA, 1998.
- [Tsai 87] R. Tsai. “A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses”. *IEEE Journal of Robotics and Automation*, Vol. 3, No. 4, pp. 323–344, August 1987.
- [Tsun 04] Y. Tsunashima, T. Sakae, Y. Shioyama, K. Kagei, T. Terunuma, A. Nohtomi, and Y. Akine. “Correlation between the respiratory waveform measured using a respiratory sensor and 3D tumor motion in gated radiotherapy”. *International Journal Radiation Oncology Biology Physics*, Vol. 60, No. 3, pp. 951–958, 2004.
- [Ue 06] H. Ue, H. Haneishi, H. Iwanaga, and K. Suga. “Nonlinear motion correction of respiratory-gated lung SPECT images”. *IEEE Transactions on Medical Imaging*, Vol. 25, No. 4, pp. 486–495, 2006.
- [Ulri 10] C. Ulrich, C. Schaller, J. Penne, and J. Hornegger. “Evaluation of a Time-of-Flight based respiratory motion management system”. In: H.-P. Meinzer, T. M. Deserno, H. Handels, and T. Tolxdorff, Eds., *Bildverarbeitung für die Medizin*, Berlin, 2010.
- [Veda 03] S. S. Vedam, V. Kini, P. Keall, V. Ramakrishnan, H. Mostafavi, and R. Mohan. “Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker”. *Medical Physics*, Vol. 51, No. 3, pp. 505–513, 2003.
- [Vite 03] A. Viterbi. “Error bounds for convolutional codes and an asymptotically optimum decoding algorithm”. *IEEE Transactions on Information Theory*, Vol. 13, No. 2, pp. 260–269, January 2003.
- [Wang 08] L. Wang, J. Traub, S. Heining, S. Benhimane, E. Euler, R. Graumann, and N. Navab. “Long Bone X-Ray Image Stitching Using Camera Augmented Mobile C-Arm”. In: D. N. Metaxas, L. Axel, F. G., and G. Székely, Eds., *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2008, 11th International Conference, New York, NY, USA, September 6-10, 2008, Proceedings, Part II*, pp. 578–586, Springer, 2008.
- [Wu 04] H. Wu, G. C. Sharp, B. Salzberg, D. Kaeli, H. Shirato, and S. B. Jiang. “A finite state model for respiratory motion analysis in image guided radiation

- therapy”. *Physics in Medicine and Biology*, Vol. 49, No. 23, pp. 5357–5372, 2004.
- [Xu 98] Z. Xu, R. Schwarte, H. Heinol, B. Buxbaum, and T. Ringbeck. “Smart Pixel - Photometric Mixer Device (PMD) / New System Concept of a 3D-Imaging-on-a-Chip”. In: *5th International Conference on Mechatronics and Machine Vision in Practice*, pp. 259–264, 1998.
- [Zhan 00] Z. Zhang. “A flexible new technique for camera calibration”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 11, pp. 1330–1334, nov 2000.
- [Zhan 03] T. Zhang, H. Keller, M. O’Brien, T. Mackie, and B. Paliwal. “Application of the spirometer in respiratory gated radiotherapy”. *Medical Physics*, Vol. 30, No. 12, pp. 3165–3171, 2003.

List of Tables

3.1	Correlation of tumor/organ motion with the respiratory signal; 3-D: three - dimensional, AP: anterior - posterior, CT: computed tomography, MRI: magnetic resonance imaging, pts: patients, s: second(s), SI: superior - inferior [AAPM 06]	20
3.2	Results of the two error measures for various datasets.	28
3.3	Results for the sample sequence: NN-NN-DS-NN-SF-SF	30
4.1	Overview of advantages and disadvantages of different 3-D surface acquisition technologies. ⁽¹⁾ There is a strong dependency on the observed object. (Full circle represents the best.)	40
5.1	Overview of advantages and disadvantages of different patient positioning systems. (Full circle represents the best. In the case of IGRT one has to count the additional process of surgery as well.)	66
5.2	RMS of the Euclidean error of the computed translation with respect to the ground truth for lateral (x), longitudinal (y) and vertical (z) displacements. ⁽¹⁾ all measurements are in mm.	76
5.3	RMS of the Euclidean error of the computed rotations with respect to the ground truth. ⁽¹⁾ translation measurements are in mm, rotation measurements in °.	76
6.1	Overview of advantages and disadvantages of different respiratory motion gating systems. (Full circle is best. In the case of complexity, a full circle means, that the system is less complex.)	82
6.2	Experimental results showing average, median, variance of the correlation ANZAI to ToF signal for 13 test persons; ⁽¹⁾ average abdominal breathing percentage	91

