universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Customized source-detector trajectory optimization for CBCT using Simulated Annealing algorithm "

verfasst von / submitted by

Hannah Jungreuthmayer BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2023 / Vienna, 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 876
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Master's degree Physics
Betreut von / Supervisor:	Univ.-Prof. Dr. Christoph Dellago

Note of thanks

I want to sincerely thank my external supervisors Prof. Sepideh Hatamikia from the Danube Private University (DPU) and Prof. Wolfgang Birkfellner from the Medical University of Vienna for taking me in as their Master's student. They provided me with a topic for my thesis and a workplace at the Medical University of Vienna. I had great help from them in writing this thesis and I was even given the opportunity to present my work at two conferences. They also allowed me to attend a summer school on biomedical imaging, from which I was able to gain very valuable knowledge. Both of them were always really nice and helpful and without them this project would not have been possible.

Furthermore I am grateful to my colleagues, especially S M Ragib Shahriar Islam. He helped me to get started with the thesis' topic, which was rather new to me in the beginning.

I also want to express my gratitude to the Austrian Center for Medical Innovation and Technology (ACMIT) for financing my project.

Last but not least, I am very thankful to my family and friends for providing me with the best support system anyone could have.

Abstract

Cone Beam Computed Tomography (CBCT) has become a routine clinical imaging modality in interventional radiology. CBCT images are usually created by rotating an x-ray source and a detector in a circular trajectory around the patient. Thereby projections are taken from which an image can be reconstructed. However, this standard approach has some limitations, including limited field of view (FOV), the need of a high projection number (i. e. a high radiation dose for the patient), and that it is not patient specific. Furthermore the procedure needs a considerable amount of space, which is often not available in the operating room.

Extended Field of View (FOV) CBCT is of great clinical importance for many medical applications, especially for cases where the Volume of Interest (VOI) is outside the standard FOV. In this study, FOV extension is investigated by optimizing customized source-detector CBCT trajectories using the Simulated Annealing (SA) algorithm, a heuristic search optimization algorithm. The SA algorithm explores different elliptical trajectories within a given parameter space, attempting to optimize image quality in a given VOI. Kinematic constraints (due to collisions of the imager with the patient or other medical devices) are taken into account when designing the trajectories. A digital XCAT body phantom and the geometry of Philips Allura Xper C-arm were considered for simulation. The TIGRE toolbox and the Universal Quality Index (UQI) were used for image reconstruction and image quality assessment, respectively. The results showed that the proposed trajectories could achieve a UQI at the respective VOIs (UQI: 0.9148, 0.9681, 0.9632 and 0.9273) which had a considerably better image quality compared with the circular trajectories (UQI: 0.5960, 0.4892, 0.4798 and 0.7179). In addition, for 3 out of 4 cases a notable FOV extension (26.17%, 22.59%, 0.92% and 44.57%) was achieved.

Furthermore an attempt was made to optimize the recorded projections for reconstructing a specific VOI when reducing the projection number. Reducing the projection number in general results in an image quality decrease, but selecting projections with the most information about a specific VOI could still lead to a diagnostic image quality. The utilized optimization algorithm was again Simulated Annealing and regular spaced projections served as comparison to the optimized projections. The projection selection optimization shows visible improvement, although the achieved UQI enhancements (from 0.8050 to 0.8992 and from 0.8805 to 0.8882) are small, compared to the ones achieved by trajectory optimization. Finally the proposed projection selection optimization was applied to real projection data. Because the data was not ideal for this experiment, the visual improvement and the UQI enhancement from regular spaced projections (from 0.9260 to 0.9284) is rather small.

The experimental results have shown that our proposed customized trajectories can lead to an extended FOV and enable improved visualization of anatomical structures in extreme positions while taking into account the available kinematic constraints. Furthermore, it was shown that SA can be utilized to optimize recorded projections when reducing their amount. This study offers a new approach to improve the diagnostic capabilities of CBCT imaging, thus providing valuable insight into improving patient care.

Zusammenfassung

Cone Beam Computed Tomography (CBCT) hat sich zu einer Routinebildgebungsmethode in der interventionellen Radiologie entwickelt. CBCT-Bilder werden in der Regel durch kreisförmige Rotation einer Röntgenquelle und eines Detektors um den Patienten erstellt. Bei der Rotation werden Projektionen aufgenommen, aus denen danach ein Bild rekonstruiert werden kann. Allerdings hat diese Standardtechnik einige Einschränkungen, darunter ein begrenztes Sichtfeld, die Notwendigkeit einer hohen Anzahl von Projektionen (d. h. eine hohe Strahlendosis für den Patienten) und dass die Methode nicht patientenspezifisch ist. Darüber hinaus erfordert das Verfahren eine gewisse Menge Platz, die in einem Operationssaal oft nicht verfügbar ist.

Ein erweitertes Sichtfeld ist für viele medizinische Anwendungen des CBCTs von großer klinischer Bedeutung, insbesondere für Fälle, in denen das Interessengebiet außerhalb des standardmäßigen Sichtfeldes liegt. In dieser Studie wurde Sichtfelderweiterung durch die Optimierung von Quelle-Detektor-Trajektorien unter Verwendung des Simulated Annealing (SA) Algorithmus untersucht, einem heuristischen Optimierungsalgorithmus. Der SA-Algorithmus erforscht verschiedene elliptische Trajektorien innerhalb eines gegebenen Parameterbereichs und versucht, die Bildqualität in einem bestimmten Interessengebiet zu optimieren. Kinematische Einschränkungen (aufgrund von Kollisionen des Bildgebungsgeräts mit dem Patienten oder anderen medizinischen Geräten) wurden bei der Optimierung der Trajektorien berücksichtigt. Für die Simulationen wurden ein digitales XCAT-Phantom und die Bildgebungsgeometrie des Philips Allura Xper C-Arms verwendet. Die TIGRE Toolbox und der Universal Quality Index (UQI) wurden für die Bildrekonstruktion bzw. die Bewertung der Bildqualität benutzt. Die Ergebnisse zeigen, dass die vorgeschlagenen Trajektorien einen UQI (0.9148, 0.9681, 0.9632 und 0.9273) in den jeweiligen Interessengebieten erreichen konnten, der im Vergleich zu den kreisförmigen Trajektorien (UQI: 0.5960, 0.4892, 0.4798 und 0.7179) eine bedeutend bessere Bildqualität aufweist. Zusätzlich wurde für 3 von 4 Fällen auch eine bemerkenswerte FOV-Erweiterung erreicht (26.17%, 22.59%, 0.92% und 44.57%).

Darüber hinaus wurde versucht die für die Rekonstruktion eines spezifischen Interessengebiet aufgenommenen Projektionen zu optimieren, wenn deren Anzahl reduziert werden soll. Generell führt eine Reduktion der Projektionszahl zu einer Verringerung der Bildqualität. Werden aber jene Projektionen mit dem meisten Informationsgehalt über ein bestimmtes Interessengebiet ausgewählt, kann dennoch eine diagnostische Bildqualität erreicht werden. Der verwendete Optimierungsalgorithmus war erneut Simulated Annealing, wobei Projektionen mit regelmäßigen Abständen als Vergleich zu den optimierten Projektionen dienten. Die Optimierung der Projektionen zeigt eine sichtbare Verbesserung, obwohl die erreichten UQI-Verbesserungen (von 0.8050 auf 0.8992 und von 0.8805 auf 0.8882) im Vergleich zu denen der Trajektorienoptimierung gering sind. Schließlich wurde die vorgeschlagene Optimierung der Projektion auf echte Projektionsdaten angewandt. Da die verwendeten Daten für dieses Experiment nicht ideal waren, ist die visuelle Verbesserung und die UQI Erhöhung (von 0.9260 auf 0.9284) gegenüber regelmäßig verteilten Projektionen eher gering.

Die experimentellen Ergebnisse haben gezeigt, dass unsere vorgeschlagenen Trajektorien zu einem erweiterten Sichtfeld führen können. Weiters wurde eine verbesserte Visualisierung anatomischer Strukturen in peripheren Positionen demonstriert, wobei gewisse kinematische Einschränkungen berücksichtigt wurden. Darüber hinaus wurde gezeigt, dass SA dazu genutzt werden kann, Projektionen zu optimieren, wenn ihre Anzahl reduziert werden soll. Diese Studie bietet eine neue Vorgehensweise zur Verbesserung der diagnostischen Möglichkeiten der CBCT-Bildgebung und liefert somit wertvolle Einblicke in die Verbesserung der Patientenversorgung.

Contents

1 X-Ray radiation

1.1 Historical aspects

In 1895 the physicist Wilhelm C. Röntgen conducted experiments with cathode rays. A potential difference was applied across a glass "discharge tube", which was partially evacuated. He observed that light was emitted from scintillators some distance away and concluded that the cause had to be some unknown radiation produced by his experiments. He therefore named them x-rays. Röntgen quickly discovered that they had the ability to penetrate various materials and that they could be recorded by photographic plates. One image he recorded was the radiograph of a hand, which is shown in Fig. 1.

Figure 1: Radiograph of a hand taken by Röntgen (1895). The hand probably belonged to his wife. (Adapted from [1], Creator: W.C. Röntgen)

Several countries explored x-rays as medical tools within a month of their discovery. Shortly after, Poincaré demonstrated that x-rays were produced when cathode rays hit the glass discharge tube. In 1897 Thompson found that the cathode rays were actually negatively charged particles (electrons) [1].

There are also some Austrian scientists who made important contributions to the usage of x-rays in medicine. In 1896 Röntgen sent recorded radiographs to about a dozen of his colleagues. One of them was Franz-Serafin Exner, an Austrian physicist and former student of Röntgen. Exner was a convivial man and helped in spreading the word about Röntgens discovery [2]. Johann Puluj was an Austrian-Hungarian physicist and electrical engineer and conducted experiments with cathode rays even before Röntgens discovery. After x-rays were discovered, Puluj focused on their application to the clinical environment and therefore was a pioneer in medical imaging [3]. Eduard Haschek and Otto Theodor Lindenthal from Vienna opacified the vessels of an amputated hand in 1897. They developed a vascular contrast agent, consisting of a mixture of cinnabar, chalk and vaseline and thereby invented angiography [4].

1.2 Production

Producing medical images with x-rays requires a source that generates enough x-rays in a short period of time. Furthermore it is important that x-rays are generated in a reproducible fashion and that the standards for safety and economy are met. There are several options for generating x-rays including nuclear reactions, radioactive isotopes and particle accelerators. However, only certain particle accelerators, named x-ray tubes, meet the criteria mentioned before [1]. The main components of a modern x-ray tube as devised by Coolidge are shown in Fig. 2.

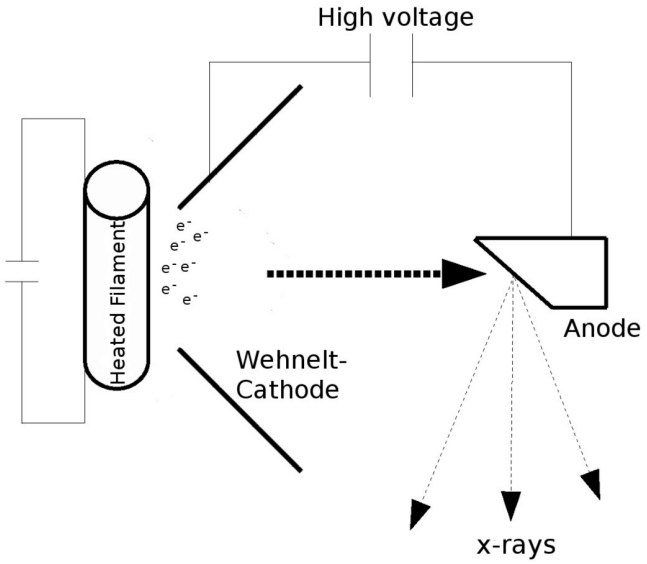


Figure 2: Main components of an x-ray tube. (Adapted from [5], Creator: W. Birkfellner)

Electrons are released by a heated filament (cathode) and accelerated onto a target (anode) by a high voltage U. The term 'tube current' therefore describes the stream of electrons. When the electrons interact with the target material, x-rays are released. By means of collimation a useful x-ray beam can be produced. Inside the glass envelope a vacuum is maintained in order to prevent interaction of electrons with gas molecules and to prevent the filament from burning [1].

The achieved kinetic energy of the electrons by the end of the acceleration path equals $E_{kin} = U \cdot e$, where e is the elementary charge of the electron. The emitted x-ray radiation therefore has a maximal energy of

$$E_{max} = e \cdot U = h \cdot \frac{c}{\lambda}, \tag{1}$$

with h being Planck's constant and c being the speed of light. Eq. 1 is called Duane-Hunt law. The wavelength λ thus equals

$$\lambda = \frac{c \cdot h}{e \cdot U}. \tag{2}$$

Plugging in typical voltage U values used in diagnostics (30 to 150 kV) results in an x-ray wavelength of 8 to 41 pm, which is 4 orders of magnitude smaller than visible light [6].

In the diagnostic energy range there are mostly three types of electron interactions with the anode material. The first one is scattering without energy transfer (Thomson scattering). How-

ever, this does not lead to the formation of x-rays. The two other interactions, which contribute to x-ray formation are described in the following chapters [6].

1.2.1 Bremsstrahlung

An electron passing a nucleus can be deflected with reduced speed. That means energy is released during the encounter, in this case as electromagnetic radiation (x-rays). The interaction is therefore called inelastic [1]. The energy of the produced photon is equal to the difference between the energy of the incoming electron E_e and the energy of the deflected electron E'_e. The latter depends on the distance from the positively charged nucleus. The closer the electron gets, the higher the x-ray energy will be, therefore the x-ray spectrum is continuous. Fig. 3 shows a schematic of the process that generates this radiation, known as bremsstrahlung [6].

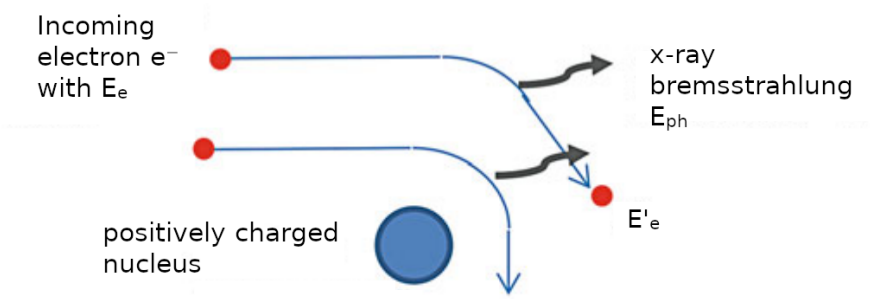


Figure 3: Depending on the distance of the electron from the nucleus, its deflection varies. The resultant x-ray radiation has an energy of $E_{ph} = E_e - E'_e$. The closer the electron comes to the nucleus, the higher the x-ray energy E_{ph}. (Adapted from [6], Creator: R. Ringler. The descriptions were translated to English.)

X-ray photons can have an energy up to the total kinetic energy of the electron [1]. In general, more low energy photons are generated by this process, but these photons are also more likely to be attenuated (or filtered) by the anode and tube material. In contrast to low energy photons, high energy photons are more likely to pass through a material without interaction [1, 6]. This results in a spectrum similar to the one in Fig. 4.

A higher voltage U results in a higher maximal energy of the photons and therefore influences the spectrum. Another factor influencing the spectrum is the heating current (see Fig. 2). A higher heating current results in a higher x-ray intensity, caused by a higher number of electrons being accelerated towards the anode [6]. The likelihood of producing bremsstrahlung depends on Z^2 (atomic number squared) of the absorbing material. In simple terms this means that the bigger the target, the likelier is an interaction. The relative shape of the bremsstrahlung's spectrum (Fig. 4) does not vary with Z [1].

If the electrons have a low energy, photons are radiated mainly perpendicular to the particle motion. For higher electron energies the radiation angle narrows [1].

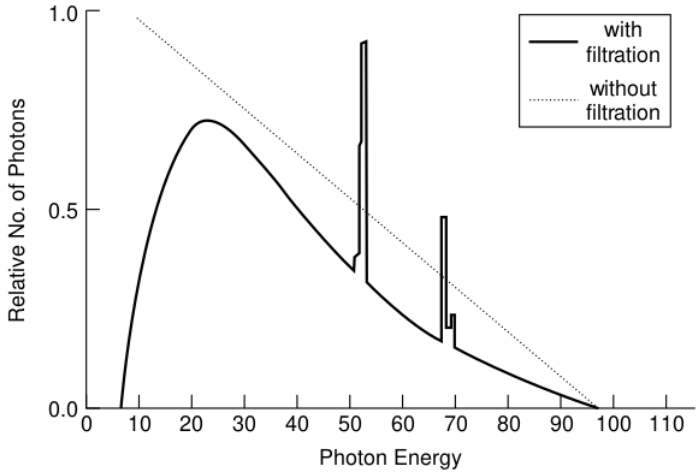


Figure 4: Typical shape of a bremsstrahlung spectrum, including characteristic x-rays (see Ch. 1.2.2). The dotted line shows the theoretical spectrum without the filtering effect of anode and tube material. (Adapted from [1], Creators: W. R. Hendee and E. R. Ritenour)

1.2.2 Characteristic X-Rays

If an incident electron with sufficient energy hits an electron belonging to a nucleus of the anode material, the hit electron is ejected from its respective shell. The necessary energy for this process depends on the material and the shell, where the electron is located. As a result, the atom is ionized and for a short period of time there is an empty space in this shell. The empty space is usually immediately occupied by an electron of a higher energy shell [6].

There are two different ways on how to compensate for the difference in binding energy between the two shells. On way is to release a photon carrying the respective energy (=characteristic x-rays). This process is schematically illustrated for K- and L-shell in Fig. 5.

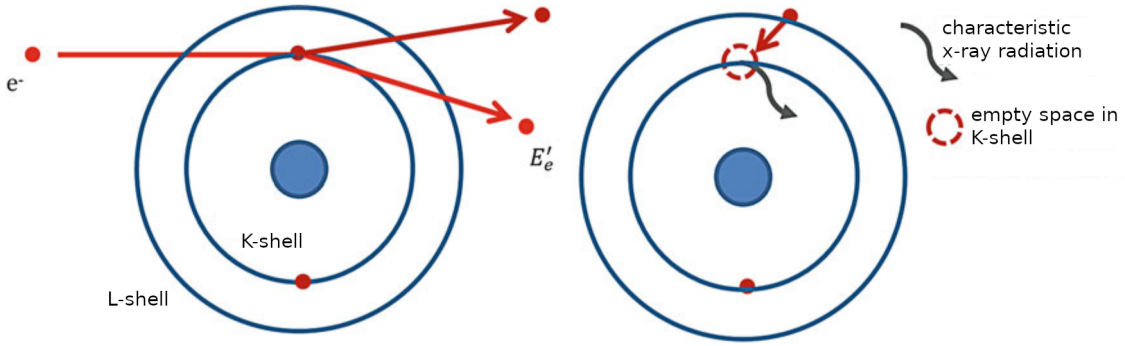


Figure 5: An electron from the K-shell is knocked out of the atom by an incident electron. The empty space in the shell is filled by an electron from a higher energy shell. The difference in binding energy is carried away by a photon. (Adapted from [6], Creator: R. Ringler. The descriptions were translated to English.)

After the electron transition, there is now a vacancy in the higher energy shell, which is in turn can be filled by an electron from another (even higher) shell and so on. Therefore there is a cascade of electron transitions, with a range of characteristic photon energies. Since the

binding energy difference between neighbouring shells decreases further away from the nucleus, transitions to shells beyond the M shell do not release enough energy to be considered x-rays [1, 6].

The second way to release the energy from the electron transition is by transferring it to another electron, which is ejected as a result. These electrons are called an Auger electrons and they are usually released from the same shell where the transitioned electron originated from. Therefore its kinetic energy can be calculated by the binding energy of the inner shell minus twice the binding energy of the outer shell. There is no way of knowing which of these two processes will follow an electron transition, but the probability of characteristic emission (called fluorescence yield) in general increases with atomic number [1].

A material often utilized in anodes for medical purposes is tungsten, because of its high atomic number and high melting point. In tungsten there are four high probability transitions that lie within the energy spectrum used in diagnostic radiology. Two transitions from L- to K-shell and two transitions from M- to K-shell. The electron transitions in tungsten and the resulting discrete energy spectrum (overlayed with the continuous bremsstrahlung spectrum) are shown in Fig. 6 and Fig. 7. The characteristic radiation in tungsten only makes up about 10% of the total x-ray

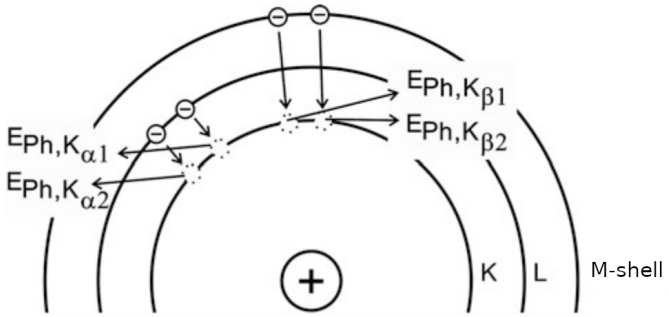


Figure 6: Main electron transitions for producing characteristic x-rays with energies E_{Ph} in tungsten. The different transitions are named $K_{\alpha1}$, $K_{\alpha2}$, $K_{\beta1}$ and K_{β}. (Adapted from [6], Creator: R. Ringler. The descriptions were translated to English.)

