TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Informatik
I–16 / Computer Aided Medical Procedures

Robotic freehand SPECT Imaging

José Francisco Gardiazabal Schilling

"Experience without theory is blind, but theory without experience is mere intellectual play."

Immanuel Kant

Abstract

This work is part of the joint developments in the field of Nuclear Imaging by the Chair of Computer Aided Medical Procedures at Technische Universtät München and the Department of Nuclear Medicine at Klinikum Rechts der Isar starting in 2005. Within this collaboration freehand SPECT was born, a new imaging modality that generates 3D gamma reconstructions with mobile devices. Thus it can provide easy detection of labeled structures as well as guidance during biopsy and resection in an interventional setup.

Freehand SPECT has however limitations. Due to the freehand acquisition, the expertise of the operator has a great impact on the quality of the reconstruction. Furthermore, the lack of repeatability makes systematic adjustments very challenging. This thesis therefore presents a flexible imaging setup using a robotic arm (UR5 of Universal Robots) in order to tackle those challenges.

Beyond these, in a robot-assisted freehand, the robot can perform the same experiment an indefinite number of times, which makes it possible to see how variables like collimator type or acquisition speed play a role in the reconstruction quality. Thus it brings advantages within the improvement and understanding of freehand SPECT.

Against our expectations, the results of this work showed that the quality of the robotically acquired reconstructions were not significantly better than the ones produced by handheld human acquisitions. Together with the arrival of a better detector – the CrystalCam mini gamma camera of Crystal Photonics – as well as an exhaustive characterization and faster reconstruction algorithms, this work lead to a major step forward as the imaging results were improved to such an extent that entirely new possibilities opened up.

Whereas the original freehand SPECT – which this thesis is based on – had been developed to guide the biopsy and resection of sentinel lymph nodes in breast cancer, the robotic setup presented here renders it possible to also image smaller and more complex structures such as the thyroid, in which even nodules with no tracer uptake (cold nodules) become visible.

Moreover, by co-calibrating the robotic arm holding the mini gamma camera with a C-arm CT, a first flexible SPECT-CT prototype was developed. This resulting combination of inherently co-registered anatomical and functional information allows for a very intuitive and semi-automatic workflow for a use in an interventional setup.

As a conclusion this work extends the application possibilities of freehand SPECT and provides also means for a thorough analysis of the nature of this new imaging technology. With the tools developed here go thus beyond a theoretical improvement but bear the possibility of making freehand SPECT a successful approach in selected routine clinical applications, in particular in setups were flexibility and guidance are demanded.

Zusammenfassung

Diese Dissertation ist Teil einer Forschungskollaboration zwischen dem Lehrstuhl für Informatikanwendungen in der Medizin an der Technischen Universität München und dem Institut für Nuklearmedizin am Klinikum Rechts der Isar, welche bereits seit 2005 besteht. Im Rahmen dieser Kollaboration wurde "freehand SPECT" ins Leben gerufen, eine neue nuklearmedizinische Bildgebungsmodalität, die es erlaubt, mit mobilen Gammadetektoren 3D-Bildrekonstruktionen zu erstellen. Somit ermöglicht es eine einfachere Detektion markierter Strukturen sowie eine interventionelle Navigation während Biopsien und Resektionen.

Freehand SPECT hat jedoch gewisse Limitierungen. Wegen der freihändigen Aufnahme hat die diesbezügliche Expertise des Operateurs einen hohen Einfluss auf die Qualität der Bildrekonstruktion. Zudem macht die fehlende Wiederholbarkeit systematische Verbesserungen zu einer großen Herausforderung. In der vorliegenden Arbeit wird daher ein flexibles Bildgebungsverfahren unter Verwendung eines Roboterarms vorgestellt (UR5 von Universal Robots).

Darüber hinaus kann der Roboter in diesem roboterassistierten Freihandscan das gleiche Experiment beliebig oft wiederholen, was es ermöglicht, den Einfluss von Variablen wie Kollimatortyp und Aufnahmegeschwindigkeit auf die Rekonstruktionsqualität systematisch zu untersuchen. Damit bringt die Roboterassistenz große Vorteile bezüglich der Verbesserung und des Verständnisses von freehand SPECT mit sich.

Entgegen unserer Erwartungen zeigten die ersten Ergebnisse, dass die Qualität der robotisch aufgenommenen Rekonstruktionen nicht signifikant besser war als von Aufnahmen, die von Menschen durchgeführt wurden. Durch die Verwendung eines besseren Detektors – der CrystalCam mini Gammakamera von Crystal Photonics – sowie einer ausführlichen Charakterisierung des Aufnahmesystems und schnellerer Rekonstruktionsalgorithmen konnte ein großer Schritt nach vorne gemacht werden, da die Bildgebungsergebnisse in einer solchen Art verbessert wurden, dass sich komplett neue Anwendungsmöglichkeiten auftaten.

Während das ursprüngliche freehand SPECT – auf welchem diese Arbeit basiert – entwickelt wurde um Biopsien und Resektionen von Wächterlymphknoten bei Brustkrebs zu navigieren, macht es der hier vorgestellte robotische Aufbau möglich, Bilder von kleineren und komplexeren Strukturen wie der Schilddrüse aufzunehmen, in welchen sogar Schilddrüsenknoten ohne Traceraufnahme (kalte Knoten) sichtbar werden.

Durch eine gemeinsame Kalibrierung des Roboterarms, welcher die mini-Gammakamera hält, mit einem C-Bogen CT wurde der erste mobile SPECT-CT Demonstrator entwickelt. Die sich daraus ergebende Kombination von inhärent koregistrierten anatomischen und funktionalen Bildinformationen erlaubt einen hochintuitiven und semi-automatischen Workflow möglich, welcher in Interventionen integriert werden kann.

In der Schlussfolgerung erweitert die vorliegende Arbeit die Anwendungsmöglichkeiten von freehand SPECT und stellt die Mittel für eine eingehende Analyse der Natur dieses neuen Bildgebungsverfahrens zur Verfügung. Die entwickelten Methoden und Werkzeuge stellen nicht nur eine theoretische Verbesserung dar, sondern ermöglichen es auch, freehand SPECT zu einem erfolgreichen Ansatz für spezifische klinische Routineanwendungen zu machen, in denen Flexibilität und Navigation benötigt werden.

Acknowledgements

It all started at the end of 2007, when I visited Germany for work. On that occasion I decided to visit my good friend, Thomas Wendler, who was starting his PhD with Prof. Nassir Navab. He brought me to IFL and showed me his research, and what other people were doing at the lab. He was clearly trying to convince me of how cool it was to work on medical research and how you can really help people, instead of wasting your time on something completely useless, like Astronomy, field in which I was working at the time. Somehow he succeded and organized a position for me, so on June 2008 I moved to Munich and started working at Nassir's chair.

First of all I have to thank my family, for all their support during this whole process, from the life in Chile to helping me out to survive in Germany.

Maca, who has been a fundamental pillar in my life all these years in Germany. We met almost by chance, one and a half months before I moved here. Without her, many things would not exist, including this thesis and our amazing son, Javier, who has brought endless happiness to our lives since his arrival on December 2015.

I would like to also thank my two supervisors, Nassir Navab and Sibylle Ziegler. Nassir, you welcomed me in your chair without knowing me, you always took care of us and never ceased to motivate me. I will always be grateful for that. Sibylle, thank you for having your door always open, and for your ideas and insight for this work. These where essential in making it what it is today.

Thomas, the catalyzer of everything, thank you for your friendship and unconditional help and support in this process (including correcting this thesis).

Tobias, thank you very much for being a friend as well as a supervisor, for your infinite patience and for pushing me to do my best. Without your help, I doubt that this work would have any value.

My thanks go also to many people at CAMP: Martina Hilla, the backbone of the chair, without whom the chair wouldn't be the same. My friends and colleagues Asli Okur, Philipp Matthies, Jakob Vogel and Marco Esposito, for their friendship, the fruitful discussions and the uncountable work hours (and lack of sleep) together.

I would also like to thank Christoph Hennersperger, Sebastian Pölsterl and Richard Brosig, besides for their friendship, which I cherish beyond our situation at Nassir's chair, for their willingness on whatever new projects appeared.

Contents

Introduction

Medical doctors, up until mid 19th century, had to rely only on their senses to diagnose a patient. A bone fracture would be checked by observing and touching the patient, and diabetes would be diagnosed by checking if the urine tasted sweet.

The discovery of X-Rays and its application to medical diagnostics opened a new era in medicine, an era where it was possible to see what was inside the human body without having to open it, and to plan an intervention knowing the patient's internal anatomy, understanding the risks of it and evaluating the expected outcome. This revolution has not ceased and until today new imaging techniques based on different physical phenomena are produced. The goal has remained the same, to help physicians achieve better diagnosis.

On the interventional side, imaging techniques were not so popular, mainly because of the constraints inside the operating room. It is not possible to have big machines and, besides, to wait long for the results. Today fortunately this has changed thank to the advances in electronics, mechanics and computing capabilities, which are immensely superior to what was possible twenty or thirty years ago.

Nowadays it is possible to have anatomical information of the patient in real time inside of the operating room. For example, a patient can be re-positioned in order to have better access to a certain structure (organ, bone, etc.), after which new images can be obtained so that the surgeon can proceed with the intervention supported by this new information. This concept, known as Image Guided Surgery, is currently the trend and the basis of techniques like endoscopic and minimally invasive surgery.

The imaging devices used in nuclear medicine are, on the contrary, fairly big, and the acquisition time for diagnostic images is also long enough to make them unpractical as an intraoperative modality, therefore limiting its use to diagnostic imaging. Some years ago, around 2007, a group at Technische Universität München started with the idea of having tracked handheld detectors with the aim of generating nuclear images inside the operating room with a small footprint device. This device, called freehand SPECT, the starting point for the work here presented.

The original motivation was to evaluate the possible improvements for freehand SPECT by combining with a robotic arm. That way, it would be possible to generate repeatable experiments, which would help us better understand what produces a good or a bad image acquisition.

Later on, this motivation shifted towards the development of a robotic freehand SPECT device, since substantial improvements could be observed with the replacement of the original single pixel detectors by a calibrated small gamma camera. The potential of this new setup determined the evaluation of different acquisition protocols and tomographic reconstruction methods, which would help optimize the process and maximize the quality of the outcome.

In parallel, the idea of using the output of a C-Arm CT as reference to guide the robotic freehand SPECT acquisition was successfully tested, opening the door for the first prototype of an interventional SPECT-CT device.

The present work is divided in two parts: the first one introduces the context in which this research is situated, presenting the fields of medical and nuclear medical imaging, to follow with freehand SPECT and theoretical basis. Chapter 5, Robotics, serves as introduction to the field in which this work is set. The second part aims to explain the contributions made and how these can be extended and integrated with the current medical workflow.

Part I

State of the Art

Chapter 1

Introduction to Medical Imaging

Medical imaging is a collection of techniques designed to provide graphical representations (images) of anatomy and functions of parts of the human body. This representations include data in multiple spatial dimensions, and also changing in time. In general, the notation to describe them is to say the number of dimensions (0D to 3D) and append "+t" when the modality describes also a time variant signal. Examples of this are electrocardiograms (0D+t), radiographs (2D), ultrasound (2D, 2D+t, 3D, 3D+t), Magnetic Resonance (3D, 3D+t) and Computer Tomography (3D, 3D+t).

The aim of these imaging techniques is usually to provide physicians a better understanding of the underlying anatomy and process, so they can perform a better diagnosis. The current trend is to use also similar techniques inside the operating or interventional theaters, to provide guidance and feedback for physicians during the procedure and with this improve performance and optimize the results of the treatment.

It is important to note that Medical Imaging is not a goal *per se*, but a way to improve the outcome of the patient. A new imaging technique has to improve the workflow of the physician, e.g. making it shorter or cheaper, reducing the radiation levels or increasing the accuracy, otherwise it will not be useful.

The medical imaging domain is very wide, with hundreds if not thousands of image modalities, using very different physical phenomena to obtain information. Some imaging concepts in the scope of this thesis will be explained, and from these, most common image modalities will be explained in more detail.

1.1 Basic Imaging Concepts

Imaging modalities use several physical principles to generate graphical representations, therefore these could be very different, even tough they show the same anatomical structure. It is in general the physician who decides which image modality to use, depending on his/her own judgment, and several other constraints, like device availability, acquisition time, radiation damage, cost, etc. The decision of which modality to use is not arbitrary, but it is based on the current information on the patient health and his/her medical history (anamnesis), and of course the experience and knowledge of the patient.

A physician in the emergency unit will in general rely more on modalities like ultrasound and X-rays, since those can generate very fast images, needed for a fast diagnosis. On the other hand, planned events like pre-surgery diagnostic imaging in general rely in the best possible imaging acquisitions for that particular pathology, and those can last several minutes, even hours.

Before introducing the image modalities, some basic concepts are presented here.

1.1.1 Anatomical Imaging

Anatomical imaging, as the name says, groups the image modalities that show structures of the body. These are used to locate anomalies within the body, like broken bones, internal bleeding, oversized or undersized organs or structures, etc. Ultrasound, radiographs, computer tomography and magnetic resonance are examples of anatomical imaging devices.

1.1.2 Functional Imaging

Functional imaging, in contrast to anatomical imaging, depicts how the body is functioning with respect to certain variable, for example distribution of glucose consumption, blood flow, etc. This imaging approach usually uses a contrast agent, that is then visualized in the final image. Examples of these contrast agents are radioactive sugar (^{18}F-FDG), X-ray opaque agents (iodine, barium) and structure-specific bonding molecules (^{111}In-PSMA). SPECT and PET, described in the next chapter, are examples of functional imaging devices. Since in many cases functional imaging is very difficult to interpret without an anatomical reference, it is today very common to see devices that generate such images bundled with an anatomical imaging device.

1.1.3 Ionizing Radiation

An important criteria for the physician about the imaging device is the probability of harming the patient during the process. Modalities like ultrasound and MR, given the nature of the imaging process, are considered harmless and can be used without major concerns.

On the other hand, modalities that use X-rays for imaging or radioactive compounds can be harmful for the patient[1]. The reason is that the underlying imaging process uses high energy photons (gamma rays or X-rays) or beta particles that can damage the tissues while traveling through the body.

Ionizing radiation comprises particles and photons carrying enough energy to liberate electrons from atoms or molecules, thereby ionizing them. Particles moving at relativistic speed[1] and photons over $4 - 10\,eV$ (ultraviolet or higher) are considered ionizing. See the electromagnetic spectrum chart in Figure 1.1.

There are several types of ionizing radiation, but the most common sources for medical imaging are X-rays or gamma rays, and positrons.
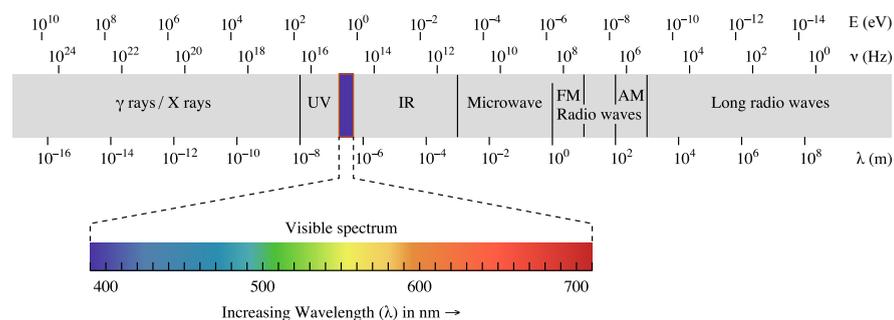
Figure 1.1: Electromagnetic spectrum. The three scales, energy (in eV), frequency (in Hz) and wavelength (in m) are depicted.[2]

1.1.4 Tomography

Tomography comes from the Greek τόμος (*tomos*), slice and γράφω (*graph*), to carve or to write[3].

Tomography is a technique to obtain higher dimensional data (3D, 3D+t) from several lower dimensional projections (2D, 1D). The projections contain partial information from the volume. These projections, with a proper model of the acquisition device and protocol are used to reconstruct the higher dimensional model.

---

[1]Relativistic speed is not strictly defined, but it is considered when Newtonian mechanics are not accurate enough. A reasonable boundary is at about $0.1\,c$, that translates in roughly 1% error in speed and masses using Newtonian physics.

The most known example is the CT (Computer tomography), that uses a rotating X-ray source and detector that rotate around the object of interest, obtaining projections of the volume over a detector, that are then used to compute a 3D reconstruction. Advanced versions of this, including some extra mathematical models, can generate a 3D volume set depicting the movement of the object of interest (e.g. a heart beating) during the projections. This produces a time sequence of 3D reconstructions (3D+t), where it is possible to see the movement in three dimensions.

1.2   Description of common imaging devices

There is an incredible number of imaging devices used for medical purposes and new ones appear constantly in the market, providing new or improved techniques. For this reason, it is impossible to list here all the existing ones. However, a brief description of common devices currently in mainstream usage and will be provided.

1.2.1   Electrocardiography (ECG)

This modality measures and records the electrical activity of the heart[4]. It uses one or more electrodes placed normally on the chest, which measure the variation in the voltage around the place where they are located with respect of a reference one. Each electrode value is then plotted with respect of time, and this response is then analyzed by the physicians.

Since the cardiac signal is quasi-periodic [5], and it is possible to discern the cardiac phases from the electrode plots, the ECG is also used in different imaging modalities, like CT or MRI as a synchronization signal (ECG gating) to reduce the artifacts generated with the cardiac movement.

An ECG plot can be seen in Figure 1.2.

1.2.2   Radiography

Together with the ECG, radiograms are among the oldest imaging techniques used in medicine. The German physicist Wilhelm Röntgen is known as the person who discovered X-rays in 1895, and systematically studied them and its interaction with matter. X-rays penetrate matter and, depending on the energy of each photon, are absorbed with an increasing probability depending on the density and amount of matter they need to cross.

Figure 1.2: Electrocardiogram (ECG), showing the output of several electrodes mounted around the body. *Image courtesy Dr. Ludwig Hennersperger.*

This principle, coupled with a device able to record the X-rays that were not absorbed, generate a 2D image that shows a projection of the inner structure of the imaged object (anatomical imaging). These absorption images (radiograms) are extremely useful for physicians to visualize, for example, a bone fracture, liquid in the lungs, etc., since different components (soft tissue, bone, air) absorb different amounts of X-rays. The main limitations of radiograms are the fact that soft tissues have little difference in absorption among them, so the contrast among them is fairly limited, and the overlap of the structures makes the interpretation more difficult.

A radiography image can be seen in Figure 1.3.

Figure 1.3: Radiography of a bone fracture (right clavicle). *Image courtesy Klinikum Bogenhausen.*

Radiographs are classified as ionizing radiation imaging devices, since they use an X-ray source for imaging.

1.2.3   Computer Tomography (CT)

Computer tomography is a technique used to obtain 3D tomographic reconstructions using X-Ray projections. The first device was originally developed by Sir Godfrey Hounsfield, and the fundamental design remains the same until today.

The machine consists on an X-ray source and a planar detector (currently semiconductor detectors, replacing film that was used in the past), with the object to be imaged in between those. The source and the detector rotate around the object, capturing several X-ray projections. This projections are transferred to a computer, where they are used to compute a tomographic reconstruction of the object. The mathematical part of this reconstruction process is explained in chapter 4, and a sample image can be seen in Figure 1.4.

The computer tomography was the first technique that allowed the physicians to see the body of the patient and understand a particular anatomy before surgery, allowing them to make a much better diagnosis, anticipate possible problems and to plan accordingly. It generated a revolution in diagnostic medicine, and its mathematical principle was adapted for several other modalities, including SPECT and PET.

The main disadvantage of CT is the radiation that the patient receives. Compared to a standard radiogram, a CT delivers about 100 to 400 times higher radiation dose, mainly because it takes this same number of projections, and each projection is a normal radiogram. This can be reduced by using better reconstruction algorithms that can cope with lower dose radiograms.


Figure 1.4: 3D rendered model of a CT data, showing the same bone fracture as Figure 1.3. *Image courtesy Klinikum Bogenhausen.*

1.2.4   Ultrasonography

Ultrasonography is an imaging technique that uses high frequency mechanical waves (commonly $2 - 18\,MHz$) [6] that are induced in the region of interest from the skin. The tissue response time and amplitude (known as echo) is recorded and used to form an image.

Figure 1.5: Prenatal diagnostic ultrasonography. *Image courtesy Dr. Andrea Wendler.*

The most common image configuration is called "B mode", where an array of transducers is used simultaneously to generate a beam, that produces a 2D transverse image, as seen in Figure 1.5. The frequency selected is usually a trade off between resolution and image penetration. The fast image generation process is used to produce one ultrasound image after the other, making this a 2D+t modality. It is also possible to use a moving 1D detector (with respect of the probe) to sweep and generate a spatial sequence that can be seen as a 3D ultrasound (assuming no patient movement). A newer generation of 2D ultrasound arrays allow the direct acquisition of 3D data, without artifacts produced by the movement of the patient.

It is also possible to measure if a certain structure is moving towards the probe or away from it using the Doppler principle. This is used, for example, to see if there is flow in an artery during an emergency diagnostic procedure.

The main advantage of ultrasonography is the ability to create anatomical images extremely fast, cost-effectively and without harming the patient or physician[7].

The most important disadvantage is the fact that the images are subjective, and need an expert to interpret them correctly. Other issues are the need of very complex (and usually proprietary) imaging filters to improve the image quality. The quality of the images is

particularly low when there is a mismatch in density of the tissue, like bones or gases (lungs or gastrointestinal track). The fact that the transducer needs to be in contact with the skin and needs a certain pressure leads to deformation in the anatomy, so it is also difficult to combine with different imaging modalities.

1.2.5   Magnetic Resonance Imaging (MRI / MR)

Magnetic Resonance Imaging is an anatomical imaging technique that is used to visualize soft tissue in the body with much higher contrast than a CT.
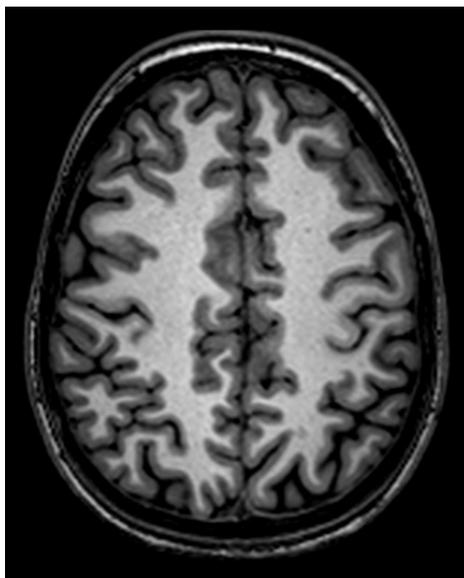
Figure 1.6: Slice of a brain MR reconstruction. The MR acquisition protocol produces a reconstruction capable of differentiate the gray and white matter. *Image courtesy Aslı Okur-Kuru.*

The working principle is fairly complex, but the important part is that the image is constructed using strong magnetic fields that align the spin of certain nuclei (mainly ^{1}H, ^{13}C), and the recovery time from this excited state to its relaxed state is measured and used to generate the images. This means that it is possible to obtain 3D reconstructed data from those signals without the need of ionizing radiation. It is also possible to image different nuclei by varying the acquisition protocol (frequency and amplitude of the magnetic fields) and combine the acquisition to maximize the contrast between the desired tissues (e. g. the gray and white matter in the brain). An example of a brain magnetic resonance can be seen in in Figure 1.6.

The main disadvantages of the MR devices are the long and noisy acquisitions and the running cost of the devices. Metallic devices have to be kept away, and persons with pacemakers or ferromagnetic implants cannot be safely scanned.

1.2.6 Scintigraphy

This modality produces a 2D projection image, similar to a radiogram, but generating a functional image. The image is generated by administering a radioactive compound to the patient, waiting a known time until the tracer reaches the organs or regions of interest and collecting the photons emitted by the compound in a 2D detector. The compound and the administration method (oral, intravenous) depend on the type of image required.

To give a rough anatomical reference in scintigraphy, it is common to put the patient between the detector and a low dose flood-source (flat radioactive plate), so it is possible to see the silhouette of the patient in the final image, as shown in Figure 1.7a.
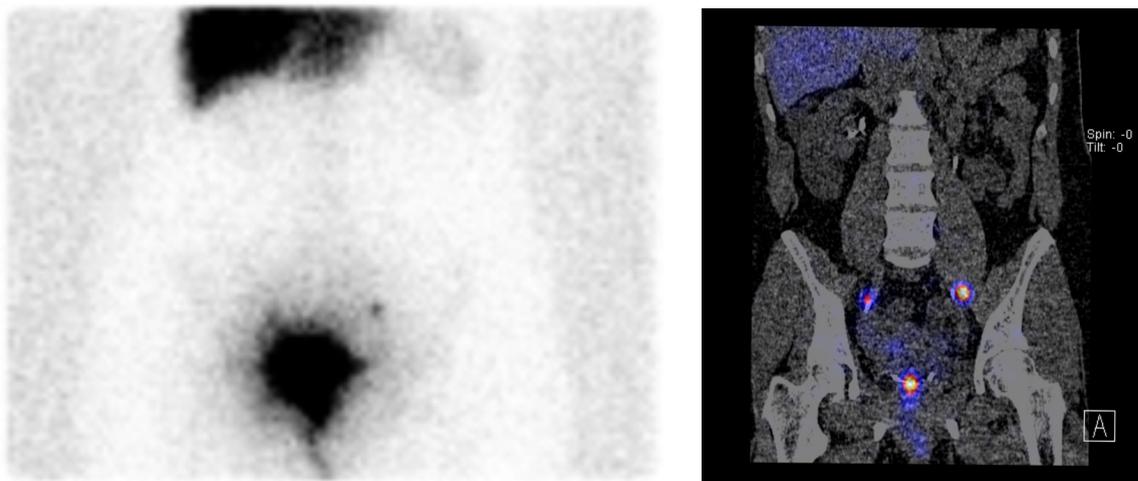
(a) Gynecological scintigraphy. The anatomical silhouette is visible by putting a planar low dose radioactive source behind the patient.

(b) Slice of a SPECT image overlaid with anatomical information from CT.

Figure 1.7: Scintigraphy and SPECT slice from a patient. In both cases, the injection site, two visible lymph nodes (upper right and upper left of the injection site) and the liver are seen. *Images courtesy Klinikum Rechts der Isar.*

Since the physical principle is based on a radioactive compound, this imaging technique is classified as ionizing.

1.2.7 Single photon emission computer tomography (SPECT)

SPECT is a tomographic modality, similar to CT, but it generates a functional 3D volume of a radioactive compound administered in the same way as in scintigraphy. Similar to the CT, the way to generate the volume is to acquire several projections by rotating a 2D detector around the patient. This projections are then used together with the mathematical model of the device to generate a tomographic reconstruction.

Since this 3D volume is functional, sometimes it is very difficult to understand the image without anatomical references (bones, organs, etc.). It is currently a standard practice to

combine the SPECT device with a CT, to provide this reference in one machine, called SPECT-CT.

It is also possible to improve the reconstruction quality of the SPECT image by using the acquired CT image as a map to correct for the absorption of the of the gamma rays in the body.

A SPECT-CT image of the same patient as shown in Figure 1.7a is shown in Figure 1.7b.

1.2.8   Positron emission computer tomography (PET)

Positron emission tomography produces a 3D tomographic volume, similar to SPECT. The main difference here is the imaging principle. For PET, a positron emitter compound is used as contrast agent. When the radioactive nucleus decays, a positron is emitted, and this travels a short distance before it interacts with an electron, annihilating and generating two $511\,keV$ gamma rays travelling in (almost) perpendicular directions.

The PET machine has then a ring with detectors around the patient. The ring collects the $511\,keV$ gamma rays with a very high temporal resolution. This allows to infer when two detected events might come from from the same annihilation, narrowing down its location to (approximately) a line segment that connects both detections.

Similar to SPECT, PET devices are also currently integrated with a CT device to provide anatomical reference of the generated images. Very new devices are coupled with a magnetic resonance device (MR) instead of the CT. A PET-MR image is shown in Figure 1.8.
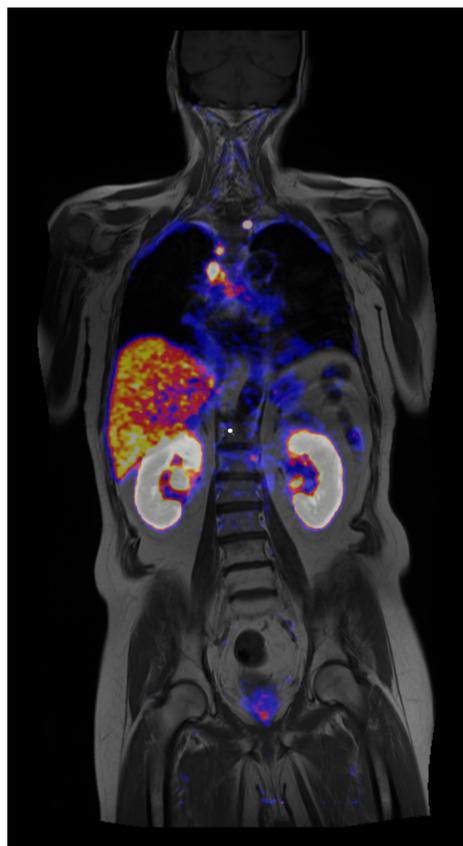

Figure 1.8: Slice of a PET scan, using MR as an anatomical reference. *Image courtesy Klinikum Rechts der Isar.*

Chapter 2

Medical Nuclear Imaging

Medical nuclear imaging comprises all kinds of contrast-based (functional) imaging techniques where the contrast is a radioactive compound (radiotracer). The objective of nuclear imaging is to generate functional images to help physicians in their diagnosis, as guidance inside the operating theater, follow-up of treatments, intervention planning, etc.

2.1 Why Nuclear Imaging?

Nuclear imaging uses radiotracers that are, by definition, radioactive, and therefore generate ionizing radiation, which is harmful for living beings. So the question is, why would you use a technique that is harmful for the patient?

The reason is that most of the time there is no other way to visualize the problem of the patient. The physician is then the responsible of weighting the potential damage to the patient versus the positive outcome the visualization might provide. Nuclear imaging is then used when there is a reasonable suspicion of a disease that cannot be imaged by a harmless technique, and where the result of a better diagnosis outweighs the possible radiation damage.

To understand the reasons why radioactive compounds allow the visualization of processes, it is necessary to look at their chemical properties. In general, the molecules used for carrying the radioactive atoms have a small molar mass, so the molecules themselves are very compact, so they can have higher affinity and penetrate tissues with a much higher probability. This makes the needed amount of tracer injected to the patient lower, since a higher percentage of it will end at the correct location for imaging.

The second reason is the high sensitivity of the PET and SPECT machines, down to $10^{-12}mol/L$, compared to other modalities suitable for humans, like MRI, that work in

the order of $10^{-3} - 10^{-5} mol/L$[8]. That makes the amount of tracer required in the region of interest much smaller than other modalities. These two concepts, higher penetration and high sensitivity, align with the principle stated by De Hevesy, the father of the isotope tracer and nuclear medicine, that the concentration of the tracer has to be as small as possible in order not to interfere with the biological processes.