List of Figures

1.1	Age-adjusted cancer death rates (per 100,000, age adjusted to the 2000 US standard population) for lung and bronchus cancer, US, 1930-2006 [Amer 10].	2
1.2	Schematic overview of a conventional radiotherapy workflow. The process usually starts with the acquisition of a CT, followed by the treatment planning. This initial procedure is usually done only once. Afterwards, the treatment sessions require patient positioning and also a motion management strategy, depending on the location of the tumor.	2
1.3	LINAC with rotating gantry and patient table. Source: based on linear particle accelerator model (Google Sketchup 3-D Warehouse.)	3
2.1	A schematic overview of the 4-D CT process. At each table position images for many respiratory phases are acquired (images of the same respiratory phase are indicated by colors). Afterwards, the time synchronized images are sorted and put together to a 3-D CT volume sequence.	8
2.2	Clinical target volumes, according to the ICRU Reports 50 and its supplement 62 [Doug 94]. The ITV/PTV regions, which should be reduced to a minimum are indicated in red.	9
2.3	Screenshot of a treatment planning system from RaySearch and the Massachusetts General Hospital (MGH). The distribution of the applied radiation is displayed as an color-coded overlay. Source: http://gray.mgh.harvard.edu/	10
2.4	Phase gating. Left: Moving tumor at different points in time. Right: Corresponding breathing cycle with a gate applied on t_0 , t_1 and t_4	13
2.5	External Gating: An external surrogate (red box) is used to measure e.g. chest or abdominal movement. A correlation between the tumor movement and the surrogate signal is assumed. Internal Gating: A fiducial marker (red circle) is implanted near the tumor. The information gained through tracking the fiducial marker is used as a surrogate respiratory signal. . . .	14
3.1	Left: Depicts the coronal slices of planning 4-D CT data sets with visible contour lines. Right: Demonstrates two frames of a fluoroscopic acquisition overlaid with the outlines of the projected structures from the planning 4-D dataset [Kham 07].	21

List of Figures

3.2	Schematic overview of the similarity matrix based on 4-D CT DRRs $I_i^{(p)}$ and a sequence of fluoroscopic images $I_j^{(pt)}$ with a corresponding respiratory signal $s_j^{(pt)}$. The last row (red box) indicates an additional row for "no-match" situation.	23
3.3	Left: Depicts the state transition network for $L = 5$. Thicker lines denote higher transition probability, whereas red lines denote the transition to an undefined state. Right: Example for a state transition matrix \mathbf{A} . Darker entries show higher probabilities to stay in this state.	25
3.4	Left: Breathing cycle, represented by different states, where all states are known. Right: An unknown state is included, which causes the LINAC to stop irradiating.	27
3.5	In the top row of the figures (left) and (right) a 4-D CT frame with matching fluoroscopic frame with overlaid contours for phantom and patient datasets is depicted, respectively. The bottom row shows the corresponding image similarity matrix for each case [Kham 07].	28
3.6	Similarity matrix for a specific sequence (NN-NN-DS-NN-SF-SF). Between frame 65 and 73 there is no information available as there are no corresponding 4-D CT volumes. The green dots in the upper image indicate corresponding frames, where the red line shows the computed path. . .	30
4.1	Examples of ToF camera models. Left side: CamCube from PMD Technologies GmbH, Germany (http://www.pmdtec.com). Right side: SR4000 from MESA Imaging AG, Switzerland (http://www.swissranger.ch) . . .	32
4.2	ToF camera data. Upper left: amplitude image (red pixels indicate pixels not considered in the 3-D reconstruction. Lower left: distance image. Right: 3-D reconstruction of data.	33
4.3	Relation of a phase shift to the distance at a fixed modulation frequency of $f_{mod} = 20$ MHz. A single wave is of length 2π or 15 m. Given a phase shift ϕ_d , the phase shift is proportional to the distance d	35
4.4	ToF principle for various distances at a modulation frequency of 20 MHz. Dashed green line indicates border of first non-ambiguity range for this modulation frequency. If the object is beyond this line, the distance cannot non-ambiguously be determined.	38
4.5	Setup for optical calibration. A checkerboard is placed in front of the camera at different distances and orientations.	42
4.6	Setup for distance calibration. Boards with different reflectivity (gray values) are placed at different distances (1.0m, 1.2m, 1.4m) in front of the camera. The acquired distances are used to perform a distance calibration.	43
4.7	Computation of κ by the theorem on intersecting lines.	45
4.8	Computation of distance correction for a measured z_k value. The two planes represent the look-up table samples, with the pixel position in one direction and the intensity samples in the other direction, for different positions (e.g. 1.0 m and 1.2 m). κ^{high} and κ^{low} are computed using a spline interpolation. The distance offset $\kappa^{(ip)}$ is computed using a linear interpolation.	46
4.9	Gain of distance correction for three different samples.	47