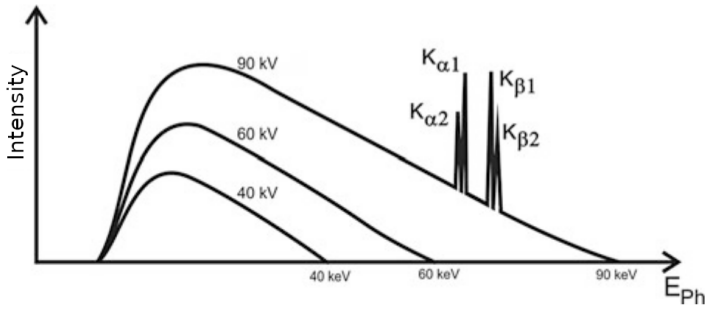


Figure 7: Characteristic x-ray spectrum resulting from the electron transitions shown in Fig. 6. The continuous bremsstrahlung spectrum is shown as well. (Adapted from [6], Creator: R. Ringler. The descriptions were translated to English.)

intensity and therefore has minor relevance compared to bremsstrahlung. For molybdenum, which

is used as anode material in mammography, the characteristic radiation has a higher significance [6].

1.3 Interactions of X-Rays in the body

There are three possible outcomes when an x-ray interacts with matter. The photon can be absorbed (energy transfer to atoms of the target material), scattered or it can pass the material without interaction. Absorption and scattering are commonly together referred to as attenuation. For a monoenergetic narrow x-ray beam, that does not contain scattered photons, the number of photons I penetrating a material of thickness ds is described by the Beer-Lambert law

$$I = I_0 e^{-\mu \cdot ds}, \tag{3}$$

where I_0 is the number of photons before the medium and μ is the linear attenuation coefficient. μ varies with x-ray energy, effective atomic number (weighted average of atomic numbers in a composite material) of the absorber and the density of the absorber [1]. The main physical processes that contribute to attenuation are photoelectric absorption and Compton scattering. In general the photoelectric effect dominates at low photon energies and the Compton effect is more important for higher energies [7]. Coherent scattering, pair production and photodisintegration are negligible in diagnostic radiology [1].

In a photoelectric interaction, the energy of the x-ray is totally transferred to an electron in the inner shell of an atom. This electron is ejected with a kinetic energy, that equals the photon energy $E_{ph} = h\nu$ minus the binding energy E_B of the electron. The process is illustrated in Fig. 8.

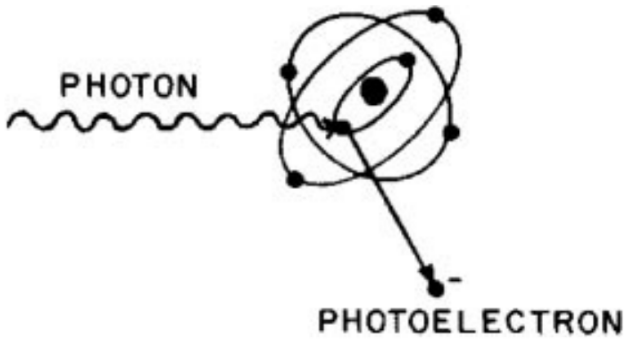


Figure 8: Photoelectric absorption of a photon. The photon completely disappears after the interaction and is replaced by a so called photoelectron. (Adapted from [1], Creators: W. R. Hendee and E. R. Ritenour)

While in a photoelectric interaction the photon is completely absorbed, in Compton scattering only part of the photon energy is transferred to a free or loosely bound electron in the material. The resulting kinetic energy of the electron corresponds to the energy lost by the photon, assuming that the electrons binding energy is negligible. A Compton interaction results in a wavelength change $\Delta\lambda = 0.00243(1\text{-cos }\Phi)$ nm of the photon, where Φ is its scattering angle. The resultant

photon wavelength equals $\lambda' = \lambda + \Delta\lambda$, with λ being the initial wavelength. Compton scattering is illustrated in Fig. 9 [1].

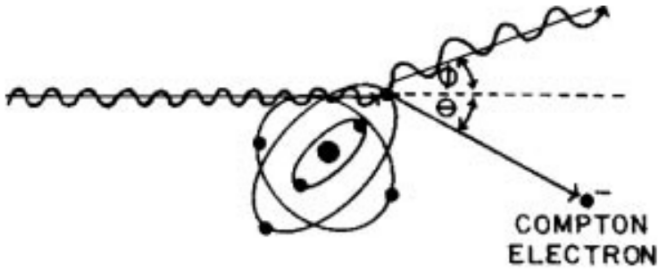


Figure 9: Illustration of Compton scattering, where the photon is scattered by an angle Φ. The electron ejection angle is Θ, with respect to the incident photon direction. (Adapted from [1], Creators: W. R. Hendee and E. R. Ritenour)

X-ray energies in medical imaging are between 30 and 150 keV. In this energy range Compton scattering is the most pronounced interaction and it depends on the electron density ρ_e as

$$\rho_e = \rho N_0 \frac{Z}{A}, \tag{4}$$

where ρ is the mass density, A is the atomic mass number, Z is the atomic number and N_0 is Avogadro's number. Since $\frac{Z}{A}$ is constant $(1/2)$ for most elements in the human body (hydrogen, oxygen, nitrogen, carbon) the x-ray image shows primarily a map of mass density. The Z dependence in the image comes from the photoelectric interaction, which is pronounced for elements with higher Z (bone, metallic implants, contrast agents) [8].

Uncharged particles, such as x-rays, are indirectly ionizing. This means that they do not deposit energy in a medium via direct Coulomb interactions, like charged particles do [7]. However, photoelectric interaction and Compton scattering result in electrons being ejected from their orbitals. These electrons are charged particles and can therefore interact with the shell electrons from atoms in the medium directly and ionize them [7, 9].

Cells that are exposed to ionizing radiation can suffer from biological damage, resulting from the physical effects of radiation on its atoms and molecules. The most critical target in the cell is the DNA molecule. Directly ionizing radiation (high linear energy transfer) can ionize or excite the target via Coulomb interaction, leading to a chain of events eventually causing biological damage. For indirectly ionizing radiation (low linear energy transfer) other molecules and atoms are lead to produce free radicals. They can be extremely reactive and damage the target through diffusion in the cell. Irradiated cells have different fates including mutation, division delay, apoptosis, reproductive failure, transformation, genomic instability or they can stay unchanged [7].

2 Computed Tomography (CT)

2.1 Introduction

Computed tomography (CT) enables physicians to learn about their patients anatomy and potentially reveals the presence of diseases or injuries in a highly detailed fashion. Before CT devices were integrated into the clinical environment, doctors had to do exploratory surgeries to diagnose patients symptoms [8]. In Fig. 10 the basic working principle of a CT device is shown.

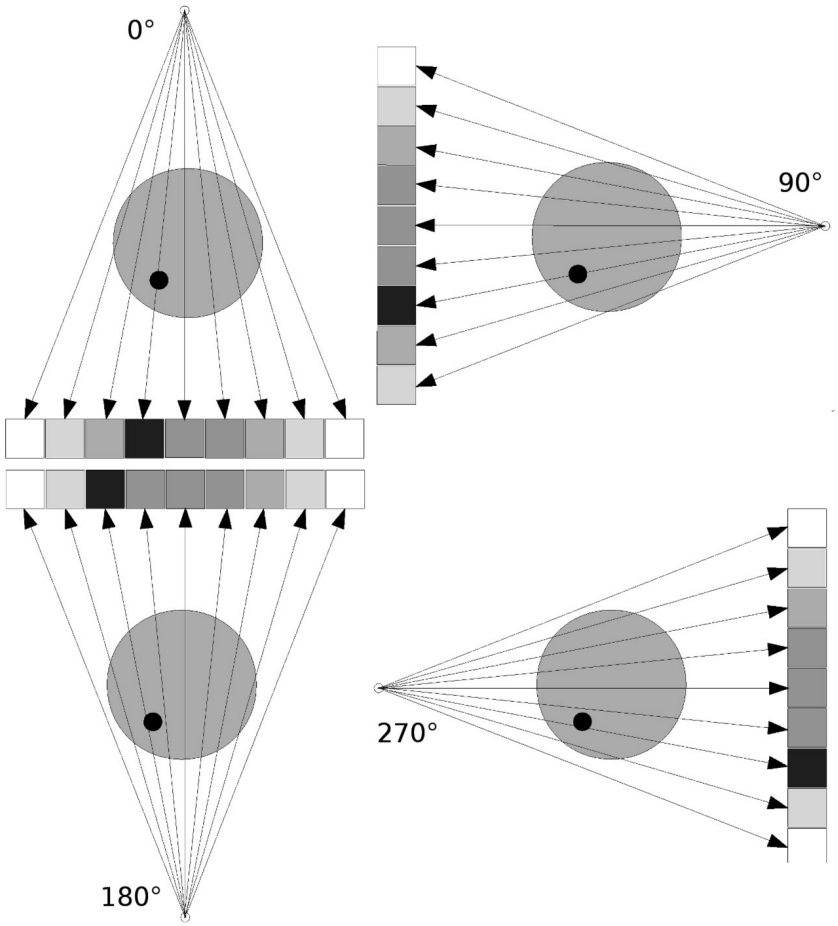


Figure 10: Working principle of a CT. Detector and x-ray source rotate around an object. Depending on the rotation angle the attenuation profile changes. (Adapted from [5], Creator: W. Birkfellner)

An x-ray tube rotates around the patient and a patient slice is measured by a collimated fan beam of x-rays originating from the tube. A detector on the opposite side measures its remaining intensity after crossing the body. This process is repeated for different angles of the source-detector system [5]. Those projections taken at different angles can be depicted in so called sinograms, as shown in Fig. 11. The name sinogram comes from the fact that spots seem to wander across the detector. Johann Radon showed back in 1917 that a two-dimensional image can be reconstructed from a set of these projections (see Ch. 2.2)[5].

The grey values in the image are proportional to the linear attenuation coefficient and their unit is given in Hounsfield units (HU). The Hounsfield unit of air is -1000 HU and the Hounsfield

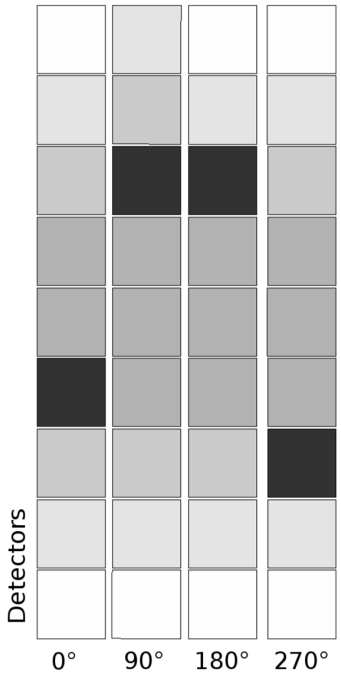


Figure 11: The intensity profiles from Fig. 10 shown as a function of detectors and rotation angle. This is called sinogram because the dark spot seems to wander across the detector. (Adapted from [5], Creator: W. Birkfellner)

unit of water is 0 HU. Many metals have more than 1000 HU [5]. Therefore, one advantage of CT images for diagnostics nowadays is that they are quantitative, making correct diagnosis easier. One example is lesions in the lung, which tend to be benign when they are calcified. This increases the CT number compared to cancerous lesions. Another example is bone mineral density measurement, which can be used as a predictor for fracture risk. The physical dimensions in CT data sets are quantitatively accurate as well, which is a result of the well defined imaging geometry. This helps in the planning and performing of certain treatments [8].

Several generations of CT machines have been invented over the years. They are summarized in Fig. 12 [5]. If a volumetric image is desired, many slices have to be reconstructed, which can be done by a so called "step and shoot" technique. Another possibility is the helical (or spiral) CT, where the detector and the x-ray tube rotate around the patient, while the table is translated continuously. This leads to a helical path of the focus. An advantage of helical CT is the short examination time. To reach a higher volume coverage in shorter periods of time, multi-detector spiral CT was developed. With these systems several slices can be acquired at once. Nonetheless detector design and development of suitable reconstruction algorithms remain a challenge here. Cone beam CTs (CBCTs) are similar, but they have two-dimensional planar detector arrays. Multi-slice detectors can also be interpreted as such, however they still scan large volumes sequentially and they have a curved detector. For CBCTs axial motion of the patient table is usually not necessary, because the x-ray beam has the shape of a 3D cone instead of a 2D fan. Furthermore cone beam detectors are flat and not curved like the ones used for multi-slice CT [5].

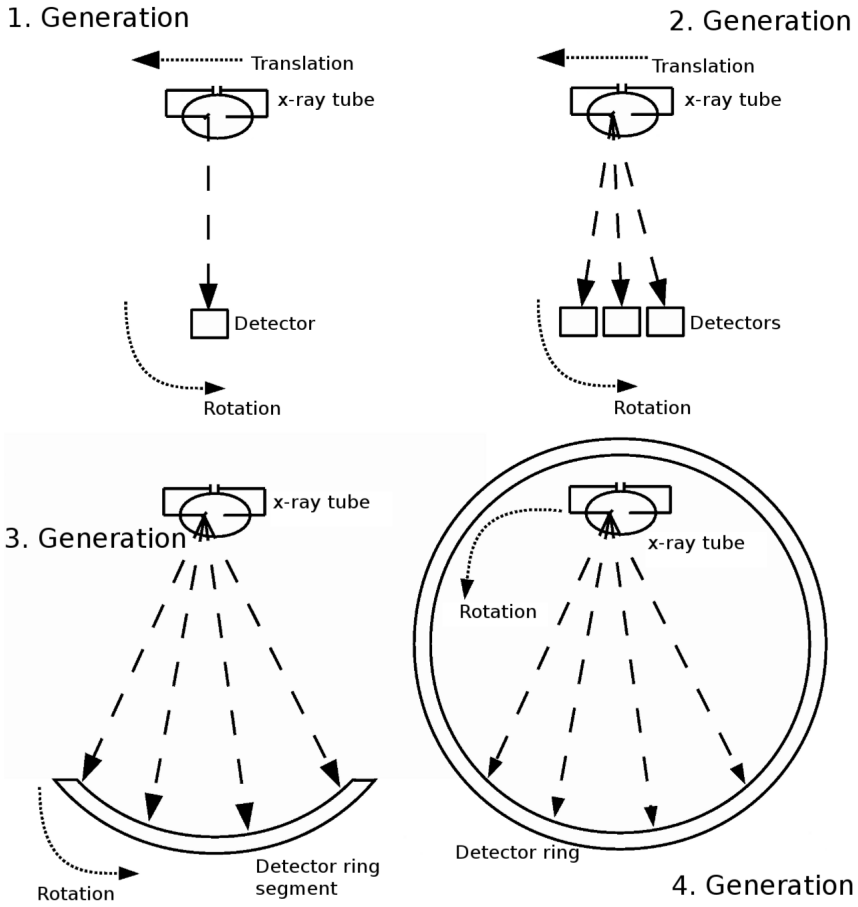


Figure 12: First generation CTs had a sharply collimated x-ray beam and only one single detector element. To record an attenuation profile the x-ray tube and the detector had to be translated and then rotated. Second generation CTs recorded the data similarly, but they had more detectors and the x-ray beam had less collimation and was thus in the form of a fan. Third generation scanners have no translational movement. Data is acquired by a rotation of the x-ray tube and the curved detector. Fourth generation systems have a stationary ring detector with a rotating x-ray tube. However, because of their lower signal to noise ratio (SNR) [10] and technological advances of the third generation scanner, they have vanished. (Adapted from [5], Creator: W. Birkfellner)

2.2 CT Reconstruction

2.2.1 Radon transform

The attenuation of an x-ray passing a volume element with attenuation coefficient μ and length ds can be described by the Beer-Lambert law in Eq. 3. I_0 and I in Eq. 3 are the intensities before and after the volume element. When the material varies along the x-ray path, different attenuation coefficients μ_i have to be considered

$$I = I_0 e^{-\mu_1 \cdot ds} e^{-\mu_2 \cdot ds} ... e^{-\mu_n \cdot ds} = I_0 e^{-\sum_i \mu_i \cdot ds} \approx I_0 e^{-\int \mu(s) ds}. \tag{5}$$

Rearranging Eq. 5 results in

$$\int \mu(s) ds \approx \sum_i \mu_i \cdot ds = log \frac{I_0}{I}. \tag{6}$$

Therefore, the integral can be computed by knowing the x-ray intensities before and after the volume it passes. This fact is used by CT reconstruction, which means deriving the density/attenuation distribution from the measured $\frac{I_0}{I}$ [5, 11].

There are different versions of CT-scanners (see Fig. 12), but to get a simple mathematical model the following description is confined to first generation ones. They produce parallel projections through the volume of interest at different angles φ. The projections are taken in one plane as can be seen in Fig. 13.

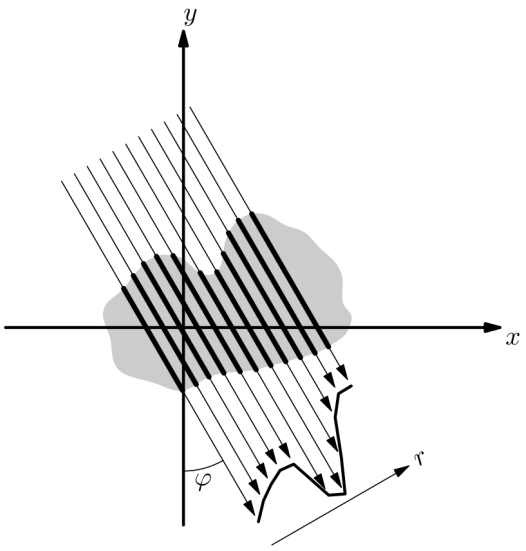


Figure 13: Principle of a first generation CT scanner. X-rays with the same angle φ but different distances r are sent through an object. Then the angle is changed. The arrow length represents the attenuation. (Adapted from [5], Creator: M. Figl)

A slice of the volume can be described as a distribution of attenuation coefficients $\mu(x, y)$. The rays sent through, have different distances r from the origin and different angles φ. To parameterize the line along which the x-ray beam is attenuated, the Hesse normal form (illustrated in Fig. 14) is used

$$\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix} = x\cos\varphi + y\sin\varphi = r. \tag{7}$$

The line of attenuation can be thought of as the rotation of a line $\binom{r}{t}$ that is parallel to the y-axis

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \cdot \begin{pmatrix} r \\ t \end{pmatrix} = \begin{pmatrix} r\cos\varphi - t\sin\varphi \\ r\sin\varphi + t\cos\varphi \end{pmatrix}, t \in \mathbb{R}. \tag{8}$$

The observed attenuation at the detector can now be calculated by the line integral

$$\int_{\binom{x}{y}\cdot\binom{\cos\varphi}{\sin\varphi}=r} \mu(x, y)ds = \int_{-\infty}^{\infty} \mu(r\cos\varphi - t\sin\varphi, r\sin\varphi + t\cos\varphi)dt, \tag{9}$$

using the parameterization introduced in Eq. 8. The function $\mu(x, y)$ can thus be transformed

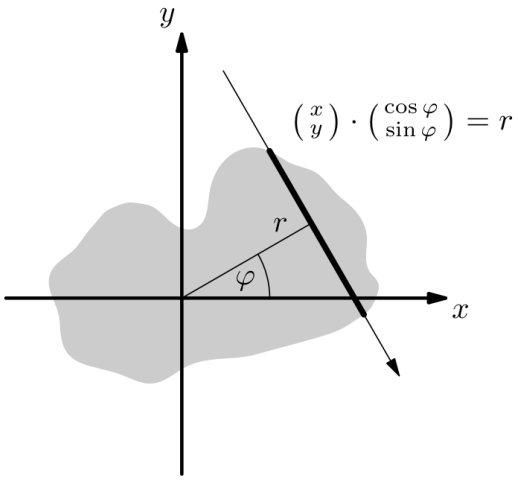


Figure 14: The x-ray beam can be described by its orthogonal distance r from the origin and the angle of its normal vector. (Adapted from [5], Creator: M. Figl)

to another function $R_f(\varphi, r)$ defined on [0,2π]$\times\mathbb{R}$:

$$\begin{aligned} \mu(x,y) &\to R_f(\varphi, r) \\ R_f(\varphi, r) &\coloneqq \int_{\binom{x}{y}\cdot\binom{\cos\varphi}{\sin\varphi}=r} \mu(x,y)ds. \end{aligned} \tag{10}$$

This is called Radon transform. It is the overall attenuation in the ray direction. To sum up, a CT measures the attenuation along paths of different orientation and reconstructs the density distribution of the volume with this information [5].

2.2.2 Algebraic reconstruction

A discretized image can be represented as a big matrix of squares like in Fig. 15.

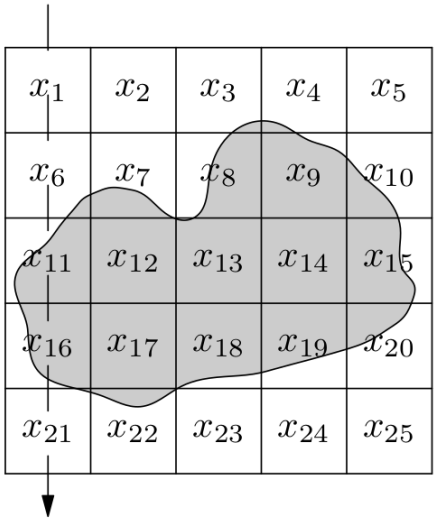


Figure 15: Discretization of an imaging volume. (Adapted from [5], Creator: M. Figl)

The task is to calculate the image values x_i, $i \in 1,...n^2$ out of the attenuation values they give rise to. Or in other words, the image has to be reconstructed from the Radon transform

$R(\varphi, r)$. When using for example n parallel beams and 180 angle steps, this results in $180 \times n$ linear equations. Since there are n^2 unknown x_i, the problem can be reformulated using a $(180 \times n) \times (n \times n)$ matrix S such that

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & ... & 1 & 0 & 0 & ... \\ 0 & 1 & 0 & ... & 0 & 1 & 0 & ... \\ 0 & 0 & 1 & ... & 0 & 0 & 1 & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \end{pmatrix}}_{S} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ ... \\ x_{n^2} \end{pmatrix} = \begin{pmatrix} R(1,1) \\ R(1,2) \\ ... \\ R(1,n) \\ R(2,1) \\ ... \\ R(180,n) \end{pmatrix}. \tag{11}$$

Matrix S is called system matrix and it describes the image formation. The elements of one matrix line are the coefficients of the x_i, so it describes which image elements contribute to which Radon transform. Additional weighting factors could be multiplied with the elements of S in order to account for the way the x-ray beam crosses a pixel. Depending on the image size, more angle steps need to be taken in order to be able to solve the system of equations [5].