2.2   How Nuclear Decay Works

Nuclear imaging works by capturing and recording information carried by the particles emitted from the decay of radioactive nuclei, or the subsequent interactions of those particles.

A radioactive nuclei is an unstable nuclei, that can spontaneously transform into a more stable one, and emit particles (for example alpha and beta radiation) or photons (gamma radiation, characteristic X-ray lines) in the process [3].

The time it takes for a nucleus to produce an emission is unknown, and it is not possible to influence it. This means that, if you have many molecules with a radioactive atom, all of them will decay independently. If the radioactive atom is the same in those molecules, the probability of each atom to decay within certain time is also the same. This is the concept behind a Poisson process, and it has been successfully tested that nuclear emission in fact follows this definition[9].

Since it is possible to understand the nuclear emission as a Poisson process, it is possible to measure certain properties. The most important one is the concept of half life, that is the time it takes for half of the sample to decay. This time is constant for a certain isotope and it can go from millions of years (e.g. ^{180}W) to fractions of attoseconds (e.g. ^{12}O). Useful radioisotopes in Nuclear Imaging usually decay in the range of a minute to a couple of days.

The equation that describes the behaviour of the radioactive decay is the following:

$$N(t) = N_0 \cdot e^{-\lambda \cdot t} \tag{2.1}$$

where N_0 is number of radioactive atoms at $t = 0$, N is the number of atoms at time t, λ is the decay constant of this particular isotope and t is the time passed since there were N_0 atoms. From the equation it is possible to see that the number of radioactive atoms decays exponentially, and if one takes two points in time it is also possible to calculate the average decays per second. The SI unit to measure the decays per second is the Becquerel (Bq).

An example to understand the importance of this equation is the following: Let's assume we have in $t = 0$, 10^{12} atoms of ^{99m}Tc. The half life of ^{99m}Tc is $6.01\,h$. One second after

we will have roughly $9.99968 \cdot 10^{11}$ atoms. So, approximately $32 \cdot 10^6$ atoms decayed in the first second, that means we have an emission of $32\,MBq$.

In practice, with the measurement of the number of decays within certain time and knowing the half life of the isotope, it is possible to know how many radioactive atoms are still in the sample, how much it was originally, how long is is necessary to wait until the emission is below certain threshold (e.g. to safely dispose it), etc.

With several emission counts on time it is also possible to fit the exponential decay curve and calculate the half life of the isotope.

2.3   Nuclear Emission

There are basically two types of nuclear emission, particles and energy. Particle emission, as its name says, occurs when an atom expels a particle. This particle can be, depending on the atom, alpha, beta, proton or neutron. Those emissions occur in the nucleus of the atom, and this changes the nuclear configuration. For example, when a ^{18}F atom emits a β^+ particle, a proton is converted into a neutron, and the resulting atom is ^{18}O.

The energy emission consist when an atom emits a photon. This photon can come from the nucleus of the atom, in the form of gamma rays(γ). There are other sources of gamma rays not coming from the atomic nucleus, which are also described below.

A brief description of the most common emissions can be summarized as it follows:

2.3.1   Alpha emission

Alpha radiation consists in the expulsion of two protons and two neutrons bound together. Alpha particles are highly ionizing, and have very little penetration. For example in the case of ^{223}Ra, used in bone cancer treatment, it emits alpha particles with a kinetic energy of $5.78\,MeV$ and a speed of $0.05\,c$. That penetrates less than $100\,\mu m$ of tissue. Alpha particles inside the body are extremely destructive, so they are not used directly for imaging.

2.3.2   Beta emission

Beta radiation consists in electrons (β^-) or positrons (β^+) expelled from nuclei at high speed. When an electron is expelled[1], it travels a short distance (relatively low penetration) until most of its kinetic energy is dissipated, and then is absorbed by an atom. There is

---

[1] An electron antineutrino $\bar{\nu}_e$ is also expelled and one neutron is converted to a proton in the nucleus.

not much of this process that can be used to generate images, except the bremsstrahlung radiation, explained later in this chapter.

On the other hand, when a positron is expelled[2], it also travels a short distance dissipating its kinetic energy, until it interacts with an electron. This interaction between the positron and an electron is called annihilation, where both particles are destroyed, and two gamma rays of $511\,keV$ are created, which travel in opposite directions[3]. It is possible to detect those gamma rays, as explained in the next section, and generate an image of the distribution of the annihilations.

2.3.3   Gamma emission

Gamma emission from a nuclei can occur in conjunction of other alpha and beta emission, or as a product of the transition of a higher energy state to a lower energy state in a nucleus. A gamma ray (γ) is a high energy photon (same as an X-ray), so it travels at the speed of light, and its energy depends on the specific decay process.

Examples of those are ^{90}Y, that decays to ^{90}Sr emitting a β^- (electron), and in $0.01\,\%$ of the cases it produces also a $1.7\,MeV$ γ-ray. Also, ^{99m}Tc decays to ^{99}Tc with a $140\,keV$ γ-ray.

The penetration of gamma rays depends on their energy, but in the range used in nuclear imaging (roughly between $30\,keV$ to $511\,keV$), most of them can pass through the imaging body and be detected outside. This property makes gamma rays the most suitable type of radiation for nuclear imaging.

2.3.4   Other Gamma sources

There are three other gamma sources suitable to be detected that don't come directly from a nuclear emission, but they are directly (generally) related to a previous decay. The first one is the β^+ annihilation mentioned in the beta emission, that is used as operating principle of the PET machines (briefly described in the previous chapter).

The second source of γ-rays is called bremsstrahlung (decelerated radiation), and it is produced on every interaction of β particles while losing their kinetic energy after being expelled from the nuclei. This radiation can have any energy between zero to the current kinetic energy of the particle, and there can be several interactions before the electron is absorbed

---

[2]An electron neutrino $\bar{\nu}_e$ is also expelled and one proton is converted to a neutron in the nucleus.

[3]It is important to consider here the conservation of energy and momentum. The directions are exactly 180° and the energy of the gamma rays is $511\,kEv$ if the sum of their kinetic energies before the annihilation is zero. If that is not the case, the kinetic energy will be distributed between the two gamma rays.

or the positron is annihilated, which makes this radiation difficult to discriminate against background radiation. It is anyway possible to use this radiation to generate images in certain medical interventions[10].

The third source of gamma emission occurs in a process called electron capture and it occurs in atoms with excess of protons (compared to neutrons). The mechanism is basically that the nucleus absorbs an electron, and converts one proton to a neutron, emitting a photon (gamma ray)[4]. ^{123}I is an example of electron capture, decaying to an excited state of ^{123}Te, that immediately decays to an stable state, producing (87% of the time) a $159\,keV$ gamma ray.

Since the electron is usually taken from the lower energy orbit, there will be a missing electron there. One electron from a higher energy orbit will eventually drop to this free orbit, releasing part of its energy in the form of a photon. Since the energy of the electrons is fixed in the orbits, the change from one orbit to another also produces a photon with a fixed energy. Those photons are called characteristic X-rays, and are not necessarily produced by an electron capture process, but by any process that removes an inner shell electron.

2.4   How Nuclear Detection Works

As presented in the previous section, most of the time in nuclear imaging the goal is to detect gamma rays. The detection is done mostly from outside the body with special detectors placed around it.

When a gamma ray is expelled, it travels in a straight line until it interacts with matter. The interaction itself is also a probabilistic event, and the main variables that control this probability are the energy of the gamma ray, the atomic mass of the atoms in the material, the density and the length traveled by the ray through the material. The correct physical equations are outside the scope of this chapter, but in general a good rule of thumb is that the higher the energy of the photon, the longer it needs to travel in a material to interact, and at the same time, the higher the nuclear density of the material, the higher is the probability of interaction with a gamma ray. Table 2.1 shows some concrete examples of this relationship.

Some materials are particularly interesting for the detection process, since it is possible to measure when an interaction occurs. The NaI(Tl) is an example of this. This compound is a crystal that emits visible photons when there is an interaction with a gamma ray (this interaction is called scintillation). NaI(Tl) is a transparent crystal, so it is possible to count

---

[4]and an electron neutrino $\bar{\nu}_e$ is expelled

Energy	Water	Lead	NaI(Tl)
$50\,keV$	$164\,mm$	$0.1\,mm$	$0.3\,mm$
$100\,keV$	$272\,mm$	$0.3\,mm$	$1.0\,mm$
$150\,keV$	$251\,mm$	$0.6\,mm$	$1.8\,mm$
$200\,keV$	$234\,mm$	$1.0\,mm$	$3.2\,mm$
$300\,keV$	$217\,mm$	$2.5\,mm$	$7.7\,mm$
$400\,keV$	$211\,mm$	$4.5\,mm$	$13.8\,mm$
$500\,keV$	$210\,mm$	$6.7\,mm$	$20.7\,mm$

Table 2.1: Required thickness of a material to absorb 50% of gamma rays, versus gamma ray energy.

the visible photons by coupling it with a visible light detector (photomultiplier), making a gamma ray detector. There are also more advanced detectors, like CdZnTe crystals, where the crystal itself is part of the measuring circuit, and the detection is done directly (without generating visible photons).

Having a device able to detect gamma rays is half of the problem, the other half is to know from where those photons come, to be able to generate images that show how the radioactive compound emitting the gamma rays is distributed. To solve the problem of knowing where the gamma rays come from, there are basically two approaches. In scintigraphs and SPECT machines, the detectors are shielded with a material capable of absorbing the gamma rays, and a small window is left opened, so the detected gammas come from the side of the opening. This is called collimator, and the size of the opening defines how precise the localization of the photons is (a small opening means that the gammas can only come from a certain angle) and also how many gammas can be detected (a bigger opening means more gammas can reach the crystal and be detected). There is a compromise to be made when selecting the aperture geometry, between resolution, sensitivity and the amount of material needed for blocking the unwanted gamma rays. PET Machines use a different principle, to be discussed in Section 2.6.3.

2.5   Imaging the functional behavior

The objective of having a functional imaging modality is to visualize information that is not really available in an anatomic image, or not easily discernible. For that, the chemistry of the process to image is vital. Once the process is understood, it is then necessary to design a chemical compound that can take part of the process and, at the same time, it does not interfere with the process.

For example, ^{18}F-FDG (Fludeoxyglucose 18F) is a radioactive glucose analogue which, once injected in the human body, is consumed as glucose, so it will be more concentrated on

places where there is high energy consumption. It is normal to expect higher concentrations in the brain, liver and bladder (superfluous glucose gets disposed by the body over urine and as a result accumulates in the bladder), but if some other areas show higher concentrations of ^{18}F-FDG, the reason should be studied.

Some other compounds, like ^{68}Ga-HBED-PSMA (hydroxybenzyl ethylenediamine) or ^{111}In-PSMA, Prostate-Specific Membrane Antigens[11][12], as their names say, are radioactive molecules designed to attach specifically to the prostate membrane, and with even higher probability[13] to cancerous tissue in the prostate. The first molecule has a ^{68}Ga atom that is a β^+ emitter (suitable for PET imaging), whereas the ^{111}In, present in the second molecule, decays by electron capture with γ emissions at $171\,keV$ and $245\,keV$, suitable for SPECT imaging.

2.6 Imaging machines

As named before, the most common machines used in nuclear medicine are the Scintigraphy, SPECT and PET. The working principle of those devices is presented here.

2.6.1 Scintigraphy

This machine, as briefly mentioned in Section 1.2.6, produces 2D images of the radioactivity distribution of the observed region, and it is used mainly as a diagnostic machine. The device has various names, gamma camera, scintillation camera or Anger camera, and the technique itself is called scintigraphy. This name derives from Latin "scintilla", that means spark.
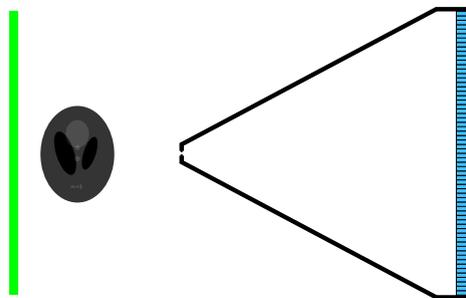
Figure 2.1: Schematic drawing of a scintigraphy. The machine comprises an array of crystals (cyan), surrounded by shielding in the form of a pinhole collimator (black). The Shepp-Logan[14] phantom represents the region of interest, and the green bar represents the low dose background radiation source, to generate the silhouette.

The machine has a two-dimensional array of gamma detectors with collimator(s), that is big enough to cover the region of interest in the patient. In general, they use scintillation

crystals coupled with photodetectors. In many scintigraphy machines it is also possible to exchange the collimator array, depending on the needs of the acquisition. In general, a more restrictive collimator gives better spatial resolution, at expense of the sensitivity (less collected photons).

The patient is positioned next to the detector, and the camera remains static with respect of the patient during the acquisition. The result is a 2D projection of the radioactive distribution within the region of interest.

The images obtained from nuclear devices have no anatomical references, so that in order to provide some, a low dose planar source is usually positioned on the other side of the patient, to produce a small amount of events in the detector forming a silhouette that the physician can then use as reference.

A schematic drawing is presented in Figure 2.1

2.6.2　SPECT

The acquisition principle of a SPECT machine is very similar to the scintigraphy one. There is one or more planar 2D detector arrays positioned around the patient, able to acquire a 2D projection of the radioactive distribution. The first big difference is that the detectors rotate around the patient, obtaining sequences of 2D projections from different perspectives. Those projections, plus information on position and orientation, are then used to generate a 3D reconstruction of the radioactive distribution within the patient. The reconstruction process is described in Chapter 4.
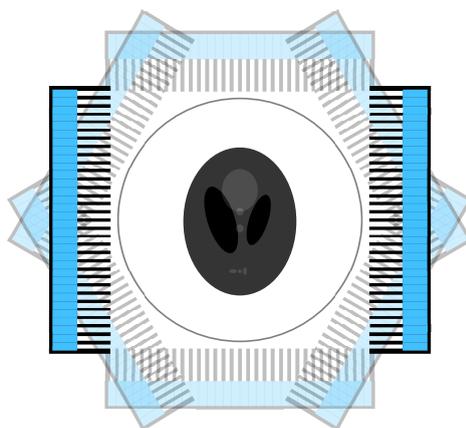
Figure 2.2: Schematic drawing of a SPECT machine. The detectors, two planar surfaces (in this case) depicted in cyan are shielded and have parallel hole collimators (depicted in black). The maximum scanning area is represented with a grey circle and inside is the region of interest, depicted with the Shepp-Logan phantom[14]. The shadowed versions of the detectors show how they are rotated around the patient.

The patient, in general, can also be moved with respect to the detectors, so it is possible to cover a bigger region of interest than what the detectors can see. To provide anatomical references, SPECT machines today come with an integrated CT machine.

A SPECT machine is mechanically much more complex than a scintigraphy one, since it needs to rotate around the patient. This extra complexity (and cost) is compensated with the ability to produce 3D distribution images, that provide the physician with better information, which consequently help to take better decisions. In many routine scans SPECT machines have displaced scintigraphy as the default imaging device.

A schematic drawing is presented in Figure 2.2

2.6.3 PET

PET machines, in a similar way to SPECT machines, produce a 3D reconstruction of the radioactive tracer distribution inside the human body. The tracers are different, and the machines and physical principles are also different.

The objective of the PET machine is to detect a pair of gamma rays coming from the same positron annihilation. Those photons have an energy of $511\,kEv$ and travel in (almost) opposite directions, at the speed of light.
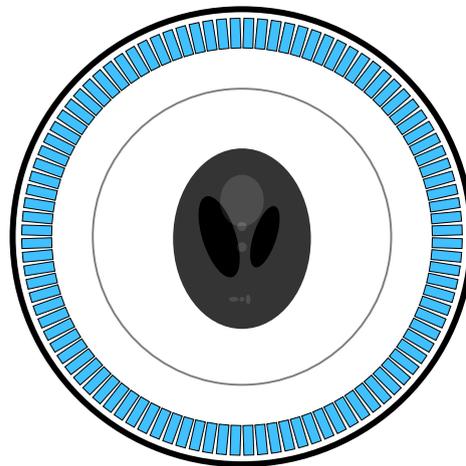
Figure 2.3: Schematic of a PET machine. In this drawing, the crystals are arranged forming a circular ring. The region of interest, depicted with the Shepp-Logan phantom[14] is inside the maximum scanning area (depicted with a grey circle).

The idea behind this is that it is possible to narrow down the volume from where the annihilation occurred to a straight line between the points where the gamma photons were detected, without needing to use collimators. Moreover, if the detection of the gamma photons can be constrained in time (the detection process is done in some fractions of nanoseconds), the line itself can be reduced to a shorter segment. In practice, the line is a

volumetric structure, since it is not really possible to pinpoint the detection of the positions exactly.

The PET machine itself comprises one or several crystal rings forming a crystal cylinder. The crystals can be segmented or continuous, and can also have more than one level of depth (one cylinder crystals within a second cylinder of crystals). One big advantage is that no collimator is needed in PET, so no photons are lost in the collimator.

In general, PET machines are simpler in terms of mechanical construction (no need to rotate the detectors around the patient), but the detectors and the electronics associated are much more complex.

A schematic drawing is presented in Figure 2.3

Chapter 3

Freehand SPECT Imaging

3.1 Introduction and background

The contributions presented in this thesis are based in the technology developed by a group of researchers at CAMP (computer aided Medical Procedures) and the department of Nuclear Medicine of Klinikum Rechts der Isar, starting in 2005[15]. Part of this group of researchers founded their own company, SurgicEye GmbH[16], and multiple ideas, problems and solutions were discussed among those three groups. Several people also worked in both academical and industrial groups, so sometimes is difficult to separate what was done in one or the other.

This chapter explains the technology itself and, except where indicated, it applies both to commercial devices manufactured by SurgicEye and to the experimental setups done in Klinikum Rechts der Isar and CAMP.

3.2 Motivation

SPECT-CT machines provide a very useful anatomical image (CT) registered with a functional nuclear modality (SPECT). Those combined images are used on a daily basis by physicians to perform diagnosis and therapy monitoring/follow-up on patients[17][18].

In the intra-operative scenario, in contrast to the diagnostic imaging, it is not really possible to have big sized machines, since the working space is limited and the surgeon needs room to perform the intervention. In the same line, for complex procedures the surgeon needs guidance systems that are able to provide fast and reasonably accurate results. Diagnostic image quality is not really needed, since the diagnosis was already done at this point. However in general new images are needed, since the pose of the patient during the operation is usually

different to that of the diagnostic images. Furthermore, as the intervention proceeds, the anatomy is changed: organs are displaced, tissue is removed, blood vessels are rerouted, etc. This translates in movement of the internal structures and soft tissue deformation[19][20]. Additionally, the surgeon may need to be sure that the resection volume was completely taken out, so additional imaging during and after the procedure would be beneficial, and sometimes it is mandatory.

There are solutions that integrate PET/CT and MR in the operating room, like the Advanced Multimodal Image-Guided Operating (AMIGO) suite[21], but the cost, the amount of space needed, plus the fact that the patient needs to be moved constantly to acquire new images are major disadvantages.

A different approach was presented in 2007 using single pixel directional detectors coupled with optical tracking[22], to generate 3D reconstructions of the radioactivity distribution, much like a SPECT machine, albeit with lower quality but with a much smaller footprint. This setup, named freehand SPECT, resulted in a tailored solution for intraoperative nuclear imaging.

One of the first targeted applications was the surgical procedure called the sentinel lymph node biopsy (SLNB), where the physician try to resect the first lymph node draining a tumor[23]. The main reason is that if cancerous cells are spreading, it is highly probable that the first lymph node(s) that drain to the tumor (the sentinel lymph node(s)) will have some cancerous cells inside (at least for most solid tumor entities). This procedure is done in the case of breast cancer, where the lymph nodes are located typically in the axillary region. This region contains usually 20 to 30 nodes, as shown in Figure 3.1

So, the task for the physician in SLNB is to resect the correct lymph node. In the past, all lymph nodes were resected, but since approximately 20% of the liquid coming from the circulatory system travels back through the lymphatic system, the resection of all the nodules produced liquid accumulation in the arm (lymphedema), that leads to pain and lack of strength.

The traditional way to deal with the identification of the correct lymph node is to inject a traceable solution around the tumor, and follow it to the lymph nodes. The usual compounds are *Methylene blue* or *Patent Blue V*, both commonly running under the name of *blue dye*, and ^{99m}Tc colloid. The advantages of blue dye are its high optical contrast with respect to the tissue and the fact the it is considered innocuous to a vast majority of patients (Patent Blue V is also the food colorant E131). Its main disadvantage is that its visibility is limited when there is even a small layer of tissue on top, so it is not possible to use as a guidance for the resection. Furthermore, blue dye may drain and washout very fast in some patients making its detection only possible within a short window of time.
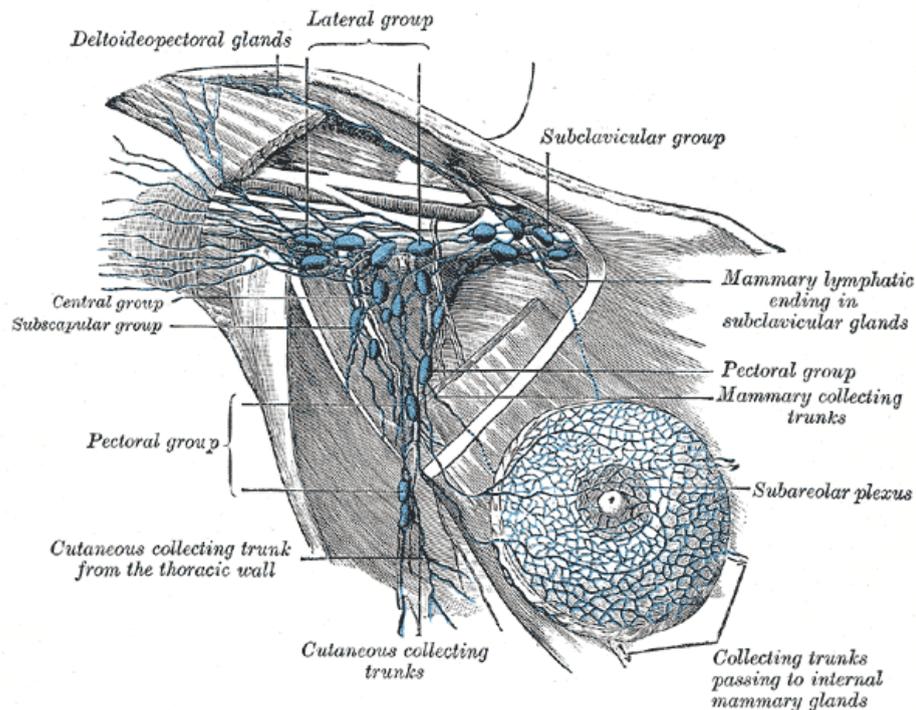
Figure 3.1: Lymph nodes in the axillary region.[24].
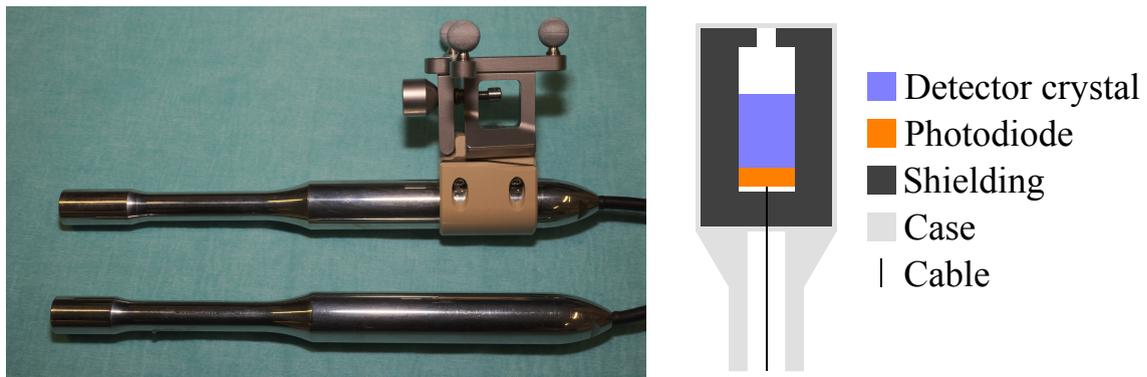
(a) Gamma probes, the topmost with an attached optical tracking target for freehand SPECT

(b) Schematic of the internal structure of the gamma probe.

Figure 3.2: Gamma probe picture and schematic drawing.

The ^{99m}Tc colloids are, on the other hand, radioactive and can be detected before the resection, so it can be used to plan the incision and the extent of the surgical procedure. To detect the radiation emitted intraoperatively, an instrument called gamma probe is used. This gamma probe comprises basically a detector and some shielding, to make the detection directional. An image and an schematic drawing can be seen in Figure 3.2. The gamma probe is typically set to beep every time a certain number of events are detected (one, ten, one hundred, etc.), so the surgeon can use the acoustic feedback to orient the gamma probe in the direction of maximum emission and therefore locate the sentinel lymph node(s) where the radioactive colloids accumulate.

The use of blue dye in combination with ^{99m}Tc colloids as a dual tracer allows the surgeon to navigate using the gamma probe and confirm visually the dyed lymph node. This combination provides high concordance and high identification rate of the sentinel lymph nodes[25].

This procedure provided a first entry point to test the idea of freehand SPECT. The idea is to re-use the current gamma probe detectors, but adding an optical tracking target, so it would be possible to know its position and orientation. Based on this information, 3D images could be reconstructed providing nuclear imaging to the surgeon at any time during the surgical procedure.

3.3   Working principle

3.3.1   Acquisition

The gamma probe is tracked, in this case using an optical tracking, by attaching an optical target to it, as seein in Figure 3.2a. A second tracking target is generally attached to the patient (on top of the sternum for the SLNB) to have a local reference. This allows to re-position the tracking device, in order to improve the acquisition, without losing the previously acquired data. The gamma probe is then used to scan over the region of interest, trying to maximize the coverage from all possible directions. In practice, this is usually done by scanning the surface of the region of interest covering it from different angles to provide enough angular information for the reconstruction of a 3D image.

Figure 3.3: A freehand SPECT cart, from SurgicEye, commercially sold under the name declipse© SPECT.

The acquisition itself requires practice and training to produce satisfactory scans. The scanning quality tends to vary significantly from a novice to an expert and this directly impacts the reconstruction[26].

This is a crucial step, since it is not possible to have a good reconstruction with a bad acquisition. What defines a good or a bad acquisition has been studied[26], and in general the region of interest (where the tumor is) has to be seen from several different angles to

provide enough information for the reconstruction algorithms. A more detailed, mathematical explanation to obtain an optimal acquisition trajectory a priori was developed by Vogel et al.[27][28]

3.3.2 Calibration

To interpret the acquired data properly, it is necessary to measure or model several aspects of the setup.

The tracking system provides the position and orientation of the tracking targets, like the one shown in Figure 3.2a, mounted on a gamma probe. To generate the data needed for the reconstruction, the position and orientation of the tip of the gamma probe, where the detector is placed, have to be calculated. Since the tracking target is rigidly mounted on the gamma probe, and the gamma probe itself is also rigid, their mechanical transformation is fixed, and can be estimated from the mechanical drawings, or directly measured, using the tracking system and a second target mounted on the tip.

The gamma probe produces a signal every time an event is detected, but this in itself is not enough to define its origin, i.e. the position in space where the source that originated the event is placed in space. It is necessary to have an idea of the spatial probability distribution of radioactivity with respect to the detector. This is, given an event detected, what is the spatial probability from where this event could have come from. This probability map can be either inferred from the mechanics of the probe and physical models, simulated by software, or directly obtained using a radioactive point source and measure the response of the gamma detector at different locations with respect of the source.

It is essential that the readings of the tracking and the gamma detector are properly synchronized, since both of them (especially the tracking system) introduce a delay in the measurements. This can be solved using a temporal calibration, in a similar way this is done in tracked ultrasound[29].

3.3.3 Reconstruction

The reconstruction problem is extensive, and therefore will be explained in more detail in Chapter 4, in particular in Section 4.4. Here we will assume the reconstruction process as a black box that gives us the correct radioactivity distribution (a 3D image) given the input of a gamma reading and positions/orientations and the previously acquired or computed models.

3.4   Visualization of reconstructed data

Visualizing the data is particularly important for intra-operative usage, where the physicians are making decisions based on this data. The visualization of the reconstructed 3D images is done by using a co-registered optical camera, so it is possible to overlay the results with the image live view, making it much more user friendly and intuitive.

The camera setup, that consists on one central optical camera and two infrared camera/illuminator pairs on the sides, is presented in Figure 3.4, the infrared system being the optical tracking hardware.

Figure 3.4: Close-up to the optical and tracking cameras from SurgicEye's declipseSPECT.

With this setup it is possible to present a real-time image, coming from the visible camera, with the overlaid reconstructed nuclear imaging. This augmented view, shown in Figure 3.6, enables the surgeon to locate the functional imaging with respect to the anatomical imaging. The idea of combining different datasets in one view is not new[30], in fact that is the concept behind SPECT-CT, PET-CT and PET-MR machines, but the use of augmented reality, that is, co-register and fuse a real time acquired image with precomputed data (in this case, the tomographic reconstruction), provides an easy and intuitive guidance for the physician[31].

### 3.4.1   Intraoperative Use Examples

An exemplary operation theater is shown in Figure 3.5. As it is possible to see, the working space is fairly limited, but the low footprint of the cart allows its use without complications.

A close-up of the area during the scan is depicted in Figure 3.6. On the left, it is possible to see the injection site (the big purple blob), and two lymphatic nodules (the smaller purple blobs), overlaid on top of the optical camera image. The right image shows a second reconstruction after the nodules were resected.

The overlaid visualization is presented in Figure 3.6.

Figure 3.5: Physician in the operation room searching the sentinel lymph node with freehand SPECT[32].

Figure 3.6: Intra-operative freehand SPECT reconstruction, overlaid on the optical camera image. Left: pre-excision, Right: post-excision. The red circles indicates two active nodules in the axilla that were detected and and afterwards resected during the procedure[33].

Chapter 4

The Tomographic Reconstruction Problem

4.1 Introduction

In Physics, problems are modeled as a set of equations that explain the relationship between cause and results (observations). Those problems can, in general, be separated in two classes: the ones that calculate the results given the causes (direct problem), and the ones that try to figure out the causes given the results (inverse problem).

An example to illustrate these different problems would be, if we have an X-Ray machine, where we have a set of equations that model the machine, and we put an object we know its structure. The direct problem is, in this case, to calculate the radiogram (X-Ray image), given the structure of the object, the model of the X-Ray machine, and how the object is placed with respect of the machine. The solution is, in general, unique. However most of the time we have the X-Ray machine, so we can acquire the radiograms, but what we don't know is the object itself.

In this case, going from a set of radiograms and determining the structure of the object is the inverse problem and, most likely, what we want to solve. Those problems do not have unique solutions, some of them have infinite solutions or no solution at all, so the methods used to deal with those inverse problems try to find the best solution, given some extra constraints.

