Universiteit
Antwerpen

Faculteit Wetenschappen
Departement Fysica

Model-based iterative reconstruction algorithms for computed tomography

---

Modelgebaseerde iteratieve reconstructiealgoritmes voor computertomografie

Proefschrift voorgelegd tot het behalen van de graad van

Doctor in de Wetenschappen

aan de Universiteit Antwerpen, te verdedigen door

Geert VAN EYNDHOVEN

Promotoren:
Prof. dr. Jan Sijbers
Prof. dr. K. Joost Batenburg

Antwerpen, 2015

Doctoral committee

Prof. dr. Sara Bals
Prof. dr. K. Joost Batenburg
Prof. dr. Jan Sijbers
Prof. dr. Wim Van Roose

Other jury members

Dr. Anders Kaestner
Prof. dr. Johan Nuyts

Contact information:

✉ **Geert Van Eyndhoven**
iMinds-Vision Lab, Dept. of Physics
University of Antwerp (CDE)
Universiteitsplein 1, Building N1.11
B-2610 Wilrijk, Antwerpen, Belgium
☎ +32(0)3 265 24 49
📱 +32 (0)494 75 37 74
📧 `vegeert@hotmail.com`
🌍 `http://visielab.uantwerpen.be/people/geert-van-eyndhoven`

Table of Contents

Samenvatting

Computertomografie (CT) is een niet-invasieve beeldvormingsmethode waarbij het inwendige van een object gevisualiseerd wordt uitgaande van een aantal X-stralenfoto's die vanuit verschillende hoeken werden opgenomen. De techniek kan toegepast worden bij verschillende modaliteiten, gaande van een typische X-stralen CT scanner tot elektronenmicroscopie en synchrotrons.

Vaak zijn er slechts weinig X-stralenfoto's beschikbaar, waardoor de beeldreconstructie bemoeilijkt wordt en waardoor artefacten verschijnen in het CT beeld. Beperkte data problemen komen voor in verschillende CT toepassingen. In medische CT is het opportuun om slechts een beperkt aantal projecties op te nemen om de door de patiënt geabsorbeerde stralingsdosis zo laag mogelijk te houden. In elektronenmicroscopie kan het monster vaak slechts over een beperkt hoekbereik geroteerd worden wegens mechanische beperkingen en is het aantal projectiehoeken vaak beperkt om schade aan het monster (veroorzaakt door de elektronenstraal) te vermijden. Bij dynamische CT is het tijdsinterval om voldoende projectiedata over het volledige angulaire bereik op te nemen vaak groot in vergelijking met het tijdsinterval waarin substantiële veranderingen plaatsvinden binnen het gescande object. Dit impliceert dat, om onscherpe reconstructies te vermijden, slechts een beperkt aantal projecties opgenomen kunnen worden per tijdspunt, waardoor er een reconstructieprobleem met beperkte data ontstaat.

In deze thesis worden verschillende verbeterde CT beeldreconstructiealgoritmes voorgesteld, allen met het doel nauwkeurige reconstructies te genereren op basis van slechts een beperkt aantal X-stralenfoto's. De verbetering wordt bereikt door gebruik te maken van lokale modellen van het te scannen object. Een betere modellering van het object en/of reconstructieprobleem resulteert in een verbeterde kwaliteit van het CT beeld. In het eerste deel van deze thesis, bestaande uit enkel hoofdstuk 2, wordt deze aanpak geïmplementeerd voor stationaire poreuze objecten. In het tweede deel verandert de focus naar dynamische CT, waarbij de voorgestelde aanpak toegepast wordt op structureel veranderende objecten, tomografie van vloeistofstromen en perfusie CT, in respectievelijk hoofdstuk 3, 4 en 5. Het derde deel bevat de algemene conclusies en appendices. In wat volgt zullen de verschillende hoofdstukken verder toegelicht worden.

Hoofdstuk 1 – Introductie

In dit hoofdstuk wordt de lezer bekend gemaakt met de noodzakelijke technische achtergrond en de notatie die gebruikt wordt doorheen de thesis. Zowel stationaire als dynamische CT komen aan bod.

Deel I: Stationaire CT

Hoofdstuk 2 – Reconstructiemethode voor verbeterde porositeit kwantificatie van poreuze materialen

In dit hoofdstuk wordt een op voorkennis gebaseerde reconstructiemethode geïntroduceerd voor poreuze materialen. Het algoritme wordt gevalideerd in de context van elektronenmicroscopie, een vaak gebruikte beeldvormingstechniek voor onderzoek naar het verband tussen de structuur en de eigenschappen van nanomaterialen.

Met een conventionele werkwijze is het kwantitatief interpreteren van elektronentomografieresultaten niet eenvoudig. In het bijzonder wordt de accurate kwantificatie van de poreuze ruimte bemoeilijkt door artefacten die geïntroduceerd worden in zowat elke stap van de conventionele werkwijze, i.e., acquisitie, reconstructie, segmentatie en kwantificatie. Bovendien vereisen de meest voorkomende conventionele werkwijzen subjectieve input van de gebruiker.

In dit hoofdstuk wordt het PORES algoritme ("*POre REconstruction and Segmentation*") geïntroduceerd, het is een op maat gemaakte, integrale aanpak voor de reconstructie, segmentatie en kwantificatie van poreuze materialen aan de hand van CT. In de PORES werkwijze wordt een reconstructie berekend door simultaan de poriën te classificeren en tegelijkertijd de overblijvende regio te reconstrueren door de fout ten opzichte van de opgemeten projectiedata te minimaliseren. Deze reconstructie kan dan onmiddellijk gebruikt worden in de overblijvende stappen van het PORES algoritme, wat resulteert in nauwkeurige individuele kwantificatie van de poriën en accurate statistieken over de porieverdeling binnen het volledige sample. Het voorgestelde algoritme werd uitgebreid gevalideerd met zowel gesimuleerde als opgemeten experimentele data.

Deel II: Dynamische CT

Hoofdstuk 3 – Regiogebaseerde iteratieve reconstructie

In dit hoofdstuk wordt een reconstructiemethode voorgesteld voor lokaal in de tijd veranderende objecten, i.e., objecten die dynamische veranderingen ondergaan in slechts een lokale regio binnen het object. De focus is in dit hoofdstuk gericht op structureel veranderende objecten.

In tegenstelling tot reconstructiemethodes voor continu veranderende objecten, waar de continuïteit tussen reconstructies op verschillende tijdspunten opgelegd kan worden via een deformatiemodel of regularisatie, berekenen huidige methodes voor het reconstrueren van discreet of structureel veranderende objecten doorgaans onafhankelijke reconstructies op de verschillende tijdspunten. De kwaliteit van zulke reconstructies is gelimiteerd door het feit dat er slechts een beperkt aantal projecties beschikbaar zijn per tijdspunt.

Vertrekkende vanuit de observatie dat er in het gescande object regio's bestaan die onveranderd blijven doorheen de tijd, wordt in dit hoofdstuk een reconstructiemethode voorgesteld die deze regio's automatisch bepaalt en deze kennis gebruikt in een algebraïsche reconstructiemethode. Het voorgestelde algoritme werd gevalideerd met gesimuleerde data en experimentele μCT-data. Uit de validatie-experimenten blijkt de geschiktheid van het voorgestelde algoritme om structureel veranderende objecten preciezer te reconstrueren in vergelijking met huidige technieken.

De in dit hoofdstuk voorgestelde techniek vormt de basis voor de dynamische reconstructiealgoritmes uit hoofdstuk 4 en 5.

Hoofdstuk 4 – Een iteratief CT reconstructiealgoritme voor het beeldvormen van snelle vloeistofstromen

Het onderzoek van vloeistofstromen in vaste media door middel van CT-beeldvorming heeft vele toepassingen, onder andere in de olie-industrie en in biomedisch en milieu-onderzoek. Om bewegingsartefacten te vermijden moeten huidige experimenten vaak beperkt worden tot het beeldvormen van trage vloeistofstromen. Dit is een groot nadeel aan de techniek in zijn huidige vorm.

In dit hoofdstuk wordt een nieuwe iteratieve reconstructietechniek voorgesteld die verbeterde temporele/spatiale resolutie toelaat bij het beeldvormen van vloeistofstromen door vaste media. Het voorgestelde algoritme maakt gebruik van voorkennis op twee manieren. Analoog aan hoofdstuk 3 wordt er in de eerste plaats aangenomen dat het in de tijd veranderende object bestaat uit stationaire (het vaste medium) en dynamische regio's (de vloeistofstroom). Ten tweede wordt het attenuatieverloop in een specifiek voxel binnen de dynamische regio gemodelleerd aan de hand van een stuksgewijs constante functie, wat in overeenstemming is met de werkelijke oprukkende scheidingslijn tussen de vloeistof en de lucht.

Kwantitatieve en kwalitatieve resultaten van verschillende simulatie experimenten en een experiment met neutronentomografiedata tonen aan dat, in vergelijking met huidige state-of-the-art methodes, het voorgestelde algoritme toelaat reconstructies te berekenen op basis van een substantieel minder aantal projecties zonder kwaliteitsverlies. Hieruit volgt dat de voorgestelde techniek toelaat de temporele resolutie substantieel op te drijven en bijgevolg snellere vloeistofstromen te beeldvormen.

Hoofdstuk 5 – Iteratieve reconstructie voor lage dosis bij cerebrale perfusie CT

Cerebrale perfusie CT (PCT) is een krachtige techniek voor het niet-invasief beeldvormen van hemodynamische informatie in de hersenen. Stralingsblootstelling ten gevolge van het herhaaldelijk scannen van de hersenen gedurende het perfusieproces is echter een groot nadeel. In de literatuur zijn reeds verschillende reconstructiealgoritmes voorgesteld die toelaten de geleverde dosis te verlagen terwijl de reconstructiekwaliteit behouden blijft. Deze methodes zijn ofwel afhankelijk van een vooraf opgenomen beeld van hoge kwaliteit (waardoor opnieuw extra dosis geïntroduceerd wordt) of leggen perfusiespecifieke modelaannames op over het ganse reconstructiedomein (waarbij er geen rekening gehouden wordt met de specifieke eigenschappen van verschillende lokale regio's binnen de hersenen).

In dit hoofdstuk wordt het Local Enhancement Steered Tomography (LEST) algoritme voorgesteld. In het algoritme wordt aangenomen dat de hersenen regio's bevat die niet veranderen in de tijd (de schedel en alles daarbuiten) en bovendien wordt de temporale relatie tussen reconstructies op verschillende tijdspunten in de dynamische regio's (weefsel en bloedvaten) benut. Daarenboven wordt de vorm van de arteriële input functie (AIF) onafhankelijk geoptimaliseerd op basis van de projectiedata op tussentijdse iteraties.

Het LEST algoritme werd uitgebreid gevalideerd met gesimuleerde en klinische data en de resultaten werden vergeleken met vaak gebruikte en state-of-the-art methodes. Kwantitatieve en kwalitatieve resultaten tonen aan dat het LEST algoritme, in vergelijking met de andere methodes, toelaat de radiatiedosis substantieel te verlagen met behoud van beeldkwaliteit.

Deel III: Conclusies en Appendices

Hoofdstuk 6 – Conclusies

In dit hoofdstuk wordt een samenvatting en algemene conclusies gegeven omtrent het werk dat voorgesteld werd in deze thesis.

Appendix

In de appendix wordt een klein zijproject omtrent een algoritme voor automatische schaduwkunstcreatie toegelicht. Schaduwkunst draait om het uitzonderlijke artistieke effect waarbij één of meerdere verrassende schaduwen geworpen worden vanuit een ogenschijnlijk niet gerelateerd object. De meeste schaduwkunstartiesten maken zulke beeldwerken op basis van trial-and-error, een methode die zeer moeilijk wordt als het doel is om meerdere onderling verschillende schaduwen te creëren vanuit eenzelfde object. In deze appendix wordt een uniek computer algoritme (geïnspireerd door CT-reconstructiealgoritmes) voorgesteld dat de mogelijkheid heeft zulke driedimensionale objecten automatisch te berekenen, gegeven een aantal gewenste schaduwen.

Summary

Computed Tomography (CT) is a powerful tool for non-destructive imaging in which an object's interior is visualized by reconstructing a set of projection images. The technique can be applied in various modalities, ranging from a typical X-ray CT scanner to electron microscopy and synchrotron beamlines.

Often, only limited projection data is available, which makes the reconstruction process more difficult and results in reconstruction artifacts if standard techniques are employed. Limited data problems can arise in a variety of applications. In medical CT, the acquisition of only a limited number of projections is beneficial to reduce the radiation dose delivered to the patient. In electron tomography, the sample can only be rotated over a limited tilt range due to mechanical constraints and the number of acquisition angles is often relatively small to avoid beam damage. In dynamic CT, the time to acquire sufficient projection data over the full angular range is often long in comparison to the time interval in which substantial changes inside the scanned sample occur. This implies that, in order to avoid blurry reconstructions due to the time-varying nature of the sample, only a limited amount of projections can be acquired per time frame, resulting in a limited data problem.

In this thesis, various improved reconstruction algorithms are proposed, all with the goal of achieving adequate image quality with only few projection images. The improvement is mainly due to the introduction of local models, specific to the problem at hand, into the reconstruction process. By more accurately modelling the sample, the reconstruction problem becomes more determined, which generally results in improved reconstruction quality. In the first part of this thesis, which consists solely of Chapter 2, this approach is implemented for stationary porous samples. Next, in part II, the focus changes to dynamic CT, where the approach is applied to structurally changing samples, fluid flow tomography and perfusion CT, in Chapter 3, 4 and 5, respectively. Part III contains the conclusions and the appendices. In what follows, the various chapters of this thesis are summarized.

Chapter 1 – Introduction

This chapter introduces the reader to the necessary technical background and notation of this thesis. Both classical and dynamic CT are discussed.

Part I: Stationary CT

Chapter 2 – Reconstruction method for improved porosity quantification of porous materials

In this chapter, a prior knowledge based reconstruction algorithm for porous samples is introduced. It is validated in the electron tomography setting, which is currently a versatile tool to investigate the connection between the structure and properties of nanomaterials.

With standard approaches, a quantitative interpretation of electron tomography results is still far from straightforward. Especially accurate quantification of pore-space is hampered by artifacts introduced in all steps of the processing chain, i.e., acquisition, reconstruction, segmentation and quantification. Furthermore, most common approaches require subjective manual user input.

We introduce the PORES algorithm (“*POre REconstruction and Segmentation*”); it is a tailor-made, integral approach, for the reconstruction, segmentation, and quantification of porous materials. In the PORES processing chain, a reconstruction is calculated by simultaneously classifying the interior region to the pores while reconstructing the remaining region by reducing the error with respect to the acquired projection data. This reconstruction can be directly plugged into the remaining processing chain of the PORES algorithm, resulting in accurate individual pore quantification and full sample pore statistics. The proposed approach was extensively validated on both simulated and experimental data, indicating its ability to generate accurate statistics of porous materials.

Part II: Dynamic CT

Chapter 3 – Region-based iterative reconstruction

In this chapter, a reconstruction method is introduced for locally time-varying objects, that is, objects that change dynamically only in a local region inside the object. The focus of this chapter is on structurally changing objects.

In contrast to reconstruction methods for continuously changing objects, where the continuity between reconstructions at different time points can be enforced by a deformation model or regularization techniques, current methods for the reconstruction of discretely or structurally changing objects usually calculate independent reconstructions at different points in time. The quality of such reconstructions is limited by the fact that only a small number of projections are available at each time point, thereby also limiting the time resolution.

Starting from the observation that there exist regions within the scanned object that remain unchanged over time, we introduce an iterative optimization routine that can automatically determine these regions and incorporate this knowledge in an algebraic reconstruction method. The proposed algorithm was validated on simulation data and experimental μCT data, illustrating its capability to reconstruct structurally changing objects more accurately in comparison to current techniques.

The technique introduced in this chapter, forms a basis for the dynamic reconstruction algorithms in Chapter 4 and 5.

Chapter 4 – An iterative CT reconstruction algorithm for fast fluid flow imaging

The study of fluid flow through solid matter by CT imaging has many applications, ranging from petroleum and aquifer engineering to biomedical, manufacturing and environmental research. To avoid motion artifacts, current experiments are often limited to slow fluid flow dynamics. This severely limits the applicability of the technique.

In this chapter, a new iterative CT reconstruction algorithm for improved temporal/spatial resolution in the imaging of fluid flow through solid matter is introduced. The proposed algorithm exploits prior knowledge in two ways. Firstly, analogously to Chapter 3, the time-varying object is assumed to consist of stationary (the solid matter) and dynamic regions (the fluid flow). Secondly, the attenuation curve of a particular voxel in the dynamic region is modeled by a piecewise constant function over time, which is in accordance with the actual advancing fluid/air boundary.

Quantitative and qualitative results on different simulation experiments and a real neutron tomography dataset show that, in comparison to state-of-the-art algorithms, the proposed algorithm allows reconstruction from substantially fewer projections per rotation without image quality loss. Therefore, temporal resolution can be substantially increased and thus fluid flow experiments with faster dynamics can be performed.

Chapter 5 – Iterative reconstruction for low-dose cerebral perfusion CT

Cerebral perfusion X-ray computed tomography (PCT) is a powerful tool for noninvasive imaging of hemodynamic information throughout the brain. However, because PCT requires the brain to be imaged multiple times during the perfusion process, radiation dose is of major concern. Various reconstruction algorithms that allow for lowering the dose while maintaining image quality have been proposed in the literature. These methods either depend on a prior high quality image (which again introduces extra dose) or enforce perfusion specific model assumptions globally (not taking into account the specific properties of different regions throughout the reconstructed volume).

In this chapter, we propose the Local Enhancement Steered Tomography (LEST)

method. It assumes that the brain volume has stationary regions over time (bone and void space) and exploits the temporal relation between the different time frames in the dynamic region (brain vessels and tissue). Furthermore, the shape of the arterial input function (AIF) is independently optimized based on the projection data at intermediate iterations.

The LEST algorithm is extensively validated with simulation and real clinical experiments and its performance is compared to commonly used methods and the state-of-the-art prior image constrained compressed sensing (PICCS) method. Quantitative and qualitative results show that LEST is able to substantially reduce the radiation dose while maintaining image quality in comparison to these methods.

Part III: Conclusions and Appendices

Chapter 6 – Conclusions

This chapter summarizes and provides general conclusions about the work presented in this thesis.

Appendix

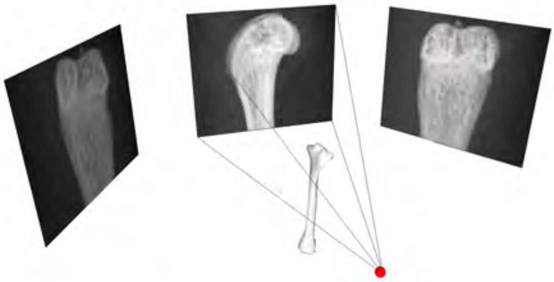
In the appendix, a small side-project about automatic algorithms for shadow art creation is described. Shadow art involves the exceptional artistic effect of surprising shadows that are shed from a seemingly unrelated sculpture. Most shadow artists construct these sculptures in a trial-and-error-fashion, a method that becomes particularly difficult if the goal is to shed multiple distinct shadows from a single object. In our approach, a unique computer algorithm (inspired by CT reconstruction techniques) was designed, which has the ability to calculate such three-dimensional objects automatically, given a set of desired shadows.

1

Introduction

Contents

In this chapter, a short introduction to computed tomography (CT) and X-ray physics is given. CT is an advanced imaging technique that allows for non-invasive visualization of the interior of a scanned object. In Fig. 1.1, the typical steps involved in CT imaging are described. In a first step, see Fig. 1.1a, X-ray projection images (also known as Röntgen photos or radiographs) are acquired at different angles from an object of interest. All steps involved in this acquisition process are described in Section 1.1. The second step in CT imaging is the reconstruction step, which is conceptually visualized in Fig. 1.1b and thoroughly explained in Section 1.2. In Section 1.3, the concepts of the previous section are extended to dynamic CT, where the object is no longer assumed to be stationary throughout the acquisition process. Finally, the introduction is finalized in Section 1.4 by enumerating some important applications of CT imaging.



(a) The first step: Projection data is acquired at different angles from an object under interest.

(b) The second step: From the projection data, the unknown attenuation values of the object are to be calculated.

Figure 1.1: A conceptual visualization of the steps involved in CT imaging.

This introductory chapter is a basic introduction to tomography. A more detailed overview of tomography can be found in [1, 2].

1.1 Acquisition process

In this section, a brief introduction to the CT acquisition process is given. First, in Section 1.1.1, the necessary X-ray physics are discussed. A description of the law of Beer-Lambert, which is a simple model for the data acquisition step, is given in Section 1.1.2. Finally, in Section 1.1.3, diffferent projection geometries are discussed.

1.1.1 X-rays: generation, matter interaction and detection

X-rays are a form of electromagnetic radiation with an energy range between 100 eV and 100 keV. They were first discovered by Wilhem Röntgen in 1895, who named it X-radiation to signify the (at that time) unknown type of radiation [3]. X-rays can be described with a wave model or a particle model. In this thesis, X-rays will be mostly modelled with a particle model, i.e., they consist of discrete bundles or packets of energy, referred to as *photons* or X-ray *quanta*.

X-ray generation

In medical and μCT scanners, X-rays are typically generated within a vacuum tube. A standard vacuum tube consists of a cathode and an anode, over which a high voltage is applied. Thermionic emission sets electrons free from the cathode, accelerating them through the high voltage such that they hit the anode surface at high speed. When the fast electrons enter the anode surface, multiple interactions take place, resulting in a conversion of the electron kinetic energy into X-ray radiation and heat. The emitted X-rays have a spectrum that typically resembles the spectrum illustrated in Fig. 1.2 [4, 1]. The shape of the spectrum can be explained by the physical mechanisms that are responsible for the production of X-rays in the X-ray tube:

- Fast electrons can be diffracted and slowed down once they enter the anode surface. Due to the charged particles being decelerated (often multiple times), electromagnetic waves (in our case: X-rays) are radiated in a continuous range of energies. This phenomenon is known as *bremsstrahlung* and corresponds to the smooth part of the spectrum in Fig. 1.2. The amount of deceleration is directly linked to the energy level of the emitted X-ray photon. If an electron directly collides into the nucleus of an anode atom, all the elec-

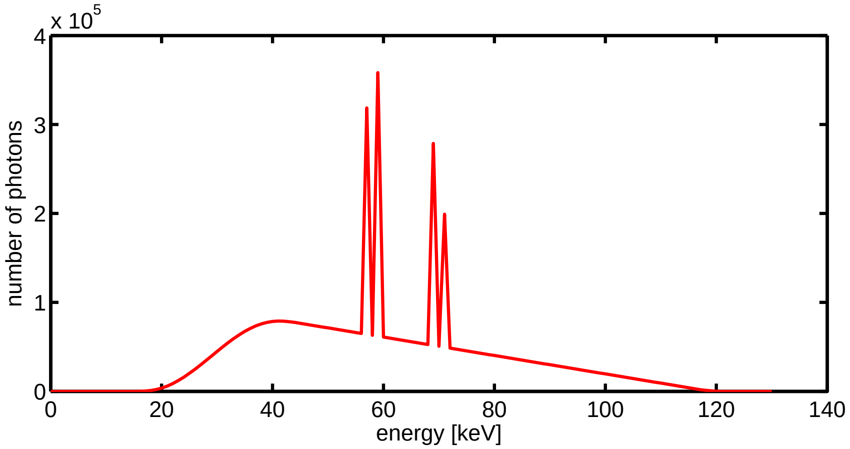


Figure 1.2: A typical conceptual representation of the X-ray spectrum in a medical CT scanner with an acceleration voltage of 120 keV.

tron's energy is conversed into X-ray radiation, a process that corresponds to the upper energy limit in the emitted spectrum (see Fig. 1.2).

- X-rays can also be generated from a direct interaction of fast electrons with the inner shell electrons of the anode's surface. If an inner shell electron is kicked out of the atom by a collision with a fast electron, an outer shell electron will take its place, a process that is accompanied by the emission of a photon. Since the binding energy difference between the same two shells is always constant, the emitted X-ray quanta resulting from this process can only have energy at a few discrete values. This creates a few sharp peaks in the emitted X-ray spectrum, known as *characteristic emission*.

The area on the anode that is hit by the fast electrons and from which the X-rays are emitted is called the *focal spot*. To obtain high resolution radiographs, a small focal spot size is desirable. The effect of the focal spot size is illustrated in Fig. 1.3. In this figure, the *penumbra* is conceptually visualized, it is a blurry region at the edge of the scanned object's projection, which is due to partial absorption/illumination of the X-rays originating from the source with a larger spot size. Since only 1% of the kinetic energy of the fast electrons is conversed into X-rays while the remaining 99% goes into heat, the maximum heat capacity of the focal spot area is the major limiting factor for the focal spot size. If the heat delivered during a single exposure exceeds the focal spot heat capacity, the anode surface can melt. For this reason, a common technique is to continuously rotate the anode, thereby spreading the heat over a larger surface. Other, more advanced, techniques exist

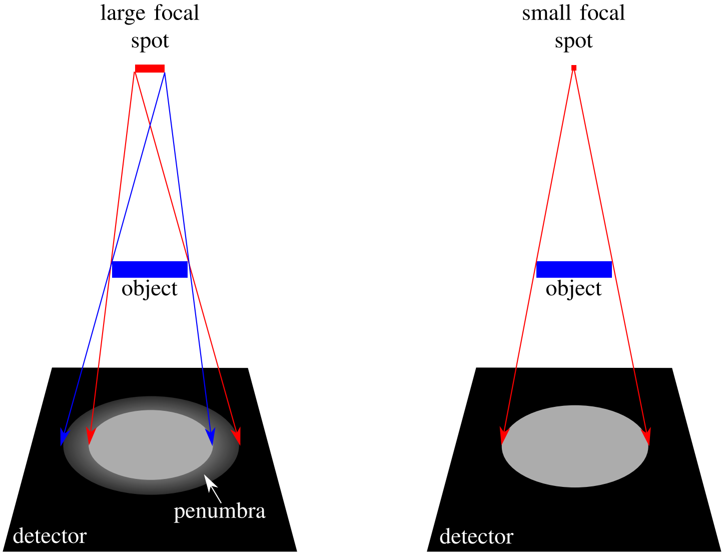


Figure 1.3: Illustration of the effect of a large focal spot size (left part of the figure) in comparison to a small focal spot size (right part of the figure).

as well, such as the liquid metal jet anode, where a continuous flow of liquid metal replaces the solid anode [5].

In synchrotron facilities, X-rays are produced in a different manner [6]. The path of the high energy electrons that are contained within the storage ring is bended (corresponding to a radial acceleration), resulting in the production of X-rays. Among many other, a great advantage in synchrotron facilities is that monochromatic X-ray beams (i.e., X-rays of a single energy level) can be produced at a high photon flux. To generate a proper photon flux in medical and lab-based CT systems, one has to work with polychromatic X-ray beams, consisting of photons within a continuous range of energies (see Fig. 1.2).

In the remainder of this thesis, the intensity of an X-ray beam (which is defined as the energy arriving at a surface of one square meter each second) is denoted by $I(\eta, E)$, where η represents the distance travelled along the X-ray path and E the energy bin. Often, a monochromatic X-ray beam is assumed, in which case the intensity is proportional to the number of photons and its intensity is denoted simply by $I(\eta)$. The X-ray beam intensity at the source position is denoted by $I(0) = I_0$.

X-ray matter interaction

As X-rays pass through an object, various scattering and absorption events result in a decreased intensity at the end of the object [7, 1]. This decreased intensity is described via the object's attenuation coefficient μ, which models all physical mechanisms that lead to attenuation of the X-ray beam. The radiation intensity of a monochromatic beam after passing a distance $\triangle\eta$ through a thin slice of homogeneous material with attenuation coefficient μ is described as follows:

$$\begin{aligned} I(\eta+\triangle\eta) &= I(\eta) - \mu(\eta) I(\eta) \triangle\eta & (1.1) \\ &= I(\eta)\left(1-\mu(\eta)\triangle\eta\right). & (1.2) \end{aligned}$$

These equations have two intuitively different interpretations. In Eq. 1.1, one can observe that the difference in intensity after passing through the thin slice is proportional to the attenuation coefficient and the distance travelled through the slice. Another interpretation is given by Eq. 1.2, where $(1-\mu(\eta)\triangle\eta)$ can be seen as the probability of a single photon of passing through the thin slice and $I(\eta)\left(1-\mu(\eta)\triangle\eta\right)$ as the expected number of photons that pass through the slice.

X-ray detection

A crucial part of a CT system is the detection of the transmitted X-rays, i.e., the beam intensity I after it passed through the object. Various types of X-ray detection technologies are described in [1]. Conventional X-ray detectors integrate the total number of photons in each detector pixel over a short period of time (the *exposure time*), without obtaining information about the energy of individual photons. An alternative is energy-resolved photon counting with dual energy detection [8] or multiple energy thresholds [9], providing the additional capability of counting individual photons based on their detected energies.

1.1.2 The law of Beer-Lambert

The law of Beer-Lambert is an important law that is often exploited in the theory of computed tomography. It relates the attenuation of light (in our case X-rays) to the properties of the material through which the light is propagating. The Beer-Lambert law states that a monochromatic X-ray with radiation intensity I_0 that propagates a distance s through a material with distance dependent attenuation coefficient $\mu(\eta)$ has a remaining intensity given by

$$I(s) = I_0 e^{-\int_0^s \mu(\eta) d\eta}. \tag{1.3}$$

The law is conceptually visualized in Fig. 1.4.

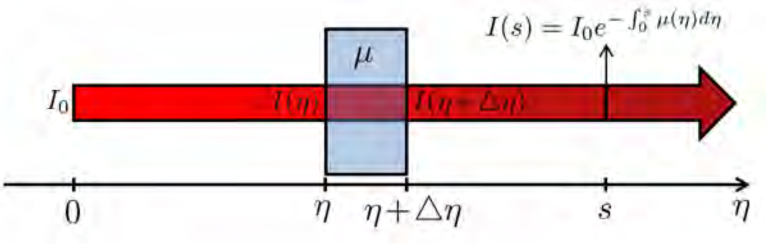


Figure 1.4: Illustration of the Beer-Lambert law.

The law of Beer-Lambert can be derived as follows. Let η be the axis parallel to the direction of the X-ray beam (see Fig. 1.4). Rearranging the terms in Eq. 1.1 gives

$$\frac{I(\eta+\triangle\eta) - I(\eta)}{\triangle\eta} = -\mu(\eta)I(\eta), \tag{1.4}$$

which results in

$$\frac{dI}{d\eta}(\eta) = -\mu(\eta)I(\eta) \tag{1.5}$$

after taking the limit for $\triangle\eta \to 0$. Eq. 1.5 is an ordinary linear differential equation which can be solved by rearranging the terms and multiplying the equation with the integrating factor $e^{\int_0^\eta \mu(\eta')d\eta'}$, leading to

$$\frac{dI}{d\eta}(\eta)e^{\int_0^\eta \mu(\eta')d\eta'} + \mu(\eta)I(\eta)e^{\int_0^\eta \mu(\eta')d\eta'} = 0. \tag{1.6}$$

Applying the product rule backwards, this simplifies to

$$\frac{d}{d\eta}\left(I(\eta)e^{\int_0^\eta \mu(\eta')d\eta'}\right) = 0. \tag{1.7}$$

Integrating both sides of this equation, leads to

$$\int_0^s \frac{d}{d\eta}\left(I(\eta)e^{\int_0^\eta \mu(\eta')d\eta'}\right) d\eta = \int_0^s 0 \; d\eta \tag{1.8}$$

$$\Leftrightarrow \quad I(s)e^{\int_0^s \mu(\eta')d\eta'} - I(0)e^{\int_0^0 \mu(\eta')d\eta'} = 0 \tag{1.9}$$

$$\Leftrightarrow \quad I(s)e^{\int_0^s \mu(\eta')d\eta'} = I(0) \tag{1.10}$$

$$\Leftrightarrow \quad I(s) = I(0)e^{-\int_0^s \mu(\eta')d\eta'}. \tag{1.11}$$

This completes the derivation, since Eq. 1.11 is equivalent to Eq. 1.3.

The law of Beer-Lambert is particularly of interest because it can be transformed to a linear relation between the measured data and the attenuation coefficients of the object. That is, Eq. 1.3 can be rewritten as

$$-\ln\left(\frac{I(s)}{I_0}\right) = \int_0^s \mu(\eta)d\eta. \tag{1.12}$$

If the incoming beam intensity I_0 is known, which is almost always the case in practical applications, the left hand side of Eq. 1.12 is fully known. The process of dividing the measured intensity by the incoming beam intensity I_0 is known as *flat field correction.* In the remainder of this thesis it is assumed that the available projection data is always in its preprocessed form $-\ln(I(s)/I_0)$.

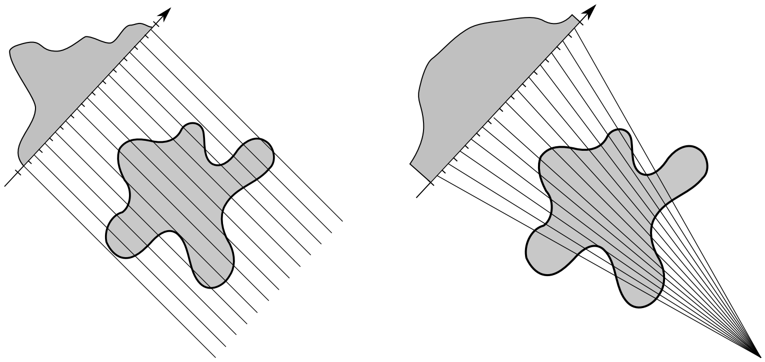
1.1.3 Projection geometries

A projection geometry refers to the set-up and position of detector and source. Without intending to cover a full oversight of all possible projection geometries, some common projection geometries are described here.

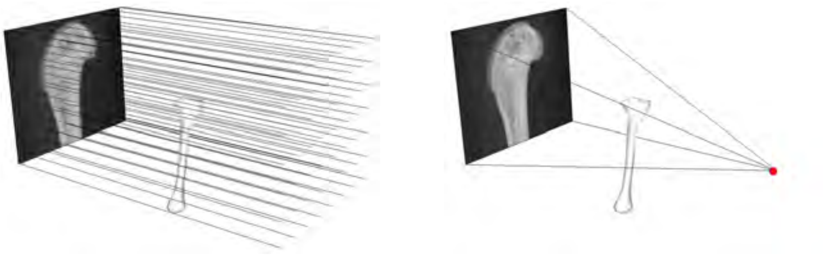
In the 2D case, a *parallel beam* geometry and *fan beam* geometry are the most common. In the parallel beam geometry all rays in a single projection are parallel to each other, as is visualized in Fig. 1.5a. In a fan beam geometry, all rays start from a single point source and reach the detector in equidistantly spaced detector points, see Fig. 1.5b. In 3D, the parallel beam (see Fig. 1.5c) and *cone* beam (see Fig. 1.5d) are the 3D analogue of the 2D parallel beam and fan beam geometry, respectively. If cone beam projections are acquired in a circular trajectory, the term *circular* cone beam geometry is utilized. A *helical* cone beam geometry refers to cone beam projections that are acquired in a helical trajectory (a geometry that is quite common in medical CT).

1.2 Reconstruction methods

The goal of a reconstruction method is to find the distribution of attenuation coefficients within the scanned object based on the measured projection data (see Fig. 1.1b). In what follows, different types of reconstruction methods are described. These can be roughly subdivided into three classes. *Analytical reconstruction methods* (Section 1.2.1) model the object's attenuation coefficients as a function of its spatial coordinates and exploit various analytical properties of the forward projection model in order to generate a reconstruction. In *algebraic reconstruction methods* (Section 1.2.2), the object is modelled on a discrete pixel/voxel grid and the reconstruction problem is reduced to a large system of linear equations. In a final class of methods, i.e., *statistical reconstruction methods* (Section 1.2.3),



(a) 2D parallel beam **(b)** 2D fan beam



(c) 3D parallel beam **(d)** 3D cone beam

Figure 1.5: Illustration of different projection geometries for the 2D and 3D case.

various statistical properties of the acquisition process are exploited and typically some likelihood function, incorporating these statistical properties, is optimized.

In what follows, all methods are described for the 2D case, i.e., a 2D object from which one-dimensional projections are acquired. For algebraic and statistical methods, the extension to the 3D case is straightforward, but will not be described here. The more difficult extension to 3D for analytical methods will also not be discussed here, since this thesis deals with algebraic reconstruction methods.

1.2.1 Analytical reconstruction methods

In the analytical approach, the object's attenuation coefficients are described as a function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ that maps the spatial coordinate (x, y) to its corresponding local attenuation coefficient μ. In Section 1.2.1.1 and Section 1.2.1.2, the Radon transform and the Fourier slice theorem are introduced. The latter makes a remarkable connection between the analytical projections and the two-dimensional

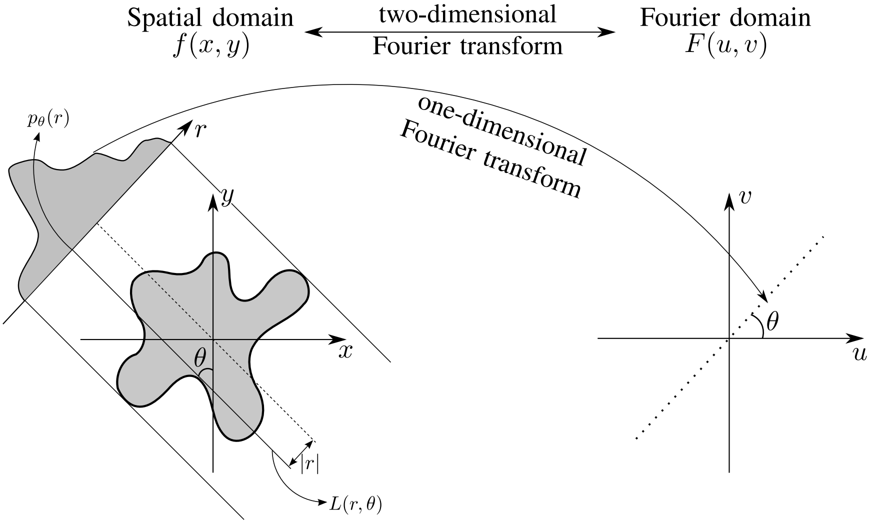


Figure 1.6: Schematic overview of the analytical projection model and the Fourier slice theorem. The one dimensional Fourier transform of the projection data at angle θ yields a slice through the Fourier domain, which is perpendicular to the projection direction.

Fourier transform of $f(x,y)$. These concepts lead to the Filtered Backprojection (FBP) reconstruction algorithm in Section 1.2.1.3. This section is finalized by briefly discussing some other analytical reconstruction methods in Section 1.2.1.4.

1.2.1.1 Radon transform

The projection process in tomography consists of straight rays traversing the object f at a certain angle and a certain distance from the center of the detector to the ray (this is illustrated in the left part of Fig. 1.6). A particular line $L(r,\theta)$ at a counter-clockwise angle θ from the y-axis and at a signed distance r from the origin is defined as follows:

$$L(r,\theta) = \{(x,y) \in \mathbb{R}^2 \mid x\cos\theta + y\sin\theta = r\}. \tag{1.13}$$

Remember that the log- and flat-field-corrected projections are theoretically given by a line integral of the attenuation coefficients, see Eq. 1.12. Therefore, a particular projection value $p_\theta(r)$ is defined as the line integral through $f(x,y)$ over the line $L(r,\theta)$:

$$p_\theta(r) = \int_{L(r,\theta)} f(x,y)ds. \tag{1.14}$$

The Radon transform $\mathcal{R}$ is the transformation that maps the object function $f(x,y)$ to the complete set of projection values, i.e.,

$$f(x,y) \overset{\mathcal{R}}{\leftrightarrow} \{p_\theta(r) \mid \theta \in [0,\pi[, r \in \mathbb{R}\}. \tag{1.15}$$

Eq. 1.15 implies that $(\mathcal{R}f)(\theta,r) = p_\theta(r)$.

1.2.1.2 Fourier slice theorem

The Fourier slice theorem for two dimensions is conceptually visualized in Fig. 1.6. The theorem states that the one-dimensional Fourier transform of a parallel beam projection of $f(x,y)$ at a certain angle is exactly the same as the slice through the Fourier transform $F(u,v)$ which is perpendicular to the projection direction. The Fourier slice theorem is stated more precisely in the following theorem:

Theorem. *Let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R} : (x,y) \mapsto f(x,y)$ be a two-dimensional function and define its projection $p_\theta(r) : \mathbb{R} \to \mathbb{R} : r \mapsto p_\theta(r)$ as in Eq. 1.14. Denote the two-dimensional Fourier transform of $f(x,y)$ as $F(u,v)$ and the one-dimensional Fourier transform of $p_\theta(r)$ as $P_\theta(q)$. Then the following equality holds:*

$$F(u,v)\big|_{\substack{u=q\cos(\theta)\\ v=q\sin(\theta)}} = P_\theta(q). \tag{1.16}$$

Proof. The proof of the Fourier slice theorem is straightforward after assuming that $\theta = 0$. This can be assumed without loss of generality. Indeed, if the theorem applies for $\theta = 0$, then the theorem is also valid for any $\theta \neq 0$, since a rotation is the spatial domain corresponds to exactly the same rotation in Fourier space. It is hence sufficient to prove that $F(q,0) = P_0(q)$.

The line integral $p_0(r)$ corresponds to

$$p_0(r) = \int_{L(r,0)} f(x,y)ds = \int_{-\infty}^{\infty} f(r,y)\,dy \tag{1.17}$$

since $L(r,0)$ is the line $x = r$. The Fourier transform of $f(x,y)$ is

$$F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,e^{-2\pi i(xu+yv)}\,dx\,dy. \tag{1.18}$$

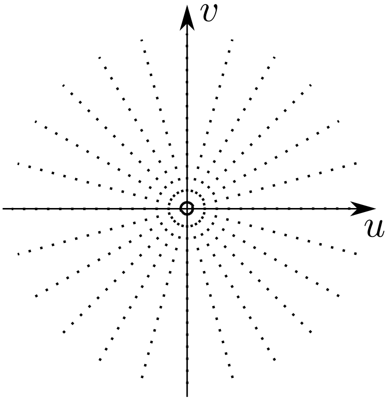


Figure 1.7: Parallel beam projections yield a radial filling of values in the Fourier space.

Evaluated in $u = q$ and $v = 0$ this becomes

$$\begin{aligned} F(q,0) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, e^{-2\pi ixq}\, dxdy \qquad (1.19) \\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x,y)\, dy\right] e^{-2\pi ixq} dx \qquad (1.20) \\ &= \int_{-\infty}^{\infty} p_0(x)\, e^{-2\pi ixq} dx \qquad (1.21) \\ &= P_0(q). \qquad (1.22) \end{aligned}$$

This concludes the proof. □

This theorem has some important consequences. In theory, if enough projections $p_\theta(r)$ can be acquired, the Fourier domain of the object can be fully sampled. A simple inverse Fourier transform could hence suffice to reconstruct the object function $f(x, y)$. In practice however, projections are given in a finite number of detector pixels and can only be acquired at a finite number of angles. Assuming a parallel beam projection geometry (i.e., all rays in a single projection image are parallel to each other), the Fourier space of the object $f(x, y)$ is radially sampled, as is illustrated in Fig. 1.7. To use the fast Fourier transform (FFT) algorithm [10] to perform the inverse Fourier transform in an actual implementation, the radially sampled Fourier space must be resampled on a regular grid. Therefore, interpolation is necessary. However, this introduces large interpolation errors in the higher frequency part of the spectrum, since samples are only sparsely available in this region. These high spatial frequencies correspond to fine details in the object function, and since they are less accurately represented after the resampling, image quality is seriously degraded. In the next section, an analytical method that compensates for the fact that points in Fourier space are more densely sampled near the origin is introduced.

1.2.1.3 Filtered Back Projection (FBP)

The Filtered Back Projection (FBP) reconstruction method is based on the following analytical formula:

$$f(x,y) = \int_0^{\pi} \left\{ \int_{-\infty}^{\infty} P_\theta(q)|q|e^{2\pi iq(x\cos\theta + y\sin\theta)}dq \right\} d\theta. \tag{1.23}$$

Before proving this formula, its different components are explained. As can be observed from Eq. 1.23, the FBP formula gives rise to a simple two step approach for calculating a reconstruction of the scanned object based on the measured projection data:

1. Filter the projection data $p_\theta(r)$ by multiplying its Fourier transform $P_\theta(q)$ with $|q|$ and calculating the inverse Fourier transform. This step corresponds to the inner integral in Eq. 1.23.
2. For a particular location in the image domain (x, y), sum up all the filtered projection data that corresponds to the lines $x\cos\theta + y\sin\theta$ with $\theta \in [0, \pi]$. This step corresponds to the outer integral in Eq. 1.23.