In general, algebraic reconstruction can be done by computing the system matrix S and solving $Sx = r$, Eq. 11, iteratively. S can be calculated by knowing the CT geometry. When there a more equations than unknowns, there will most probably not be a solution that satisfies all equations at once. The solution for x can be generalized by minimizing $||Sx = r||$, but it is not necessarily unique. All other solutions x' can be expressed as $x' = x + w$, with $Sw = 0$ [5].

There is also a different approach than algebraic reconstruction, namely filtered backprojection, which is shortly explained in Ch. 2.2.3. It needs less processing time and memory but often leads to increased image noise, which can be corrected with higher X-ray tube settings. However, this results in a higher radiation dose for the patient [5, 12]. Since computational speed is becoming less of a problem nowadays, the focus of this thesis will be on algebraic iterative reconstruction.

While analytic reconstruction algorithms, like filtered backprojection, are an "open loop" system, iterative algorithms can be considered a "closed loop" system. In each iteration the projection data from the currently estimated image is compared to the measured projection data. The found data discrepancies are backprojected to the image space, meaning they are used to update the current image [13]. A schematic sketch of this general procedure is shown in Fig. 16. The iterative algorithm that is used for this thesis is SIRT (Simultaneous Iterative Reconstruction Technique). Recall Eq. 11, where S represents the (forward) projection operation. The back projection to the image space can be represented by its transpose S^T. The basic principle of SIRT is alternating forward and back projection with the following update equation:

$$x^{(t+1)} = x^{(t)} + CS^T R(r - Sx^{(t)}). \tag{12}$$

R and C in Eq. 12 are diagonal matrices with $R_{ii} = 1/\sum_j S_{ij}$ and $C_{jj} = 1/\sum_i S_{ij}$. They

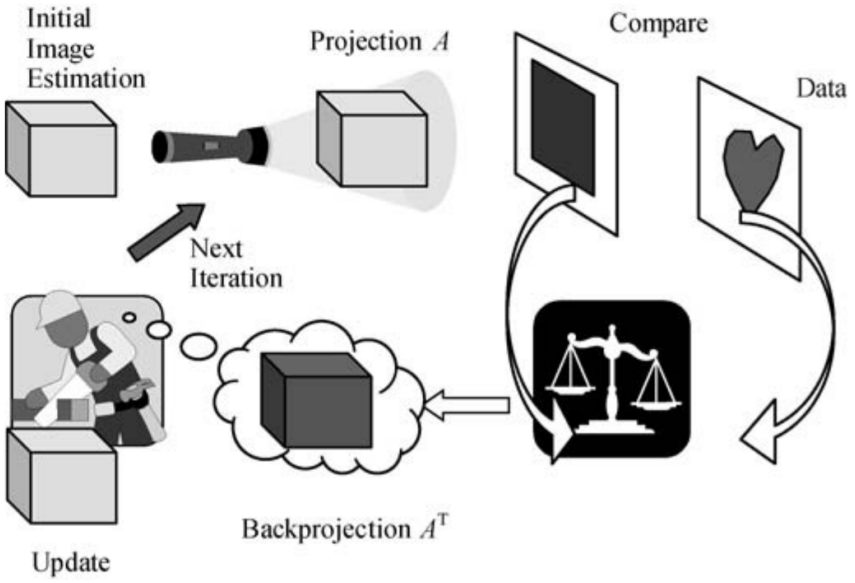


Figure 16: Basic principle of iterative image reconstruction algorithms. The projections of an initial image estimate are compared with real projection data. The discrepancies between the two are used to update the estimated image further. (Adapted from [13], Creator: G. L. Zeng)

compensate for the number of pixels that are hit by each ray and the number of rays that hit each pixel. Update Eq. 12 first forward projects the current image estimate $x^{(t)}$ and then subtracts the result from the original projections (Radon transform r) in $r - Sx^{(t)}$. The difference is backprojected with S^T, weighted with R and C. The obtained correction factor $CS^T R(r - Sx^{(t)})$ is then added to the current image estimate. This process is repeated until a satisfactory result is reached [14].

2.2.3 Filtered backprojection

Another common way to reconstruct an image from projection data is filtered backprojection. The process is illustrated in this chapter by considering an empty 2D plane with a small non-empty dot at an arbitrary position in the plane. A schematic is shown in Fig. 17. As mentioned before, a detector rotates around the origin and records projections of the object at different angles. A projection is formed by a line integral along the direction orthogonal to the detector (**a** in Fig. 17). If the point is not included in this integral, the projection value is zero. The value on the detector is the sum of all activities along the x-ray path, so we do not know where along this path the point is located. One way to overcome this problem is distributing the activity equally everywhere along the projection path (**b** in Fig. 17). Doing this for more projections angles result in Fig. **c**. This is called backprojection. Using all projection angles results in an image similar to **d** in Fig. 17. So backprojection results in a blurred version of the original image. To get rid of this blurring effect negative wings around the signal are introduced before backprojection. This is shown in Fig. 17 **e** and is called (high-pass) filtering. The result is a clear image similar to Fig. **f**. The whole image reconstruction procedure is called filtered backprojection [13].

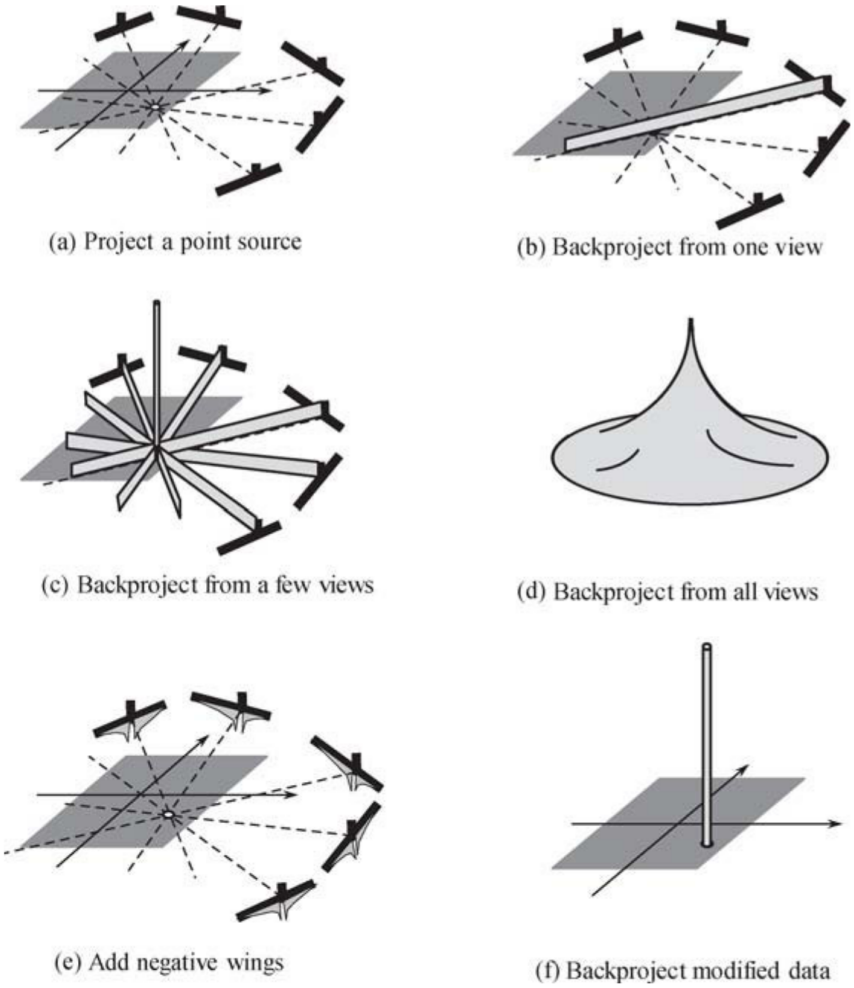


Figure 17: Unfiltered and filtered backprojection of a point source. (Adapted from [13], Creator: G. L. Zeng)

3 Cone-beam CT (CBCT)

3.1 CT vs. CBCT

In CBCT imaging is done by rotating a gantry with fixed detector and x-ray source. The cone shaped x-ray beam is directed through the volume of interest onto an opposite lying flat detector. This is illustrated in Fig. 18 **A**. During the gantry rotation planar projections of the field of view (FOV) are recorded, sometimes in a partial, but most of the time in a full arc [16].

In traditional CT a fan-shaped beam that moves on a helical path acquires image slices and thereupon stacks them together to get a 3D image. This means that all slices need a scan and 2D reconstruction of their own [16]. One important advancement in CT scanners is the development of X-ray detectors that enable simultaneous acquisition of more than one slice at once and therefore a faster image acquisition. This is done by utilizing several detector elements that are separated in z direction (see Fig. 18 **B**). CT systems like this are called Multi-Detector Row CTs (MDCTs). Even faster image acquisition without the need for higher rotational speed can be achieved with dual source CT. Two X-ray tubes and corresponding detectors are used with an offset of 90°. A system like this can for example be useful for imaging the motion of a heart with

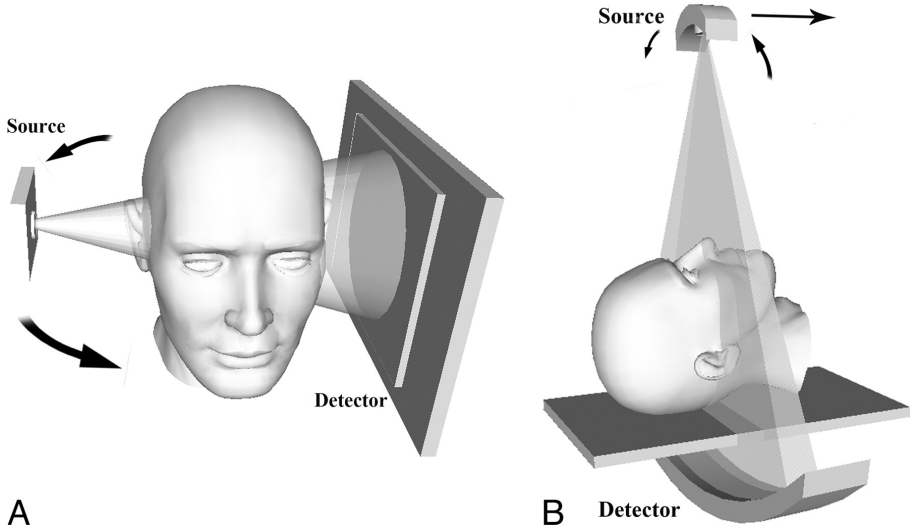


Figure 18: **A** Image acquisition geometry of cone-beam CT. **B** Image acquisition geometry of MDCT (right). (Adapted from [15], Copyright © American Society of Neuroradiology)

high temporal resolution. The main challenge here is cross-scattered radiation between the two sources and detectors. Another possibility is to operate both x-ray tubes at different energies. The two relevant interactions of x-rays in the body are Compton scattering and Photoelectric effect. They have a different dependency on the atomic number Z and the photon energy E. Therefore a material can be characterized by its change of attenuation when using two different energies. Clinical application for this technique are for example tissue characterization and calcium quantification [10].

In contrast to CT, CBCT needs only one rotational sequence for the entire FOV, since CBCT exposure comprises the whole FOV, as shown in Fig. 18 **A**. Therefore CBCT systems occupy less space than traditional CTs. Furthermore, they are less expensive. However, the main disadvantage of such a system is the limitation in image quality, especially with a larger FOV. This is related to a higher amount of recorded scatter radiation, which results in more noise and lower contrast [16].

3.2 Geometry of a CBCT device

A sketch of the standard CBCT geometry is shown in Fig. 19. The image is reconstructed at the isocenter on an XYZ coordinate system. The detector, which records the projections, has a different coordinate system UV. Similar to a CT, CBCT estimates the patient attenuation $\mu(\vec{r})$ at a position $\vec{r} = (x, y, z)$ from the projections. Beer's law for CBCT geometry can be expressed as

$$p(u, v, \alpha) = p_0 e^{-\int \mu(\vec{r_0}(\alpha)+\lambda\hat{l})d\lambda}. \tag{13}$$

Here $p(u, v, \alpha)$ is the projection value for rotation angle α. $\vec{r_0}(\alpha)$ is the source location and $\lambda \in [0, \sqrt{u^2 + v^2 + D^2}]$, where D is the distance between source and detector. $\hat{l}$ is a unit vector along the direction between source and detector coordinates u and v [17].

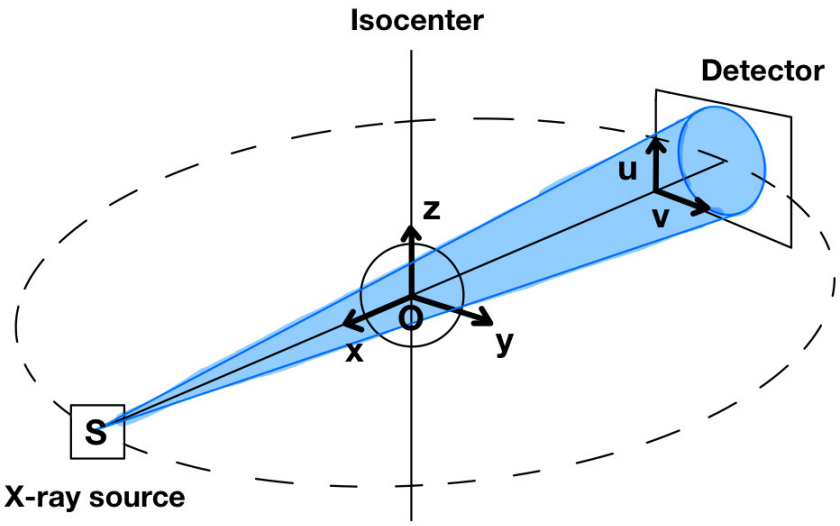


Figure 19: Sketch of CBCT geometry and circular source-detector trajectory.

There are again two types of algorithms that can be utilized for CBCT image reconstruction: analytic methods and iterative algorithms. FDK algorithm is analytic and it is advantageous in some cases, however with limited projection data its performance is poor. Iterative algorithms in contrast can produce good image quality with theoretically insufficient projection data [5, 17].

3.3 CBCT image production

3.3.1 Acquisition configuration

An x-ray source does either a full rotational scan or single partial scan, while a flat detector moves synchronously around a fixed fulcrum. Each projection is made sequentially. Theoretically, exposing the patient to a constant beam of radiation allows sampling the attenuation along the beam trajectory. However, this results in a higher radiation exposure and does not contribute to the image formation. An alternative is a pulsed x-ray beam, that coincides with the sampling of the detector [16]. Modern scanners use tube current modulation that alters the strength of the beam depending on the signal to noise ratio (SNR). The purpose is to maintain a constant SNR in the image despite varying thicknesses of the object while avoiding excessive irradiation. The reasoning is that the human body is usually less wide in anterior-posterior direction compared to the transverse direction and therefore it takes fewer x-ray photons for imaging [18].

The size of the field of view (FOV) depends on the detector dimensions, the projection geometry of the beam and the beam collimation. Collimation ensures that the radiation exposure is limited to the volume of interest (VOI). Therefore it is important that an optimal FOV is chosen for each patient, keeping the region to be imaged in mind. Extended FOV scanning (for larger VOIs) is difficult to realize because large-area detectors are expensive and not really available. Expanding the scan volume has been done by adding two rotational arcs for producing a single volume. Another method for achieving a wider FOV is applying an offset to the detector position

and collimating the beam in an asymmetric way. With this method only half of the patient can be scanned. A sketch of this arrangement is shown in Fig. 20 [16].

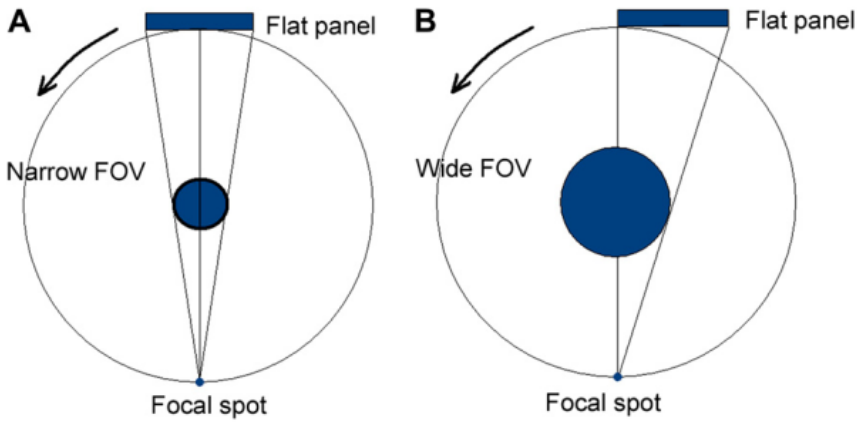


Figure 20: Method for extending the FOV with a flat panel detector. (A) Conventional arrangement. The central ray of the beam goes through the middle of the object to the center of the detector. (B) Alternative method with a shifted detector location and an asymmetric x-ray beam. (Adapted from [16], Creator: SOREDEX, Helsinki, Finland)

During the scan, projection images are recorded at certain degree intervals. Thus the projection data consists of a specified amount of images. The number of images depends on the frame rate, which is the number of images per second, the extent of the trajectory arc and the rotational speed. In general, more projections provide more information, meaning higher contrast, spatial resolution and signal-to-noise ratio. The result is a smoother image. The downside of recording more projections is a longer scanning time, a longer image reconstruction time and a higher radiation dose for the patient. Considering the "as low as reasonably achievable" (ALARA) principle, the projection number should be minimized while producing diagnostic image quality. Within a clinical setting, where there is a limit to scanning time and detector readout speed, the number of available viewing angles is usually limited to several hundred [16].

Another thing to consider is the extent of the trajectory arc. Most of the time a scan arc of 360° is used. This is done to improve the signal-to-noise ratio for adequate 3D reconstruction using the FDK algorithm. Nevertheless it is technically possible to reduce scanning arc trajectory completeness and still reconstruct a volume. Such images suffer from greater noise and reconstruction interpolation artefacts, but they have a shorter scan time and the procedure is easier to perform [16].

3.3.2 Image Detection

Flat-panel imagers are the standard in today's CBCT systems. There are two types of flat panel x-ray detectors: direct conversion and indirect conversion systems. In principle, detectors utilized in radiographical applications convert the x-ray photon energy into an electrical signal, that can be related to the magnitude of exposure in the respective pixels. In direct conversion systems, the x-ray energy is converted in a single layer of photoconductor. Various materials, like amorphous selenium, lead iodide and cadmium zinc telluride, can be used as photoconductors. When an x-ray

impinges onto the photoconductor conversion plate, an electron-hole pair is created. An electrical field, that is applied between the two surfaces of the photoconductor forces the separation of the electron-hole pair. They are then transferred to opposite surfaces. Each surface is actually a thin-film transistor (TFT) array, that can determine the amount of charge present. The TFT array translates to the array of pixels in the image, so the number of pixels equals the number of transistors. This detector type is mainly used in mammography. In indirect conversion systems a scintillator layer converts x-ray photons to visible light, because it is easier to detect. Materials used for this purpose are amorphous silicon and cesium iodide. The produced visible light can be again read out by a photoconductor. In indirect conversion systems the photoconductor consists of either CCDs or amorphous silicon coupled to a TFT array [1, 6].

Flat panel images provide a good performance, but they have limitations that are related to linearity of response to the spectrum of radiation, uniform response of the detector and deficient pixels. Therefore these detectors must be linearized and calibrated accordingly. Bad pixels are usually replaced by the average of neighbor pixels [16].

Smaller pixels record more detail (i.e. higher spatial resolution), but they capture fewer x-ray photons and thus cause more noise in the image. This means that CBCT imaging with smaller pixels requires a higher radiation dose for the patient [16].

The voxels produced from the CBCT data determine the resolution of the image. The dimension of the voxels primarily depends on detector pixel size, which is in submillimeter range. Therefore CBCT devices provide isotropic voxel resolutions, unlike conventional CT, where the voxel size also depends on the distance the table moves for every rotation [16].

3.3.3 Image reconstruction

After recording the projections, a volumetric data set must be created from them. Since there are usually several hundreds of projection images, each consisting of more than one million pixels (12 to 16 bits of data), the reconstruction is computationally expensive. The required time varies depending on acquisition parameters like voxel size, number of projections, and hardware as well as software. In order to complement patient flow, the reconstruction should be done in a reasonable time [16].

Raw CBCT detector images show some spatial variations. This is because of the spatial variability of the physical properties of the switching elements and photodiodes in the flat panel detector. In addition the x-ray sensitivity of the scintillator layer also varies. Furthermore, even high-quality detectors display some pixel imperfections or defects. Before reconstructing the image, those inhomogeneities have to be accounted for. After correcting the projection images, they have to be related to each other. This can be done by constructing sinograms, as shown in Fig. 11. Afterwards the sinograms are converted into 2D CT slices by a reconstruction filter algorithm. FDK algorithm is the most widely used reconstruction algorithm for cone-beam data. Once all

the slices are gathered, they are recombined to form a 3D image [16].

3.4 Applications of CBCT

3.4.1 Dentistry

Panoramic radiographs are often used in dental radiology, however, like all planar 2D projections, they have some limitations including magnification, distortion, superimposition and misinterpretation of structures. There has been some effort in the direction of 3D radiographic imaging, and although CT is available, dental applications are not common. The reasons are cost, large device size and dose considerations [16].

CBCT has facilitated the transition from 2D to 3D image acquisition in dentistry, because of its ability to produce good images of structures with high intrinsic contrast. This makes it particularly interesting for imaging osseous structures in the craniofacial region. Another advantage is the rapid scan time, which reduces artefacts due to subject movement. The reason is that CBCT records all projections in one single rotation. Furthermore beam collimation limits the area that is exposed to radiation to the VOI. This means that an optimal FOV can be chosen for each patient. The radiation dose in CBCT is therefore lower than in conventional CT. The radiation dose of a single panoramic x-ray is even lower, but only provides a 2D image. CBCT provides submillimeter isotropic voxel resolution, which enables high precision in all dimensions. This can be of advantage for orthodontic analysis and implant site assessment. Additionally the images can also be sectioned nonorthogonally, because of the isotropy. Some of these reasons, enable CBCT to not only be useful for diagnosis, but also for image guidance of surgical and operative procedures [16].