In general, the problem can be modeled as a mapping:

$$r = M(x) \tag{4.1}$$

Where r are the results (radiograms in our example) and M is our model (the mapping) that relates our object x with the obtained results. In theory, what we would like to have is M^{-1}, so we could compute:

$$x = M^{-1}(r) \tag{4.2}$$

In general, M^{-1} does not exist, but there is a variety of methods to cope with that. Those methods can be classified in two groups, analytical and iterative (series expansion) solvers.

4.2   Analytical solvers

Analytical solvers try to estimate a M^{-1} by making some some extra assumptions and also simplifying the model, so an analytical inverse is possible. The big advantage of those methods is that, for a fixed model M, there is a defined inverse M^{-1} that can be pre-computed and just applied to new measurements r to estimate the corresponding objects x. This means, in the calculation of M^{-1} there is no discretization, and this is only done at the very end, when the reconstruction model is implemented.

The advantages of these methods are that the solutions have a closed form, which makes them very fast, and can provide good results under certain conditions. The disadvantages are that they cope poorly with noise (compared with iterative methods) and that only certain problems can be solved with the'se methods. For example, a CT with a gantry and a regular sampling is a solvable problem, but a handheld nuclear acquisition (like freehand SPECT) is not. This is the reason why these methods are not covered here in more detail.

Methods that belong to this category are the Fourier Transform and (Filtered) Back-projection.

One good, state of the art application where these can be used is in low dose CT, as starting point for the iterative methods. Since they give an initial result that is clearly better than a flat zero image, it allows the iterative method to compute faster (iterate less) by starting from an initial approximate result.

4.3   Iterative solvers

Iterative solvers have a different approach to the problem. The idea of these methods is that they start discretizing the problem, converting it into a linear system that is then solved to obtain the reconstruction.

4.3.1   Discretization

What we would like to have is a mapping $f : V \to \mathbb{R}$, that is the physical property we want to estimate in space[34]. Here we assume our property is a scalar, like the radiation distribution in SPECT or the attenuation coefficient of an object in CT.

The first step is to discretize f with a finite set of basis functions b_i:

$$\hat{f} = \sum_{i \in I} x_i \cdot b_i \tag{4.3}$$

Then $\vec{x} = \{x_i\}$ is the quantity we want to reconstruct. A common reconstruction basis is setting b_i as 1 if the coordinates are inside the i-th pixel, and 0 if not.

If we map $\vec{r} = \{r_j\}_{j \in J}$ as the measurements (results) acquired and $\mathcal{M}_j$ as our model that relates $\hat{f}$ with r_j, we have:

$$r_j = \mathcal{M}_j \cdot f \approx \mathcal{M}_j \cdot \hat{f} = \sum_{i \in I} x_i \cdot \mathcal{M}_j \cdot b_i \tag{4.4}$$

If we also map $y_j = \mathcal{M}_j \cdot f$ as our physical measurements, and denoting $a_{ji} = \mathcal{M}_j \cdot b_i$ gives us $A = \{a_{ij}\}_{i \in I, j \in J}$, called *System matrix*.

Our equation system is then:

$$\vec{y} = A \cdot \vec{x} \tag{4.5}$$

Each row of our system matrix corresponds then to one measurement in our system, and each column represents the different contributions of a particular voxel to the results on each measurement. A single element of our matrix corresponds to the contribution of a particular voxel in our dataset to the corresponding measurement.

We try to determine x (using an iterative solver method), and with that we calculate $\hat{f}$ using (Equation 4.3).

4.3.2   Iterative solver methods

There are multiple iterative solver methods, exploiting different properties of the data. The first, and probably most obvious approach, would be to calculate the pseudo-inverse of A. This is commonly done, for example, using SVD (Single Value Decomposition), and the solution is equivalent to $min_x \|A \cdot x - y\|_2$, the least squares solution to the problem. This

solution is, in general, not very good, since the A matrix is ill-conditioned in almost any setup.

There is a large variety of methods to solve those linear systems, but for our type of problem, the most interesting are the ones that take into account the probabilistic nature of the problem. A nuclear emission is modeled as a random, Poisson distributed event[35], and for that there are methods that model this behavior. In particular two of them, MLEM and List Mode EM, are suited for the type of data we need to reconstruct.

Maximum Likelihood Expectation Maximization (MLEM) was introduced by L.A. Shepp and Y. Vardi in 1982[36], and assumes that each voxel in the discretized volume is a Poisson-distributed emitter, which is the accepted model of radioactive emission[9]. The convergence proof of the algorithm was done some years after[37]. The update formula of the iterative algorithm is

$$x_i^{k+1} = x_i^k \cdot \frac{1}{\sum\limits_{j=0}^{J} a_{ij}} \cdot \sum\limits_{j=0}^{J} \frac{y_j \cdot a_{ij}}{q_j} \qquad q_j = \sum\limits_{i=0}^{I} x_i^k \cdot a_{ij} \qquad (4.6)$$

The equation has three multiplicative terms, first one, x_i^k, are the values from the previous iteration (k). The second term can be seen as a coverage factor of the voxel, that is, how well the measurements included that particular voxel.

The third factor is the backward projection, and it is the correction factor applied to the voxel. From this correction factor, $\vec{q}$ the forward projection, or the solution of the direct problem, given the current model we have. So, if $\vec{q} = \vec{y}$, then $\vec{x}^k = \vec{x}^{k+1}$.

A common starting point is $\vec{x}^0 = \vec{1}$, and it is the one used in our experiments. Some other problems, like CT can use different starting points, like the result of a Filtered Backprojection.

List Mode EM (LMEM) was introduced by Barret and Parra[38] and shares several aspects with MLEM, including the modeling of the voxel as Poisson emitters. The main difference in the usage of the algorithms is that MLEM counts the events y of the j-th measurements, whereas List Mode creates a measurement every time an event is received. The update formula in Barret and Parra is not the same as the one obtained by Byrne[39], since they are using slightly different assumptions in their models. In our studies we used Byrne's approach, so our update formula for List Mode is:

$$x_i^{k+1} = x_i^k \cdot \frac{1}{\sum\limits_{j=0}^{L} a_{ij}} \cdot \sum\limits_{j=0}^{J} \frac{a_{ij}}{q_j} \qquad q_j = \sum\limits_{i=0}^{I} x_i^k \cdot a_{ij} \qquad (4.7)$$

The update formulas are very similar, but they have some key differences. The number of measurements J is different in both, since in MLEM this is independent of the number of events (individual detections) received and it is arbitrarily set to, for example, every m seconds at certain position. In the case of List Mode, it corresponds to the number of events acquired. For the same reason, there is no y_j term in List mode, since it is by definition 1.

The second term of the equation has the sum with a different index L. This is to indicate that what is being summed here are in fact not the same values of the system matrix A, but as explained before, this is is the model of the voxel coverage in our acquisition, and therefore independent of the events detected, as in MLEM.

For math inclined readers, a longer description of both MLEM and LMEM algorithms is shown in Appendix A.

4.4 MLEM and LMEM for freehand SPECT

MLEM and List Mode EM algorithms were conceived for PET acquisitions, where you have a fixed gantry. In freehand SPECT, this gantry concept does not really exist, but the algorithms work in a similar fashion, with some small reinterpretations. The descriptions here apply to the workflow used in CAMP and Klinikum Rechts der Isar, and not necessarily to the one followed by SurgicEye.

The key element is the system matrix construction. If we go back to equations Equation 4.6 and Equation 4.7, they are both composed of three multiplicative terms: the first element is the previous iteration, and both the second and third terms are computed in terms of the system matrix.

4.4.1 Modeling the Detector

As described before, a gamma detector consists of a material that interacts with the gamma rays and can measure or count the number of events it detects, and a partial shielding that blocks the gamma rays, so the detected radiation comes from a certain direction.

Since the plan is to measure the incoming radiation in the detector and then produce a 3D reconstruction that shows the correct radiation distribution map, it is extremely important to know, given a detected event, where this might have come from. This is what is needed to model, the spatial probability of the detector.

There are several ways to model the detector(s)[40]. In the case of a gamma probe, with a single crystal and a symmetric shielding around the detector, making a physical model based

on the geometry an the physical properties of the detector and the shielding, is one good possibility. When the model is not that symmetric (e.g. in the case of a slanted probe) or there is an array of detectors, it is possible to simulate emission in many different locations in space, and generate a simulated model.

The third possibility is to directly measure the response of the probe with a radioactive point source. If the point source and the gamma detector are set in a positioning table that can move one with respect of the other, it is possible to exhaustively measure the response.

All three alternatives have advantages and disadvantages. The main advantage of the model is that it is a continuous model, so it is resolution independent, and it is very compact in terms of storage. The main disadvantage is that it does not scale, so it is useful only for detectors that are simple enough. The third problem is that it is a model, so there might be discrepancies between the model and the real construction of the detector.

The second alternative, the simulation, can cope with very complex models, and the quality of the simulation is bound to the amount of computing power available[41]. It is possible to improve the model afterwards, since the calculations are done based on Monte Carlo simulations, where multiple model instances can be combined. The main disadvantages of this approach is that it is very time consuming to set up accurate models to compute, and the computation itself is in general also very expensive. This approach also suffers from the inaccuracies that appear when the model is not equal to the real device.

The third approach, to directly measure the detector, has the advantage that it works for very complex devices, and that since the characterization is based on the real device, there are no discrepancies between the measured device and the real device, because they are the same (the obvious limitation is that applies only to that particular device). The disadvantages of this approach are the need of hardware (positioning table, lead shielding), and radioactive point sources to produce an accurate model, plus a calibrator to know its position with respect to the detector (or some calibration). These are also time consuming (the model of the gamma camera used in the following chapters took about a month to be acquired), and it is difficult to improve an already acquired model, since it is not mechanically easy to re-position the complete setup as it was originally.

Each modeling alternative has a use case. The analytical model makes sense for simple detectors, the simulation makes a lot of sense for the device manufacturers, since they have access to the mechanical model of their own machines, and the direct measurement makes sense when doing a simulation is not really possible, or there is a single one unit to characterize. A combination of the simulation and partial measurement is also used to see if a device is, for example, mechanically misaligned, or if there is an error in the modeling.

To compute the model in the case of the simulation and the measurement, the assumption is that the same number of events N is generated at each sampling point. For the direct measurement, this is not possible to do, but staying on each position the same amount of time (compensated by the decay, if the half life isotope is short), gives the same number of expected events, since the nuclear emission is Poisson distributed.

If the number of detected events on each position is n_i, N_i/N is the detection ratio on each position. This detection ratio, normalized over the sampled volume (assuming no detection outside the sampled volume) gives the spatial probability of a detected emission, that is the model of the detector.

4.4.2   Building a System Matrix

The system matrix is the mapping between the known measurements and the unknown volume. To generate it, a model of the detectors is needed, and also the spatial location of these during the acquisition.

The reconstructions done in Chapter Chapter 6 were done using a geometrical model of the single detector, but for Chapters Chapter 7, Chapter 9 and Chapter 8, a directly measured look-up table model of the 2D gamma detector array (gamma camera), was used.

Assuming the canonical volume discretization presented in Equation 4.3, the system matrix is constructed by projecting the look up table (or the model) corresponding to a certain measurement, from the position and orientation of the detector in that measurement.

In the MLEM update formula, the equalization parameter is basically the sum over the measurements of the system matrix, so once we have a way to compute it, we just need to iterate once over it, summing up the values and storing them. In the case of LMEM, this is a bit different, since our system matrix has an entry every time an event is detected, and this does not represent the acquired volume anymore. To overcome this, the voxel coverage is not computed from the system matrix, but from a matrix equivalent to the one used in MLEM, which uses every position and detector, but only the sum over the positions L is computed.

4.4.3   Filtering the System Matrix

Since the system matrix is generated using the actual trajectory, it is possible that some rows are completely zero. This happens when the detector is in a position that covers no

voxel. Those zero rows are eliminated. Furthermore, a threshold $t_r > 0$ is used also to remove measurements with little relevance, since such measurements jeopardize the numerical stability of the solution.

It is possible to also have voxels that were never covered given the scanned trajectory, i.e. the detector never had them in the field of view during the scan. Those voxels are also discarded, and again, a threshold $t_c > 0$ is used, so voxels with little coverage, prone to generate artifacts, are removed[26].

The third filtering uses a model of the detector, and any voxel that is in contact with the detector is removed. The assumption here is that there is only air in those voxels, so they do not belong to the region of interest, since the detector was physically at that location[42]. The exception here is if there is deformation of the tissues if the detector is pressed against the patient or phantom. If that occurs, this method will certainly produce artifacts the reconstruction, so in general contact between the detector and the scanned body is avoided.

4.4.4   Iterating

With a filtered system matrix in place, the next step is to iterate to get the reconstruction. The question is when to stop iterating. Both MLEM and LMEM converge at infinity, but in general we would like to have a solution in a finite amount of time.

MLEM and LMEM are not considered fast converging algorithms, so in general multiple iterations have to be done to reach a good solution. Those algorithms resolve the low frequency components of the images first, and the high frequency components (finer detail) appear with more iterations.

One would think that the more iterations the better for better reconstruction results, so it would be possible to iterate as much as possible to maximize quality. The problem here is the noise that has mostly high frequency components, and with many iterations it tends to be overestimated, making the reconstruction extremely noisy and generating artifacts.

To visualize the effect of the number of iterations versus the reconstructed image, a set of planes taken from a reconstructed thyroid phantom, shown in Chapter 8, is shown in Figure 4.1.

4.4.5   Filtering the Reconstruction

Since not all parts of the reconstruction converge at the same speed[43][44], it is difficult sometimes to find a good compromise between the overall image quality. A good approach
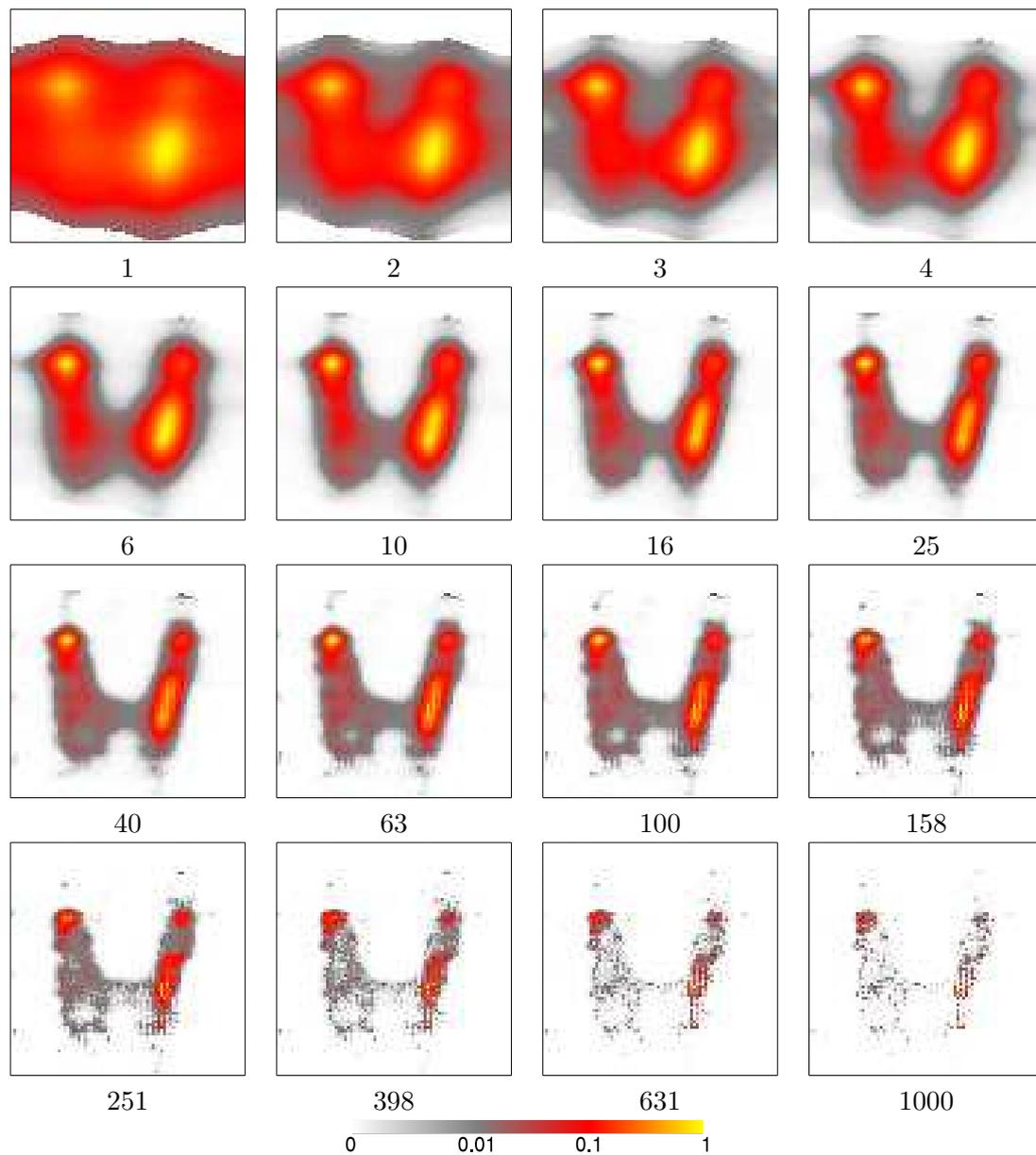
Figure 4.1: Cross section of thyroid phantom versus iterations. The number of iterations are geometrically spaced. A good solution should be able to resolve the cold and hot nodules, without over-fitting the results. A good compromise would be around 63 to 100 iterations.

has been to over-iterate a bit, and then to apply a Gaussian Filter over the reconstruction, with a small kernel[45].

Figure 4.2 shows selected reconstruction cross sections from Figure 4.1, with one voxel Gaussian filter. Structures like the hot and cold nodules are preserved, but most of the artifacts in the reconstruction are removed.
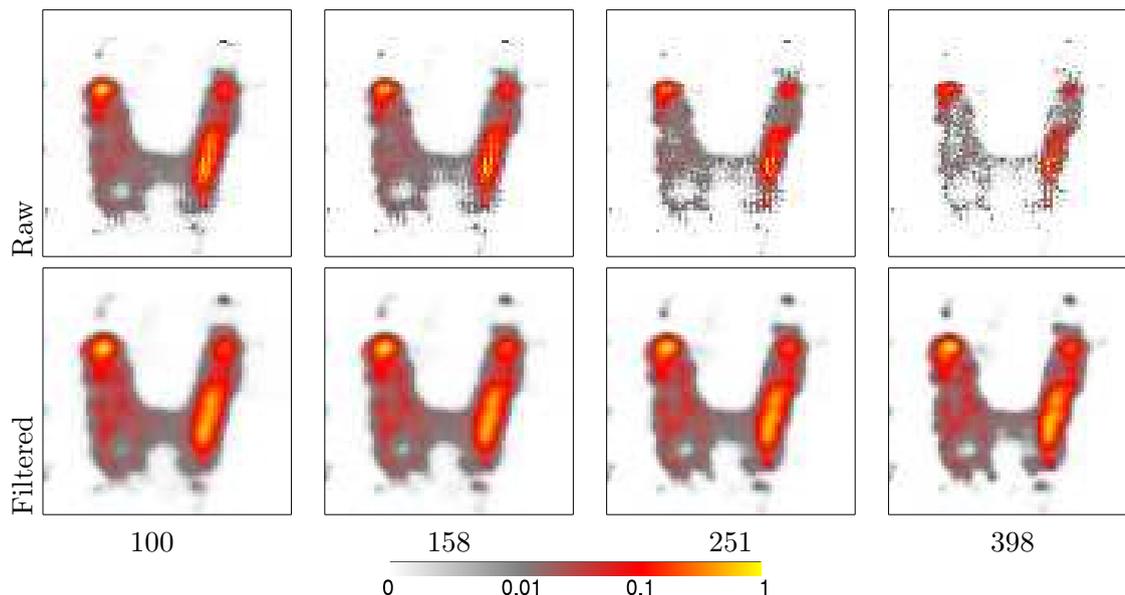
Figure 4.2: Cross section of thyroid phantom versus iterations, selected from Figure 4.1. The results of the first row are not filtered (identical to Figure 4.1) and the results in the second row were filtered with 1 voxel Gaussian blur. Here the best compromise would be around 158 iterations.

4.4.6 Cost

The size of the system matrix (in bytes) is the number of measurements, multiplied by the number of detectors, multiplied by the number of voxels in the case of MLEM, and the number of detected events multiplied by the number of voxels in the case of LMEM, both multiplied by the size (in bytes) of each number. For some measurements, our system matrix was around a terabyte in size, so it is not feasible to have those numbers in RAM. To overcome this, the system matrix was computed "on the fly", so every time a value was needed, it was re-computed.

In general, as shown in Chapter 8, if the acquisition follows a step and shoot pattern, that is, move, acquire without moving and then move again to the next position, MLEM results in smaller problem size. On the other hand, if the acquisition is done moving continuously, LMEM results in a small problem size. The reconstruction quality is fairly similar for both algorithms.

MLEM and LMEM algorithms are good candidates for parallel computation (like general purpose GPUs or Intel MIC devices)[46][47], since they need to be executed for each voxel in the region of interest, and the result of one voxel does not affect the results in the others in the current iteration.

Chapter 5

Robotics

5.1 Introduction

Robotics is a branch of engineering that deals with the design and construction of a very wide spectrum of autonomous and semi-autonomous machines. From very simple x-y positioning table to fairly complex vehicles like the Curiosity Mars Rover in Figure 5.1a, they share the ability of performing complex series of actions (movements, grasping, etc.) in an automated way. This ability, plus the usual high accuracy in their movements and the possibility of programming and/or remote controlling them, makes them ideal to be used in repetitive tasks and in constrained spaces, as well as tasks performed in areas which are dangerous or simply inaccessible for humans.

In several industries, like car manufacturing or food processing, many types of robots are used, in activities such as palletizing, packaging, welding, painting, etc., as the one shown in Figure 5.1b. In general, industrial robots are set to perform fixed sequences of movements, with very little reaction to the environment (in general, limited to safely stopping the process).

On the other extreme of the spectrum, the Curiosity Mars Rover is located in an unknown environment, at approximately 225 million kilometers away, where any command sent from Earth takes several minutes to reach the vehicle, making any type of remote control unfeasible. This makes the Rover a much more complex device, with multiple sensors, a software pipeline capable of evaluating the measurements, defining the risks, making decisions, doing corrections on the fly, an so on.
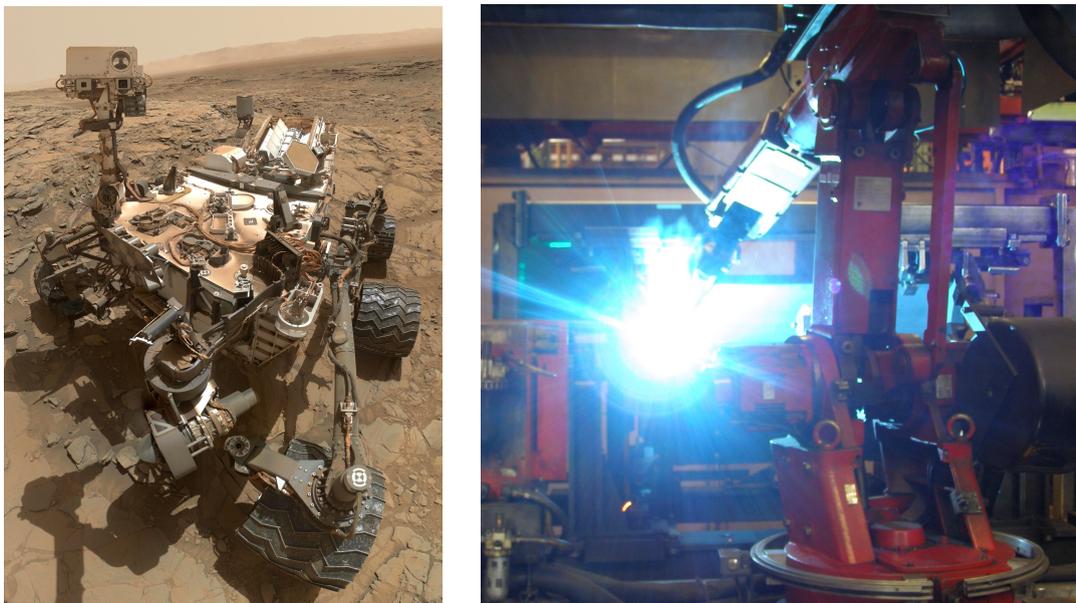
(a) The Curiosity Mars Rover (Copyright NASA).[48]   **(b)** Welding robot at the manufacturing plant of Kinshofer GmbH in Germany.[49]

Figure 5.1: Examples of robots in their environment.

5.1.1   Medical Robots

In medicine, robots can be classified in two categories according to their usage: surgical robots and imaging robots. In both cases, the safety of the patient and medical personnel is the first priority, so systems are designed with this in mind. The majority of medical robots consist of one or several robotic arms, a series of rigid bodies connected by an actuator (see Section 5.2 , with a construction more similar to the industrial robots than the Curiosity Rover.

In general, surgical robots are semi-autonomous machines that perform surgical tasks under full control of by a human. One of the main advantages of having the robot performing the surgery is the fact that it is possible to map the human movements to the robot's, which can be linearly scaled, so large human movements can be translated to small (and very precise) robotic movements. Other advantages are the possibility of filtering the movements and implementing tremor correction, providing a platform to perform remote surgery. The most known example here is the da-Vinci Surgical System, from Intuitive Surgical[50]. This robot is used in minimally invasive surgery, with a separate surgeon console and effector arms, and it is used in prostate surgery, cardiac valve repair, gynecological procedures, eye surgery and throat procedures[51][52].

Projects like ROBOCAST aim to have autonomous robots performing certain surgical tasks, commanded (but not controlled) by the surgeon, like catheter positioning, biopsy sampling, etc[53]. The challenges in this area are multiple, and there is still no solution to many of

those problems, being the most important ones the safety of the patient and personnel, and the guarantee of the outcome.

RoboDoc is an example of the problems an automated surgical robot can have. This system was designed for orthopedic use, it was used for knee and hip replacement and it had a success comparable with surgeons[54][55]. The fact that the system was automated but not perfect ended up in a class lawsuit in Germany that took the product out of the global market.

The second type, medical imaging robots, acquire data to create diagnostic or intraoperative images and support the surgeon making decisions or while performing interventions[56]. Many imaging robots do not need to be in direct contact with the patient, since the adopted modality does not require it, for example, X-Rays or detected radioactive emission[57]. In the case of ultrasound robotics, there is direct contact between the ultrasound probe, rigidly mounted on the robot, and the patient. The robot therefore exerts a force on the skin of the patient equivalent to a human operator.

Imaging robots are less challenging than surgical robots, but nonetheless they have to generate results that justify their operation. The Artis zeego Imaging Suite[58] from Siemens is a robotic C-Arm CT, designed to be used intraoperatively, and can create full body CT inside the operating room without disrupting the workflow.

As an example, freehand ultrasound compounding is a very challenging topic[59], but by replacing the operator by a robotic arm, it is possible to produce a scan with a regular sampling, making better quality ultrasound volumes[60].

This work focuses mainly in robotic arms since, as mentioned early in this chapter, most medical imaging robots are based on this type of construction.

5.2   Modeling a Robotic Arm

A robotic arm, like the one shown in Figure 5.2 can be seen as a series of rigid bodies (links) connected by rotational articulations (joints) or translational actuators. This configuration allows these parts to move with respect to the other (by the rotation or translation of a link with respect to the previous one). The idea of this type of construction is that, by moving this joints or linear actuators, the tip of the robot can be positioned in space (within the mechanical constrains) with respect of the robot base. This joint configuration is known as the robot pose, and the pose describes unambiguously the relationship between the robot base and the tip, while the inverse problem, known as inverse kinematics, does not have

a single solution. The number of joints plus linear actuators in this type of robotic arms define, in general, the degrees of freedom (DOF) the robot has.
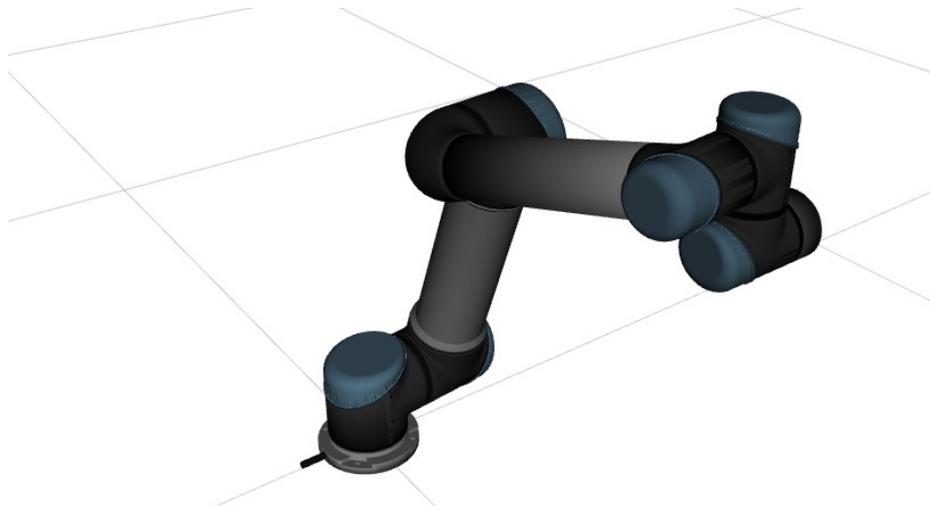
Figure 5.2: Robotic arm inside a virtual planning environment.

Imagining a tool on the tip of the robot (a detector, a scarpel, etc.) that is required to be positioned in space, it can be shown[61] that 6 degrees of freedom are required for an arbitrary positioning, 3 translational (X, Y, Z) and 3 rotational (roll, pitch, yaw), within its dexterous workspace. Therefore, a robotic arm with 5 DOF or less cannot bring an object to an arbitrary position in space, and one with 7 or more DOF can achieve that with multiple poses (redundancy).

This redundancy can be seen also in the human body. It is possible to fix one hand against an object, and move the elbow by rotating the shoulder and the wrist. The final pose, shoulder to hand, does not change even tough it is possible to move the elbow.

To successfully have a robotic arm moving from a pose A to a pose B smoothly, an interpolation between the old and new pose is, in general, not possible, since this could originate collisions between different parts of the of the robot, or with object in the environment. A joint trajectory, that is the succession of joint values in time must be planned taking the cartesian space in consideration, as well as the mechanical limitations of the arm and the environment.

This joint trajectory is computed in several stages, described here:

5.2.1 Kinematics

To understand the relationship between the robot base and its tip, and how the movements of the joints or actuators affect its positioning, a mechanical model of the robot in necessary.

This model is usually derived using the Denavit-Hartenberg parametrization. They proposed the first minimal parameter representation that describes unambiguously the location of joints and links in space[62]. Without going into the details of the modeling, using this parametrization it is possible to construct a series of matrices that, multiplied together, result into the robot base to tip transformation of the current pose, as in the following kinematic chain:

$$[T] = [Z_0] \cdot [X_1] \cdot [Z_1] \cdot [X_2] \cdot [Z_2] \cdot [X_2] \cdot ... \cdot [X_{n-1}] \cdot [Z_n]$$

The matrices $[T]$, $[X_i]$ and $[Z_i]$ are all orientation-preserving Euclidean transformation matrices, that means, they are all $4x4$ matrices, where the upper left $3x3$ matrix is a pure rotation matrix without reflection, the upper right $1x3$ matrix corresponds to the translation in X, Y and Z, and the lower $4x1$ matrix correspond to the vector $[0, 0, 0, 1]$[63].