This approach can be turned into a practical algorithm, keeping the following in mind:

- Since projection data is acquired at a finite number of detector pixels, and thus only available at discrete locations, the Fourier and inverse Fourier transform are performed with the FFT algorithm. Also, the formula is only evaluated at discrete locations in the spatial domain, typically on the pixel-coordinates of a pixel grid. Therefore, in a practical implementation, the entire formula in Eq. 1.23 is discretized by changing the integrals to sums.
- Projection data needs to be acquired over the full angular range $[0, \pi]$, corresponding to the outer integral in Eq. 1.23. Also, the scanned object must be fully inside the field of view, so that $p_\theta(r)$ and hence also $P_\theta(q)$ is available on its entire domain, thereby ensuring that the inner integral in Eq. 1.23 can be calculated. If these assumptions are violated, the reconstructed image will contain artifacts.
- In a practical implementation, the FBP formula Eq. 1.23 is never evaluated individually at different locations in the spatial domain. Typically, the FBP

formula is evaluated simultaneously at all pixel coordinates by first filtering the projection data with the high pass filter $|q|$ (in Fourier space) and subsequently summing all *backprojections* of each projection onto the pixel grid. A backprojection simply places the values of the projection $p_\theta(r)$ (with r describing the signed distances to all rays in a single projection defined by the fixed angle θ) at all pixels that coincide with the line $x\cos\theta + y\sin\theta = r$. This implies that the FBP algorithm as described here is only suitable for a parallel beam geometry.

The application of the high-pass filter $|q|$ compensates for the high density sampling ratio in the low frequency domain and the low density sampling ratio in the high frequency domain (see Fig. 1.7). In the context of FBP, the high pass filter $|q|$ is usually referred to as the “ramp” filter, because of its shape in the Fourier domain.

The FBP formula can be derived as follows. First, the rectangular coordinate system (u, v) over which is integrated in the inverse Fourier transform formula, i.e.,

$$f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)e^{2\pi i(xu+yv)}\, dudv, \tag{1.24}$$

is changed to a polar coordinate system (q, θ). This is achieved by making the substitution $u = q\cos\theta$ and $v = q\sin\theta$, which results in

$$dudv = \begin{vmatrix} \cos\theta & -q\sin\theta \\ \sin\theta & q\cos\theta \end{vmatrix}\, dqd\theta = q\, dqd\theta. \tag{1.25}$$

The inverse Fourier transform in Eq. 1.24 can now be expressed in polar form as

$$f(x,y) = \int_0^{2\pi}\int_0^{\infty} F(q\cos\theta, q\sin\theta)e^{2\pi iq(x\cos\theta + y\sin\theta)} q\, dqd\theta. \tag{1.26}$$

The integral in Eq. 1.26 can be split into two parts by integrating θ from 0 to π and from π to 2π and further rewritten as follows:

$$\begin{aligned} f(x,y) &= \int_0^{\pi}\int_0^{\infty} F(q\cos\theta, q\sin\theta)e^{2\pi iq(x\cos\theta + y\sin\theta)} q\, dqd\theta \\ &+ \underbrace{\int_0^{\pi}\int_0^{\infty} F(q\cos(\theta+\pi), q\sin(\theta+\pi))e^{2\pi iq(x\cos(\theta+\pi) + y\sin(\theta+\pi))} q\, dqd\theta}_{\substack{= \int_0^{\pi}\int_0^{\infty} F(-q\cos\theta,\, -q\sin\theta)e^{2\pi i(-q)(x\cos\theta + y\sin\theta)} q\, dqd\theta \\ = \int_0^{\pi}\int_{-\infty}^{0} F(q\cos\theta,\, q\sin\theta)e^{2\pi iq(x\cos\theta + y\sin\theta)} (-q)\, dqd\theta}} \\ &= \int_0^{\pi}\int_{-\infty}^{\infty} F(q\cos\theta, q\sin\theta)|q|e^{2\pi iq(x\cos\theta + y\sin\theta)}\, dqd\theta. \end{aligned} \tag{1.27}$$

Substituting the Fourier slice theorem formula (see Eq. 1.16) into Eq. 1.27 leads to

$$f(x, y) = \int_0^\pi \int_{-\infty}^\infty P_\theta(q)|q|e^{2\pi iq(x\cos\theta + y\sin\theta)}\, dqd\theta, \tag{1.28}$$

which completes the derivation of the FBP formula Eq. 1.23.

1.2.1.4 Other analytical reconstruction methods

Various variants on the FBP algorithm have been proposed in literature. Most variants focus on redesigning the ramp filter $|q|$ in Eq. 1.23. One approach is to train a neural network that learns an optimal filter for a certain class of objects [11]. Another approach consist of approximating algebraic methods (see Section 1.2.2) by selecting a proper FBP filter [12]. The FBP algorithm described in Section 1.2.1.3 is suitable for a 2D parallel beam geometry. For fan beam, another formula and approach must be followed [2]. For the 3D case, the Feldkamp (FDK) algorithm is a common choice for reconstructing from circular cone beam projections [13]. For helical cone beam CT, other algorithms are available [14].

1.2.2 Algebraic reconstruction methods

In the algebraic approach, the object's attenuation coefficients are modelled by a discrete representation of the object function $f(x, y)$, typically on a pixel grid. The discretization process is presented in Section 1.2.2.1. A well-known algebraic reconstruction method, the simultaneous iterative reconstruction technique (SIRT), is derived in Section 1.2.2.2. Next, some other algebraic reconstruction methods are discussed in Section 1.2.2.3. Finally, some practical implementation issues are addressed in Section 1.2.2.4.

1.2.2.1 Discretization

A typical approach is to represent the scanned object on a pixel (or, in the 3D case, voxel) grid. This is conceptually visualized in Fig. 1.8, which is the discrete analogue of the left part of Fig. 1.6. In the discretization on a pixel grid, it is assumed that the object has a constant attenuation value within each pixel. Assuming the pixel grid consists of N pixels, the object function $f(x, y)$ is approximated as $f(x, y) \approx \sum_{j=1}^{N} x_j\phi_j$ with ϕ_j the pixel basis function for the jth pixel which is zero outside the pixel and one inside and x_j the constant attenuation value within the jth pixel. The basis functions ϕ_j do not necessarily need to be pixel basis functions. Other choices are possible as well, e.g., generalized Kaiser-Bessel functions (also known as blobs) [15]. Since the object is modelled by a finite number of attenuation values, it can be represented as a column vector $\boldsymbol{x} = (x_j) \in \mathbb{R}^N$.

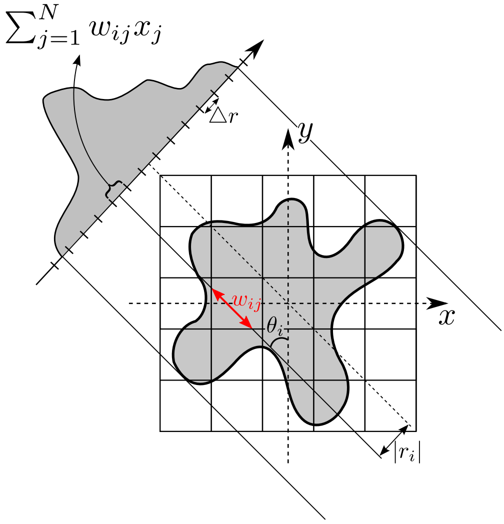
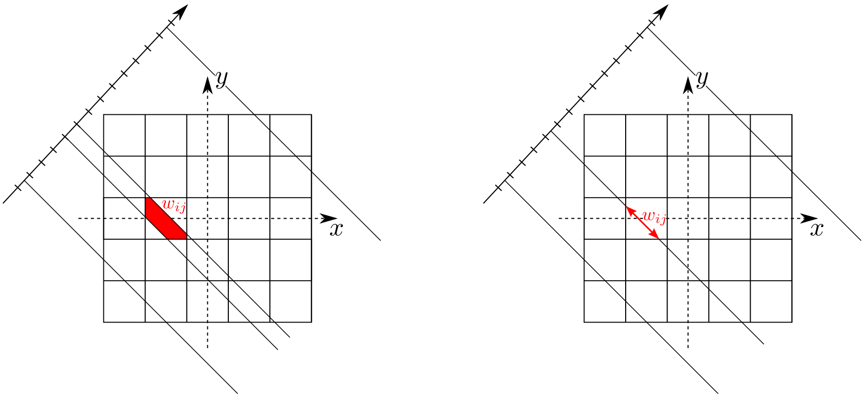


Figure 1.8: Illustration of the discrete representation of the object and the projection. In this image, the contribution w_{ij} of pixel j to the projection value with index i is represented as the ray-intersection length of projection line i with pixel j.

In practice, the measured projection data is also discrete, it consists of a finite number of measured projection values, each one corresponding to a specific detector pixel at a specific angle. Let M denote the total number of measured projection values for all angles, which are log-corrected and ordered in a vector $\boldsymbol{p} = (p_i) \in \mathbb{R}^M$. Denote θ_i as the counter-clockwise angle from the y-axis and r_i as the signed distance from the origin to the center of the detector pixel corresponding to p_i. Following Eq. 1.14, each projection value p_i can be modelled as

$$p_i = \int_{-\frac{\triangle r}{2}}^{\frac{\triangle r}{2}} p_{\theta_i}(r_i + r')\, dr' = \int_{-\frac{\triangle r}{2}}^{\frac{\triangle r}{2}} \int_{L(r_i,\theta_i)} f(x,y)\, ds dr', \tag{1.29}$$

where $\triangle r$ represents the detector pixel width. In algebraic reconstruction methods, the forward projection model of Eq. 1.29 is approximated by $p_i \approx \sum_{j=1}^{N} w_{ij}x_j$, where w_{ij} represents the contribution of pixel j to the projection value with index i. This is also illustrated in Fig. 1.8. The complete projection data $\boldsymbol{p}$ can then be simulated by $\boldsymbol{W}\boldsymbol{x}$, where $\boldsymbol{W} = (w_{ij}) \in \mathbb{R}^{M\times N}$ is a sparse matrix that collects all weights w_{ij}. The weights can be calculated in a variety of ways. The most precise calculation involves a strip-kernel which is visualized in Fig. 1.9a, where the weight w_{ij} is equal to the fractional area of the jth pixel intercepted by the ith ray. A computationally faster approximation is given by the line-kernel, where the weight w_{ij} equals the ray-intersection length of the ith ray with the jth pixel



(a) strip-kernel **(b)** line-kernel

Figure 1.9: Illustration of two different approaches to calculate the weights of the forward projection matrix $\boldsymbol{W}$.

(see Fig. 1.9b), or by a linear-kernel [16] (also known as Joseph's method), where the contribution of the ray to the projection value is determined by linearly interpolating between the two nearest pixels of the intersection of the ray and the row or column.

Directly solving the system of linear equations $\boldsymbol{W}\boldsymbol{x} = \boldsymbol{p}$ for an exact solution $\boldsymbol{x}$ is typically infeasible, since noise and discretization effects render the system of linear equations inconsistent. Therefore, algebraic methods typically minimize the *projection distance* $||\boldsymbol{W}\boldsymbol{x} - \boldsymbol{p}||$ for some norm $||\cdot||$.

In this thesis, the projection data acquired from a 2D object is usually represented as a *sinogram*. A sinogram collects the projections from a 2D object in a matrix, where columns typically represents the different detector pixels and the rows the different projection angles. An example of a sinogram acquired from the Shepp-Logan phantom [17] is given in Fig. 1.10.

1.2.2.2 Simultaneous Iterative Reconstruction Technique (SIRT)

The simultaneous iterative reconstruction technique (SIRT) is an algebraic reconstruction algorithm known to converge to a solution of

$$x^* = \arg\min_{\boldsymbol{x}} \left(||\boldsymbol{W}\boldsymbol{x} - \boldsymbol{p}||^2_{\boldsymbol{R}}\right), \tag{1.30}$$

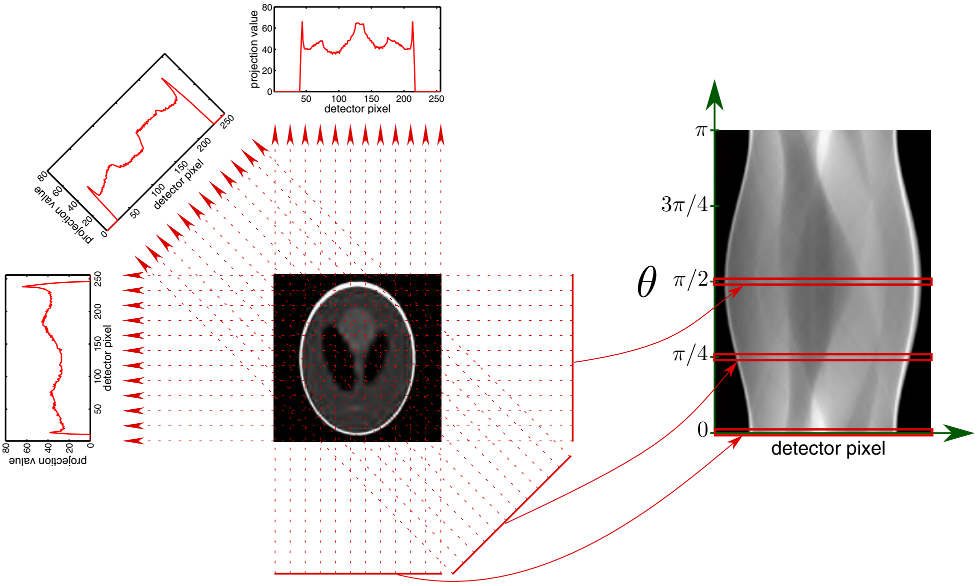


Figure 1.10: An example of a sinogram acquired with a parallel beam geometry from the Shepp-Logan phantom. The projections acquired at $\theta = 0$, $\theta = \pi/4$ and $\theta = \pi/2$ are explicitly shown and their connection with the sinogram is illustrated.

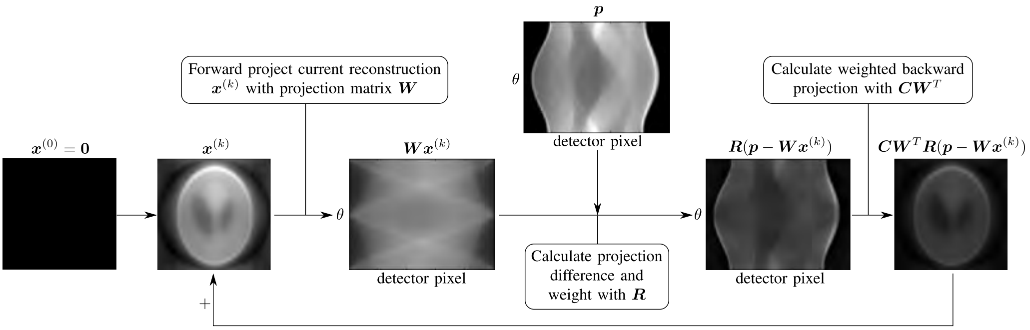


Figure 1.11: The SIRT update process illustrated with the Shepp-Logan phantom.

where $\boldsymbol{R} = (r_{ij}) \in \mathbb{R}^{M \times M}$ is the diagonal matrix with inverse row sums of the projection matrix $\boldsymbol{W}$ (its diagonal elements are given by $r_{ii} = 1/\sum_j w_{ij}$) and $||\boldsymbol{W}\boldsymbol{x} - \boldsymbol{p}||^2_{\boldsymbol{R}} = (\boldsymbol{W}\boldsymbol{x} - \boldsymbol{p})^T\boldsymbol{R}(\boldsymbol{W}\boldsymbol{x} - \boldsymbol{p})$ [1, 18, 19]. Starting from an initial reconstruction $\boldsymbol{x}^{(0)} = \boldsymbol{0}$, the SIRT algorithm iteratively updates the reconstruction as follows:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R}(\boldsymbol{p} - \boldsymbol{W}\boldsymbol{x}^{(k)}) \ , \tag{1.31}$$

where $\boldsymbol{C} = (c_{ij}) \in \mathbb{R}^{N \times N}$ is defined as the diagonal matrix with the inverse column sums of $\boldsymbol{W}$ (i.e., $c_{jj} = 1/\sum_i w_{ij}$). Before demonstrating the connection between Eq. 1.30 and Eq. 1.31, the iterative update in Eq. 1.31 is analysed more closely. The update in Eq. 1.31 is illustrated in Fig. 1.11 and consists of the following steps:

1. Starting from the current estimate $\boldsymbol{x}^{(k)}$, a forward projection $\boldsymbol{W}\boldsymbol{x}^{(k)}$ is simulated.

2. The *projection difference* $\boldsymbol{p} - \boldsymbol{W}\boldsymbol{x}^{(k)}$ is calculated. This difference indicates where and how much the simulated projection data $\boldsymbol{W}\boldsymbol{x}^{(k)}$ is different from the measured projection data $\boldsymbol{p}$. It quantifies the reconstruction quality of the current estimate $\boldsymbol{x}^{(k)}$.

3. The projection difference $\boldsymbol{p} - \boldsymbol{W}\boldsymbol{x}^{(k)}$ is weighted with the inverse row sum matrix $\boldsymbol{R}$. Intuitively, projection value differences corresponding to rays with a long intersection length with the pixel grid of the reconstruction domain get a small weight and projection value differences corresponding to rays that intersect only shortly with the reconstruction domain get a large weight. A large weight (and hence short ray-intersection length) indicates that the projection difference in the particular detector pixel has a large influence in the update. This makes sense, since that ray intersects with only few pixels in

the reconstruction domain, and thus the projection difference in the corresponding detector pixel is a good indication for how to update the pixels on that particular ray. If a ray intersects a lot of pixels, the projection difference is less trustworthy since the projection value is based on the (weighted) sum of lots of attenuation values along the ray.

4. The backprojected weighted projection difference $\boldsymbol{W}^T\boldsymbol{R}(\boldsymbol{p}-\boldsymbol{W}\boldsymbol{x}^{(k)})$ is calculated. That is, all projection differences are "smeared back" over the reconstruction domain by simply assigning the projection difference value of each ray to all pixels along that ray, weighted with the ray-pixel-intersection length.

5. The result is weighted with the inverse column sum matrix $\boldsymbol{C}$. For a particular pixel, this corresponds to dividing the pixel value of the update by the combined length of all rays that intersected that pixel. The resulting update reflects how the reconstruction should be updated in order to reduce the projection difference.

6. The update is added to the current reconstruction and the algorithm continues from the first step until a certain stopping criterion is reached. The stopping criterion can consist of stopping the algorithm after a predefined number of iterations or after the projection distance has reached a certain threshold value.

The connection between Eq. 1.30 and Eq. 1.31 can be made as follows. The normal equations for Eq. 1.30 are given by

$$\boldsymbol{W}^T\boldsymbol{R}\boldsymbol{W}\boldsymbol{x}=\boldsymbol{W}^T\boldsymbol{R}\boldsymbol{p}. \tag{1.32}$$

Preconditioning Eq. 1.32 with $\boldsymbol{C}$ and rewriting the equations gives

$$\begin{aligned}
&\boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R}\boldsymbol{W}\boldsymbol{x}=\boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R}\boldsymbol{p} && (1.33)\\
\Leftrightarrow\quad &(\boldsymbol{I}-(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R}\boldsymbol{W}))\boldsymbol{x}=\boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R}\boldsymbol{p} && (1.34)\\
\Leftrightarrow\quad &\boldsymbol{x}=(\boldsymbol{I}-\boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R}\boldsymbol{W})\boldsymbol{x}+\boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R}\boldsymbol{p} && (1.35)\\
\Leftrightarrow\quad &\boldsymbol{x}=\boldsymbol{x}+\boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R}(\boldsymbol{p}-\boldsymbol{W}\boldsymbol{x}). && (1.36)
\end{aligned}$$

It follows that Eq. 1.31 is a fixed point iteration scheme derived from Eq. 1.36. It can be proven that this scheme converges (for a proof the reader is referred to [18]).

The SIRT algorithm in Eq. 1.31 can also be written explicitly for each component of $\boldsymbol{x}$, i.e., each pixel's attenuation value x_j:

$$x_j^{(k+1)} = x_j^{(k)} + \frac{1}{\sum_{i=1}^{M} w_{ij}} \sum_{i=1}^{M} \left(\frac{w_{ij}(p_i - \sum_{h=1}^{N} w_{ih} x_h^{(k)})}{\sum_{h=1}^{N} w_{ih}} \right) \quad \forall j \tag{1.37}$$

1.2.2.3 Other algebraic reconstruction methods

In this section, some other algebraic reconstruction methods are discussed. It should be noted that the algorithms listed here are merely a small grasp of all available algebraic reconstruction methods.

Algebraic Reconstruction Technique (ART) [20]: In the ART method, the update formula is given by

$$x_j^{(k+1)} = x_j^{(k)} + \frac{w_{ij}(p_i - \sum_{h=1}^{N} w_{ih} x_h^{(k)})}{\sum_{h=1}^{N} w_{ih}^2} \quad \forall j, \tag{1.38}$$

with i the index of the particular projection value that is utilized in iteration k. As can be observed from Eq. 1.38, ART processes a single projection value at a time to update the current reconstruction. In a geometric interpretation, the update formula represents the orthogonal projection of the current reconstruction on the hyperplane defined by the ith equation in $\boldsymbol{W}\boldsymbol{x} = \boldsymbol{p}$. Also note the difference with SIRT (Eq. 1.37), where all projection values are employed simultaneous to generate an update for the current reconstruction. The order in which the projection values are selected has been studied extensively in the past, and various schemes with different properties are available [21, 22]. Generally, ART converges faster but is less stable with respect to noise.

Simultaneous Algebraic Reconstruction Technique (SART) [23]: In SART, all projection values in a single projection are processed together in a single iteration. Let I_θ denote the set of all indices corresponding to projection values in a single projection at angle θ, then the update formula for SART is given by

$$x_j^{(k+1)} = x_j^{(k)} + \frac{1}{\sum_{i \in I_\theta} w_{ij}} \sum_{i \in I_\theta} \left(\frac{w_{ij}(p_i - \sum_{h=1}^{N} w_{ih} x_h^{(k)})}{\sum_{h=1}^{N} w_{ih}} \right) \quad \forall j. \tag{1.39}$$

Notice that this algorithm is almost exactly the same as SIRT, with the only difference that SIRT processes all projection data simultaneous in one iteration while SART handles one projection at a time. Modifications to SART have been developed where more than one projections are processed at a time [24].

Conjugate Gradient Least Squares (CGLS) [25]: In the CGLS method, the conjugate gradient method is applied to solve the following sparse least-squares problem:

$$x^* = \text{argmin}_{\boldsymbol{x}} \left(||\boldsymbol{W}\boldsymbol{x} - \boldsymbol{p}||_2^2 \right). \tag{1.40}$$

Notice that the norm in the object function of Eq. 1.40 is the standard 2-norm whereas in SIRT a weighted norm is minimized (see Eq. 1.30).

Regularization methods: Various types of algorithms introduce a regularization term $U(\boldsymbol{x})$ in the objective function. The optimization problem then becomes

$$x^* = \text{argmin}_{\boldsymbol{x}} \left(||\boldsymbol{W}\boldsymbol{x} - \boldsymbol{p}||^2 + \lambda U(\boldsymbol{x}) \right), \tag{1.41}$$

where $\lambda > 0$ is the regularization parameter that controls the strength of the regularization. The regularization term $U(\boldsymbol{x})$ typically reflects some prior knowledge about the scanned object $\boldsymbol{x}$. Among many options, popular choices for $U(\boldsymbol{x})$ include the total variation penalty [26], smoothness priors [27] and the non-local means prior [28, 29].

Algorithms for special cases: Many algebraic algorithms have been tailored specifically to certain applications. These algorithms benefit from exploiting various types of prior knowledge about the scanned object, in the sense that artifacts can be reduced or high quality reconstructions can be reconstructed from only few projections. In what follows, a few examples of many hundreds of specific case algebraic algorithms are enumerated.

If the object consist of only a few discrete grey levels, the discrete algebraic reconstruction technique (DART) has shown great potential for practical applications [30, 31]. In fact, an entire domain named “discrete tomography” deals with the reconstructions of such objects [32].

Often, a segmentation of the reconstruction is of interest. Various algorithms combine the segmentation and the reconstruction step into one algorithm, thereby simultaneously improving both steps [33, 34].

Specific algorithms have been developed in medical imaging for reducing metal artifacts, such as the normalized metal artifact reduction (NMAR) [35] and the frequency split metal artifact reduction (FSMAR) [36] algorithms.

In region of interest (ROI) tomography, the goal is to reconstruct only a small region inside the object. Often, certain assumptions are made about the region of the object surrounding the ROI, thereby greatly improving image quality [37, 38, 39].

Another approach consists of changing the acquisition protocol from a step-and-shoot approach (i.e., stopping the gantry at each acquisition angle to acquire a projection) to a continuous gantry rotation. This can be modelled in the reconstruction algorithm, resulting in algorithms which are particularly useful for ROI tomography [40].

1.2.2.4 Practical implementation

Algebraic reconstruction methods impose major computational and memory demands. First of all, storing the entire forward projection matrix $\boldsymbol{W}$ into memory is almost impossible in practical applications, since real projection data can contain hundreds of projections containing millions of pixels and reconstructed volumes/images can contain even more voxels/pixels. Therefore, the values of the matrix $\boldsymbol{W}$ are typically calculated on-the-fly. Secondly, applying the update operation in Eq. 1.37 pixel by pixel is practically infeasible. For reasonable computation times a GPU implementation is desirable. In this thesis, all forward and backward projection operations were implemented using the ASTRA toolbox for tomography [41, 42, 43].

1.2.3 Statistical reconstruction methods

In this section, a short introduction to statistical reconstruction methods is given. First, in Section 1.2.3.1, it is explained that the number of detected X-ray quanta follow a Poisson distribution. Next, in Section 1.2.3.2, the expectation maximization (EM) algorithm is derived. Finally, a class of more stable statistical algorithms is described in Section 1.2.3.3.

1.2.3.1 Poisson distributed noise

There are three components in X-ray imaging that can be modelled via statistical processes (the reader is referred to [1] for more details):

- In a typical detector, X-ray photon detection is a statistically independent process. Therefore, the detection of multiple photons follows a Binomial distribution.
- The absorption of photons inside the object are also guided by Binomial statistics.
- The number of photons that leave the source follows a Poisson distribution.

Because Binomial selection of a Poisson process yields another Poisson process, the overall statistics can be modelled via a Poisson distribution:

$$P(I_i|I_i^*) = \frac{(I_i^*)^{I_i}}{I_i!} e^{-I_i^*} \quad i = 1, \ldots, M, \tag{1.42}$$

where I_i denotes the number of measured X-ray quanta at the ith detector pixel and I_i^* the expected number of detected X-ray quanta[1]. From Eq. 1.42, it follows that

$$P(\boldsymbol{I}|\boldsymbol{I}^*) = \prod_{i=1}^{M} \frac{(I_i^*)^{I_i}}{I_i!} e^{-I_i^*}, \tag{1.43}$$

where $\boldsymbol{I}$ and $\boldsymbol{I}^*$ are the column vectors collecting all the values I_i and I_i^* for all i, respectively.

1.2.3.2 Expectation Maximization (EM)

The underlying idea of the expectation maximization (EM) algorithm [44, 45, 46] is to maximize the likelihood of acquiring the observed data while varying the expectation values of the attenuation coefficients of the scanned object. It can be derived as follows. Substituting the law of Beer-Lambert (i.e., the discrete version of Eq. 1.3) into Eq. 1.43 results in

$$L(\boldsymbol{x}^*) := P(\boldsymbol{I}|\boldsymbol{x}^*) = \prod_{i=1}^{M} \frac{\left(I_0 e^{-\sum_{j=1}^{N} w_{ij} x_j^*}\right)^{I_i}}{I_i!} e^{-I_0 e^{-\sum_{j=1}^{N} w_{ij} x_j^*}}, \tag{1.44}$$

where $\boldsymbol{x}^*$ denotes the expected attenuation values for the scanned object. The product in the *likelihood* function $L(\boldsymbol{x}^*)$ can be rewritten as a sum by taking its logarithm to produce the *log likelihood* function

$$l(\boldsymbol{x}^*) := \ln\left(L(\boldsymbol{x}^*)\right) = \sum_{i=1}^{M} \left(I_i \ln(I_0) - I_i \sum_{j=1}^{N} w_{ij} x_j^* - \ln(I_i!) - I_0 e^{-\sum_{i=1}^{N} w_{ij} x_j^*} \right). \tag{1.45}$$

Note that the maximization of Eq. 1.45 or Eq. 1.44 yields the same result, since the logarithm is a monotonically increasing function. Therefore, maximizing $P(\boldsymbol{I}|\boldsymbol{x}^*)$ (which is the goal of the EM algorithm) is equivalent to solving the following optimization problem:

$$\hat{\boldsymbol{x}}^* = \arg \max_{\boldsymbol{x}^* \geq \boldsymbol{0}} \left(l(\boldsymbol{x}^*) \right). \tag{1.46}$$

If the log likelihood function $l(\boldsymbol{x}^*)$ is concave, a global maximum is guaranteed. The concavity of $l(\boldsymbol{x}^*)$ can be shown by proving the equivalent condition that the Hessian of $l(\boldsymbol{x}^*)$ is negative semi-definitive. The reader is referred to [45] for a

---

[1] This introduces a slight abuse of notation since I_0 was already defined in Section 1.1.1. I_0 still denotes the radiation intensity at the source (in this case expressed in number of photons), whereas I_i ($i = 1, \ldots, M$) denotes the number of measured photons in the ith detector pixel. Throughout this section, it is assumed that I_0 is the same for all detector pixels. However, all derivations can be easily adapted to account for a detector pixel dependent I_0 value.

proof of this statement. Since a global maximum exists, the Karush-Kuhn-Tucker (KKT) conditions must be fulfilled. The inequality constraint in Eq. 1.46 can be written in a more standard form for the KKT conditions as $g_k(\boldsymbol{x}^*) := -x_k^* \leq 0$ for all $k = 1, \ldots, N$. The KKT conditions guarantee the existence of the KKT multipliers $\lambda_k \in \mathbb{R}$ $(k = 1, \ldots, N)$ such that

$$(1) \quad \nabla l(\hat{\boldsymbol{x}}^*) - \sum_{k=1}^{N} \lambda_k \nabla g_k(\hat{\boldsymbol{x}}^*) = 0 \tag{1.47}$$

$$(2) \quad g_k(\hat{\boldsymbol{x}}^*) \leq 0 \text{ for all } k = 1, \ldots, N \tag{1.48}$$

$$(3) \quad \lambda_k \geq 0 \text{ for all } k = 1, \ldots, N \tag{1.49}$$

$$(4) \quad \lambda_k g_k(\hat{\boldsymbol{x}}^*) = 0 \text{ for all } k = 1, \ldots, N. \tag{1.50}$$

Condition (1), (2), (3) and (4) are typically referred to as the stationarity condition, the primal feasibility condition, the dual feasibility and the complementary slackness condition, respectively. The first condition in Eq. 1.47 is equivalent to

$$\lambda_k = -\frac{\partial l}{\partial x_k^*}(\hat{\boldsymbol{x}}^*). \tag{1.51}$$

Substituting Eq. 1.51 in the third and fourth condition (see Eq. 1.49 and Eq. 1.50) yields

$$(3) \quad \frac{\partial l}{\partial x_k^*}(\hat{\boldsymbol{x}}^*) \leq 0 \tag{1.52}$$

$$(4) \quad \hat{x}_k^* \frac{\partial l}{\partial x_k^*}(\hat{\boldsymbol{x}}^*) = 0. \tag{1.53}$$

By calculating the partial derivative of l with respect to x_k^*, Eq. 1.53 can be rewritten as

$$\hat{x}_k^* \frac{\partial l}{\partial x_k^*}(\hat{\boldsymbol{x}}^*) = \hat{x}_k^* \left(I_0 \sum_{i=1}^{M} w_{ik} e^{-\sum_{j=1}^{N} w_{ij} \hat{x}_j^*} - \sum_{i=1}^{M} I_i w_{ik} \right) = 0 \tag{1.54}$$

$$\Leftrightarrow \quad \hat{x}_k^* I_0 \sum_{i=1}^{M} w_{ik} e^{-\sum_{j=1}^{N} w_{ij} \hat{x}_j^*} - \hat{x}_k^* \sum_{i=1}^{M} I_i w_{ik} = 0 \tag{1.55}$$

$$\Leftrightarrow \quad \hat{x}_k^* = \frac{\hat{x}_k^* I_0}{\sum_{i=1}^{M} I_i w_{ik}} \sum_{i=1}^{M} w_{ik} e^{-\sum_{j=1}^{N} w_{ij} \hat{x}_j^*}. \tag{1.56}$$

In terms of the log-corrected projection data $p_i = -\ln\left(\frac{I_i}{I_0}\right)$, Eq. 1.56 can be rewritten as

$$\hat{x}_k^* = \hat{x}_k^* \frac{\sum_{i=1}^M w_{ik} e^{-\sum_{j=1}^N w_{ij}\hat{x}_j^*}}{\sum_{i=1}^M \; w_{ik} e^{-p_i}}. \tag{1.57}$$

From Eq. 1.57, a fixed point iteration scheme can be derived. Starting from a non-zero reconstruction $\boldsymbol{x}^{(0)} = (x_k^{(0)})$, each individual pixel is updated as

$$x_k^{(n+1)} = x_k^{(n)} \frac{\sum_{i=1}^M w_{ik} e^{-\sum_{j=1}^N w_{ij} x_j^{(n)}}}{\sum_{i=1}^M \; w_{ik} e^{-p_i}} \qquad k = 1, \ldots, N. \tag{1.58}$$

The iterative algorithm in Eq. 1.58 is the expectation maximization (EM) algorithm. In summary, the EM algorithm consists of the following steps:

1. Starting from the current estimate $\boldsymbol{x}^{(n)}$, simulate the projection data with $e^{-\sum_{j=1}^N w_{ij} x_j^{(n)}}$ for $i = 1, \ldots, M$.

2. Backproject the simulated projection data. For the pixel with index k, this is achieved by processing all rays, and for each ray adding the corresponding simulated projection value multiplied with the intersection length of that particular ray with pixel k, i.e.,

$$\sum_{i=1}^M w_{ik} e^{-\sum_{j=1}^N w_{ij} x_j^{(n)}}. \tag{1.59}$$

3. Also backproject the measured projection data, i.e., e^{-p_i} for $i = 1, \ldots, M$. In the kth pixel, this corresponds to

$$\sum_{i=1}^M \; w_{ik} e^{-p_i}. \tag{1.60}$$

4. Compare the backprojected simulated projection data to the backprojected measured projection data by calculating the ratio

$$\frac{\sum_{i=1}^M w_{ik} e^{-\sum_{j=1}^N w_{ij} x_j^{(n)}}}{\sum_{i=1}^M \; w_{ik} e^{-p_i}}. \tag{1.61}$$

5. This ratio is multiplied in a pixel-by-pixel manner with the current reconstruction $\boldsymbol{x}^{(n)}$ to produce an improved reconstruction $\boldsymbol{x}^{(n+1)}$. The algorithm returns to step 1 and terminates based on some stopping criterion.

In practice, the EM algorithm of Eq. 1.58 is rather unstable. Often some regularization term is added to the objective function in Eq. 1.46 in order to improve the algorithm's stability. This is discussed in the following section.

1.2.3.3 Maximum A Posteriori (MAP) reconstruction

In contrast to the EM algorithm, where the probability $P(\boldsymbol{I}|\boldsymbol{x}^*)$ is maximized, the posterior probability $P(\boldsymbol{x}^*|\boldsymbol{I})$ is maximized with respect to $\boldsymbol{x}^*$ in the maximum a posteriori (MAP) approach [47]. With Bayes' theorem, the posterior probability $P(\boldsymbol{x}^*|\boldsymbol{I})$ can be written in function of the conditional probability $P(\boldsymbol{I}|\boldsymbol{x}^*)$:

$$P(\boldsymbol{x}^*|\boldsymbol{I}) = \frac{P(\boldsymbol{I}|\boldsymbol{x}^*)P(\boldsymbol{x}^*)}{P(\boldsymbol{I})} = P(\boldsymbol{x}^*)\prod_{i=1}^{M}\frac{P(I_i|\boldsymbol{x}^*)}{P(I_i)}. \tag{1.62}$$

The product in Eq. 1.62 can be transformed into a sum by taking the logarithm

$$\ln\left(P(\boldsymbol{x}^*|\boldsymbol{I})\right) = \ln\left(P(\boldsymbol{x}^*)\right) + \sum_{i=1}^{M}\left\{\ln\left(P(I_i|\boldsymbol{x}^*)\right) - \ln\left(P(I_i)\right)\right\}. \tag{1.63}$$

Maximization of Eq. 1.63 over $\boldsymbol{x}^*$ is equivalent to minimizing $-\ln\left(P(\boldsymbol{x}^*|\boldsymbol{I})\right)$, i.e.,

$$\hat{\boldsymbol{x}}^* = \arg\min_{\boldsymbol{x}^*\geq\boldsymbol{0}}\left(\sum_{i=1}^{M}\left\{-\ln\left(P(I_i|\boldsymbol{x}^*)\right)\right\} - \ln\left(P(\boldsymbol{x}^*)\right)\right), \tag{1.64}$$

where the irrelevant terms in the objective function were ignored. With the result in Eq. 1.45 and by ignoring all irrelevant terms, the minimization problem in Eq. 1.64 can be rewritten as

$$\hat{\boldsymbol{x}}^* = \arg\min_{\boldsymbol{x}^*\geq\boldsymbol{0}}\left(\sum_{i=1}^{M}\left(I_i\sum_{j=1}^{N}w_{ij}x_j^* + I_0 e^{-\sum_{i=1}^{N}w_{ij}x_j^*}\right) - \ln\left(P(\boldsymbol{x}^*)\right)\right). \tag{1.65}$$

A MAP reconstruction method solves the optimization problem in Eq. 1.65. The term $-\ln\left(P(\boldsymbol{x}^*)\right)$ typically represents the state of knowledge about the object $\boldsymbol{x}^*$. It is often referred to as the regularization term $U(\boldsymbol{x}^*) = -\ln\left(P(\boldsymbol{x}^*)\right)$ (see also the discussion on regularization methods in Section 1.2.2.3). With this notation, the MAP optimization problem becomes

$$\hat{\boldsymbol{x}}^* = \arg\min_{\boldsymbol{x}^*\geq\boldsymbol{0}}\left(\sum_{i=1}^{M}\left(I_i\sum_{j=1}^{N}w_{ij}x_j^* + I_0 e^{-\sum_{i=1}^{N}w_{ij}x_j^*}\right) + U(\boldsymbol{x}^*)\right). \tag{1.66}$$

A connection with the optimization problem in Eq. 1.41 can be made as follows. First, rewrite Eq. 1.66 by introducing $p_i^* = \sum_{j=1}^{N} w_{ij} x_j^*$:

$$\hat{\boldsymbol{x}}^* = \arg \min_{\boldsymbol{x}^* \geq \mathbf{0}} \left(\sum_{i=1}^{M} \left(I_i p_i^* + I_0 e^{-p_i^*} \right) + U(\boldsymbol{x}^*) \right). \tag{1.67}$$

Next, calculate a second order Taylor expansion of $I_i p_i^* + I_0 e^{-p_i^*}$ around the log-corrected measured projection values p_i, which results in (see [1] for a detailed calculation)

$$\hat{\boldsymbol{x}}^* = \arg \min_{\boldsymbol{x}^* \geq \mathbf{0}} \left(\frac{1}{2} (\boldsymbol{p} - \boldsymbol{W}\boldsymbol{x}^*)^T \boldsymbol{D} (\boldsymbol{p} - \boldsymbol{W}\boldsymbol{x}^*) + U(\boldsymbol{x}^*) \right), \tag{1.68}$$

with $D = \text{diag}(I_1, I_2, \ldots, I_M)$. The only difference with Eq. 1.41 is the weighting by the diagonal matrix $\boldsymbol{D}$. Again, various choices can be made for the regularization term $U(\boldsymbol{x}^*)$ [26, 27, 28, 29]. The optimization problem in Eq. 1.68 or Eq. 1.66 can be optimized with a variety of algorithms [47, 27]. In general, algorithms based on the MAP principle are more stable than EM-like algorithms.

1.3 Dynamic computed tomography

In Section 1.1 and Section 1.2, the classical tomography model was described, which assumes the scanned object to remain unaltered throughout the entire data acquisition process. This assumption is no longer valid in dynamic computed tomography, where projections are acquired from a time-varying object. Each projection hence corresponds to a different instance of the time-varying object, as is illustrated in Fig. 1.12.

The most well-known application of dynamic CT can probably be found in medical CT. The motion of the heart and/or lungs cause tissue to deform during the imaging process, making the tomography problem a dynamic one [48]. In gated CT, projections are sorted into several phase bins and a reconstruction is generated separately for every separate phase bin. The sorting is typically performed with an external breathing or electrocardiogram (ECG) signal [49, 50, 51]. In order to improve reconstruction quality, the correlation of reconstructions at adjacent phases can be exploited by temporal regularization [52, 53]. Approaches without gating typically incorporate motion models into the reconstruction algorithm [54, 55]. Another medical dynamic CT application is perfusion CT, where a contrast bolus is injected into the patient's blood stream and the local concentration changes in an organ of interest (brain, lung, liver, etc.) are monitored [56]. The same type of dynamic problems are naturally also encountered in small animal imaging [57].

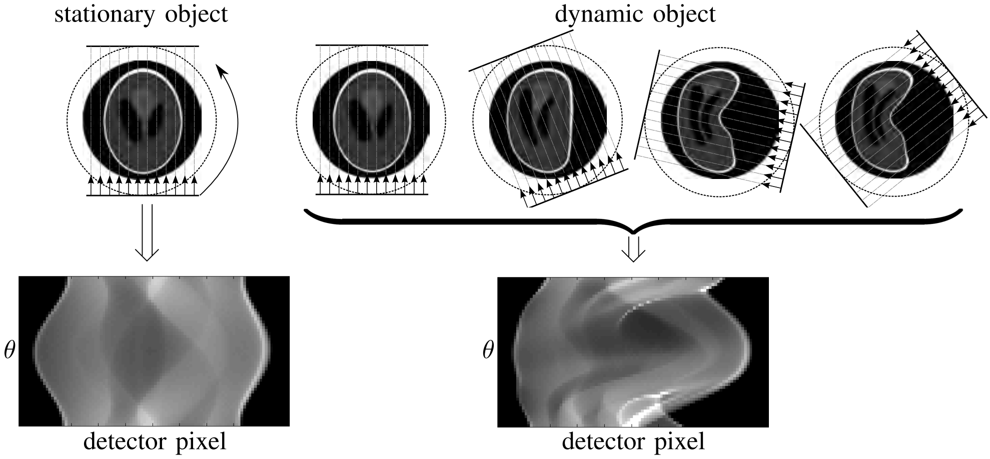


Figure 1.12: The effect of a time-varying object on the sinogram, illustrated with the Shepp-Logan phantom. In the left part of the figure, the sinogram corresponding to a stationary object is visualized. In the right part of the figure, the effect of a deformation in the same object on the sinogram is illustrated.

Dynamic CT problems are also frequently encountered in material research. In pressure tests, an incrementally increasing pressure is applied to a sample, while simultaneously monitoring the changes inside the sample [58, 59, 60]. Rigid sample motion, causing the projections to be misaligned, is a typical problem encountered in electron microscopy [61]. In fluid flow tomography, a (porous) sample is either immersed with fluid or drained, while continuously visualizing the internal dynamic processes [62, 63, 64].

Other examples include sample changes (while scanning) due to radiation damage [65, 66], the monitoring of rooth growth over time [67], imaging the solar corona [68], the investigation of micro-structural changes during development of internal flesh browning of apples [69], etc.

The remainder of this section focusses on introducing some basic concepts and notation that will be utilized in later chapters of this thesis. The literature will be covered more in depth in the introductory section of each chapter that concerns dynamic CT.

When dealing with time-varying objects, an extra dependency on time is introduced. Therefore, the object's attenuation coefficient function is represented by $f(x, y, t)$. The projection value model of Eq. 1.29 hence also includes the time dependency, i.e.,

$$p_i = \int_{-\frac{\triangle r}{2}}^{\frac{\triangle r}{2}} \int_{L(r_i,\theta_i)} f(x,y,t_i)\, ds dr', \tag{1.69}$$

where t_i represents the point in time at which the ith projection value was acquired.