3.4.2 C-arm CBCT in interventional radiology

The term C-arm CBCT is used to describe a CBCT system, that is mounted on a C-arm. A device like that (Philips Allura FD20 Xper) is shown in Fig. 21 [17]. Flat-panel detectors are used by almost all commercially available C-arm CBCT systems [17].

The term interventional radiology is used for minimally invasive surgical procedures that are guided by medical imaging [19]. C-arm CBCT is suitable for interventional applications for several reasons. One being its ability to rotate and translate in arbitrary directions, allowing the patient to remain stationary during the interventional procedure. Another reason is that a single rotation allows collecting a full projection data set from which a reconstruction in submillimeter resolution can be generated. Furthermore, state of the art flat panel detectors allow detection of structures with low contrast. Another advantage is the low dose compared to CT [17].

There are many clinical applications for interventional 3D imaging. It is used for example in the neuro-interventional field, especially for aneurysm classification. Additionally CBCT can potentially be utilized in surgical planning in orthopaedics and head, neck, abdominal or thoracic

Figure 21: Philips Allura FD20 Xper C-arm. (Adapted from [17], Creator: S. Hatamikia)

surgery. In dentistry, CBCT is used for maxillofacial and oral surgery [17].

Before acquiring the CBCT data, some steps have to be taken. First, the patient is positioned in a way that the VOI is visible in all 2D projections. Then, a safety test is done by rotating the C-arm around the patient in order to find possible collisions during the actual scan. The last step before recording the actual image is to take an x-ray that initializes the automatic exposure control (AEC). With this, the exposure parameters can be adjusted in a way that the detector always receives a constant entrance dose, which means the dose delivered to the patient can be monitored [17].

3.4.3 CBCT in image guided therapy

Image guided therapy is when imaging is used before or during radiotherapy, with the goal of improving target accuracy and thus allowing more accurate delivery of dose. CBCT offers volumetric imaging during radiotherapy sessions and has replaced 2D based image guided radiotherapy. The most common use in a clinical setup is patient positioning, which can replace immobilization frames and fiducial markers. An advanced application of CBCT is treatment adaption, when anatomical changes occur during treatment (for example due to breathing motion) [17].

3.5 Limitations of CBCT

Limitations in CBCT are related to detector sensitivity, projection geometry and contrast resolution. Images often lack the usefulness and quality/resolution of conventional CT. The quality is affected by noise and poor soft tissue contrast. The higher noise comes from the fact that scattered radiation is recorded by pixels, which do not reflect the original x-ray beam path. In

CT this effect is less pronounced because no area detector is used and it is operated at higher tube currents. The scattered radiation also significantly reduces the soft-tissue contrast and can cause streak artefacts and wrong density values [16, 20].

Furthermore, the divergence of the beam causes the Heel effect, which produces a greater signal-to-noise ratio (SNR) on the cathode side compared to the anode side. The result is nonuniformity of the incident x-rays and thus nonuniformity in the absorption [16, 20].

Another common problem in CBCT imaging is artefacts. They are defined as visible structures in the reconstructed image that are not present in object. Their cause is discrepancies between the physical imaging process and its simplified mathematical modelling [20].

Extinction artefacts or "missing value" artefacts are caused by highly absorbing materials. Those materials result in intensity values that are either zero or close to zero in the respective detector elements and consequently no attenuation can be computed. When trying to reconstruct the image severe artefacts are induced [20].

One of the more prominent sources of artefacts is beam hardening. Lower energy photons are substantially absorbed when passing through an object, while higher energy photons can still pass. This means that there are more low energy photons present in the spectrum, than actually recorded on the detector (i. e. the beam is hardened). Dark streaks are caused by this non-linear error (relatively higher energy on the detector) in the recorded data [20].

The exponential edge gradient effect appears when there are sharp edges between structures. The cause is that the measured intensity is averaged over a finite beam width, while the mathematical models assume zero width for reconstruction. The finite width is affected by the detector pixel size, the focal spot and the imaging geometry. Typical artefacts arising from this effect are streak tangents to sharp edges in beam direction [20].

Aliasing artefacts are common in CBCT as well. The Nyquist sampling theorem requires that for complete reconstruction of a continuous signal, the frequency of sampling must be larger than twice the highest frequency in the signal. For imaging this means that the pixel size must be small enough, otherwise it would result in undersampling. Another factor causing aliasing artefacts in CBCT is the divergence of the cone beam. Voxels close to the source will be crossed by more rays than the ones closer to the detector, as shown in Fig. 22. This means that the cone divergence itself causes undersampling. Another cause for aliasing artefacts can be crude interpolation methods for calculating the beam path length inside a certain voxel. Aliasing artefacts represent themselves as line patterns, that usually diverge towards the outer part of the reconstructed image. This is called Moire pattern and is shown in Fig. 23. Better interpolation schemes can be used to avoid these types of artefacts [20].

Ring artefacts are caused by uncalibrated or defect detector elements. Because of the usu-

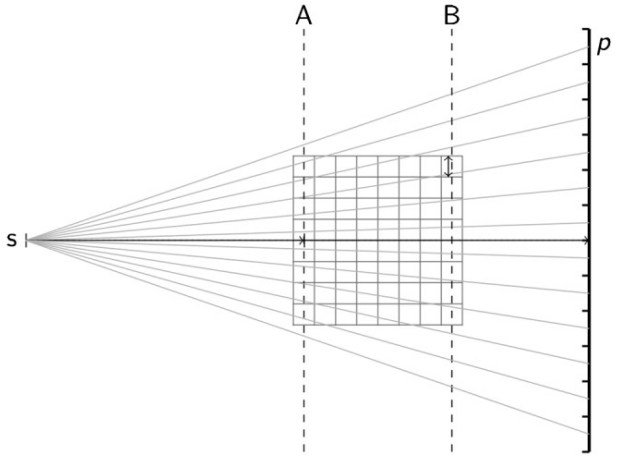


Figure 22: The divergence of the cone beam results in undersampling itself. Slice A is crossed by more rays per voxel than slice B. (Adapted from [20], Copyright © 2011 The British Institute of Radiology)

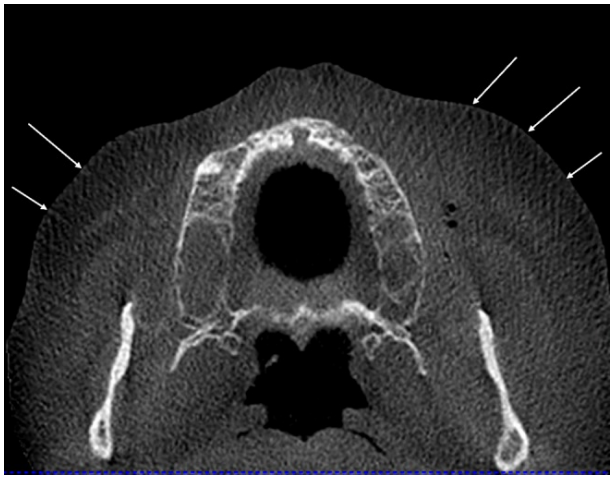


Figure 23: Typical aliasing or Moire patterns in CBCT, resulting from undersampling. (Adapted from [20], Copyright © 2011 The British Institute of Radiology)

ally circular trajectory, these inconsistencies appear as rings centered around the axis of rotation [20].

Motion or misalignment artefacts are either caused by misalignment of source, detector or object, or by movement of the patient. The reconstruction does not account for these problems, which results in a faulty image. The higher the resolution, the more likely it is that motion artefacts appear. The reason is that for smaller voxel sizes, the required movement to place structures out of the 'right' voxels is smaller. Fixation of the patient and great mechanical stability help in avoiding these artefacts [20].

3.6 Conventional CBCT source-detector trajectory limitations

The conventional way to record a CBCT is by rotating source and detector around the patient in a circular trajectory. Thereby 2D projections are acquired at equal intervals from which a 3D image can be reconstructed. However, the circular trajectory brings some limitations [17]. For the reconstruction of a 2D Slice in CT it makes sense that source and detector orbit a common

isocenter. However, for analysing 3D structures this trajectory is not ideal, because it does not generate enough information for the whole object. Tuy's sufficiency condition for cone-beam data states says that any plane that goes through a point in the object to be imaged must have an intersection with the source trajectory, for precise reconstruction [21]. It is illustrated in Fig. 24. Because of the simplicity of the circular trajectory it is still the standard, even though only points

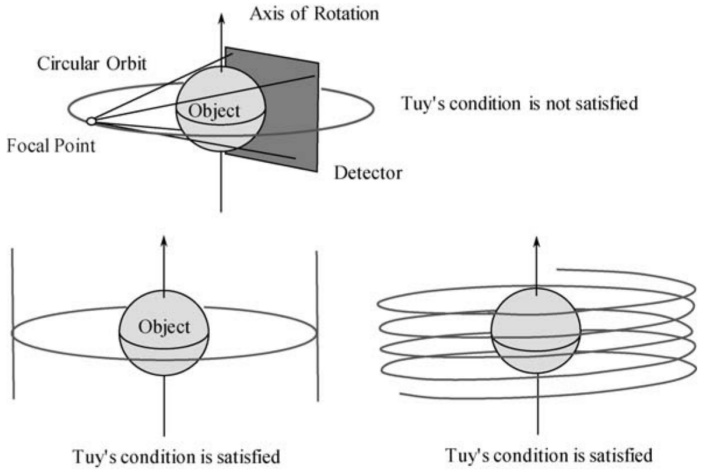


Figure 24: Illustration of Tuy's cone-beam data-sufficiency condition. The circular source trajectory does not satisfy Tuy's sufficiency condition. The other two trajectories shown in contrast theoretically generate enough information for precise reconstruction. (Adapted from [21], Creator: G. L. Zeng)

inside the trajectory plane are fully sampled [17].

Another limitation is the radiation dose obtained by the patient. A high number of projections is needed to reconstruct a volume properly from a standard circular source-detector trajectory, which leads to a considerable radiation dose for the patient. Therefore it would be beneficial to decrease the dose needed for these procedures [17].

Furthermore a wide angular range is needed to acquire the projections. Usually CBCTs either do full (360°) or a half (180°) rotations. However, sometimes only small sets of projections can be gathered due to limited available space, e.g. during surgery. In these cases the feasible angular range is less than 180° if collisions are to be avoided [17].

Circular trajectory also implies that the rotation can only be done around the cranio-caudal axis of the patient. Information from oblique viewing angles can thus not be accessed. If a certain target region has to be reconstructed, this can be unfavorable [17].

C-arm CBCT is an important tool in interventional radiology. Nevertheless, there is a major drawback with those devices. Commercially, the available field of view (FOV) is limited by the x-ray detector size, which is often not large enough to obtain information about peripheral anatomical areas. Especially for patients of larger size this limitation is crucial, when the volume of interest (VOI) is located outside of the standard FOV. There are some commercial solutions to this problem, including larger detector sizes, merging two images and increasing the space be-

tween x-ray tube and detector. The drawback of these solutions is that they lead to an increased radiation dose [17].

Initially, non-conventional or non-isocentric trajectories were used to address FOV and sampling issues in interventional CBCT. In the last decade a number of studies have been published with a broad variety of goals and methods and have reported various clinical advantages of non-standard trajectories [22].

The interventional field provides an ideal environment for customizing source-detector trajectories specifically to a patient, because a prior medical image usually exists. Novel robotic interventional systems provide high flexibility in movement, offering a breadth of alternative CBCT trajectories. As mentioned before, one limiting factor for CBCT in more image-guided procedures and radiation therapies is the small FOV. Clinically this means that long (e.g. spine, aorta, head and neck) and wide (e.g. pelvis, liver, thorax) body parts can not be captured with a single CBCT image. When leaving the hardware components the same, extended FOV imaging can be accomplished by non-standard trajectories. Extending the lateral FOV with a C-arm CBCT has been solved by rotated detector, offset detector (Fig. 20) and utilizing a dual isocenter. Pre-programmed elliptical isocenter trajectories are also available to extend the lateral FOV. However, increasing the lateral coverage reduces the longitudinal coverage [22]. These solutions are not patient specific and may therefore not provide an optimal trajectory for a certain VOI. Leading candidates for extending the longitudinal FOV in a clinical setting are multiple parallel circles, reverse helix (standard helix is not feasible, because CBCT systems are usually limited to a finite rotation) and the line-ellipse-line trajectory. They are schematically shown in Fig. 25 [22].

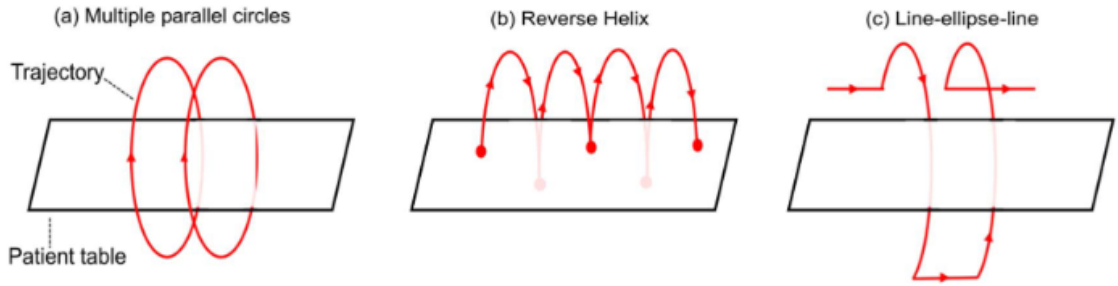


Figure 25: Different methods for extending the longitudinal FOV in a clinical setting. (Adapted from [22], Creator: S. Hatamikia)

The aim of this thesis is to investigate source-detector trajectory optimization by means of a heuristic optimization algorithm, which will be discussed in Ch. 4. The goal is to customize trajectories for imaging a certain predefined target region optimally and to extend the FOV, taking into account kinematic constraints. Another goal is to reduce the amount of projections and still get a diagnostic image quality. This is done by optimizing the isocenter position and imaging angle when recording the projections.

4 Global heuristic optimization algorithms

Heuristic optimization algorithms exist in a broad variety of forms. These methods are usually easy to implement, understand and use. They are not restricted to any specific mathematical problems and therefore can solve different tasks. Heuristic optimization algorithms have the advantages of being fast and not having the need for a large memory space. Therefore they can solve complex and large problems in a reasonable time. The disadvantage is that they do not guarantee finding the optimal solution, but only a near-optimal solution, in contrast to exact optimization algorithms. Furthermore they give no information on how close the obtained solution is from the optimal one. In an unlucky case, finding a near-optimal solution can require a large amount of time [23, 24].

There are different classification schemes for heuristic optimization algorithms. They can for example be nature or non-nature inspired, they can be memory-less or have some sort of memory usage and they can be deterministic or stochastic. Furthermore, they can utilize single-solution-based search or population-based search, meaning either one or more solutions are transformed or manipulated during optimization. Lastly, there are iterative and greedy heuristic optimization algorithms. Iterative algorithms start with a complete solution and transform it at each iteration, while greedy algorithms start from an empty solution. Then at each step some decision is made until a complete solution is obtained [23, 24].

All in all, heuristic methods allow tackling large-size problems in a short amount of time, while delivering satisfactory results. They are used for many applications including engineering design, machine learning, system modeling and planning in routine problems [23, 24].

An example of heuristic optimization is genetic algorithms. They can solve both unconstrained and constrained optimization problems based on natural selection, which is the process that is responsible for biological evolution. The algorithms modify a population of solutions in order produce better ones. This is done by selecting solutions from the current solution population, that are supposed to serve as parents for the next generation and produce offspring solutions. With the number of generations getting higher, the population becomes fitter and evolves towards an optimum. There a three main groups of rules which determine how the next generation is created from the current population. The first group comprises selection rules. They choose the solutions, called parents, that produce the next generation. In general the selection is stochastic but can depend on the solutions' scores. The second group of rules is responsible for crossover, meaning how two parents form children for the subsequent generation. The last set of rules are mutation rules. They apply random changes to the parents, which serve as children afterwards. Some parents also survive until the next generation without any change. These are called elite children. Genetic algorithms can solve optimization problems with objective functions not well suited for traditional techniques [25]. Similar to genetic algorithms many other heuristic optimization techniques exist, such as Simulated Annealing.

4.1 Simulated Annealing

The principle of Simulated Annealing algorithm is an analogy with thermodynamics. The molecules of a liquid metal are free to move independent of each other at high temperatures. Cooling the liquid slowly, results in loss of thermal motility. Slow cooling often enables the formation of ordered structures (pure crystals), meaning individual atoms lining up over great distances in all directions. The crystal state is the minimum energy state of the system and amazingly enough nature is able to find it during slow cooling. If a liquid metal is cooled down in a faster way, the minimum energy state is not reached. The result is often an amorphous or polycrystalline state with a higher energy. The essence of this process is therefore slow cooling, allowing enough time for atom redistribution as they lose motility. This is what annealing means and it is important for finding a low energy state [26].

Other optimization algorithms often go immediately downhill from the starting point as far as possible. However, this often results in only finding a local and not a global minimum. Simulated Annealing has its own way of dealing with this problem. The Boltzmann probability distribution,

$$\text{Prob}(E) \sim exp(-E/kT), \tag{14}$$

describes a system in thermal equilibrium at temperature T. It's energies E are distributed probabilistically according to Eq. 14. This means that even at low temperatures, the system can theoretically still be in a high energy state. Therefore there is a chance that the system can get out of a local energy minimum and thus find a better, more global, minimum. So the system can go uphill and downhill in the energy landscape, but at lower temperatures it becomes less likely that that a significant uphill movement occurs. k in Eq. 14 is the Boltzmann constant and relates temperature to energy. Assuming that a thermodynamic system changes its state from energy E_1 to E_2 with a probability $p = exp[-(E_2 - E1)/kT]$, if $E_2 < E_1$ the probability would be greater than 1. In such cases the probability is arbitrarily set to 1, i. e. the system always takes this option. The general scheme is to always take downhill steps, while only sometimes taking uphill steps [26].

To apply this method to problems other than thermodynamic systems some elements need to be provided: The possible system configurations x, a way of making random changes in the system configuration $x + \Delta x$, an objective function $f(x)$ (which should be minimized) and a control parameter T alongside an annealing schedule that steadily reduces it. There are many different ways of how to provide these elements [26].

5 Programs & Materials

5.1 TIGRE

TIGRE is an open-source toolbox for 3D tomographic reconstruction. It focuses on iterative reconstruction that runs on the GPU for improved speed. It can be built in MATLAB or Python and therefore combines high level abstraction with the performance of CUDA. TIGRE allows

flexible CT geometry including Cone Beam, Parallel Beam and C-arm CT but also any other geometry. It contains many reconstruction algorithms like FDK (Feldkamp, Davis and Kress), which is classical filtered backprojection for cone-beam geometry. SART (Simultaneous Algebraic Reconstruction Technique), OS-SART (Ordered Subset - SART) and SIRT (Simultaneous Iterative Reconstruction Technique) algorithms are implemented too and mainly differ by the number of projections used to simultaneously update the image according to equation 12. TIGRE also contains more complex algorithms like ASD-POCS (Adaptive Steepest Descent Projection Onto Convex Set), which works well when the data is particularly noisy [27].

5.2 MATLAB Global Optimization Toolbox

The Simulated Annealing algorithm from the MATLAB Global Optimization Toolbox is used for this thesis [28]. It can solve both unconstrained and bound-constrained optimization problems. One iteration begins with generating a new point on the available parameter space randomly. The distance of this point from the last generated point follows a probability distribution dependent on temperature. The distance is determined by multiplying a random number in the interval [0;1] by the current temperature. Since arbitrarily small steps are not feasible, integer constraints are applied by rounding the result. The code for applying integer constraints is from [29]. The step direction is determined by multiplying the result by either 1, 0 or -1. Steps that would disregard the set boundary conditions are modified by being backprojected into the feasible region [29]. All new points with an objective lower, but also, with a certain probability, points with an objective higher than the current point are accepted. The latter avoids that the algorithm gets stuck in local minima during early iterations. The temperature is controlled by the annealing schedule and is steadily decreased, while the best found point so far is stored. The initial SA temperature is set to 100 and it is lowered by a factor of 0.95 each iteration [30]. With decreasing temperature, the algorithms search reduces its extent. After a certain amount of points were accepted, reannealing takes place, meaning that the temperature is raised again and the process is repeated to potentially find a better solution this time. The algorithm stops after a stopping criterion is reached [31, 32].

5.3 XCAT digital phantom

The XCAT digital phantom models whole-body anatomies for both male and female adults in a highly detailed way. It contains many different parameters for changing the anatomy and can for example perform respiratory or cardiac motions. An image that illustrates what a slice of the phantom can look like is shown in Fig. 26 [33].

5.4 Materials

The utilized MATLAB version is 19.4.0 (R2023a) [34] and the employed GPU is a NVIDIA TITAN Xp. Furthermore, version 11.8 of CUDA and version 2.4 of TIGRE is used [27, 35].

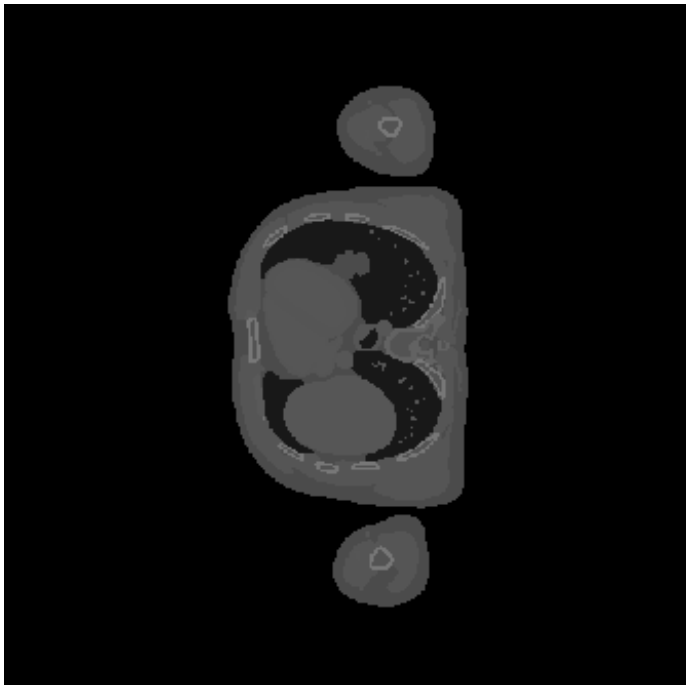


Figure 26: Sample slice of the XCAT digital body phantom [33].