The origin of the system coordinates for the robot is known as the base, and the tip of the robot is the end of the last link, so the transformation $[T]$ represents the base to tip transformation.

The matrices $[X_i]$ represent the actuators of the robot, and they are dynamically computed depending on the current value of the actuator. Each $[X_i]$ transformation has a single parameter p_i that defines the matrix. In the case of a rotational joint, the angle is the parameter; for linear actuators, the extension of the actuator is the parameter.

$[Z_i]$ represents the mechanical linkage between $[X_{i-1}]$ and $[X_i]$. Those transformations are fixed, and they are computed from the robot design schematics (or provided with them).

The calculation of the transformation $[T]$ (or segments of it) from $[X_i]$ and $[Z_i]$ is known as the direct kinematics.

The inverse problem, known as inverse kinematics, consists on finding one or more set of parameters p_i, that define the $[X_i]$ matrices, to obtain a given base to tip transformation. The parameters have to also fulfill the mechanical constrains of the particular robotic arm, including maximum extension of the linear actuators and limitations in the rotation of the joints.

This problem is clearly more complex, and it is in general solved by a dedicated algorithm. Currently probabilistic sampling algorithms are the state of the art, since deterministic methods cope with the dimensionality of the problem. A robot with 6 joints an 0.1°precision has $(360° * 0.1)^6 = 10^21$ possible configurations in a toroidal space. Actual robots have greater precision, hence it is unfeasible to completely explore the solution space.

There is a long list of ready-to-use solvers, but most of them have multiple configuration parameters for them, so choosing the best for a particular application is not always an easy task. There is also a code generator that is capable of autonomously delivering closed form solutions for generic robots[64].

In general, finding the robot base to tip transformation for 6 degrees of freedom (3 translational and 3 rotational) is a solved problem, but it is also possible to solve for less DOF. This is interesting in the case of a symmetric detector mounted on the robotic arm, where the rotation about its axis is irrelevant (as long as it is known), so finding a suitable solution with less degrees of freedom allow different planning strategies to be used (minimizing movement, dealing better with other spatial restrictions, etc.).

Another example is a distance measuring laser mounted on the tip of the robot. Since those devices are rotation independent, and the total distance will be given by the final robot pose plus the laser measurements, only four degrees of freedom are needed (two rotations and two translations), assuming constant accuracy of the measuring device.

5.2.2   Collision detection

The first step to move the tip of the robotic arm from a current pose A to a new pose B, is to know if this pose exists. Inverse kinematics solve that, but still the new joint configuration can lead to a collision with the environment or with itself. To properly handle this, it is necessary to have a 3D model of the robot, where it is possible to position each segment that connects the joints separately. With this, plus a model of the environment, each possible solution from the inverse kinematics is tested, until a valid one is found.

This is a fairly expensive step in terms of computational effort, that has to be repeated multiple times during the trajectory planning pipeline. A reasonably way to speed up this calculations is to pre-compute the joint ranges where the robot collides with itself, so then it is just a matter of searching if the desired configuration is feasible or not.

The collisions with the environment cannot be pre-computed unless the environment is static, but in general this is a very unrealistic assumption. It is, however, possible to check for collisions at runtime with, for example, an RGBD camera, but it is still needed to re-compute the trajectory and steer the robot away.

5.2.3 Path Planning

Knowing that the current pose A and the final pose B are valid configurations, the next step is to find a path that connects A and B, given that all possible poses in between are also valid. This is an NP hard problem[65][66], but there are multiple heuristic approaches.

One basic approach is to first try to reach B from A by linearly moving the actuators (a straight line in joint space). To do so, the path is divided in very small steps, and the collision detection check is calculated. If that succeeds, a solution has been found. When that is not the case, an intermediate point position M is inserted in the path, and then A to M and next M to B are checked. If both trajectories are possible, a solution has been found. If any of them are not valid, a new pose is inserted within the unresolved interval(s), until there is a solution.

This algorithm does not sound very clever, but in general it finds reasonable solutions for many cases. To improve the outcome, it is possible to try several M poses in parallel. It is common also to add a timeout, so if there is no solution, the current state is discarded, and new M poses are generated.

Several other algorithms are smarter variations of this one, changing how M is selected. Just to mention one interesting variation of this is to find the first collision point C in the path from A to B, and then try to surround the boundary of the obstacle until it reaches a point D that is closer to the distance between C to B. If this is found, then the algorithm continues solving from D and B. If it is not possible to find D, there is no solution available between A and B[67].

It is also important that the planning routines that select the trajectories of the robot avoid the "Wrist flip", the equivalent to a Gimbal lock in robotics (loss of one degree of freedom at singular points). This happens, for example, when two non-consecutive joints are aligned (e.g. 1st and 3rd joint), and the joint in the middle (2nd) needs to spin in 180°to keep the trajectory.

5.2.4 Dynamics (Trajectory planning)

Finding a path connecting A and B is a geometric problem. A path is unfortunately not a complete solution to the problem if there is no time component in the trajectory planning.

A trajectory is a path plus the timing specification of each point. Finding feasible trajectories, so that they respect the physical limitations of the actuators of the robot requires knowledge of the masses and inertias of the system, and other forces involved, like gravity, friction, tip pressure (in ultrasound robotics)[68].

A dynamic modeling of the robot, plus torque sensors in the robot enable the system to implement a fine-grained closed control loop, where it is possible to adjust a path on the fly and, for example, keep a pressure constant (for ultrasound) or move away if contact with the patient is detected.

The robot dynamics are normally modeled using Lagrangian mechanics, since they allow easier modeling of the problem[68]. The optimum solution, so, the fastest trajectory following a given collision-free path, that fulfills the mechanical constrains of the robot (maximum acceleration, torque and speed) is a solved problem[69][70].

The dynamics problem are outside the scope of this thesis, since the robot used for the experiment did not allow the user to interact with their internal model and did not have torque sensors. The only available control is to set a maximum speed and acceleration of the joint rotations.

5.3   Universal Robots UR5

The work presented in the following chapters was done using a Universal Robot UR5, shown in Figure 5.3. This robotic arm has 6 degrees of freedom, can lift 5 kg, and can work in a maximum radius of 850 mm. Our version did not have any torque sensor, so any dynamic calculation operates only in open loop (no feedback).

Figure 5.3:   Universal Robots UR5, with a cart and control unit.

This device is considered a "collaborative robot", meaning that it can operate around humans, since the force it can exert is limited.

The UR5 was designed as an industrial arm, meant to do pick and place. The tip can be positioned with 0.1 mm precision, but no angular precision is stated.

The robot was not designed to be used as a medical device, but for phantom use was perfectly acceptable. It has several big limitations we needed to work around during operation. The

first one is that the two joints, the 3rd and the 5th are physically too close, so it is extremely easy to make the robot self collide. Figure 5.4 shows one common example of a joint configuration that produces a self collision. That also made the algorithms to compute path and trajectory planning very slow, since they had to test a huge amount of combinations until they could find a working solution.

The second limitation was that it is not really possible to abort a command sent to the robot using their real time controller and API, the only possibility is to hit the emergency switch, making an abrupt interruption, and the need to re-initialize the system (and having to rotate all the joints about 60 degrees).

Taking this in consideration, most experiments were designed to have a static setup, where only the robot was moving, to minimize the risk of damage to the robot and the surrounding equipment, including phantoms filled with radioactivity. For the experiments in Chapter 8, the movements were pre-calculated, since the movement speed and the trajectory needed to be carefully controlled.

In total, 13 publications have been written using the results from the experiments performed with this robotic arm in Klinikum Rechts der Isar, 12 of them related to nuclear medicine, and one hybrid robotic tracking and electromagnetic tracking.

Figure 5.4: Close-up to the joint configuration of the UR5 showing a self collision.

Part II

Contribution

Chapter 6

Robotic freehand SPECT using Gamma Probes

6.1 Introduction

The idea of moving towards a robotic solution for freehand SPECT came from the user driven nature of this modality, making it very difficult to systematically repeat experiments. This is particularly complex when trying to optimize the parameters of an acquisition. In this case an experiment has to be repeated several times trying to vary only one parameter at a time (e.g. collimator, scanning speed, etc.), in order to evaluate how a certain variable impacts on the reconstruction results. A robotic arm holding the probe, while keeping the rest of the setup intact, was the starting point of our Robotic freehand SPECT journey.

The first challenge of this experimental modality was to try to reproduce a human scan with comparable quality. The second one was to see if the system as a whole was able to perform decay compensation, that is, to produce equivalent reconstructions after some important portion of the radioactive tracer has decayed.

The work discussed here belongs to the paper "First flexible robotic intra-operative nuclear imaging for image-guided surgery", presented at "Information Processing in Computer Assisted Interventions (IPCAI)" in 2013[71].

6.2 Setup and Methods

The setup used in this chapter consists of a Universal Robots UR5, described in Section 5.3, holding a single pixel gamma detector (HiSens, Crystal Photonics, Berlin, Germany[72])
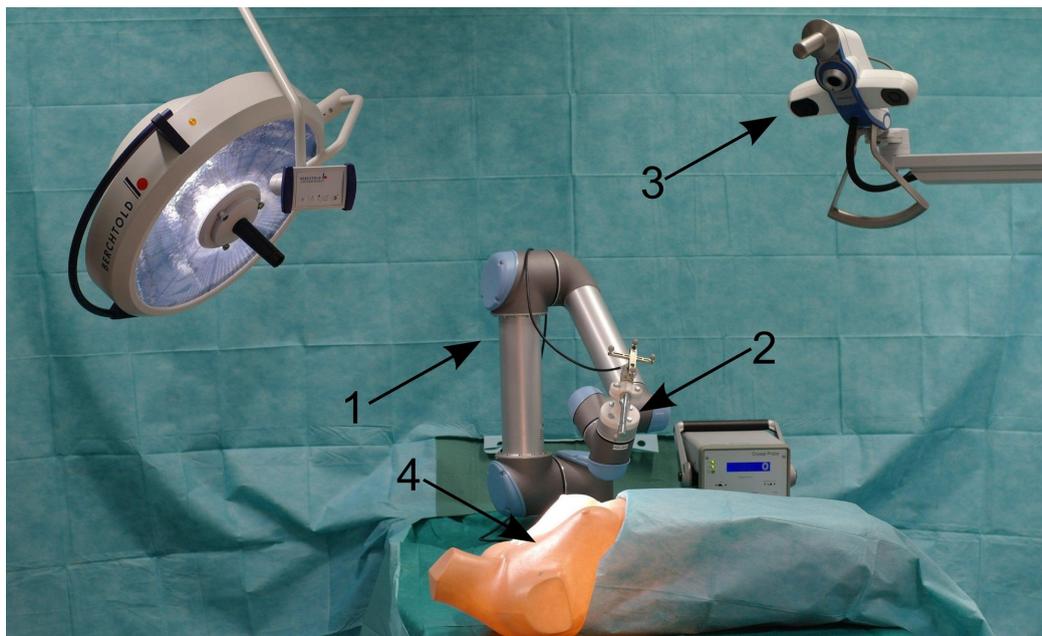
Figure 6.1: Experimental setup showing the (1) robotic arm, (2) gamma detector, (3) optical tracking system and (4) phantom[71].

with their 60° collimator for radioactivity measurements. An optical tracking system (Polaris Vicra, Northern Digital, Waterloo, ON, Canada[73]) and optical tracking targets are used to determine the position of the gamma detector with respect to the region of interest. The setup is shown in Figure 6.1.

Calibration   The spatial relation between the gamma detector tip and the tracking target was determined using a tracked cylindrical calibrator, provided by SurgicEye[16], fitting around the detector. Since the location of the detector crystal inside the probe is known with respect to the tip, the relationship between the tracker of the calibrator and the tracker of the probe allows to compute the tracker-to-crystal transformation. The relation between gamma detector tracking target and robot hand was determined automatically using hand-eye calibration [74][75]. Thus, the relation between the detection crystal and the robot hand can be computed. Figure 6.2 shows the different coordinate systems involved.

Data Acquisition   The events detected by the gamma probe, combined with the tracking information from the optical tracking and the robot, and the timestamps are recorded during the acquisition. The gamma detector and the robot are sampled at $60\,Hz$, and the optical tracking system at $20\,Hz$.

Robot Trajectories   Two phantoms were scanned during the experiments, an ex-vivo phantom and a translucent plastic phantom, see Figure 6.3. In order to generate scanning
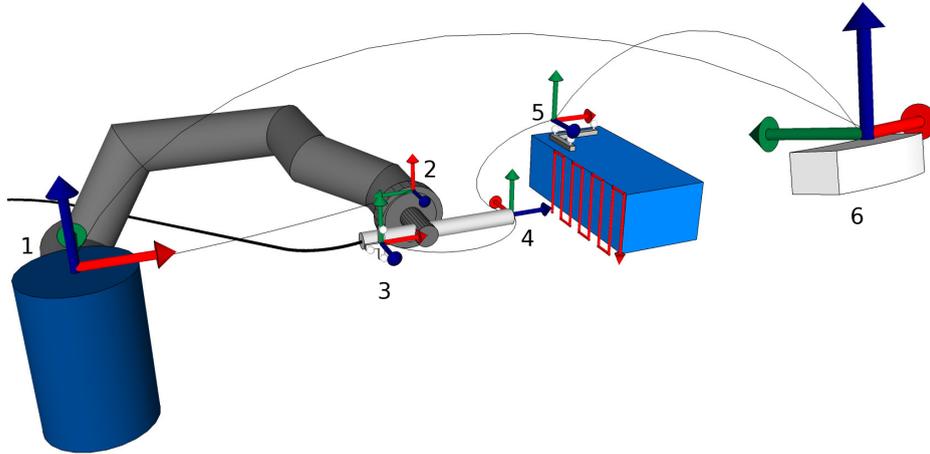
Figure 6.2: Coordinate systems: (1) robot base, (2) robot hand, (3) detector target, (4) detector tip, (5) phantom target and (6) optical tracking system. All transformations can either be precalibrated (detector calibration, hand-eye calibration), measured intra-operatively (robot, optical tracking system), or computed from the precalibrated or measured ones[71].

trajectories for the ex-vivo phantom, human scans were used to compute two types of robotic trajectories.
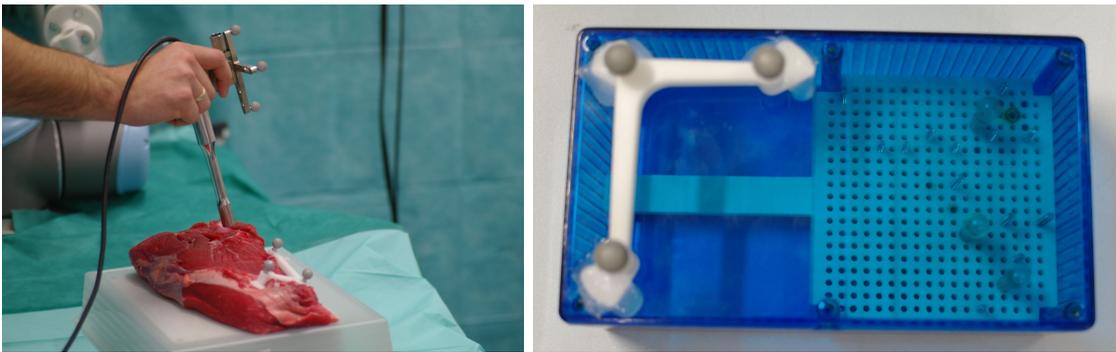
Figure 6.3: Setup with ex-vivo phantom (left), and plastic phantom (right). In both cases the optical tracking target is identical to the one used for clinical patients. The hole raster of the plastic phantom enables reproducible placement of radioactive seeds[71].

The first type of scan, called *Path Follow*, takes the positions recorded during the human scan and down-sample them to $5\,Hz$, which is then executed by the robot at constant speed. This will produce a similar trajectory as the human one, with a slightly different movement speed.

The second type of scan, called *Area Cover*, the convex hull of the human scan is computed and a raster scan is generated within these bounds, which is then executed by the robot at constant speed.

The different trajectories can be seen in Figure 6.4. In both cases, the orientation of the probe is approximately the average orientation measured in the human scan.
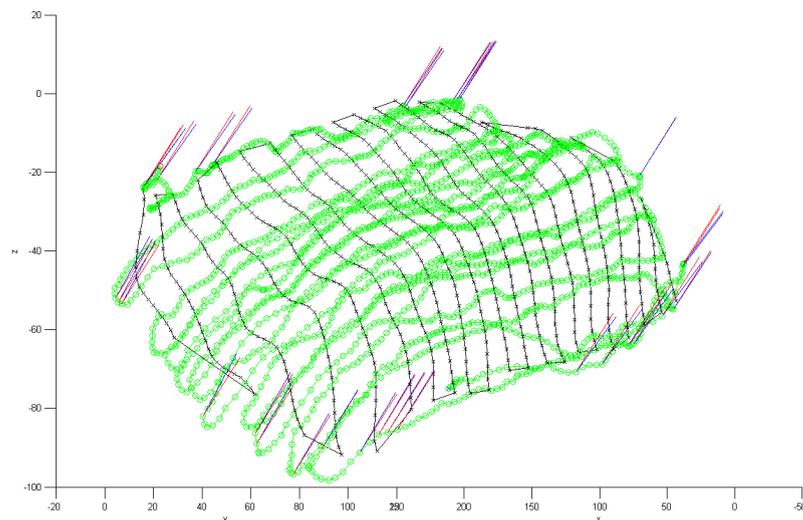
Figure 6.4: Area cover scan (black crosses, probe orientation in red) synthesized from human input (green circles, orientation in blue)[71].

To scan the plastic phantom, since the geometry is known and regular, the scanning pattern was a basic raster scan with constant speed over three orthogonal faces of the phantom (closer to the spheres).

Decay compensation   As radioactivity decays over time, the exposure time for each measurement has to be extended, in order to achieve similar photon detection statistics and thus comparable image quality. This becomes particularly relevant once a half-life or more of the radioactive tracer has passed, which in the clinical case for the commonly used Technetium-99m (^{99m}Tc) is $6.01\,h$. In this case, adjusting the acquisition speed of the robot accordingly can compensate for the decay.

Reconstruction   After acquisition, the 3D activity distribution in the volume of interest was reconstructed from the recorded data, using MLEM, described in Chapter 4 and custom detection models for the gamma detector [76]. The voxel size was $2.5\,mm$ for the ex-vivo phantom and $1.5\,mm$ for the plastic phantom. The reconstruction volumes were $17.5 \times 17.5 \times 12.5\,cm^3$ and $13.5 \times 11.25 \times 4.5\,cm^3$ respectively.

6.3   Experiments and Results

Two sets of experiments were conducted, the first set to show that robotic acquisitions can perform consistent and reproducible scans, while the second set shows the usefulness of decay compensation.

For the first set of experiments, the ex-vivo phantom was used, with three radioactive seeds ($1.5\,ml$ with a solution of $1.5\,MBq$ of ^{99m}Tc each). The phantom was scanned by one expert operator and two novices, twice. The best scan of each operator was selected, and then used to generate both a path follow scan and an area cover scan. The path follow scan was performed by the robot three times, and the area cover scan was performed once.

Since the radioactive seeds were located at different depths, the plane containing all three hotspots was extracted from the 3D reconstruction using PCA (Principal Component Analysis) for visualization purposes, see Figure 6.5.
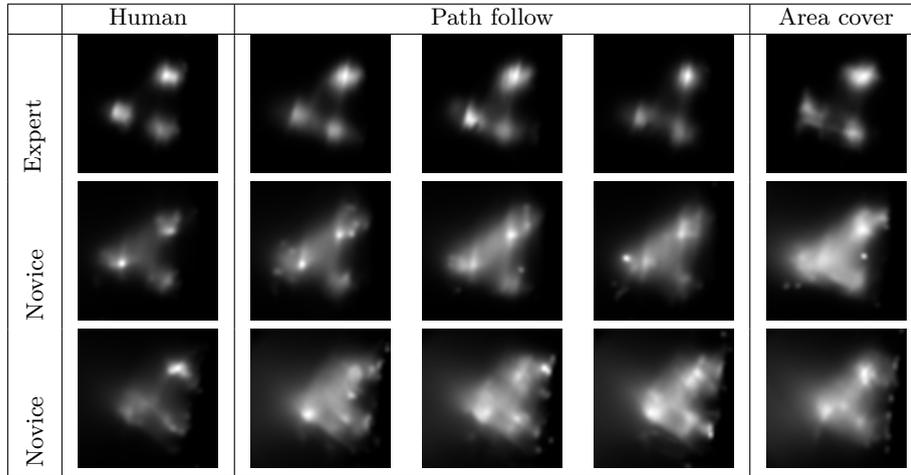
Figure 6.5: Slices through reconstructed images from human operator, robot following human path (three scans), and robot following synthetic path covering area of interest[71].

Figure 6.6 shows an intensity profile across two rightmost hotspots (from Figure 6.5) for the human expert scan and for the robotic scans.

NCC	Human 1-2	Robot 1-2	Robot 2-3	Robot 1-3
Expert	0.942	0.990	0.979	0.989
Novice	0.709	0.980	0.971	0.980
Novice	0.896	0.966	0.973	0.965

Table 6.1: Normalized cross correlation between two operator scans, versus the robotic scan pairs from the path follow trajectory[71].

As a measure of reproducibility, the normalized cross correlation (NCC) was computed between the reconstructed volumes projected along their Z axis, from the two repeated human and the three repeated robotic scans with the path follow trajectory, see Table 6.1.

For the second set of experiments, the plastic phantom containing three spherical seeds ($250\,\mu l$ with a solution of $500\,kBq$ of ^{99m}Tc each) was used.

A raster scan over three orthogonal sides of the phantom was performed, and used as reference (time t). At $t+2$ hours, $t+6$ hours and $t+18$ hours, the same scan was repeated
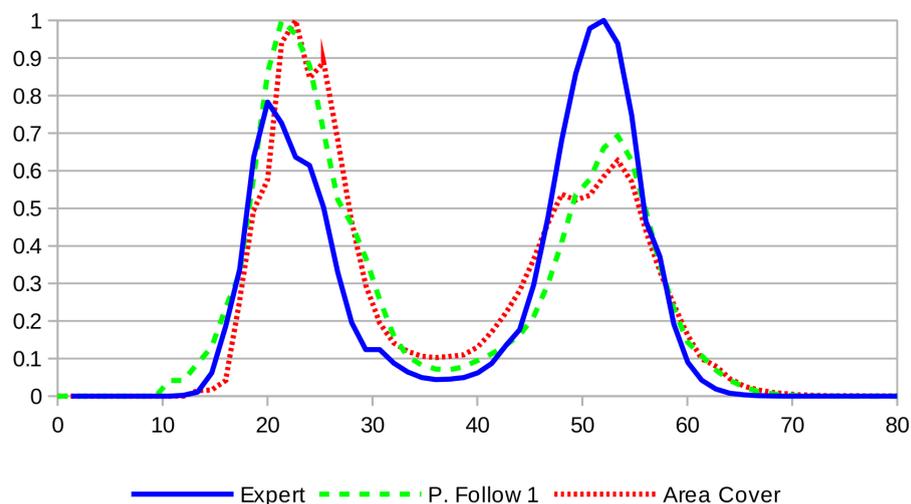
Figure 6.6: Normalized intensity profiles across the rightmost two hotspots for expert and robotic scans. Distance in mm[71].

two times each, once using the original acquisition speed, and once with the speed adjusted to compensate for radioactive decay. In order to evaluate the compensation, again the NCC was computed between the subsequent and the original reconstruction, see Figure 6.7 in the same way as Table 6.1.
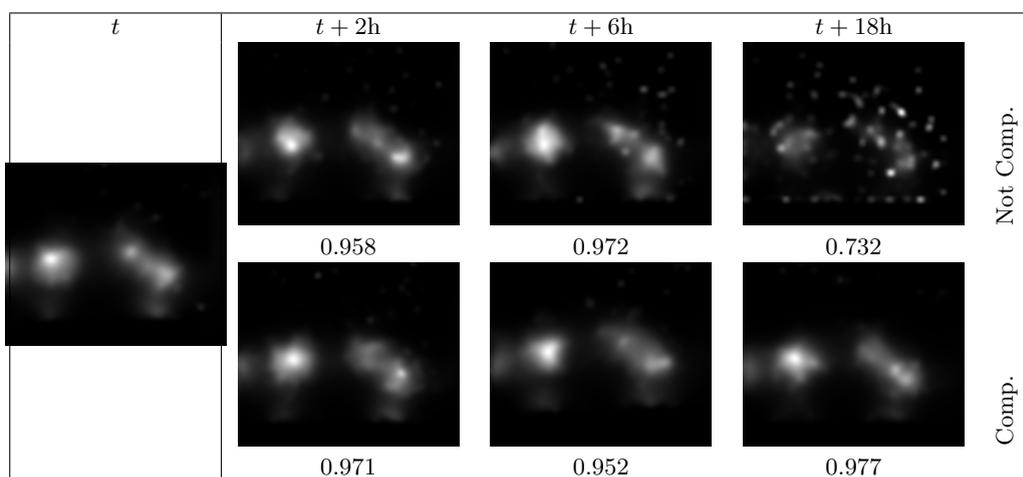

Figure 6.7: Decay compensation experiment. In each column, both images were acquired at approximately the same time $t + x$. Images in upper row were acquired without decay compensation, images in lower row were acquired with decay compensation (i.e. suitably adjusted slower scanning speeds)[71].

6.4   Discussion

The procedure to generate SPECT images from freehand acquisitions using a single detector probe is very challenging. Since the acquisition time is shorter, the detector area is much

smaller, the total number of photons is considerably less than in a conventional SPECT machine. On top of that, the viewing angle of the probe is very broad, limiting the reconstruction capabilities of the setup. Their uses, on the other hand are very different. The SPECT machine is used for the diagnostics, and it requires a reasonably high image quality. The freehand acquisition, on the other hand are performed for guidance purposes, so the exact amplitude of the hot spots are less relevant, assuming that is possible to detect and separate them.

Figure 6.5 and Figure 6.6 show that robotic scans can provide image quality and hotspot separability close to a human expert operator. Furthermore, all robotic scans were highly reproducible with an NCC of over 95%, see Table 6.1. The distances between the hotspots shown in Figure 6.6 appear fairly constant, an indication of the reproducibility of the scans. It is important to note that the path-follow scans are an approximation of the real human scan, which for simplicity and hardware limitations were not performed at the same human speed, and small angle variations during the scan were simply not reproduced my the robotic arm. Those limitations can explain the differences between the human scans and the robotic ones, but what is important to notice here is the reproducibility of the results obtained by the robot arm, as shown before, something not really possible for human operators.

In the decay compensation experiment, after 18 hours, the non-compensated scan does not yield a meaningful image, whereas the compensated scan is still comparable to the original acquisition. It is important to notice, however, that the scan time increased 8-fold in order to have comparable statistics, i.e. from 5 to 40 minutes. This is particularly important because a 40 minute scan is not feasible with human operators, but easily doable with the robotic arm. This experiment is particularly relevant to the clinical workflow of our medical partners. The common procedure in sentinel lymph node biopsy for breast cancer is to inject ^{99m}Tc to the patient during noon, acquire a scintigraphy image for lymphatic mapping, and in the next morning the patient undergoes surgery. Such a decay compensation with robotic imaging can provide more reliable images for incision planning compared to the ones achieved by much shorter freehand scans.

6.5   Conclusion

In this work we presented the first setup for flexible robot controlled intraoperative functional imaging with a first evaluation of its performance. This combination enables patient-specific, flexible imaging in the operating room, which could be integrated to the current surgical workflow and provide guidance for surgeons.

The results do not show a real improvement over expert human scans, which is rather disappointing. On the other hand, it is evident that the robot can be used to optimize the acquisition workflow, given the consistency in the reconstruction results.

Chapter 7

Step and Shoot Robotic freehand SPECT using Gamma Camera

7.1 Introduction

As shown in the previous chapter, the idea of having a robot holding a gamma detector to generate freehand SPECT reconstructions looked promising, albeit the previous results did not show a real quality improvement. The main limitation was, in our eyes, the detector. A single pixel detector with a fairly broad angle collimator (60°) is not able to resolve small structures and has a small collecting area.

Therefore, the next logical step was to replace the gamma probe. At that point we had access to a 16 by 16 pixels mini gamma camera with a collimator of approximately 11° and a much larger collecting area. With that we proceeded to generate a characterization of the detector using a ^{57}Co point source and a positioning table.

As acquisition modality, we chose *step and shoot* since using a camera with 256 detectors instead of a single one resulted in a much bigger memory consumption. This fact restricted us in the number of poses we could use for reconstruction.

The work presented here belongs to the paper "Mini Gamma Cameras for Intra-operative Nuclear Tomographic Reconstruction", published at "Medical Image Analysis (MedIA)" in 2014[77].
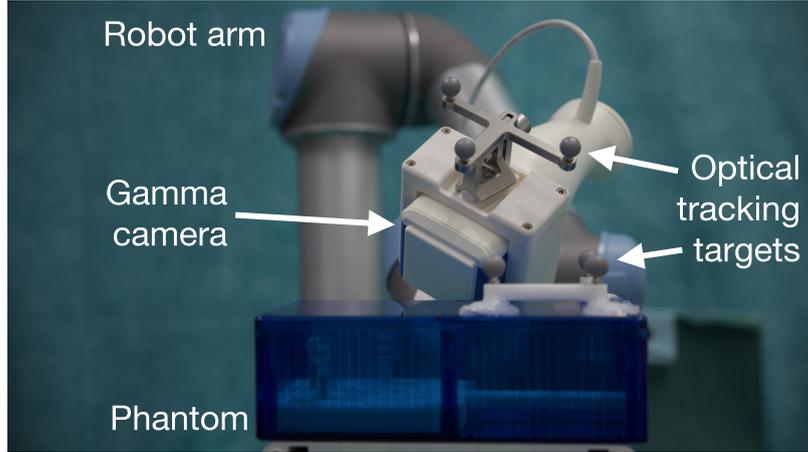
Figure 7.1: Setup with optically tracked gamma camera attached to robot arm, scanning the phantom[77].

7.2 Methods

The imaging setup consists of three parts: the mini gamma camera, the robot arm and the optical tracking system, shown in Figure 7.1. The gamma camera was rigidly mounted on the robotic arm using a custom-made holder.

In the following, we describe the methods to connect these three parts together to enable nuclear tomographic reconstruction.

7.2.1 Measurements

The robot arm moves the camera to several measurement *poses* $\mathbf{y}_k = (\mathbf{p}_k, \mathbf{o}_k) \in \mathbb{R}^3 \times \mathbb{R}^3$ (*positions* $\mathbf{p}_k$ and *orientations* $\mathbf{o}_k$) distributed around a *volume of interest* $V \subset \mathbb{R}^3$. At each position, the camera acquires the detected events for each camera pixel, during a fixed *exposure time interval* ($t_{\exp}$), giving a certain number of counts per second (cps) per camera pixel. We use the unique index j to refer to such a *single* detector reading, and the respective measured value is denoted as m_j.