In the algebraic setting, the dynamic object is typically represented as a time series of images $\boldsymbol{x}_r \in \mathbb{R}^N$, where each $r \in \{1,\ldots,R\}$ is the index referring to a particular point in time (i.e., a *time frame*) and R is the total number of time frames. The entire time series is represented by the vertical concatenation of $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_R$, i.e., $\tilde{\boldsymbol{x}} := (\boldsymbol{x}_1^T, \boldsymbol{x}_2^T, \ldots, \boldsymbol{x}_R^T)^T \in \mathbb{R}^{RN}$. To reconstruct this time series of images, projection data is acquired for each time frame by rotating source and detector multiple times around the object, or, equivalently, by rotating the object itself in between a fixed source and detector. Standard approaches then typically reconstruct the object at each time frame individually solely based on the projection data corresponding to a single 180° or 360° rotation. For every $r \in \{1,\ldots,R\}$, let $\boldsymbol{p}_r \in \mathbb{R}^M$ be the measured projection data corresponding to the r^{th} time frame. Define $\boldsymbol{W} \in \mathbb{R}^{RM\times N}$ as the forward projection matrix that models all projection angles and $\boldsymbol{W}_r \in \mathbb{R}^{M\times N}$ as the submatrix of $\boldsymbol{W}$ that models the forward projection for the r^{th} time frame. Furthermore, let the full vector of measured projection data $\tilde{\boldsymbol{p}} \in \mathbb{R}^{RM}$ be the vertical concatenation of $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_R$ and $\tilde{\boldsymbol{W}}$ the block diagonal matrix consisting of blocks $\boldsymbol{W}_1, \boldsymbol{W}_2, \ldots, \boldsymbol{W}_R$. In summary, we have introduced the following notations:

$$\tilde{\boldsymbol{W}} := \begin{bmatrix} \boldsymbol{W}_1 & 0 & \cdots & 0 \\ 0 & \boldsymbol{W}_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{W}_R \end{bmatrix} \in \mathbb{R}^{RM\times RN}, \qquad \boldsymbol{W} := \begin{bmatrix} \boldsymbol{W_1} \\ \boldsymbol{W_2} \\ \vdots \\ \boldsymbol{W_R} \end{bmatrix} \in \mathbb{R}^{RM\times N} \tag{1.70}$$

and

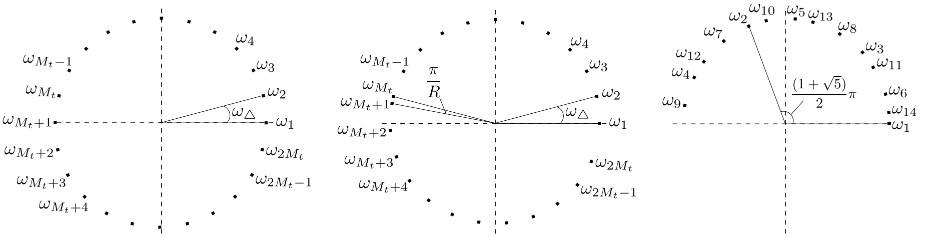
$$\tilde{\boldsymbol{p}} := \begin{bmatrix} \boldsymbol{p}_1 \\ \boldsymbol{p}_2 \\ \vdots \\ \boldsymbol{p}_R \end{bmatrix} \in \mathbb{R}^{RM}, \quad \tilde{\boldsymbol{x}} := \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_R \end{bmatrix} \in \mathbb{R}^{RN} \quad . \tag{1.71}$$

Analogously to the reconstruction problem for the static case, the goal in dynamic tomography is to find a reconstruction $\tilde{\boldsymbol{x}}$ that minimizes

$$||\tilde{\boldsymbol{W}}\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{p}}|| \tag{1.72}$$

for some norm $||\cdot||$. If only few projections are available per time frame, finding $\tilde{\boldsymbol{x}}$ such that Eq. 1.72 is minimal represents an ill-posed problem. This is mainly due to the large null space of the forward operator $\tilde{\boldsymbol{W}}$ and the noise in the measured projection data $\tilde{\boldsymbol{p}}$. Therefore, directly minimizing Eq. 1.72 for $\tilde{\boldsymbol{x}}$ (e.g., by applying SIRT separately to the projection data $\boldsymbol{p}_r$ for $r = 1, \ldots, R$) will typically result in a solution $\tilde{\boldsymbol{x}}$ which is fully dominated by noise. This problem can be alleviated by imposing constraints/models on the reconstruction $\tilde{\boldsymbol{x}}$, resulting in a smaller solution space, an approach which is employed in this thesis. The connection between the different time frames $\boldsymbol{x}_r$ (with $r = 1, \ldots, R$) is modelled, thereby basically reducing the degrees of freedom for the reconstruction algorithm in a model-compliant manner. The more accurately modelled reconstruction problem results in solutions that are less dominated by artifacts and more accurately represent the true underlying solution. This rationale is the basic development principle for the reconstruction algorithms that are introduced in the chapters on dynamic CT, i.e., Chapter 3, 4 and 5.

As was suggested in the previous paragraph, modelling the connection between the different time frames can be beneficial. If the information from reconstructions at different time frames is combined, it becomes important to avoid redundant information as much as possible. This can be achieved by changing the conventional acquisition angle selection schemes to more advanced. In what follows, three different ways of selecting the acquisition angles for the projection data corresponding to each time frame are introduced. Let M_t denote the number of projections acquired per time frame and M_d the number of detector elements per projection (this implies that the total number of projection values per time frame is given by $M = M_t M_d$). Next, denote the projection angle for the lth projection as ω_l ($l = 1, \ldots, M_t R$). This notation allows us to define the following three angle selection schemes, which are visualized in Fig. 1.13:



(a) Conventional decomposition **(b)** Binary decomposition **(c)** Golden ratio decomposition

Figure 1.13: Illustration of three different approaches to select the acquisition angles in dynamic CT.

1. *Conventional decomposition*: For each 180° rotation, the same (equiangular) projection angles are selected. This approach is the most widely used and suffices if the scanned object is reconstructed independently at each time frame. The angles are selected as follows

   $$\omega_l = (l-1)\omega_{\triangle} \quad l = 1, \dots, M_t R, \tag{1.73}$$

   where $\omega_{\triangle} = \pi / M_t$. This angle selection scheme is illustrated in Fig. 1.13a.

2. *Binary decomposition*: Analogously to the conventional decomposition, the acquisition angles in each 180° rotation are chosen equiangularly. However, a small angular shift is applied to the starting angle after each 180° rotation. The first M_t equiangular projection angles are chosen as $\omega_l = (l-1)\omega_{\triangle} \in [0, \pi - \omega_{\triangle}]$ for $l = 1, \dots, M_t$. The other projection directions are defined as

   $$\omega_{kM_t+l} = k(\pi - \omega_{\triangle}) + k\frac{\pi}{R} + (l-1)\omega_{\triangle} \tag{1.74}$$

   for $l = 1, \dots, M_t$, $k = 1, \dots, \lfloor R/M_t \rfloor$ and $kM_t + l \leq M_t R$. The angles $\omega_1, \omega_2, \dots, \omega_{2M_t}$ are schematically displayed in Fig. 1.13b. The small incremental step $k\frac{\pi}{R}$ in Eq. 1.74 ensures that there exist no two projection directions that are equal modulo π. This is beneficial for reconstruction algorithms that aim at combining information from different time frames. On the one hand, using the same projection angle (modulo π for parallel beam and modulo 2π for fan beam) more than once increases the signal-to-noise ratio (SNR). On the other hand, this will introduce redundant angular information, since the rays going through stationary regions (that is, regions inside the object that do not change over time) give the same projection values (up to noise). This situation is avoided by choosing the projection angles via Eq. 1.74. Furthermore, Eq. 1.74 guarantees that each subsequent M_t projection directions cover a range of approximately 180°, thus avoiding limited view artifacts.

3. *Golden ratio decomposition*: In the "golden ratio" scanning scheme, source and detector are rotated over a fixed angular step of $\triangle\omega = \pi(1+\sqrt{5})/2$ radians to determine the next acquisition angle [70, 71]. More precisely, the projection angle ω_l ($l = 1, \dots, RM_t$) is defined by

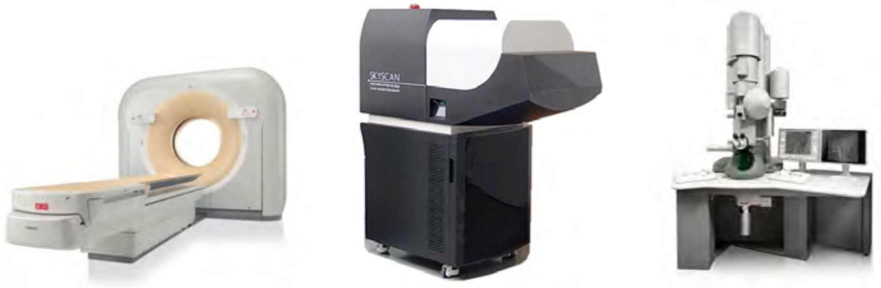
   $$\omega_l = \left((l-1)\frac{(1+\sqrt{5})}{2}\pi \right) \bmod \pi \ .$$

   In the binary decomposition, the user must select the number of projections per time frame before the experiment starts. The golden ratio scanning scheme is more flexible in the sense that it allows the user to select an arbitrary number of projections per time frame after the data acquisition, while still approximately covering equiangular positions over the entire angular range for each time frame [71]. This allows the user to balance the temporal and spatial resolution a posteriori, which is a useful property in many applications. Furthermore, the golden ratio decomposition also ensures that the same projection angle is never selected twice.

1.4 Applications

In this section, a short overview of different CT applications is given.

- **Medical CT**: The most well-known application of CT is in the medical sector, where medical CT devices allow for non-destructively imaging various parts of the patient's anatomy. CT is applied to virtually every part of the human body, including cerebral [72], dental [73], cardiac [74], abdominal [75] and pulmonary [76] imaging. It is also utilized during image-guided surgery [77] and radiation therapy [78]. An example of a typical scanner for medical CT application is given in Fig. 1.14a.



(a) Medical CT: the Philips Ingenuity CT Family scanner

(b) Biomedical CT: the Bruker-microCT SKYSCAN 1176 μCT scanner

(c) Material science: the FEI Tecnai G2 electron microscope

Figure 1.14: Three different CT imaging modalities. The resolution limit varies with the scanning modality: ranging from milimeter resolution scale in the medical CT scanner, to micrometer and nanometer resolution in the μCT scanner and electron microscope, respectively.

- **Biomedical CT**: In biomedical research, the effect of various diseases and drugs are tested by monitoring small animals by means of CT imaging [57]. A typical μCT scanner for biomedical application is presented in Fig. 1.14b.
- **Materials**: The discovery, study and design of materials (i.e., material science) has benefited greatly from computed tomography technology [79]. Besides regular μCT scanners, electron microscopes [80, 81, 82, 83] and synchrotron facilities are popular imaging modalities for material science. A typical electron microscope is visualized in Fig. 1.14c. Besides assessing the internal structure normally, the materials are typically tested under various circumstances, e.g., with pressure tests [84] or thermal cycling [85]. The techniques are applied in geological sciences [86] and petroleum research [87].
- **Other**: In industry, CT scanning also finds applications for inspection of products, including flaw detection, metrology and reverse engineering applications [88, 89]. Another application can be found in automatic control and explosive detection in luggage [90, 91].

  In an exciting application on a large scale, CT even makes it possible to reconstruct the corona of the sun [68] and the surface of the earth [92].

References

[1] T. M. Buzug, *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT.* Springer-Verlag, 2008.

[2] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging.* Society of Industrial and Applied Mathematics, 2001.

[3] W. C. Röntgen, “Über eine neue art von strahlen,” *Annalen der Physik*, vol. 300, no. 1, pp. 1–11, 1898.

[4] H. A. Duisterwinkel, J. K. van Abbema, M. J. van Goethem, R. Kawachimaru, L. Paganini, E. R. van der Graaf, and S. Brandenburg, “Spectra of clinical ct scanners using a portable compton spectrometer,” *Medical Physics*, vol. 42, no. 4, pp. 1884–1894, 2015.

[5] O. Hemberg, M. Otendal, and H. Hertz, “Liquid-metal-jet anode electron-impact X-ray source,” *Applied Physics Letters*, vol. 83, no. 7, pp. 1483–1485, 2003.

[6] H. Winick, *Synchrotron radiation sources: a primer*, vol. 1. World Scientific, 1995.

[7] J. T. Bushberg, J. A. Seibert, E. M. Leidholdt, and J. M. Boone, *The essential physics of medical imaging*, ch. 3. Lippincott Williams & Wilkins, 3 ed., 2011.

[8] F. Kelcz, P. M. Joseph, and S. K. Hilal, “Noise considerations in dual energy CT scanning,” *Medical Physics*, vol. 6, no. 5, pp. 418–425, 1979.

[9] J. P. Schlomka, E. Roessl, R. Dorscheid, S. Dill, G. Martens, T. Istel, C. Bäumer, C. Herrmann, R. Steadman, G. Zeitler, A. Livne, and R. Proksa, “Experimental feasibility of multi-energy photon-counting k-edge imaging in pre-clinical computed tomography,” *Physics in Medicine and Biology*, vol. 53, no. 15, p. 4031, 2008.

[10] M. Heideman, D. Johnson, and C. Burrus, “Gauss and the history of the fast fourier transform,” *ASSP Magazine, IEEE*, vol. 1, pp. 14–21, October 1984.

[11] D. M. Pelt and K. J. Batenburg, “Fast tomographic reconstruction from limited data using artificial neural networks,” *IEEE Transactions on Image Processing*, vol. 22, pp. 5238–5251, Dec 2013.

[12] K. J. Batenburg and L. Plantagie, “Fast approximation of algebraic reconstruction methods for tomography,” *IEEE Transactions on Image Processing*, vol. 21, pp. 3648–3658, Aug 2012.

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

[14] H. Kudo, T. Rodet, F. Noo, and M. Defrise, “Exact and approximate algorithms for helical cone-beam CT,” *Physics in Medicine and Biology*, vol. 49, no. 13, p. 2913, 2004.

[15] S. Matej and R. Lewitt, “Image representation and tomographic reconstruction using spherically-symmetric volume elements,” in *Nuclear Science Symposium and Medical Imaging Conference, 1992., Conference Record of the 1992 IEEE*, pp. 1191–1193 vol.2, Oct 1992.

[16] P. M. Joseph, “An improved algorithm for reprojecting rays through pixel images,” *IEEE Transactions on Medical Imaging*, vol. 1, pp. 192–196, Nov 1982.

[17] L. A. Shepp and B. F. Logan, “The Fourier reconstruction of a head section,” *IEEE Transactions on Nuclear Science*, vol. 21, pp. 21–43, June 1974.

[18] J. Gregor and T. Benson, “Computational analysis and improvement of SIRT.,” *IEEE Transactions on Medical Imaging*, vol. 27, pp. 918–24, Jan. 2008.

[19] M. Jiang and G. Wang, “Convergence studies on iterative algorithms for image reconstruction,” *IEEE Transactions on Medical Imaging*, vol. 22, pp. 569–579, May 2003.

[20] R. Gordon, R. Bender, and G. T. Herman, “Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography,” *Journal of Theoretical Biology*, vol. 29, no. 3, pp. 471 – 481, 1970.

[21] H. Guan and R. Gordon, “A projection access order for speedy convergence of ART (algebraic reconstruction technique): a multilevel scheme for computed tomography,” *Physics in Medicine and Biology*, vol. 39, no. 11, p. 2005, 1994.

[22] H. Guan and R. Gordon, “Computed tomography using algebraic reconstruction techniques (ARTs) with different projection access schemes: a comparison study under practical situations,” *Physics in Medicine and Biology*, vol. 41, no. 9, p. 1727, 1996.

[23] A. H. Andersen and A. C. Kak, “Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm,” *Ultrasonic Imaging*, vol. 6, no. 1, pp. 81–94, 1984.

[24] H. M. Hudson and R. S. Larkin, “Accelerated image reconstruction using ordered subsets of projection data,” *IEEE Transactions on Medical Imaging*, vol. 13, pp. 601–609, Dec 1994.

[25] C. C. Paige and M. A. Saunders, “LSQR: An algorithm for sparse linear equations and sparse least squares,” *ACM Transactions on Mathematical Software*, vol. 8, pp. 43–71, Mar. 1982.

[26] A. Chambolle, “An algorithm for total variation minimization and applications,” *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1-2, pp. 89–97, 2004.

[27] J. Tang, B. E. Nett, and G.-H. Chen, “Performance comparison between total variation (TV)-based compressed sensing and statistical iterative reconstruction algorithms,” *Physics in Medicine and Biology*, vol. 54, no. 19, p. 5781, 2009.

[28] X. Zhang, M. Burger, X. Bresson, and S. Osher, “Bregmanized nonlocal regularization for deconvolution and sparse reconstruction,” *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 253–276, 2010.

[29] Y. Lou, X. Zhang, S. Osher, and A. Bertozzi, “Image recovery via nonlocal operators,” *Journal of Scientific Computing*, vol. 42, no. 2, pp. 185–197, 2010.

[30] K. J. Batenburg, S. Bals, J. Sijbers, C. Kübel, P. A. Midgley, J. C. Hernandez, U. Kaiser, E. R. Encina, E. A. Coronado, and G. Van Tendeloo, “3D imaging of nanomaterials by discrete tomography,” *Ultramicroscopy*, vol. 109, no. 6, pp. 730–740, 2009.

[31] K. J. Batenburg and J. Sijbers, “DART: a practical reconstruction algorithm for discrete tomography,” *IEEE Transactions on Image Processing*, vol. 20, no. 9, pp. 2542–2553, 2011.

[32] G. T. Herman and A. Kuba, *Advances in discrete tomography and its applications.* Springer Science & Business Media, 2008.

[33] W. van Aarle, K. J. Batenburg, and J. Sijbers, "Optimal threshold selection for segmentation of dense homogeneous objects in tomographic reconstructions," *IEEE Transactions on Medical Imaging*, vol. 30, pp. 980–989, April 2011.

[34] T. Roelandts, K. J. Batenburg, E. Biermans, C. Kübel, S. Bals, and J. Sijbers, "Accurate segmentation of dense nanoparticles by partially discrete electron tomography," *Ultramicroscopy*, vol. 114, pp. 96 – 105, 2012.

[35] E. Meyer, R. Raupach, M. Lell, B. Schmidt, and M. Kachelrieß, "Normalized metal artifact reduction (NMAR) in computed tomography," *Medical Physics*, vol. 37, no. 10, pp. 5482–5493, 2010.

[36] E. Meyer, R. Raupach, M. Lell, B. Schmidt, and M. Kachelrieß, "Frequency split metal artifact reduction (FSMAR) in computed tomography," *Medical Physics*, vol. 39, no. 4, pp. 1904–1916, 2012.

[37] E. Y. Sidky, D. N. Kraemer, E. G. Roth, C. Ullberg, I. S. Reiser, and X. Pan, "Analysis of iterative region-of-interest image reconstruction for X-ray computed tomography," *Journal of Medical Imaging*, vol. 1, no. 3, p. 031007, 2014.

[38] T.-C. Hsung and D. Lun, "New sampling scheme for region-of-interest tomography," *IEEE Transactions on Signal Processing*, vol. 48, pp. 1154–1163, Apr 2000.

[39] K. P. Anoop and K. Rajgopal, "Image reconstruction with laterally truncated projections in helical cone-beam CT: Linear prediction based projection completion techniques," *Computerized Medical Imaging and Graphics*, vol. 33, no. 4, pp. 283 – 294, 2009.

[40] J. Cant, W. J. Palenstijn, G. Behiels, and J. Sijbers, "Modeling blurring effects due to continuous gantry rotation: Application to region of interest tomography," *Medical Physics*, vol. 42, no. 5, pp. 2709–2717, 2015.

[41] W. J. Palenstijn, K. J. Batenburg, and J. Sijbers, "Performance improvements for iterative electron tomography reconstruction using graphics processing units (GPUs)," *Journal of Structural Biology*, vol. 176, no. 2, pp. 250 – 253, 2011.

[42] W. J. Palenstijn, K. J. Batenburg, and J. Sijbers, "The ASTRA tomography toolbox," in *13th International Conference on Computational and Mathematical Methods in Science and Engineering, CMMSE*, 2013.

[43] W. van Aarle, W. J. Palenstijn, J. De Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers, "The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, vol. 157, pp. 35 – 47, 2015.

[44] A. Rockmore and A. Macovski, "A maximum likelihood approach to transmission image reconstruction from projections," *IEEE Transactions on Nuclear Science*, vol. 24, pp. 1929–1935, June 1977.

[45] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," *Journal of Computer Assisted Tomography*, vol. 8, no. 2, pp. 306–16, 1984.

[46] K. Lange, M. Bahn, and R. Little, "A theoretical study of some maximum likelihood algorithms for emission and transmission tomography," *IEEE Transactions on Medical Imaging*, vol. 6, pp. 106–114, June 1987.

[47] K. Lange and J. Fessler, "Globally convergent algorithms for maximum a posteriori transmission tomography," *IEEE Transactions on Image Processing*, vol. 4, pp. 1430–1438, Oct 1995.

[48] S. A. Nehmeh and Y. E. Erdi, "Respiratory motion in positron emission tomography/computed tomography: a review," in *Seminars in nuclear medicine*, vol. 38, pp. 167–176, Elsevier, 2008.

[49] S. S. Vedam, P. J. Keall, V. R. Kini, H. Mostafavi, H. P. Shukla, and R. Mohan, "Acquiring a four-dimensional computed tomography dataset using an external respiratory signal.," *Physics in Medicine and Biology*, vol. 48, pp. 45–62, Jan. 2003.

[50] D. A. Low, M. Nystrom, E. Kalinin, P. Parikh, J. F. Dempsey, J. D. Bradley, S. Mutic, S. H. Wahab, T. Islam, G. Christensen, D. G. Politte, and B. R. Whiting, "A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing," *Medical Physics*, vol. 30, no. 6, p. 1254, 2003.

[51] K. Nieman, "Cardiac CT in clinical practice," in *Cardiac CT*, pp. 31–39, Springer, 2014.

[52] X. Jia, Y. Lou, B. Dong, Z. Tian, and S. Jiang, "4D computed tomography reconstruction from few-projection data via temporal non-local regularization.," *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, vol. 13, pp. 143–50, Jan. 2010.

[53] H. Gao, J.-F. Cai, Z. Shen, and H. Zhao, "Robust principal component analysis-based four-dimensional computed tomography.," *Physics in Medicine and Biology*, vol. 56, pp. 3181–98, June 2011.

[54] S. Rit, D. Sarrut, and L. Desbat, "Comparison of analytic and algebraic methods for motion-compensated cone-beam CT reconstruction of the thorax.," *IEEE Transactions on Medical Imaging*, vol. 28, pp. 1513–25, Oct. 2009.

[55] G. Van Eyndhoven, J. Sijbers, and K. J. Batenburg, "Combined motion estimation and reconstruction in tomography," in *Computer Vision ECCV 2012. Workshops and Demonstrations* (A. Fusiello, V. Murino, and R. Cucchiara, eds.), vol. 7583 of *Lecture Notes in Computer Science*, pp. 12–21, Springer Berlin Heidelberg, 2012.

[56] K. A. Miles and M. R. Griffiths, “Perfusion CT: a worthwhile enhancement?,” *The British Journal of Radiology*, vol. 76, no. 904, pp. 220–231, 2003.

[57] S. J. Schambach, S. Bag, L. Schilling, C. Groden, and M. A. Brockmann, “Application of micro-CT in small animal imaging,” *Methods*, vol. 50, no. 1, pp. 2–13, 2010.

[58] B. Shi, Y. Murakami, Z. Wu, J. Chen, and H. Inyang, “Monitoring of internal failure evolution in soils using computerization X-ray tomography,” *Engineering Geology*, vol. 54, pp. 321–328, Oct. 1999.

[59] E. N. Landis, T. Zhang, E. N. Nagy, G. Nagy, and W. R. Franklin, “Cracking, damage and fracture in four dimensions,” *Materials and Structures*, vol. 40, pp. 357–364, July 2006.

[60] J. Elliott, A. Windle, J. Hobdell, G. Eeckhaut, R. Oldman, W. Ludwig, E. Boller, P. Cloetens, and J. Baruchel, “In-situ deformation of an open-cell flexible polyurethane foam characterised by 3d computed microtomography,” *Journal of Materials Science*, vol. 37, no. 8, pp. 1547–1555, 2002.

[61] Z. Jing and F. Sachs, “Alignment of tomographic projections using an incomplete set of fiducial markers,” *Ultramicroscopy*, vol. 35, no. 1, pp. 37–43, 1991.

[62] M. Andrew, B. Bijeljic, and M. J. Blunt, “Pore-scale contact angle measurements at reservoir conditions using X-ray microtomography,” *Advances in Water Resources*, vol. 68, no. 0, pp. 24 – 31, 2014.

[63] A. Polak, D. Elsworth, H. Yasuhara, A. S. Grader, and P. M. Halleck, “Permeability reduction of a natural fracture under net dissolution by hydrothermal fluids,” *Geophys. Res. Lett.*, vol. 30, no. 20, 2003.

[64] D. Kazantsev, G. Van Eyndhoven, W. R. B. Lionheart, P. J. Withers, K. J. Dobson, S. A. McDonald, R. Atwood, and P. D. Lee, “Employing temporal self-similarity across the entire time domain in computed tomography reconstruction,” *Philosophical Transactions of the Royal Society A*, vol. 373, no. 2043, 2015.

[65] E. a. Stern, Y. Yacoby, G. T. Seidler, K. P. Nagle, M. P. Prange, A. P. Sorini, J. J. Rehr, and A. Joachimiak, “Reducing radiation damage in macromolecular crystals at synchrotron sources.,” *Acta crystallographica. Section D, Biological crystallography*, vol. 65, pp. 366–74, Apr. 2009.

[66] P. K. Luther, “Sample shrinkage and radiation damage of plastic sections,” in *Electron Tomography* (J. Frank, ed.), ch. 1, pp. 17–48, New York: Springer, second ed., 2007.

[67] S. R. Tracy, J. a. Roberts, C. R. Black, A. McNeill, R. Davidson, and S. J. Mooney, “The X-factor: visualizing undisturbed root architecture in soils using X-ray computed tomography.,” *Journal of experimental botany*, vol. 61, pp. 311–3, Jan. 2010.

[68] M. D. Butala, R. J. Hewett, R. A. Frazin, and F. Kamalabadi, "Dynamic three-dimensional tomography of the solar corona," *Solar Physics*, vol. 262, no. 2, pp. 495–509, 2010.

[69] E. Herremans, P. Verboven, E. Bongaers, P. Estrade, B. E. Verlinden, M. Wevers, M. L. Hertog, and B. M. Nicolai, "Characterisation of braeburn browning disorder by means of x-ray micro-ct," *Postharvest Biology and Technology*, vol. 75, pp. 114 – 124, 2013.

[70] T. Kohler, "A projection access scheme for iterative reconstruction based on the golden section," in *Nuclear Science Symposium Conference Record*, vol. 6, pp. 3961–3965, 2004.

[71] A. Kaestner, B. Munch, T. Pavel, and L. Butler, "Spatiotemporal computed tomography of dynamic processes," *Optical Engineering*, vol. 50, no. 12, p. 123201, 2011.

[72] E. G. Hoeffner, I. Case, R. Jain, S. K. Gujar, G. V. Shah, J. P. Deveikis, R. C. Carlos, B. G. Thompson, M. R. Harrigan, and S. K. Mukherji, "Cerebral perfusion CT: Technique and clinical applications 1," *Radiology*, vol. 231, no. 3, pp. 632–644, 2004.

[73] A. Gahleitner, G. Watzek, and H. Imhof, "Dental CT: imaging technique, anatomy, and pathologic conditions of the jaws," *European radiology*, vol. 13, no. 2, pp. 366–376, 2003.

[74] K. Nieman, "Cardiac CT in clinical practice," in *Cardiac CT*, pp. 29–37, Springer, 2014.

[75] R. Zeman, S. Fox, P. Silverman, W. Davros, L. Carter, D. Griego, D. Weltman, S. Ascher, and C. Cooper, "Helical (spiral) CT of the abdomen," *American Journal of Roentgenology*, vol. 160, no. 4, pp. 719–725, 1993.

[76] J. U. Schoepf, M. A. Kessler, C. T. Rieger, P. Herzog, E. Klotz, S. Wiesgigl, C. R. Becker, D. N. Exarhos, and M. F. Reiser, "Multislice CT imaging of pulmonary embolism," *European radiology*, vol. 11, no. 11, pp. 2278–2286, 2001.

[77] G. A. Brown, M. C. Willis, K. Firoozbakhsh, A. Barmada, C. L. Tessman, and A. Montgomery, "Computed tomography image guided surgery in complex acetabular fractures," *Clinical orthopaedics and related research*, vol. 370, pp. 219–226, 2000.

[78] D. A. Jaffray, J. H. Siewerdsen, J. W. Wong, and A. A. Martinez, "Flat-panel cone-beam computed tomography for image-guided radiation therapy," *International Journal of Radiation Oncology*Biology*Physics*, vol. 53, no. 5, pp. 1337 – 1349, 2002.

[79] J. Banhart, *Advanced tomographic methods in materials research and engineering.* Oxford University Press New York, 2008.

[80] P. A. Midgley and R. E. Dunin-Borkowski, “Electron tomography and holography in materials science,” *Nature materials*, vol. 8, no. 4, pp. 271–280, 2009.

[81] B. Goris, S. Bals, W. Van den Broek, E. Carbó-Argibay, S. Gómez-Graña, L. M. Liz-Marzán, and G. Van Tendeloo, “Atomic-scale determination of surface facets in gold nanorods,” *Nature materials*, vol. 11, no. 11, pp. 930–935, 2012.

[82] Z. Saghi and P. A. Midgley, “Electron tomography in the (s)tem: from nanoscale morphological analysis to 3d atomic imaging,” *Annual Review of Materials Research*, vol. 42, pp. 59–79, 2012.

[83] S. Van Aert, K. J. Batenburg, M. D. Rossell, R. Erni, and G. Van Tendeloo, “Three-dimensional atomic imaging of crystalline nanoparticles,” *Nature*, vol. 470, no. 7334, pp. 374–377, 2011.

[84] N. Lenoir, M. Bornert, J. Desrues, P. Bésuelle, and G. Viggiani, “Volumetric digital image correlation applied to X-ray microtomography images from triaxial compression tests on argillaceous rock,” *Strain*, vol. 43, no. 3, pp. 193–205, 2007.

[85] Y. Guan, Y. Gong, W. Li, J. Gelb, L. Zhang, G. Liu, X. Zhang, X. Song, C. Xia, Y. Xiong, H. Wang, Z. Wu, and Y. Tian, “Quantitative analysis of micro structural and conductivity evolution of Ni-YSZ anodes during thermal cycling based on nano-computed tomography,” *Journal of Power Sources*, vol. 196, no. 24, pp. 10601 – 10605, 2011.

[86] V. Cnudde and M. N. Boone, “High-resolution X-ray computed tomography in geosciences: A review of the current technology and applications,” *Earth-Science Reviews*, vol. 123, pp. 1–17, 2013.

[87] A. N. Golab, M. A. Knackstedt, H. Averdunk, T. Senden, A. R. Butcher, and P. Jaime, “3D porosity and mineralogy characterization in tight gas sandstones,” *The Leading Edge*, vol. 29, no. 12, pp. 1476–1483, 2010.

[88] J. P. Kruth, M. Bartscher, S. Carmignato, R. Schmitt, L. De Chiffre, and A. Weckenmann, “Computed tomography for dimensional metrology,” *CIRP Annals-Manufacturing Technology*, vol. 60, no. 2, pp. 821 – 842, 2011.

[89] L. D. Chiffre, S. Carmignato, J.-P. Kruth, R. Schmitt, and A. Weckenmann, “Industrial applications of computed tomography,” *CIRP Annals - Manufacturing Technology*, vol. 63, no. 2, pp. 655 – 677, 2014.

[90] R. Ying, Zhengrong Naidu and C. R. Crawford, “Dual energy computed tomography for explosive detection,” *Journal of X-ray Science and Technology*, vol. 14, no. 4, pp. 235–256, 2006.

[91] P. Jin, E. Haneda, K. Sauer, and C. A. Bouman, “A model-based 3D multi-slice helical CT reconstruction algorithm for transportation security application,” in *Second International Conference on Image Formation in X-Ray Computed Tomography, Salt Lake City, Utah, USA*, 2012.

[92] G. Nolet, *Seismic tomography: with applications in global seismology and exploration geophysics*, vol. 5. Springer Science & Business Media, 2012.

Part I

Stationary CT

2

Reconstruction method for improved porosity quantification of porous materials

Contents

This chapter has been published as:
G. Van Eyndhoven, M. Kurttepeli, C. J. Van Oers, P. Cool, S. Bals, K. J. Batenburg, and J. Sijbers, *"Pore REconstruction and Segmentation (PORES) method for improved porosity quantification of nanoporous materials"*, Ultramicroscopy, vol. 148, pp. 10–19, 2015.

2.1 Introduction

Although electron tomography provides valuable three-dimensional visualizations of the sample under interest, accurate quantification of pore sizes in nanoporous materials remains a difficult problem, especially if the pores are irregularly shaped.

Quantification of nanoporous materials is important in many applications in the field of sorption/separation or catalysis, in which size selectivity often plays an important role [1, 2]. This makes a reliable and accurate knowledge of the pore size distribution indispensable.

Microporous ($d < 2nm$) and mesoporous ($2nm < d < 50nm$) materials are usually characterized by N2-sorption experiments at a temperature of $77K$ [3]. However, the quantification of the pore size distribution based on these measurements is carried out using models that assume a regular pore size, i.e., cylindrical or slit-shaped pores. Unfortunately, no model is available for materials with irregular pores. In this publication, an alternative and reliable approach to determine the pore size in nanoporous materials is proposed.

Transmission electron microscopy (TEM) is an ideal technique to investigate nanoporous materials at a local scale, but conventional TEM is limited to providing two-dimensional (2D) projections of a three-dimensional (3D) microscopy sample [4]. To measure the pore size distribution, a 3D representation of the sample is required, which can be obtained using electron tomography. This technique combines the information of a tilt series of 2D TEM images in a 3D voxel-based reconstruction [5]. The quality of the 3D reconstruction is of critical importance, since it influences further quantification. Computing accurate reconstructions from TEM projection images with classical analytical algorithms such as filter back projection (FBP) [6] or algebraic algorithms like the simultaneous iterative reconstruction technique (SIRT) [7] is a difficult task, mainly because of two issues. First, the limited tilt range of the sample (usually about $\pm 75°$) causes elongation of the 3D reconstruction and smearing of the voxel values, often referred to as the "missing wedge" artifact. Secondly, the reconstruction quality also depends on the number of TEM projection images, which is often relatively small to avoid beam damage, especially for sensitive materials.

It has been shown recently that the quality of a 3D reconstruction can be improved by incorporating prior knowledge in the reconstruction process. Assuming that the sample contains just a few a priori known compositions, each occurring in homogeneous regions, the discrete algebraic reconstruction technique (DART) has been able to strongly reduce missing wedge artifacts [8, 9]. The partially discrete algebraic reconstruction technique (PDART) exploits the existence of dense homogeneous particles of which the grey value is known by incorporating this knowledge in the reconstruction algorithm, resulting in more accurate reconstruction quality

[10]. Other methods minimize the total variation of the reconstruction, where the sample is assumed to have a sparse gradient, i.e., the number of boundary pixels in the sample is relatively small compared to the total number of pixels [11, 12]. However, the prior knowledge assumptions incorporated in the reconstruction algorithms of the previous examples are not always applicable to nanoporous materials, since the reconstruction may consist of a continuous range of grey values with non-sparsity of the gradient image. In this chapter, we propose an approach that exploits a different kind of prior knowledge, which is related uniquely to porous materials: the existence of many local regions of void space.

After the reconstruction step, individual pores can be extracted. To that end, a segmentation step should be applied to separate the pores from the material matrix. Manually or automatically selecting global thresholds can produce satisfactory results if there is a clear separation between the background and the material matrix [13]. However, due to reconstruction artifacts, this separation is not straightforward in practice. In particular for the segmentation of pores that are small compared to the voxel size, this approach is error prone. It can therefore be expected that further analysis of the pores with individual pore statistics such as size, orientation, eccentricity, etc. will be strongly influenced by the results of the two previous steps, i.e., reconstruction and segmentation.

To overcome the limitations discussed above, we present a tailor-made, integral approach, for the reconstruction, segmentation, and quantification of porous nanomaterials: the PORES (“*POre REconstruction and Segmentation*”) algorithm. The PORES data processing chain outperforms conventional approaches, since it is optimized for nanoporous structures. The PORES processing chain starts by calculating a porous sample specific reconstruction with the new SUPPRESS (“*Simultaneous Update of Pore Pixels by iterative REconstruction and Simple Segmentation*”) algorithm. SUPPRESS reduces artifacts by exploiting prior knowledge about the porous structure of the material, while automatically classifying the interior of the pores. The PORES method continues by applying a watershed algorithm directly to the reconstruction, resulting in accurate segmentation of the pores. This segmentation permits accurate quantification of individual pores, which is employed to generate full sample pore statistics.

The PORES method is described in Section 2.2, starting with the reconstruction algorithm in Section 2.2.1 followed by the segmentation and quantification approach in Section 2.2.2. In Section 2.3, the method is validated with both simulation and real experiments. The chapter is concluded in Section 2.4.

2.2 Method

This section describes the entire PORES algorithm, which is displayed in the flowchart in Fig. 2.1. It consists of two parts: the reconstruction algorithm (described in Section 2.2.1 and displayed in the uppermost part of the flowchart in Fig. 2.1) and the segmentation and quantification (described in Section 2.2.2 and displayed in the bottommost part of the flowchart in Fig. 2.1).

2.2.1 Reconstruction

In this section, a novel reconstruction technique is described, which will be referred to as the SUPPRESS algorithm ("*Simultaneous Update of Pore Pixels by iterative REconstruction and Simple Segmentation*"). It exploits a prior that comes naturally for porous materials: the existence of many local regions of void space.

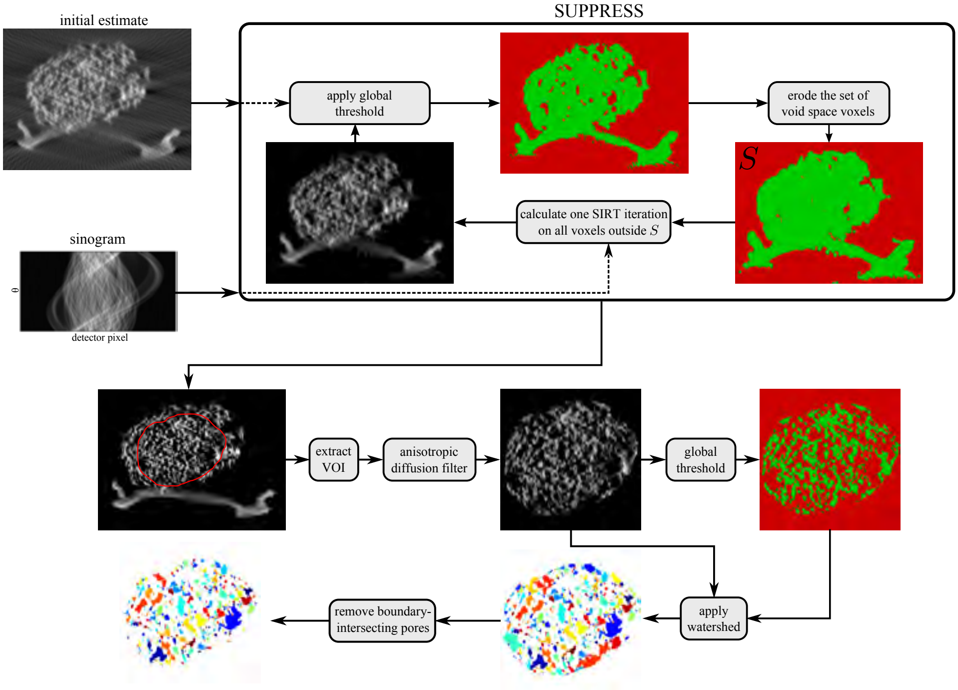


Figure 2.1: Flowchart of the entire PORES algorithm, which consist of the SUPPRESS reconstruction algorithm and the segmentation protocol. Red pixels indicate the estimate for void space, indicated by S in the flowchart. The SIRT update is calculated on the all pixels in the complement of S, i.e., on the green pixels in the illustration.

Algorithm 1 Pseudo-code for SUPPRESS algorithm

```
$\boldsymbol{x}^{(0)} \leftarrow$ initial SIRT reconstruction
for $k = 1, \dots, \#\text{iterations}$ do
    $S \leftarrow$ segment $\boldsymbol{x}^{(k-1)}$ by applying the global threshold $\tau$
    $S \leftarrow \text{erode}(S)$
    $\boldsymbol{x}^{(k)} \leftarrow \boldsymbol{x}^{(k-1)} + \boldsymbol{C}\boldsymbol{W}_S^T\boldsymbol{R}_S(\boldsymbol{p} - \boldsymbol{W}_S\boldsymbol{x}^{(k-1)})$
end for
```

The SUPPRESS algorithm is visualized in the uppermost part of the flowchart in Fig. 2.1 and pseudo-code is available in Algorithm 1.

In SUPPRESS, iterative update steps are combined with the prior knowledge that voxels inside void space, i.e., pores and background, should be homogenous and have a lower grey value than the support material. All steps in SUPPRESS are displayed in the uppermost part of the flowchart in Fig. 2.1. The algorithm starts by generating an initial SIRT reconstruction. Next, a conservative set of void space voxels, S, is estimated with the following two steps. First, the current reconstruction is segmented by a global thresholding operation, i.e., all pixels with a grey value smaller than a specified global threshold τ are selected as possible candidates for the S set. In our approach, τ was chosen as the grey value halfway between the grey value of void space and the smallest grey value of the support material. Next, this set is eroded, removing voxels at the boundary. This results in the set S, which is a conservative estimate for the void space voxels. The erosion operation reduces the chance that voxels of the material matrix are incorrectly classified as void space voxels (misclassified voxels are found typically on the edge between pore-space and material-space), which is essential for the next step in the algorithm. Based on the assumption that no material is present in the region defined by S, the reconstruction is then continued by applying a SIRT iteration solely to the voxels that belong to the complement of S, while keeping the voxels in S fixed at a grey level of 0 (i.e., no material). Mathematically, this corresponds to

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \boldsymbol{C}\boldsymbol{W}_S^T\boldsymbol{R}_S(\boldsymbol{p} - \boldsymbol{W}_S\boldsymbol{x}^{(k)}), \tag{2.1}$$

where k is the iteration number, $\boldsymbol{W}_S$ is obtained from the standard projection matrix $\boldsymbol{W}$ by putting the columns corresponding to the voxels in the set S to zero, $\boldsymbol{R}_S$ is the inverse row sums matrix corresponding to $\boldsymbol{W}_S$ and $\boldsymbol{C}$ the inverse column sum matrix corresponding to $\boldsymbol{W}$ (see Section 1.2.2.2). This procedure of identifying void space voxels and applying SIRT iterations to the remaining voxels is repeated until a fixed number of iterations is reached. In Appendix 2-A, the SUPPRESS algorithm is illustrated on a phantom at different intermediate iterations and the reader is provided with extra intuition in order to understand the different steps of the algorithm.

The key strength of SUPPRESS lies in the fact that the pores (and background) are identified during the reconstruction process itself. This results in a more refined estimate for void space at every iteration, which in turn will also results in more accurate reconstruction quality on the voxels not belonging to void space. This is due to the fact that the fixing of all voxels in S to the grey level of 0 actually decreases the number of variables in $\boldsymbol{Wx} = \boldsymbol{p}$, while maintaining the same number of equations as the original system, resulting in faster convergence and more accurate reconstruction quality.

The erosion and threshold parameters of the SUPPRESS algorithm control how many voxels are included in the S set. A large erosion size or a small threshold will result in a smaller intermediate void space estimate S, reducing the probability of falsely classifying material-space voxels as void space voxels during iterations, thereby increasing the algorithm's robustness. On the other hand, it is beneficial to include as many voxels into the S set as possible (without classifying material-space voxels as void space voxels), because this decreases the number of variables in $\boldsymbol{Wx} = \boldsymbol{p}$ even more, thereby giving faster convergence and more accurate reconstruction quality. Selecting the optimal parameters can be done with automatic procedures such as projection distance minimization schemes [14] or by histogram-based clustering methods such as Otsu's method [15]. This is, however, out of the scope of this chapter, and therefore these parameters were selected manually.

Also note that the void space estimate S will not contain all voxels corresponding to the pores; it will typically contain fewer pixels due to the erosion operation. This conservative estimate for the pore-space voxels is, however, not the final segmentation of the pores; it rather serves as an input for the segmentation of the individual pores in the next step which is described in the next section.

2.2.2 Segmentation and quantification

In the approach we present here, we did not only optimize the 3D reconstruction algorithm, but also the segmentation process. Our segmentation methodology is displayed in the bottommost part of the flowchart in Fig. 2.1. First, a volume of interest (VOI) is manually indicated. Next, an anisotropic diffusion filter [16, 17] (with the diffusion constant function as it is proposed in [17]) is applied to the reconstruction to reduce noise without compromising the edges. Since the SUPPRESS reconstruction contains little or no missing wedge artifacts, the pore-space can be segmented by global thresholding. Extracting individual pores is a crucial step, since the pore statistics depend strongly on a good inter-pore separation. A well-known approach for the removal of the artificially introduced connectivity in the case of regularly shaped pores, consists of applying the watershed algorithm to a distance transform of the segmented image [18]. Since this method is in-adequate for irregularly shaped pores, it needs to be adapted. The SUPPRESS reconstruction provides a conservative estimate of the set of voxels interior to the pores, i.e., the subset of S that does not correspond to the background. To separate individual pores, the watershed algorithm can then be applied directly on the filtered SUPPRESS reconstruction, by flooding regions starting at the grey value corresponding to void space and stop flooding at the global threshold value that was used for the pore-space segmentation. In the last step of the segmentation procedure all pores that coincide with the boundary of the VOI are removed. Once the segmentation has been computed, the equivalent spherical diameter [19] is determined for each individual pore, providing a quantitative measurement of the pore size distribution.

Note that the global threshold for final pore-space segmentation should typically be chosen larger than the threshold value in the SUPPRESS algorithm. The threshold in SUPPRESS should be chosen smaller, to prevent material space pixels to be classified as pore-space pixels during reconstruction.