6 Methods

6.1 Trajectory optimization

CBCT geometry with a source-detector distance of 1195 mm and a source-origin distance of 810 mm (compare with Fig. 19) is used for the simulations. The detector has a size of 512x512 pixels and the pixel size is 0.7413 mm x 0.7413 mm.

For this project, elliptical isocenter trajectories are chosen to extend the field of view. The ellipses have several parameters, namely the two axes lengths, x-, y-, z-offset from the origin and projection number, as shown in Fig. 27.

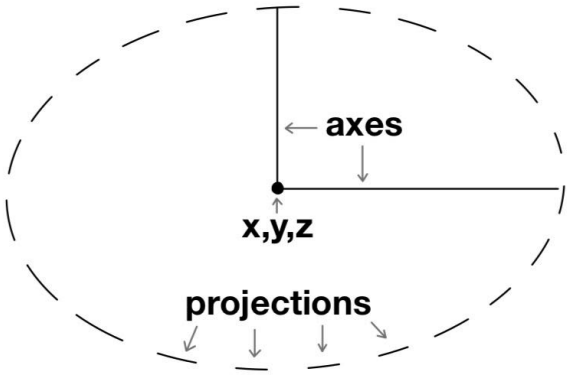


Figure 27: Illustration of the ellipse parameters used for trajectory optimization.

These parameters are optimized by Simulated Annealing, or in other words, they define the available parameter space. The boundary conditions for the shifts are 100 mm and -100 mm. The two axes of the ellipse can be between 1 mm and 100 mm. Since the resolution of the ellipse

parameter space can as a matter of fact not be arbitrarily small in reality, integer constraints are applied. Furthermore 360 projections are used for reconstruction, so this parameter is set to a constant value during optimization. Kinematic constraints are introduced as well. If an ellipse is chosen, that includes infeasible points, they are set to the maximum possible value. The maximal possible values correspond to the maximal shifts of 100 mm and -100 mm in x and y direction. The z value sets the vertical coordinate of the ellipse.

For the simulations an XCAT digital body phantom is used. Lesions are placed at different peripheral positions in the phantom's lung, which would not be fully sampled with the standard circular trajectory. The goal is to reconstruct the regions around the lesions optimally and to extend the FOV. For the final image reconstruction the SIRT algorithm with 100 iterations from the TIGRE toolbox is employed. During optimization process only 20 SIRT iterations are used for a faster process (more SIRT iterations result in a longer reconstruction time and better image quality).

The objective function is based on the universal quality index (UQI), which is a measure for image quality. The index is designed to take into account loss of correlation, luminance distortion and contrast distortion, which makes it perform better than the mean squared error. If $\mathbf{x} = \{x_i | i = 1, 2, ..., N\}$ and $\mathbf{y} = \{y_i | i = 1, 2, ..., N\}$ are the test and original image signals, the UQI can be calculated as

$$\text{UQI}(\mathbf{x}, \mathbf{y}) = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\sigma_x^2 + \sigma_x^2)[\bar{x}^2 + \bar{y}^2]}, \tag{15}$$

with

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i, \quad \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i, \tag{16}$$

$$\sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2, \quad \sigma_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y})^2, \tag{17}$$

and

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y}). \tag{18}$$

UQI can be between 1 and -1, where 1 indicates the best image quality and -1 indicates the worst [36]. Simulated Annealing goes through different trajectories and tries to optimize the ellipse parameters based on UQI in a specified volume of interest (VOI). To be more precise, the objective function value which is minimized during optimization is 1-UQI (in order to maximize UQI). As a stopping criterion for Simulated Annealing 1000 objective function evaluations are chosen.

To approximate FOV extension another image quality metric is utilized, namely structural similarity index measure (SSIM). It assesses the image quality based on luminance, contrast and structure. Using the same variable names as for UQI, the formula for SSIM can therefore be expressed as a multiplication of the luminance term $l(\mathbf{x}, \mathbf{y})$, the contrast term $c(\mathbf{x}, \mathbf{y})$ and the

structure term $s(\mathbf{x}, \mathbf{y})$ as

$$\mathrm{SSIM}(\mathbf{x}, \mathbf{y}) = [l(\mathbf{x}, \mathbf{y})]^{\alpha}[c(\mathbf{x}, \mathbf{y})]^{\beta}[s(\mathbf{x}, \mathbf{y})]^{\gamma}, \tag{19}$$

where α,β and γ are larger than 0. These parameters adjust the relative importance of the components and are usually set to 1. The luminance term is defined as

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\bar{x}\bar{y} + C_1}{\bar{x}^2 + \bar{y}^2 + C_1}, \tag{20}$$

where C_1 is a constant that avoids instability when the remainder of the denominator is close to zero. The contrast term is defined as

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}. \tag{21}$$

Similar to C_1, C_2 is a constant that avoids instability of the term. Lastly, the structure term is defined as

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}, \tag{22}$$

where constant C_3 is usually set to $C_2/2$. The whole formula for the SSIM is therefore

$$\mathrm{SSIM}(\mathbf{x}, \mathbf{y}) = \frac{(2\bar{x}\bar{y} + C_1)(2\sigma_{xy} + C_2)}{(\bar{x}^2 + \bar{y}^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \tag{23}$$

Setting the two constants C_1 and C_2 to zero actually returns an expression that is equal to the formula for the UQI (Eq. 15) [37].

To approximate the FOV extension the 3D images from the circular and the optimized reconstruction are split into many small cubes of side length 10 voxels. Then the SSIM value is calculated for all of the cubes that do not only have zero intensity values in the original image. The cubes that lie in the background (black region) are therefore ignored because they would distort the result (we do not care about the image quality in the background). A comparison of the number of cubes with an SSIM value above a certain threshold gives a measure for the FOV extension. The threshold is chosen as 0.98, because it showed a good compromise for image quality.

Trajectory optimization is also attempted for a bigger VOI, which comprises a large part of the thorax. To allow more flexibility in isocenter movement for this task, the set boundary conditions for the elliptical trajectory are altered. The long and short axis can now be between 1 mm and 200 mm, the x-, y- and z-offset can be between -200 mm and 200 mm and the kinematic constraints hinder movements outside of the region $\pm$200 mm in each direction. Other than that the same procedure as before is repeated. The goal is to reconstruct the lung without truncation artefacts.

6.2 Projection selection optimization

Another desired outcome of trajectory optimization is dose reduction. This can be achieved by reducing the amount of taken projections. However, this also reduces image quality. Within this project an attempt is made to optimize projection selection when trying to reduce their amount.

This is done by parameterizing the trajectory and the projection selection at the same time and then giving the parameters to Simulated Annealing for optimization. First an elliptical isocenter trajectory (similar to Ch. 6.1) including 360 projections is chosen, but then only some of those projections are kept for image reconstruction. The elliptical trajectories can again have a long and short axis between 1 mm and 100 mm and an x-, y- and z-offset between -100 mm and 100 mm. Kinematic constraints prevent isocenter movements outside of the region between -100 mm and 100 mm in each direction.

The number of projections is reduced to 60. Regular spaced projections on the optimized trajectories found before (in Ch. 6.1) are selected as an initial value for the optimization. Furthermore they also serve as comparison for the optimized projection selection. The final optimized image is again reconstructed with 100 SIRT iterations, while during the optimization process only 20 SIRT iterations are utilized. Since less projections are used, the image reconstruction is faster, which results in a faster optimization process.

The projection selection parameterization includes 6 parameters. The desired projections are split equally in three parts. If the number of desired projections is not divisible by three, the last projection set gets a different number of projections. This number is calculated by rounding down the result of the division, multiplying it by two and subtracting it from the number of desired projections. All of these three projection sets can have a different starting point and a different projection spacing. Furthermore, one projection can not be selected twice. In that case, the next available projection is kept. The boundary conditions for all starting points are between 1 and total number of projections and the three spacings can be between 1 and 10. Examples of how projection selection can look like are shown in Fig. 28. Optimizing those parameters by means of heuristic optimization algorithms allows for optimal image acquisition while reducing patient dose.

6.3 Projection selection optimization on real data

To confirm that the proposed projection selection optimization does not only work with a digital phantom, the optimization was repeated using real projection data. CBCT geometry with a source-detector distance of 1195 mm and a source-origin distance of 810 mm (compare with Fig. 19) is used for the simulations. The detector has a size of 512x364 pixels and the pixel size is 0.7413 mm x 0.7413 mm. The projections were created in a standard circular trajectory, with the same imaging geometry mentioned for the simulations.

First a ground truth was created from all projections, which served as digital phantom dur-

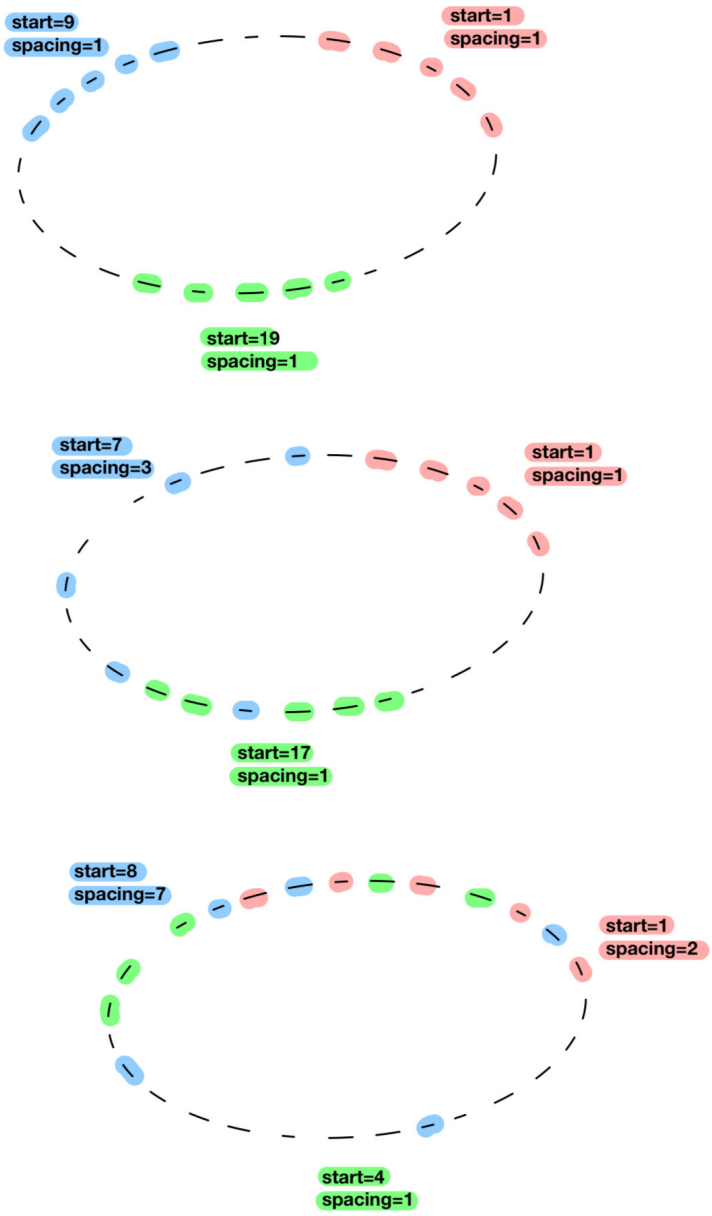


Figure 28: Illustration of the projection selection parameters used for optimization. The black markers correspond to the isocenter positions where projections were recorded. The red, blue and green markers indicate the three projection sets that are used for image reconstruction. All of them are parameterized by a starting point and a projection spacing. One projection can not be selected twice for reconstruction. In that case, the next available projection is selected.

ing the optimization. Then projections were simulated on this phantom, using only a sixth of the original viewing angles (taken from the real projection data). The same projection selection parameterization as in Ch. 6.2 was used, this time with a static isocenter. Then the UQI value between the reconstructed image from the selected simulated projections and the ground truth was computed in a certain VOI. The optimization aimed to select the projection angles, containing the highest amount of information about a predefined VOI and thereby increase the UQI in the VOI. The Simulated Annealing algorithm was again utilized for the optimization and 1000 objective function evaluations were chosen as stopping criterion. During optimization 20 SIRT

iterations were used for image reconstruction. As a last step, the best projection angles were then used to reconstruct the image from the real projections recorded at these angles using 100 SIRT iterations. The region around the tumor was selected as VOI.

Because there is no isocenter shift in the data there is not as much room for improvement as for the digital phantom. Therefore the achieved improvement might be small compared to the improvements achieved before. For this reason the same procedure as before is repeated, but with 100 SIRT iterations during optimization, in order to already have a more accurate result during the process (more SIRT iterations result in a better reconstructed image). Because this results in higher computational demands, Simulated Annealing was restricted to 300 objective function evaluations for reasons of time.

As a next step kinematic constraints were applied to the projections angles, in order to simulate avoiding potential collisions of the detector with surrounding equipment or the patient. Projections angles in the range of ± 15 degrees were excluded from the search space. The remaining projections angles were then given to Simulated Annealing for projection selection optimization, when reducing the number of projections to a third. Regular spaced projections on the remaining search space served as initial value for the optimization and as comparison for the optimized projection angles. For a more accurate search 100 SIRT iterations were carried out during optimization and the stopping criterion for SA was chosen as 300 objective function evaluations.

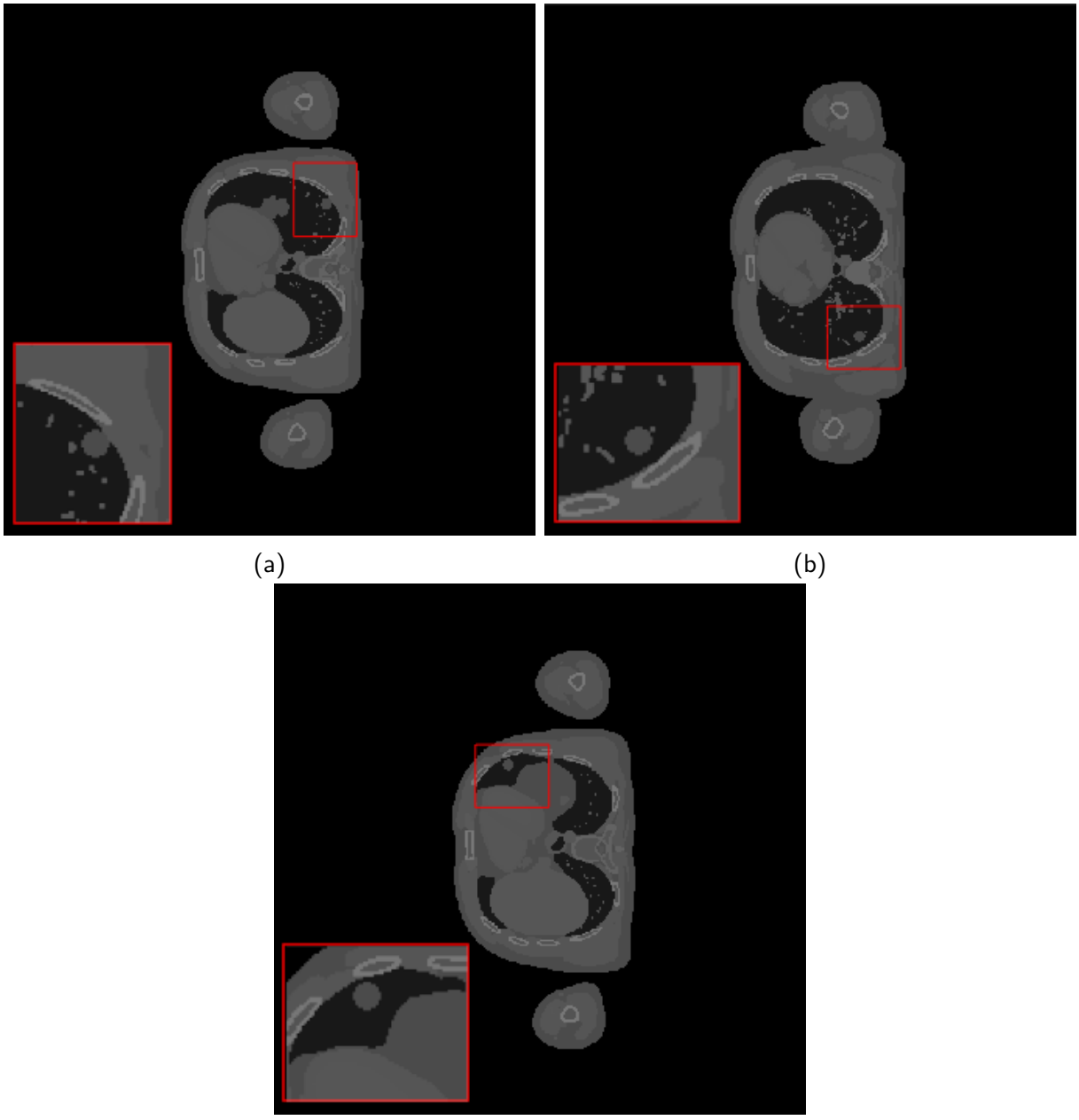
7 Results

7.1 Trajectory optimization

Three lesions are placed in different peripheral parts of the phantom's lung, as can be seen in Fig. 29. The three respective VOIs are marked with a red rectangle. A circular source-detector trajectory leads to Fig. 30. Parts of the lesions are visible, but the quality is poor. The red rectangles mark the VOI.

Optimized Trajectories lead to Fig. 31. The found isocenter trajectories are plotted in Fig. 32. Since kinematic constraints, that allow a maximum possible shift of ± 100mm in each direction only, are chosen, the ellipses get flat in some regions. The red markers indicate collision avoidance. The trajectories have different vertical coordinates, which is not visible in the plot.
The simulated annealing optimization processes (which show the current objective function value (1-UQI) plotted over the course of time) are depicted in Fig. 33. In Fig. 34 the best objective function value until a certain iteration number is plotted. The last point corresponds to the final UQI value.

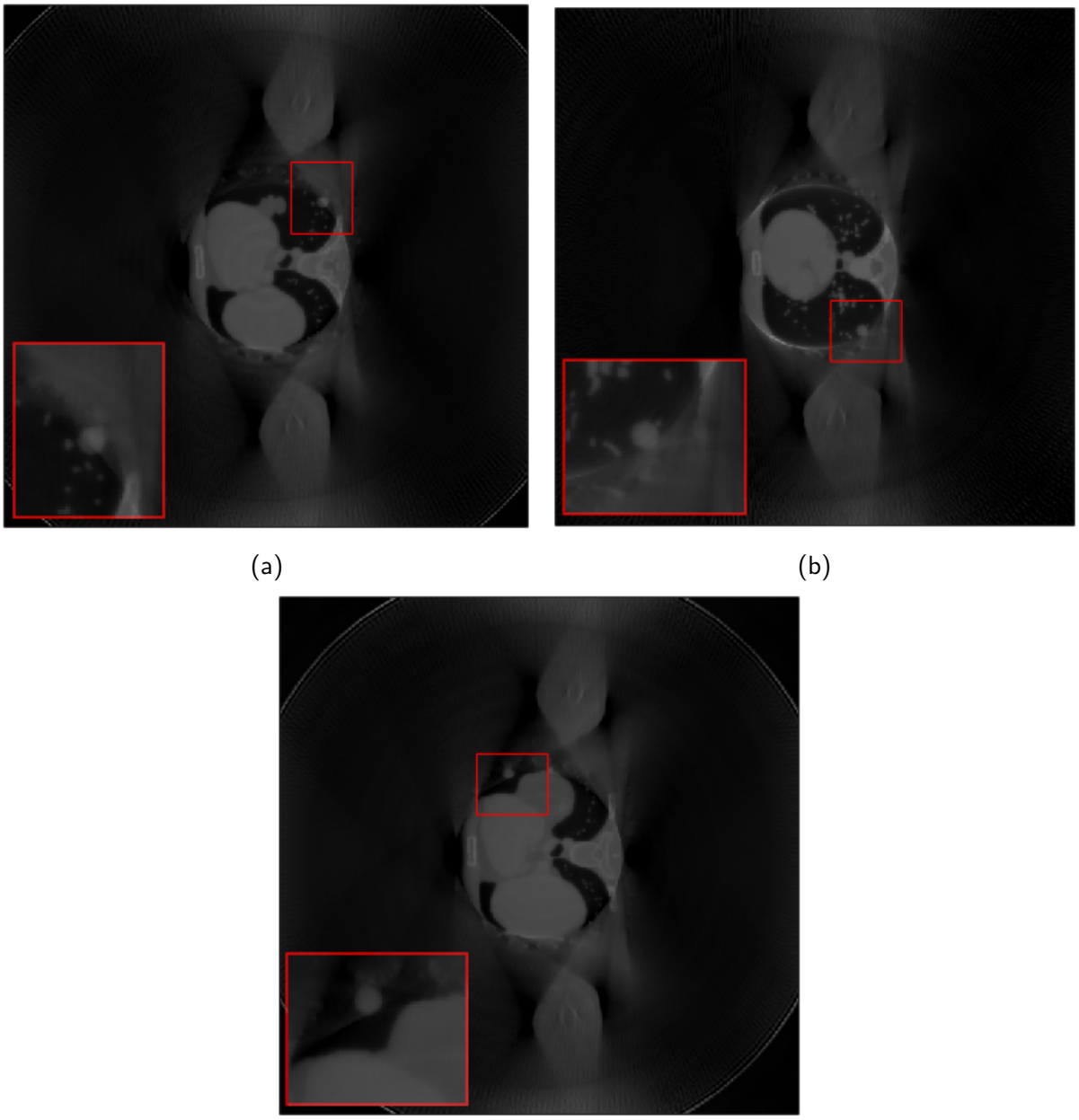
FOV extensions are illustrated in Fig. 35. The markers indicate the location of the cubes where a high enough SSIM value is achieved. The obtained results are listed in Tab. 1. They include