The number of measurement poses (k) as well as their distribution in relation to the activity to be reconstructed has a high impact in the quality of the images. Due to the constraints in time and space in intra-operative scenarios, the measurement poses will be very limited in number as well as sparsely distributed.

7.2.2 Statistical tomographic reconstruction

As explained in Chapter 4, the aim of the imaging setup is to recover the *radioactivity distribution* $f : V \subset \mathbb{R}^3 \to [0, \infty)$. We discretize this still unknown function f as

$$f(\cdot) \;\approx\; \widehat{f}(\cdot) \;=\; \sum_i x_i b_i(\cdot), \tag{7.1}$$

where $b_i : V \to [0, \infty)$ denotes a user-defined voxel basis of V, and $\mathbf{x} = (x_i) \in \mathbb{R}^n$ the vector of unknown coefficients. Once $\mathbf{x}$ is estimated from the measurements, the synthesis operator (Equation 7.1) will yield the approximated activity signal $\widehat{f}$. For simplicity, we just use the index i to refer to voxel b_i.

The process of detector j detecting emissions from voxel i is modeled using a Poisson process, denoting the detection probability as

$$a(i, j) = P[\text{detected in } j \mid \text{emitted from } i]. \tag{7.2}$$

The detector readings m_j are thus interpreted as independently distributed Poisson random variables, with expectation

$$E(m_j) = \sum_i x_i \cdot a(i, j).$$

The reconstruction is performed using MLEM, described in more detail in Section 4.3, the update equation is:

$$x_i^{q+1} = x_i^q \cdot \frac{1}{\sum_j a(i,j)} \sum_j \frac{m_j \cdot a(i,j)}{\sum_i x_i \cdot a(i,j)}$$

with an initial value of $\mathbf{x}^0 = \mathbf{1} \in \mathbb{R}^n$.

7.2.3 Modeling of the mini gamma camera

A crucial component of the reconstruction process is knowledge of the detection probabilities $a(i, j)$ from equation (Equation 7.2), also denoted as the model of the mini gamma camera. While there are several ways of estimating the $a(i, j)$ as explained in Section 4.4.1, in this work we measure an approximation in a calibration step by directly recording the detector response to a point source. The chosen method of direct measurement has the advantage that the exact geometry of the camera does not have to be known, and that it accounts for

mechanical differences between the design model and the real setup, for example a misaligned collimator with respect to the camera pixels.

As the camera measurement poses $\mathbf{y}_k$ (resulting in the detector readings j) are arbitrary, and the discretization of the volume of interest V into voxels i is variable, the mini gamma camera model has to be flexible and detailed enough to account for this. Consequently, we decompose $a(i, j)$ by decoupling the actual calibration measurement, the *detector-relative* measurement probability $p(\ell)$, from the discretization of V by a transformation $\mathcal{T} : (i, j) \mapsto \ell$ mapping the pair of detector reading j and voxel position i to the detector-relative position $\ell \in \mathbb{R}^3$, that is

$$a(i, j) = p\big(\mathcal{T}(i, j)\big).$$

The two components p and $\mathcal{T}$ are explained in more detail in the following.

Detector-relative measurement probability
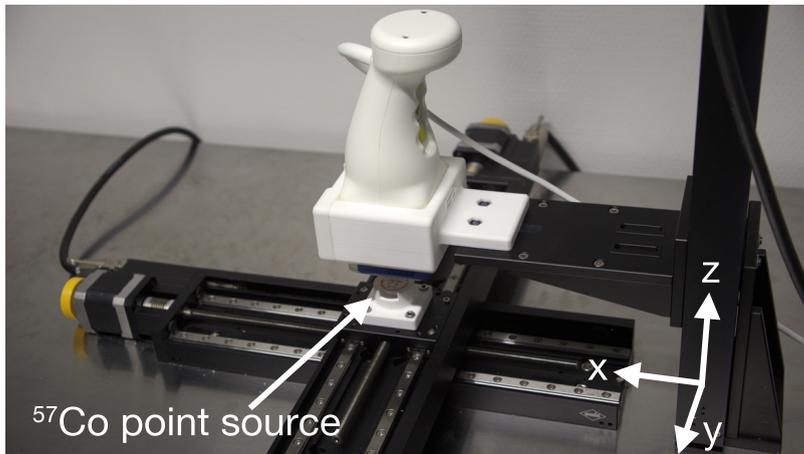
Figure 7.2: Calibration setup for acquisition of the mini gamma camera model p. The camera is mounted to a positioning table to systematically measure a radioactive ^{57}Co point source in front of the camera in different positions[77].

The detector-relative measurement probability $p : \mathbb{R}^3 \to [0, 1]$ is the camera's probability to detect a decay event occurring at location $\ell \in \mathbb{R}^3$, where ℓ is defined relative to the center of the camera's sensor plate. We call the latter the *origin* of the detector. To approximate this mapping, a radioactive ^{57}Co point source is measured systematically in front of the camera using a precision positioning table, see Figure 7.2. At each position $\ell^* \in \mathbb{R}^3$, camera readings are acquired with a statistically significant exposure time $\widetilde{t}_{\exp}$. Figure 7.3 shows p for a single example detector pixel.

(a) Mini gamma camera

(b) Measurement probabilities p
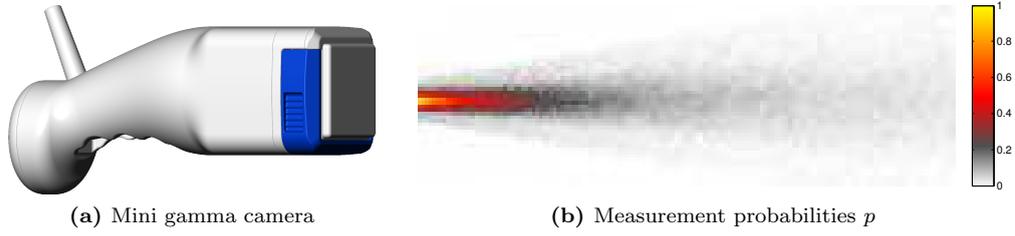
Figure 7.3: Plane cut showing the camera model p of a center pixel of the mini gamma camera. Contributions of the different voxels in front of the camera collimator to the selected camera pixel are shown[77].

Coordinate system calibration

In order to obtain the measurement probability $a(i, j)$ for a voxel i during acquisition j as required for tomographic reconstruction, the voxel position first needs to be computed relative to the detector's origin using $\mathcal{T}$. Then, the probability for this point can be extracted from the discretized measurement p using interpolation.

We obtain $\mathcal{T}$ by relating the origin of the *volume* of interest V with the *detector* origin. We define canonical coordinate systems for both of them, where the detector coordinate system is chosen consistent with the camera model p – centered on the detector front face, with x and y axes aligned to the camera's pixels, and the z axis pointing away from the camera, see Figure 7.4. Then, the crucial part of $\mathcal{T}$ can be written as an Euclidean transformation $^{volume}\mathbf{T}_{detector} \in \mathbb{R}^{4\times4}$.

Our setup is using an optical tracking system for spatial localization, with tracking targets attached to the phantom (where the *volume* of interest is) and the *camera*. Thus, given a static transformation $^{camera}\mathbf{T}_{detector}$ from camera tracking target to its detector coordinate system, the full transformation reads

$$^{volume}\mathbf{T}_{detector} = {}^{volume}\mathbf{T}_{camera} \cdot {}^{camera}\mathbf{T}_{detector},$$

where the varying transformation $^{volume}\mathbf{T}_{camera}$ is provided by the tracking system.

The last bit, $^{camera}\mathbf{T}_{detector}$, is finally obtained using a simple calibration scheme: As shown in Figure 7.4, an optically tracked pointer tool is used to mark the corners of the housing at the front of the camera. Based on the known design of the camera, the full transformation to the *detector* coordinate system can then be computed.
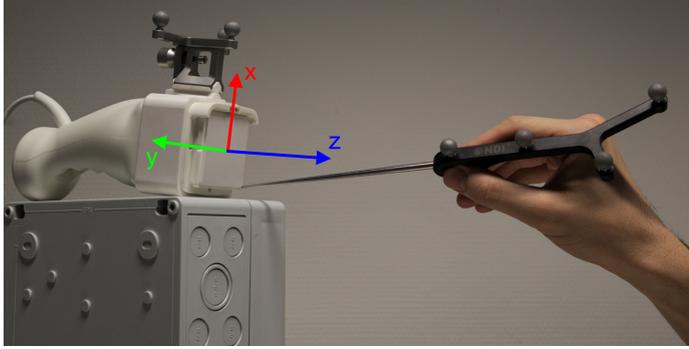
Figure 7.4: A pivoting tool is used to obtain the transformation matrix $^{camera}\mathbf{T}_{detector}$ between the optical tracking target attached to the gamma camera and the detector coordinate system (overlaid on the gamma camera)[77].

7.3   Experiments

7.3.1   Hardware setup

The mini gamma camera employed is a CrystalCam by Crystal Photonics, Germany[72]. The detector of this gamma camera consists of a $40 \times 40\,mm^2$ CdZnTe crystal which has 16×16 pixels. The collimator measures $11.15\,mm$ in length and is made of lead for side shielding and tungsten septa to separate the pixels. The 256 square holes are centered in front of the detector pixels and have sizes of $2.16 \times 2.16\,mm^2$ each. The energy resolution is less than 7% at $140\,keV$.

The camera is attached to the hand of a robot arm (UR5, Universal Robots, Denmark) via a custom mount, see Figure 7.1.

A tracking target is mounted to the camera housing for tracking by an optical tracking system (Polaris Spectra, Northern Digital Inc., Canada[73]). An additional tracking target is used to define the reconstruction volume V, in our case this target is glued to a phantom (see Section 7.3.3).

A computer running Linux is used to control and read out all the hardware. For synchronization and precise timing, the system was set up with a custom low-latency kernel with a timer frequency of $1000\,Hz$.

7.3.2   Gamma camera model

The detector-relative measurement probabilities are acquired separately using a high precision positioning table (OWIS, Germany) with a very low repeatability error ($< 15\,\mu m$), see Figure 7.2. The mini gamma camera is attached to the z-axis of this positioning table via a custom mount. A solid ^{57}Co point source is attached to the x and y axes of the positioning

table and stepped systematically in front of the mini gamma camera. ^{57}Co has a very similar peak energy ($122\,keV$) to ^{99m}Tc ($140\,keV$) but a much longer half-life (271 days, versus 6 hours), making it the preferred choice for this experiment. The energy window was set at 10% around the energy peak.

The camera model p consists of roughly 300,000 samples in a three dimensional grid, where each sample is acquired with an exposure time $\widetilde{t}_{\exp} = 7\,s$. The first 186,000 samples (within $50\,mm$ distance from the camera front, i.e. in z-direction) are sampled isotropically at $1\,mm$. The remaining samples (between $50\,mm$ and $150\,mm$ distance from the camera front, i.e. in z-direction) are sampled isotropically at $2\,mm$.

7.3.3 Phantom

A phantom containing three hollow spheres was created to mimic scenarios found in sentinel lymph node biopsy procedures [cf. 78]. The spheres have comparable sizes to sentinel lymph nodes for example in the axilla (each with a volume of $0.25\,ml$ and inner diameter of $8.6\,mm$), and are screwed to a grid contained in a plastic box, with outer dimensions $190 \times 110 \times 61\,mm^3$ and usable inner dimensions $92 \times 97 \times 41\,mm^3$.

An optical tracking target to define the *volume* coordinate system is fixed to the lid of the plastic box, as shown in Figure 7.5a. The three spheres are filled with a ^{99m}Tc solution, $1.2\,MBq$ in sphere (1), $0.5\,MBq$ in sphere (2) and $1.4\,MBq$ in sphere (3), cf. Figure 7.5a. Based on the experience in our university hospital, these are realistic post-injection uptakes for axillary sentinel lymph nodes in breast cancer patients, at a total local injection of $100 - 150\,MBq$ of ^{99m}Tc-nanocolloid around the breast nipple, where the majority of the injected radiation stays. All spheres are located at the same height z from the bottom grid of the phantom, the distances $d_{m \leftrightarrow n}$ between the spheres m and n are $d_{1 \leftrightarrow 2} = 36.40\,mm$, $d_{1 \leftrightarrow 3} = 40.31\,mm$, and $d_{2 \leftrightarrow 3} = 62.65\,mm$.

7.3.4 Parameters for reconstruction

The reconstruction volume V is set to $84 \times 100 \times 40\,mm^3$, approximately the usable inner space of the box phantom described in Section 7.3.3. The discretization of V is performed using $1\,mm^3$ isotropic voxels, yielding 336,000 voxels in total.

The acquired data sets contain up to 5376 measurements (256 camera pixels at 21 different poses).

(a) Phantom from top   **(b)** Scanning scheme
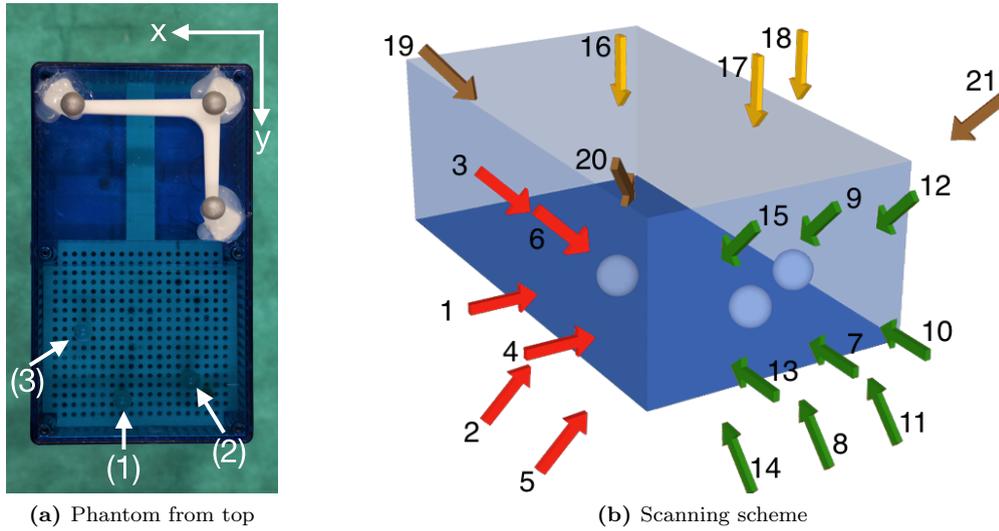
Figure 7.5: Semi-transparent phantom from top view with the white optical tracking target fixed to the lid and the three spheres filled with radioactive ^{99m}Tc solution inside, positions indicated by the arrows. The phantom was scanned from several perspectives, from sides x (red), y (green), and z (yellow), as well as from three diagonal poses (brown)[77].

All tomographic reconstructions are performed using MLEM, stopping early after 20 iterations to avoid detrimental noise effects. No further post-processing is applied to the resulting images.

7.3.5 Experiment design

To generate the data for the experiments, the robot arm is set to acquire data from 21 poses around the phantom, as shown in Figure 7.5b. To mimic an intra-operative scenario with very limited access, for example in the abdomen, we define two pose sets covering one side of the phantom only, once with the camera oriented orthogonal to the respective side (poses 7, 10 and 13), and once oriented orthogonal and diagonal (poses 7-15). For brevity we denote these two pose sets as 1S (one side) and 1SD (one side with diagonals). In scenarios like thyroid imaging, access from two perpendicular sides may be possible. Hence the next two pose sets (2S and 2SD) include scans from two sides, again once scanned with only orthogonal orientations (poses 1, 4, 7, 10, and 13) and once with orthogonal and diagonal orientations (poses 1-15). Finally, the last two sets 3S (this is also similar to the one used in [79]) and 3SD are scanned from three perpendicular sides (poses 1, 4, 7, 10, 13, 16-18 and poses 1-21, respectively). They represent clinical scenarios as found for example in sentinel lymph node biopsy procedures for breast. The distances between the center of the closest sphere and the tip of the camera vary between $3\,cm$ and $15\,cm$, comparable to clinical scenarios.

The robot stops at each pose for an exposure time of $10\,s$, and the events detected by the gamma camera are recorded at each pose. These six different sets of poses are acquired, as summarized in Table 7.1, with varying amounts of exposure times $t_{\exp} \in \{1, 2, 3, 5, 10\}\,s$.

Pose set	Poses	Description
1S	7, 10, 13	One side, orthogonal scans
1SD	7-15	One side, orthogonal and diagonal scans
2S	1, 4, 7, 10, 13	Two sides, orthogonal scans
2SD	1-15	Two sides, orthogonal and diagonal scans
3S	1, 4, 7, 10, 13, 16-18	Three sides, orthogonal scans
3SD	1-21	Three sides, orthogonal and diagonal scans

Table 7.1: Different pose sets used for experiments[77].

The first experiment (we refer to it as *Experiment A*) studies the influence of the chosen acquisition poses on the quality of the reconstructed image. Hence we evaluate all six pose sets from Table 7.1, using an exposure time $t_{\exp} = 10\,s$ to have sufficient detection statistics. This will provide an expectation of the potential imaging quality in the intra-operative scenarios outlined above.

The second experiment (we refer to it as *Experiment B*) is aimed at understanding the influence of the exposure time on the reconstruction quality. For smooth workflow integration, this time should be minimized, while for best detection statistics, exposure time per pose should be maximized. We choose the three pose sets 1SD, 2SD and 3SD (cf. Table 7.1) and reconstruct them using exposure times $t_{\exp} \in \{1, 3, 5, 10\}\,s$ for each pose. This will provide a first guideline what exposure times will be clinically useful.

In a third *Experiment C*, we fix the total acquisition duration to approximately $45\,s$ to simulate an intra-operative workflow with limited time available for imaging, but freedom in choosing poses. We select again the three pose sets 1SD, 2SD and 3SD (cf. Table 7.1), and reconstruct them using exposure times of $5\,s$, $3\,s$ and $2\,s$, respectively. This experiment will reveal whether more acquisition poses or longer exposure times are more important for image quality.

7.3.6 Evaluation

In order to evaluate the reconstruction results, we process the reconstructed volumetric images of each experiment to extract information on the hot spots and compute several quality measures.

First, we run a conservative thresholding step discarding intensities lower than 1% of the global maximum intensity of the respective image. Subsequently, the number of reconstructed hot spots, their respective centroids, and their uptake values based on the results of an automatic 3D region growing algorithm are computed. The algorithm used is similar

to the one in the work of [80], where they segment the hot spots in SPECT images to identify lymph nodes. Our algorithm also uses local maxima in the reconstructed volume as seed points for a recursive region growing approach to determine the hot spot candidates. Afterwards it iterates over the candidates and deletes the ones which were already segmented. From these segmented hot spots, those corresponding to the ground truth spheres (true positives), and the regions not forming a part of any sphere are labeled as artifacts (false positives). Both the case of several reconstructed hot spots corresponding to a ground truth sphere and the case of a ground truth sphere having no corresponding reconstruction (false negative) did not occur in any of the experiments. Based on this data, we compute several quality measures:

Average localization error

We compute the 3D distances of the reconstructed centroids to the ground truth centroids of the hot spots. The average of these centroid errors evaluates the spatial accuracy of the reconstruction as a whole.

Artifacts

We evaluate the artifacts in the reconstruction, that is the segmented hot spots not corresponding and not connected to a ground truth hot spot (false positive). Minimizing the number of artifacts and the accumulated intensity in those artifacts is crucial for intra-operative use, since high intensity artifacts may lead to false conclusions. Two measures were evaluated for each image: the number of segmented artifacts present in the reconstruction, as well as the accumulated uptake assigned to all artifacts combined (as a percentage of the total activity reconstructed).

Sørensen-Dice similarity coefficient

The Sørensen-Dice coefficient QS [81, 82] is a metric used for comparing the similarity of two datasets. In our case, we compare the volumes of all three spheres $\bar{f}_{gt}$ as ground truth with the reconstructed and binary segmented spheres $\bar{f}_{seg}$. The formula

$$QS = \frac{2 \cdot \int_V \bar{f}_{gt} \cdot \bar{f}_{seg} \, dV}{\int_V \bar{f}_{gt} \, dV + \int_V \bar{f}_{seg} \, dV} \tag{7.3}$$

indicates the amount of overlap ranging from 0 to 1.

(a)	(b)
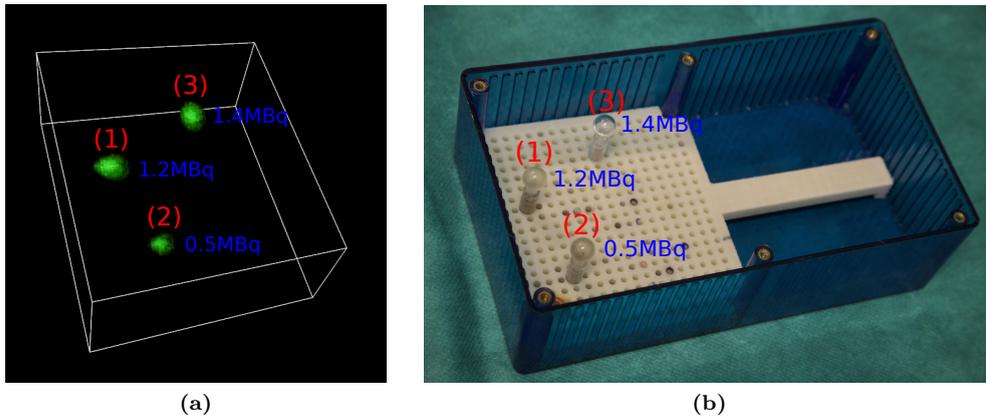
Figure 7.6: *Left:* Volume rendering of the reconstructed image (pose set 3SD, $t_{\exp} = 10\,s$). *Right:* Photograph of the box phantom without lid from a similar perspective, showing the three spheres containing radioactive ^{99m}Tc solution. The spheres are marked with their respective number (red) and ground truth activity (blue)[77].

Ratio of uptake values

Finally, we compute the relative uptake in each hot spot, that is the relative activity reconstructed per sphere as a percentage of the total activity reconstructed. Ideally, this should be equal to the ratio of the ground truth activities per hot spot as a percentage of the total activity in all spheres.

7.4 Results and discussion

7.4.1 Results

To give a general impression of the reconstructed images, Figure 7.6b shows an example reconstruction of the phantom (using measurement pose set 3SD, $t_{\exp} = 10\,s$), next to a photograph of the phantom.

Experiment A

Here we evaluate the influence of the chosen acquisition poses on the results. Figure 7.7 shows the reconstruction results of *Experiment A* as projection images along the z-axis as well as the y-axis for all six pose sets using $t_{\exp} = 10\,s$, while Table 7.2 lists the quality measures for each reconstruction. The results of pose sets 1S and 2S show elongations and artifacts along the scanning directions. Adding more sides (3S) or diagonal orientations (1SD, 2SD, 3SD) allows to increasingly eliminate the elongations and artifacts and to confine the reconstructed activity in the expected areas.

(a) Pose set 1S (one side, orthogonal scans)

(b) Pose set 2S (two sides, orthogonal scans)

(c) Pose set 3S (three sides, orthogonal scans)

(d) Pose set 1SD (one side, orthogonal and diagonal scans)

(e) Pose set 2SD (two sides, orthogonal and diagonal scans)

(f) Pose set 3SD (three sides, orthogonal and diagonal scans)
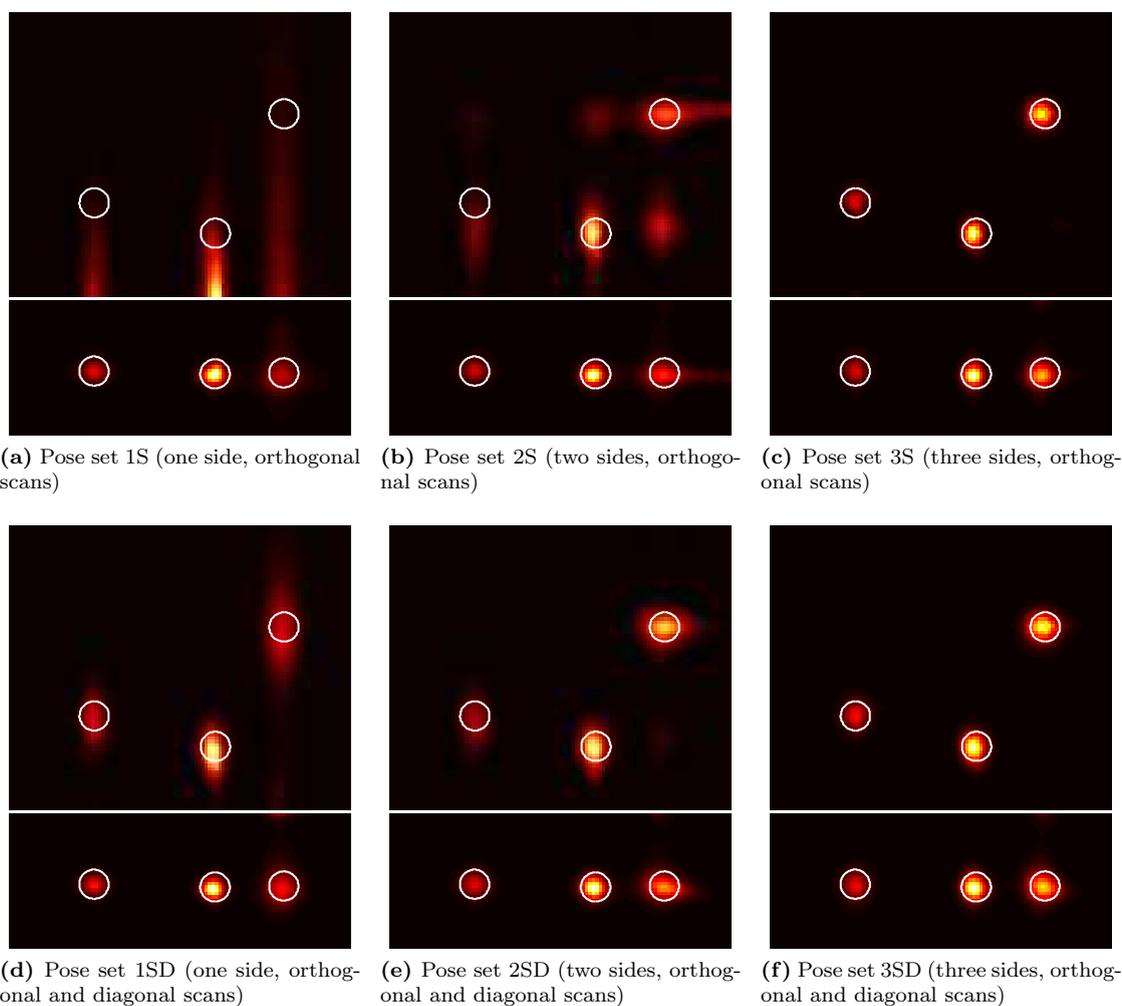
Figure 7.7: Results of *Experiment A* shown as projections along the z-axis (top row) and y-axis (bottom row). The projections are individually normalized for better visualization. The ground truth is depicted as an overlay of white circles representing the spheres of the phantom[77].

Experiment B

Here we evaluate the influence of the exposure time per pose for the pose sets 1SD, 2SD and 3SD using exposure times ranging from $1\,s$ to $10\,s$. The quality measures *average localization error* and *ratio of uptake values* do not show a significant difference (up to 10% for pose set 1SD, less than 5% for pose sets 2SD and 3SD). A more visible impact of the exposure time is shown in the *artifact* quality measure where longer exposure times reduce the number of the artifacts and their uptake, see Figure 7.8.

Experiment C

Using a fixed total acquisition duration of approximately $45\,s$, we evaluate the pose sets 1SD, 2SD and 3SD with corresponding exposure times t_{exp}. The results in Table 7.3 show,

Pose set	Activity uptake			Artifacts		Localization error	Sørensen-Dice coeff.	Data set # poses
	# 1	# 2	# 3	#	Upt.			
1S	67%	20%	13%	2	0%	29.1 mm	0.03	3
2S	33%	14%	37%	5	15%	3.0 mm	0.45	5
3S	39%	16%	42%	7	2%	1.0 mm	0.78	8
1SD	42%	19%	37%	4	3%	1.6 mm	0.48	9
2SD	40%	17%	42%	3	1%	0.9 mm	0.70	15
3SD	39%	17%	44%	1	0%	0.7 mm	0.82	21
GT	**39%**	**16%**	**45%**					

Table 7.2: *Experiment A:* quality measures for all six pose sets (using $t_{exp} = 10\,s$). From left to right: ratio of uptake values for the three spheres, number and uptake of artifacts, average localization error, Sørensen-Dice similarity coefficient, data set information (number of poses). The last row displays the ground truth (GT) uptake values for comparison[77].
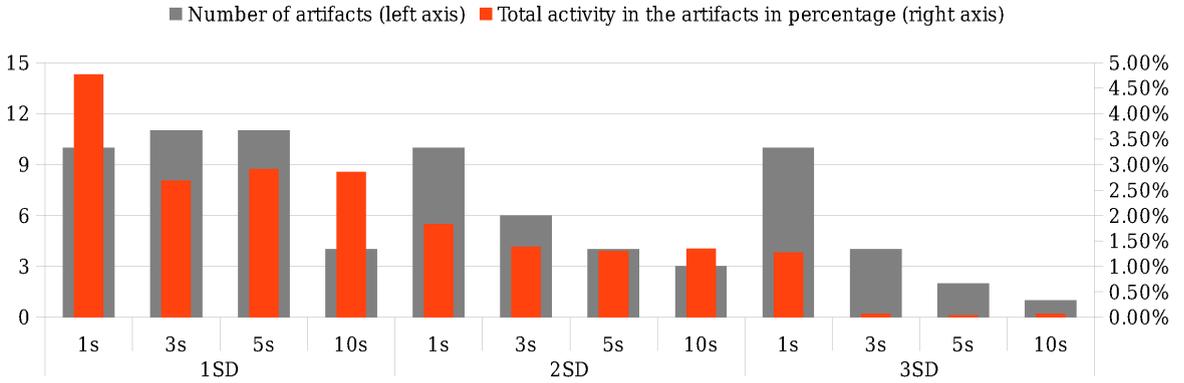
Figure 7.8: *Experiment B:* influence of the exposure time on the artifacts. Gray: total number of artifacts in volume (using left axis); Orange: percentage of activity in the artifacts compared to total reconstructed activity (using right axis)[77].

Pose set	Activity uptake			Artifacts		Localization error	Data set		Total duration
	# 1	# 2	# 3	#	Upt.		# poses	t_{exp}	
1SD	42%	19%	36%	11	3%	1.6 mm	9	5 s	45 s
2SD	40%	17%	42%	6	1%	1.0 mm	15	3 s	45 s
3SD	39%	17%	44%	2	0%	0.7 mm	21	2 s	42 s
GT	**39%**	**16%**	**45%**						

Table 7.3: *Experiment C:* quality measures for a fixed total acquisition duration of approximately $45\,s$ using selected pose sets. From left to right: ratio of uptake values for the three spheres, number and uptake of the artifacts, average localization error, and data set information (number of poses, exposure time t_{exp}, total acquisition duration). The last row displays the ground truth (GT) uptake values for comparison[77].

that more poses achieve better results compared to longer exposure times.