4.10 Left: surface without distance correction. Right: surface after correction..	48
5.1 Setup to detect the treatment couch coordinate system. The x , y , z coordinate system and the rotational part around the z -axis is shown in red. The checkerboard has to be visible by the ToF camera all time. This calibration is a one-time calibration.	51
5.2 Determination of x , y and z axes of the coordinate system. The corresponding vectors (green) are computed and averaged (red).	52
5.3 Determination of rotation center, by optimizing c_{3D}	54
5.4 Overview: The red lines indicate the three principal axes ρ_0, ρ_1, ρ_2 . The three golden sections $\gamma_0, \gamma_1, \gamma_2$ are shown as yellow lines. Furthermore all intersection points ι_k and the centroid μ are also depicted.	56
5.5 Refinement of the bounding boxes. Left: Detection of the pelvic region with the corresponding histogram below. μ is the origin of the coordinate system. The search interval d_1 has a length of 30 cm in each direction with a binning size of 4.8 cm. Right: Detection of the neck with the corresponding histogram below. γ_1 is the upper golden section intersection. The search interval d_0 has a length of 10 cm with a binning size of 3.2 cm.	57
5.6 Full body acquisition using a ToF sensor including bounding boxes (side view and top view). On the left side information about the different bounding boxes is shown. V shows the volume of the bounding box in cm^3 , I the isocenter of the bounding box in mm and B the bounds of the actual bounding box in mm. On the lower right side, some basic information about the dimensions of the patient is displayed.	58
5.7 Coin calibration: A coin is placed on a box on the patient table. The coin is aligned in the isocenter of the C-arm. Afterwards, the corresponding 3-D coordinate of the coin position can be determined in the ToF point cloud. This calibration is a fast and fairly efficient alternative to the more complex calibration presented in the beginning of this chapter.	60
5.8 Phantom evaluation: Two upper left images: 0° and 90° projections of manually aligned head (gold standard). Two lower left images: 0° and 90° projections of automatically aligned head using the proposed method. Two upper right images: 0° and 90° projection difference images of manually and automatically aligned head. Two lower right images: 0° and 90° projections of manually and automatically aligned head (automatically aligned projections are superimposed in red).	61
5.9 The BrainLAB NOVALIS system. An infrared camera is used to track reflective marker on the patients' skin. Source: http://www.brainlab.com .	63
5.10 The BrainLAB ExacTrac x-Ray 6D system. Showing the x-ray sources and the corresponding flat panel detectors as well as the LINAC and the patient lying on the couch. Source: http://www.brainlab.com	64
5.11 VisionRT System. The system uses an active illumination to project a speckle pattern on the patient. Using this pattern, the 3-D surface can be reconstructed. Based on this reconstruction, the system can compute a respiratory signal. Source: http://www.visionrt.com	65