(a) (b)

(c)

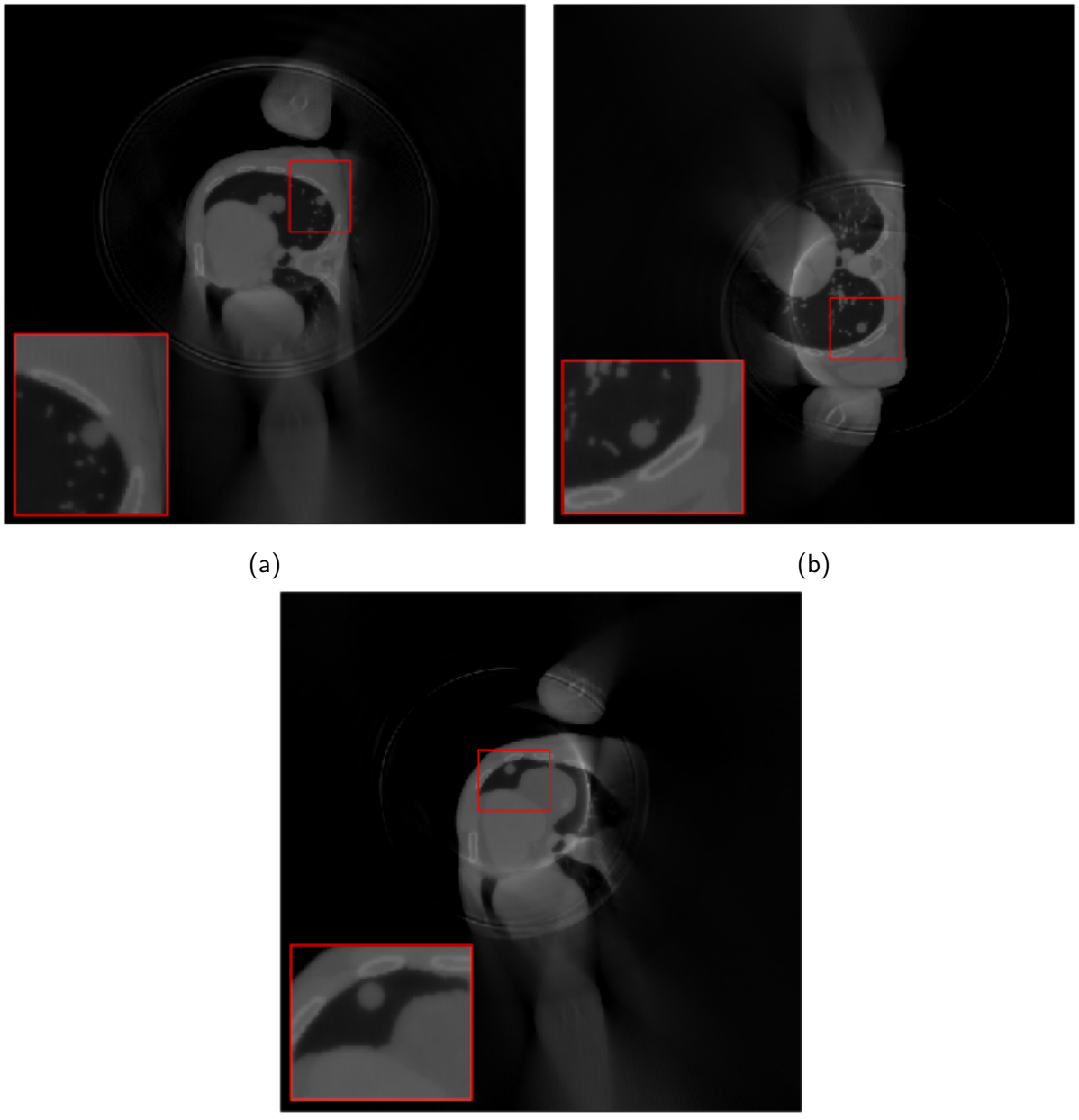
Figure 29: Three lesions in different peripheral parts of the phantom's lung. The VOIs are marked with a red rectangle.



(a) (b)

(c)

Figure 30: Reconstructed images from circular source-detector trajectories.



(a) (b)

(c)

Figure 31: Reconstructed images from optimized trajectories.

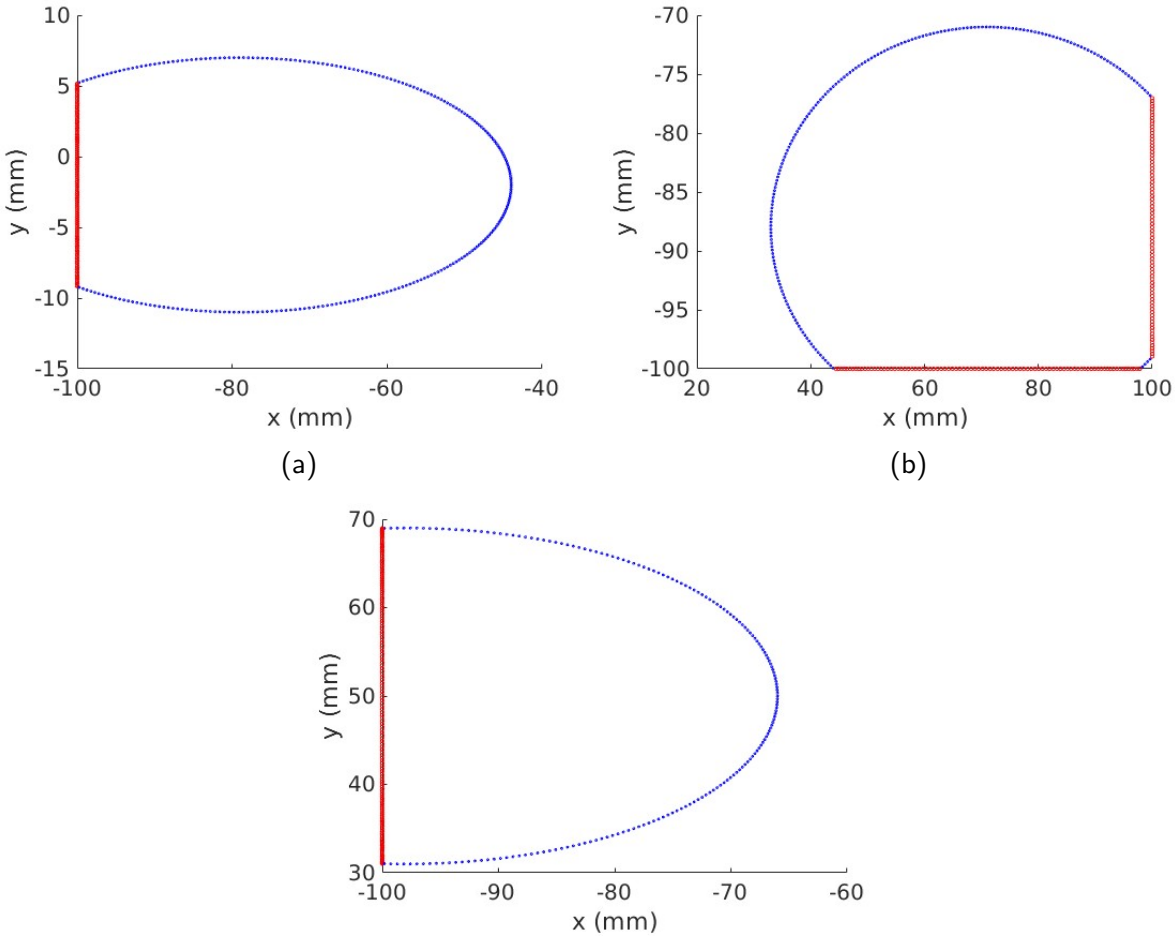
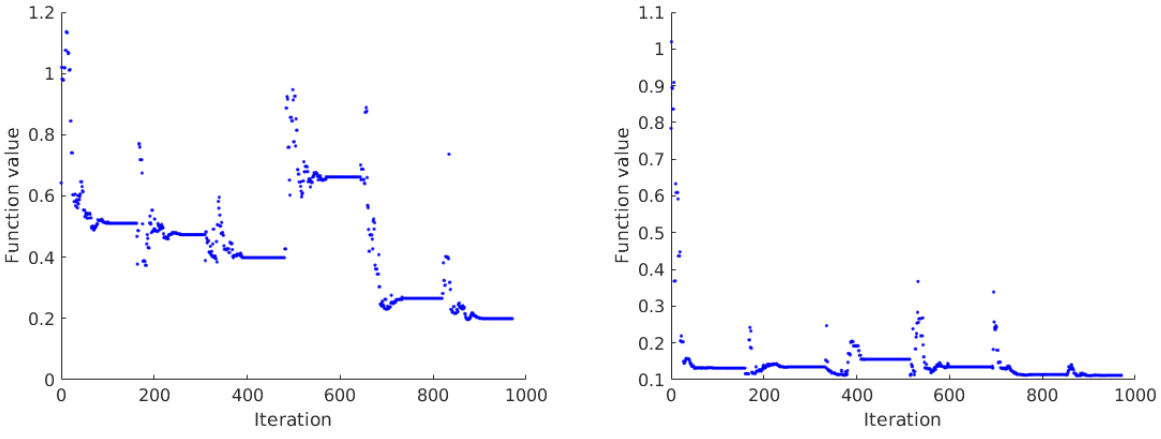
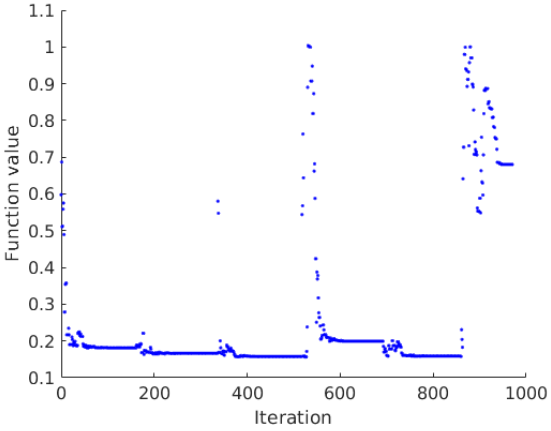


Figure 32: Optimized isocenter trajectories. The red markers indicate collision avoidance.



(a)

(b)



(c)

Figure 33: Simulated Annealing optimization processes for the three lesions. The current objective function value (1-UQI) is plotted over the course of time.

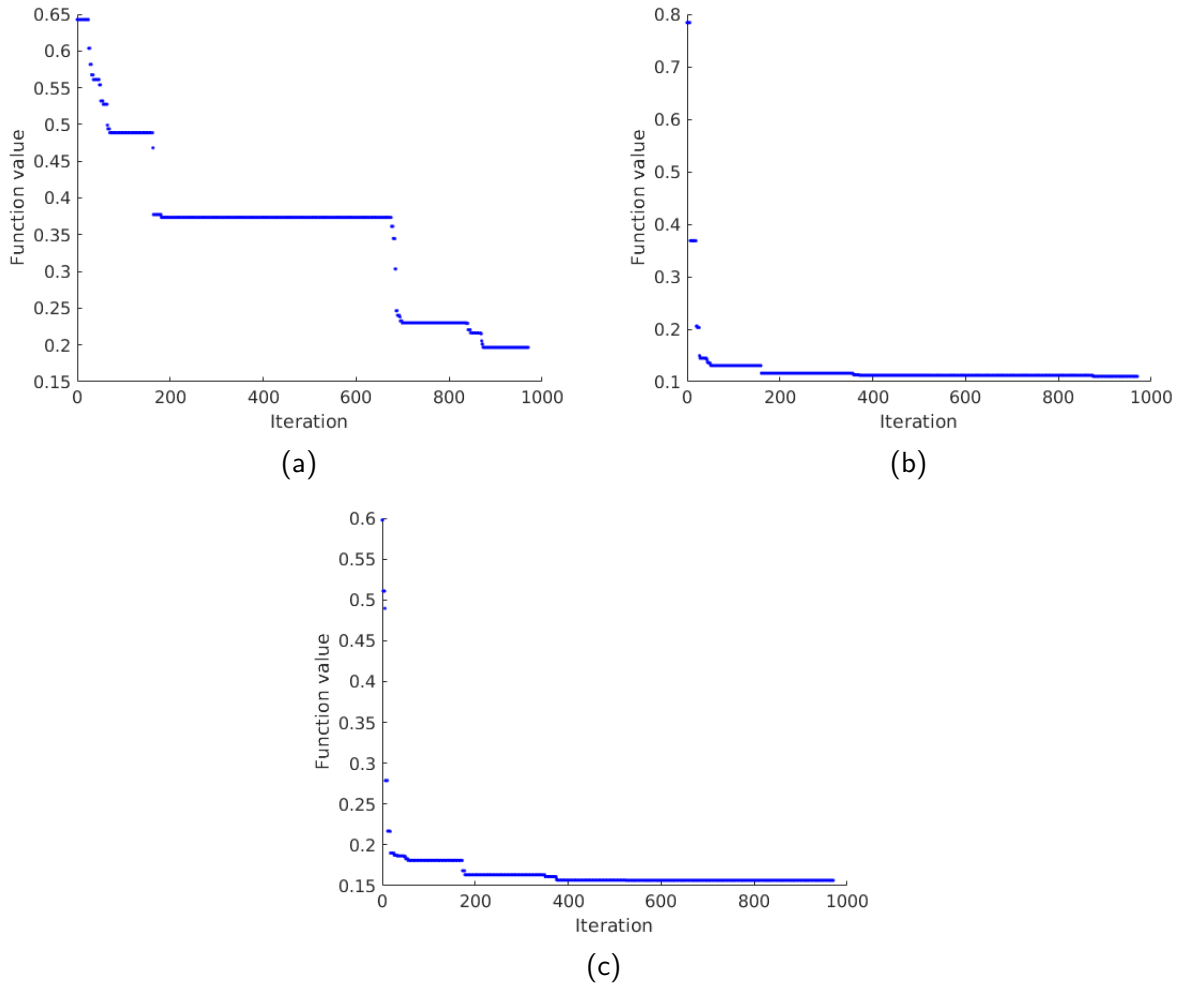


Figure 34: Best objective function value until a certain iteration number. The last point corresponds to the final UQI value.

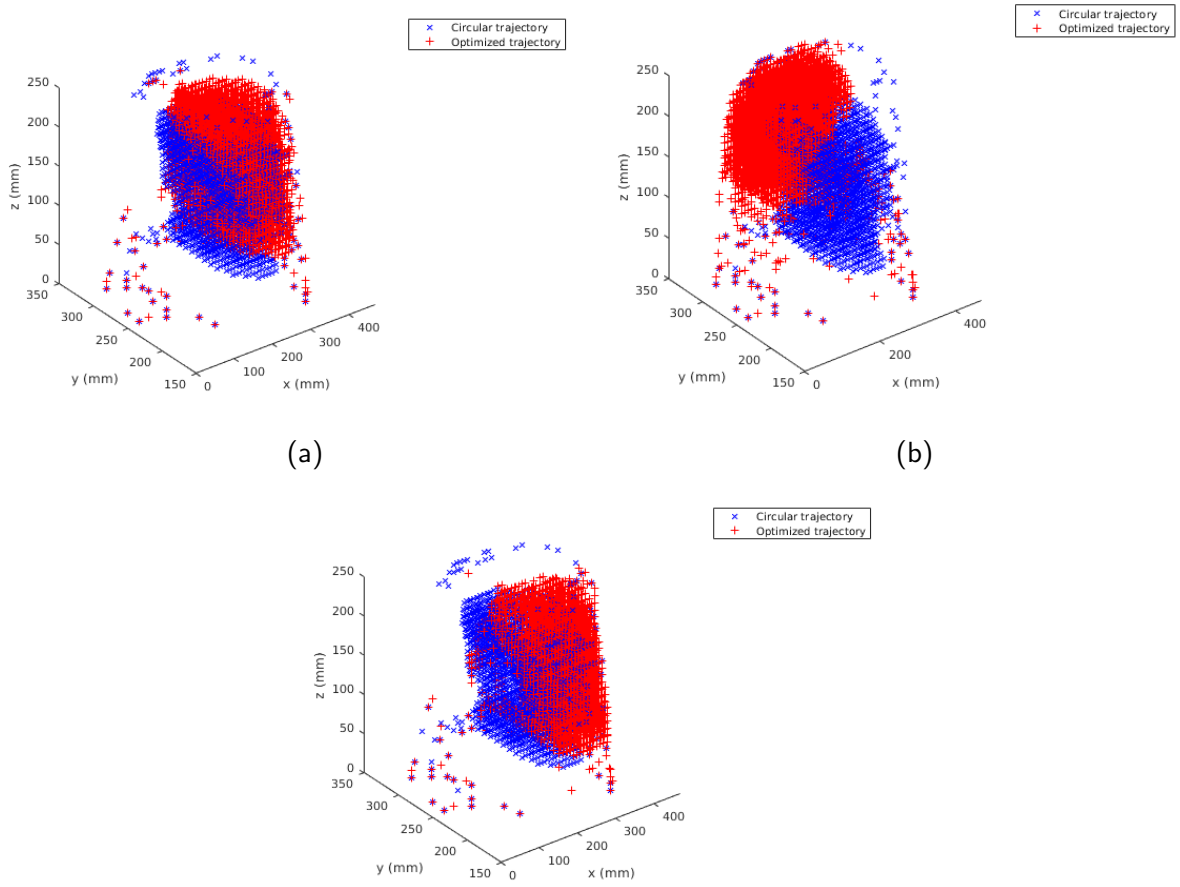


Figure 35: Illustration of FOV extension. The blue markers depict the region where an SSIM value above 0.98 is reached for standard circular trajectory. The red markers show the same for the optimized trajectory. One marker corresponds to one cube.

the UQI values in the VOIs from both the circular and the optimized reconstruction as well as the achieved FOV extensions.

		VOI 1	VOI 2	VOI 3
UQI	Circular trajectory	0.5960	0.4892	0.4798
	Optimized trajectory	0.9148	0.9681	0.9632
FOV extension		26.17%	22.59%	0.92%

Table 1: Results including UQI values achieved at three VOIs for both circular and optimized trajectories as well as FOV extension percentage.

The results from the attempt with a larger VOI are shown in Fig. 36. The corresponding numerical results are shown in Tab. 2.

		VOI
UQI	Circular trajectory	0.7179
	Optimized trajectory	0.9273
FOV extension		44.57%

Table 2: Results including UQI values achieved at the larger VOI for both circular and optimized trajectory as well as FOV extension percentage.

7.2 Projection selection optimization

The reconstructed images for the first lesion are shown in Fig. 37. Image **a** shows the reconstruction from regular spaced projections on the optimized trajectory found in Ch. 7.1. In **b** the image obtained from optimized trajectory and projection selection is shown. The corresponding trajectories are shown in Fig. 38.

The reconstructed image for the second lesion is shown in Fig. 39. The reconstruction is based on regular spaced projections on the optimized trajectory found in Ch. 7.1. No better trajectory was found during optimization. The isocenter trajectory is shown in Fig. 40.

The reconstructed images for the third lesion are shown in Fig. 41. Image **a** shows the reconstruction from regular spaced projections on the optimized trajectory found in Ch. 7.1. In **b** the image obtained from optimized trajectory and projection selection is shown. The corresponding trajectories are shown in Fig. 42.

Tab. 3 summarizes the results from the projection selection optimization. UQI values in the three VOIs from both regular spaced projections as well as optimized projections are listed. For comparison, the UQI value in the VOI obtained by using all projections for reconstruction is shown too.