7.4.2   Discussion

As shown in the previous section, it is possible to segment and locate hot spots with high accuracy, given that the measurements cover the phantom sufficiently. The exposure time per pose also plays a role, but it is less crucial when considering the overall reconstruction quality.

Experiment A shows that the number of sides covered by the projections have great impact on the reconstruction quality. Having diagonal views on the sides helps improving the quality of the reconstruction, making Figure 7.7f the best reconstruction. The improvements made by adding diagonal views is also clearly less crucial than the improvement gained by adding poses from another, orthogonal side. Having poses from three orthogonal sides plus diagonal views provides the best image quality. Unfortunately, this type of configuration is not always possible within the operating room, due to the patient anatomy and limited access to the region of interest. Consequently, studying constrained scenarios is key to understanding the feasibility of the technology within actual clinical settings. Judging from the results, it is clear that the 2SD case can also provide acceptable reconstructions.

Experiment B shows the existence of a clear dependency between the exposure time on each pose and the quality of the reconstructed images. The *average localization errors* and *ratios of uptake values* do not improve much, but the *artifact* quality measure improves by at least 50% when increasing the exposure time from $1\,s$ to $10\,s$, independent of the number of poses covered. But not only the number of artifacts is important, also the accumulated activity plays a role. For the acquisition sets 1SD and 2SD, this measure does not improve much with longer exposure time per pose, whereas for set 3SD it improved significantly. But since in set 3SD the number of artifacts and their total activity is already quite low, this improvement has to be put into perspective.

Experiment C clearly shows that the main factor for improved image quality is the number of distinct poses. All our quality measures improve within the fixed amount of approximately $45\,s$ scanning time: The *ratio of uptake values* approaches the ground truth values when adding more poses, the *number of artifacts* and their *uptake* reduces significantly and the *localization error* also improves by more than 50%.

Experiments B and C show that increased exposure times do not improve the quality of the reconstructions very much. In particular, it is much more effective to distribute the available time to cover more poses, and in doing so increase the angular coverage of the region of interest.

Increasing the number of poses has a direct influence in the reduction of the artifacts, as long as the three orthogonal views are still covered. Furthermore, the activity percentages in the segmented hotspots in all our reconstructions correspond nicely with the actual radioactivity in the spheres, with only slight deviations.

Compared to [79], the hardware setup was upgraded and software changes were introduced. The localization errors of the acquisition scan similar to pose set 3S are much lower in this work, and the acquisition times were reduced tenfold in the shortest scans.

We believe those improvements are mainly caused by a much better, high-resolution and exhaustive characterization of the gamma camera's measurement model. The voxel size was also reduced from about $16\,mm^3$ to $1\,mm^3$, and as a result we can now measure the error with much higher accuracy than before.

7.5 Conclusion

The results show that it is possible to obtain accurate reconstructions with less than one millimeter average localization error in less than one minute of acquisition. The key parameter for image quality was shown to be the number of significantly different measurement poses.

The tested configurations were comparable to an intraoperative sentinel lymph node biopsy, where the results obtained show that this setup can be used as guidance for the surgeon to accurately locate and resect the radioactive lymph nodes.

Compared to the results from the previous chapter using the gamma probe, the localization error of the point sources is about one order of magnitude lower. This opened us the door for more challenging reconstruction scenarios, presented in the next chapter.

Chapter 8

Continuously Moving Robotic freehand SPECT using Gamma Cameras

8.1   Introduction

As seen in the previous chapter, the results obtained using robotic gamma camera acquisitions for freehand SPECT reconstructions meant a huge step forward in terms of image quality. With respect to the gamma probe acquisitions, the improvement in terms of accuracy made us believe that we could use this technology in more complex scenarios.

To be able to reconstruct extended sources with a small camera, it is necessary to scan the objects from many poses, which produces fairly big matrices for the reconstruction. This, combined with the fact that the time between scanning positions is wasted and that *step and shoot* with a gamma camera is not really feasible for handheld acquisitions, made us search for new reconstruction methods.

List Mode Expectation Maximization appeared as the solution for this problem. This method allowed us to have continuous trajectories, so no more wasted acquisition time. This also makes it possible to have handheld acquisitions again if needed. The size of the matrices and, therefore, the computation time were also significantly reduced when doing a continuous acquisition.

The work presented here belongs to the paper "Flexible Mini Gamma Camera Reconstruction of Extended Sources using Step and Shoot and List Mode", published at "Medical Physics" [83].
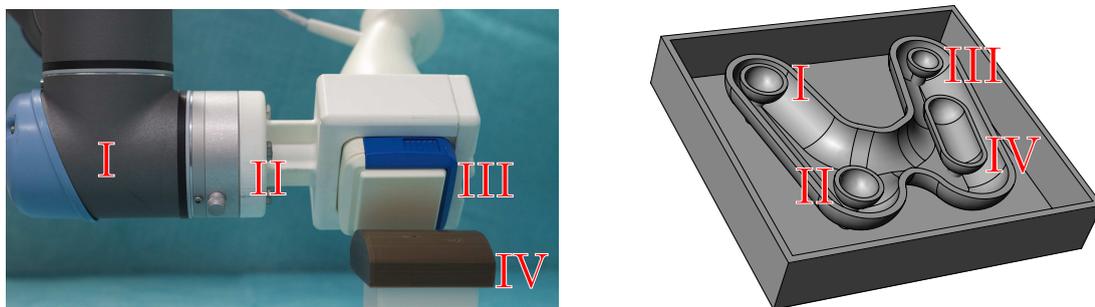
(a) Complete setup with part of the robotic arm (I), gamma camera holder (II), gamma camera (III) and thyroid phantom (IV).

(b) Rendering of the thyroid phantom (bottom half). Two medium-sized nodules (I, II) are on the left, a small (III) and a big one (IV) on the right.

Figure 8.1: Hardware setup and thyroid phantom[83].

8.2   Hardware Setup

Our experiments aim at comparing different acquisition and reconstruction modes for thyroid imaging using a flexible mini gamma camera setup. In order to have equivalent, realistic data in each of the settings, our setup consists of three components: A custom designed thyroid phantom, a mini gamma camera, and a robotic arm guiding the camera (see Figure 8.1).

The mini gamma camera is a CrystalCam by Crystal Photonics[72], described in detail in Section 7.3.1. We selected an energy window of $\pm 5\%$ around the $140\,keV$ gamma peak of ^{99m}Tc. The gamma camera characterization procedure corresponds to the one presented in Section 7.2.

The robot, a Universal Robots[84] UR5 (see Section 5.3) , and the camera are rigidly coupled by a custom-designed adapter. The main purpose of the robot is to have accurate position tracking and reproducible trajectories.

8.3   Phantom

Many commercial phantoms for scintigraphy are intended for planar acquisition only. Our custom-printed phantom is designed to better mimic a human thyroid, both in terms of three-dimensional shape as well as volume. A rendering of the model and a picture of the complete setup are shown in Figure 8.1.

The phantom consists of a chamber shaped like a thyroid, with four inner chambers resembling nodules (numbered I through IV as marked in Figure 8.1b), inside a box-like enclosure.

The phantom is mostly symmetric, with small cylindrical support structures ($2\,mm$ diameter) that hold the thyroid and inner chambers in place with respect to the enclosure. The

(a) Extracted coronal section (in blue), transverse sections (in yellow and red), and profile extraction location (in green).

(b) Mechanical drawing.
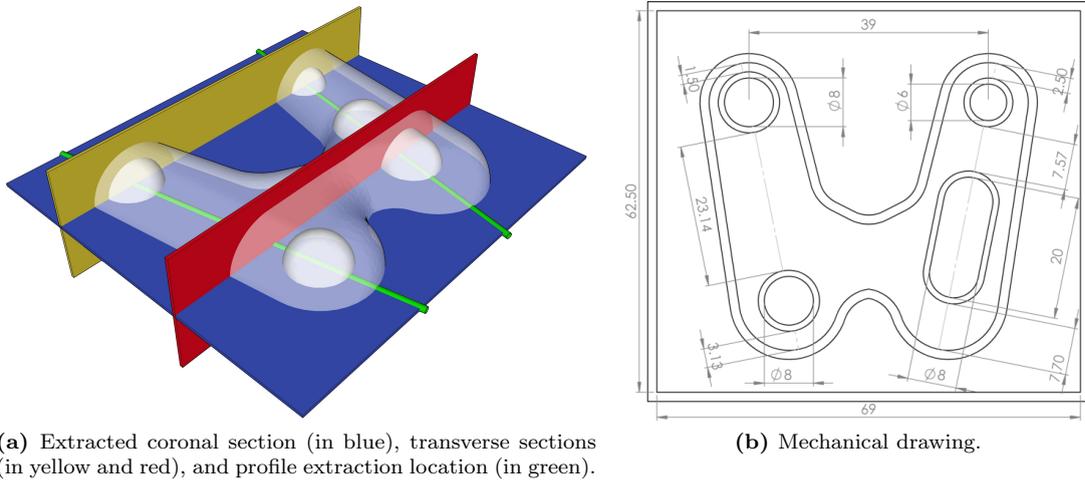
Figure 8.2: Graphical and mechanical description of the thyroid phantom[83].

support structures at the top of the phantom have a hole of $0.5\,mm$ diameter that connects the nodule chambers with the outside of the phantom, enabling the user to fill the chambers with a syringe. Additional hollow support structure enable separate filling of the thyroid body and the outer chamber. The outer bottom part of the enclosure is rectangular, while the top part is rounded, to mimic the neck region better.

The total volume of the thyroid (including the inner chambers and the support material) is $15.5\,ml$, the effective volume is $12.6\,ml$. The smallest chamber (III) has a volume of $0.11\,ml$ ($6\,mm$ inner diameter), the medium sized ones (I, II) $0.27\,ml$ ($8\,mm$ inner diameter), and the bigger elongated one (IV) $0.87\,ml$ ($8\,mm$ diameter half spheres connected by a cylinder, with a total length of $20\,mm$). Figure 8.2b shows a mechanical drawing of the phantom, including the relevant dimensions.

8.4   Experiment Design

With this setup in place, two datasets have been acquired. The first one consists of a *"step and shoot" trajectory (SAS)* around the phantom. In such a setting, the camera is moved to a certain pose, at which measurements are acquired only while the camera remains in place. Afterwards, the camera is moved to the next pose, to resume measuring there. The second dataset is a *"continuous" trajectory* where the camera is moving smoothly while continuously measuring.
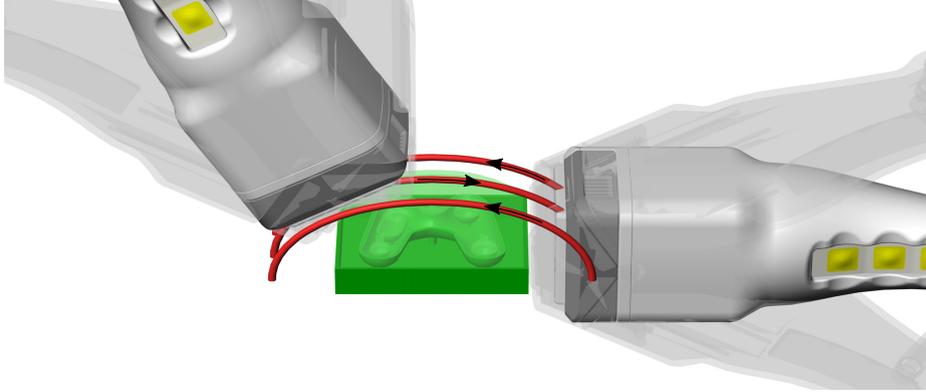
Figure 8.3: Rendering of the gamma camera, showing the scanning positions with respect to the phantom (in green). Measurements were acquired at each position along the trajectory (in red) using three tilts at 0° and ±30°, in plane with respect to the trajectory. In total, three half-elliptical movement paths were performed over the thyroid to achieve full coverage[83].

8.4.1 Datasets

For the SAS case, three parallel half-elliptic trajectories were computed around the phantom, and every 3° over the ellipse, the camera was first positioned oriented perpendicular to the phantom, and then tilted by 30° and −30° as second and third poses, see Figure 8.3. We measured three half-ellipses of 61 stopping points each. Considering the three perspectives at every stopping point, the total number of poses amounts to 549 for SAS. At each pose the acquisition time was $0.6\,s$, yielding an effective acquisition time of $329\,s$. The movement of the camera took $654\,s$ in total. We only consider gamma rays detected while the setup is stationary.

For the continuous case, a similar trajectory was generated. However, this time, we were continuously measuring gamma rays during movement without stopping the movement. We used the same ellipses with the same ±30° tilts as basis for the trajectory. The complete continuous trajectory was executed in $293\,s$.

8.4.2 Reconstruction

With these two datasets and the two expectation-maximization reconstruction algorithms, ML-EM and LM-EM (see Chapter 4), a total of four combinations have been investigated: For *step and shoot ML-EM*, the static poses were extracted, and all respective detections were accumulated into pixel-wise counts, i.e. binned according to poses. For *step and shoot LM-EM*, the same trajectory with static poses was used, and the selected events are identical to the ones used in the previous dataset, but in this case they are represented as a list of events without any binning. For *continuous ML-EM*, the data from the continuous trajectory was temporally binned into intervals of $48\,ms$, similar to the refresh rate of an

optical tracking system [73] as used commercially[16]. For *continuous LM-EM*, we used the continuously acquired detections directly as a list of events.

For all reconstructions, we used a volume of interest of $80 \times 80 \times 50 \, mm^3$, with $2 \, mm$ isotropic cubic voxels.

For all reconstructions we used 166 iterations of the respective algorithm. This number was selected as the maximum number of iterations such that all experiments achieve a likelihood difference between subsequent iterations of less than $5 \cdot 10^{-6}$.

Additionally, we used a simple attenuation correction built in the system matrix, assuming that the phantom completely consists of water. No scatter correction was used. The reconstructions were post-processed with an isotropic Gaussian filter with a kernel size of $1 \, mm^3$ (0.5 voxels).

8.4.3   Radioactivity and Simulated Lower Uptake

The nodules I, III and IV (hot nodules) were loaded with a solution with the same concentration of ^{99m}Tc, yielding a total of $3 \, MBq$ of activity. This solution was then diluted 1:10, and used to fill the thyroid chamber, resulting in a total activity of $6 \, MBq$ for the whole phantom. The outer chamber was filled with water, and nodule II (cold nodule) was left empty.

In our university clinic, the patients receive an injection of $67 \, MBq$ on average, and the mean uptake in the thyroid is 1.7%, resulting in approximately $1.1 \, MBq$. In general, the injected value is relatively constant, but the uptake is very patient-dependent. That results in about 10% of the patients having an uptake of less than $300 \, kBq$ and about 5% of the patients having an uptake of over $6 \, MBq$.

To understand how the uptake influences the result of the reconstructions, and to better cover the uptake range of patients, simulated lower uptake experiments were performed using the acquisitions with the phantom loaded as described before.

Therefore, we consider two additional virtual settings, 20% of the total activity (equalling $1.2 \, MBq$, approximately the average in our clinic) and 5% ($0.3 \, MBq$, a lower bound covering almost 90% of our patients). We employed statistical rejection sampling, using the temporal histogram of the original detections to reject measurements. Thus we produce a temporal histogram with approximately the same shape, but with the amplitude scaled down to the desired acceptance factor [85]. To better understand the effect of the statistical subsampling, we produced multiple noise realizations for each dataset.

In total, we repeated each of the four reconstruction experiments once with the full data (*high uptake*), 16 times with 20 % (*medium uptake*) and 16 times with 5 % (*low uptake*) of the originally acquired observations.

8.4.4   Planar Scintigraphy

Planar scintigraphy is the current main imaging modality for thyroid diagnosis. To emulate the results of a scintigraphy, which is using a significantly bigger, stationary gamma camera, our camera was positioned orthogonally in overlapping 21 positions that were afterwards stitched together to generate one full image. At each position the camera was held for $3\,s$, the images were up-sampled, the results in the overlapping regions were averaged, and the image was down-sampled again to $1\,mm$ isotropic square pixels.
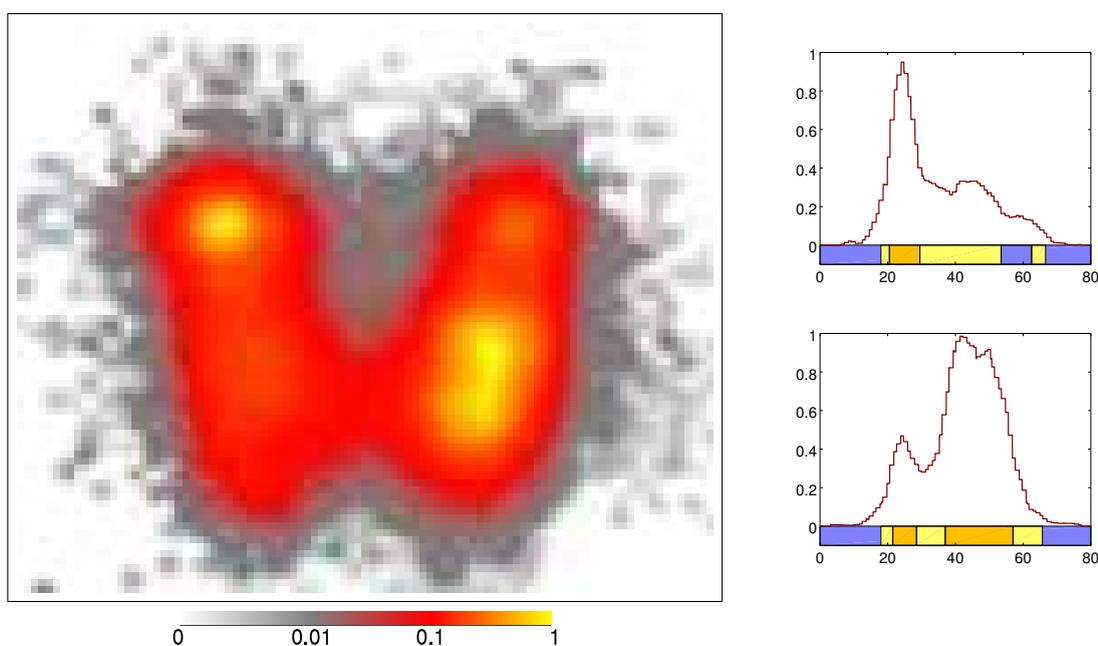
Figure 8.4: Planar scintigraphy and line profiles of the thyroid phantom, created using the gamma camera in 21 stationary positions, averaging the overlapping areas. The profiles were extracted from the planar image in a similar fashion as the ones presented in Figure 8.7[83].

8.5   Results

Figure 8.4 shows the planar scintigraphy of the thyroid phantom acquired using our setup. In comparison, Figure 8.5 shows the three-dimensional reconstructions using both ML-EM and LM-EM for the "step and shoot" (SAS) and the "continuous" (Cont.) datasets when using all the original detections, i.e. the high uptake data set. The first row shows the equivalent coronal center cross-sections through the reconstructed radioactivity distributions. The

transverse cross-sections shown in the second row are positioned to contain the centers of nodules I and III, and the transverse cross-sections in the third row are positioned to show the isthmus connecting both thyroid lobes, above nodules II and IV. The location of the coronal cross-section and the two transverse planes with respect to the thyroid phantom is indicated in Figure 8.2a. All intensities are scaled to the interval $[0, 1]$ for each reconstruction individually, and the colors are represented in a nonlinear fashion in order to highlight low contrast regions.

As seen in Figure 8.5, the two algorithms (ML-EM and LM-EM) yield very similar reconstruction results for the same dataset, extending even to the shape of artifacts. In all settings, the thyroid is reconstructed and hot and cold nodules are discernible. However, the isthmus between the lobes is not reconstructed very well when using the SAS trajectory, erroneously bulging away from the center.
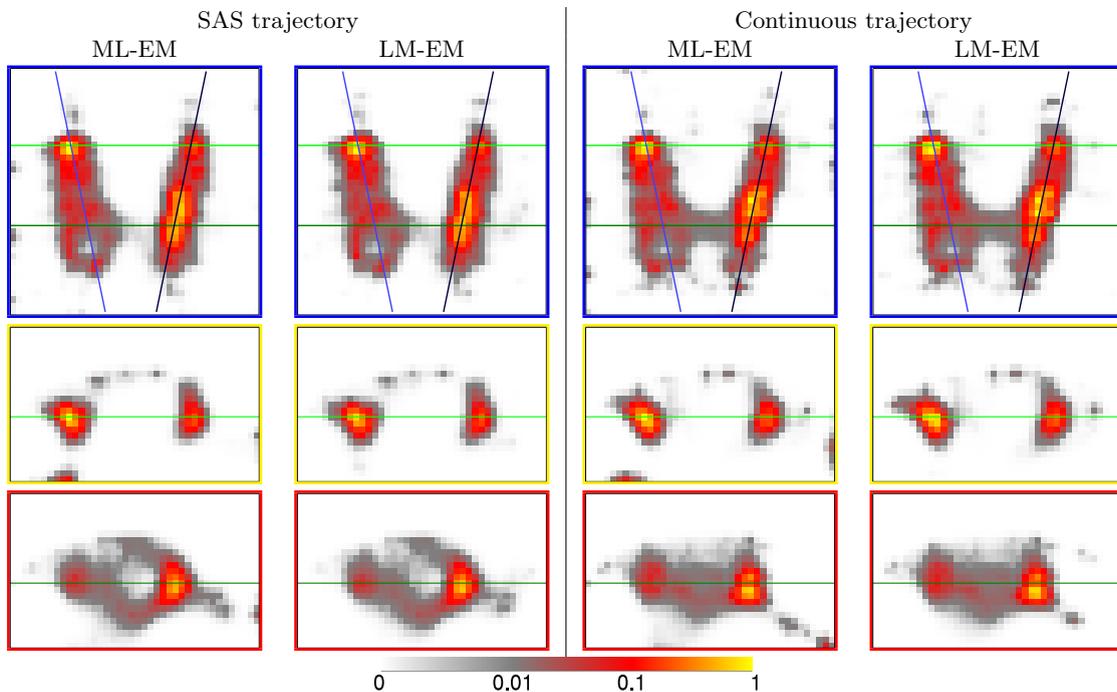
Figure 8.5: Reconstruction results of all four methods using the *high uptake* (6 MBq) data set. Shown are coronal cross-sections (top row) as well as two transverse cross-sections (middle and bottom row). The locations of the cross-sections is illustrated in Figure 8.2a. All intensities are scaled to the interval $[0, 1]$ for each reconstruction individually[83].

Figure 8.6 compares the same cross-sections for the high uptake data set (6 MBq), one representative simulated medium uptake data set (1.2 MBq, corresponding to the average patient in our clinic), and one representative simulated low uptake data set (300 kBq, corresponding to less than 10% of the patients in our clinic regardless of pathology). This time we only show the results using the SAS trajectory and ML-EM reconstruction as well as the continuous trajectory and LM-EM reconstruction, as these are the only practically relevant methods.

With decreasing uptake the image quality deteriorates, as expected. In particular, for both methods the reconstruction of the thyroid background loses homogeneity with decreasing uptake, while maintaining the overall characteristics as described for the high-uptake experiments.
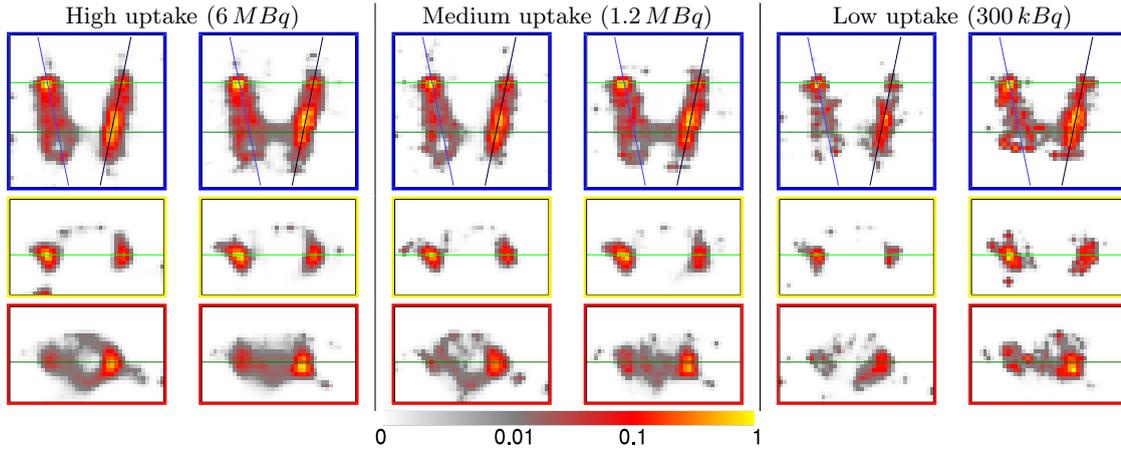
Figure 8.6: Reconstruction results of data sets with high, medium, and low uptake using the SAS trajectory and ML-EM (left columns) and the continuous trajectory and LM-EM (right columns). Shown are coronal cross-sections (top row) as well as two transverse cross-sections (middle and bottom row). The locations of the cross-sections is illustrated in Figure 8.2a. All intensities are scaled to the interval $[0, 1]$ for each reconstruction individually[83].

Figure 8.7 shows line profiles extending in sagittal direction through the nodules in both the left and right lobe of the thyroid phantom, respectively. The profile in the left lobe passes through the centers of nodules I and II and the profile in the right lobe passes through the centers of nodules III and IV, as indicated in Figure 8.2a. For reconstructions using the SAS trajectory and ML-EM, the profiles are marked in red, for reconstructions using the continuous trajectory and LM-EM, the profiles are marked in green.

Numerical magnitudes for quantitative comparison are given in Figure 8.8 for all experiments. The Contrast Recovery Coefficient (CRC) is presented for each nodule, calculated as $\frac{c_{r1}/c_{r2}-1}{c_1/c_2-1}$, where c_{r1} is the mean concentration at the nodule location, c_{r2} is the mean value of the background, and c_1/c_2 is the ground truth ratio between them. In the perfect case, the CRC has a value of 1, a value smaller or bigger than 1 indicates under- or over-estimation, respectively.

In detail, the CRC was calculated using the following procedure: The location of each nodule was extracted from the phantom's drawings, and then used to select the corresponding voxels from the reconstruction, minus a small border (0.5 voxels, or $1\,mm$). c_{r1} is then computed as the average activity of these voxels, where partially selected voxels at the borders were weighted accordingly. To obtain the mean background activity c_{r2}, two spheres of the same size as the medium nodules were extracted on each side of the thyroid, where it was certain that the area belonged to the background, and then the contained activity was averaged.
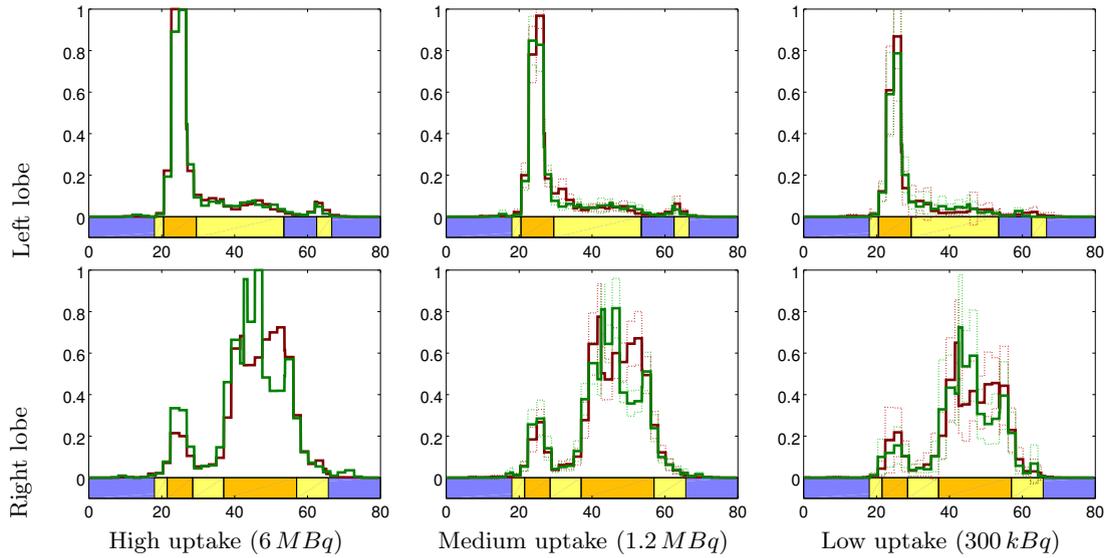
Figure 8.7: Profile plots for the reconstructions computed from high, medium, and low uptake data. For the medium and low uptake, the plot shows the average value of the realizations, and the dotted profiles correspond to one standard deviation away from the mean. The profiles for the reconstructions using SAS trajectory and ML-EM are shown in *red*, the ones using continuous trajectory and LM-EM are shown in *green*. The ground truth is shown as a colored bar below the profile; *orange* for 100% activity (hot nodules), *light yellow* for 10% activity (background) and *blue* for 0%(cold nodule and areas outside the phantom). The profile in the left lobe passes through the centers of nodules I and II and the profile in the right lobe passes through the centers of nodules III and IV, as indicated in Figure 8.2a[83].

In general, for high uptake, the four combinations of the two datasets and the two reconstruction algorithms yield very similar results: The CRC of the large hot nodule IV is overestimated (as being too radioactive), and the ratios of all other nodules are slightly underestimated. For medium uptake, ML-EM SAS also overestimates the ratio of the more separate hot nodule I. For low uptake, the results are somewhat more diverse, and in particular, the large hot nodule IV is over-estimated in all cases. Comparing all settings, the CRC of cold nodule II is always underestimated, i.e. reconstructed as it is less radioactive than in reality.
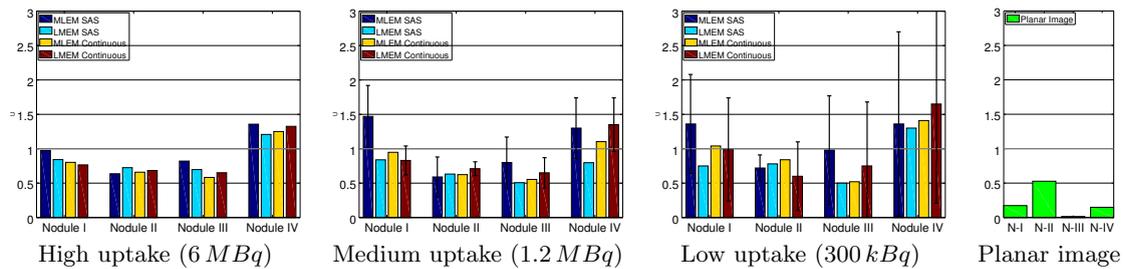
Figure 8.8: Contrast recovery coefficients for nodules I to IV in the thyroid phantom, computed for each of the four reconstruction methods using the high, medium, and low uptake data sets. As comparison, the contrast recovery coefficients obtained from the planar image are presented on the right[83].