Also note that the size of the global threshold parameter affects the size of the pore-space. However, because the SUPPRESS algorithm results in an accurate reconstruction with a clear distinction between void space and material-space, the associated pore-space segmentation is less sensitive to changes in the global threshold parameter (in comparison to other reconstruction algorithms such as SIRT or FBP).

2.3 Experiments and results

In this section, a range of experiments to evaluate our approach and its corresponding results are discussed. First, in Section 2.3.1, the TEM acquisition set-up for an aluminosilicate sample is described. Next, various simulation experiments are reported in Section 2.3.2. In Section 2.3.3, different figures of merit for the validation of our approach are introduced. In Section 2.3.4, the results of all experiments are reported. Finally, the PORES algorithm is applied to the real data in Section 2.3.5.

2.3.1 Material and acquisition: aluminosilicate sample

The material under study is an amorphous mesoporous aluminosilicate with a wormhole-like pore structure and irregularly shaped pores. The exact experimental conditions and material specifications have been elaborated in [20]. The TEM samples were prepared by applying drops of ethanol suspension of the powder sample on a carbon coated copper grid. The tomographic imaging was performed using high-angular annular dark-field scanning transmission electron microscopy

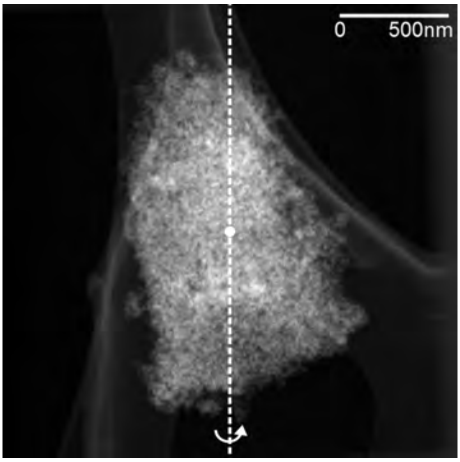
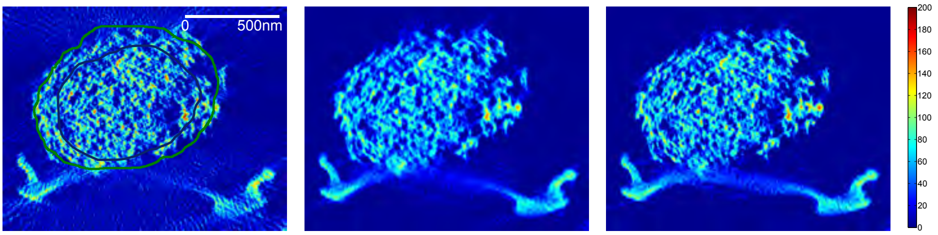


Figure 2.2: Example HAADF-STEM projection image (at 0° tilt angle) of the aluminosilicate sample. The tilt axis is indicated with the dashed line.

(HAADF-STEM) [21, 22]. A tilt series was acquired using a FEI Tecnai G2 electron microscope operated at 200 kV in combination with model 2020 Fischione Instruments tomography holder and the FEI XPlore3D software package. The HAADF-STEM images were acquired at the detector inner and outer collection angles of 72 mrad and 227 mrad with convergence semi-angle of 10 mrad. Each HAADF-STEM image contains 1024×1024 pixels, which have a 2.06 nm inter-pixel distance. The tilt series was collected over an angular range of $\pm 74^\circ$ with 2° increments and is displayed in Fig. 2.2. The projection images were aligned with an iterative cross-correlation algorithm together with a manual tilt axis adjustment implemented in FEI Inspect3D software [23]. The STEM image values were shifted in order to have a zero grey value corresponding to void space. To this end, a region where the electron beam clearly encountered void space was manually indicated in every STEM projection image and subsequently the average of the detector pixel values in this region was subtracted from the STEM image values. Reconstructions were calculated on a $1024 \times 1024 \times 1024$ voxel grid of voxel size 2.06 nm^3 in a slice-by-slice manner. The FBP, SIRT and SUPPRESS reconstructions are displayed in Fig. 2.3.

2.3.2 Material and acquisition: simulation phantoms

In this section, various simulation phantoms are described, each of which was chosen specifically to validate certain aspects of the SUPPRESS algorithm and



(a) (b) (c)

Figure 2.3: A central slice of reconstructions calculated with different algorithms. (a) FBP with small and large volume of interest (VOI) indicated. (b) SIRT with 300 iterations and a positivity constraint. (c) SUPPRESS with 100 initial SIRT iterations, 200 SUPPRESS iterations and a disk with a two pixel radius as morphological erosion operator.

Figure 2.4: Five examples of the 100 simulation phantoms yielding different pore concentrations.

the complete PORES approach. A first experiment is a direct validation of the SUPPRESS algorithm by a transmission tomography experiment with various simulation phantoms that contain different pore-space structures (see Section 2.3.2.1). In the second experiment, the nanoporous aluminosilicate TEM sample is simulated and the entire PORES processing chain is validated (see Section 2.3.2.2). For the final simulation experiment HAADF-STEM data was simulated using the CASINO software package (see Section 2.3.2.3), thereby introducing realistic noise into the experiment.

2.3.2.1 First set of simulation phantoms

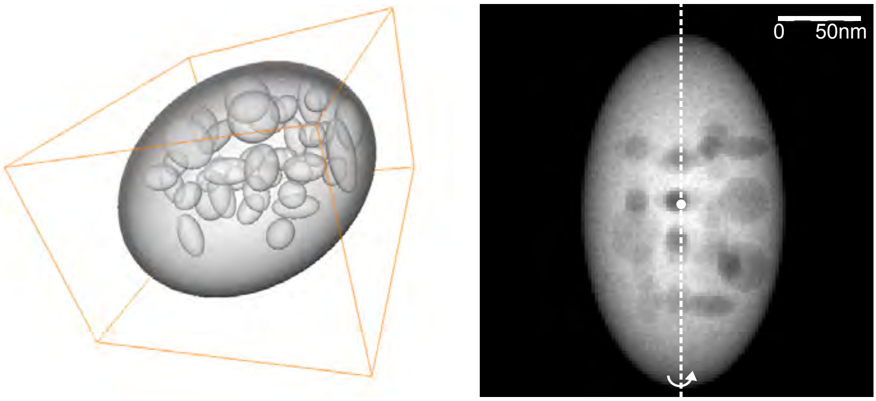
For validating the SUPPRESS algorithm for a series of different pore-space structures, 100 phantoms similar to the 5 phantoms shown in Fig. 2.4 were generated. Since these are 2D phantoms, mimicking a slice of a 3D object, their corresponding projections are 1D. The 100 phantoms were generated by introducing random pores in a fixed material phantom of cylindrical shape. The random pores were created by performing a global thresholding operation on different instances of 2D Perlin noise [24]. In total, 70 equiangular projections between ±72° were simulated with a strip kernel [25] and a higher resolution version of the phantom, i.e., on a 256×256 isotropic pixel grid. Poisson distributed noise was applied to the simulated projection data, i.e., each individual noise-free projection value was replaced by a value sampled from the Poisson distribution defined by an expectation value (and variance) equal to the noise-free projection value. Reconstructions were calculated on a 128×128 isotropic pixel grid and with a linear projection model [25].

2.3.2.2 Second simulation phantom

Direct validation of the SUPPRESS reconstruction algorithm and the subsequent segmentation and quantification of the pore size distribution on the real TEM data of the aluminosilicate is difficult, since no underlying accurate reference image is available. Therefore, a simulation phantom similar to the aluminosilicate and corresponding simulated projection data was created as follows. First, from the HAADF-STEM series of the aluminosilicate, a SUPPRESS reconstruction of one of the more central slices was calculated on a 1024×1024 pixel grid. The SUPPRESS reconstruction parameters were 100 initial SIRT iterations, 200 SUPPRESS iterations and a disk with a two pixel radius as morphological erosion operator. This reconstruction is displayed in Fig. 2.3c. Starting from this reconstruction, void space surrounding the sample was manually indicated and assigned a zero grey value. Subsequently, pores were segmented using the watershed algorithm as described above. The resulting pore-space pixels were also set to zero. The resulting 1024×1024 reference image is displayed in Fig. 2.7a. Based on the reference image, artificial projection data was generated along the same 75 projection angles as the real tilt series. Poisson distributed noise was applied to the projection data. With this approach, a reference image is available, and hence an elaborate validation can be performed.

2.3.2.3 Third simulation phantom

To validate the SUPPRESS algorithm under more realistic noise conditions, HAADF-STEM projection data was simulated with the CASINO Monte Carlo simulation software [26, 27] over an angular range of ±90° with 2° increments. The created sample is displayed in Fig. 2.5a and consists of pores of ellipsoid shape with different lengths for the semi-principal axes. The material matrix was set to contain weight fractions of 0.6% Al, 46.0% Si, 0.3% Na and 53.1% O. Reasonable values for these weight fractions were determined by applying an electron probe microanalyzer (EPMA) to the aluminosilicate sample (see Section 2.3.1). The experi-



(a) (b)

Figure 2.5: (a) The third simulation phantom. The outer contours are defined by an ellipsoid with semi-principal-axis lengths of 78 nm, 108 nm and 61.8 nm. (b) An example HAADF-STEM projection image (at 0° tilt angle) that was generated from the third simulation phantom with the CASINO software. The tilt axis is indicated by the dashed line.

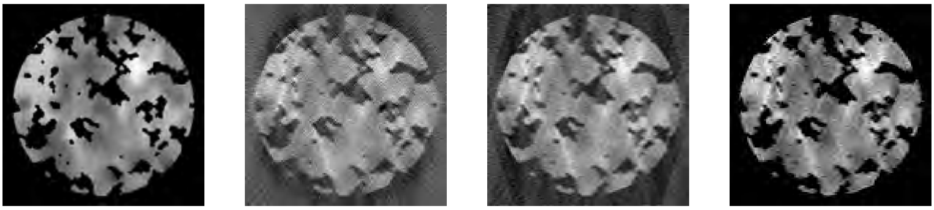
ment was set up with a 200 keV-microscope with a high-angular annular dark-field detector that collects electrons scattered between 72 and 227 mrad. Furthermore, a beam semi-angle of 10 mrad was assumed and 60000 electrons were simulated per detector pixel. Each simulated HAADF-STEM projection image was acquired by probing the porous sample with a pixel size of 2.06 nm^2 in a 116×116 grid. A simulated HAADF-STEM projection image is displayed in Fig. 2.5b. Reconstructions were calculated in a slice-by-slice fashion on a $116 \times 116 \times 116$ isotropic voxel grid.

2.3.3 Figures of merit

For validation, two different measures were utilized. A first figure of merit is the relative root mean square error (RRMSE). It is defined as

$$RRMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{\boldsymbol{x}}(i) - \boldsymbol{x}(i))^2}{\sum_{i=1}^{N}(\boldsymbol{x}(i))^2}} \quad , \tag{2.2}$$

where $\hat{\boldsymbol{x}} \in \mathbb{R}^N$ denotes the calculated reconstruction and $\boldsymbol{x} \in \mathbb{R}^N$ denotes the phantom used to generate the data. Other image quality measures such as the structural similarity index [28] were considered as well. However, because these



(a) Simulation phantom **(b)** FBP **(c)** SIRT **(d)** SUPPRESS

Figure 2.6: A simulation phantom and several different reconstructions. The SUPPRESS reconstruction has less artifacts than the FBP and the SIRT reconstruction.

measures gave similar results and the RRMSE measure is widely spread (thereby making a direct comparison of results possible), we chose to focus on the RRMSE measure throughout this thesis.

Finally, to validate the quality of the pore-space segmentation, the Number of Misclassified Pixels (NMP) was calculated, which is defined as the number of pixels that were falsely classified (with respect to the reference image) as either pore or material, reported in percentage of the total number of image pixels.

These validation measures were calculated on the pixel or voxel grid of the phantom, hence, whenever necessary, the reconstructions were upsampled to the resolution of the phantom.

2.3.4 Results of the simulation experiments

2.3.4.1 First set of simulation phantoms

The first set of simulation phantoms (of which 5 examples are displayed in Fig. 2.4) was reconstructed with FBP, SIRT and SUPPRESS. The SIRT algorithm was applied with 200 iterations and a positivity constraint. The SUPPRESS algorithm performed 100 initial SIRT iterations followed by 100 iterations of the main loop. The erosion operator was chosen to be a disk of 1 pixel radius. One particular instance of the simulation phantom and a FBP, SIRT and SUPPRESS reconstruction are shown in Fig. 2.6(a-d), respectively. Fig. 2.6 allows for a first visual assessment of the reconstruction results. The SUPPRESS has more detail and has little influence of the limited angular range over which the projection data was acquired, whereas the FBP and SIRT reconstruction clearly suffer from missing wedge artifacts, which would hamper further analysis. In total, the experiment was repeated 100 times, each time with a different phantom instance. The average

	FBP	SIRT	SUPPRESS
NMP	$10.75\% \pm 0.67\%$	$6.36\% \pm 0.91\%$	$4.12\% \pm 0.88\%$
RRMSE	0.394 ± 0.010	0.303 ± 0.019	0.245 ± 0.019

Table 2.1: Validation measures for experiment with the first simulation phantom, reported as *mean±std*.

results over all experiments are summarized in Table 2.1. Since the RRMSE assesses the image quality directly and NMP assesses the segmentation directly, it can be concluded from Table 2.1 that SUPPRESS performs better than SIRT and FBP.

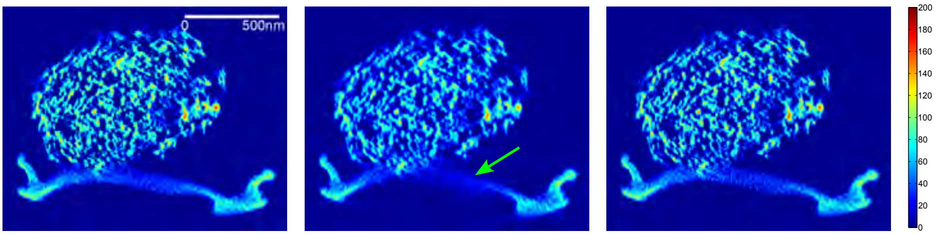
2.3.4.2 Second simulation phantom

From the simulated projections of the aluminosilicate simulation phantom, FBP, SIRT and SUPPRESS reconstructions were computed. The SUPPRESS reconstruction was calculated with the same parameters as described in Section 2.3.2.2 and the SIRT reconstruction was calculated using a positivity constraint and 300 iterations. The phantom and the SIRT and SUPPRESS reconstruction are displayed in Fig. 2.7a, 2.7b and 2.7c, respectively. It is clear that the SIRT reconstruction suffers from missing wedge artifacts. This is especially visible on the carbon grid, indicated by the green arrow on top of the SIRT reconstruction in Fig. 2.7b, which is smeared out in the vertical direction. Furthermore, the SIRT reconstruction has captured less details in comparison to the SUPPRESS reconstruction, which is clearly illustrated by observing the difference images with respect to the reference image in Fig. 2.8.

The experiment as described above was repeated 100 times (every time with new instances of the Poisson distributed noise) and the calculated statistics were averaged over all experiments. The results are summarized in Table 2.2. These quantitative results confirm the visual comparison that was made in Fig. 2.8: The SUPPRESS reconstruction results in the lowest NMP, indicating its ability to accurately capture pore-space, and the lowest RRMSE, indicating that the re-

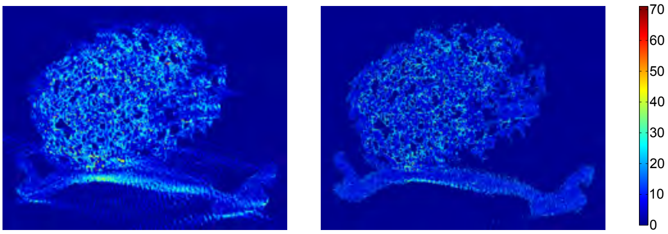
	FBP	SIRT	SUPPRESS
NMP	$1.75\% \pm 0.02\%$	$1.45\% \pm 0.03\%$	$1.37\% \pm 0.03\%$
RRMSE	1.0900 ± 0.0016	0.2900 ± 0.0002	0.2280 ± 0.0002

Table 2.2: Validation measures for experiment with the second simulation phantom, reported as *mean±std*.



(a) (b) (c)

Figure 2.7: Visual comparison between (a) the reference image, (b) the SIRT reconstruction and (c) the SUPPRESS reconstruction. The missing wedge artifacts, indicated on the SIRT reconstruction by the green arrow, are less pronounced in the SUPPRESS reconstruction.



(a) (b)

Figure 2.8: Absolute difference images between the reference image (displayed in Fig. 2.7a) and (a) the SIRT reconstruction and (b) the SUPPRESS reconstruction (displayed in Fig. 2.7b and Fig. 2.7c, respectively). Because of missing wedge artifacts, more erroneous pixels are observed in the SIRT reconstruction.

construction with the highest quality is generated by the SUPPRESS algorithm. Also, the full PORES procedure was applied and hence individual pores were extracted with the procedure as described in Section 2.2.2. Again, after repeating the experiment 100 times, an average histogram of equivalent circular diameter (being the 2D analog of the equivalent spherical diameter) was composed, shown in Fig. 2.9. For this histogram, all pore-sizes were considered, even pores corresponding to one single pixel. In practice, these measurements should not be considered, since they are inaccurate because the size of the feature is comparable to the pixel size. In this experiment, however, we were able to compare all pore sizes because of the availability of a ground truth reference image. The histogram shows that the estimation of small pores based on a regular SIRT reconstruction

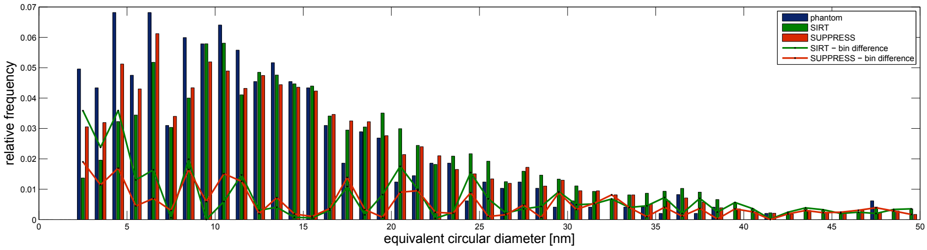


Figure 2.9: Result of PORES procedure in terms of a histogram of equivalent circular diameters. In the figure, SUPPRESS refers to the regular PORES procedure where the SUPPRESS reconstruction is used as input for the segmentation processing steps and SIRT refers to the PORES procedure where the SIRT reconstruction was used as input for the segmentation processing steps. The green and red curves indicate the absolute difference in relative frequency per bin, for the SIRT reconstruction and SUPPRESS reconstruction, respectively.

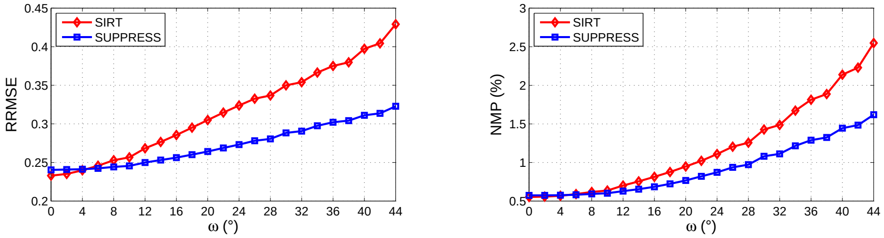


Figure 2.10: RRMSE (left) and NMP (right) as a function of missing wedge size ω for the third simulation phantom. For this experiment, the projections were simulated with the Monte Carlo method from the CASINO software package.

performs significantly worse than based on the SUPPRESS reconstruction.

2.3.4.3 Third simulation phantom

Tilt series with missing wedge were generated based on the tilt series that was simulated over the full angular range of ±90° with 2° increments. Subsets were taken from the full angular range dataset, representing the angular ranges $\pm(90-\omega)^\circ$ with 2° increments, where ω represents the size of the missing wedge. The SUPPRESS reconstructions were calculated with 100 initial SIRT iterations, 200 SUPPRESS iterations and a four pixel radius disk as morphological erosion operator. SIRT reconstructions were calculated using a positivity constraint and 300 iterations. Calculating the RRMSE and NMP as a function of ω results in Fig. 2.10. From the RRMSE plot in Fig. 2.10, one can notice that for a small missing wedge (i.e.,

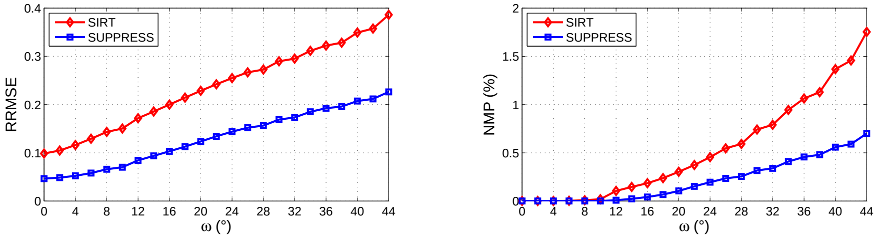
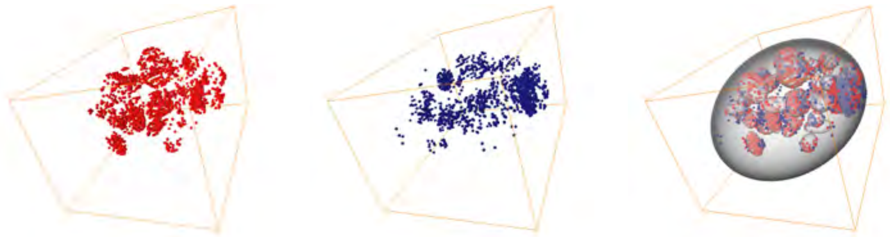


Figure 2.11: RRMSE (left) and NMP (right) as a function of missing wedge size ω for the third simulation phantom. For this experiment, the projections were simulated with a simple linear model and without noise.

$\omega < 4°$) the reconstruction quality of SIRT is slightly better than for the SUPPRESS reconstruction. This can be contributed to the fact that in SIRT the noise in the projection images is redistributed over the entire reconstruction domain, whereas the SUPPRESS reconstruction has to distribute it over the smaller set of voxels outside void space, which can result in a slightly larger RRMSE value in comparison to a SIRT reconstruction for small missing wedge values. Indeed, if the same experiment is repeated with projections generated from the phantom with a simple linear model and without noise (see Fig. 2.11), the SUPPRESS algorithm no longer suffers from this problem and clearly outperforms SIRT for every missing wedge size. Although, for the experiment with the realistic Monte Carlo simulated projections, the reconstruction quality of SIRT is slightly better in comparison to SUPPRESS for a small missing wedge (which is quantified by the RRMSE plot in Fig. 2.10), the segmentation quality (which is quantified by the NMP in Fig. 2.10) is the same. If the missing wedge increases (i.e., $\omega > 4°$) both reconstruction quality (RRMSE) and pore-space segmentation (NMP) is better for the SUPPRESS reconstruction than for SIRT reconstruction. With this increasing missing wedge size, the advantage of SUPPRESS in terms of reconstructing from projection data containing a missing wedge becomes increasingly apparent and the noise effects are no longer dominating.

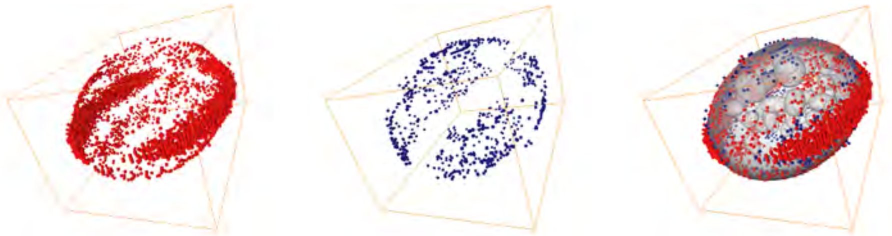
As an example, the result of the pore-space segmentation of SIRT and SUPPRESS reconstructions for a missing wedge of 28° is displayed in Fig. 2.12. To generate Fig. 2.12, the phantom (see Fig. 2.5a) was first voxelized onto the same $116 \times 116 \times 116$ voxel grid as the reconstructions. From the resulting phantom representation, the pore-space was compared to the segmented pore-space based on the SIRT and SUPPRESS reconstructions. Voxels that were misclassified in both segmentations (a total amount of 0.85% of all voxels) are not visualized, since they do not indicate the difference between both segmentations. From Fig. 2.12, it



(a) The uniquely misclassified voxels (0.13% of the total number of voxels) of the SIRT-based segmentation interior to the sample's edge (i.e., near the pores).

(b) The uniquely misclassified voxels (0.08% of the total number of voxels) of the SUPPRESS-based segmentation interior to the sample's edge (i.e., near the pores).

(c) Uniquely misclassified voxels of (a) and (b) in one visualization. The phantom of Fig. 2.5a is superimposed to clearly indicate where exactly the misclassifications are located.



(d) The uniquely misclassified voxels (0.28% of the total number of voxels) of the SIRT-based segmentation near the sample's edge and in the background.

(e) The uniquely misclassified voxels (0.04% of the total number of voxels) of the SUPPRESS-based segmentation near the sample's edge and in the background.

(f) Uniquely misclassified voxels of (d) and (e) in one visualization. The phantom of Fig. 2.5a is superimposed to clearly indicate where exactly the misclassifications are located.

Figure 2.12: Visualization of misclassified voxels of both the SIRT-based segmentation (red) and the SUPPRESS-based segmentation (blue) for the third simulation phantom with projection data containing a missing wedge of $\omega = 28^\circ$. Voxels that were misclassified by both pore-space segmentations (i.e., based on the SIRT and SUPPRESS reconstructions), representing a total amount of 0.85% of all 116^3 voxels, are not visualized, since they do not indicate the difference between the two methods. Red voxels refer to voxels that were misclassified uniquely by the segmentation based on the SIRT reconstruction while blue voxels refer to voxels that were uniquely misclassified by the SUPPRESS reconstruction. From the figures, it is obvious that the missing wedge (resulting in the horizontal smearing) has far more influence on the SIRT-based segmentation than on the SUPPRESS-based segmentation.

becomes clear that the difference in misclassified voxels is caused primarily by the missing wedge artifacts, since the horizontal smearing clearly has more influence on the SIRT-based segmentation than on the SUPPRESS-based segmentation.

2.3.5 Results for the aluminosilicate sample

The entire PORES 3D processing chain of reconstruction, segmentation and quantification was also applied to the full HAADF-STEM tilt series, i.e., all slices were reconstructed and the resulting 3D reconstruction was used for further processing. The SIRT and SUPPRESS reconstruction were calculated with the same parameters as for the second simulation phantom (see Section 2.3.2.2 and 2.3.4.2). A visual comparison for the FBP, SIRT and SUPPRESS reconstructions can be made in Fig. 2.3 (a)-(c). In this figure, it is noticeable that the carbon grid (on which the aluminosilicate sample was mounted) is smeared out in the vertical direction due to the missing wedge for the FBP and SIRT reconstruction, which is no longer the case in the SUPPRESS reconstruction. Also, the SUPPRESS reconstruction appears sharper in comparison to the SIRT and FBP reconstruction. To assess the robustness of the entire PORES processing chain, the equivalent spherical diameters in both a larger VOI and a smaller VOI were calculated. The shape of the VOI is indicated on top of the FBP reconstruction of Fig. 2.3a. Equivalent spherical diameters corresponding to a single voxel volume (i.e., an equivalent spherical diameter of 2.56 nm) were discarded, because these measurements are inaccurate in practice. The resulting histograms are displayed in Fig. 2.13. It is obvious that the histograms for the small VOI and the large VOI are highly similar. We can conclude that the quantification based on the segmentation of the SUPPRESS reconstruction is robust. As an illustration, the histogram obtained with a global thresholding operation applied on a basic SIRT reconstruction (a method that is employed often in practice) is displayed in Fig. 2.14. The histograms in this figure were generated by calculating relative frequencies on bins placed around integer pore diameters and fitting a smooth curve through it. Pore diameters corresponding to a single voxel (i.e., pore diameters of 2.56 nm) were removed from the results. It follows from all previous validation experiments (see Section 2.3) that the histogram obtained from a globally thresholded SIRT reconstruction is less accurate than the histogram obtained with the PORES algorithm. Furthermore, a comparison to the pore size distribution characterized by a nitrogen sorption experiment can also be done in Fig. 2.14. This nitrogen sorption measurement has been performed on a Quantachrome Quadrasorb SI unit, after degassing the sample under high vacuum conditions for a duration of 16 h at 473 K. Subsequently, the pore size distribution has been determined by applying the Barret-Joyner-Halenda (BJH) method on the desorption branch of the nitrogen sorption isotherm. It is clear that the PORES method is in better agreement with

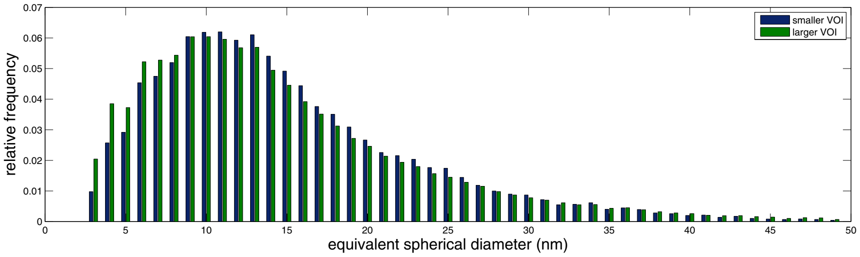


Figure 2.13: Histogram of equivalent spherical diameter for a larger VOI and a smaller VOI obtained from with the PORES method (based on the SUPPRESS reconstruction). The histograms have a large similarity, indicating the robustness of the quantification method.

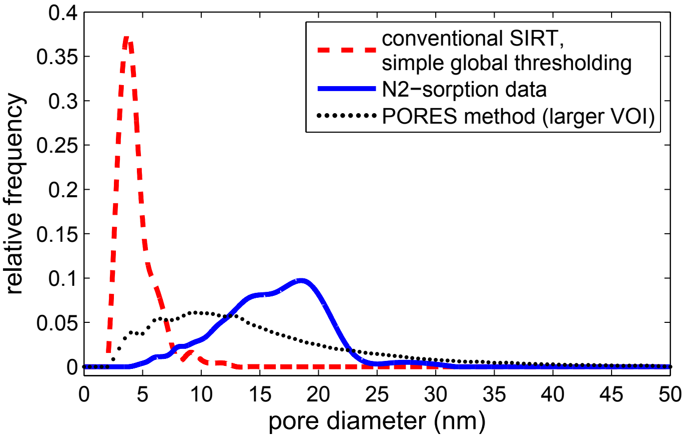


Figure 2.14: Comparison of histograms obtained with different methods: the red dashed curve is the histogram based on an estimate of pore-space that was calculated by applying a global thresholding operation to a conventional SIRT reconstruction, the blue curve is the histogram obtained with a nitrogen sorption experiment and the black dotted curve was calculated with the pores method.

the experimental nitrogen sorption data than the histogram obtained by a global thresholding operation applied on a SIRT reconstruction, although still a distinct discrepancy is detected between the PORES and the N2 sorption method. This can be appointed to the complicated unordered structure of the porous sample. Indeed, the pore size distribution based on the nitrogen sorption measurement has been calculated by the BJH model, assuming that the porous structure only contains cylindrical pores, which is not in full accordance with the actual situation.

Unfortunately, no better alternative is possible, since no calculation model has yet been developed to determine the pore size distribution based on nitrogen sorption measurements of an unordered aluminosilicate structure.

2.4 Discussion and conclusions

In conclusion, the PORES algorithm was proposed; it is an integral approach for the reconstruction, segmentation and quantification of nanoporous materials. As the proposed processing chain is tailored specifically for nanoporous materials, accurate quantification becomes possible. The first step, i.e., the SUPPRESS reconstruction, significantly reduced missing wedge artifacts in the reconstruction by the incorporation of prior knowledge in the reconstruction algorithm. Individual pores were reliably extracted, allowing for quantification by calculating individual pore statistics. The SUPPRESS and PORES algorithm were extensively validated with different experiments, varying in sample properties and the way in which data was simulated. In contrast to the standard N2-sorption method for determination of the pore size distribution, the PORES method does not assume cylindrical or slit-shaped pores. Furthermore, besides providing an overall pore size distribution (the result of the N2-sorption experiment), our method also allows for quantification of individual pores. Beside pore size, any other quantification is also possible, e.g., eccentricity, orientation, perimeter, etc. Furthermore, the information about interconnectivity between nanopores can also be extracted, which is important to improve mass transport and catalytic effectiveness in nanomaterials.

The algorithm for generating the iterative updates within the SUPPRESS algorithm, in our case SIRT, can be replaced by any other iterative algorithm. In this work we focused on SIRT, due to its ability to reconstruct a wide variety of objects, without making too many assumptions about the object at hand. If, for example, the object is expected to have a sparse gradient, iterative algorithms for minimizing the total variation could be more appropriate. Or in the case of severely noisy data, a maximum a posteriori (MAP) type of algorithm could possibly be a better choice. Also note that in the current implementation of the SUPPRESS algorithm, a pixel either belongs to pore-space or not. A possible addition would be to assign probabilities to each pixel, indicating the likelihood of the particular pixel belonging to pore-space. This information could be utilized while updating the reconstruction and further improved while iterations continue.

Appendix 2-A

In this appendix, the different steps in the SUPPRESS algorithm are illustrated on the phantom displayed in Fig. 2.15. From this 200 × 200 phantom, 40 parallel

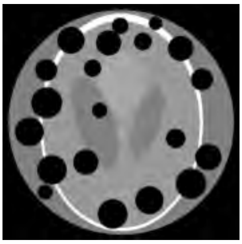


Figure 2.15: The phantom that was utilized in the flowchart of Fig. 2.16.

beam projections were simulated with a strip kernel over an angular range of ±75°. Next, the SUPPRESS reconstruction was calculated on a 100×100 pixel grid with a linear projection kernel, 200 iterations and a two-pixel radius disk as morphological erosion element. The different iterations are visualized in Fig. 2.16. In Fig. 2.16, the different steps in the SUPPRESS algorithm are visualized at iteration 1, 4, 8, 20 and 50. The first step of an iteration always consist of calculating a segmentation of pore-space based on a global threshold τ. Next, the S set, i.e., the conservative estimate of void space pixels, is calculated by applying an erosion operation on the segmented image of the previous step. An iteration is ended by calculating a SIRT update on all pixels that are not included in S. As is illustrated in Fig. 2.16, the estimate of pore-space becomes better as iterations continue. This is due to the fact that the reconstruction quality improves, which has a positive effect on the segmentation of void space as well, which again results in a better reconstruction, ans so on. After 50 iterations, almost all the pores have been found. The final segmentation result of this experiment is visualized in Fig. 2.17, where pore-space is segmented from both the SUPPRESS and a SIRT reconstruction by applying a global threshold on the final reconstruction result.

References

[1] F. Schuth, K. S. Sing, and J. Weitkamp, *Handbook of Porous Solids*, vol. 1. Wiley, 2002.

[2] G. Ertl, H. Knoezinger, F. Schuth, and J. Weitkamp, *Handbook of Heterogeneous Catalysis*, vol. 1. SIAM, 2008.

[3] S. Lowell, J. E. Shields, M. A. Thomas, and M. Thommes, *Characterization of porous solids and powders: surface area, pore size and density*, vol. 16. Kluwer Academic Publishers, 2004.

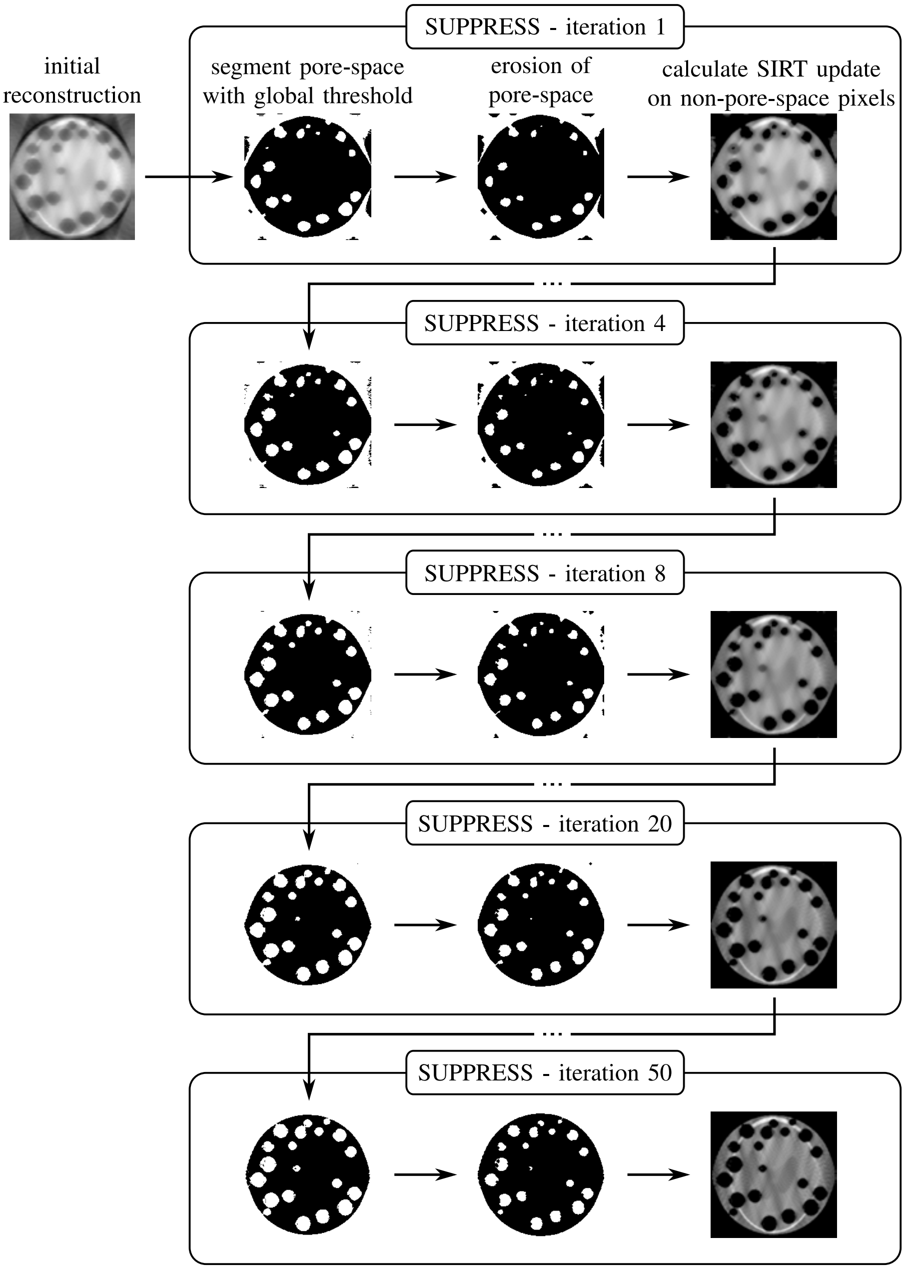
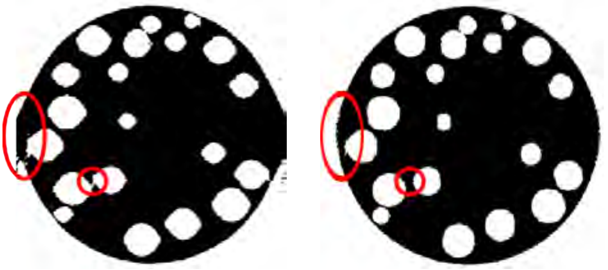
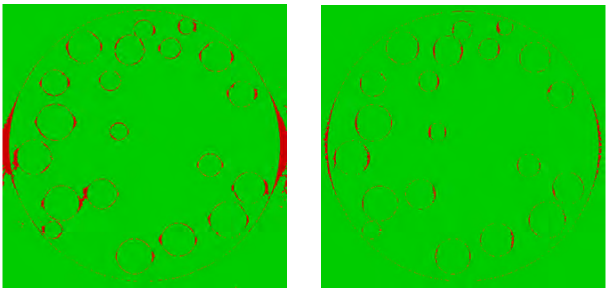


Figure 2.16: Flowchart of the SUPPRESS algorithm in which different iterations are displayed.



(a) SIRT segmentation **(b)** SUPPRESS segmentation



(c) Misclassified pixels in the SIRT-based segmentation **(d)** Misclassified pixels in the SUPPRESS-based segmentation

Figure 2.17: The segmentation results for the experiment reported in Appendix 2-A. (a) The SIRT-based segmentation is highly influenced by the missing wedge artifacts, as can be seen by observing the parts within the superimposed red circles. (b) The SUPPRESS-based segmentation is more accurate. (c-d) The misclassified pixels (green pixels were correctly classified and red pixels were wrongly classified) for the SIRT-based and the SUPPRESS-based segmentation indicate that the SUPPRESS reconstruction is more appropriate for the segmentation task.

[4] L. Reimer and H. Kohl, *Transmission electron microscopy: physics of image formation*, vol. 36. Springer, 2008.

[5] M. Weyland and P. A. Midgley, "Electron tomography," *Materials Today*, vol. 7, no. 12, pp. 32–40, 2004.

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

[7] J. Gregor and T. Benson, "Computational analysis and improvement of SIRT.," *IEEE Transactions on Medical Imaging*, vol. 27, pp. 918–24, Jan. 2008.

[8] K. J. Batenburg and J. Sijbers, "DART: a practical reconstruction algorithm for discrete tomography," *IEEE Transactions on Image Processing*, vol. 20, no. 9, pp. 2542–2553, 2011.

[9] K. J. Batenburg, S. Bals, J. Sijbers, C. Kübel, P. A. Midgley, J. C. Hernandez, U. Kaiser, E. R. Encina, E. A. Coronado, and G. Van Tendeloo, "3D imaging of nanomaterials by discrete tomography," *Ultramicroscopy*, vol. 109, no. 6, pp. 730–740, 2009.

[10] T. Roelandts, K. J. Batenburg, E. Biermans, C. Kübel, S. Bals, and J. Sijbers, "Accurate segmentation of dense nanoparticles by partially discrete electron tomography," *Ultramicroscopy*, vol. 114, pp. 96 – 105, 2012.

[11] B. Goris, W. Van den Broek, K. J. Batenburg, H. Heidari Mezerji, and S. Bals, "Electron tomography based on a total variation minimization reconstruction technique," *Ultramicroscopy*, vol. 113, pp. 120–130, Feb. 2012.

[12] R. Leary, Z. Saghi, P. A. Midgley, and D. J. Holland, "Compressed sensing electron tomography.," *Ultramicroscopy*, vol. 131, pp. 70–91, Aug. 2013.

[13] E. Biermans, L. Molina, K. Batenburg, S. Bals, and G. Van Tendeloo, "Measuring porosity at the nanoscale by quantitative electron tomography," *Nano letters*, vol. 10, no. 12, pp. 5014–5019, 2010.

[14] W. van Aarle, K. J. Batenburg, and J. Sijbers, "Optimal threshold selection for segmentation of dense homogeneous objects in tomographic reconstructions," *IEEE Transactions on Medical Imaging*, vol. 30, pp. 980–989, April 2011.

[15] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, pp. 62–66, Jan 1979.

[16] G. Gerig, O. Kubler, R. Kikinis, and F. a. Jolesz, "Nonlinear anisotropic filtering of MRI data.," *IEEE Transactions on Medical Imaging*, vol. 11, pp. 221–32, Jan. 1992.

[17] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 629–639, Jul 1990.

[18] F. Meyer, "Topographic distance and watershed lines," *Signal processing*, vol. 38, no. 1, pp. 113–125, 1994.

[19] B. Jennings and K. Parslow, "Particle size measurement: the equivalent spherical diameter," *Proceedings of the Royal Society of London A*, vol. 419, no. 1856, pp. 137–149, 1988.

[20] C. J. Van Oers, M. Kurttepeli, M. Mertens, S. Bals, V. Meynen, and P. Cool, "Zeolite β nanoparticles based bimodal structures: Mechanism and tuning of the porosity and zeolitic properties," *Microporous and Mesoporous Materials*, vol. 185, pp. 204–212, Feb. 2014.

[21] P. Midgley and M. Weyland, "3d electron microscopy in the physical sciences: the development of z-contrast and EFTEM tomography," *Ultramicroscopy*, vol. 96, no. 3, pp. 413 – 431, 2003.

[22] S. Pennycook, "Z-contrast transmission electron microscopy: direct atomic imaging of materials," *Annual Review of Materials Research*, vol. 22, no. 1, pp. 171–195, 1992.

[23] R. H. M. Schoenmakers, R. A. Perquin, T. F. Fliervoet, and W. Voorhout, "High resolution, high throughput electron tomography reconstruction," *Microscopy and Microanalysis*, vol. 11, pp. 312–313, 8 2005.

[24] K. Perlin, "Improving noise," *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 681–682, 2002.

[25] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Society of Industrial and Applied Mathematics, 2001.

[26] H. Demers, N. Poirier-Demers, D. Drouin, and N. de Jonge, "Simulating stem imaging of nanoparticles in micrometers-thick substrates," *Microscopy and Microanalysis*, vol. 16, no. 6, p. 795, 2010.