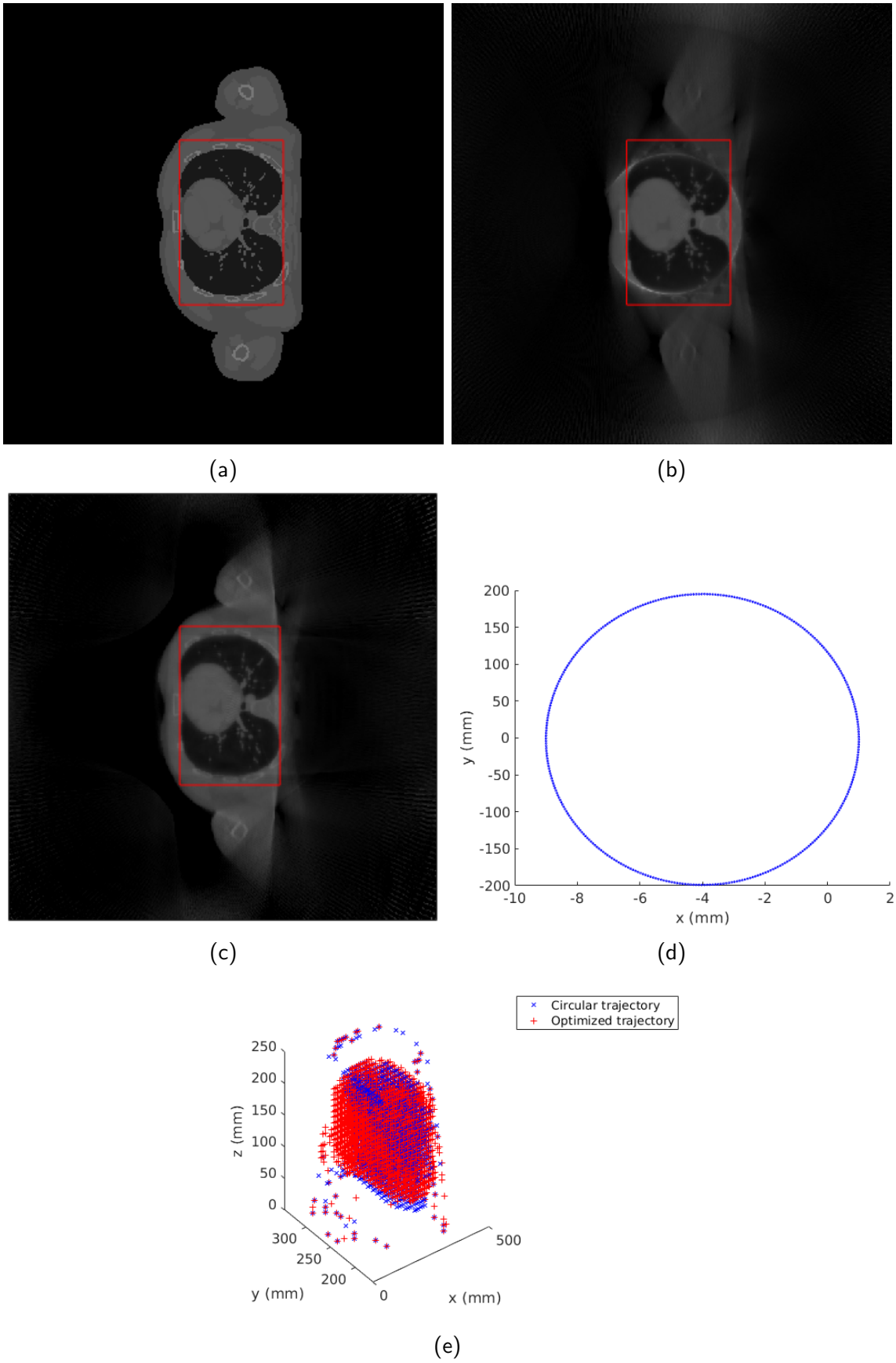
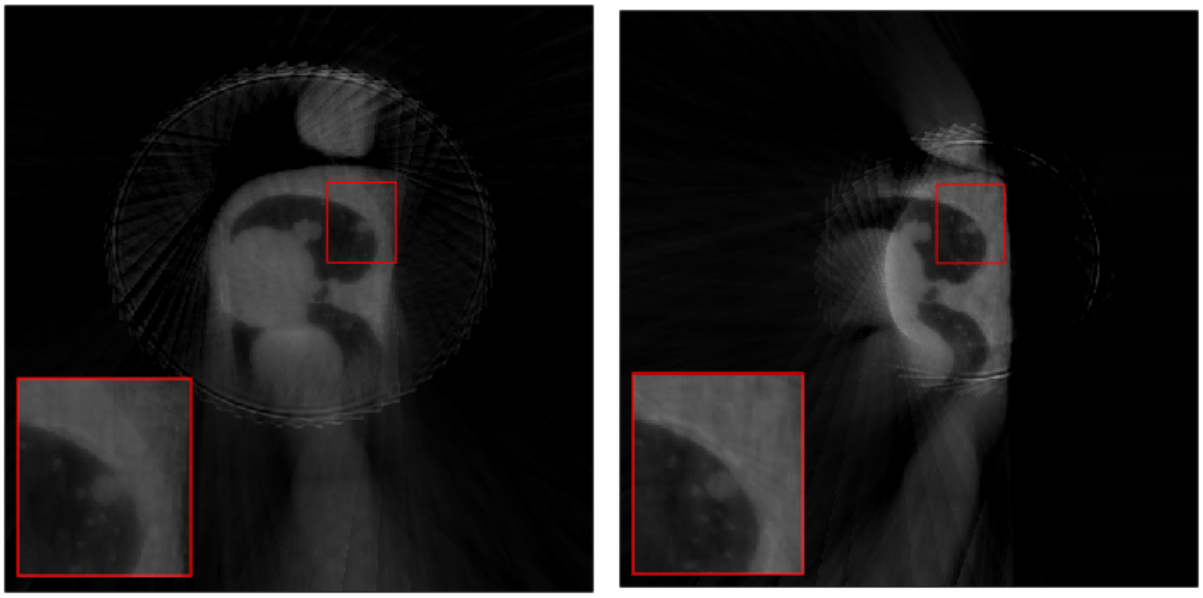
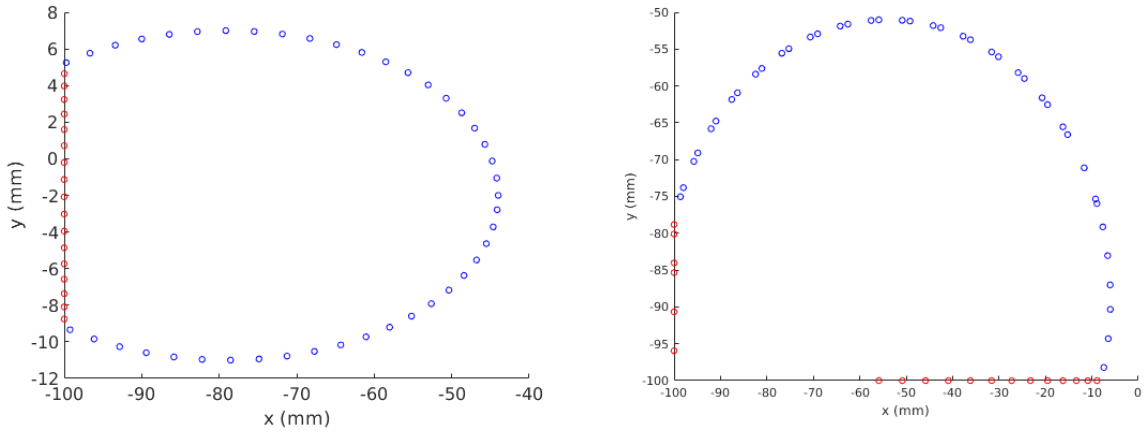


Figure 36: (a) The selected VOI in the phantom is marked with a red rectangle. (b) Reconstruction of the phantom with a standard circular trajectory. (c) Reconstruction of the phantom with the optimized trajectory. (d) Optimized elliptical trajectory for the selected VOI in the phantom. Note that the scales of the x and y axis are different. (e) Illustration of the FOV extension. The location of the cubes with a high SSIM are indicated for both circular and optimized trajectory.



(a) (b)

Figure 37: (a) Reconstructed image from regular projection spacing on the optimized trajectory found in Ch. 7.1 (b) Reconstructed image from optimized projections.



(a) (b)

Figure 38: (a) Regular projection spacing on the optimized trajectory found in Ch. 7.1 (b) Optimized trajectory and projection selection.

	VOI 1	VOI 2	VOI 3
Regular projection spacing	0.8050	0.9080	0.8805
Optimized projection spacing	0.8992	-	0.8882
All projections	0.9148	0.9681	0.9632

Table 3: Results including UQI values achieved at the three VOIs for both regular spaced projections as well as optimized projection selection. The UQI value obtained by using all projections for reconstruction is also shown for comparison.

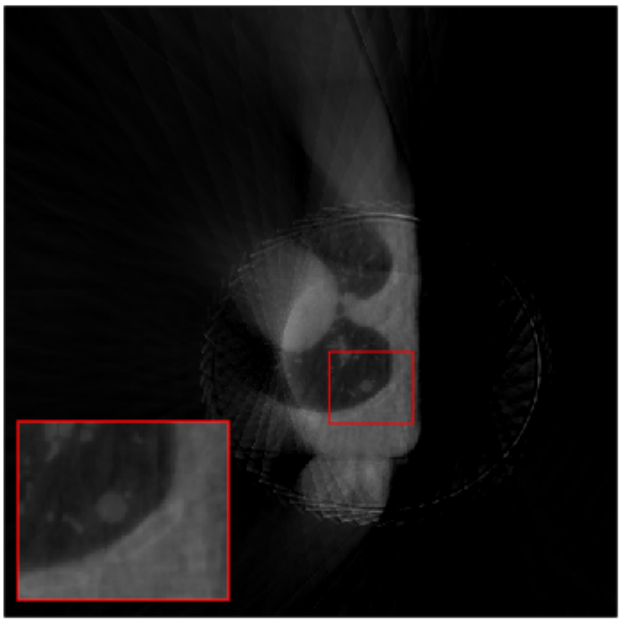


Figure 39: Reconstructed image from regular projection spacing on the optimized trajectory found in Ch. 7.1. No better reconstruction was found during optimization.

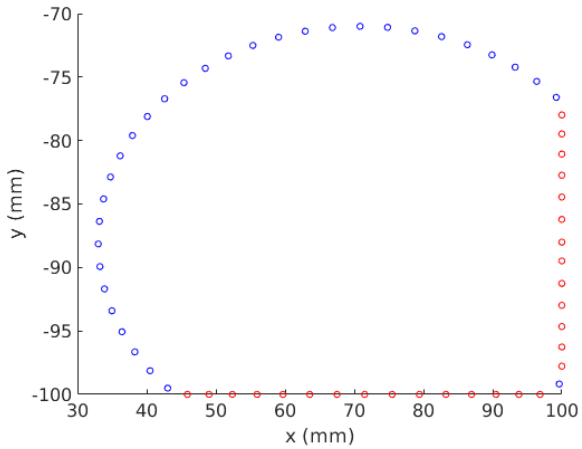
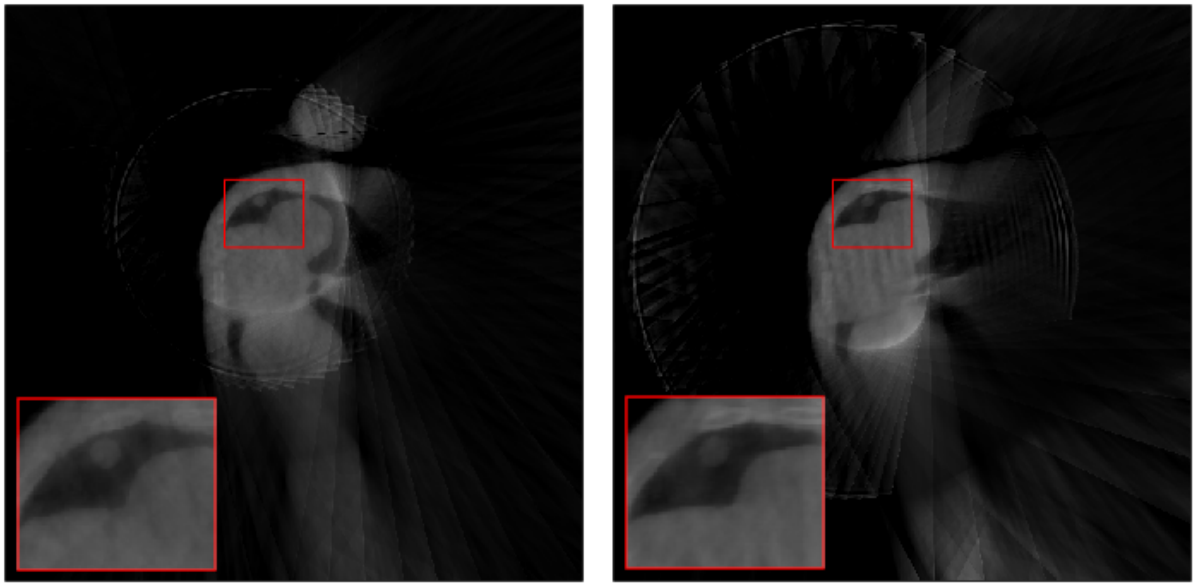
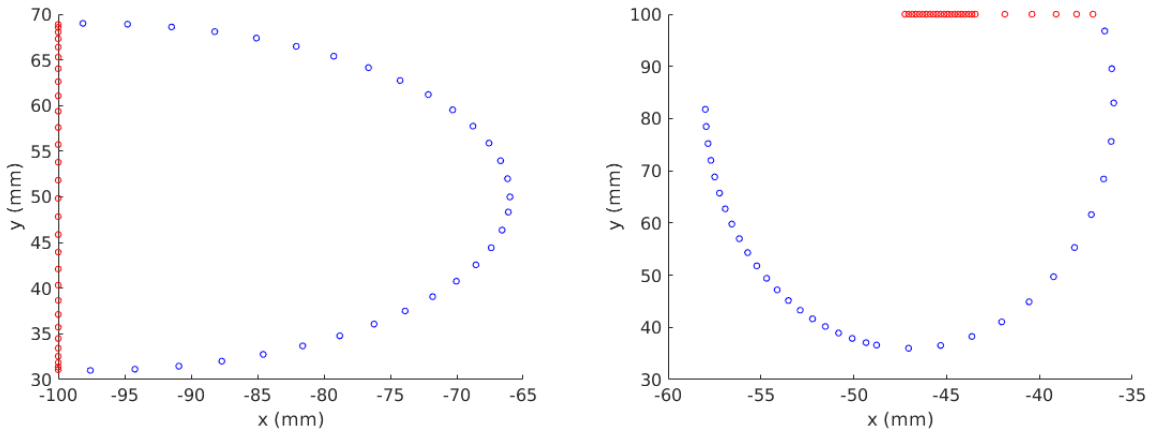


Figure 40: (a) Regular projection spacing on the optimized trajectory found in Ch. 7.1. No better trajectory was found during optimization.



(a) (b)

Figure 41: (a) Reconstructed image from regular projection spacing on the optimized trajectory found in Ch. 7.1. (b) Reconstructed image from optimized projections.

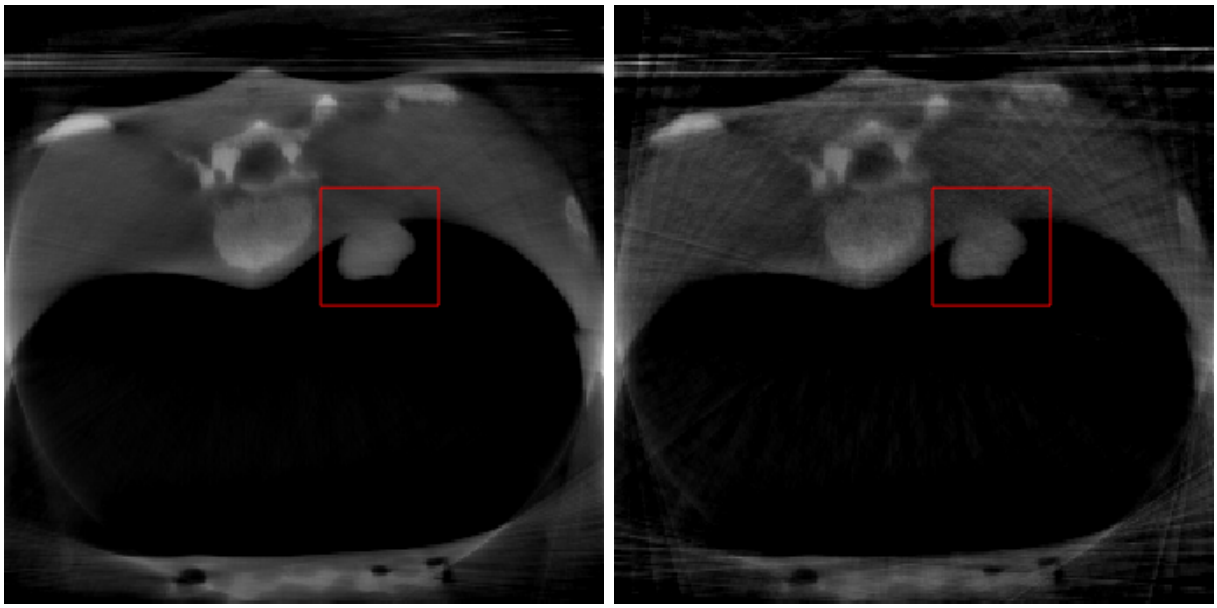


(a) (b)

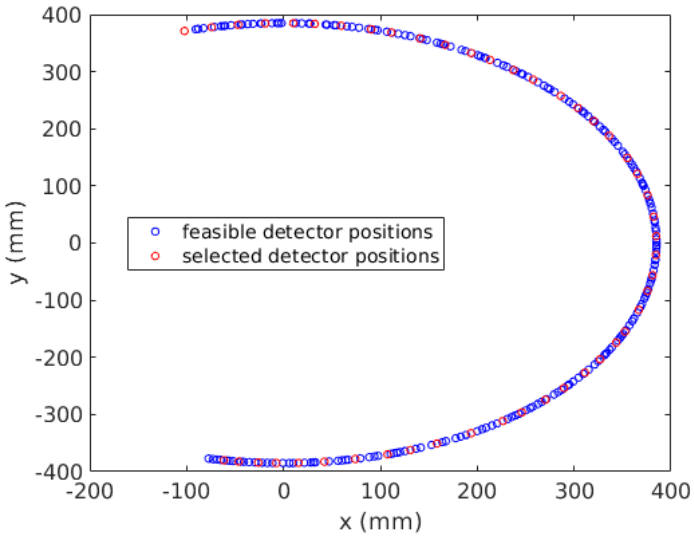
Figure 42: (a) Regular projection spacing on the optimized trajectory found in Ch. 7.1. (b) Optimized trajectory and projection selection.

7.3 Projection selection optimization on real data

For the unconstrained real data the number of desired projections is set to a sixth of the original projection number. No better projection selection than regular spaced projections was found during optimization (for both 20 and 100 SIRT iterations during optimization). The resulting image is shown in Fig. 43 **b** and the UQI value in the VOI is 0.9922. Figure 43 **c** shows the



(a) (b)



(c)

Figure 43: (a) Image obtained from the reconstruction using all projections (ground truth). The VOI (tumor) is marked with a red rectangle. (b) Reconstruction using a sixth of original projections. The projections are regularly spaced because no better projection spacing was found during optimization. (c) Detector trajectory. The red circles correspond to the utilized viewing angles. Note that the scales on the x- and y-axis are different.

corresponding trajectory and **a** shows the ground truth image, which is generated by using all projections for reconstruction.

For the constrained real data the number of desired projections is set to a third of remaining

projections. Fig. 44 shows the feasible detector positions after applying kinematic constraints in **a**, the regular spaced projections on the feasible region in **b** and the optimized detector positions, i. e. viewing angles, in **c**. Fig. 45 **a** shows the reconstructed image using all projections (ground truth). The VOI (rib) is marked with a red rectangle. The reconstructed images from the regular spaced projections and the optimized projections in the feasible region are shown in **b** and **c**.

The achieved UQI values in the VOI for both regular spaced and optimized projections are shown in Tab. 4.

		VOI
UQI	Regular spaced projections	0.9260
	Optimized projections	0.9284

Table 4: Results from the projection selection optimization using the real projection data after applying kinematic constraints. UQI values for both regular spaced and optimized projections are shown.

8 Discussion

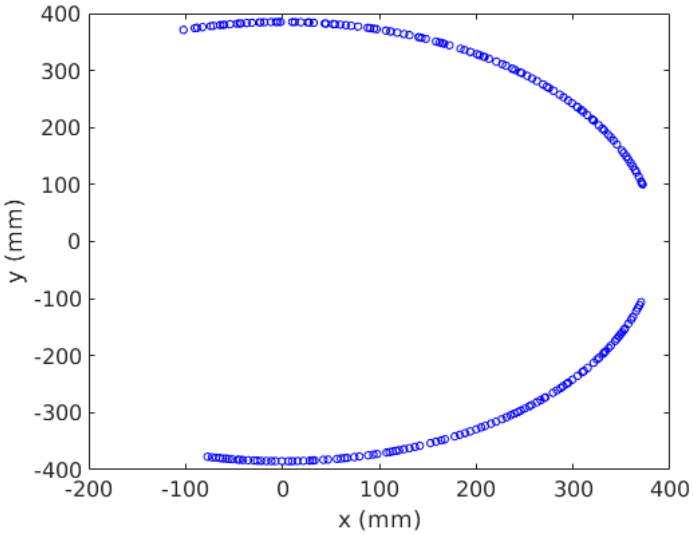
This study is a continuation of [38], where the authors proposed on the fly trajectory optimization for CBCT image acquisition. Their protocol is able to react to kinematic constraints. They selected a specific VOI from a pre-existing diagnostic CT and simulated different trajectories taking into account kinematic constraints. The best trajectory is based on an image quality metric in the VOI. However, they used mainly brute force methods for optimization.

The novel aspect represented in this study is the application of a heuristic optimization algorithm (Simulated Annealing) to CBCT trajectory optimization, with the goal of customized FOV extension, image quality improvement in a specific VOI, collision avoidance and dose reduction.

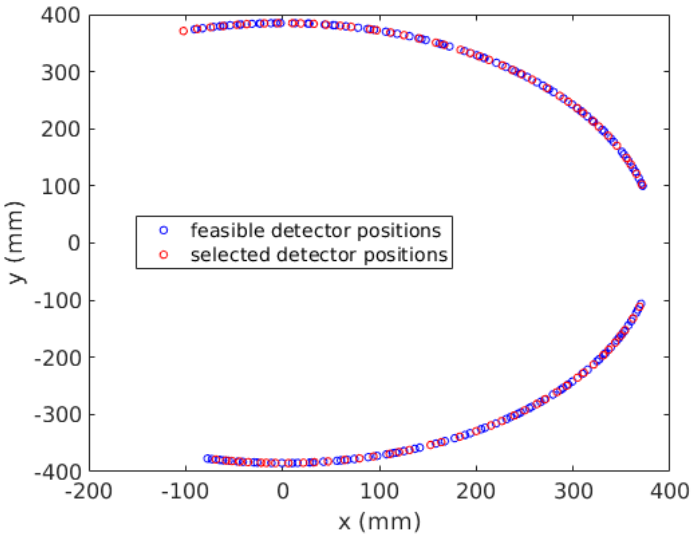
Heuristic optimization algorithms have the advantage of finding an at least near-optimal solution in a short amount of time, which makes them interesting for trajectory optimization [23]. For this study, the Simulated Annealing (SA) algorithm was utilized, however other algorithms (e. g. genetic algorithms) could be investigated as well. There are different options and parameters for Simulated Annealing, therefore it is possible that a better, more efficient optimization could be achieved by changing them.

It should be noted that the proposed approach requires a registered preoperative CT for trajectory optimization design. This can be done for example based on some initial projections and 2D/3D registration, which is a method for mapping the imaging coordinate system to the patient space [39].