5.12 Processing pipeline: Each target, reference, and moving target is processed the same way. After applying a bilateral filter [Toma 98, Pari 09], a variance filter is applied. This filter is necessary to eliminate so-called "flying pixel". Before the 3-D body points are segmented from the background, a temporal filtering is applied as well. Using a point-to-plane ICP both targets are registered in the end. Note: The "gap" in the phantom is due to its mechanical parts.	66
5.13 Comparison of preprocessing steps: Original ToF data on the left side (with flying pixels), pre-processed data on the right side	67
5.14 Comparison of distance metrics (Source: [Bron 08]).	69
5.15 Schematic overview of the mechanical motion of the thorax and the abdomen (sideview).	73
5.16 Plaster cast phantom used to simulate respiratory motion. Left: surface of the plaster cast. Right: ToF acquisition of the phantom.	74
5.17 Schematic overview of the test setup. Three gauges with an accuracy of $\frac{1}{100}$ mm and a range of 10 mm are attached to the patient table. One gauge for each table direction.	75
5.18 Exemplary registration result: Color coded distance between transformed source and target dataset (range: 0 mm - 5 mm).	75
 6.1 AZ-733V Anzai Belt	79
6.2 Test Person with the CPX spirometer and the preVent Pneumotach.	80
6.3 Screenshot of the software of the RPM system. The green box represents the detected box, which provides the respiratory curve. Based on this curve, the gating intervals can be set. Source: http://www.varian.com	81
6.4 Schematic overview of ToF-based respiratory motion detection. The algorithm is divided into three parts. First an independent detection of the patient table plane is performed. This plane is then used to segment the patient from the background in 3-D. To compute a respiratory signal two methods are available. The first one computes the distance of a plane to the camera, whereas the second measures the volume between the table plane and the thorax and abdomen, respectively.	82
6.5 Segmented 3-D point cloud including two axis with superimposed regions of interest (red). The partitioning of I into n parts uses the x-axis for the first division in two disjunct parts and equally subdivides the rest parallel to the y-axis.	83
6.6 Example of the histogram computation: 3-D reconstruction of three pairwise perpendicular walls and the significant part of the computed histogram representing the angles α_x and α_y of the surface normals. Bright red spots in the histogram indicate a high number of normals with the same orientation. The three planar walls yield three clearly observable peaks in the histogram. Note: Parts of the colored best fitting planes are hidden due to noise of the ToF signal and spatial averaging (near edge regions).	84

6.7	Example of the histogram computation: 3-D reconstruction of a table rear side (red) and the significant part of the computed histogram representing the angles of surface normals to x-axis and the y-axis. The red point in the histogram indicates the obviously dominant surface of the table rear side.	85
6.8	Overview of the test setup for the evaluation of the respiratory gating system. The system can be used to evaluate surface-based motion detection systems.	88
6.9	Evaluation of the thorax movement in ToF camera direction. Left: Respiratory frequency of 11 min^{-1} , amplitude of 17.8 mm, correlation factor of 0.94. Right: Respiratory frequency of 14 min^{-1} , amplitude of 1.5 mm, correlation factor of 0.65.	89
6.10	Correlation factor of the ToF signal and the phantom signal, Left: amplitude, right: frequency	90
6.11	On the left: 3-D model of the automatically segmented upper part of the body of a patient. The colored planes are best fitting planes used to determine two independent respiratory signals (red: chest; green: abdomen). On the right: Derived 1-D signals. Vertical axis: time in seconds. Horizontal axis: motion amplitude of the planes representing the respiratory motion during breathing of chest (red) and abdomen (green) as well as the ANZAI belt signal (gray).	91

List of Algorithms

1	Verification	26
2	Inverse patient positioning algorithm	59
3	Classical ICP algorithm used e.g. in the VTK (www.vtk.org) toolkit. A quaternion based estimator is used.	68
4	Pre-computation of closest points on a voxel grid	70
5	Determination of a closest point during live mode	70
6	Customized ICP algorithm for 6 DoF	72
7	Table detection algorithm	86