[27] H. Demers, N. Poirier-Demers, A. R. Couture, D. Joly, M. Guilmain, N. de Jonge, and D. Drouin, "Three-dimensional electron microscopy simulation with the CASINO Monte Carlo software," *Scanning*, vol. 33, no. 3, pp. 135–46, 2011.

[28] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

Part II

Dynamic CT

3

Region-based iterative reconstruction

Contents

This chapter has been published as:
G. Van Eyndhoven, K. J. Batenburg, and J. Sijbers, *"Region-based iterative reconstruction of structurally changing objects in CT"*, IEEE Transactions on Image Processing, vol. 23, no. 2, pp. 909–919, 2014.

3.1 Introduction

In computed tomography (CT), most reconstruction techniques assume that the object does not deform during the acquisition of projection data. If the scanned object is subject to deformation or structural changes, however, these techniques are no longer adequate and the reconstructed image will suffer from artefacts such as blurring.

Current approaches to account for deformation of the object can be classified into two categories. In a first class of methods, which we refer to as the *deformation-based techniques*, a deformation model is incorporated in the reconstruction process. Affine transformations can be modeled directly by adjusting the projection data [1] and subsequently using a standard reconstruction algorithm. Many a priori known invertible deformation models can be compensated for by modifying classical algorithms like FBP or SIRT [2, 3]. If no a priori deformation model is available, the deformation parameters can be estimated using a series of different techniques [1, 4, 5, 6, 7, 8]. A second class of methods generates reconstructions using different subsets of all available projection data, assuming that each of these subsets contains data acquired from a motionless object. We refer to this type of methods as *subset-reconstruction techniques.* In medical imaging, this technique is usually known as phase binning [9], where the subsets are obtained by ordering the projection data per phase, assuming a periodic motion. Alternatively, different subsequent scans can be performed in order to obtain the subsets, assuming the object remained unchanged during each scan. Examples of this technique can be found in non-destructive material tests [10], soil structure and water retention studies [11] or observations of root growth [12]. Reconstruction quality of standard subset-reconstruction techniques can be improved by enforcing similarity among the subsets with regularization strategies [13]. Another improvement can be made if a prior, high quality reconstruction is available, e.g. based on projection data acquired from the object before the changes have commenced. From this prior reconstruction, projections can be simulated and subtracted from the projection data acquired from the time-varying object. The changing volume can then be reconstructed by applying a sparse reconstruction technique to the resulting projection difference [14].

The deformation-based technique is only applicable to problems where continuous deformations deteriorate reconstruction quality, whereas the subset-reconstruction technique can only be applied to periodic motion or under the assumption that the object remains unchanged during the acquisition of projection data for each subset. The latter assumption is a severe restriction on the time resolution of the reconstruction, since reconstructing the scanned object per subset implies the acquisition of sufficient projection data per subset in order to avoid limited data reconstruction artifacts, which in turn implies that the acquisition time per subset is far from negligible. This limits the applicability of these techniques, as they cannot be applied to reconstruction setups where objects are subject to structural, discontinuous changes that happen in a faster time frame than the acquisition time needed for each subset. This type of changes occurs in a wide variety of reconstruction problems, e.g., the scanning of beam-sensitive samples, where regions in the object are damaged by the X-ray beam [15], microstructural investigations of solidification [16] and problems in the field of non-destructive testing, where sudden discontinuous cracks are formed inside the object [17, 10].

A straightforward dynamic imaging acquisition method that employs the subset-reconstruction technique, is an acquisition in which the source and detector rotate in a circular orbit multiple times around the object. In each time window for which the full angular range is covered, a 3D image is reconstructed. This approach is extensively used in medical imaging to obtain real-time reconstructions [18]. The same scanning technique could, however, be used in μCT, synchrotron tomography and other advanced lab CT setups as well. In such an approach, it is assumed that the object remains stationary during each time window that covers the full angular range. Evidently, this time window also defines the temporal resolution. Increasing the temporal resolution (by decreasing the number of projection angles or by decreasing the radiation exposure time needed to acquire projection data for one angular direction) will result in reconstructions of poor quality. This limits the temporal resolution for such a subset-reconstruction method from a hardware point of view. From a computational point of view, however, there is still room for improvement, as is illustrated by the techniques introduced in this chapter.

In many problems where structural changes complicate the reconstruction process, there is redundancy between projection data acquired at different time points. This redundancy is exploited by the techniques introduced in this chapter; we propose an iterative method that generates accurate reconstructions using limited projection data by assuming the existence of regions inside the object that remain constant over time. Since less projection data is needed, time resolution increases. Regularization methods are less suitable for handling structural changes, as they assume that every region in a time-varying object is similar over time, which is certainly not the case for structurally changing objects. Unlike regularization methods, the proposed method enforces similarity by combining iterative update steps over different projection-subset reconstructions. Also, in contrast to the approach introduced in [14], the proposed method does not depend on the prior knowledge of a high quality image of the object before the structural changes initiated. In this chapter, a region-based SIRT (rSIRT) algorithm is developed, where these stationary regions are assumed to be known. Next, the rSIRT algorithm is improved and incorporated in an optimization routine that automatically determines the stationary regions inside the object. The developed method was validated on simulation data and on data obtained from a controlled experimental μCT setup. Results show that the proposed method reduces the number of projections and thus a significant increase in time resolution is achieved.

In Section 3.2, the rSIRT algorithm is introduced, together with the iterative routine that incorporates the rSIRT algorithm in a projection distance minimization scheme that automatically estimates the parameters defining the stationary region. Simulation and experimental results are reported in Section 3.3. The chapter is concluded in Section 3.4.

3.2 Methods

In this section, we introduce the rSIRT algorithm, a modified version of the well known SIRT algorithm that combines traditional SIRT update steps over different regions and points in time. The rSIRT algorithm assumes that stationary regions within the object are given a priori. As in practice, this assumption does not hold, we propose a methodology to automatically estimate the stationary regions: a B-spline model for region description, together with the rSIRT algorithm, are incorporated in a weighted projection distance minimization scheme, that automatically calculates the parameters describing the stationary regions.

3.2.1 rSIRT

The scheme to calculate a single rSIRT iteration is displayed in the flowchart of Fig. 3.1. The phantom of Fig. 3.4a-d is used for visualization purposes. The scanned object is assumed to consist of stationary regions and regions that change over time, i.e., variable regions. Define $S \subset \{1, \ldots, N\}$ as the set of pixel indices that correspond to the stationary regions and define V as its complement, i.e., $V = \{1, \ldots, N\} \backslash S$ is the set of pixel indices corresponding to the object's variable regions. Let $\boldsymbol{I}_V \in \{0,1\}^{N \times N}$ be the binary diagonal matrix representing the operator that sets all pixels belonging to the stationary region to 0. Its diagonal elements are given by

$$\boldsymbol{I}_V(j,j) = \begin{cases} 1 & \text{if } j \in V \\ 0 & \text{otherwise} \end{cases} . \tag{3.1}$$

Define the binary diagonal matrix $\boldsymbol{I}_S \in \{0,1\}^{N \times N}$ analogously. It is the operator setting all pixels in the variable regions to 0. In this chapter, we let each time frame $r \in \{1, \ldots, R\}$ correspond to a single projection image taken at a specific acquisition angle ω_r (thus, $M_t = 1$, see Section 1.3). Therefore, the projection matrix $\boldsymbol{W}_r$ ($r \in \{1, \ldots, R\}$), introduced in Section 1.3, corresponds solely to the

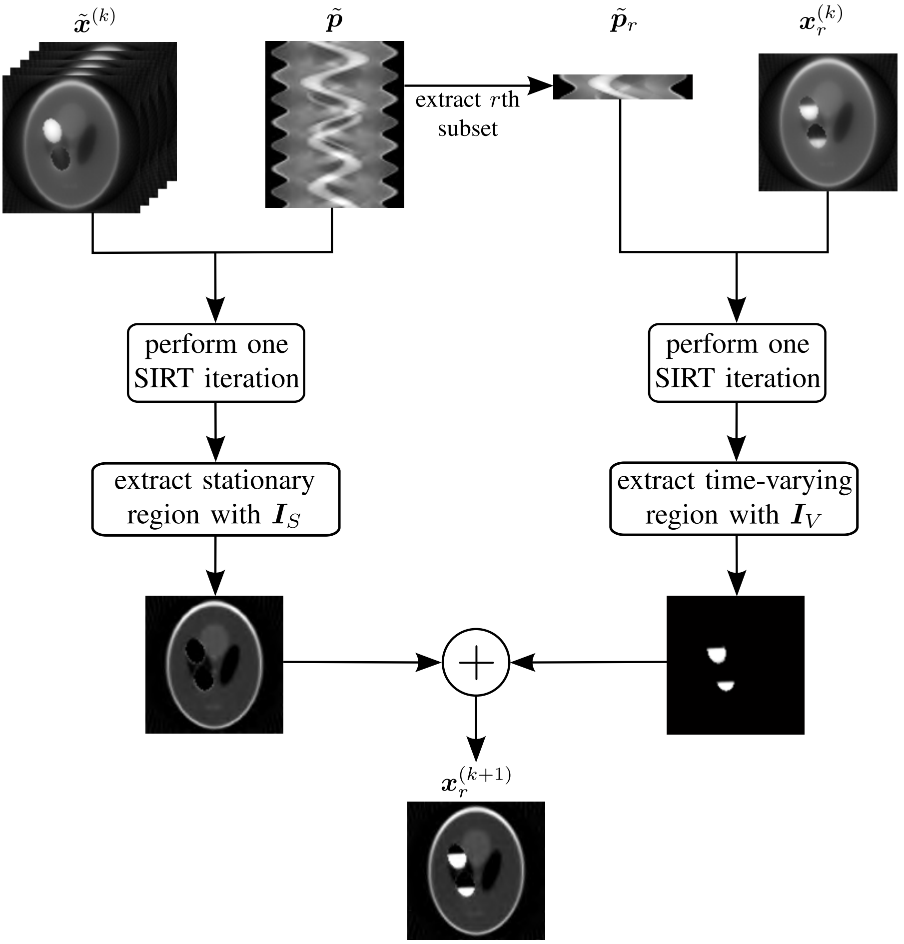


Figure 3.1: Flowchart of a single rSIRT iteration for the rth time frame. Starting from a previous estimate $\tilde{\boldsymbol{x}}^{(k)}$ and the projection data $\tilde{\boldsymbol{p}}$, the rSIRT algorithm calculates for every time index r a new estimate $\boldsymbol{x}_r^{(k+1)}$.

projection along the direction ω_r. Define a *window size* $n_w \in \mathbb{N}\backslash\{0\}$ and use this number to define the projection matrix

$$\tilde{\boldsymbol{W}}_r = \begin{bmatrix} \boldsymbol{W}_{r-\lfloor n_w/2 \rfloor} \\ \boldsymbol{W}_{r-\lfloor n_w/2 \rfloor + 1} \\ \vdots \\ \boldsymbol{W}_{r+\lceil n_w/2 \rceil - 1} \end{bmatrix} \tag{3.2}$$

for each time index r with $\lfloor n_w/2 \rfloor + 1 \leq r \leq R - \lceil n_w/2 \rceil + 1$. The matrix $\tilde{\boldsymbol{W}}_r \in \mathbb{R}^{n_w M_d \times N}$ represents the projection operator that projects along the n_w directions centered around projection angle ω_r. Define

$$\tilde{\boldsymbol{W}}_r = \begin{bmatrix} \boldsymbol{W}_1 \\ \vdots \\ \boldsymbol{W}_{n_w} \end{bmatrix} \tag{3.3}$$

for $r < \lfloor n_w/2 \rfloor + 1$ and

$$\tilde{\boldsymbol{W}}_r = \begin{bmatrix} \boldsymbol{W}_{R-n_w+1} \\ \vdots \\ \boldsymbol{W}_R \end{bmatrix} \tag{3.4}$$

for $r > R - \lceil n_w/2 \rceil + 1$. Next, let $\tilde{\boldsymbol{p}}_r \in \mathbb{R}^{n_w M_d}$ denote the projection data corresponding to the projection directions as they are encoded in $\tilde{\boldsymbol{W}}_r$. Finally, define $\tilde{\boldsymbol{R}}_r \in \mathbb{R}^{n_w M_d \times n_w M_d}$ as the diagonal matrix with inverse row sums of $\tilde{\boldsymbol{W}}_r$ and $\tilde{\boldsymbol{C}}_r \in \mathbb{R}^{N \times N}$ as the diagonal matrix with inverse column sums of $\tilde{\boldsymbol{W}}_r$. The introduced notations allow us to describe the rSIRT algorithm as the following iterative process:

$$\begin{aligned} \boldsymbol{x}_r^{(k+1)} = \boldsymbol{x}_r^{(k)} \quad & + \quad \boldsymbol{I}_S \boldsymbol{C} \boldsymbol{W}^T \boldsymbol{R} (\tilde{\boldsymbol{p}} - \tilde{\boldsymbol{W}} \tilde{\boldsymbol{x}}^{(k)}) \\ & + \quad \boldsymbol{I}_V \tilde{\boldsymbol{C}}_r \tilde{\boldsymbol{W}}_r^T \tilde{\boldsymbol{R}}_r (\tilde{\boldsymbol{p}}_r - \tilde{\boldsymbol{W}}_r \boldsymbol{x}_r^{(k)}). \end{aligned} \tag{3.5}$$

This update needs to be calculated for every $r \in \{1, \ldots, R\}$ before incrementing the iteration count k. Basically, Eq. 3.5 calculates two update steps. A traditional SIRT update step for the stationary region using all available projection data $\tilde{\boldsymbol{p}}$, corresponding to the left hand side of the flowchart in Fig. 3.1, and an update step for the variable region using only the projection data centered around the current time index r, corresponding to the right hand side of the flowchart in Fig. 3.1.

The rSIRT algorithm with initial estimate $\tilde{\boldsymbol{x}}^{(0)} = \boldsymbol{0}$ is representable as a linear operator applied to the projection data $\tilde{\boldsymbol{p}}$, which we denote as $\tilde{\boldsymbol{x}}^{(K)} = \boldsymbol{S}_K \tilde{\boldsymbol{p}}$, where K denotes the number of rSIRT iterations. This can be proven in an analogous manner to the proof in [19], where the regular SIRT algorithm is proven to be a linear operator on the projection data. The details of this derivation can be found in Appendix 3-A of this chapter.

Note that for a practical implementation, the update term in Eq. 3.5 for the stationary region $\boldsymbol{I}_S\boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R}(\tilde{\boldsymbol{p}} - \tilde{\boldsymbol{W}}\tilde{\boldsymbol{x}}^{(k)})$ needs to be calculated only once, since it is exactly the same for every time index r.

Also note that the rSIRT algorithm as presented in Eq. 3.5 is based on a *"sliding window"* approach for selecting subsets in the projection data. This means that a time frame r corresponds to a single acquisition angle, and that for the reconstruction of the dynamic part of that particular time frame, projection data centered around the corresponding acquisition angle, i.e., $\tilde{\boldsymbol{p}}_r$, is considered.
In a *"constant window"* approach, each time frame is associated with multiple acquisition angles (i.e., $M_t > 1$, see Section 1.3). In the constant window approach, projection data corresponding to neighbouring time frames is not considered for the reconstruction of the current time frame and the rSIRT algorithm can be written as

$$\begin{aligned}\boldsymbol{x}_r^{(k+1)} = \boldsymbol{x}_r^{(k)} &+ \boldsymbol{I}_S\boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R}(\tilde{\boldsymbol{p}} - \tilde{\boldsymbol{W}}\tilde{\boldsymbol{x}}^{(k)}) \\ &+ \boldsymbol{I}_V\boldsymbol{C}_r\boldsymbol{W}_r^T\boldsymbol{R}_r(\boldsymbol{p}_r - \boldsymbol{W}_r\boldsymbol{x}_r^{(k)}),\end{aligned} \tag{3.6}$$

where $\boldsymbol{R}_r \in \mathbb{R}^{M\times M}$ and $\boldsymbol{C}_r \in \mathbb{R}^{N\times N}$ are defined as the diagonal matrices with inverse row sums and inverse column sums of $\boldsymbol{W}_r$, respectively. The difference between the sliding window approach and the constant window approach is illustrated in Fig. 3.2. In this chapter, the sliding window approach is followed.

3.2.2 Region inconsistency minimization

In this section, we introduce a projection distance minimization scheme that uses the rSIRT algorithm to estimate the stationary region automatically.

For describing the variable region, a B-spline based closed curve model is used. More details can be found in Appendix 3-B of this chapter. The coordinates of the B-spline closed curve control points are ordered in a parameter vector $\boldsymbol{\alpha}$, describing the degrees of freedom of our optimization routine. Using this region model, we present a measure that indicates how likely a stationary region is to occur. The measure itself is introduced first, while its different components are explained later. The measure is named *region Inconsistency* (rI) and is defined as

$$rI_{\lambda,\mu_1,\mu_2}(\boldsymbol{\alpha}) = ||\tilde{\boldsymbol{W}}\boldsymbol{S}_K^{\boldsymbol{\alpha}}\tilde{\boldsymbol{p}} - \tilde{\boldsymbol{p}}||^2_{\bar{\boldsymbol{R}}^{\lambda}_{\boldsymbol{\alpha}}} + \mu_1 P_1(\boldsymbol{\alpha}) + \mu_2 P_2(\boldsymbol{\alpha}) \quad . \tag{3.7}$$

The first term in Eq. 3.7 is a data fidelity term. Its calculation is displayed schematically in the uppermost part of the flowchart in Fig. 3.3. Since the linear operator

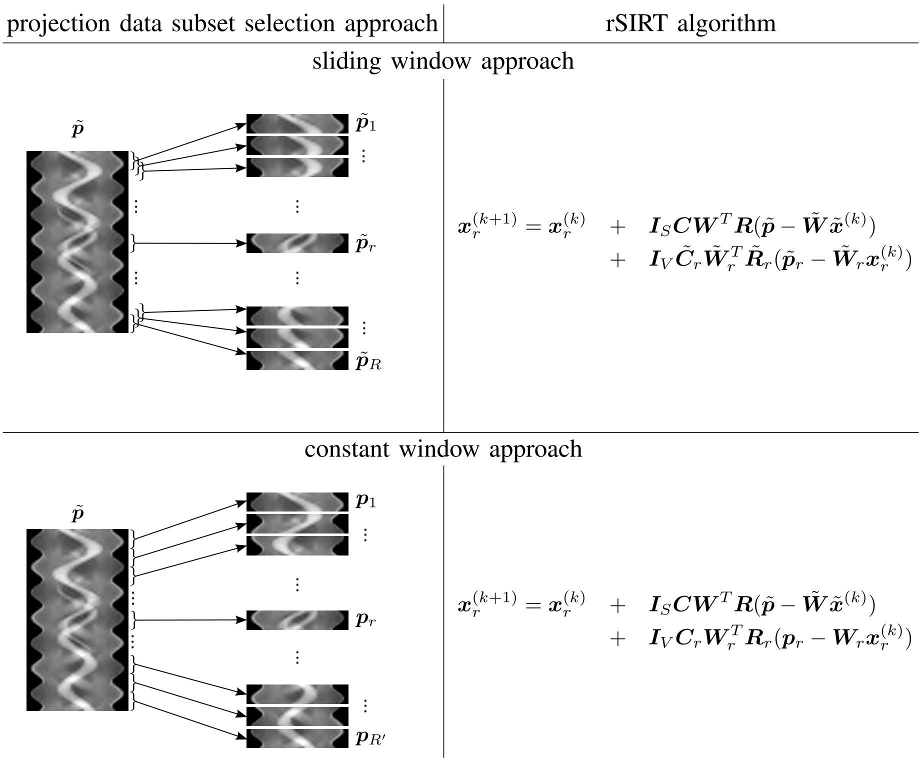


Figure 3.2: Comparison of the rSIRT algorithm for two different projection data subset selection approaches. In the sliding window approach, the dynamic region of each time frame is reconstructed based on projection data centered around each acquisition angle. In the constant window approach, the projection data is subdivided into mutually disjoint subsets in a sequential manner. In the visualization, the number of subsets (or time frames) in the constant window approach is denoted by R', thereby explicitly indicating that, for the same projection data $\tilde{\boldsymbol{p}}$, this number is different than the number of subsets in the sliding window approach (denoted by R in the visualization).

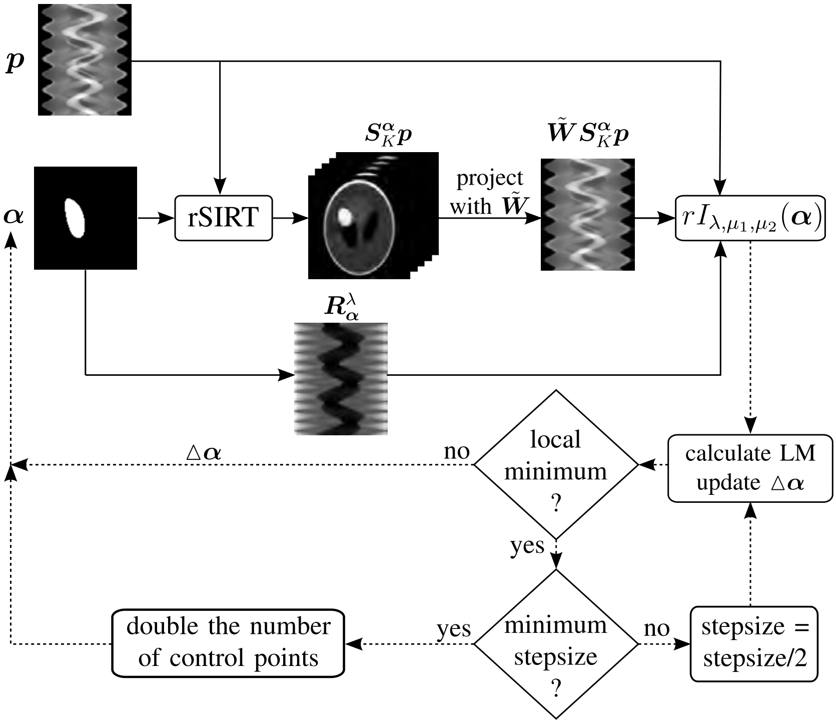


Figure 3.3: Flowchart of the optimization strategy. The solid arrows indicate the calculation of the region inconsistency, while the dotted arrows illustrate the optimization strategy.

describing K iterations of rSIRT depends on the variable region's shape, we denote it by $\boldsymbol{S}_K^{\boldsymbol{\alpha}}$. The rSIRT reconstruction $\boldsymbol{S}_K^{\boldsymbol{\alpha}}\tilde{\boldsymbol{p}}$ is forward projected with the projection matrix $\tilde{\boldsymbol{W}}$ and compared to the original data $\tilde{\boldsymbol{p}}$ through the norm $||\cdot||^2_{\bar{\boldsymbol{R}}_{\boldsymbol{\alpha}}^{\lambda}}$. Instead of weighting this norm with the classical inverse row sum matrix $\boldsymbol{R}$ (see Section 1.2.2.2), it is weighted with a normalized version of an adjusted inverse row sum diagonal matrix $\boldsymbol{R}_{\boldsymbol{\alpha}}^{\lambda}$, where the ray intersection lengths through pixels belonging to the variable region are multiplied with a factor $\lambda > 1$. Hence the diagonal elements of $\boldsymbol{R}_{\boldsymbol{\alpha}}^{\lambda}$ are defined as

$$\boldsymbol{R}_{\boldsymbol{\alpha}}^{\lambda}(i,i) = 1/\sum_j \left((1 + \boldsymbol{I}_V(j,j)(\lambda - 1))w_{ij}\right). \tag{3.8}$$

By normalizing $\boldsymbol{R}_{\boldsymbol{\alpha}}^{\lambda}$ such that it has the same mean along its diagonal as $\boldsymbol{R}$, the normalized diagonal matrix $\bar{\boldsymbol{R}}_{\boldsymbol{\alpha}}^{\lambda}$ is formed. Basically, this weight matrix reflects the higher confidence in measurements corresponding to rays that mainly intersected the stationary region. A more detailed description of $\bar{\boldsymbol{R}}^{\lambda}_{\boldsymbol{\alpha}}$ can be found in Appendix 3-C of this chapter. The term $P_1(\boldsymbol{\alpha})$ in Eq. 3.7 gives a penalty to self-intersecting curves; it simply counts the number of times the curve intersects itself. The term $P_2(\boldsymbol{\alpha})$ gives a penalty to regions that exceed the boundaries of the reconstruction domain. It is given by the sum of the Euclidean distances of each control point outside the reconstruction domain to the closest point inside the reconstruction domain, which is assumed to be confined within $[-1, 1] \times [-1, 1]$. The severity of the penalties defined by $P_1(\boldsymbol{\alpha})$ and $P_2(\boldsymbol{\alpha})$ are controlled by the parameters μ_1 and μ_2.

An adapted variant of the Levenberg-Marquardt (LM) algorithm [20] is used for minimizing rI, which is illustrated in the lower part of the flowchart in Fig. 3.3. The rI function is a non-convex function that, due to noise and discretization effects, has a non-smooth, coarse landscape on a small scale. This coarseness can make the finite difference approximation of the objective function's gradient inaccurate. Therefore, finite differences are calculated with a parameter increment (referred to as the *stepsize*), starting from a larger initial stepsize that is halved every time the solver reaches a local minimum. When the stepsize becomes smaller than a specified threshold, the stepsize is reinitialized and the number of spline control points is doubled by applying Boehm's formula for knot insertion [21], thus providing more degrees of freedom to the solver by this multi-resolution approach on the parameter vector $\boldsymbol{\alpha}$.

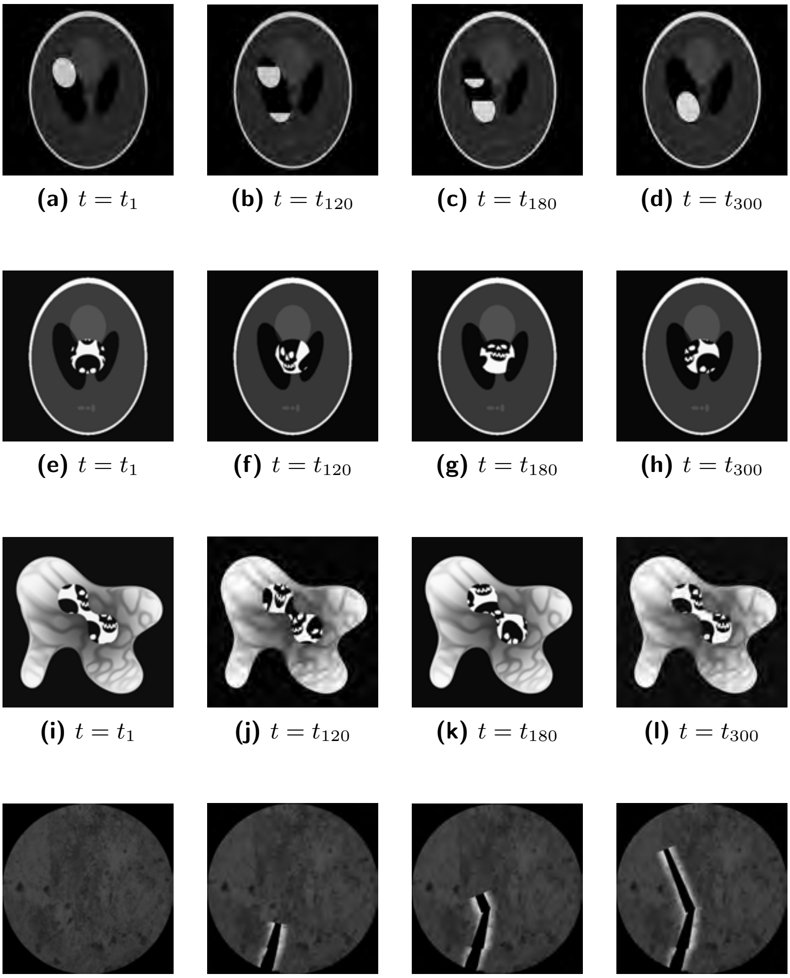
3.3 Experiments

In this section, rSIRT and the region inconsistency minimization are validated on numerical and experimental data.

3.3.1 Numerical simulations

Consider the phantoms in Fig. 3.4. Phantom 1 is displayed in Fig. 3.4a-d. It is a modified Shepp-Logan phantom. The phantom resembles a sample in which fluid flows from one chamber into another during the acquisition of the CT data. Phantom 2 (Fig. 3.4e-h) and phantom 3 (Fig. 3.4i-l) are artificial phantoms in which a structurally varying region in the form of a circle and an eight-like shape are present in the object. Finally, phantom 4, displayed in Fig. 3.4m-p, represents an object in which a crack is formed.

Projections were simulated with the binary decomposition scanning protocol (see Section 1.3) where we took the window size to be $n_w = 30$. We simulated 300 projections using a strip kernel and a higher resolution version of the phantom, i.e., on a 500×500 isotropic pixel grid, while the algorithm calculates reconstructions



(a) $t = t_1$ **(b)** $t = t_{120}$ **(c)** $t = t_{180}$ **(d)** $t = t_{300}$

(e) $t = t_1$ **(f)** $t = t_{120}$ **(g)** $t = t_{180}$ **(h)** $t = t_{300}$

(i) $t = t_1$ **(j)** $t = t_{120}$ **(k)** $t = t_{180}$ **(l)** $t = t_{300}$

(m) $t = t_1$ **(n)** $t = t_{120}$ **(o)** $t = t_{180}$ **(p)** $t = t_{300}$

Figure 3.4: Each row represents a different phantom. Each column represents a different point in time. (a-d) Phantom 1: a modified Shepp-Logan phantom. The changes in the phantom resemble a liquid flowing from one chamber of the object into another. (e-h) Phantom 2: a modified Shepp-Logan phantom. In the middle of the phantom a structure with skulls is changing over time. (i-l) Phantom 3: blob-shaped phantom with an eight-like variable region in the middle. (m-p) Phantom 4: image of material in which a crack is formed.

on a 100 × 100 isotropic pixel grid. This way, inverse crimes [22] are avoided, i.e., using the same model for generating data as for testing the designed algorithm. Also, Poisson distributed noise was applied to the projection data assuming an incoming beam intensity of 10000 (photon count) per detector pixel.

For validation, three different measures were utilized. A first figure of merit is the Root Mean Square Error (RMSE). Denote the calculated reconstruction by $\hat{\boldsymbol{x}}$ and the phantom used to generate the data by $\tilde{\boldsymbol{x}}$. Then the RMSE is given by

$$RMSE = \sqrt{\langle(\hat{\boldsymbol{x}} - \tilde{\boldsymbol{x}})^2\rangle}, \tag{3.9}$$

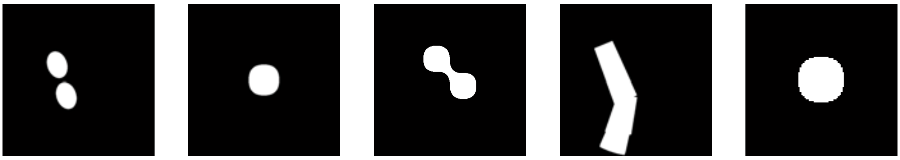
where the squaring should be interpreted pointwise and where $\langle\cdot\rangle$ represents the average. We upsampled the calculated reconstruction 5 times in order to perform the RMSE calculations on the 500 × 500 grid. A second measure is the projection distance, defined as

$$PD = ||\tilde{\boldsymbol{W}}\hat{\boldsymbol{x}} - \tilde{\boldsymbol{p}}||_2. \tag{3.10}$$

Finally, for the validation of the region inconsistency minimization, we also look at the relative Number of Misclassified Pixels (rNMP), i.e., the number of misclassified variable region pixels (with respect to the ground truth variable region) divided by the total number of pixels belonging to the ground truth variable region. The rNMP was also calculated on the original 500 × 500 grid by scaling up the optimized region.

The performance of rSIRT was compared to two other reconstruction methods. A first method is the *conventional* method. It calculates a regular SIRT reconstruction per subset of the projection data $\tilde{\boldsymbol{p}}_r$ with the corresponding projection matrices as they are defined in Eq. 3.2, Eq. 3.3 and Eq. 3.4. Also, a regular SIRT reconstruction was used for comparison, i.e., a SIRT reconstruction using all the projection data $\tilde{\boldsymbol{p}}$ where the object was regarded as stationary through time (see Section 1.2.2.2).

First, a validation of the rSIRT algorithm is presented. Since the phantoms are simulated, the stationary region is known and can be used as prior knowledge for rSIRT. The ground truth regions are displayed in Fig. 3.5a-d. The RMSE after 100 iterations with an initial zero estimate was calculated for all phantoms and for each of the reconstruction methods. The results are summarized in the columns "*SIRT*", "*Conventional*" and "*rSIRT*" in Table 3.1. In this table, the column *rSIRT* refers to the rSIRT reconstruction with the ground truth stationary region. In order to illustrate the convergence properties of rSIRT, more detailed plots of the results for phantom 4 are presented in Fig. 3.6 and Fig. 3.7. Fig. 3.6b indicates that all methods reduce the projection distance as the iteration number increases. As the reconstruction per time point for the conventional method must only match with a subset of all projection data, its projection distance decreases

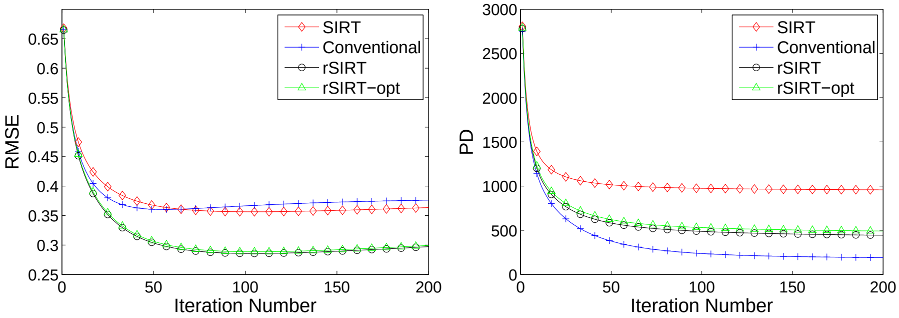


(a) Phantom 1 **(b)** Phantom 2 **(c)** Phantom 3 **(d)** Phantom 4 **(e)**

Figure 3.5: (a-d) Ground truth for the time-varying region for all simulation phantoms. (e) Initial estimate for simulation experiments with the region inconsistency minimization algorithm. In all figures, white pixels belong to the time-varying region and black pixels belong to the stationary region.

	SIRT	*Conventional*	*rSIRT*	*rSIRT-opt*
Phantom 1	0.13379	0.12291	0.09767	0.09793
Phantom 2	0.12508	0.13702	0.11247	0.11257
Phantom 3	0.12993	0.13627	0.11038	0.11068
Phantom 4	0.36335	0.36543	0.28545	0.28867

Table 3.1: RMSE after 100 iterations for all simulated phantoms (rows) and for different methods (columns).



(a) **(b)**

Figure 3.6: The convergence of the different methods for phantom 4. (a) RMSE as a function of iteration number. (b) PD as a function of iteration number.

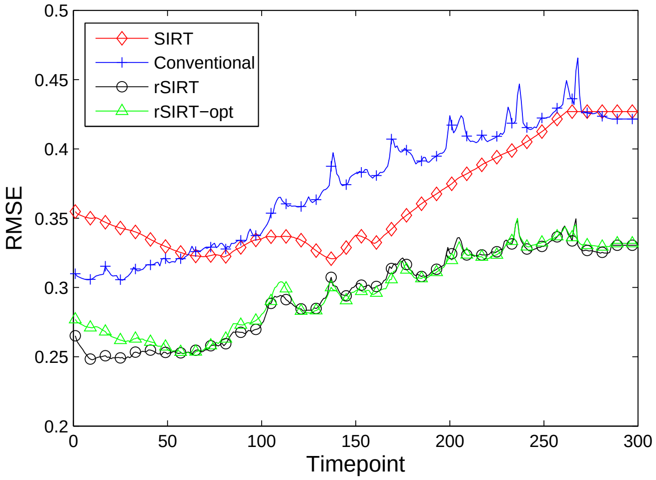


Figure 3.7: RMSE per point in time for reconstructions after 100 iterations for phantom 4.

the fastest. However, since the RMSE assesses the image quality directly, it can be concluded from Fig. 3.6a that rSIRT has superior performance. Also, SIRT is unable to capture the object's dynamics, as it generates one reconstruction for the entire projection data set. The conventional method can capture some dynamics, but reconstruction quality is severely degraded. The rSIRT algorithm combines the better properties of the conventional method and SIRT, and reconstructs the object with improved image quality and time resolution. The RMSE at each time instance of the reconstructions after 100 iterations is displayed in Fig. 3.7, confirming the previous statements.

Also, the region inconsistency minimization was validated on the simulated data. The parameter setup was $\lambda = 10$, spline degree $k = 2$, $\mu_1 = 1$, $\mu_2 = 1$, $n_w = 30$ and $K = 30$. The initial estimate for the variable region is displayed in Fig. 3.5e. Other initial estimates (e.g., not intersecting with the ground truth variable region) may affect convergence speed, but typically produce the same result. The number of control points was doubled three times, resulting in a final region descriptor of 32 control points. In the last optimization round, i.e., when using 32 control points, the number of rSIRT iterations K was set to 60 to obtain a more accurate result. The output of the region inconsistency minimization algorithm — the stationary region — was used as input for the rSIRT algorithm with 100 iterations. The RMSE of this rSIRT reconstruction is tabulated in the column "*rSIRT-opt*" in Table 3.1. The rNMP as a function of iteration number is displayed in Fig. 3.8. The resulting variable region estimates and its misclassified pixels are visualized

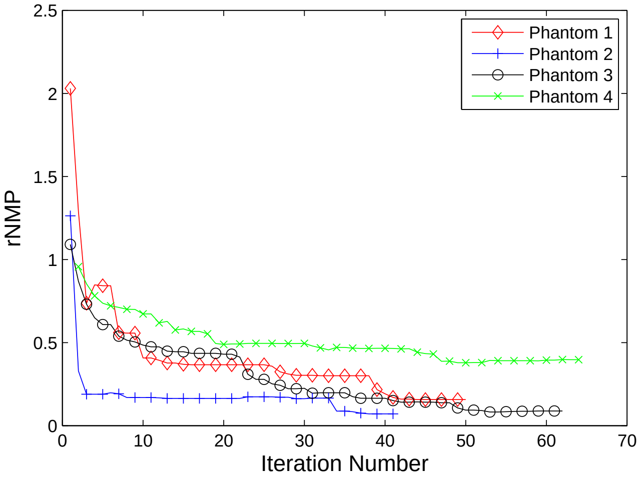
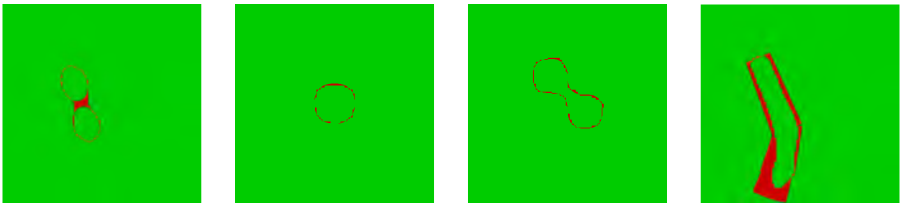


Figure 3.8: rNMP as a function of iteration number of the region inconsistency minimization algorithm.



(a) Phantom 1, rNMP $= 0.157$

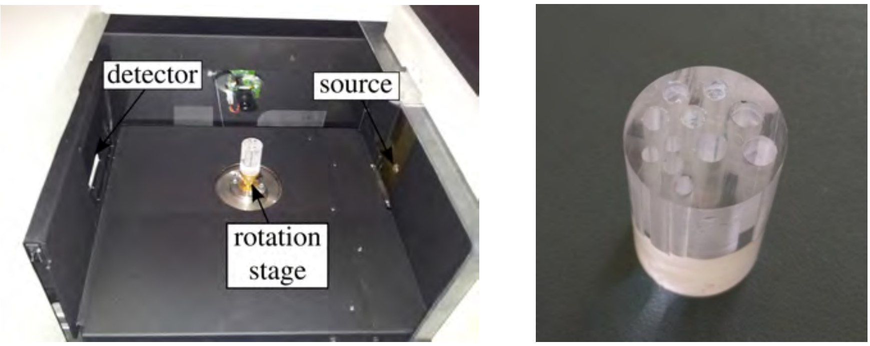
(b) Phantom 2, rNMP $= 0.0712$

(c) Phantom 3, rNMP $= 0.0879$

(d) Phantom 4, rNMP $= 0.397$

Figure 3.9: Region estimate after region inconsistency minimization. Red indicates misclassified pixels, green indicates correctly classified pixels. The corresponding rNMP is also indicated for every region estimate.

in Fig. 3.9. These results indicate that the region inconsistency minimization algorithm is able to estimate a stationary region that closely approximates the ground truth stationary region, resulting in an almost identical RMSE for the rSIRT reconstruction based on the estimated region and the ground truth region. Despite the region estimate for phantom 4 being less accurate (see Fig. 3.9d), Fig. 3.7 indicates that the RMSE is still significantly improved. Also, Fig. 3.6a illustrates that the convergence properties remain almost unaltered.



(a) The plexiglas phantom placed inside the SkyScan 1172 μCT scanner

(b) A close-up picture of the plexiglas phantom

Figure 3.10: The plexiglas phantom and the experimental set-up.

The calculation of 30 rSIRT iterations in the experimental setup as described above (i.e., reconstructing on a 100 × 100 pixel grid and using projection data consisting of 300 projection angles), takes about 5 seconds with an unoptimized GPU implementation. Depending on the number of iterations needed by the region inconsistency minimization algorithm, it takes about 2-4 hours to calculate a variable region estimate. However, this computational time could be significantly reduced by parallelizing the rI function evaluations, which are needed for the gradient calculation.

3.3.2 Experimental data

A sequence of cone beam projection images of a plexiglas (Polymethyl Methacrylate) resolution phantom was acquired in a Bruker-microCT SkyScan 1172 μCT scanner. The experimental set-up and the phantom are displayed in Fig. 3.10. Each radiograph was acquired using a source voltage of 80 kV, a source current of 124 μA, a 0.5 mm Al filter and an exposure time of 360 ms. The object has been scanned over the full angular range of 360° with a 0.6° tilt increment, this resulted in a total of 600 log-corrected projection images per scan. To increase the SNR, the original detector pixels of size 9.01 μm were downsampled by a factor 4 in both dimensions, resulting in projection images of dimension 666×1000 with detector pixel size 36.04 μm. Reconstructions were performed on a 1000×1000 isotropic pixel grid of the slice on the optical axis. Having a horizontal cone-beam angle of 9.52° and a source to origin distance of 130.21 mm, this results in a pixel size of 21.69 μm. Since ring- and beam-artifact correction goes beyond the focus of this thesis, we have preprocessed the projection images to correct these artifacts using the standard Bruker-microCT NRecon software package.

A cross section on the optical axis of the plexiglas phantom is displayed in Fig. 3.14a. The 2 cm diameter plexiglas cylinder was drilled with three 4 mm diameter holes, four 3 mm diameter holes and four 2 mm diameter holes. We created 2 datasets, each consisting of 600 projection images of the phantom. The first dataset is a regular scan of the phantom. Next, one of the pores was filled with water and the second dataset was acquired with exactly the same parameter setup as for the acquisition of the first dataset. A ground truth for each of the datasets was created by calculating a SIRT reconstruction using 50 iterations and the full set of 600 projections. Subsequently the reconstruction was segmented using the method of Otsu [23]. The obtained ground truth for the first and the second dataset are displayed in Fig. 3.14a and Fig. 3.14f, respectively.

The first two projection datasets were combined into a single projection dataset as if the hole was filled with water during the data acquisition. The combination of the projection datasets is illustrated in Fig. 3.11. The first 75 projections were

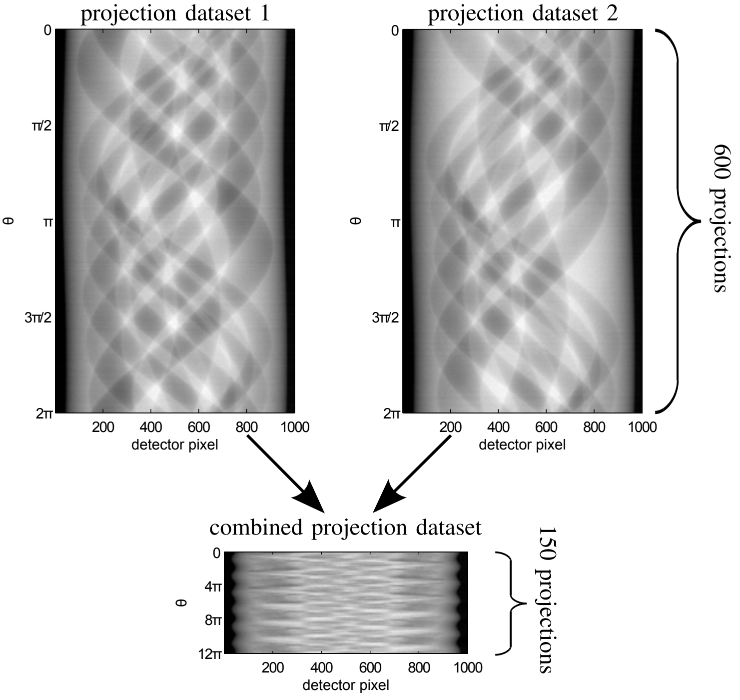


Figure 3.11: Illustration of the combination of the two datasets, consisting of 600 projections each, into one single dataset, consisting of only 150 projections.

taken from the first dataset, where the projections were ordered analogously to the binary decomposition scanning protocol (see Section 1.3) with window size $n_w = 25$. The only difference is that we made sure that each n_w directions correspond to approximately 360°, since each of the projection datasets corresponds to a full 360° angular range. The next 75 projections were taken from the second dataset, and ordered in the same manner as the first dataset. This resulted in a projection dataset consisting of 150 angular directions, where source and detector have rotated a total of 6 times around the object in a binary decomposition scanning protocol.