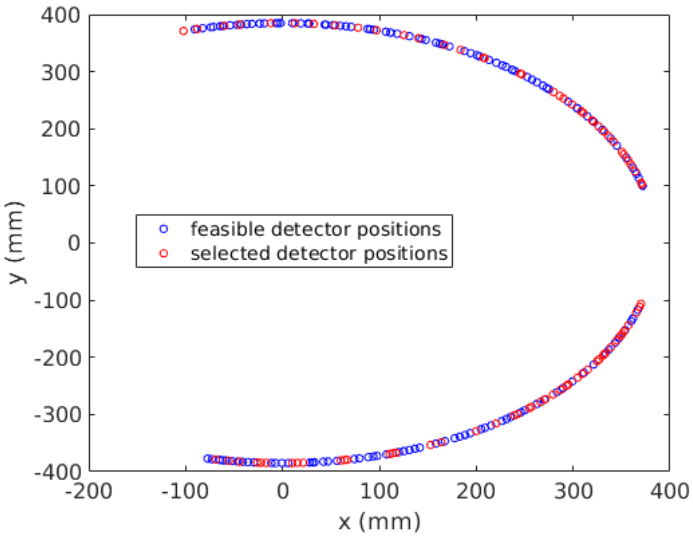
The TIGRE toolbox contains many reconstruction algorithms. SIRT algorithm was chosen because of its fast convergence. An image reconstruction is a computationally expensive process,



(a)

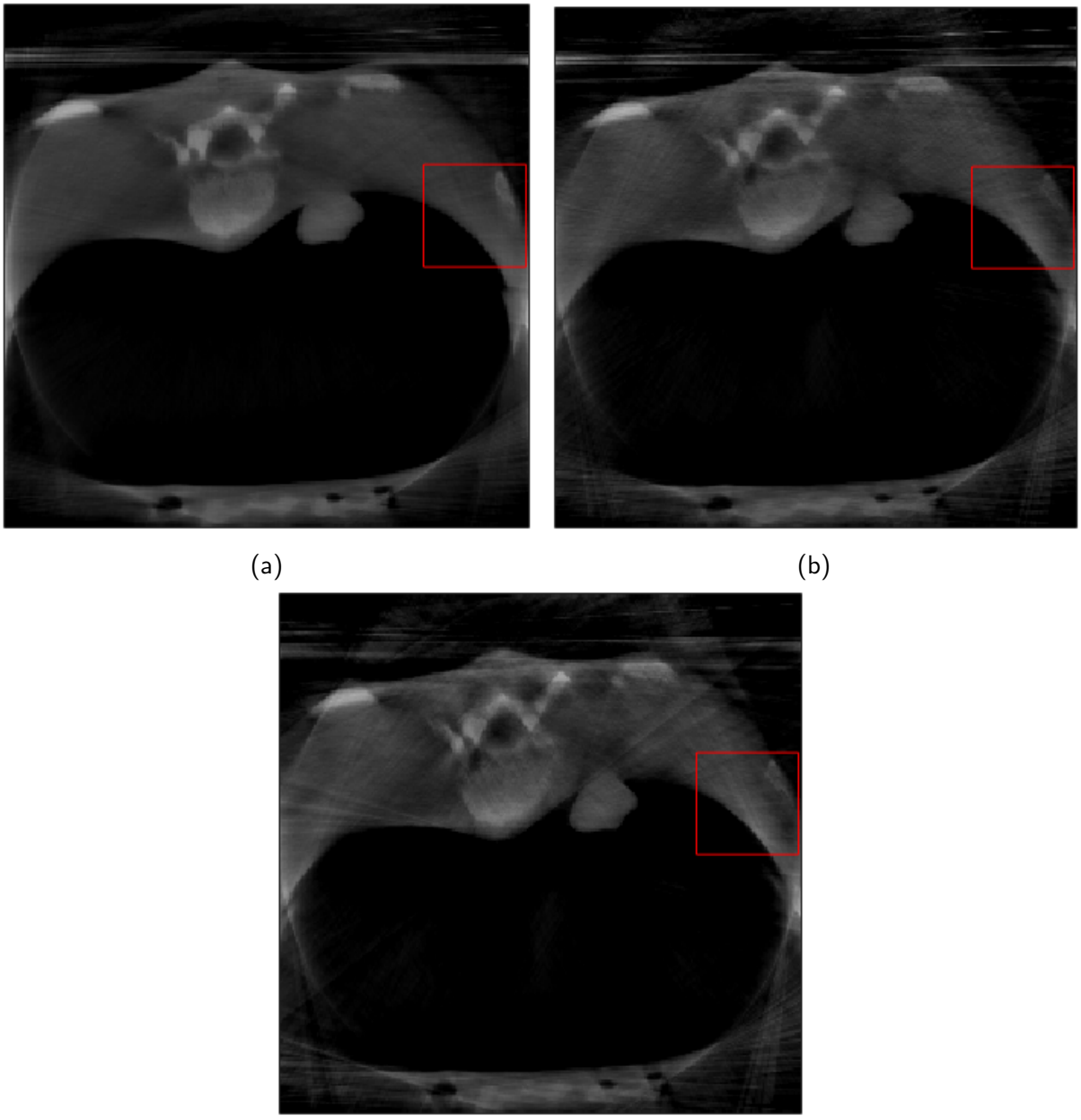


(b)



(c)

Figure 44: (a) Feasible detector positions. (b) Regular spaced projections in the feasible region. (c) Optimized viewing angles. The red circles correspond to the selected projections. Note that the scales on the x and y axis are different.



(a) (b)

(c)

Figure 45: (a) Reconstructed image using all projections (ground truth). The VOI is marked with a red rectangle. (b) Reconstructed image from regular spaced projections in the feasible region. (c) Reconstructed image from optimized viewing angles.

and doing it many times during the Simulated Annealing optimization process takes a long time in MATLAB (in the order of a few hours to a day). Some parts of TIGRE exist in C++, but they are not open source. In C++ the image reconstruction and therefore the optimization would be considerably faster. With a setup like this more sophisticated reconstruction algorithms could be used for better results.

Elliptical trajectories with kinematic constraints were chosen to extend the FOV in this study, however any other trajectory parameterization is possible and might lead to better results. The goal of the trajectories is to shift x-ray sampling to peripheral regions of the body, which would be missed or undersampled with a standard circular trajectory. For the optimization, the resolution of the ellipse parameter space was reduced by applying integer constraints, since arbitrarily small steps are not feasible. This also avoids that the algorithm gets stuck in local minima quickly. After 1000 objective function evaluations in the trajectory optimization process Simulated Annealing is told to stop. With more evaluations, even better results could be achieved. For this experiment it was decided to stop the algorithm at this point for reasons of time.

From the calculated UQI values in the selected VOIs (Tab. 1) it is clear that peripheral body parts can be reconstructed better with non standard source-detector trajectories. The truncation artefacts in the VOI from the standard circular trajectories are not present anymore. This is especially useful for larger patients, where certain body parts are not in the regular FOV. The digital phantom used for these simulation has a height of 2015 mm. Its chest has a long axis of 374 mm and a short axis of 267 mm.

The relevant part of the image for a doctor would be the volume of interest, because this is the region that was optimized. Even though the FOV was extended outside of the VOI, there are no real constraints on this area, which is why some parts can be displayed with a lower image quality than in the standard image. If a larger diagnostic image is desired the size of the VOI has to be increased, which is attempted here too. Nevertheless it was shown that in principle elliptical trajectories lead to a FOV extension and can be used for further studies. The artefacts in Fig. 31 are truncation artefacts.

For the larger VOI, the lung is reconstructed without artefacts. The achieved UQI value is lower than the average optimized UQI value for the smaller VOIs. The reason is probably that some x-ray sampling is shifted to the periphery of the phantom, which results in less sampling in the center. Therefore the overall image quality is less than when sampling only a small VOI. The calculated VOI extension is notable as well, which suggests that the proposed approach works for different VOI sizes.

For the calculation of the FOV the SSIM image quality metric was used instead of the UQI. The reason was that the cubes where a high SSIM value was achieved seemed to correlate better with the regions that were visually well reconstructed. Therefore this method provided a better estimation for the FOV, than when doing the same with the UQI. It is possible that the SSIM

would also be a better fit for the objective function than UQI, although the obtained optimization results (using UQI) look promising.

Fig. 35 shows that the region which is well reconstructed in the optimized images is shifted compared to the standard images. It shows the locations of the cubes with an SSIM value higher than the selected threshold. By comparing the number of high SSIM cubes in the optimized and standard image a measure for the FOV extension was approximated. In the thesis only 2D slices of the phantom are shown, however this value accounts for the FOV extension in the whole 3D phantom. The achieved FOV extension is not really targeted. The objective function is only based on the image quality in the VOI, which is indeed enhanced, however for the remaining phantom parts there is no condition on the image quality.

For the third lesion no notable FOV extension was calculated. As already mentioned, since there are no real conditions on the image quality in regions outside the VOI it makes sense, that the proposed method does not always lead to a considerable FOV extension. However, for the other two lesions a notable FOV extension was achieved. It should be mentioned that the calculated values are only an approximation and depend both on the selected cube size and SSIM threshold.

One limitation for CBCT in interventional procedures is the limited available space in the surgery setup. Larger patients can hinder the system from moving freely too. In this thesis kinematic constraints were considered by only allowing isocenter movements within a certain range and thus providing a more realistic scenario. Several techniques for collision detection have already been investigated in radiotherapy, including computer-aided design (CAD) systems and optical detection based on a laser camera [22]. One possibility for future studies would be to incorporate those into the trajectory optimization process.

As mentioned before, the number of projections necessary to properly reconstruct a CBCT image is high and delivers a substantial radiation dose to the patient. Since repetitive CBCT scans are required for image-guided procedures and daily pretreatment patient alignment images are required in radiation therapy, accumulated dose has become a concern. It is desirable to reduce patient dose, while still maintaining a diagnostic image quality. Different approaches have been investigated to reduce the dose, including filters, optimizing scan parameters, using statistical reconstruction and projection reduction. The latter has been investigated in this thesis. The idea was to keep only those projections, which offer the highest information content about a specific VOI [22]. The most intuitive way of selecting for example only half of projections is by just keeping every other projection. Therefore the goal was to achieve an image quality higher than it would be achieved with regular projection spacing. The heuristic optimization algorithm was again utilized to achieve the best selection of projections, based on the parameterization explained in Ch. 6.2.

For the first lesion optimized projections lead to a better reconstruction of the right corner

of the VOI, compared to the regular spaced projections. Furthermore the rib and the boundary between lung and surrounding tissue are more recognizable in the optimized image. Even though it is visually arguable if the bright spots in the lung were reconstructed better, the overall image quality has improved and the image seems smoother.

For the second lesion no better set of projections (than regular spaced projections) was found during the optimization. With more SA iterations one could maybe still achieve a better trajectory. However, the visual image quality is already comparable to the optimized images of lesion 1 and 3, so in this case the trajectory found before apparently already was a good fit for this task.

The image reconstruction from optimized projections for the third lesion in general looks smoother than the reconstruction from regular spaced projections. Furthermore the rib and the boundary between lung and surrounding tissue are again reconstructed better. Therefore the overall image quality was improved by the proposed method.

Image quality variation over different regions is common in circular CBCT. Reasons are different location, different attenuation path of the x-rays and different voxel values. These effects are still present in optimized trajectories, which is probably the reason for different relative UQI enhancements compared to the initial trajectories [38].

As a general note, image quality metrics attempt to make statements about the image quality. However, they do not necessarily always align with our visual perception. It is not always straightforward why certain images have a specific image quality metric. The applied image quality metrics in this case were UQI and SSIM and both did not always correlate ideally with visual perception. The idea of making statements about the image quality solely based on an image quality metric can in some cases be problematic. With a better measure for image quality, closer to human vision, the results could be improved further.

Another possible continuation of this study is reducing metal artefacts. If the imaging target is close to metal objects like implants, surgical tools, needles or other radiopaque structures, the reconstructed image can suffer from poor image quality. Artefacts arise from discrepancies between the actual physical processes in image formation and the assumed reconstruction model. Trajectories could be optimized by heuristic optimization algorithms in order to avoid projections, which would deteriorate the image quality near the metal object [22].

Finally, an attempt was made to optimize selected projections from a set of real CBCT projections that was acquired with a circular source-detector trajectory. The projection selection optimization using 20 SIRT iterations during optimization could not find a better image reconstruction than the one achieved by regular projection spacing. The main problem is probably that there are no isocenter shifts in the projection data, which limits the possibilities for optimization greatly. Because the possible improvement in image quality is therefore subtle it was decided to already carry out 100 SIRT iterations during the optimization, instead of doing only 20 SIRT

iterations. The quality of the reconstructed image rises with the number of SIRT iterations, therefore the algorithm might be more likely to find a better projection selection in this way. However, this increases the computational time, which is why the allowed objective function evaluations were restricted to 300. Also, for real data (including all physical effects) it might be more important to already have a better reconstruction during optimization in order to find the best set of projections. However, no better projection selection than regular spaced projections was found also with this setup.

Because no better projection selection was found for the real data, some random kinematic constraints were applied by not allowing specific projection angles, namely angles in the range between $\pm 15^{\circ}$. As VOI a peripheral part of the phantom, including a rib, was chosen. 100 SIRT iterations were again carried out during optimization for a more accurate result. The achieved improvement after 300 SA objective function evaluations is small, but noticeable. As mentioned before, the data is not ideal, because there are no isocenter shifts in the recorded projections. With better data, which was not feasible within this project, the achieved improvement would probably be more significant.

It should be noted that within this thesis only simulations were done and no experiments have been conducted to see if the proposed approach works in a clinical setup. If not, adjusting parameters or changing the optimization approach could still lead to an improvement in the clinical setup as well. Simulated results for UQI improvement, FOV extension (Tab. 1 and Tab. 2) and projection selection (Tab. 3 and Tab. 4) seem promising and should be investigated further.

As already mentioned, there is no experimental validation for the proposed approach yet because it did not align with the time frame of this thesis. However, a procedure that utilizes the proposed optimization method could look like this: A prior CT of the patient anatomy has to be available. On this CT image trajectories can be simulated in order to optimize the image quality in a certain region, i. e. the VOI. The reconstruction from the simulated projections should already have a high quality during optimization, which is why it is important to make the reconstruction faster. This means implementing the proposed method in C++, otherwise it is not helpful in a clinical environment. The found trajectory can then be implemented on the real device when imaging the patient with a CBCT. The geometries of the real CBCT device and the simulated device have to be identical.

9 Conclusion

This study shows that the Simulated Annealing algorithm can efficiently be used for optimizing source-detector trajectories in CBCT imaging. To be more precise, non-isocentric trajectories optimized by Simulated Annealing enable imaging of peripheral or large predefined volume of interests, which would not be fully reconstructed with the standard circular trajectory. Furthermore those non-isocentric trajectories led to a notable field of view extension in three out of four

cases and took into account available kinematic constraints. The proposed projection selection optimization method using Simulated Annealing provides the opportunity to achieve better image quality than achieved by regular spaced projections. The approach was also tested on real projection data. Since the available data did not include isocenter shifts, it was not ideal for projection selection optimization. However, a subtle improvement could still be achieved using the proposed method. In order to apply Simulated Annealing in a clinical setup, more research has to be done, especially towards making the optimization process faster. To sum up, this thesis deals with common problems in CBCT like FOV extension, image quality improvement, collision avoidance and dose reduction using a heuristic optimization algorithm. It presents an approach to personalizing CBCT imaging and thus improving patient care.

References

[1] Hendee, W. R. *Medical imaging physics* (Wiley-Liss, New York, NY, 2002), 4. edn.

[2] Pietzsch, J. Perspectives: A helping hand from the media. URL `https://www.nobelprize.org/prizes/physics/1901/perspectives/`. Accessed: 2023-11-24.

[3] Těšínská, E. Johann Puluj (1845–1918): his career and the "invisible cathode rays" (2020). URL `https://www.iucr.org/news/newsletter/volume-28/number-2/physicist-and-electrician-of-ukrainian-roots-johann-puluj-18451918-his-professional-career-and-the-invisible-cathode-rays`. Accessed: 2023-11-24.

[4] Van Tiggelen, R. The rise of contrast-enhanced roentgenology: An illustrated and chronological overview. *JBR-BTR (Bruxelles)* **100**, 102–102 (2016).

[5] Birkfellner, W. *Applied Medical Image Processing: A Basic Course* (CRC Press, 2016). URL `https://books.google.at/books?id=iVjOBQAAQBAJ`.

[6] Schlegel, W., Karger, C. P. & Jäkel, O. *Medizinische Physik : Grundlagen – Bildgebung – Therapie – Technik* (Springer Berlin Heidelberg Imprint: Springer Spektrum, Berlin Heidelberg, 2018).

[7] Rosenberg, I. Radiation oncology physics: A handbook for teachers and students. *British Journal of Cancer* **98**, 1020–1020 (2008).

[8] Boone, J. M. & McCollough, C. H. Computed tomography turns 50. *Physics today* **74**, 34–40 (2021).

[9] Zanzonico, P. Physics in nuclear medicine. *Medical Physics* **41** (2014).

[10] Flohr, T. Ct systems. *Current Radiology Reports* 52–63 (2013).

[11] Hounsfield, G. N. Computerized transverse axial scanning (tomography): Part 1. description of system. *The British Journal of Radiology* **46** (1973).

[12] Fareed, A. *et al.* Impact of iterative reconstruction vs. filtered back projection on image quality in 320-slice ct coronary angiography: Insights from the core320 multicenter study. *Medicine (Baltimore)* **96**, e8452–e8452 (2017).

[13] Zeng, G. *Medical image reconstruction: A conceptual tutorial* (2010).

[14] Gregor, J. & Benson, T. Computational analysis and improvement of sirt. *IEEE Transactions on Medical Imaging* **27**, 918–924 (2008).

[15] Miracle, A. & Mukherji, S. Conebeam ct of the head and neck, part 1: Physical principles. *American journal of neuroradiology : AJNR* **30**, 1088–1095 (2009).

[16] Scarfe, W. C. & Farman, A. G. What is cone-beam ct and how does it work? *Dental Clinics of North America* **52**, 707–730 (2008). URL `https://www.sciencedirect.com/science/article/pii/S001185320800044X`. Accessed: 2023-12-09.

[17] Hatamikia, S. *Patient specific source-detector trajectory optimization for Cone Beam Computed Tomography*. Dissertation, Medical University of Vienna (2020).

[18] Bushberg, J. T., Seibert, J. A., Leidholdt, E. M. & Boone, J. M. *The essential physics of medical imaging* (Lippincott, Williams Wilkins, Philadelphia, 2002).

[19] UVA Health. What is interventional radiology? (2019). URL `https://blog.radiology.virginia.edu/interventional-radiologist-definition/`. Accessed: 2023-10-04.

[20] Schulze, R. *et al.* Artefacts in cbct: a review. *Dento-maxillo-facial radiology* **40**, 265–273 (2011).

[21] Zeng, G. *Medical image reconstruction: A conceptual tutorial* (2010).

[22] Hatamikia, S. *et al.* Source-detector trajectory optimization in cone-beam computed tomography: a comprehensive review on today's state-of-the-art. *Physics in medicine biology* **67**, 16 (2022).

[23] Wang, F.-S. & Chen, L.-H. *Heuristic Optimization*, 885–885 (Springer New York, New York, NY, 2013).

[24] Talbi, E.-G. *Metaheuristics: from design to implementation* (Wiley, Hoboken, NJ, 2009).

[25] MathWorks. What is the genetic algorithm. https://de.mathworks.com/help/gads/what-is-the-genetic-algorithm.html. Accessed: 2023-07-06.

[26] *Numerical recipes in C: the art of scientific computing* (Cambridge Univ. Press, Cambridge, 2002), 2. edn.

[27] Biguri, A., Dosanjh, M., Hancock, S. & Soleimani, M. Tigre: a matlab-gpu toolbox for cbct image reconstruction. *Biomedical Physics & Engineering Express* **2**, 055010 (2016).

[28] Inc., T. M. Global optimization toolbox (r2023a) (2023). URL `https://www.mathworks.com`. Accessed: 2023-12-09.

[29] Hosseinzadeh, S. Integer/discrete optimization with simulated annealing (2019). URL `https://de.mathworks.com/matlabcentral/fileexchange/72539-integer-discrete-optimization-with-simulated-annealing`. Accessed: 2023-12-09.

[30] MathWorks. Simulated annealing options. https://de.mathworks.com/help/gads/simulated-annealing-options.html. Accessed: 2023-11-13.

[31] MathWorks. Simulated annealing. https://de.mathworks.com/discovery/simulated-annealing.html. Accessed: 2023-07-06.

[32] MathWorks. How simulated annealing works. https://de.mathworks.com/help/gads/how-simulated-annealing-works.html. Accessed: 2023-07-06.

[33] Segars, W. P., Sturgeon, G., Mendonca, S., Grimes, J. & Tsui, B. M. W. 4d xcat phantom for multimodality imaging research. *Medical physics (Lancaster)* **37**, 4902–4915 (2010).

[34] The MathWorks Inc. Matlab version: 9.14.0 (r2023a) (2023). URL `https://www.mathworks.com`. Accessed: 2023-12-09.

[35] NVIDIA, Vingelmann, P. & Fitzek, F. H. Cuda, release: 11.8 (2022). URL `https://developer.nvidia.com/cuda-toolkit`. Accessed: 2023-12-09.

[36] Wang, Z. & Bovik, A. A universal image quality index. *IEEE signal processing letters* **9**, 81–84 (2002).

[37] Wang, Z., Bovik, A., Sheikh, H. & Simoncelli, E. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* **13**, 600–612 (2004).

[38] Hatamikia, S. *et al.* Toward on-the-fly trajectory optimization for c-arm cbct under strong kinematic constraints. *PloS one* **16**, e0245508–e0245508 (2021).

[39] Birkfellner, W. *et al.* A faster method for 3d/2d medical image registration—a simulation study. *Physics in Medicine Biology* **48**, 2665 (2003). URL `https://dx.doi.org/10.1088/0031-9155/48/16/307`.