Finally, Table 8.1 provides the time required for trajectory execution and the number of

detected emissions for each experiment. In addition to these values, the average time required to perform one iteration of the reconstruction algorithm is specified (a total of 166 iterations was performed for each reconstruction), using our own custom software package executed on a dual Intel Xeon (E5-2687W) machine with $64\,GB$ of RAM.

| Data Set | high uptake | | medium uptake | | low uptake | |
Trajectory	SAS	Cont.	SAS	Cont.	SAS	Cont.
trajectory execution (s)	654	293	654	293	654	293
detected emissions	323859	313671	65056	63352	16358	15725
ML-EM iteration (s)	34	459	34	454	35	453
LM-EM iteration (s)	140	146	42	43	12	13

Table 8.1: Time required (in seconds) for trajectory execution as well as number of detected emissions for each experiment. Additionally, average time required (in seconds) for one iteration of the reconstruction algorithms (ML-EM and LM-EM) for each experiment[83].

8.6 Discussion

In our experiments we compare two acquisition protocols, *"step and shoot"* (SAS) and *"continuous"*, together with two reconstruction methods, binned-data ML-EM and list-mode LM-EM. While the SAS trajectory lends itself quite well to binned-data ML-EM, and conversely the list-mode LM-EM fits perfectly to the continuous trajectory, the two other combinations SAS/LM-EM and continuous/ML-EM are feasible as well and produce reconstructions with very similar characteristics overall: As shown in Figure 8.5 for the high uptake data, all three hot nodules are clearly visible in all four methods, and it is also possible to infer the cold nodule with each method. The contrast recovery coefficients for the same data set are also fairly consistent across all four methods, as seen in the left column of Figure 8.8. Compared to the planar image in Figure 8.4, the cold nodule visibility is the biggest improvement in the reconstructions. Additionally, the continuous trajectory reconstructions resolve the isthmus between the two lobes of the thyroid phantom better, yielding a significantly more accurate representation of that area. We hypothesize that this is caused by the increased coverage of the central area, as the gamma camera sensor, which is too small to image the entire isthmus area at once, can acquire more different view points while continuously moving as opposed to the few viewpoints acquired by the SAS trajectory.

For the data set with medium uptake (middle columns of Figures Figure 8.6 and Figure 8.8), the outcome is similar to the one of the high uptake data, but slightly more noise is visible in the images, as expected from the reduction of counts. In fact, for all noise realizations, the standard deviation of the background activity increases by 79% compared to the high uptake data. The four nodules are qualitatively recovered clearly, including the cold nodule.

The corresponding CRC values of all nodules are very close to the high uptake values, with the exception of SAS ML-EM, which is now markedly overestimating nodule I.

In the low uptake results (right columns of Figures Figure 8.6 and Figure 8.8) the noise is even more dominant. Over all noise realizations, the standard deviation of the background activity now increases by 109% compared to the high uptake data. Qualitatively, the cross-sectional images shown no longer allow a clear determination of which regions represent cold and hot nodules, except for the big nodule IV, which is still visible. When taking into account the full three-dimensional reconstructed image, however, nodule I and nodule II (the cold nodule) can be inferred for example from the line profiles, see Figure 8.7. This is also reflected by the CRC values in Figure 8.8. Since the background values in these images are close to zero, the calculation of the contrast values are dominated by these regions, explaining the highly overestimated values, in particular for nodule IV.

In terms of measurement time (see Table 8.1), the continuous trajectory is clearly preferable, as it allows continuous movement while acquiring data. In our example, actual measurement time was cut in half compared to SAS, while still offering comparable detection statistics.

In terms of computational effort (see again Table 8.1), the combination of SAS trajectory and binned data MLEM performs very fast, as there is a limited number of static detection poses, and computational complexity is roughly proportional to this number. While the combination of SAS and list-mode LM-EM is feasible, yielding comparable image quality, there is a big drawback in terms of computational effort, except for the case of very low detection statistics. When using the continuous trajectory, the combination with binned-data MLEM is again feasible, yielding comparable image quality, but due to the huge number of bins required to accurately represent the data of the moving detector, the computational effort for reconstruction is needlessly high. The LM-EM method is clearly the better choice for continuously acquired data. It results in fast reconstruction times, and since the computational effort is bound to the number of events and not camera poses, it can also result in reconstruction times faster than SAS ML-EM for low uptakes.

In general, all four presented methods are suitable for mini gamma reconstructions of the proposed thyroid phantom. However, the continuous trajectory achieves a better coverage of the phantom, resulting in better resolved images in the center region of the phantom, while also providing significantly faster acquisition speeds. Meanwhile, for continuous trajectories, the LM-EM reconstruction method is a natural fit, yielding comparable images to ML-EM but significantly faster reconstruction times.

Overall, our proposed approach using mini gamma cameras for three-dimensional imaging provides image quality between scintigraphy and SPECT, while allowing significantly faster

acquisition times. In particular, our proposed approach is more easily affordable and enables the visualization of cold nodules, as demonstrated in case of our thyroid phantom.

We note that the field of view of the mini gamma camera is limited and usually cannot cover the entire region of interest at once. As a result, additional uptake, such as in salivary glands, might be in the field of view only for some of the recorded view points, potentially leading to artifacts in the reconstructed image.

Another important point to consider for clinical implementation is the tracking and guidance of the camera. One option is to remove the robot, to add a tracking system of sufficient accuracy for localization of the detector, such as outside-in [73] or inside-out [86] tracking, and to leave the movement of the camera entirely to a human operator. This approach is currently used in radio-guided surgery on an open situs[26, 87]. Naturally, this implies a continuous trajectory and list-mode data LM-EM reconstruction. The main disadvantage of the hand-held approach is the loss of repeatability and the weight of the gamma camera that has to be borne by the human operator.

While using a robotic arm eliminates these disadvantages, it requires developments to ensure the required safety for patient applications. Collaborative medical robotic systems that interact with a human operator in order to provide nuclear imaging in addition to ultrasound[88], or that can autonomously acquire ultrasound images on human probands [89] are a precursor for the wider introduction of robotic imaging. Additionally, a robot enables the incorporation of optimized trajectories based on the scanning geometry [28] to even further reduce acquisition times. A more elaborate option are acquisition optimization schemata, where, for example, the robot adjusts the scanning speed to provide a simultaneously good spatial coverage and good detection statistics, generating patient-specific scans. Robotic nuclear imaging systems could find their application in real-time imaging of the radioactivity distribution during radioembolization of the liver[90], or the application of radioguidance to laparoscopic surgeries[91].

8.7   Conclusion

In this work we have presented a way to acquire and reconstruct extended phantoms using a mini gamma camera and a robotic arm. For the tests we generated two datasets, a *step and shoot* and a *continuous* scan, following the same trajectory around a thyroid phantom with multiple nodules. Both datasets were then sub-sampled to simulate lower activity, and afterwards the original and the new datasets were reconstructed with both ML-EM and LM-EM algorithms.

The reconstruction quality in all four cases produced satisfactory results for all the nodules. This is specially relevant with cold nodules, since it is not possible to detect them in the scintigraphy simulation, albeit its clinical relevance.

The combination of continuous trajectory and LM-EM emerged as the clear favorite. The continuous trajectory enabled a better coverage of the phantom, improving the reconstruction results and reducing the acquisition time. The results of the reconstruction algorithms were fairly equivalent, but for the continuous acquisition LM-EM generated the results much faster.

Chapter 9

Combining Robotic freehand SPECT and C-Arm CT

9.1 Introduction

As presented in Chapter 7, the reconstruction quality of robotic freehand SPECT with a Gamma Camera was clearly improved with respect to the gamma probe case. In this line, the idea of having an intraoperative SPECT device based on this Robot plus Gamma Camera combination appeared as feasible. The goal here was to create a first prototype of an *intaoperative SPECT-CT* by using the results of a C-Arm CT to guide a robotic SPECT acquisition and afterwards combine the results of both devices.

The main challenges that we faced were, on the one hand, to solve the co-registration of the systems and, on the other, the surface extraction from a C-Arm CT and on-the-fly planning, based on this extraction and the selected region of interest.

The work presented here is part of the paper "Towards personalized interventional SPECT-CT imaging", presented at "Medical Image Computing and Computer Assisted Interventions (MICCAI)" in 2014[90].

9.2 Clinical Motivation

One particular example of an interventional application where a device like this is highly desirable, is the brachytherapy of unresectable liver tumors, like hepatocellular carcinoma (HCC). Brachytherapy in the form of radioembolization (Selective Internal Radiotherapy – SIRT) is an alternative to classical chemoembolization or chemotherapy [92][93]. In this

case, microspheres loaded with ^{90}Y are injected into the hepatic arteries using a catheter for selective internal irradiation of the tumor cells. ^{90}Y mainly undergoes β^- decay, emitting electrons that are absorbed within at most 11 mm of tissue, making it possible to inject very high doses within a single treatment. It is, however, crucial to ensure that the radioactive compound remains confined to the injection site in order to irradiate only the surrounding cancer tissue. Any leakage to other parts of the liver or other organs, for example through a shunt to the lung, must be prevented.

The current practice is to inject ^{99m}Tc-MAA through a catheter first, which has demonstrated good prognostic value for the ^{90}Y distribution [94]. The patient is then transferred from the radiology department, where the catheterization and injection take place, to the nuclear medicine department for a whole-body SPECT scan, usually combined with CT, to monitor the ^{99m}Tc distribution. This process takes four to six hours on average and requires, if the result is positive, a second intervention in the radiology department with another catheterization for the actual radioembolization with ^{90}Y itself.

As shown by this example, the use of a diagnostic device incurs organizational complexity and prolongs the duration of the intervention. For this reason, we propose a novel approach to interventional SPECT/CT imaging consisting of a C-arm CT scanner and a robot-controlled gamma camera, where the latter's trajectory is optimized based on the patient's anatomy as extracted from the C-arm CT data. This approach enables clinicians to perform interventions like radioembolization on a single site and during a single procedure, substantially reducing patient stress and time commitment of the medical personnel. On top of this, the core imaging process also benefits from this fusion, as the required data can be collected very efficiently in a patient-specific manner, leading to equivalent medical information significantly faster and with similar accuracy as currently available in the state-of-the-art whole-body scenario.

9.3   Materials and Methods

9.3.1   Overview

As shown in Figure 9.1, we extend an angiography suite by placing a robotic arm with a mini gamma camera next to the C-arm. Using appropriate calibration techniques as detailed in Section 9.3.2, we can co-registered them, and acquire all relevant data, X-ray transmission and gamma radiation, with respect to a single, common coordinate frame.

With this setup in place and calibrated, we first acquire a CT image of the volume of interest. Based on it, we extract the convex hull of the patient's (or phantom's) surface and compute

an optimal trajectory for acquiring the SPECT images from minimal, but safe distances. Finally, we record the emission data by moving the gamma camera along the trajectory using the robotic arm and reconstruct the tracer distribution using likelihood-based tomographic reconstruction, as described in Chapter 4.
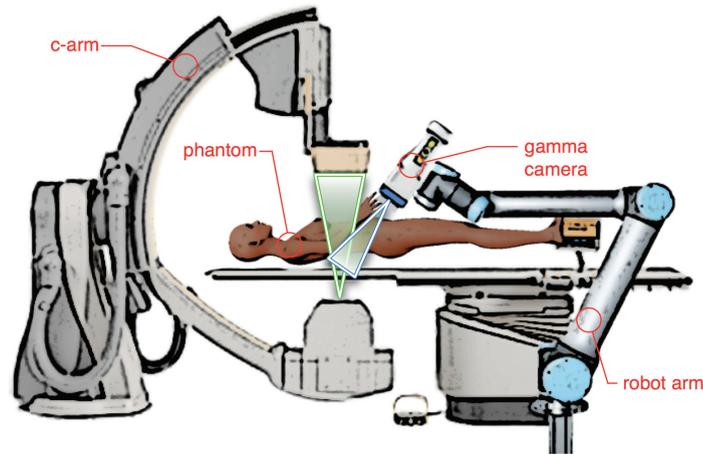
Figure 9.1: Combined C-arm CT and robotic SPECT setup in the operating room[90].

9.3.2 Coordinate systems calibration

Since we have fairly limited access to the C-Arm in terms of configuration or coordinate systems, the most reasonable approach is to use the C-Arm coordinate system, and calibrate the robotic arm to the C-Arm.

For this purpose, we mount a custom-designed calibration target, shown in Figure 9.2a, to the gamma camera (replacing the collimator), as shown in Figure 9.2b. The target contains special CT marker spheres in a well-defined pattern, that can be segmented easily in the CT volume. Using point-based registration as suggested by Umeyama [95], we obtain the transformation between the CT image and the calibration target, which allow us to compute the robot-base-to-CT-image transformation, thanks to the forward kinematics of the robot.

9.3.3 Trajectory planning

The first step of the core acquisition protocol consists of recording the X-ray component of the joint signal by rotating the C-arm over 180° around the region of interest and reconstructing it into a 3D volume.

In a second step, we compute a model of the patient's (or phantom's) surface from the CT image. As X-ray transmission through the surrounding air is significantly higher than

(a) Calibration tool

(b) Calibration setup

Figure 9.2: a) CAD drawing of the calibration tool and b) setup used to map the C-arm coordinate system to the robot coordinate system. The spheres attached to the fins appear prominently in the CT volume and can be easily segmented[90].

through tissue, it is possible to distinguish both regions by applying a threshold. The convex hull is then a suitable mesh-model for the surface.

Finally, the trajectories are generated as parallel scan-lines normal to the convex hull of the patient (up to a certain security distance, 1 cm in this case), as shown in Figure 9.3. Note that, unlike a fixed gantry in diagnostic SPECT, the camera poses are much closer for a personalized trajectory with better detection statistics.
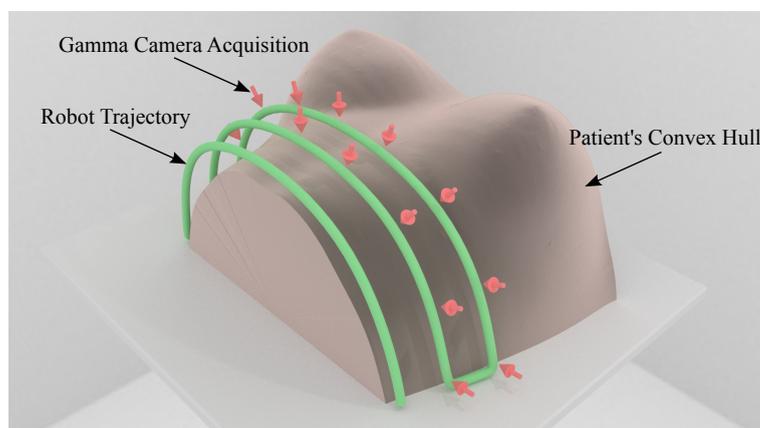
Figure 9.3: Schematic representation of an in-progress personalized trajectory for the robot-guided gamma camera along the convex hull of the patient/phantom as extracted from CT. The green line represents the planned robot trajectory, that follows the convex hull plus the safety margin (1 cm), the red arrows indicate positions where the gamma camera acquires images[90].

9.3.4   Image reconstruction

As shown in previous chapters, the reconstruction was done using MLEM, given that the data was acquired using a Step and Shoot procedure. The gamma camera characterization procedure, the same as the one used in the previous chapters, is explained in Section 7.2.

The X-ray CT reconstruction is directly obtained from the software solution of the C-arm manufacturer, and is based on Feldkamp's variant of Filtered Back-Projection [96]. As the SPECT measurements have been collected relative to the C-arm's coordinate frame, the two volumetric images are inherently aligned.

9.4   Experiments

9.4.1   Setup

The imaging setup consists of a C-arm CT and a mini gamma camera mounted on a robotic arm, as illustrated in Figure 9.1. The C-arm CT is a Philips Allura Xper FD20 set up in an angiography room. The mini gamma camera (CrystalCam, Crystal Photonics, Germany) is described in Section 7.3.1. The camera is mounted on a robotic arm (UR5, Universal Robots, Denmark) placed next to the C-arm, see Figure 9.4b.

We use software provided by ROS INDUSTRIAL (USA) that we extend with our own collision meshes, a resampling of the existing visual meshes and a modification of the joint limits in order to improve its inverse kinematics.

9.4.2   Experimental procedure

Verification experiment

In a first experiment, we verify the quality of our calibration procedure. We attach a single ^{57}Co point source to a torso phantom and obtain a CT image. Based on the location of the point source segmented from the CT image, we instruct the robot with the gamma camera to acquire a single image centered on the point source. The expected result is a gamma camera image with the point source showing as a hotspot in the center.

Reconstruction experiment

In a second experiment, we place a human torso phantom with a liver model made out of candle gel (Ceraflex N 530 transparent, Wachs-und Ceresinfabriken, Germany) on the exam table. Liver arteries are imitated by plastic tubes ($4\,mm$ diameter), filled with $4\,MBq$ of ^{99m}Tc.

We obtain a CT image of the torso and compute the trajectory for the SPECT acquisition according to the procedure described in section Section 9.3.3. The emission data recorded

by the robot-positioned gamma camera is reconstructed as an image inherently co-registered with the CT image. We expect to see the liver arteries in the reconstructed SPECT image embedded in the anatomical information from the CT.

9.5   Results

9.5.1   Verification experiment

Figure 9.4a shows a planar scintigraphy of the ^{57}Co source, acquired with the gamma camera instructed to point its center at the source (see Figure 9.4b), based on the source position as detected from the C-arm CT as well as the coordinate system calibration. The relocalization error, i.e. the offset from the hotspot's center of mass to the center of the image, is 1.6 camera pixels (corresponding to 3.9 mm).

(a) Gamma camera image of point source   (b) Photo of gamma camera acquiring image shown in (a)
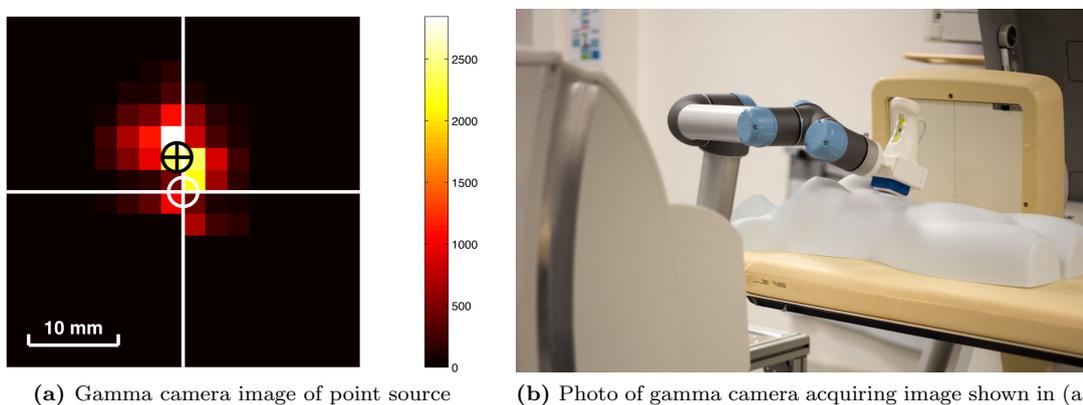
Figure 9.4: Result of calibration verification experiment[90].

9.5.2   Reconstruction experiment

Figure 9.5a shows a 3D rendering of the robotic SPECT image of our phantom. The branching blood-vessels simulated by three activity-filled tubes are clearly distinguishable. The co-registered robotic SPECT/C-arm CT slice image in Figure 9.5b shows that the offset between the tubes in the SPECT and the CT image is minimal.

9.6   Discussion

Calibration of the C-arm and gamma camera coordinate systems is a vital part of the imaging protocol. Our first experiment validates our approach, as the relocalization error is within the acceptable error limit for interventions. In addition to this, we performed a

(a) Robotic SPECT   **(b)** Slice of fused SPECT/CT
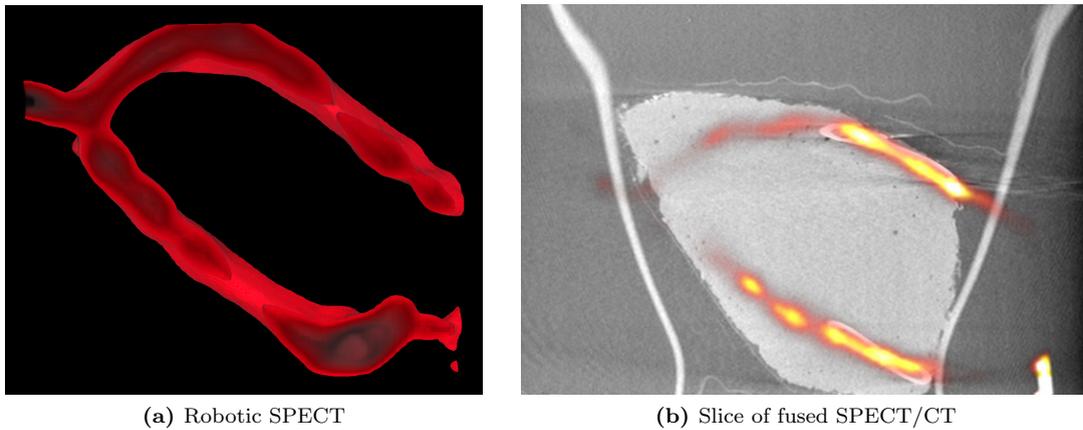
Figure 9.5: (a) 3D rendering of the SPECT image of the liver phantom with radioactive blood vessels and (b) slice of fused robotic SPECT/C-arm CT[90].

comparable experiment under guidance of an optical tracking system, and only achieved considerably higher errors.

The second experiment confirms the possibility of acquiring bi-modal interventional robotic SPECT and C-arm CT images in a clinical environment, with a realistic liver phantom. We were able to extract a patient-specific acquisition trajectory based on the convex hull of the torso and record a SPECT image that is naturally aligned to the C-arm CT image.

Our setup is limited by the C-arm that does not allow any customization of the acquisition protocol. A C-arm CT mounted on a robotic arm [97] would overcome this limitation, provided that the respective control interface is accessible. It would further allow to define personalized C-arm CT acquisition trajectories as suggested by Stayman et al. [98]. Also, the spectrum-based trajectory optimization for SPECT as proposed by Vogel et al. [28] could be considered. Both would lead to fully personalized interventional SPECT/CT acquisition.

9.7   Conclusion

This paper presents the first prototype for an interventional SPECT/CT scanner. In recent time, the medical community has shown considerable interest in such a device, and simulation studies have been published, for instance by Bowsher et al. [99], where a much heavier robot is used, which makes relocation or removal rather difficult. In contrast, our device is very light, thus meeting the accessibility and flexibility requirements of clinical practice.

The main advantages of an interventional SPECT/CT, as the one proposed here, are a reduction of the duration and complexity of the interventions, plus the ability to continuously monitor the progress in an easy and flexible way (for instance, by acquiring a bremsstrahlung

SPECT image [100]). The latter can also be used to quickly detect anomalies in the procedure such as leaks, thus enabling effective quality control.

The suggested method is ready for further investigation in the operating room, opening the path for alternative applications, for example in the head and neck region.

Conclusions

This work presented a small overview of image modalities, in order to put nuclear medicine in context with respect of other common imaging approaches. The goal was to highlight the crucial role it has by allowing physicians to understand how a certain structure inside the body is working (functional imaging) versus how it looks (anatomical imaging).

The core of of the first part of the thesis is an explanation of the working principles of nuclear medicine and, in particular, freehand SPECT, together with the mathematical models used to generate tomographic images. An introduction to robotics close the states of the art chapters.

The second part shows clear milestones we reached by trying to make freehand SPECT more predictable in terms of results and aim to understand how different parameters affect the reconstruction quality.

The first acquisition and reconstruction attempts, shown in Chapter 6 cannot be really seen as a success, but more like a turning point. The results showed that there was no real gain by using a robot instead of a human when scanning with a single pixel detector.

The real advantages of the robot appeared by using the $16x16$ pixel gamma camera with its characterization (Chapter Chapter 7), combination that made it possible to localize point sources with sub-millimetric precision, an order of magnitude better than previous reconstructions using a single pixel detector. It is also possible to observe how different views of the camera with respect to the phantom define how much can be resolved by the reconstruction algorithms. The latter is consistent with the experience of limited angle tomography.

The combination of the robotic freehand SPECT with a C-Arm CT (Chapter 9) showed us that this setup is feasible and makes sense, since the C-Arm image may be used to direct the robot to a certain region of interest, reducing the scanning and reconstruction time. It also made us aware of several limitations in our setup, being the most important one the camera size. Scanning a human-sized liver (about 1.6 L) with a small $40x40$ mm^2 detector proved

to be a real challenge, and the reconstruction quality also suffered. Those experiments were also our first ones using extended phantoms, instead of point sources as hot nodules.

After these, two facts were clear: the targets to scan had to be proportional to the detector size and, by scanning in a step-and-shoot manner, we were wasting time and limiting the reconstruction capabilities of our setup. That pushed the idea of moving to continuous acquisitions and also to move from the well known MLEM reconstruction algorithm to List Mode (Chapter 8), to efficiently make use of the new scans. For the first time in our setups, our phantom had different radioactive isotope concentrations in different chambers, making it more realistic, but also more difficult to reconstruct.

The present work took several years and, in the process, I was able to draw several key conclusions of freehand SPECT, robotics, detectors and reconstructions, highlighted here:

- Freehand SPECT makes sense inside the operating theater, and gamma probes (single pixel detectors) make only sense if it is not possible to use gamma cameras, like if imaging through natural orifice or minimally invasive surgeries.

- Robotic freehand SPECT makes procedures much more complicated, in terms of logistics and safety for the patient so that, unless those aspects are properly solved (including a better robotic arm than the one used in this thesis), the application on real patients is not feasible. On the other hand, the improvement in the image quality with respect to the handheld version, and the fact that it is possible to guarantee a certain coverage (spatial and angular) of the region of interest, makes it a more accurate and reliable option.

- The idea of having a robotic SPECT-CT inside the operating theater still makes sense for me, but looking at the past experience, a much bigger detector is needed (about six to ten times the area of the mini gamma camera). A collaboration with the C-Arm CT manufacturer would make the integration task much easier.

- In terms of scanning protocols, continuous scanning is clearly more attractive than step-and-shoot, both in terms of acquisition efficiency and angular coverage. The fact that the sensor is continuously acquiring while moving maximizes both of them.

- This continuous trajectory makes sense only with List Mode, otherwise the reconstruction times are prohibitive but, even in the case of List Mode reconstructions, a GPU implementation of the algorithm is mandatory. For freehand SPECT in general, this would be for me the preferred reconstruction in every case.

- Moving from the current pre-planned trajectories towards a more flexible scan, that takes advantage of the force feedback integrated in the Kuka robots, could be the way to have a safe robotic acquisition inside the operating room.

- The trajectory optimization from Jakob Vogel [28], computed during the acquisition, can produce minimal scanning trajectories that guarantee coverage and reconstruction quality, might further improve the acquisition and reconstruction times without risking the reconstruction quality.

What started like a side project of freehand SPECT, as a way to better understand how to scan, what collimators to use, etc., is today a different way to do freehand SPECT and, in some cases, the better way, at least from the side of reconstruction quality. There are, of course, many open challenges, most of them from the robotics point of view, that I am sure will be solved by the next generation of PhD students.

Appendix A

Reconstruction Methods

This appendix describes the fundamentals of both Maximum-Likelihood Expectation-Maximization and List-Mode Expectation-Maximization algorithms, used to generate the reconstructions in the second part of this thesis.

The work presented here is part of the paper "Flexible Mini Gamma Camera Reconstruction of Extended Sources using Step and Shoot and List Mode", published at "Medical Physics" [83] in 2016.

Fundamentals

As a non-negative quantity, the mean number of emissions within the volume $V \subset \mathbb{R}^3$ can be considered a function $f : V \to \mathbb{R}_0^+ = [0, \infty)$, discretized as

$$f(\cdot) \approx \widehat{f}(\cdot) = \sum_{j=1}^{J} x_j b_j(\cdot), \tag{A.1}$$

with $b_j(\cdot)$ denoting spatial basis functions, in our case isotropic cubic voxels, and $\mathbf{x} = (x_j)_{j=1}^{J}$ representing their respective coefficients – voxel-wise mean number of emissions, so to say. The latter are initially unknown, and will be computed during reconstruction. An approximation of f can later be recovered via (A.1). For simplicity, we just use index j to refer to voxel b_j.

Reconstruction itself is based on observation of nuclear decay events. Following a "preset time" acquisition protocol (rather than "preset counts"), both the *total number of detected*

emissions N as well as the individual occurrences are random. In particular, N is a realization of a Poisson distributed random variable with

$$\overline{N} = \sum_{j=1}^{J} x_j d_j, \qquad (A.2)$$

denoting its expectation. (For brevity we assume unit measurement times.) d_j is the *detection sensitivity* of voxel j, i.e. the probability that an emission from voxel j was detected, or in other words, a number giving an estimate for how well a voxel is covered during the measurements. Considering our flexible detector, the robotically controlled mini gamma camera, this magnitude can vary noticeably over j.

We use two different recording schemes – *binned-mode* for time-integrated count values, and *list-mode* for individually detected emissions. Each of them leads to a specific expectation maximization algorithm for reconstruction, to be outlined below. In both cases, the detector is moved robotically during acquisition of data, and we mark the different *detector poses* using the index $t \in \{1, \ldots, T\}$.

Furthermore, we need to know the *general measurement probability*

$$P_{tj} := P[\text{emission from voxel } j \text{ detected at detector pose } t] \qquad (A.3)$$

that, for example, can be measured by systematic long-term observation of the imaging setup. In traditional tomographic reconstruction nomenclature, P can be called *system matrix*, describing the characteristics of the imaging setup independent of actual distributions. We consider it known for now; the calibration procedure[77] is outlined in Chapter 7. Note that P relates to independent emissions and is hence valid for all recording modes. Also consider that geometrical information such as position and orientation of the gamma camera with respect to the volume of interest is indirectly contained within P.

Based on P, it is also possible to give an estimate for the voxel-wise detection sensitivity,

$$d_j = \sum_{t=1}^{T} P_{tj} \qquad (A.4)$$

Unrelated to actual emissions, d_j contains purely geometrical information only, describing how well voxel j is observed.

A.1 Maximum-Likelihood Expectation-Maximization

A common approach is to store counts $y_i \in \mathbb{N}_0 := \{0, 1, \ldots\}$ of emissions integrated over certain time frames. This corresponds to temporal binning, and may happen natively in the detector hardware, or retrospectively in software. The number of detector poses T then equals the number of measured time-integrated counts $\mathbf{y} = (y_i)_{i=1}^I$, and we can use the indices t and i interchangeably.