Next we applied the region inconsistency minimization to the resulting combined projection dataset. The used parameters were $\lambda = 10$, spline degree $k = 2$, $\mu_1 = 1$, $\mu_2 = 1$, $n_w = 25$ and $K = 25$. The initial variable region estimate is displayed in Fig. 3.13a. The intermediate variable region estimates, i.e., after each iteration of the region inconsistency minimization algorithm, were used to calculate the rNMP as well as the RMSE of the associated rSIRT reconstruction. These results are displayed in Fig. 3.12. Convergence was reached after 33 iterations. The final variable region estimate was used to generate an rSIRT reconstruction, which is referred to as "*rSIRT-opt*". The RMSE of this reconstruction with respect to the ground truth was calculated and compared to the results of alternative methods in Table 3.2. The ground truth variable region, the final variable region estimate and the misclassified pixels are displayed in Fig. 3.13. A visual comparison of the different reconstructions can be done in Fig. 3.14. Visual as well as numerical results indicate the advantage of the region inconsistency minimization algorithm. The SIRT reconstruction in Fig. 3.14b and Fig. 3.14g is exactly the same for every time point, which makes it useless for studying structural changes within the scanned object. The conventional method's reconstruction of Fig. 3.14c and Fig. 3.14h is able to image changes through time, but suffers from severe streak artifacts because it is based on limited projection data per time point reconstruction. In contrast, the rSIRT reconstruction of Fig. 3.14d and Fig. 3.14i is able to simultaneously capture the structural changes and maintain image quality, because it combines projection data over different time points within the stationary region. This improves image quality in the stationary region and in the variable region, since the update step in rSIRT is based on difference between the simulated projections of the current reconstruction and the measured projection data.

SIRT	*conventional*	*rSIRT*	*rSIRT-opt*
1.6752e-04	2.0325e-04	1.4597e-04	1.4821e-04

Table 3.2: RMSE after 40 iterations for the plexiglas phantom.

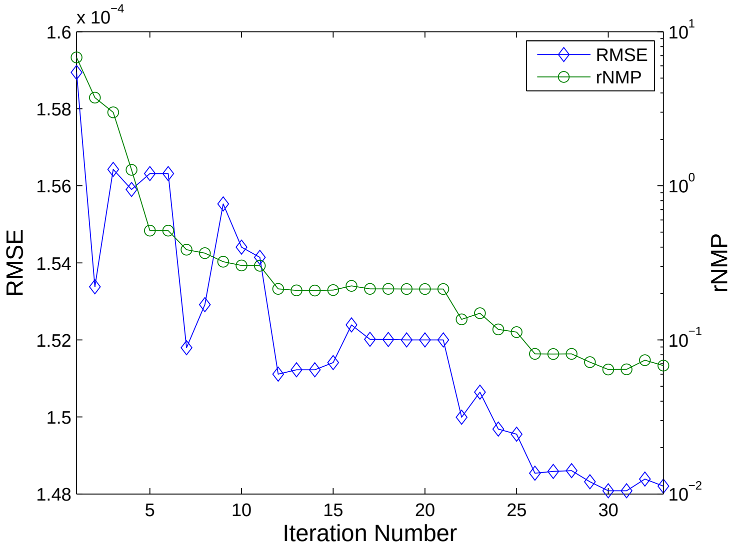
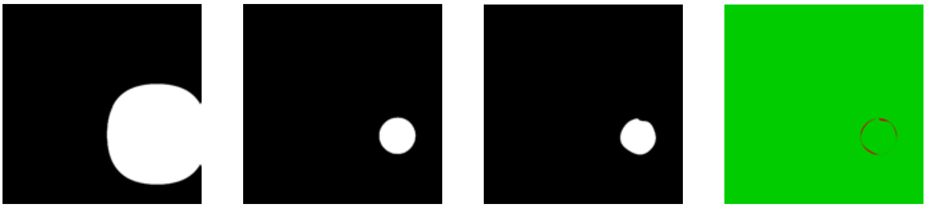
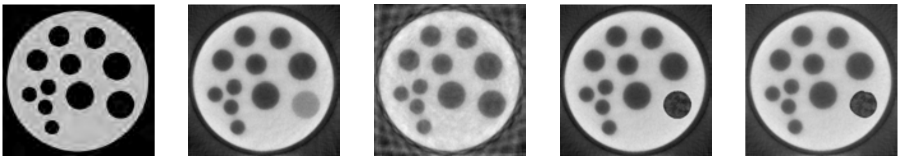


Figure 3.12: RMSE (left vertical axis) and rNMP (right vertical axis) in function of iteration number for the region inconsistency minimization applied to the plexiglas resolution phantom. The RMSE values are calculated on the reconstruction generated with 40 rSIRT iterations and the intermediate variable region estimate.

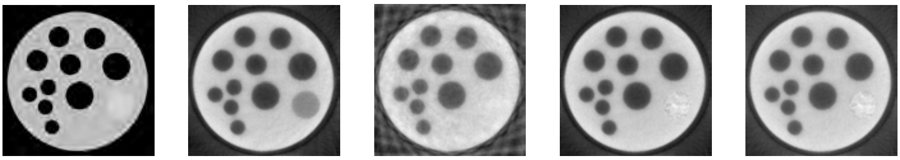


(a) Initial estimate **(b)** Ground truth **(c)** rI minimization result **(d)** Misclassified pixels, $\mathrm{rNMP} = 0.0681$

Figure 3.13: Region inconsistency minimization results. In (a)-(c), the initial estimate, the ground truth and the region inconsistency minimization result are shown, respectively. White pixels indicate the time-varying region and black pixels indicate the stationary region. In (d), the misclassified pixels are visualized, red pixels are misclassified and green pixels are correctly classified.



(a) *Ground truth* at time instance t_i with $i \leq 75$ **(b)** *SIRT* - t_{30} **(c)** *conventional* - t_{30} **(d)** *rSIRT* - t_{30} **(e)** *rSIRT-opt* - t_{30}



(f) *Ground truth* at time instance t_i with $i > 75$ **(g)** *SIRT* - t_{120} **(h)** *conventional* - t_{120} **(i)** *rSIRT* - t_{120} **(j)** *rSIRT-opt* - t_{120}

Figure 3.14: Ground truth and reconstructions at t_{30} and at t_{120} with 40 iterations for the different methods.

3.4 Discussion and conclusion

In general, reconstructing structurally or discretely time-varying objects based on tomographic data is a difficult problem. Popular methods either reconstruct the object independently at different time points using a subset of all projection data or assume the changes to be continuous. The first method suffers from artifacts introduced by the lack of projection data per reconstruction, and the latter cannot be applied to structurally or discretely time-varying objects, as these changes are no longer representable by a continuous deformation model.

In this chapter, we have presented the novel rSIRT algorithm. It can be used for the reconstruction of time-varying structurally changing objects when there exist regions within the object that remain stationary through time. There are two intuitive aspects of the rSIRT algorithm that illustrate its ability to create accurate reconstructions. On the one hand, rSIRT guarantees accurate image quality in the stationary region, since the iterative update step for this region is based on all available projection data, in contrast to the conventional method, where the stationary (and time-varying) region is reconstructed based upon information available in a subset of all projection data. On the other hand, as image quality in the stationary region improves, this is propagated to the variable region to some extent, since the update step for the variable region is computed using the projection difference of the previous estimate, which is based on the projection of stationary and variable region. Naturally, the final reconstruction quality is influenced by the amount of projection data used for generating the iterative update step in the time-varying region, which is encoded by the window size. However, for any fixed window size rSIRT improves the reconstruction quality, thereby allowing shorter acquisition times per time window. Motivated by its ability to reconstruct a wide variety of objects, SIRT was the algorithm of choice for the generation of the iterative update steps in the time-varying and stationary region within the rSIRT algorithm. However, any other iterative algorithm could be employed to generate these update steps. In more specific scanning set-ups, the choice of reconstruction algorithm should be tailored to the specific properties of the object under investigation, e.g. sparsity.

As it is time consuming and not always trivial to manually indicate the time-varying regions, we have developed the region inconsistency minimization procedure, an optimization algorithm that automatically estimates the time-varying regions. The method minimizes a region inconsistency measure, where the difference between simulated and measured projection data is weighted according to a factor that essentially encodes the ray intersection length and the quality of the simulated projection based on the proportion of the ray that passed through the stationary region (see Appendix 3-C). Consequently, the region inconsistency minimization algorithm is most suited for objects where no extreme changes in reconstruction quality through the stationary region are present. However, most objects do not exhibit this kind of behavior, which makes the region inconsistency minimization method widely applicable.

Experiments with simulated and μCT data illustrate the ability of the region inconsistency minimization algorithm to find a good approximation of the time-varying regions. Also, the rSIRT algorithm was shown to have a clear advantage over popular methods when constructing structurally changing objects.

Appendix 3-A

This appendix provides more details about the linear operator $\boldsymbol{S}_K$ that represents the rSIRT algorithm. More precisely, if $\tilde{\boldsymbol{x}}^{(0)} = \boldsymbol{0}$, the rSIRT reconstruction after K iterations depends linearly on the projection data, i.e., $\tilde{\boldsymbol{x}}^{(K)} = \boldsymbol{S}_K\tilde{\boldsymbol{p}}$.

First, for $r = 1, \dots, R$, define $\boldsymbol{K}_r$ as the $N \times RN$ matrix that selects the reconstruction $\boldsymbol{x}_r$ at time index r from the full reconstruction vector $\tilde{\boldsymbol{x}}$, i.e., $\boldsymbol{K}_r\tilde{\boldsymbol{x}} = \boldsymbol{x}_r$. Also define $\boldsymbol{L}_r$ as the $n_w M_d \times RM_d$ matrix that selects the projection data $\tilde{\boldsymbol{p}}_r$ corresponding to the rth window from the full projection data $\tilde{\boldsymbol{p}}$, i.e., $\boldsymbol{L}_r\tilde{\boldsymbol{p}} = \tilde{\boldsymbol{p}}_r$.

With these definitions, Eq. 3.5 can be rewritten as

$$\boldsymbol{x}_r^{(k+1)} = \boldsymbol{M}_{l,1}\tilde{\boldsymbol{x}}^{(k)} + \boldsymbol{M}_{l,2}\tilde{\boldsymbol{p}}, \tag{3.11}$$

with

$$\boldsymbol{M}_{l,1} = \boldsymbol{K}_r - \boldsymbol{I}_S\boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R}\tilde{\boldsymbol{W}} - \boldsymbol{I}_V\tilde{\boldsymbol{C}}_r\tilde{\boldsymbol{W}}_r^T\tilde{\boldsymbol{R}}_r\tilde{\boldsymbol{W}}_r\boldsymbol{K}_r \tag{3.12}$$

and

$$\boldsymbol{M}_{l,2} = \boldsymbol{I}_S\boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R} + \boldsymbol{I}_V\tilde{\boldsymbol{C}}_r\tilde{\boldsymbol{W}}_r^T\tilde{\boldsymbol{R}}_r\boldsymbol{L}_r. \tag{3.13}$$

Set

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{M}_{1,1} & \boldsymbol{M}_{1,2} \\ \vdots & \vdots \\ \boldsymbol{M}_{R,1} & \boldsymbol{M}_{R,2} \\ \emptyset & \mathbb{1}_{RM_d} \end{bmatrix}, \tag{3.14}$$

where $\mathbb{1}_n$ represents the identity matrix of size $n \in \mathbb{N}\backslash\{0\}$. It can be easily verified that for

$$\boldsymbol{S}_K = \begin{bmatrix} \mathbb{1}_{RN} & \emptyset \end{bmatrix} \boldsymbol{M}^K \begin{bmatrix} \emptyset \\ \mathbb{1}_{RM_d} \end{bmatrix}, \tag{3.15}$$

we have $\tilde{\boldsymbol{x}}^{(K)} = \boldsymbol{S}_K\tilde{\boldsymbol{p}}$.

Appendix 3-B

A brief description of the parametric B-spline closed curve model [24, 21] is given, which is used throughout Chapter 3 for describing the variable region within the scanned object. To define the normalized B-spline closed curve of degree k with n *control points*, we first introduce $n + 2k + 1$ *knot-points* $t_{-k} < t_{-k+1} < \cdots < t_{n+k} \in \mathbb{R}$. The following recursion relations are used to define the normalized B-spline basis functions of degree k for $i = -k, -k + 1, \ldots, n - 1$:

$$\begin{aligned} N_{i,k+1}(t) &= \frac{t - t_i}{t_{i+k} - t_i}N_{i,k}(t) \\ &\quad + \frac{t_{i+k+1} - t}{t_{i+k+1} - t_{i+1}}N_{i+1,k}(t) \end{aligned} \tag{3.16}$$

$$N_{i,1}(t) = \begin{cases} 1 & \textit{if} \quad t_i \le t < t_{i+1} \\ 0 & \textit{otherwise} \end{cases} \tag{3.17}$$

The normalized B-spline $N_{i,k+1}$ can be used to define a closed curve $\boldsymbol{c}(t)$ for $t \in [t_0, t_n]$:

$$\boldsymbol{c}(t) = \sum_{i=-k}^{n-1} \boldsymbol{c}_i N_{i,k+1}(t), \tag{3.18}$$

with *control points* $\boldsymbol{c}_i \in \mathbb{R}^2$ for $i = -k, \ldots, n-1$ and $\boldsymbol{c}_i = \boldsymbol{c}_{i-n}$ for $i = n-k, \ldots, n-1$. Once the knot-points are fixed, the region within the closed curve $\boldsymbol{c}$ is completely described by the coordinates of the control points $\boldsymbol{c}_0, \ldots, \boldsymbol{c}_{n-1}$.

Appendix 3-C

In this appendix, a more detailed description of the normalized adjusted inverse row sum matrix $\bar{\boldsymbol{R}}^{\lambda}_{\boldsymbol{\alpha}}$ is given. Remember that $\bar{\boldsymbol{R}}^{\lambda}_{\boldsymbol{\alpha}}$ is calculated as the normalized version of Eq. 3.8, where the normalization refers to the fact that the diagonal elements of $\boldsymbol{R}^{\lambda}_{\boldsymbol{\alpha}}$ are multiplied with a suitable factor in order to have the same mean as the diagonal elements of the standard inverse row sum matrix $\boldsymbol{R}$.

Unlike $\boldsymbol{R}$, which contains weights that encode the ray intersection lengths with the reconstruction domain, $\bar{\boldsymbol{R}}^{\lambda}_{\boldsymbol{\alpha}}$ additionally encodes how much of the ray interacted with the variable region. It can be understood in light of the region inconsistency measure $rI_{\lambda,0,0}(\boldsymbol{\alpha}) = ||\tilde{\boldsymbol{W}}\boldsymbol{S}^{\boldsymbol{\alpha}}_{K}\tilde{\boldsymbol{p}} - \tilde{\boldsymbol{p}}||^2_{\bar{\boldsymbol{R}}^{\lambda}_{\boldsymbol{\alpha}}}$:

- If a large part of the ray passes through the variable region, its corresponding value in $\bar{\boldsymbol{R}}^{\lambda}_{\boldsymbol{\alpha}}$ will be rather small, indicating that the difference between the simulated and the measured detector value in this detector pixel is less important than others.

- If a ray passes mostly through the stationary region, the corresponding value in $\bar{\boldsymbol{R}}^{\lambda}_{\boldsymbol{\alpha}}$ will be rather large. This means that the simulated projection value should have a strong resemblance to the measured detector value.

Intuitively, rSIRT reconstructs the stationary region using *all* projection data, hence this part of the reconstruction should be more accurate and the calculated projection values corresponding to this region should have a good match with the measured projection data, which is encoded by the relatively larger weights in $\bar{\boldsymbol{R}}^{\lambda}_{\boldsymbol{\alpha}}$. On the other hand, the variable region is reconstructed using less projection data, and is therefore expected to be of lower quality in comparison to the stationary region. This is encoded in the fact that corresponding projection values receive a smaller weight. Without this adjusted weighting, optimization of rI would always result in a variable region covering the entire reconstruction domain. This can be understood as follows: The stationary regions are used to generate projections that must match up with *all* projection data, which limits the set of possible solutions for the pixels belonging to the stationary regions. On the other hand, the variable regions are used to simulate projections that must match up with only a subset of all projection data, which results in a larger set of possible solutions for the pixels in the variable region. As the size of the variable region increases, the set of possible reconstructions matching the projection data grows, which would result in a smaller value for the projection distance. This property was experimentally confirmed by the following simulation experiment. Consider the modified Shepp-Logan phantom of Fig. 3.4e-(h). A version of this phantom on a 500 × 500 pixel grid was used to generate projections over 300 time points. This projection data was used to calculate a rSIRT reconstruction with window size $n_w = 30$ on a 100 × 100 grid. Subsequently $rI_{\lambda,0,0}(\boldsymbol{\alpha})$ was calculated for different values of λ and for different variable region's sizes, resulting in Fig. 3.15. Note that $\lambda = 1$ corresponds to a weighting with the standard inverse row sum matrix $\boldsymbol{R}$.

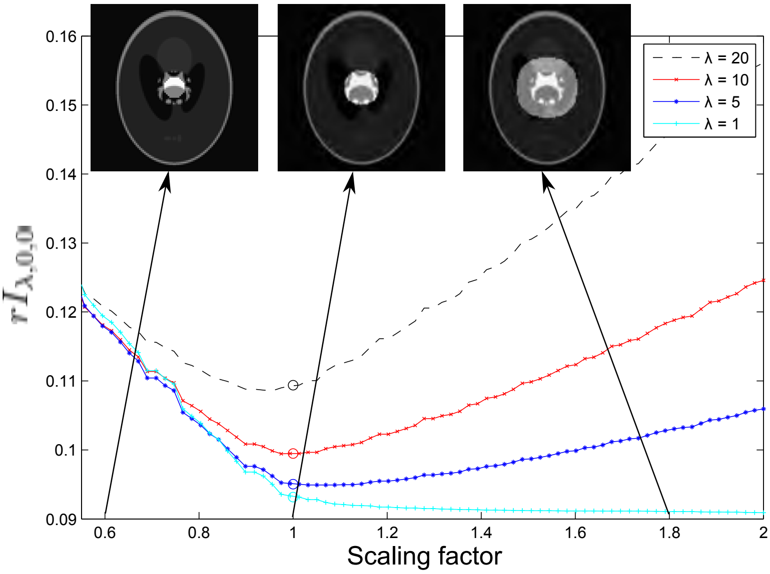


Figure 3.15: The region inconsistency measure $rI_{\lambda,0,0}(\boldsymbol{\alpha})$ for different region sizes and for $\lambda = 1, 5, 10, 20$, illustrated on the phantom of Fig. 3.4e-(h). A scaling factor 1 corresponds to the ground truth region. To illustrate the size of the variable regions for different scaling factors, we have superimposed the region for a scaling factor of 0.6, 1 and 1.8 on the phantom.

References

[1] R. Mooser, F. Forsberg, E. Hack, G. Szkely, and U. Sennhauser, “Estimation of affine transformations directly from tomographic projections in two and three dimensions,” *Machine Vision and Applications*, vol. 24, pp. 419–434, 2013.

[2] G. Van Eyndhoven, J. Sijbers, and K. J. Batenburg, “Combined motion estimation and reconstruction in tomography,” in *Computer Vision ECCV 2012. Workshops and Demonstrations*, vol. 7583, pp. 12–21, 2012.

[3] T. Li, E. Schreibmann, Y. Yang, and L. Xing, “Motion correction for improved target localization with on-board cone-beam computed tomography.,” *Physics in Medicine and Biology*, vol. 51, pp. 253–67, Jan. 2006.

[4] Q. Zhang, A. Pevsner, A. Hertanto, Y.-C. Hu, K. E. Rosenzweig, C. C. Ling, and G. S. Mageras, “A patient-specific respiratory model of anatomical motion for radiation treatment planning.,” *Medical Physics*, vol. 34, pp. 4772–81, Dec. 2007.

[5] J. Hinkle, M. Szegedi, B. Wang, B. Salter, and S. Joshi, “4D CT image reconstruction with diffeomorphic motion model.,” *Medical Image Analysis*, vol. 16, pp. 1307–16, Aug. 2012.

[6] R. Li, X. Jia, J. H. Lewis, X. Gu, M. Folkerts, C. Men, and S. B. Jiang, “Real-time volumetric image reconstruction and 3D tumor localization based on a single x-ray projection image for lung cancer radiotherapy.,” *Medical Physics*, vol. 37, pp. 2822–6, June 2010.

[7] A. Docef and M. J. Murphy, “Reconstruction of 4D deformed CT for moving anatomy,” *International Journal for Computer Assisted Radiology and Surgery*, vol. 3, pp. 591–598, Oct. 2008.

[8] J. Ehrhardt, R. Werner, A. Schmidt-Richberg, and H. Handels, “Statistical modeling of 4D respiratory lung motion using diffeomorphic image registration.,” *IEEE Transactions on Medical Imaging*, vol. 30, pp. 251–65, Feb. 2011.

[9] S. S. Vedam, P. J. Keall, V. R. Kini, H. Mostafavi, H. P. Shukla, and R. Mohan, “Acquiring a four-dimensional computed tomography dataset using an external respiratory signal.,” *Physics in Medicine and Biology*, vol. 48, pp. 45–62, Jan. 2003.

[10] J. Williams, K. Yazzie, E. Padilla, N. Chawla, X. Xiao, and F. De Carlo, “Understanding fatigue crack growth in aluminum alloys by in situ X-ray synchrotron tomography,” *International Journal of Fatigue*, pp. 6–12, June 2012.

[11] I. Riedel, E. Andò, S. Salager, P. Bsuelle, and G. Viggiani, “Water retention behaviour explored by x-ray ct analysis,” in *Unsaturated Soils: Research and Applications*, pp. 81–88, 2012.

[12] P. Gregory, D. J. Hutchison, D. B. Read, P. M. Jenneson, W. B. Gilboy, and E. J. Morton, "Non-invasive imaging of roots with high resolution X-ray micro-tomography," *Plant and Soil*, vol. 255, pp. 351–359, Aug. 2003.

[13] H. Gao, J.-F. Cai, Z. Shen, and H. Zhao, "Robust principal component analysis-based four-dimensional computed tomography.," *Physics in Medicine and Biology*, vol. 56, pp. 3181–98, June 2011.

[14] J. Lee, J. W. Stayman, Y. Otake, S. Schafer, W. Zbijewski, A. J. Khanna, J. L. Prince, and J. H. Siewerdsen, "Volume-of-change cone-beam CT for image-guided surgery.," *Physics in Medicine and Biology*, vol. 57, pp. 4969–4989, Aug. 2012.

[15] R. Henderson, "The potential and limitations of neutrons, electrons and x-rays for atomic resolution microscopy of unstained biological molecules," *Quarterly Reviews of Biophysics*, vol. 28, pp. 171–193, 1995.

[16] O. Ludwig, M. Dimichiel, and L. Salvo, "In-situ three-dimensional microstructural investigation of solidification of an Al-Cu alloy by ultrafast X-ray microtomography," *Metallurgical and Materials Transactions A*, vol. 36A, pp. 1515–1523, June 2005.

[17] P. Bésuelle, G. Viggiani, N. Lenoir, J. Desrues, and M. Bornert, "X-ray micro CT for studying strain localization in clay rocks under triaxial compression," in *Advances in X-ray Tomography for Geomaterials, GeoX2006*, pp. 35–52, ISTE, 2006.

[18] C. Kondo, S. Mori, M. Endo, K. Kusakabe, N. Suzuki, A. Hattori, and M. Kusakabe, "Real-time volumetric imaging of human heart without electrocardiographic gating by 256-detector row computed tomography," *Journal of Computer Assisted Tomography*, vol. 29, pp. 694–698, Sept. 2005.

[19] K. J. Batenburg and L. Plantagie, "Fast approximation of algebraic reconstruction methods for tomography," *IEEE Transactions on Image Processing*, vol. 21, pp. 3648–3658, Aug 2012.

[20] J. Moré, *The Levenberg-Marquardt algorithm: Implementation and theory*, vol. 630 of *Lecture Notes in Mathematics*, ch. 10, pp. 105–116. Springer Berlin / Heidelberg, 1978.

[21] W. Boehm, "Inserting new knots into B-spline curves," *Computer-Aided Design*, vol. 12, pp. 199–201, July 1980.

[22] J. Kaipio and E. Somersalo, "Statistical inverse problems: Discretization, model reduction and inverse crimes," *Journal of Computational and Applied Mathematics*, vol. 198, pp. 493–504, Jan. 2007.

[23] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, pp. 62–66, Jan 1979.

[24] C. D. Boor, "On calculating with B-splines," *Journal of Approximation Theory*, vol. 62, pp. 50–62, 1972.

4

An iterative CT reconstruction algorithm for fast fluid flow imaging

Contents

This chapter has been published as:
G. Van Eyndhoven, K. J. Batenburg, D. Kazantsev, V. Van Nieuwenhove, P. D. Lee, K. J. Dobson, and J. Sijbers, *"An iterative CT reconstruction algorithm for fast fluid flow imaging"*, IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 4446-4458, 2015.

4.1 Introduction

Dynamic computed tomography is a versatile tool for the non-invasive imaging of time-varying objects, as images collected with high temporal frequency allow the visualization of dynamic processes. The CT technique has great potential in fluid flow experiments, where the main goal is to visualize, understand and model the dynamics of the fluid over time. Recent advances in image acquisition speed are now permitting preliminary studies [1, 2, 3], but current temporal resolutions are insufficient to capture high speed behavior in low viscosity fluids. True high speed dynamic CT of multi-phase flow has potential applications across petroleum and geoscience research [4, 5, 6], in civil and environmental engineering [7, 8, 9, 10], as well as for biomedical and materials science applications.

Conventionally, each time point (also referred to as *time frame*) in a fluid flow experiment is reconstructed independently using projection data that was acquired over a full 180° (or 360°) angular range, typically using classical analytical algorithms such as FBP or algebraic algorithms like SIRT. Afterwards, the reconstructed 4D (3D+t) volume can be processed further for quantification of the fluid flow. The main issue with this conventional approach is that the dynamic process should be slow enough to ensure a nearly stable object during the acquisition of all projections at every time frame. If this assumption is violated, blurring artifacts distort the reconstructed images and further quantification becomes difficult. A straightforward approach to increase the temporal resolution is to reduce the scanning time at each time frame. This can be achieved by lowering either the number of acquired projections per time frame or the exposure time per projection. However, reducing the number of projections typically results in limited data artifacts in the reconstructed images while shortening the exposure time results in a decreased signal-to-noise ratio. This implies a trade-off between spatial and temporal resolution, which ultimately limits current fluid flow experiments to experiments with slow temporal dynamics. This is especially true for neutron tomography, which can image very low concentrations of hydrous fluids but has long projection acquisition times.

In the literature, several approaches for improving the temporal resolution in fluid flow imaging by means of an adapted reconstruction algorithm have already been suggested. Most approaches assume that an a priori high quality reconstruction of the dry stage (i.e., the sample without fluid flow) is available. The simulated projection data of this reconstruction is then subtracted from the measured projection data of the dynamic object, resulting in projection data that corresponds solely to the dynamic component in the reconstruction, i.e., the fluid flow. Myers et al. utilized this approach to iteratively reconstruct the fluid flow, while also enforcing prior knowledge about the porous nature of the material matrix and dynamics of the fluid flow, allowing for faster fluid flow imaging [11, 12, 13]. Another way of exploiting the prior image consists of minimizing a sparsity measure on the image difference between the dynamic reconstruction and the prior image. Chen et al. combined this sparsity constraint with a data fidelity term to achieve improved image quality [14]. A more statistical approach, i.e., a maximum a posteriori probability (MAP) estimation method, was presented in [15]. Other approaches do not rely on a prior high quality image, such as the spatial-temporal regularization approach based on non-local means proposed by Kazantsev et al. [16, 17] and the region-based SIRT (rSIRT) method, which assumes only prior knowledge about the location of stationary voxels and dynamic voxels [18].

In this chapter, we present a 4D reconstruction algorithm for fluid flow imaging in which specific models describing the space-time evolution of each voxel are exploited during the reconstruction process. It allows for a substantial reduction of acquired projection data per time frame (thus increasing temporal resolution) while maintaining image quality. The proposed algorithm exploits two types of prior knowledge. Firstly, following the approach of Chapter 3, a dynamic reconstruction is generated assuming the presence of stationary regions (the solid matter) and dynamic regions (the fluid flow) throughout the reconstruction domain. Secondly, corresponding to the actual physical advancing fluid/air boundary, the attenuation of a particular voxel over time in the dynamic region can typically be described by a piecewise constant (PWC) function. As such, the attenuation curves of all voxels in the dynamic region are approximated by PWC functions at intermediate iterations.

In Section 4.2, the region-based SIRT algorithm with intermediate PWC function estimation (rSIRT-PWC) algorithm is introduced. Experiments with simulation phantoms and with a real neutron tomography dataset are reported in Section 4.3. The results are discussed in Section 4.4 and the chapter is concluded in Section 4.5.

4.2 Method

The proposed rSIRT-PWC method is visualized in the flowchart of Fig. 4.1 and pseudo code is available in Algorithm 2. The SUPPRESS algorithm is visualized in the uppermost part of the flowchart in Fig. 2.1 and pseudo-code is available in Algorithm 1. It is designed to frequently exploit two important model assumptions for fluid flow imaging, which are explained in Section 4.2.1. The enforced model assumptions on the reconstruction $\tilde{\boldsymbol{x}}$ result in a more accurately modeled reconstruction problem, which ultimately leads to better reconstruction quality. Starting from an initial estimate $\tilde{\boldsymbol{x}}^{(0)} = \mathbf{0}$, rSIRT-PWC continues with the following steps:

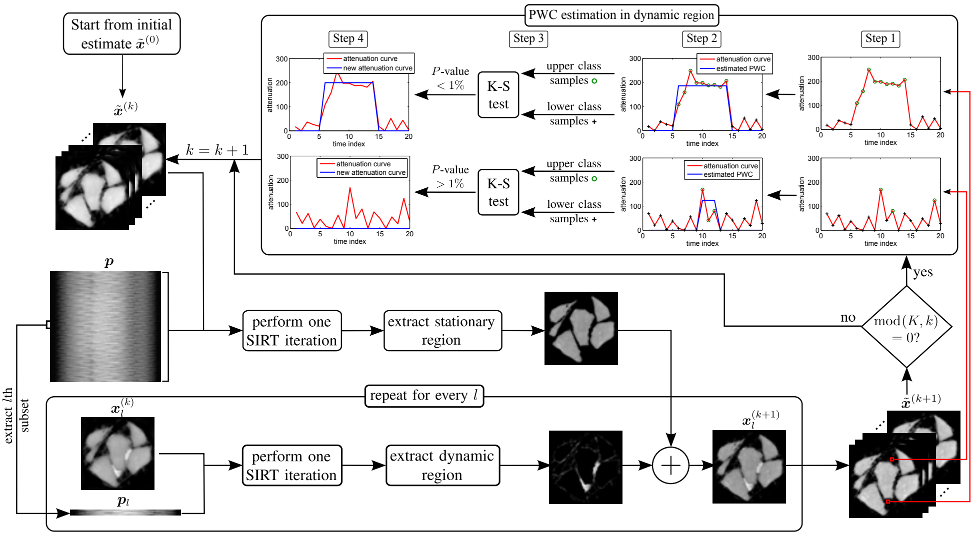


Figure 4.1: Flowchart of the rSIRT-PWC algorithm.

Algorithm 2 Pseudo-code for the rSIRT-PWC algorithm

```
$\tilde{\boldsymbol{x}}^{(0)} \leftarrow \mathbf{0}$
for $k = 0, \ldots, K_{\text{tot}} - 1$ do
    for $r = 1, \ldots, R$ do
        $\boldsymbol{x}_r^{(k+1)} \leftarrow \boldsymbol{x}_r^{(k)} + \boldsymbol{I}_S \boldsymbol{C} \boldsymbol{W}^T \boldsymbol{R}(\tilde{\boldsymbol{p}} - \tilde{\boldsymbol{W}} \tilde{\boldsymbol{x}}^{(k)})$
                        $+ \boldsymbol{I}_V \boldsymbol{C}_r \boldsymbol{W}_r^T \boldsymbol{R}_r (\boldsymbol{p}_r - \boldsymbol{W}_r \boldsymbol{x}_r^{(k)})$
    end for
    if $\mathrm{mod}\ (K, k) = 0$ then
        $\tilde{\boldsymbol{x}}^{(k+1)} \leftarrow$ replace the attenuation curves in the dynamic region
                    by appropriate PWC functions (see Section 4.2.3)
    end if
end for
```

1. First, K rSIRT iterations are executed, which is visualized in the bottom most part of the flowchart in Fig. 4.1.

2. Next, PWC functions are estimated for all voxels in the dynamic region. This is illustrated in the upper most part of the flowchart in Fig. 4.1.

3. Go back to step 1.

This process is repeated iteratively for K_{tot} iterations. The first step, i.e., the rSIRT iterations, is based on the first model assumption of Section 4.2.1 and is explained more thoroughly in Section 4.2.2. The second step in which the PWC functions are estimated is based on the second model assumption of Section 4.2.1 and is elaborated in Section 4.2.3.

In this chapter, the constant window approach for selecting subsets of the projection data is adapted (see Fig. 3.2 in Chapter 3). It is assumed that for each time frame $r \in \{1, \ldots, R\}$ multiple projection images were acquired (thus, $M_t > 1$, see Section 1.3). Therefore, the projection matrix $\boldsymbol{W}_r$ ($r \in \{1, \ldots, R\}$), introduced in Section 1.3, corresponds to the projection along M_t acquisition angles.

4.2.1 Model assumptions

In fluid flow experiments, several assumptions about the scanned object can be made.

4.2.1.1 The presence of stationary regions

The scanned object is assumed to consist of stationary regions (the solid matter) and regions that change over time, i.e., dynamic regions (the fluid flow). In mathematical terms, this assumption means that there is a set $S \subset \{1, \ldots, N\}$ of voxel indices that correspond to the stationary regions, such that $\boldsymbol{x}_r(j) = \boldsymbol{x}_{r'}(j)$ for all $j \in S$ and $r, r' \in \{1, \ldots, R\}$. This is not necessarily true for voxels in the dynamic region, of which the voxel indices belong to $V := \{1, \ldots, N\} \backslash S$. These sets ($S$ and V) can be calculated prior to the actual fluid flow experiment by generating a segmentation of the solid matter based on a reconstruction of the object before the fluid flow initiated.

In what follows, the set V is further partitioned into two distinct subsets $V = V_B \cup V_F$ with $V_B \cap V_F = \emptyset$. The set V_B contains the indices corresponding to voxels on the border between the dynamic and the stationary region and the set $V_F = V \backslash V_B$ contains indices corresponding to voxels that are fully inside the dynamic region.

4.2.1.2 Two-phase incompressible fluid flow

Since a two-phase incompressible fluid flow is imaged, voxels in the dynamic region can only contain fluid or air. Furthermore, the attenuation value of a homogeneous incompressible fluid is a fixed value (in space and time). This means that, in the dynamic region, the attenuation over time of a particular voxel, i.e., its *attenuation curve*, can be modeled by a PWC function with one fixed attenuation value for the fluid and a zero attenuation value for air, in full accordance with the actual physically advancing fluid/air boundary. That is, in the first time frame, a voxel will contain air. At a certain point in time, fluid will enter the voxel and the voxel's attenuation value will change to the attenuation value of the fluid. Once the fluid leaves the voxel, the attenuation value will return to zero. This can be modeled by a PWC function. In what follows, the attenuation value of the fluid is denoted as a_f.

4.2.2 region-based SIRT (rSIRT)

The rSIRT algorithm in the constant window approach, which was introduced in Chapter 3 and is utilized throughout this chapter, is shortly repeated here. Let $\boldsymbol{I}_V \in \{0,1\}^{N\times N}$ be the binary diagonal matrix representing the operator that sets all voxels belonging to the stationary region to 0. Its diagonal elements are given by $\boldsymbol{I}_V(j,j) = \chi_V(j)$ where χ_V is the characteristic function for the set V. Analogously, the binary diagonal matrix $\boldsymbol{I}_S \in \{0,1\}^{N\times N}$ is defined as the operator setting all voxels in the dynamic region to 0, i.e., $\boldsymbol{I}_S(j,j) = \chi_S(j)$. Finally, define $\boldsymbol{R}_r \in \mathbb{R}^{M\times M}$ and $\boldsymbol{C}_r \in \mathbb{R}^{N\times N}$ as the diagonal matrices with inverse row sums and inverse column sums of $\boldsymbol{W}_r$, respectively. The introduced notations allow us to describe the rSIRT algorithm as the following iterative process:

$$\begin{aligned}\boldsymbol{x}_r^{(k+1)} = \boldsymbol{x}_r^{(k)} \quad &+ \quad \boldsymbol{I}_S\boldsymbol{C}\boldsymbol{W}^T\boldsymbol{R}(\tilde{\boldsymbol{p}} - \tilde{\boldsymbol{W}}\tilde{\boldsymbol{x}}^{(k)}) \\ &+ \quad \boldsymbol{I}_V\boldsymbol{C}_r\boldsymbol{W}_r^T\boldsymbol{R}_r(\boldsymbol{p}_r - \boldsymbol{W}_r\boldsymbol{x}_r^{(k)}).\end{aligned} \tag{4.1}$$

This update needs to be calculated for every $r \in \{1,\ldots,R\}$ before incrementing the iteration count k. Eq. 4.1 calculates a traditional SIRT update for the stationary region using all available projection data $\tilde{\boldsymbol{p}}$ and then a second update for the dynamic region using only the projection data from the relevant time frame.

4.2.3 Piecewise constant function estimation

In the dynamic region, the attenuation curves of each voxel are replaced by PWC functions at intermediate iterations. The PWC functions are estimated with a different approach depending on the position of the specific voxel. The different types of distinguished voxel positions are displayed in Fig. 4.2.

The set V_B corresponds to voxels of type #1. The voxel types in the fully dynamic region V_F, i.e., voxel types #2, #3 and #4, are assigned during each PWC estimation step, which is explained below.

4.2.3.1 Voxels in the fully dynamic region V_F

The PWC estimation for voxels in the fully dynamic region V_F is subdivided into 4 steps (upper panel in Fig. 4.1).

In the *first step*, the time attenuation curve (TAC) is extracted for each voxel with index $j \in V_F$. It is defined by $TAC_j(t_r) = \boldsymbol{x}_r(j)$, where t_r represents the

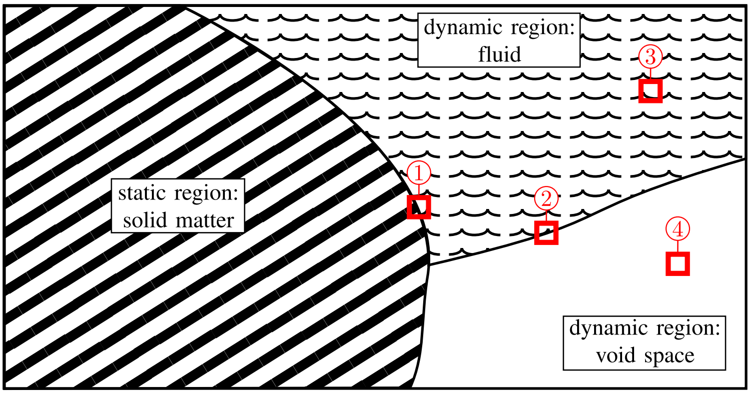


Figure 4.2: Illustration of the different types of voxels. Voxel type #1 is situated on the border between the dynamic and the stationary region. Voxels of type #3 are in the fully dynamic region and contain fluid at some point in time. Voxels of type #4 are also classified as dynamic but never contain fluid. Voxel type #2 is located on the border of the fluid and void space in the fully dynamic region.

time corresponding to the r^{th} time frame. For each of these attenuation curves, a two-class Otsu segmentation [19] of the attenuation values is calculated.

In the *second step*, the mean of the attenuation values in the upper Otsu class, defined as M_j, is utilized to define the PWC function

$$PWC_{t_a,t_b}^{0,M_j}(t) = \begin{cases} M_j & \text{if } t_a < t < t_b \\ 0 & \text{otherwise} \end{cases} , \tag{4.2}$$

where t_a and t_b are discrete parameters that can be chosen from the finite set $\{t_1, \ldots, t_R\}$, indicating the time points at which the fluid enters and exits the voxel. Note that the PWC function was defined with the mean M_j of the upper Otsu segmentation class rather than the fluid's attenuation value a_f, because iterative algorithms like SIRT typically underestimate the higher attenuation values in the first iterations. This function is used to approximate the extracted attenuation curves TAC_j by solving the following problem:

$$(t_a, t_b) = \underset{t_a,t_b}{\operatorname{argmin}} \sum_{r=1}^{R} (PWC_{t_a,t_b}^{0,M_j}(t_r) - TAC_j(t_r))^2 \quad . \tag{4.3}$$

Eq. 4.3 is solved by testing all possible combinations of t_a and t_b.

In the *third step*, a statistical test is performed to check whether or not the estimated PWC is relevant. For this purpose, the attenuation values in the original attenuation curve for which $t_a < t < t_b$ (referred to as upper class samples) and for which $(t < t_a)|(t > t_b)$ (referred to as lower class samples) are extracted. With these two sets of samples, a Kolmogorov-Smirnov (K-S) test is performed, which checks the nul hypothesis that claims that the upper and lower class samples are drawn from the same underlying continuous population [20]. This nul hypothesis is rejected at a significance level of 1%, i.e., if the associated P-value is less than 1%. With a rejection of the nul hypothesis, the two sets of samples (the upper and the lower class) are regarded as sufficiently separated and hence the estimated PWC is regarded as relevant, i.e., it does not fit the noise in the extracted attenuation curve.

In the *fourth step*, the attenuation curves are replaced by different PWC functions depending on the type of voxel (type #2, #3 or #4). To obtain the type of each voxel, the fully dynamic region V_F is subdivided using two conditions that determine if a voxel contains fluid over time:

1. The K-S test, performed in step 3, results in a rejection of the nul hypothesis at a 1% significance level. This indicates that the different classes are sufficiently separated.

2. The upper class mean attenuation value (i.e., the mean of TAC_j for time points satisfying $t_a < t < t_b$) is larger than $a_f/2$.

The region of voxels that contain fluid over time is defined by voxels that satisfy both conditions. Voxels on the border of this region are assigned type #2 and voxels that are fully contained in this region (i.e., not on the border) are assigned type #3. If the nul hypothesis is not rejected at the 1% significance level or the mean of the upper class is smaller than $a_f/2$, the voxel is assigned type #4.

Since voxels of type #2 contain both fluid and air, their attenuation curve is replaced by the PWC curve defined in Eq. 4.2. Voxels of type #3 are either fully emersed by fluid or contain only void space, therefore their corresponding attenuation curves are replaced by

$$PWC^{0,a_f}_{t_a,t_b}(t) = \begin{cases} a_f & \text{if } t_a < t < t_b \\ 0 & \text{otherwise} \end{cases} . \tag{4.4}$$

Voxels of type #4 never contain any fluid and are hence replaced by zero.

4.2.3.2 Voxels in the stationary/dynamic region V_B

The four steps of the PWC estimation procedure introduced previously are slightly adjusted. Since voxels in the stationary/dynamic region V_B can contain a combination of solid matter and fluid or air, the use of a PWC function as defined in Eq. 4.2 is incorrect. Instead, the mean of the lower Otsu class m_j is calculated in step 1 and the attenuation curves are approximated by the PWC function

$$PWC_{t_a,t_b}^{m_j,M_j}(t) = \begin{cases} M_j & \text{if } t_a < t < t_b \\ m_j & \text{otherwise} \end{cases} \tag{4.5}$$

in step 2. The same K-S test is performed in step 3. Finally, in step 4, the attenuation curves are replaced by Eq. 4.5 if the null hypothesis is rejected at the 1% significance level and by their mean if the null hypothesis is accepted.

4.3 Experiments

In this section, various experiments for the validation of the rSIRT-PWC algorithm are described. First, the figures of merit are introduced in Section 4.3.1 and the reconstruction methods to which rSIRT-PWC is compared are described in Section 4.3.2. Next, in Section 4.3.3, different simulation experiments are introduced. An experiment with real neutron tomography data is described in Section 4.3.4. The results for all these experiments are reported in Section 4.4.

4.3.1 Figures of merit

The Relative Root Mean Squared Error (RRMSE) is utilized as a quality measure and is defined as

$$RRMSE(\hat{\boldsymbol{x}}) = \sqrt{\frac{\sum_i (\hat{\boldsymbol{x}}(i) - \tilde{\boldsymbol{x}}(i))^2}{\sum_i (\tilde{\boldsymbol{x}}(i))^2}}\ , \tag{4.6}$$

where $\hat{\boldsymbol{x}}$ denotes the calculated reconstruction and $\tilde{\boldsymbol{x}}$ denotes the ground truth phantom. For some experiments, the RRMSE is inspected only in a Region of Interest (ROI), in which case the sum in Eq. 4.6 sums over all points in time and over all voxels in the specific ROI.

4.3.2 Reconstruction methods

The rSIRT-PWC method is compared to the following reconstruction methods.

- *SIRT*: The SIRT algorithm individually applied to the projection data associated with each time frame.
- *SIRT PWC*: In order to have a fair comparison, the SIRT reconstruction was also post-processed with our PWC function estimation method. We refer to this reconstruction as SIRT PWC.
- *rSIRT*: The region-based SIRT algorithm (see Section 4.2.2 and [18]).

- *Myers et al.*: The fluid flow reconstruction method introduced by Myers et al. and implemented as described in [11]. For the numerical experiments, the reconstructions are modeled on a 200×200 voxel grid. Therefore, following Myers et al., the static reconstruction was calculated with FBP based on the full set of $315 \approx 200 * \pi/2$ projections, which guarantees accurate image quality [21]. These projections were simulated from the ground truth image assuming an incoming beam intensity of $I_0 = 5 \times 10^4$ (photon count), which is 10 times the number of photons that were assumed for generating the projections of the dynamically evolving object. Also, the soft thresholding parameter and convergence tolerance (which is denoted by ϵ in [11]) were optimized by selecting those parameters that gave the lowest RRMSE with respect to the ground truth. This optimal parameter selection was repeated for each and every experiment. That is, whenever a different number of projections per time frame was utilized, new optimal parameters were selected, thereby applying the method by Myers et al. at its full strength. In the neutron tomography experiment, there is no ground truth available. Therefore, the algorithmic parameters were chosen manually by visual assessment for the neutron data. All other parameters were chosen as described in [11].

- *CGLS-NLST*: The Conjugate Gradient Least Squares method with Non-Local Spatio-Temporal penalty (CGLS-NLST) developed by Kazantsev et al. and implemented as described in [17]. The reader is referred to [17] for details about the method and its parameters. The prior image containing structural information was set to be the same static reconstruction that was utilized in the method by Meyers et al. Most algorithm parameters were adapted from [17], except for the number of iterations, the regularization parameter and the noise-dependent parameter (denoted by $MaxOuter$, β and h in [17], respectively), since these parameters are problem dependent. These three parameters were optimized for lowest RRMSE with respect to the ground truth for each and every experiment in the simulation experiments and selected manually based on visual assessment in the neutron tomography experiment. The parameter expressing the level of trust (strength of smoothing) in the dynamic data (denoted by γ in [17]) was also manually tuned in the neutron tomography experiment.

The rSIRT-PWC algorithm was implemented as described in Section 4.2 with the PWC estimation applied every 20^{th} iteration starting from the 60^{th} iteration. Every method was implemented with a positivity constraint, i.e., voxels with an attenuation value smaller than zero are set to zero after each iteration. All reconstruction methods were applied with 200 iterations, with the exception of the CGLS-NLST method, where this parameter was optimized for lowest RRMSE with respect to the ground truth image. The algorithms were implemented with the

ASTRA toolbox [22, 23, 24].

4.3.3 Numerical simulations

Two different simulation phantoms were created. In the simulation experiments, the projections were generated with a strip kernel [25] and a parallel beam geometry. The data was simulated from a higher resolution version of the phantom, i.e., on a 400×400 isotropic voxel grid. The number of detector bins was set to 200, with each detector pixel of twice the size as the voxel size in the high resolution version of the phantom. Projection angles were selected with the golden ratio scanning scheme. Poisson distributed noise was applied to the projection data, assuming an incoming beam intensity of 5×10^3 (photon count). Reconstructions were calculated on a 200×200 isotropic voxel grid and with a linear projection model [25], to avoid the "inverse crime" of generating the data with the same model as the model for calculating the reconstruction [26].

4.3.3.1 Shepp-Logan phantom

The first phantom is a Shepp-Logan type phantom in which fluid flows from one chamber into another during the acquisition of the CT data. The phantom was simulated on 20 time frames, of which frame 1, 10 and 20 are displayed in Fig. 4.3(a), (b) and (c), respectively. The mask that separates the stationary from the dynamic voxels is displayed in Fig. 4.3(d). For each time frame, 10 projections were simulated.

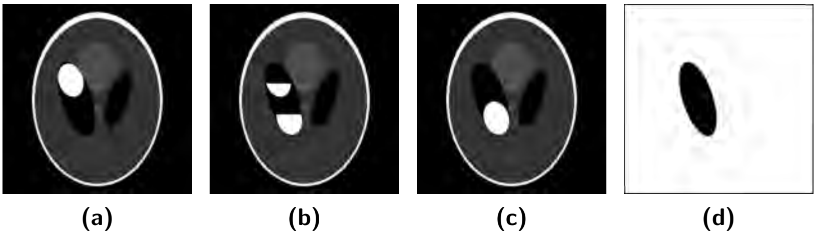
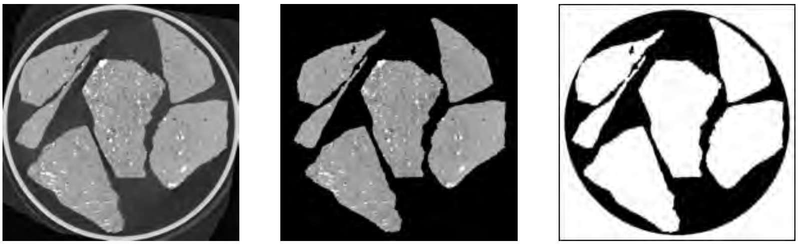

(a) (b) (c) (d)

Figure 4.3: The adjusted Shepp-Logan phantom on time frame 1, 10 and 20 (a-c) and the mask for stationary region (d), where the white and black region correspond to the stationary and the dynamic region, respectively.

4.3.3.2 Porous rock phantom

The second simulation phantom was created from a high quality FBP reconstruction of an X-ray tomography dataset of rock (porous gravel) acquired on a Nikon



(a) High quality FBP reconstruction **(b)** Processed FBP reconstruction **(c)** Mask for the stationary region

Figure 4.4: The high quality reconstruction of the porous gravel (a), its segmented version (b) and the corresponding mask for the stationary region (c).

XTH 225 ST scanner at the Manchester X-ray Imaging Facility (Fig. 4.4a). In this reconstruction, all voxels that do not belong to the rock (the void volume) were set to zero (Fig. 4.4b) and fluid flow was simulated in the void space. The time point at which the fluid enters a certain voxel was randomly generated by a properly scaled 2D Perlin noise image [27], the same approach was utilized to select the time point at which the fluid leaves a certain voxel. The stationary region (Fig. 4.4c) was defined as the region containing rock and the region outside the sample container. The phantom was simulated on 20 time frames, of which frame 1, 4, 5, 10, 15 and 20 are displayed in the first row of Fig. 4.8. In this experiment, 20 projections were simulated per time frame.

4.3.4 Neutron tomography dataset

A neutron tomography dataset was acquired at the cold neutron imaging Beamline ICON at the SINQ spallation neutron source, Paul Scherrer Institute, Switzerland. Granitic gravel particles with a 5-10 mm diameter were loaded into a 25 mm thin walled Al tube and mounted in a gravity driven flow cell. Parallel beam projection images were acquired under the golden ratio scanning scheme [28] with an exposure time of 20 seconds per projection. After a 2×2 rebinning, the dimension of the detector was given by 1023×1030 pixels with pixel size 26 μm. Reconstructions were calculated on the central slice on a 1030×1030 pixel grid, also with pixel size 26 μm. Initially, the sample was scanned in its dry stage (no fluid flow) at 154 projection angles and a reconstruction was calculated with 200 SIRT iterations (Fig. 4.5a). A mask for the stationary region (Fig. 4.5b) was extracted from this SIRT reconstruction. Next, after starting the fluid flow, another 326 “wet stage” projections were acquired. The projection data of the dry stage and the wet stage were combined to form a projection data set corresponding to 480 projection

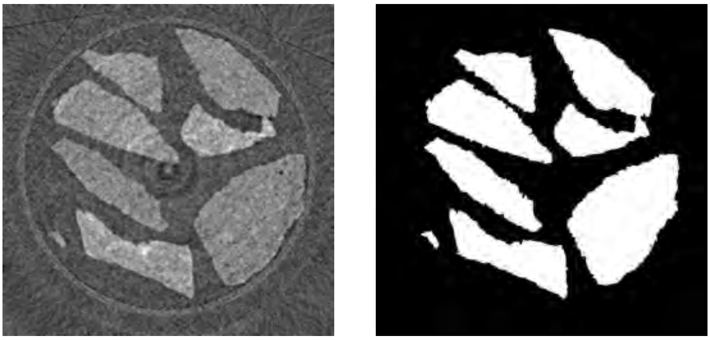

(a) Dry stage SIRT reconstruction **(b)** Mask for the stationary voxels

Figure 4.5: The dry stage reconstruction (a) and the extracted mask (b) of the gravel particles in the neutron tomography experiment.

angles. This set of projections served as the input for all evaluated reconstruction methods. Since angles were selected with the golden ratio scanning scheme, the projection data can be subdivided in arbitrary subsets which allows for selecting the number of time frames as desired.

The stationary region mask in Fig. 4.5b was utilized as prior knowledge in rSIRT, rSIRT-PWC and the method by Myers et al. The latter method and CGLS-NLST were applied with the dry stage SIRT reconstruction (Fig. 4.5a) as the prior image.

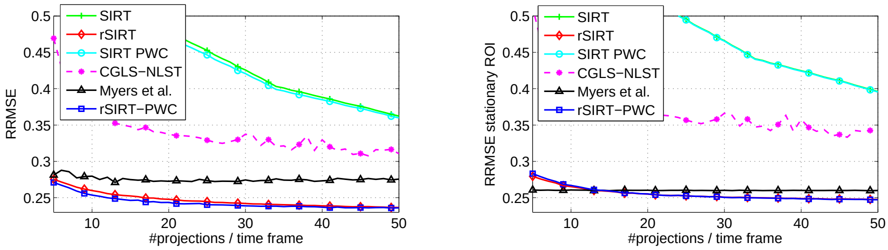
4.4 Results

4.4.1 Shepp-Logan phantom

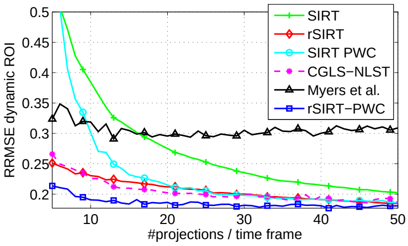
Frame 10 of the reconstructions for all algorithms are displayed in Fig. 4.6. The reconstruction method by Myers et al. is based on more projection data (i.e., the projection data to generate the prior image and the projection data acquired from the dynamically evolving object) in comparison to the other methods. A fully fair comparison is hence only possible if the rSIRT-PWC method is allowed to use this projection data as well. For this reason, the reconstruction in Fig. 4.6g was created, it is the rSIRT-PWC reconstruction generated from the combination of the projection data used in the generation of the prior image for the method by Myers et al. and the projection data acquired from the dynamically evolving object. From Fig. 4.6, it is obvious that the standard SIRT algorithm suffers from limited data artifacts. This is improved by rSIRT (Fig. 4.6c) and substantially improved

(a) Ground truth **(b)** SIRT **(c)** rSIRT **(d)** CGLS-NLST **(e)** Myers et al. **(f)** rSIRT-PWC **(g)** rSIRT-PWC (all data)

Figure 4.6: Reconstructions (displayed in time frame 10) for the simulation experiment with the adjusted Shepp-Logan phantom. All reconstructions were calculated based on 10 simulated projections per time frame.



(a) RRMSE calculated in the full reconstruction domain **(b)** RRMSE calculated in the stationary region



(c) RRMSE calculated in the dynamic region

Figure 4.7: The RRMSE as function of the number of projections per time frame for SIRT, rSIRT, the method by Myers et al. and rSIRT-PWC for the Shepp-Logan phantom.

by CGLS-NLST, the method of Myers et al. and rSIRT-PWC (Fig. 4.6(d-g)). Furthermore, the rSIRT-PWC has better image quality in the dynamic region in comparison to CGLS-NLST and the method by Myers et al. These observations are confirmed by the RRMSE values in Table 4.1. Since the reconstruction by the method of Myers et al. is based on a prior high quality reconstruction, it is

	SIRT	*SIRT PWC*	*rSIRT*	*CGLS-NLST*	*Myers et al.*	*rSIRT-PWC*
full ROI	0.5938	0.5819	0.2598	0.3806	0.2796	**0.2536**
stat. ROI	0.6414	0.6406	0.2653	0.4117	**0.2607**	0.2669
dyn. ROI	0.3845	0.3014	0.2304	0.2360	0.3187	**0.1905**

Table 4.1: RRMSE values for the Shepp-Logan phantom evaluated for different reconstructions (columns) in the full reconstruction domain (first row), in the stationary ROI (middle row) and in the dynamic ROI (last row). In this experiment, 10 projections were simulated per time frame.

obvious that it has better image quality in the stationary region.

To test the algorithms' performance with respect to the amount of available projection data, the previous experiment was repeated for a varying number of projections per time frame, while keeping all other experimental parameters the same. The RRMSE for each of these experiments is plotted as function of the number of projections per time frame in Fig. 4.7. From these plots, it is obvious that rSIRT-PWC outperforms all other reconstruction methods in almost all scenarios. One exception is the stationary region's image quality if only few projection data (less than 15 projections per time frame) is available: in this case the method by Myers et al. gives better results than the proposed rSIRT-PWC method. This is again due to the fact that the stationary region's reconstruction for the method by Myers et al. is based on more projection data (315 projections) than the rSIRT-PWC reconstruction. As soon as sufficient projection data becomes available (more than 15 projections per time frame), the rSIRT-PWC method also outperforms the method of Myers et al. with respect to image quality in the stationary region. Furthermore, by comparing the RRMSE values at 50 projections per time frame for the SIRT method and 5 projections per time frame for the rSIRT-PWC method in Fig. 4.7c, it can be observed that the rSIRT-PWC reconstruction method achieves comparable image quality in the dynamic region with up to an order of magnitude fewer projections than the conventional SIRT method.

4.4.2 Porous rock phantom

The reconstructions in time frame 1, 4, 5, 10, 15 and 20 for all algorithms are displayed in Fig. 4.8. The RRMSE values can be found in Table 4.2. It is clear, both visually from Fig. 4.8 and numerically from Table 4.2, that a conventional reconstruction approach such as SIRT suffers from limited data artifacts if only 20 projections are available per time frame. The rSIRT reconstruction can more

Figure 4.8: Ground truth images and reconstructions of the porous rock phantom showing the evolution through time. The columns represent different time frames and the rows represent the ground truth phantom and the different reconstruction methods. All reconstructions are based on 20 simulated projections in each time frame.

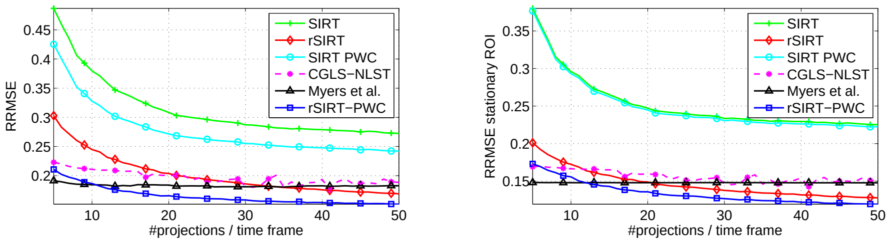
accurately reconstruct the stationary region. However, mainly because the lack of model restrictions in the dynamic region, the rSIRT reconstruction is strongly influenced by noise in the dynamic region. The CGLS-NLST method improves image quality in both stationary and dynamic region, but due to the high level of noise in the projection data and the low number of projections, the optimal smoothing parameters were rather large and hence the fine structures are partially

	SIRT	*SIRT PWC*	*rSIRT*	*CGLS-NLST*	*Myers et al.*	*rSIRT-PWC*
full ROI	0.3084	0.2716	0.2037	0.2004	0.1839	**0.1655**
stat. ROI	0.2472	0.2446	0.1488	0.1597	0.1479	**0.1357**
dyn. ROI	0.8999	0.4974	0.6358	0.4984	0.4637	**0.3330**

Table 4.2: RRMSE values for the porous rock phantom evaluated in the full reconstruction domain (first row), in the stationary ROI (middle row) and in the dynamic ROI (last row). The different columns represent different reconstruction algorithms. In this experiment, 20 projections were simulated per time frame.

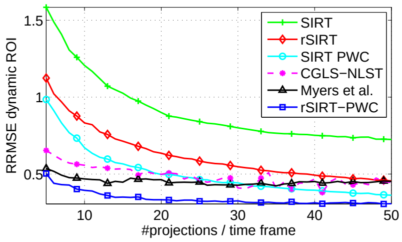
erased. The method by Myers et al. also has improved image quality in both the stationary and the dynamic region. The rSIRT-PWC reconstruction, however, has even better image quality. The fluid's dynamics are captured more correctly and the region on the edge between stationary and dynamic region has a better correspondence to the ground truth image.

The previous experiment was repeated for a varying number of projections per time frame, while keeping all other experimental parameters the same. The RRMSE as function of the number of projections per time frame can be observed in Fig. 4.9. It can be concluded that the proposed rSIRT-PWC algorithm outperforms all other algorithms. Notice that the stationary region of the reconstruction by Myers et al. has better image quality for a low number of projections per time frame (Fig. 4.9b). This is again due to the fact that this region of the reconstruction domain is based on the prior image, which was reconstructed a priori with FBP based on 315 projection images. The rSIRT-PWC algorithm does not utilize this data. However, starting from around 15 projections per time frame (which corresponds to a total of 300 projections for all 20 time frames) the rSIRT-PWC stationary region reconstruction has better quality than the reconstruction of Myers et al. At that point, rSIRT-PWC has acces to approximately an equal number of projections for reconstructing the stationary region and due to the fact that rSIRT-PWC is based on an iterative technique (which generally outperforms FBP) the reconstruction quality in the stationary region becomes better. Analogously to the Shepp-Logan experiment, it can also be observed in Fig. 4.9c that the rSIRT-PWC reconstruction method achieves comparable image quality in the dynamic region with up to an order of magnitude fewer projections than the other methods.



(a) RRMSE calculated in the full reconstruction domain

(b) RRMSE calculated in the stationary region



(c) RRMSE calculated in the dynamic region

Figure 4.9: The RRMSE as function of the number of projections per time frame for SIRT, rSIRT, the method by Myers et al. and rSIRT-PWC for the porous rock phantom.

4.4.3 Neutron tomography dataset

The projection dataset was first subdivided into 48 time frames of 10 projections each. The corresponding reconstructions generated by SIRT, rSIRT, CGLS-NLST, the method by Myers et al. and rSIRT-PWC on time frame 1, 30, 31 and 48 are displayed in Fig. 4.10. The reconstructions generated by SIRT and rSIRT are heavily influenced by both noise and artifacts in the dynamic region, as are those produced by the CGLS-NLST and Myers et al. method because the prior image does not have perfect quality in this experiment. The rSIRT-PWC reconstruction has good contrast both in the stationary and the dynamic regions.

In a second experiment, the projection dataset was respectively subdivided into 10 subsets of 48 projections each, 20 subsets of 24 projections each, 30 subsets of 16 projections each and 48 subsets of 10 projections each. The resulting reconstructions on the last time frame can be observed in Fig. 4.11. The rSIRT-PWC reconstruction (bottom row in Fig. 4.11) is the least affected by having small numbers of projections available per time frame, which illustrates the ability of rSIRT-PWC to increase the temporal resolution without affecting the image quality.

4.5 Conclusion

Capturing the high speed dynamics of fluid flow by means of CT imaging requires a short acquisition time, which can be achieved by acquiring only few projection images per time frame. However, reconstructing data from undersampled projection data is a difficult problem, and conventional approaches that reconstruct the object independently at different time frames result in images containing limited data artifacts.

In this chapter, the rSIRT-PWC algorithm was introduced, an iterative method tailored specifically to fluid flow reconstruction problems. The algorithm divides the reconstruction domain into stationary (the solid matter) and dynamic (the fluid flow) regions, and assumes the shape of the attenuation curves in the dynamic region to be piecewise constant in accordance with a physical advancing air-fluid boundary. Since the reconstruction problem is modeled more accurately, the size of the solution space is substantially reduced and the final image quality improves. Therefore, the rSIRT-PWC algorithm allows for a significant reduction in the number of projections per time frame without image quality loss. The rSIRT-PWC reconstruction method achieves comparable image quality in the dynamic region with up to an order of magnitude fewer projections than conventional methods. It therefore provides a much-needed method for probing high speed fluid dynamics.

References

[1] S. Iglauer, A. Paluszny, and M. Blunt, “Corrigendum to simultaneous oil recovery and residual gas storage: A pore-level analysis using in situ x-ray micro-tomography [fuel 103 (2013) 905914],” *Fuel*, vol. 139, no. 0, pp. 780 – 780, 2015.

[2] M. Andrew, B. Bijeljic, and M. J. Blunt, “Pore-by-pore capillary pressure measurements using X-ray microtomography at reservoir conditions: Curvature, snap-off, and remobilization of residual CO2,” *Water. Resour. Res.*, vol. 50, no. 11, pp. 8760–8774, 2014.

[3] M. Andrew, B. Bijeljic, and M. J. Blunt, “Pore-scale contact angle measurements at reservoir conditions using X-ray microtomography,” *Advances in Water Resources*, vol. 68, no. 0, pp. 24 – 31, 2014.

[4] V. Cnudde and M. N. Boone, “High-resolution X-ray computed tomography in geosciences: A review of the current technology and applications,” *Earth-Science Reviews*, vol. 123, pp. 1–17, 2013.

[5] R. Dann, M. Turner, M. Close, and M. Knackstedt, “Multi-scale characterisation of coastal sand aquifer media for contaminant transport using X-ray computed tomography,” *Environ. Earth Sci.*, vol. 63, no. 5, pp. 1125–1137, 2011.

[6] A. N. Golab, M. A. Knackstedt, H. Averdunk, T. Senden, A. R. Butcher, and P. Jaime, “3D porosity and mineralogy characterization in tight gas sandstones,” *The Leading Edge*, vol. 29, no. 12, pp. 1476–1483, 2010.

[7] M. Slavíková, F. Krejcí, J. Žemlicka, M. Pech, P. Kotlík, and J. Jakubek, “X-ray radiography and tomography for monitoring the penetration depth of consolidants in opuka the building stone of prague monuments,” *Journal of Cultural Heritage*, vol. 13, no. 4, pp. 357 – 364, 2012.

[8] S. Raneri, V. Cnudde, T. De Kock, H. Derluyn, G. Barone, and P. Mazzoleni, “X-ray computed micro-tomography to study the porous structure and degradation processes of a building stone from Sabucina (Sicily),” *Euro. J. of Mineral.*, 2015, to be published.

[9] A. Polak, D. Elsworth, H. Yasuhara, A. S. Grader, and P. M. Halleck, “Permeability reduction of a natural fracture under net dissolution by hydrothermal fluids,” *Geophys. Res. Lett.*, vol. 30, no. 20, 2003.

[10] K. Singh, R. K. Niven, T. J. Senden, M. L. Turner, A. P. Sheppard, J. P. Middleton, and M. A. Knackstedt, “Remobilization of residual non-aqueous phase liquid in porous media by freeze-thaw cycles,” *Environ. Sci. Technol.*, vol. 45, no. 8, pp. 3473–3478, 2011.

[11] G. R. Myers, A. M. Kingston, T. K. Varslot, M. L. Turner, and A. P. Sheppard, “Dynamic tomography with a priori information,” *Appl. Optics*, vol. 50, no. 20, pp. 3685–3690, 2011.

[12] G. R. Myers, A. M. Kingston, T. K. Varslot, M. L. Turner, and A. P. Sheppard, “Dynamic X-ray micro-tomography for real time imaging of drainage and imbibition processes at the pore scale,” in *Int. Symp. of the Soc. of Core Analysts, Austin, TX*, pp. 1–12, 2011.

[13] G. Myers, T. Varslot, A. Kingston, A. Herring, and A. Sheppard, “Ground-truth verification of dynamic X-ray micro-tomography images of fluid displacement,” in *Proc. SPIE, Developments in X-ray Tomography VIII*, vol. 8506, p. 85060P, 2012.

[14] G.-H. Chen, J. Tang, and S. Leng, “Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets,” *Med. Phys.*, vol. 35, no. 2, pp. 660–663, 2008.

[15] G. R. Myers, M. Geleta, A. M. Kingston, B. Recur, and A. P. Sheppard, “Improving dynamic tomography, through maximum a posteriori estimation,” in *Proc. SPIE, Developments in X-Ray Tomography IX*, vol. 9212, p. 921211, 2014.

[16] D. Kazantsev, W. M. Thompson, W. R. B. Lionheart, G. Van Eyndhoven, A. P. Kaestner, K. J. Dobson, P. J. Withers, and P. D. Lee, "4D-CT reconstruction with unified spatial-temporal patch-based regularization," *Inverse Probl. Imag.*, vol. 9, no. 2, pp. 447–467, 2015.

[17] D. Kazantsev, G. Van Eyndhoven, W. R. B. Lionheart, P. J. Withers, K. J. Dobson, S. A. McDonald, R. Atwood, and P. D. Lee, "Employing temporal self-similarity across the entire time domain in computed tomography reconstruction," *Philosophical Transactions of the Royal Society A*, vol. 373, no. 2043, 2015.

[18] G. Van Eyndhoven, K. J. Batenburg, and J. Sijbers, "Region-based iterative reconstruction of structurally changing objects in CT," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 909–919, 2014.

[19] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, pp. 62–66, Jan 1979.

[20] F. J. Massey, "The Kolmogorov-Smirnov test for goodness of fit," *J. Am. Stat. Assoc.*, vol. 46, no. 253, pp. 68–78, 1951.

[21] F. Natterer, *The mathematics of computerized tomography*, vol. 32. SIAM, 1986.

[22] W. J. Palenstijn, K. J. Batenburg, and J. Sijbers, "Performance improvements for iterative electron tomography reconstruction using graphics processing units (GPUs)," *Journal of Structural Biology*, vol. 176, no. 2, pp. 250 – 253, 2011.

[23] W. J. Palenstijn, K. J. Batenburg, and J. Sijbers, "The ASTRA tomography toolbox," in *13th International Conference on Computational and Mathematical Methods in Science and Engineering, CMMSE*, 2013.

[24] W. van Aarle, W. J. Palenstijn, J. De Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers, "The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, vol. 157, pp. 35 – 47, 2015.

[25] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging.* Society of Industrial and Applied Mathematics, 2001.

[26] J. Kaipio and E. Somersalo, "Statistical inverse problems: Discretization, model reduction and inverse crimes," *Journal of Computational and Applied Mathematics*, vol. 198, pp. 493–504, Jan. 2007.

[27] K. Perlin, "Improving noise," *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 681–682, 2002.

[28] A. Kaestner, B. Munch, T. Pavel, and L. Butler, "Spatiotemporal computed tomography of dynamic processes," *Optical Engineering*, vol. 50, no. 12, p. 123201, 2011.

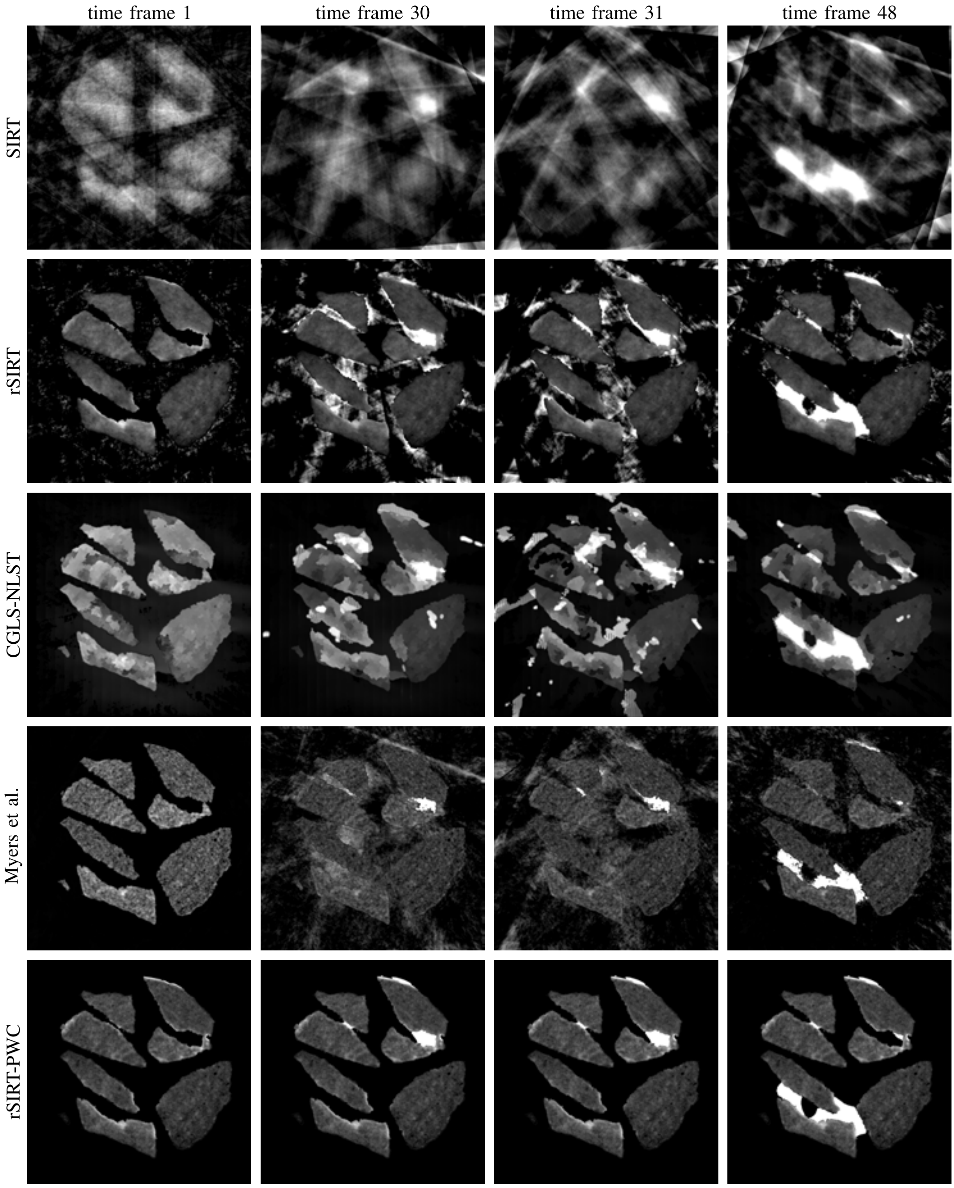


Figure 4.10: Single slice through the reconstructed neutron tomography data showing the evolution through time. The columns correspond to different time frames and the rows to the different reconstruction methods. In this experiment, the projection dataset was subdivided in 48 time frames with 10 projections each.

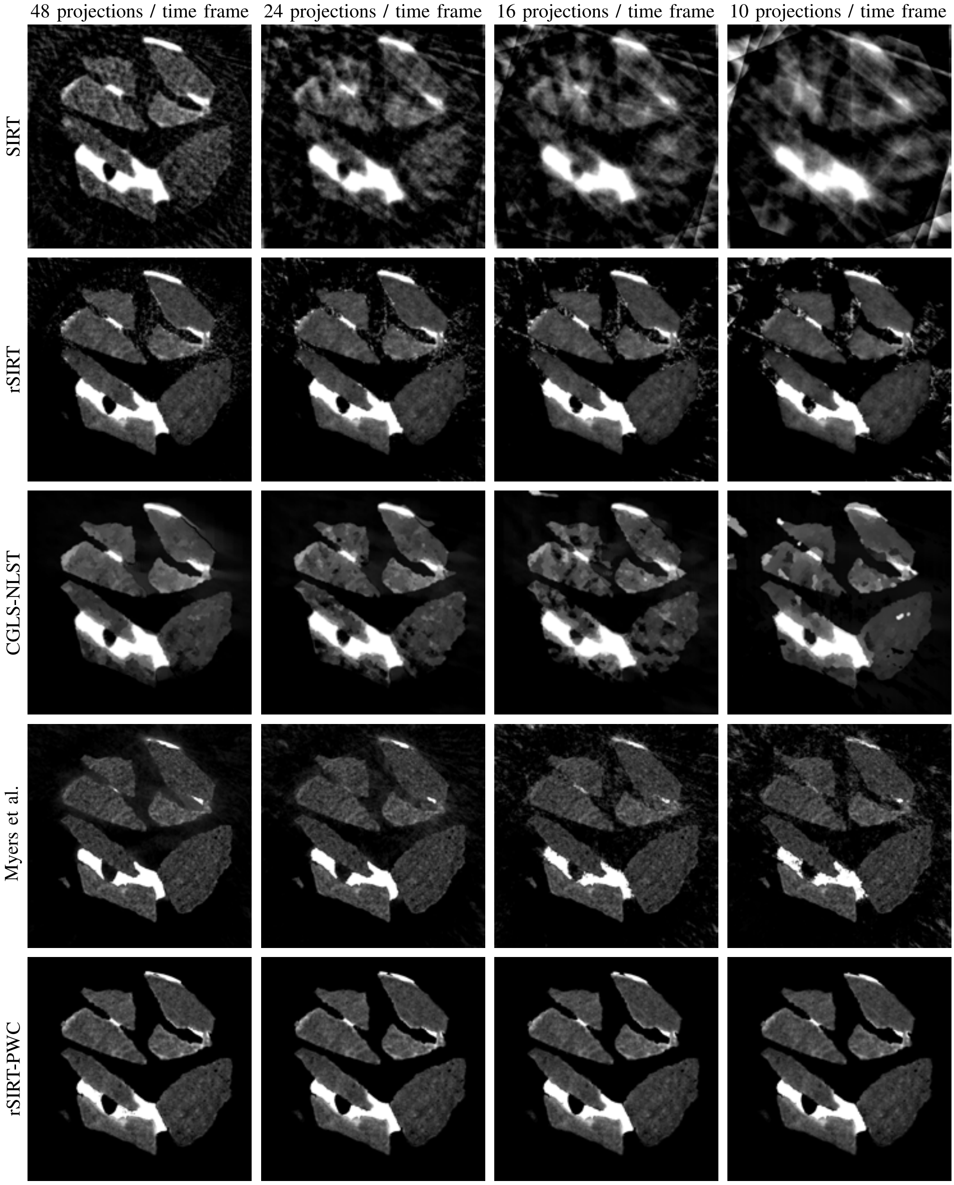


Figure 4.11: Single slice through the reconstructed neutron tomography data on the last time frame. Column labels refer to the number of projections per time frame. Row labels indicate the reconstruction method employed.

5

A local enhancement steered tomography (LEST) algorithm for low-dose cerebral perfusion CT

Contents

At the time of writing, this chapter is submitted for publication:
G. Van Eyndhoven, K.J. Batenburg, N. Buls, J. Vandemeulebroucke, J. De Beenhouwer, and J. Sijbers, *"A local enhancement steered tomography (LEST) algorithm*

for low-dose cerebral perfusion CT", IEEE Transactions on Computational Imaging, submitted, 2015.

5.1 Introduction

Cerebral perfusion computed tomography (PCT) is an imaging technique that allows the visualization and quantification of important hemodynamic information throughout tissue and vessels in the brain [1, 2, 3, 4]. After an intravenous contrast bolus injection, the brain is scanned several times and from the reconstructed time series of brain volumes, a time concentration curve (TCC) describing the evolution over time of local contrast agent concentration can be extracted at the voxel-level. Based on the TCCs, perfusion parameter maps such as cerebral blood volume (CBV), cerebral blood flow (CBF), mean transit time (MTT) and time-to-peak (TTP) can be derived [5, 6, 7].

Cerebral PCT is of particular importance in the case of an acute stroke event, a serious medical emergency that was the fourth leading cause of death and accounted for 5.2% of all deaths in the United States in 2010 [8]. During an ischemic stroke event, the blood supply to the brain is locally disturbed, resulting in rapid loss of brain function [2]. Based on the CBF and the CBV perfusion maps, the extent of a region of severely ischemic but potentially salvageable brain tissue can be identified, providing the clinician with important therapeutic information [9, 6, 7]. However, it suffers from major drawbacks. Radiation dose is by far the most important issue, since the same volume needs to be scanned several times [2, 10]. This has stimulated research on dose reduction in PCT.

Several approaches for reducing the radiation dose and/or increasing image quality have already been suggested. In a straightforward approach, the radiation dose can be decreased simply by reducing the number of projections per time frame. This will, however, give rise to limited data artifacts that may result in erroneous diagnosis. Apart from the optimization of acquisition parameters and hardware [11, 12], a large opportunity can be found in the reconstruction algorithm, which allows for reducing the number of projections per time frame without compromising image quality.

A first class of reconstruction methods assumes that a high quality anatomical image is available and enforces similarity with this image during reconstruction [13, 14, 15, 16, 17]. However, a prior high resolution scan introduces extra radiation dose. Also, the final image quality depends strongly on the prior image quality. Furthermore, a perfect registration between the prior image and the time series reconstruction is indispensable.

In a second class of methods, perfusion specific model assumptions are enforced globally over the entire reconstructed volume. This can be achieved by assuming a predefined shape for the TCCs, e.g., a linear combination of temporal basis functions or a gamma variate function, and estimating the model specific parameters during reconstruction [18, 19, 20]. These approaches enforce the model assumptions globally, while in practice distinct regions throughout the reconstructed volume can have locally different properties.

In this work, we introduce a novel iterative reconstruction algorithm for PCT. Our proposed approach, the Local Enhancement Steered Tomography (LEST) algorithm, does not depend on a high quality prior reconstruction and enforces specific model assumptions locally. Temporal relation between the reconstructions at different time frames is exploited in two different ways. Firstly, the region outside the brain (i.e., bone and void space) is forced to be stationary over time. Secondly, all brain tissue is assumed to be temporally connected through its TCC. Furthermore, the arterial TCCs are modeled by a linear combination of time-shifted gamma variate functions, of which the coefficients are optimized at intermediate iterations. As will be demonstrated, the proposed LEST algorithm significantly improves the quality of the perfusion maps for equal dose or allows for dose reduction while maintaining image quality.

This chapter is organized as follows. In Section 5.2, the model assumptions for cerebral PCT are discussed. In Section 5.3, the LEST method is introduced. Simulation and experimental results are reported in Section 5.4. The results are discussed in Section 5.5 and the chapter is summarized and concluded in Section 5.6.

5.2 Object model

In cerebral PCT the time-varying object $\tilde{\boldsymbol{x}}$ is highly correlated over time, which can be accurately described by the following two model assumptions.

Firstly, some regions (e.g., bone and void space) are known to remain stationary throughout the entire acquisition process. Hence, we assume that a set $S \subset \{1, \ldots, N\}$ of stationary pixel indices can be defined such that $\boldsymbol{x}_r(j) = \boldsymbol{x}_s(j)$ for all $j \in S$ and $r, s \in \{1, \ldots, R\}$. Then, pixels in the dynamic region are defined by $V = \{1, \ldots, N\} \backslash S$. The sets S and V can be determined a priori either by manual delineation or automatically.

Secondly, each time attenuation curve can typically be described by a linear combination of K functions: the constant function $y_K = 1$ and $K-1$ time-shifted gamma variate functions $y_1, y_2, \ldots, y_{K-1}$ [21, 19]. The latter are defined by their simplified form

$$y_k(t) = (t - t_k)^{\kappa} \exp\left(-\frac{t - t_k}{\beta}\right) \qquad k = 1, \ldots, K-1 \tag{5.1}$$

where κ and β are shape parameters and t_k is a time shift. The shape parameters and the time shifts are chosen similar to [19], such as to place $y_1, y_2, \ldots, y_{K-1}$ uniformly over the acquisition time interval. Hence we have $\boldsymbol{x}_r(j) = \sum_{k=1}^{K} a_{k,j} y_k(t_r)$

with $j \in V$, $a_{k,j} \in \mathbb{R}$ the coefficient corresponding to the k^{th} basis function for the j^{th} pixel and t_r the time point corresponding to the r^{th} time index. An example of an attenuation curve formed by the linear combination of $K = 12$ basis functions is shown in Fig. 5.1.

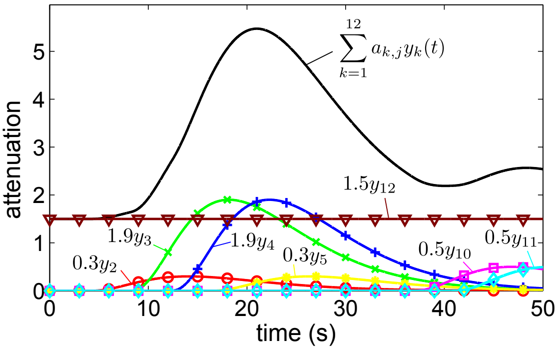


Figure 5.1: An example of a time attenuation curve $\sum_{k=1}^{12} a_{k,j} y_k(t) = 0.3y_2 + 1.9y_3 + 0.3y_5 + 0.5y_{10} + 0.5y_{11} + 1.5y_{12}$. All gamma variate basis functions y_k are scaled to have a maximum of 1 and are defined by $\kappa = 2.5$ and $\beta = 4$. Basis functions with a zero coefficient are not shown.

5.3 Methods

The proposed reconstruction processing chain is visualized in Fig. 5.2. The LEST algorithm (Fig. 5.2, left) is described in Section 5.3.1 and the intermediate gamma variate basis function optimization (Fig. 5.2, right) is explained in Section 5.3.2.

5.3.1 Local Enhancement Steered Tomography (LEST)

In this section, the Local Enhancement Steered Tomography (LEST) algorithm is introduced, an iterative reconstruction method for PCT in which the time evolution of the attenuation values is accurately modeled and local correlation of the PCT dynamics is exploited.

The LEST algorithm starts at iteration $k = 0$ from an initial time-varying reconstruction $\tilde{\boldsymbol{x}}^{(0)}$ generated by applying the standard SIRT algorithm separately to each projection data subset $\boldsymbol{p}_r$ $(r \in \{1, \ldots, R\})$. Since all attenuation values of pixels inside the stationary region S are assumed to remain constant over time, all time attenuation curves of pixels inside S are pixel-wise replaced by the constant function defined by their temporal average. Next, for every time frame $r \in \{1, \ldots, R\}$, the LEST algorithm iteratively performs three updates:

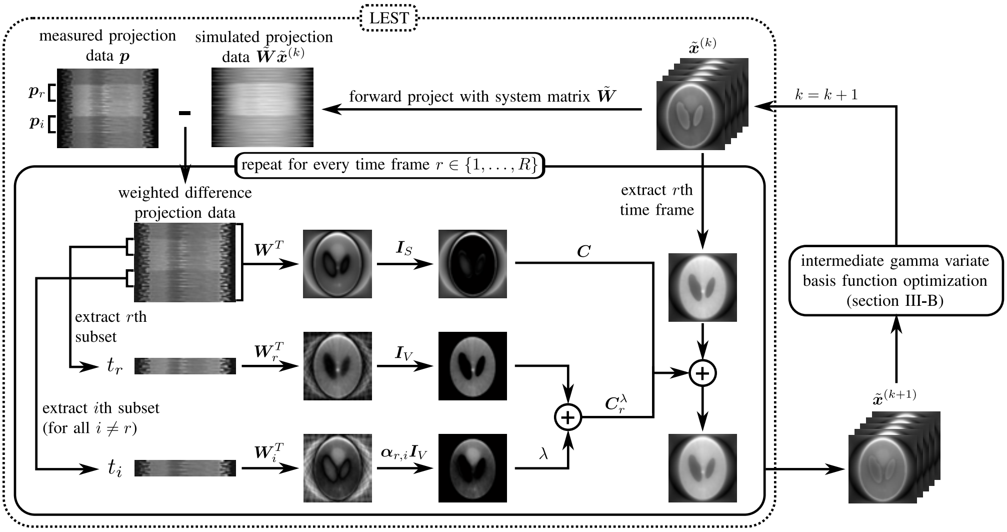


Figure 5.2: Flowchart of the LEST algorithm.

1. *Stationary region update*: This update (top row in the frame of the flowchart in Fig. 5.2) is the same for all time frames and enforces the first model assumption of Section 5.2 on the reconstruction, i.e., all pixels of the stationary region S are equal at every time frame.

2. *Local dynamic region update*: The time-varying region V is updated (middle row in the frame of the flowchart in Fig. 5.2) based solely on the projection data $\boldsymbol{p}_r$ that corresponds uniquely to the r^{th} time frame.

3. *Global dynamic region update*: The time-varying region V is updated again (bottom row in the frame of the flowchart in Fig. 5.2) based on the projection data corresponding to all time frames that differ from the r^{th} time frame.