Now we introduce the discrete random variables Y_{ij} as the number of emissions from voxel j detected at time frame i. We assume, as is typical[35] and experimentally confirmed[9], that the Y_{ij} are independent and Poisson distributed with mean $P_{ij}x_j$, using the general measurement probability P_{ij} from equation (A.3). Let further the discrete random variables $Y_i = \sum_{j=1}^J Y_{ij}$ denote the number of emissions detected at time frame i. The Y_i are then again Poisson distributed with mean

$$(P\mathbf{x})_i := \sum_{j=1}^J P_{ij}x_j \tag{A.5}$$

Our task is now to estimate $\mathbf{x} = (x_1, \ldots, x_J)$ using the *binned-mode data likelihood function*

$$L_{bin}(\mathbf{x}) = \prod_{i=1}^I \exp\left(-(P\mathbf{x})_i\right)\frac{(P\mathbf{x})_i^{y_i}}{y_i!} \tag{A.6}$$

Ignoring constants irrelevant to subsequent optimization, the corresponding *binned-mode data log-likelihood function* is

$$\ell_{bin}(\mathbf{x}) = -\sum_{i=1}^I (P\mathbf{x})_i + y_i \log(P\mathbf{x})_i + \text{const} \tag{A.7}$$

A well-known algorithm maximizing (A.7) is the *Maximum-Likelihood Expectation-Maximization*[36] (ML-EM), iteratively updating a vector initialized as $\mathbf{x}^{(0)} = \mathbf{1}$ using the following multiplicative update

$$x_j^{(k+1)} = x_j^k \cdot \frac{1}{d_j} \sum_{i=1}^I P_{ij} \frac{y_i}{(P\mathbf{x}^{(k)})_i} \tag{A.8}$$

for $j = 1, \ldots, J$. Here, $d_j = \sum_{i=1}^I P_{ij}$ again denotes the voxel-wise detection sensitivity.

A.2   List-Mode Expectation-Maximization

In general, binning may lead to aliasing artifacts, as additional measurement errors are inherently introduced during the process, and it may be advisable to treat the original measurements separately. In our case, for instance, the camera will move during the acquisition, thus invalidating the static model implicitly assumed when using time-integrated count values. Therefore, in list-mode recording settings, instead of counts, every individual detected emission (or 'event') is stored in a list, and the total number of emissions is N, a realization of the Poisson random variable introduced earlier with mean $\overline{N} = \sum_{j=1}^{J} x_j d_j$ and probability mass function

$$P(N) = \exp(-\overline{N}) \frac{\overline{N}^N}{N!} \tag{A.9}$$

Let Z denote the set of all possible emissions, and let $\{z_1, \ldots, z_N\} \subset Z$ denote the list of detected emissions. The probability density function for an emission z_n on the list, $n \in \{1, \ldots, N\}$, is

$$h(z_n) = \overline{N}^{-1} \cdot \sum_{j=1}^{J} P_{nj} x_j \tag{A.10}$$

using the general measurement probability P_{nj} from equation (A.3).

Our task is now to estimate $\mathbf{x} = (x_1, \ldots, x_J)$ using the *list-mode data likelihood function*

$$L_{list}(\mathbf{x}) = P(N) \cdot \prod_{n=1}^{N} h(z_n) \tag{A.11}$$

Again ignoring constants irrelevant to subsequent optimization, the corresponding *list-mode data log-likelihood function* is

$$\ell_{list}(\mathbf{x}) = -\overline{N} + \sum_{n=1}^{N} \log \left(\sum_{j=1}^{J} P_{nj} x_j \right) + \text{const} \tag{A.12}$$

An algorithm maximizing (A.12) similar to ML-EM has been proposed[38, 39, 101], iteratively updating a vector initialized as $\mathbf{x}^{(0)} = \mathbf{1}$ using the following multiplicative update:

$$x_j^{(k+1)} = x_j^{(k)} \cdot \frac{1}{d_j} \sum_{n=1}^{N} P_{nj} \frac{1}{(P\mathbf{x}^{(k)})_n} \tag{A.13}$$

for $j = 1, \ldots, J$, employing the short-hand notation $(P\mathbf{x})_n := \sum_{j=1}^{J} P_{nj} x_j$. Here, $d_j = \sum_{t=1}^{T} P_{tj}$ again denotes the voxel-wise detection sensitivity, which is dependent only on the detector's trajectory, i.e. the poses $t \in \{1, \ldots, T\}$, and on the choice of collimation.

We use the term *List-Mode Expectation-Maximization* (LM-EM) to refer to this method.

Appendix B

Publication List

Journal Publications

2016 **J. Gardiazabal**, P. Matthies, J. Vogel, B. Frisch, N. Navab, S. Ziegler, T. Lasser. Flexible Mini Gamma Camera Reconstruction of extended sources using Step and Shoot and List Mode. *Medical Physics*

2014 P. Matthies, **J. Gardiazabal**, A. Okur, J. Vogel, T. Lasser, N. Navab. Mini Gamma Cameras for Intra-operative Nuclear Tomographic Reconstruction. *Medical Image Analysis*

2013 J. Vogel, T. Lasser, **J. Gardiazabal**, N. Navab. Trajectory optimization for intra-operative nuclear tomographic imaging. *Medical Image Analysis*

T. Reichl, **J. Gardiazabal**, N. Navab. Electromagnetic servoing - a new tracking paradigm. *IEEE Transactions on Medical Imaging*

Conference publications

2015 **J. Gardiazabal**, B. Frisch, P. Matthies, J. Vogel, S. I. Ziegler, N. Navab, T. Lasser. List-Mode Reconstruction for Continuous Freehand SPECT Acquisitions. *IEEE Nuclear Science Symposium and Medical Imaging Conference*

S. Habert, **J. Gardiazabal**, P. Fallavollita, N. Navab. RGBDX: first design and experimental validation of a mirror-based RGBD Xray imaging system. *International Symposium on Mixed and Augmented Reality (ISMAR)*

A. Duliu, **J. Gardiazabal**, T. Lasser, N. Navab. Reproducible high-resolution Multispectral Image Acquisition in Dermatology. *European Conferences on Biomedical Optics*

J. Gardiazabal, J. Vogel, P. Matthies, M. Wieczorek, B. Frisch, N. Navab, S. I. Ziegler, T. Lasser. Fully 3D thyroid imaging with mini gamma cameras. *Fully3D*

T. Lasser, **J. Gardiazabal**, M. Wieczorek, P. Matthies, J. Vogel, B. Frisch, N. Navab. Towards 3D thyroid imaging using robotic mini gamma cameras. *Bildverarbeitung für die Medizin*

2014 B. Frisch, M. Eiber, **J. Gardiazabal**, P. Matthies, T. Maurer, A. Okur, T. Lasser, N. Navab. Challenges for Intra-Operative SPECT. *IEEE Nuclear Science Symposium and Medical Imaging Conference*

A. Okur, C. Hennersperger, J.B. Runyan, **J. Gardiazabal**, M. Keicher, S. Paepke, T. Wendler, N. Navab. fhSPECT-US Guided Needle Biopsy of Sentinel Lymph Nodes in the Axilla: Is it Feasible? *Proceedings of MICCAI*

J. Gardiazabal, M. Esposito, P. Matthies, A. Okur, J. Vogel, S. Kraft, B. Frisch, T. Lasser, N. Navab. Towards personalized interventional SPECT-CT imaging. *Proceedings of MICCAI*

P. Matthies, **J. Gardiazabal**, A. Okur, T. Lasser, N. Navab. Accuracy evaluation of interventional nuclear tomographic reconstruction using mini gamma cameras. *Hamlyn Symposium on Medical Robotics*, London, UK

J. Vogel, A. Duliu, Y. Oyamada, **J. Gardiazabal**, T. Lasser, M. Ziai, R. Hein, N. Navab. Towards Robust Identification and Tracking of Nevi in Sparse Photographic Time Series. *Proceedings of SPIE*

2013 P. Matthies, K. Sharma, A. Okur, **J. Gardiazabal**, J. Vogel, T. Lasser, N. Navab. First use of mini gamma cameras for intra-operative robotic SPECT reconstruction. *Proceedings of MICCAI*

J. Gardiazabal, T. Reichl, A. Okur, T. Lasser, N. Navab. First flexible robotic intra-operative nuclear imaging for image-guided surgery. *IPCAI*

2012 J. Vogel, T. Reichl, **J. Gardiazabal**, N. Navab, T. Lasser. Optimization of Acquisition Geometry for Intra-operative Tomographic Imaging. *Proceedings of MICCAI*

2011 M. Feuerstein, H. Heibel, **J. Gardiazabal**, N. Navab and M. Groher. Reconstruction of 3-D Histology Images by Simultaneous Deformable Registration. In *Proceedings of MICCAI*

Bibliography

[1] J. Shapiro, *Radiation Protection: A Guide for Scientists, Regulators and Physicians, 4th Edition.* 2002.

[2] "The electromagnetic spectrum (modified)." http://commons.wikimedia.org/wiki/File:EM_spectrum.svg.

[3] T. M. Buzug, *Computer Tomography.* 2008.

[4] G. Walraven, *Basic Arrhythmias, 7th Edition.* 2010.

[5] "Merriam-Webster dictionary and thesaurus." http://www.merriam-webster.com/dictionary/quasiperiodic.

[6] R. S. C. Cobbold, *Foundations of Biomedical Ultrasound.* 2006.

[7] S. Bhargava, C. Brugera, G. G. Cerri, S. H. Eik-Nes, M. R. El-Meligi, H. A. Gharbi, B. B. Goldberg, H. T. Lutz, C. R. B. Merrit, R. D. E., D. Szabunio, M. W. Wachira, H. Wanatbe, P. N. T. Wells, and X. Zhi-Zhang, "Training in diagnostic ultrasound: Essentials, principles and standards.," *WHO Technical Report Series*, 1998.

[8] C. S. Levin, "New imaging technologies to enhance the molecular sensitivity of positron emission tomography," pp. 439–467, 2008.

[9] F. Cannizzaro, G. Greco, S. Rizzo, and E. Sinagra, "Results of the measurements carried out in order to verify the validity of the poisson-exponential distribution in radioactive decay events," *The International Journal of Applied Radiation and Isotopes*, vol. 29, no. 11, pp. 649 – IN1, 1978.

[10] "Bremsstrahlung (breaking radiation), Encyclopædia Britannica.." http://www.britannica.com/EBchecked/topic/78784/bremsstrahlung.

[11] S. S. Chang, "Overview of prostate-specific membrane antigen," *Rev Urol*, vol. 6 Suppl 10, pp. S13–18, 2004.

[12] J. S. Ross, C. E. Sheehan, H. A. Fisher, R. P. Kaufman, P. Kaur, K. Gray, I. Webb, G. S. Gray, R. Mosher, and B. V. Kallakury, "Correlation of primary tumor prostate-specific membrane antigen expression with disease recurrence in prostate cancer," *Clin. Cancer Res.*, vol. 9, pp. 6357–6362, Dec 2003.

[13] D. S. O'Keefe, D. J. Bacich, and W. D. Heston, "Comparative analysis of prostate-specific membrane antigen (psma) versus a prostate-specific membrane antigen-like gene," *The Prostate*, vol. 58, no. 2, pp. 200–210, 2004.

[14] L. A. Shepp and B. F. Logan, "The Fourier Reconstruction of a Head Section," *Transactions on Nuclear Science*, vol. NS-21, 1974.

[15] T. Wendler, J. Traub, S. Ziegler, and N. Navab, "Navigated three dimensional beta probe for optimal cancer resection," in *Proceedings of MICCAI 2006* (R. Larsen, M. Nielsen, and J. Sporring, eds.), vol. 4190 of *LNCS*, (Copenhagen, Denmark), pp. 561–569, MICCAI Society, Springer, Oct. 2006.

[16] "SurgicEye GmbH." http://www.surgiceye.com.

[17] Z. Nahas, D. Goldenberg, C. Fakhry, M. Ewertz, M. Zeiger, P. W. Ladenson, R. Wahl, and R. P. Tufano, "The role of positron emission tomography/computed tomography in the management of recurrent papillary thyroid carcinoma," *Laryngoscope*, vol. 115, pp. 237–243, Feb 2005.

[18] M. S. Lenhard, A. Burges, T. R. Johnson, P. Stieber, C. Kumper, N. Ditsch, R. Linke, and K. Friese, "PET-CT in recurrent ovarian cancer: impact on treatment planning," *Anticancer Res.*, vol. 28, no. 4C, pp. 2303–2308, 2008.

[19] M. Baumhauer, M. Feuerstein, H. P. Meinzer, and J. Rassweiler, "Navigation in endoscopic soft tissue surgery: perspectives and limitations," *J. Endourol.*, vol. 22, pp. 751–766, Apr 2008.

[20] L. Maier-Hein, A. Tekbas, A. Seitel, F. Pianka, S. A. Muller, S. Satzl, S. Schawo, B. Radeleff, R. Tetzlaff, A. M. Franz, B. P. Muller-Stich, I. Wolf, H. U. Kauczor, B. M. Schmied, and H. P. Meinzer, "In vivo accuracy assessment of a needle-based navigation system for CT-guided radiofrequency ablation of the liver," *Med Phys*, vol. 35, pp. 5385–5396, Dec 2008.

[21] C. M. Tempany, J. Jayender, T. Kapur, R. Bueno, A. Golby, N. Agar, and F. A. Jolesz, "Multimodal imaging for improved diagnosis and treatment of cancers," *Cancer*, vol. 121, pp. 817–827, Mar 2015.

[22] T. Wendler, A. Hartl, T. Lasser, J. Traub, F. Daghighian, S. Ziegler, and N. Navab, "Towards Intra-operative 3D Nuclear Imaging: Reconstruction of 3D Radioactive

Distributions Using Tracked Gamma Probes," in *Springer LNCS 4792 (MICCAI)*, pp. 909–917, 2007.

[23] T. Wendler, A. Hartl, T. Lasser, J. Traub, S. Ziegler, and N. Navab, "3d intra-operative nuclear imaging for slnb in neck," *The Journal of Nuclear Medicine (JNM)*, vol. 49, no. 1, pp. 408P–d, 2008.

[24] H. Gray, *Gray's Anatomy of the Human body, 20th edition. Currently in public domain.* 1918.

[25] K. A. Kern, "Concordance and validation study of sentinel lymph node biopsy for breast cancer using subareolar injection of blue dye and technetium 99m sulfur colloid," *J. Am. Coll. Surg.*, vol. 195, pp. 467–475, Oct 2002.

[26] T. Wendler, K. Herrmann, A. Schnelzer, T. Lasser, J. Traub, O. Kutter, A. Ehlerding, K. Scheidhauer, T. Schuster, M. Kiechle, M. Schwaiger, N. Navab, S. I. Ziegler, and A. K. Buck, "First demonstration of 3-D lymphatic mapping in breast cancer using freehand SPECT," *Eur J Nucl Med*, vol. 37, no. 8, pp. 1452–1461, 2010.

[27] J. Vogel, T. Reichl, J. Gardiazabal, N. Navab, and T. Lasser, "Optimization of acquisition geometry for intra-operative tomographic imaging," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, vol. 7512, pp. 42–49, Springer, 2012.

[28] J. Vogel, T. Lasser, J. Gardiazabal, and N. Navab, "Trajectory optimization for intra-operative nuclear tomographic imaging," *Medical Image Analysis*, vol. 17, no. 7, pp. 723–731, 2013.

[29] M. J. Gooding, S. H. Kennedy, and J. A. Noble, "Temporal calibration of freehand three-dimensional ultrasound using image alignment," *Ultrasound Med Biol*, vol. 31, pp. 919–927, Jul 2005.

[30] T. Sielhorst, M. Feuerstein, and N. Navab, "Advanced medical displays: A literature review of augmented reality," *IEEE/OSA Journal of Display Technology; Special Issue on Medical Displays*, vol. 4, pp. 451–467, December 2008.

[31] N. Navab, T. Blum, L. Wang, A. Okur, and T. Wendler, "First deployments of augmented reality in operating rooms," *Computer*, vol. 99, pp. 48–55, 2012.

[32] A. Okur, R. Voigt, R. Stauder, and N. Navab, "Investigation of performance log files of freehand SPECT acquisitions for usage characteristics and surgical phase determination," in *The Hamlyn Symposium on Medical Robotics: Proceedings* (G.-Z. Yang and A. Darzi, eds.), 2014.

[33] T. Lasser, *Tomographic Reconstruction Methods for Optical and Intra-operative Functional Imaging.* PhD thesis, Technische Universität München, Fakultät für Informatik I16, June 2011.

[34] M. N. Wernick and J. N. Aarsvold, *Emission Tomography: The fundamentals of PET and SPECT.* 2004.

[35] H. H. Barrett and W. Swindell, eds., *Radiological Imaging: The Theory of Image Formation, Detection, and Processing*, ch. 3, p. 98. Academic Press, Inc., 1981.

[36] Shepp, L. A. and Vardi, Y., "Maximum Likelihood Reconstruction for Emission Tomography," *Transactions on Medical Imaging*, vol. MI-1, no. 2, pp. 113–122, 1982.

[37] Y. Vardi, L. A. Shepp, and L. Kaufman, "A statistical model for positron emission tomography," *Journal of the American Statistical Association*, vol. 80, no. 389, pp. 8–20, 1985.

[38] L. Parra and H. H. Barrett, "List-Mode Likelihood: EM Algorithm and Image Quality Estimation Demonstrated on 2-D PET," *IEEE Transactions on Medical Imaging*, vol. 7, pp. 228–235, April 1998.

[39] C. Byrne, "Likelihood Maximization for List-Mode Emission Tomographic Image Reconstruction," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 1084–1092, October 2001.

[40] A. Hartl, D. Shakir, T. Lasser, S. I. Ziegler, and N. Navab, "Detection models for freehand spect reconstruction," *Physics in Medicine and Biology*, jan 2015.

[41] M. Zvolsky, A. Cserkaszky, G. Cucciati, E. Garutti, P. Mitsakis, and B. Frisch, "Montecarlo simulation and image reconstruction for an endoscopic tofpet detector," in *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference Record*, 2006.

[42] A. Hartl, *Computational modeling of detection physics for 3D interoperative imaging with navigated probes.* PhD thesis, Technische Universität München, Fakultät für Informatik I16, November 2015.

[43] S. H. Walrand, L. R. van Elmbt, and S. Pauwels, "A non-negative fast multiplicative algorithm in 3D scatter-compensated SPET reconstruction," *Eur J Nucl Med*, vol. 23, pp. 1521–1526, Nov 1996.

[44] B. F. Hutton, H. M. Hudson, and F. J. Beekman, "A clinical perspective of accelerated statistical reconstruction," *Eur J Nucl Med*, vol. 24, pp. 797–808, Jul 1997.

[45] T. Wendler, *3D intraoperative functional imaging with navigated probes.* PhD thesis, Technische Universität München, Fakultät für Informatik I16, October 2010.

[46] F. Xu and K. Mueller, "Accelerating popular tomographic reconstruction algorithms on commodity pc graphics hardware," *IEEE Transactions on Nuclear Science*, vol. 52, pp. 654–663, June 2005.

[47] G. Pratx, G. Chinn, F. Habte, P. D. Olcott, and C. S. Levin, "Fully 3d list-mode osem accelerated with graphics processing units," in *2006 IEEE Nuclear Science Symposium and Medical Imaging Conference Record*, pp. 2196–2202, 2006.

[48] "NASA Mars Rover." http://photojournal.jpl.nasa.gov/catalog/PIA19920.

[49] "Welding robot at the manufacturing plant of Kinshofer GmbH." https://commons.wikimedia.org/wiki/File:KinshoferGmbH_welding_computer.JPG.

[50] "Intuitive Surgical, Inc." www.intuitivesurgical.com.

[51] R. J. Romero, R. Kosanovic, J. R. Rabaza, R. Seetharamaiah, C. Donkor, M. Gallas, and A. M. Gonzalez, "Robotic sleeve gastrectomy: experience of 134 cases and comparison with a systematic review of the laparoscopic approach," *Obes Surg*, vol. 23, pp. 1743–1752, Nov 2013.

[52] L. S. Canale, S. Mick, T. Mihaljevic, R. Nair, and J. Bonatti, "Robotically assisted totally endoscopic coronary artery bypass surgery," *J Thorac Dis*, vol. 5 Suppl 6, pp. S641–649, Nov 2013.

[53] A. Vaccarella, M. D. Comparetti, A. Enquobahrie, G. Ferrigno, and E. De Momi, "Sensors management in robotic neurosurgery: the ROBOCAST project," *Conf Proc IEEE Eng Med Biol Soc*, vol. 2011, pp. 2119–2122, 2011.

[54] W. L. Bargar, A. Bauer, and M. Borner, "Primary and revision total hip replacement using the Robodoc system," *Clin. Orthop. Relat. Res.*, pp. 82–91, Sep 1998.

[55] A. P. Schulz, K. Seide, C. Queitsch, A. von Haugwitz, J. Meiners, B. Kienast, M. Tarabolsi, M. Kammal, and C. Jurgens, "Results of total hip replacement using the Robodoc surgical assistant system: clinical outcome and evaluation of complications for 97 procedures," *Int J Med Robot*, vol. 3, pp. 301–306, Dec 2007.

[56] W. P. Liu, Y. Otake, M. Azizian, O. J. Wagner, J. M. Sorger, M. Armand, and R. H. Taylor, "2D-3D radiograph to cone-beam computed tomography (CBCT) registration for C-arm image-guided robotic surgery," *Int J Comput Assist Radiol Surg*, vol. 10, pp. 1239–1252, Aug 2015.

[57] A. Kaladji, A. Dumenil, G. Mahe, M. Castro, A. Cardon, A. Lucas, and P. Haigron, "Safety and accuracy of endovascular aneurysm repair without pre-operative and intra-operative contrast agent," *Eur J Vasc Endovasc Surg*, vol. 49, pp. 255–261, Mar 2015.

[58] "Siemens artis zeego imaging suite,"

[59] R. Feurer, C. Hennersperger, J. B. Runyan, C. L. Seifert, J. Pongratz, M. Wilhelm, J. Pelisek, N. Navab, E. Bartels, and H. Poppert, "Reliability of a freehand three-dimensional ultrasonic device allowing anatomical orientation at a glance: Study protocol for 3d measurements with curefab cs," vol. 2, 2012.

[60] C. Graumann, B. Fuerst, C. Hennersperger, F. Bork, and N. Navab, "Robotic ultrasound trajectory planning for volume of interest coverage," May 2016.

[61] J. Craig, *Introduction to Robotics: Mechanics and Control.* Addison-Wesley series in electrical and computer engineering: control engineering, Pearson/Prentice Hall, 2005.

[62] J. Denavit and R. S. Hartenberg, "A kinematic notation for lower-pair mechanisms based on matrices.," *Trans. of the ASME. Journal of Applied Mechanics*, vol. 22, pp. 215–221, 1955.

[63] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision.* Cambridge University Press, ISBN: 0521540518, second ed., 2004.

[64] R. Diankov, *Automated Construction of Robotic Manipulation Programs.* PhD thesis, Carnegie Mellon University, Robotics Institute, August 2010.

[65] J. H. Reif, "Complexity of the mover's problem and generalizations," in *Proceedings of the 20th Annual Symposium on Foundations of Computer Science*, SFCS '79, (Washington, DC, USA), pp. 421–427, IEEE Computer Society, 1979.

[66] J. Canny, B. Donald, J. Reif, and P. Xavier, "On the complexity of kinodynamic planning," in *Foundations of Computer Science, 1988., 29th Annual Symposium on*, pp. 306–316, Oct 1988.

[67] V. J. Lumelsky and A. A. Stepanov, "Path-planning strategies for a point mobile automaton moving amidst unknown obstacles of arbitrary shape," *Algorithmica*, vol. 2, no. 1, pp. 403–430, 1987.

[68] H. Choset, K. M. Lynch, S. Hutchinson, G. A. Kantor, W. Burgard, L. E. Kavraki, and S. Thrun, *Principles of Robot Motion: Theory, Algorithms, and Implementations.* Cambridge, MA: MIT Press, June 2005.

[69] K. Shin and N. McKay, "Minimum-time control of robotic manipulators with geometric path constraints," *IEEE Transactions on Automatic Control*, vol. 30, no. 6, pp. 531–541, 1985.

[70] J. Bobrow, S. Dubowsky, and J. Gibson, "Time-optimal control of robotic manipulators along specified paths," *The International Journal of Robotics Research*, vol. 4, no. 3, pp. 3–17, 1985.

[71] J. Gardiazabal, T. Reichl, A. Okur, T. Lasser, and N. Navab, "First flexible robotic intra-operative nuclear imaging for image-guided surgery," in *Springer LNCS 7915 (IPCAI)*, pp. 81–90, 2013.

[72] "Crystal Photonics GmbH." http://www.crystal-photonics.com.

[73] "Northern Digital Inc.." http://www.ndigital.com/medical/products.

[74] R. Tsai and R. Lenz, "Real time versatile robotics hand/eye calibration using 3D machine vision," in *Proceedings of the IEEE International Conference on Robotics and Automation*, vol. 1, pp. 554–561, 1988.

[75] R. Tsai and R. Lenz, "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration," *IEEE Transactions on Robotics and Automation*, vol. 5, pp. 345–358, June 1989.

[76] A. Hartl, S. Ziegler, and N. Navab, "Models of detection physics for nuclear probes in freehand SPECT reconstruction," in *Proceedings of IEEE Medical Imaging Conference (IEEE MIC)*, (Knoxville, Tennessee, USA), October 2010.

[77] P. Matthies, J. Gardiazabal, A. Okur, J. Vogel, M. Friebe, T. Lasser, and N. Navab, "Mini gamma cameras for intra-operative nuclear tomographic reconstruction," *Medical Image Analisys*, 2014.

[78] T. Kuehn, A. Bembenek, T. Decker, D. L. Munz, M.-L. Sautter-Bihl, M. Untch, and D. Wallwiener, "A concept for the clinical implementation of sentinel lymph node biopsy in patients with breast carcinoma with special regard to quality assurance," *Cancer*, vol. 103, pp. 451–461, February 2005.

[79] Matthies, Philipp and Sharma, Kanishka and Okur, Aslı and Gardiazabal, José and Vogel, Jakob and Lasser, Tobias and Navab, Nassir, "First use of mini gamma cameras for intra-operative robotic SPECT reconstruction," in *Medical Image Computing and Computer-Assisted InterventionMICCAI 2013*, no. 8149 in Lecture Notes in Computer Science, pp. 163–170, Springer Berlin Heidelberg, January 2013.

[80] L. Papp, N. Zsoter, C. Loh, B. Ole, B. Egeler, I. Garai, and U. Luetzen, "Automated lymph node detection and classification on breast and prostate cancer SPECT-CT images," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC*, pp. 3431–3434, 2011.

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

[82] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[83] J. Gardiazabal, P. Matthies, J. Vogel, N. Frisch, B. Navab, S. Ziegler, and L. T., "Flexible mini gamma camera reconstructions of extended sources using step and shoot and list mode," *Medical physics*, 2016.

[84] "Universal Robots A/S." http://www.universal-robots.com.

[85] C. P. Robert and G. Casella, eds., *Monte Carlo Statistical Methods*, pp. 47–51. Springer Science + Business Media, 2004.

[86] P. Matthies, B. Frisch, J. Vogel, T. Lasser, M. Friebe, and N. Navab, "Inside-Out Tracking for Flexible Hand-held Nuclear Tomographic Imaging," in *IEEE Nuclear Science Symposium*, 2015.

[87] T. Maurer, G. Weirich, M. Schottelius, M. Weineisen, B. Frisch, A. Okur, H. Kübler, M. Thalgott, N. Navab, M. Schwaiger, *et al.*, "Prostate-specific membrane antigen–radioguided surgery for metastatic lymph nodes in prostate cancer," *European urology*, vol. 68, no. 3, pp. 530–534, 2015.

[88] M. Esposito, B. Busam, C. Hennersperger, J. Rackerseder, A. Lu, N. Navab, and B. Frisch, "Cooperative robotic gamma imaging: enhancing US-guided needle biopsy," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 611–618, Springer, 2015.

[89] S. Virga, O. Zettinig, M. Esposito, K. Pfister, B. Frisch, T. Neff, N. Navab, and C. Hennersperger, "Automatic force-compliant robotic ultrasound screening of abdominal aortic aneurysms," in *IEEE/RSJ 2016 International Conference on Intelligent Robots and Systems (IROS 2016)*, 2016.

[90] J. Gardiazabal, M. Esposito, P. Matthies, A. Okur, J. Vogel, S. Kraft, B. Frisch, T. Lasser, and N. Navab, "Towards personalized interventional spect-ct imaging," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 504–511, Springer, 2014.

[91] B. Fuerst, J. Sprung, F. Pinto, B. Frisch, T. Wendler, H. Simon, L. Mengus, N. S. van den Berg, H. G. van der Poel, F. W. van Leeuwen, *et al.*, "First robotic SPECT for minimally invasive sentinel lymph node mapping," *IEEE Transactions on Medical Imaging*, vol. 35, no. 3, pp. 830–838, 2016.

[92] B. Sangro, M. Iñarrairaegui, and J. I. Bilbao, "Radioembolization for hepatocellular carcinoma," *Journal of Hepatology*, vol. 56, pp. 464–473, February 2012.

[93] B. Peynircioğlu, B. Çil, F. Bozkurt, E. Aydemir, O. Uğur, and F. Balkancı, "Radioembolization for the treatment of unresectable liver cancer: initial experience at a single center," *Diagn Interv Radiol*, vol. 16, no. 1, pp. 70–78, 2010.

[94] M. G. E. H. Lam, M. L. Goris, A. H. Iagaru, E. S. Mittra, J. D. Louie, and D. Y. Sze, "Prognostic Utility of 90Y Radioembolization Dosimetry Based on Fusion 99mTc-Macroaggregated Albumin-99mTc-Sulfur Colloid SPECT," *Journal of Nuclear Medicine*, vol. 54, pp. 2055–2061, December 2013.

[95] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 376–380, April 1991.

[96] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *Journal of the Optical Society of America A*, vol. 1, no. 6, pp. 612–619, 1984.

[97] A. Ganguly, A. Fieselmann, M. Marks, J. Rosenberg, J. Boese, Y. Deuerling-Zheng, M. Straka, G. Zaharchuck, R. Bammer, and R. Fahrig, "Cerebral CT Perfusion Using an Interventional C-Arm Imaging System: Cerebral Blood Flow Measurements," *Am J Neuroradiol*, vol. 32, pp. 1525–1531, 2011.

[98] J. Stayman and J. Siewerdsen, "Task-based trajectories in iteratively reconstructed interventional cone-beam CT," in *Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine (Fully3D)*, (Lake Tahoe, CA), pp. 257–260, 2013.

[99] J. Bowsher, S. Yan, J. Roper, W. Giles, and F.-F. Yin, "Onboard functional and molecular imaging: a design investigation for robotic multipinhole SPECT," *Medical physics*, vol. 41, p. 010701, January 2014. PMID: 24387490 PMCID: PMC3888458.

[100] H. Ahmadzadehfar, M. Muckle, A. Sabet, K. Wilhelm, C. Kuhl, K. Biermann, T. Haslerud, H.-J. Biersack, and S. Ezziddin, "The significance of bremsstrahlung SPECT/CT after yttrium-90 radioembolization treatment in the prediction of extrahepatic side effects," *European journal of nuclear medicine and molecular imaging*, October 2011. PMID: 21975832.

[101] R. H. Huesman, G. J. Klein, W. W. Moses, J. Qi, B. W. Reutter, and P. R. G. Virador, "List-mode maximum-likelihood reconstruction applied to positron emission mammography (PEM) with irregular sampling," *IEEE Transactions on Medical Imaging*, vol. 19, pp. 532–537, May 2000.

List of Figures

List of Tables