In the remaining part of this section, these different updates are explained more in depth. All updates are based on the weighted projection difference $\boldsymbol{R}(\boldsymbol{p}-\tilde{\boldsymbol{W}}\tilde{\boldsymbol{x}}^{(k)})$, which is illustrated in the top left part of the flowchart in Fig. 5.2. Since it indicates the discrepancy between the current reconstruction and the measured projection data, this projection difference can be utilized to steer the current reconstruction towards a reconstruction that is more consistent with the measured projection data.

The *stationary region update* is based on the entire projection data set $\boldsymbol{p}$ and is given by $\boldsymbol{C}\boldsymbol{I}_S\boldsymbol{W}^T\boldsymbol{R}(\boldsymbol{p}-\tilde{\boldsymbol{W}}\tilde{\boldsymbol{x}}^{(k)})$, where $\boldsymbol{I}_S \in \{0,1\}^{N\times N}$ is defined as the diagonal matrix with

$$\boldsymbol{I}_S(j,j) = \begin{cases} 1 & \text{if } j \in S \\ 0 & \text{otherwise} \end{cases} \tag{5.2}$$

which selects all pixels of the stationary region (see Chapter 3). This update is exactly the same for every time frame. Therefore, the first model assumption of Section 5.2 is enforced, i.e., all pixel values in the stationary region S are the same at all time frames.

The *local dynamic region update* is a standard SIRT update, restricted to the time-varying region. It is based on the projection data corresponding to the r^{th} time frame and is defined by $\boldsymbol{I}_V \boldsymbol{W}_r^T \boldsymbol{R}_r (\boldsymbol{p}_r - \boldsymbol{W}_r \boldsymbol{x}_r^{(k)})$, where $\boldsymbol{I}_V$ is defined analogously to $\boldsymbol{I}_S$ in Eq. 5.2 and $\boldsymbol{R}_r \in \mathbb{R}^{M \times M}$ is the diagonal matrix with inverse row sums of $\boldsymbol{W}_r$. This update directs the current reconstruction at the r^{th} time frame towards a reconstruction that is more consistent with its corresponding projection data $\boldsymbol{p}_r$.

The *global dynamic region update* is based on the SIRT updates of the other time frames and consequently also on their corresponding projection data $\boldsymbol{p}_i$ with $i \neq r$. Before adding these updates to the reconstruction of the current time frame, they must be scaled properly in a pixel-wise manner according to the ratio of each attenuation value in the current time frame r with respect to the attenuation value of the other time frame i with $i \neq r$. To this end, the diagonal matrix $\boldsymbol{\alpha}_{r,i} \in \mathbb{R}^{N \times N}$ is introduced. The diagonal entries of $\boldsymbol{\alpha}_{r,i}$ represent the ratio of the attenuation values from reconstruction $\boldsymbol{x}_r$ with respect to reconstruction $\boldsymbol{x}_i$. Its diagonal elements are defined as

$$\boldsymbol{\alpha}_{r,i}(k,k) = \frac{\sum_{j \in \mathcal{N}(k)} \bar{z}_{k,j} \boldsymbol{x}_r(j)}{\sum_{j \in \mathcal{N}(k)} \bar{z}_{k,j} \boldsymbol{x}_i(j)} \ , \tag{5.3}$$

where $\mathcal{N}(k)$ represents the 3×3 pixel neighborhood of the pixel with index k. The weights $\bar{z}_{k,j}$ are calculated for every neighborhood $\mathcal{N}(k)$ and reflect the amount of similarity with the central pixel. They are calculated as follows:

$$z_{k,j} = \begin{cases} 1/\sum_{r=1}^{R} |\boldsymbol{x}_r(k) - \boldsymbol{x}_r(j)| & \text{if } k \neq j \\ \max\limits_{l \in \mathcal{N}(k), l \neq k} \left(1/\sum_{r=1}^{R} |\boldsymbol{x}_r(k) - \boldsymbol{x}_r(l)|\right) & \text{if } k = j \end{cases} \tag{5.4}$$

$$\bar{z}_{k,j} = \frac{z_{k,j}}{\sum_{j \in \mathcal{N}(k)} z_{k,j}} \ . \tag{5.5}$$

By calculating the weights in this manner, dissimilar tissues (e.g., a tissue pixel near a bone-tissue interface) will have little to no influence on the diagonal matrix $\boldsymbol{\alpha}_{r,i}$. In summary, the global dynamic region update can be described as

$\lambda \sum_{i\in\{1,\dots,R\}\setminus\{r\}} \boldsymbol{\alpha}_{r,i}\boldsymbol{I}_V\boldsymbol{W}_i^T\boldsymbol{R}_i(\boldsymbol{p}_i - \boldsymbol{W}_i\boldsymbol{x}_i^{(k)})$), where λ $(0 < \lambda < 1)$ represents the relative importance of the projection data corresponding to the current time frame with respect to the projection data corresponding to the other time frames. The multiplication with the diagonal matrix $\boldsymbol{\alpha}_{r,i}$ ensures that the update generated from the projection date $\boldsymbol{p}_i$ $(i \neq r)$ is properly scaled, before adding it to the r^{th} time frame, thereby accounting for the attenuation changes in different types of tissue over time.

More intuitively, this update term can be regarded as a scaling operation on the attenuation curve at each pixel. After adding the scaled version of the update at a particular pixel in a particular time frame to the same pixel at all other time frames, the attenuation curve is scaled in the direction of this update. This is illustrated in Fig. 5.3. Since the r^{th} time frame at the particular pixel receives a scaled update from all other time frames, the resulting attenuation value of this pixel at the r^{th} time frame can be considered intuitively as the sum of all these scaled SIRT updates.

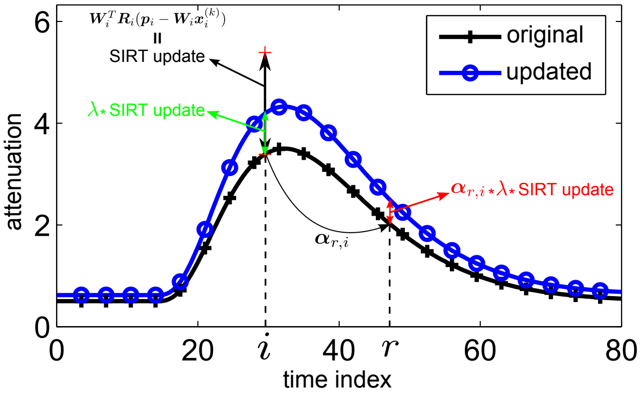


Figure 5.3: Illustration of the effect of the last update term in Eq. 5.7 on the attenuation curve. Imagine that the black curve is the attenuation curve at a particular pixel after k iterations. After adding the properly scaled SIRT update of the i^{th} time frame to all other time frames, the curve changes to the updated blue curve.

Finally, to account for the number or rays intersecting each pixel, the local and global dynamic region update must also be scaled with a proper inverse column sum matrix. Define $\boldsymbol{C}_r^{\lambda} \in \mathbb{R}^{N\times N}$ as the diagonal matrix with inverse column sums of

$$\boldsymbol{W}_r + \lambda \sum_{i\in\{1,\dots,R\}\setminus\{r\}} \boldsymbol{W}_i \ .$$

After a left multiplication of the sum of the local and global dynamic region update by $\boldsymbol{C}_r^{\lambda}$, the resulting update is added to the reconstruction of the r^{th} time frame.

Remember that the classical SIRT algorithm weights each pixel by the inverse of the sum of all ray-intersection lengths through that pixel, i.e., the inverse column sum of the projection matrix. The weighting with $\boldsymbol{C}_r^{\lambda}$ is the analog for the LEST algorithm: since the update for each pixel is based on a weighted sum of updates from all projection data, the weight for that pixel should reflect the ray-intersection lengths of all rays, weighted with their contribution λ.

The introduced notation allows for a compact description of the LEST algorithm. Starting from an initial SIRT reconstruction $\tilde{\boldsymbol{x}}^{(0)}$, the LEST algorithm continues with the following iterative process:

$$\boldsymbol{x}_r^{(k+1)} = \boldsymbol{x}_r^{(k)} + \boldsymbol{C}\boldsymbol{I}_S\boldsymbol{W}^T\boldsymbol{R}(\boldsymbol{p} - \tilde{\boldsymbol{W}}\tilde{\boldsymbol{x}}^{(k)}) \tag{5.6}$$

$$+\boldsymbol{C}_r^{\lambda}\left(\begin{array}{l} \boldsymbol{I}_V\boldsymbol{W}_r^T\boldsymbol{R}_r(\boldsymbol{p}_r - \boldsymbol{W}_r\boldsymbol{x}_r^{(k)}) \quad + \\ \lambda \displaystyle\sum_{i\in\{1,\dots,R\}\setminus\{r\}} \boldsymbol{\alpha}_{r,i}\boldsymbol{I}_V\boldsymbol{W}_i^T\boldsymbol{R}_i(\boldsymbol{p}_i - \boldsymbol{W}_i\boldsymbol{x}_i^{(k)}) \end{array}\right) \tag{5.7}$$

5.3.2 Intermediate gamma variate basis function optimization

For faster convergence and a more accurate estimate of the AIF, a least squares optimization of the arterial attenuation curves is alternated with LEST iterations. This optimization, which is described in this section, enforces the second model assumption of Section 5.2 in the artery regions of the reconstruction domain.

It is assumed that a mask for the artery region, containing L arterial pixels, is known a priori. This mask may also contain other vessel pixels, in order to speed up convergence and increase accuracy in those pixels as well. The mask can either by indicated manually or determined automatically, see [18, 16]. In what follows, the artery region is indicated manually.

For each of the L arterial pixels and K basis function, define $\tilde{\boldsymbol{y}}_{l,k}$ (with $(l,k) \in \{1,\dots,L\} \times \{1,\dots,K\}$) as the time-varying image that is zero everywhere, except in the l^{th} arterial pixel, where the function values of k^{th} basis function are assigned. The basis functions were defined in Section 5.2. Furthermore, define $\tilde{\boldsymbol{y}}_0$ to be the time-varying image that is the same as the current reconstruction $\tilde{\boldsymbol{x}}$ in all pixels except in the arterial pixels, where the attenuation curves are replaced by zero. With these definitions, a time-varying parameterized reconstruction $\tilde{\boldsymbol{y}}(\boldsymbol{a})$ can be defined as

$$\tilde{\boldsymbol{y}}(\boldsymbol{a}) = \tilde{\boldsymbol{y}}_0 + \sum_{l=1}^{L}\sum_{k=1}^{K} a_{l,k}\tilde{\boldsymbol{y}}_{l,k} \ , \tag{5.8}$$

where $\boldsymbol{a}$ is a vector containing all coefficients $a_{l,k}$, i.e., for all L arterial pixels and all K gamma variate basis functions. The intermediate optimization problem can now be written as

$$\min_{\boldsymbol{a}} \left(||\boldsymbol{p} - \tilde{\boldsymbol{W}}\tilde{\boldsymbol{y}}(\boldsymbol{a})||_2^2 + \frac{\mu}{2}\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{j\in\mathcal{M}(l)} (a_{l,k} - a_{j,k})^2 \right) , \tag{5.9}$$

where $\mathcal{M}(l) \subset \{1,\ldots,L\}$ is the set of indices corresponding to the 8-connected neighborhood of the l^{th} arterial pixel (considering only those pixels that belong to the artery region). The parameter μ controls the degree of similarity between coefficients corresponding to the same gamma variate basis function and to neighboring pixels.

Let $\boldsymbol{q}_0$ and $\boldsymbol{q}_{l,k}$ be the simulated projections of $\tilde{\boldsymbol{y}}_0$ and $\tilde{\boldsymbol{y}}_{l,k}$ respectively, i.e., $\boldsymbol{q}_0 = \tilde{\boldsymbol{W}}\tilde{\boldsymbol{y}}_0$ and $\boldsymbol{q}_{l,k} = \tilde{\boldsymbol{W}}\tilde{\boldsymbol{y}}_{l,k}$. This allows for rewriting the norm in Eq. 5.9 as

$$||\boldsymbol{p} - \tilde{\boldsymbol{W}}\tilde{\boldsymbol{y}}(\boldsymbol{a})||_2^2 = ||(\boldsymbol{p} - \boldsymbol{q}_0) - \sum_{l=1}^{L}\sum_{k=1}^{K} a_{l,k}\boldsymbol{q}_{l,k}||_2^2 . \tag{5.10}$$

The minimization problem in Eq. 5.9 can be solved by setting the partial derivatives (with respect to $a_{l,k}$) to zero, resulting in

$$\sum_{j\in\mathcal{M}(l)} (\mu a_{l,k} - \mu a_{j,k}) + \sum_{l'=1}^{L}\sum_{k'=1}^{K} (\boldsymbol{q}_{l,k}^T\boldsymbol{q}_{l',k'})a_{l',k'} = \boldsymbol{q}_{l,k}^T(\boldsymbol{p} - \boldsymbol{q}_0) . \tag{5.11}$$

Since Eq. 5.11 holds for every $(l,k) \in \{1,\ldots,L\} \times \{1,\ldots,K\}$ and Eq. 5.11 is a linear equation with respect to the coefficients in $\boldsymbol{a}$, this gives rise to the following matrix equation:

$$\boldsymbol{M}\boldsymbol{a} = \boldsymbol{b} , \tag{5.12}$$

where $\boldsymbol{M} \in \mathbb{R}^{LK\times LK}$ and $\boldsymbol{b} \in \mathbb{R}^{LK}$ are defined via the left-hand side and the right-hand side of Eq. 5.11, respectively.

Note that $\boldsymbol{M}$ and all $\boldsymbol{q}_{l,k}$ in Eq. 5.12 can be calculated before the LEST algorithm starts. This way, only the right hand side $\boldsymbol{b}$ of Eq. 5.12 needs to be updated during the intermediate optimization. To speed up the computations, Eq. 5.12 is solved for each connected component of the artery region containing a maximum of 15 pixels. The equations are solved by LU factorization with partial pivoting.

5.3.3 Automated dynamic region estimation method

The LEST algorithms needs prior knowledge about the location of stationary and dynamic pixels. For this purpose, a simple and automated dynamic region estimation method was implemented. To make the method more stable if only few projections are available per time frame, the projection data for each three subsequent time frames are joined and SIRT reconstructions are calculated based on these joined projection data subsets. Next, the difference between the maximum and the minimum attenuation value is calculated for each pixel (i.e., for each time attenuation curve). Since this difference quantifies the maximum amount of contrast fluid present in each pixel, it is thresholded to produce a first rough estimate for the dynamic mask. Finally, morphological image operations are applied to the resulting mask: after opening the mask with a two-pixel radius disk, the largest connected component is selected and all holes are filled. The resulting mask is the estimate for the dynamic region.

5.4 Experiments and results

In this section, the LEST algorithm is validated with both numerical and real data experiments. First, different figures of merit are introduced in Section 5.4.1. Next, in Section 5.4.2, simulation phantoms are introduced and LEST is compared to state-of-the-art reconstruction methods. Finally, the LEST algorithm is validated on clinical perfusion CT data in Section 5.4.3.

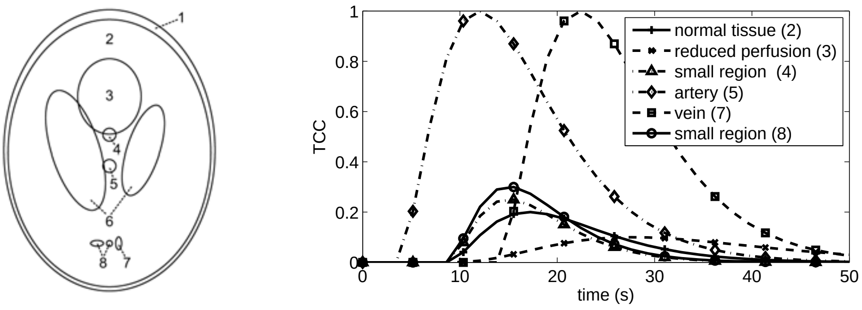
5.4.1 Figures of merit

A comparison between the different reconstruction algorithms is performed based on both the reconstructed attenuation values and the derived CBF (ml/100 ml/min) and CBV (ml/100 ml) perfusion maps. The CBF and CBV maps are calculated with the deconvolution-based truncated singular value decomposition method with a fixed threshold value of 20% of the largest singular value [5]. The estimate for the AIF was calculated as the average TCC over all pixels in the arterial mask.

The Relative Root Mean Squared Error (RRMSE) is utilized as a quality measure and is defined as

$$RRMSE = \sqrt{\frac{\sum_i (\hat{\boldsymbol{x}}(i) - \tilde{\boldsymbol{x}}(i))^2}{\sum_i (\tilde{\boldsymbol{x}}(i))^2}} \ , \tag{5.13}$$

where $\hat{\boldsymbol{x}}$ denotes the calculated reconstruction (in which case the summation index i goes over all possible pixels at all possible points in time) or perfusion map (in which case the summation index i goes over all possible pixels in the perfusion map) and $\tilde{\boldsymbol{x}}$ denotes the ground truth phantom or perfusion map. For some experiments, the RRMSE is inspected only in a Region of Interest (ROI), in which case the sum in Eq. 5.13 sums over all points in time and over all pixels in the specific ROI.



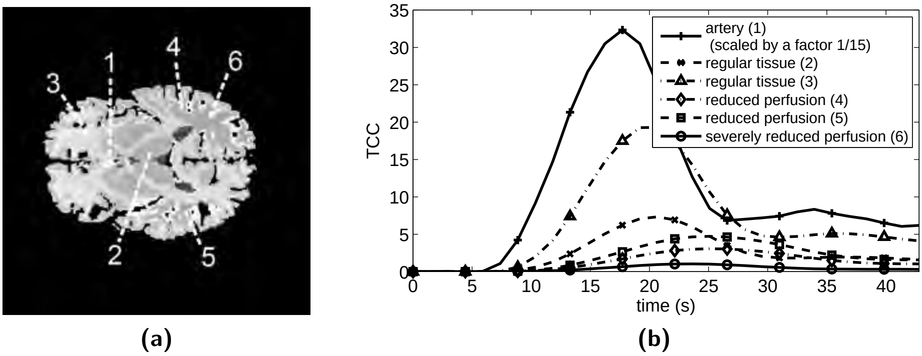
(a) (b)

Figure 5.4: The first simulation phantom is created by superimposing various concentration curves (b) on top of the standard Shepp-Logan phantom (a). The bone region (1) and the void space region (6) are unaltered. All other regions (2,3,4,5,7,8) are altered by adding the TCCs of the plot in (b) to the attenuation value of the standard Shepp-Logan phantom. These TCCs were chosen to reflect various sorts of tissue, including normal tissue (2), a region with reduced perfusion (3), smaller regions (4 and 8) and an artery (5) and vein (7) region.

5.4.2 Numerical simulations

In this section, the LEST algorithm is validated with a series of numerical experiments. The first simulation phantom, displayed in Fig. 5.4, is a standard Shepp-Logan phantom on which various TCCs are superimposed, each corresponding to a different type of tissue. All TCCs are created with Eq. 5.1, where different values were chosen for the scaling, time shift and shape parameters. The second simulation phantom is a realistic digital brain phantom that was developed by Riordan et al. and extended by Manhart et al. [18, 22, 23]. It is visualized in Fig. 5.5. In contrast to classical digital CT phantoms, this phantom has reduced sparsity in the image domain, thereby not favoring reconstruction algorithms exploiting homogeneity (e.g., minimizing total variation). Both simulation phantoms were defined on a 256×256 isotropic pixel grid.

For all experiments, the projections were sampled on angles defined by the golden ratio scanning protocol [24]. Projections were simulated with a parallel beam geometry using a strip kernel, while reconstructions were calculated assuming a linear projection kernel. Also, Poisson distributed noise was applied to the projection data assuming an incoming beam intensity of $I_0 = 2 \times 10^4$ (photon count) per detector pixel. The number of simulated detector bins was set to 128 for the Shepp-Logan and 256 for the digital brain phantom.



(a) **(b)**

Figure 5.5: The second simulation phantom is adopted from [18, 23]. Some exemplary TCCs are displayed in (b), corresponding to the indicated pixels in (a). However, all TCCs vary from pixel to pixel. This way, this phantom does not favor sparsity exploiting algorithms. It was created by starting from a CBF and CBV map (with manually indicated reduced and severely reduced tissue regions) and calculating the TCCs based on these maps. By adding random perturbations on the CBF and the extracted MTT (=CBV/CBF) map, the final TCCs have reduced sparsity characteristics.

A comparison of LEST to the conventional reconstruction algorithm SIRT and the recently proposed Prior Image Constrained Compressed Sensing (PICCS) algorithm [16, 17] is performed. The SIRT algorithm was implemented with a positivity constraint and 200 iterations for all subsequent experiments. The PICCS algorithm was implemented as described in [16, 17] and the parameter values given in [16] were adopted. All parameters that were not described in [16], i.e., the number of steepest descent iterations and the size of the steepest descent step, were optimized by selecting those parameters that yielded the lowest RRMSE with respect to the ground truth image. The prior image was calculated with the vessel-selective prior (see [16]) with the same vessel mask that was used for the intermediate gamma variate basis function optimization.

5.4.2.1 Shepp-Logan phantom

Reconstructions were calculated on a 128×128 isotropic pixel grid on 30 time frames, ranging from 0 to 50 s with a 1.724 s time step.

The LEST algorithm was applied in two different ways. In a first implementation, referred to as "*LEST-gt*", the ground truth stationary region S, defined by the void space region and the bone region (see Fig. 5.4a), was used. In the implementation referred to as "*LEST-est*", the stationary region was estimated

	SIRT	*PICCS*	*LEST-gt*	*LEST-est*
att. values	0.245	0.201	**0.200**	**0.200**
CBF	0.507	0.320	**0.282**	0.291
CBV	0.551	0.358	**0.236**	0.255

Table 5.1: RRMSE in the dynamic region of the Shepp-Logan phantom for perfusion maps and attenuation values (rows) of different reconstruction algorithms (columns). In this experiment, 20 projections were simulated per time frame.

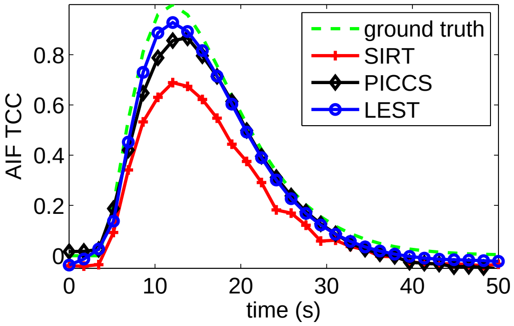
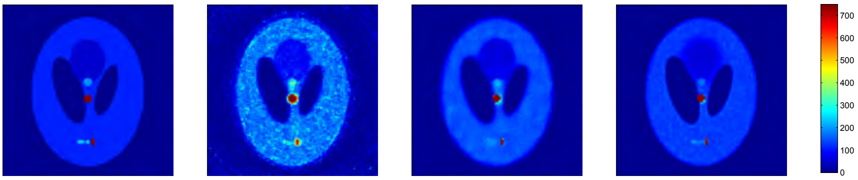


Figure 5.6: The estimated AIF for the Shepp-Logan phantom.

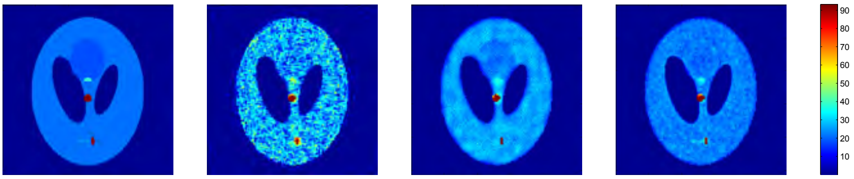
automatically with the method described in Section 5.3.3. Starting from an initial reconstruction generated with one SIRT iteration, the LEST reconstruction algorithm performed 300 iterations. LEST was set up with $\lambda = 0.4$, which was selected by visual assessment of the reconstructed images. The intermediate gamma variate basis function optimization (see Eq. 5.9) was performed in the artery and vein region at every 20th iteration. The gamma variate functions were defined as in Eq. 5.1 with $\kappa = 1.5$ and $\beta = 4$, similar to [19]. A total of 12 gamma variate basis functions were used (including the constant function). The regularization parameter in Eq. 5.9 was set to $\mu = 10$, which gives a good trade-off between data fidelity and similarity of the coefficients of the gamma variate functions in neighboring pixels.

In a first experiment, reconstructions were calculated based on 20 simulated projections per time frame. The numerical results for this experiment are summarized in Table 5.1. The estimated AIF and perfusion maps are displayed in Fig. 5.6 and Fig. 5.7, respectively. All AIFs were evaluated using the same ROI.

The simulation experiment with the Shepp-Logan phantom was repeated several times, each time for a different number of projections per time frame (while keeping the beam intensity I_0 constant). Reconstructions were generated with the



(a) CBF ground truth **(b)** CBF SIRT **(c)** CBF PICCS **(d)** CBF LEST-gt



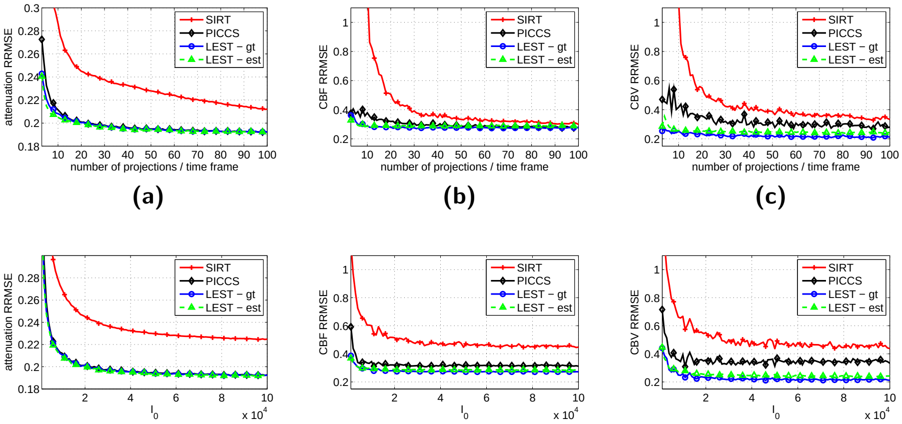
(f) CBV ground truth **(g)** CBV SIRT **(h)** CBV PICCS **(i)** CBV LEST-gt

Figure 5.7: Ground truth and calculated perfusion maps (CBF and CBV) of the Shepp-Logan phantom for the SIRT, PICCS and LEST-gt reconstructions.

SIRT, PICCS, and LEST method and the corresponding RRMSE was plotted as a function of the number of projections per time frame (see Fig. 5.8(a-c)). The same analysis is repeated for different noise levels, by varying the incoming beam intensity I_0, while maintaining 20 projection per time frame. The result is shown in Fig. 5.8(d-f).

The Shepp-Logan phantom was also utilized to check the relevancy of the intermediate optimization in the artery region (see Section 5.3.2). In Fig. 5.9, the RRMSE as a function of the number of projections per time frame is shown for the LEST method with and without the intermediate optimization. The LEST algorithm was applied with the ground truth dynamic region as prior knowledge. Furthermore, to check the effect of the LEST approach and the intermediate artery optimization separately, the optimization in the artery region was also applied at intermediate iterations in the SIRT algorithm, the resulting RRMSE values are also plotted in Fig. 5.9. The RRMSE is evaluated in both the artery region (Fig. 5.9(a)) and the dynamic region without the artery pixels (Fig. 5.9(b-c)).

Finally, to illustrate the added value of the stationary region update in LEST, the algorithm was applied with (i.e., $S \neq \emptyset$) and without (i.e., $S = \emptyset$) the stationary region update. The resulting RRMSE for the CBV map is plotted as a function of the number of projections per time frame in Fig. 5.10, the error values for the CBF map and attenuation values have a similar trend.



(a) (b) (c)

(d) (e) (f)

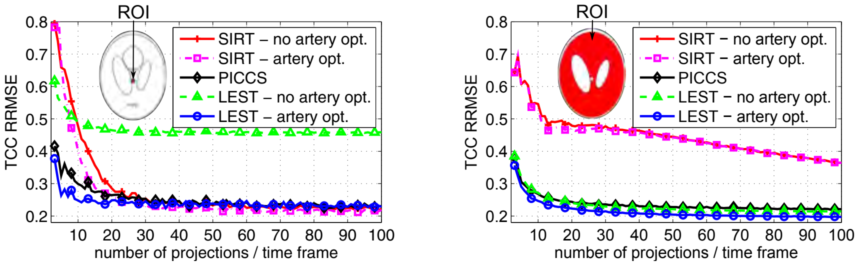
Figure 5.8: RRMSE values as a function of the number of projections per time frame (a-c) and the incoming beam intensity I_0 (d-f) for the Shepp-Logan phantom. For a given number of projections per time frame or incoming beam intensity I_0, the RRMSE in the dynamic ROI (i.e., the complement of region 6 and 1 in Fig. 5.4a) is plotted for attenuation values (a,d), CBF map (b,e) and CBV map (c,f) for the SIRT, PICCS and LEST reconstruction.

5.4.2.2 Brain phantom

Reconstructions were calculated on the same 256×256 pixel grid as the ground truth image in order to guarantee a non-sparse phantom and reconstruction. The phantom consists of 30 individual time frames, equally distributed between 0 and 43 s with a 1.476 s time increment.

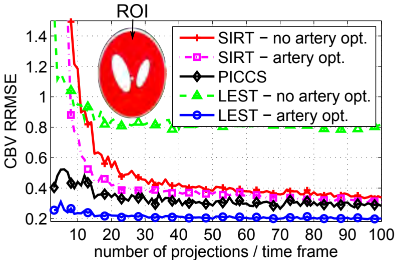
The reconstruction parameters for LEST were chosen identically to the Shepp-Logan simulation experiment. Again, "*LEST-est*" refers to the LEST implementation with automated dynamic region estimation (see Section 5.3.3) and "*LEST-gt*" refers to the LEST implementation that assumes the stationary region S, defined by the void space in Fig. 5.5, is known.

In a first experiment, 50 projections were simulated per time frame. The numerical results for this experiment are summarized in Table 5.2. The AIF estimate can be inspected in Fig. 5.11 and the perfusion maps can be compared in Fig. 5.12. The performance of all reconstruction methods for a different number of projections per time frame and a varying amount of noise can be inspected in Fig. 5.13 and Fig. 5.14. In Fig. 5.13 the RRMSE is evaluated in the full dynamic region, whereas in Fig. 5.14 the RRMSE is only evaluated in the region of reduced perfusion, thereby quantifying the ability of the different methods to identify this



(a) RRMSE of the extracted TCCs calculated in the artery region

(b) RRMSE of the extracted TCCs calculated in the dynamic region excluding the artery



(c) RRMSE of the CBV perfusion map calculated in the dynamic region excluding the artery

Figure 5.9: RRMSE values as a function of the number of projections per time frame for the Shepp-Logan phantom, evaluated in different ROIs: (a) in the artery region and (b-c) in the dynamic region excluding the arterial pixels. Both LEST and SIRT are implemented with (“*LEST/SIRT - artery opt.*”) and without (“*LEST/SIRT - no artery opt.*”) the intermediate optimization in the artery region, thereby evaluating the relevancy of this intermediate optimization.

	SIRT	*PICCS*	*LEST-gt*	*LEST-est*
att. values	0.359	0.191	**0.108**	0.110
CBF	0.656	0.208	0.092	**0.090**
CBV	0.768	0.200	0.122	**0.118**

Table 5.2: RRMSE values in the dynamic region of the brain phantom for perfusion maps and attenuation values (rows) of different reconstruction algorithms (columns). In this experiment, 50 projections were simulated per time frame.

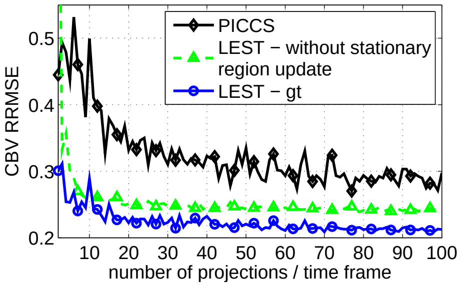


Figure 5.10: CBV RRMSE values evaluated in the dynamic region as a function of the number of projections per time frame for the Shepp-Logan phantom. The LEST algorithm was applied with and without the stationary region update.

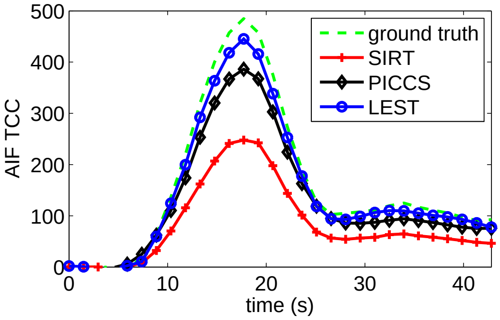


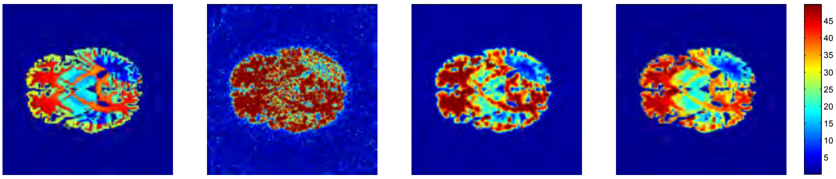
Figure 5.11: The estimated AIF for the brain phantom.

region.

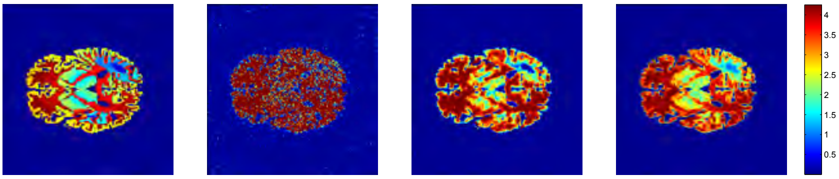
5.4.3 Clinical perfusion data

To test its performance on clinical data, the LEST algorithm was also applied on a cerebral perfusion CT dataset that was acquired with a Discovery CT750 HD (GE Healthcare) scanner. The scan was performed with a source voltage of 80 kVp and X-ray tube current of 500 mA. Following a bolus injection of 50 ml, with injection rate of 4 ml/s, 912 projections were acquired per 180° rotation of source and detector. Based on this large amount of projection data, high quality images were reconstructed on a 512×512 pixel grid at 24 time frames with a 2.8094 s inter-frame temporal distance. These high quality reconstructions are utilized as ground truth images in the subsequent experiment.

Based on the ground truth images, 250 parallel beam projections per time frame



(a) CBF ground truth **(b)** CBF SIRT **(c)** CBF PICCS **(d)** CBF LEST-gt



(f) CBV ground truth **(g)** CBV SIRT **(h)** CBV PICCS **(i)** CBV LEST-gt

Figure 5.12: Perfusion maps (CBF and CBV) calculated based on different reconstructions of the brain phantom.

were simulated. The projections were simulated with a strip kernel and poisson noise was applied, assuming a beam intensity of $I_0 = 8 \times 10^5$ (photon count) at the source. Reconstructions were calculated on a 256×256 pixel grid assuming a linear projection kernel. The artery region was indicated manually and the vessel mask was selected by a global thresholding operation on the peak attenuation value in each pixel. These masks are indicated in Fig. 5.15. The LEST reconstruction was calculated with the following parameters: 100 initial SIRT iterations, 500 LEST iterations, optimization inside the vessel mask every 50th iteration, $\mu = 1000$, $\lambda = 0.4$, gamma variate functions defined again with $\kappa = 1.5$ and $\beta = 4$.

The peak attenuation value for each pixel in the reconstructed images is displayed in Fig. 5.16. The perfusion maps can be inspected in Fig. 5.17. The numerical results are summarized in Table 5.3 and the estimated AIF is displayed in Fig. 5.18.

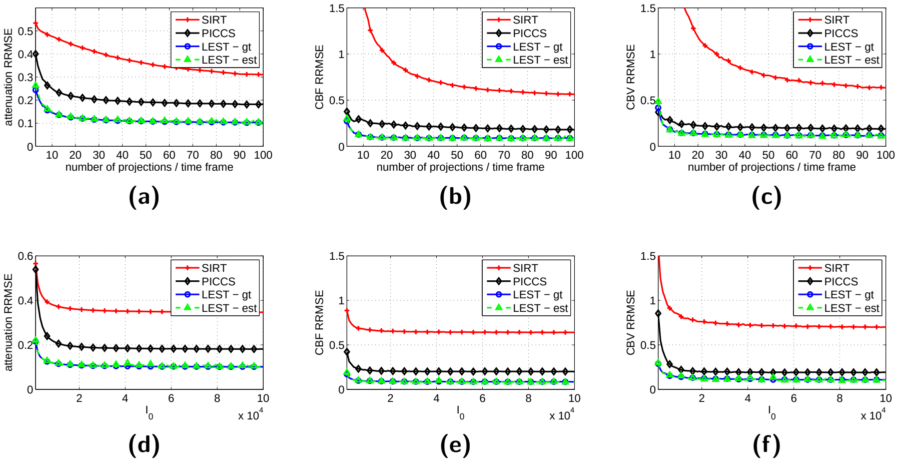


Figure 5.13: RRMSE values as a function of the number of projections per time frame (a-c) and the incoming beam intensity I_0 (d-f) for the brain phantom. For a given number of projections per time frame or incoming beam intensity I_0, the RRMSE in the dynamic ROI is plotted for attenuation values (a,d), CBF map (b,e) and CBV map (c,f) for the SIRT, PICCS and LEST reconstruction.

	SIRT	*PICCS*	*LEST*
att. values	0.013	0.017	**0.012**
CBF	0.383	0.439	**0.363**
CBV	0.632	0.732	**0.495**

Table 5.3: RRMSE in the dynamic region for the clinical perfusion data experiment for perfusion maps and attenuation values (rows) of different reconstruction algorithms (columns).

5.5 Discussion

5.5.1 Shepp-Logan phantom

From Table 5.1, it is clear that the LEST algorithm outperforms all other reconstruction methods with respect to image quality for the RRMSE metric. Since the Shepp-Logan phantom has a low total variation and the PICCS algorithm promotes such image characteristics, an adequate reconstruction quality in terms of attenuation values (first row of Table 5.1) can also be observed for the PICCS algorithm in this experiment. However, the improved AIF estimation (displayed

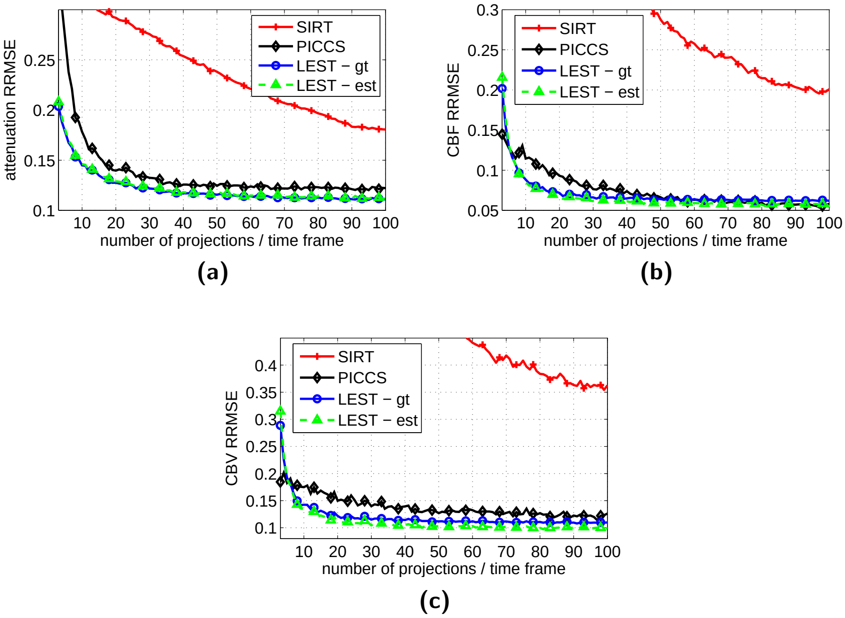


Figure 5.14: RRMSE values evaluated in the region of reduced perfusion (region 4, 5 and 6 in Fig. 5.5a) as a function of the number of projections per time frame for the brain phantom. RRMSE values are calculated for the attenuation curves (a), the CBF map (b) and the CBV map (c).

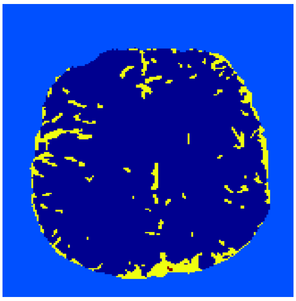
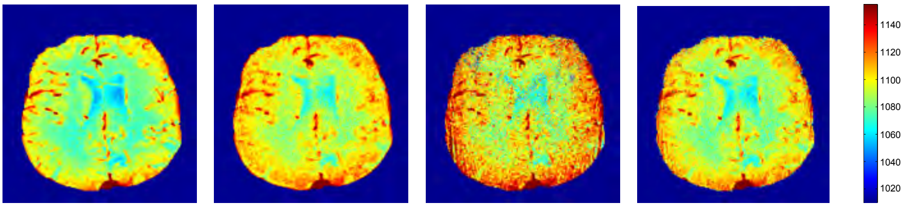
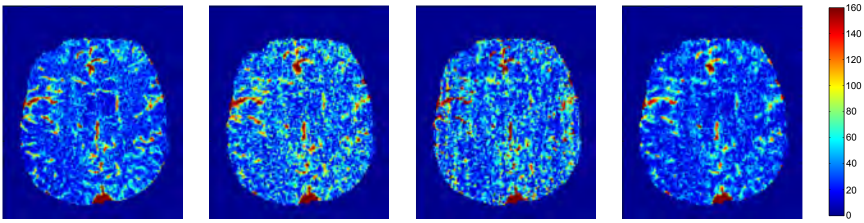


Figure 5.15: The masks that were utilized in the clinical perfusion data experiment. Light blue indicates the stationary region, blue the dynamic region, yellow the vessel mask and red the region of arterial pixels.

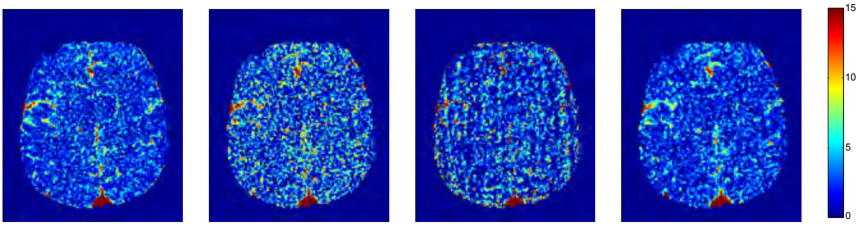


(a) Ground truth **(b)** SIRT **(c)** PICCS **(d)** LEST

Figure 5.16: The peak attenuation value in each pixel for all reconstruction methods and the ground truth data for the clinical perfusion data experiment. Only the pixels inside the dynamic region are visualized. The color range is the same for each of the figures.



(a) CBF ground truth **(b)** CBF SIRT **(c)** CBF PICCS **(d)** CBF LEST



(f) CBV ground truth **(g)** CBV SIRT **(h)** CBV PICCS **(i)** CBV LEST

Figure 5.17: CBF (units: ml/100 ml/min) and CBV (units: ml/100 ml) perfusion maps calculated based on different reconstructions of the clinical perfusion data. Only pixels inside the dynamic region are visualized.

in Fig. 5.6) results in a significant increase in image quality of the CBF and CBV maps by the LEST method. Also notice that the image quality of LEST-gt and LEST-est is highly similar, indicating that manual indication of the dynamic re-

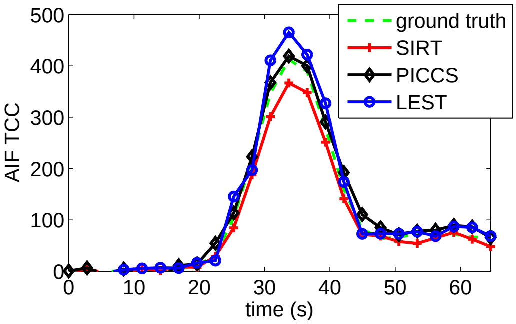


Figure 5.18: The estimated AIF for the clinical perfusion CT data experiment.

gion is unnecessary and that such a task can be easily performed by the automated estimation method proposed in Section 5.3.3.

It is noticeable that the small regions (regions 8 in Fig. 5.4a) are the most recognizable in the perfusion maps based on the LEST reconstruction, whereas they are hardly visible in the perfusion maps based on the PICCS reconstruction. This can be explained by the fact that PICCS minimizes the total variation in the spatial domain, thereby introducing the risk of erasing smaller structures. The same observation goes for the sensitivity analysis in Fig. 5.8: a significant improvement of LEST over PICCS can be noticed for all noise levels and all number of projections per time frame for the resulting CBF and CBV maps.

From the plots of Fig. 5.9(a-b), three major observations can be made. Firstly, comparing the SIRT and LEST algorithm without intermediate optimization in the artery region, one observes that the LEST algorithm improves the TCCs' quality in the dynamic region, but the estimation in the artery region is worse. Secondly, the intermediate artery optimization is a useful addition to improve the quality of the TCCs in the artery region, both for SIRT and LEST (Fig. 5.9(a)). Finally, the LEST implementation with intermediate optimization in the artery region outperforms all other methods. This last observation is emphasized even more after observing the RRMSE values for the CBV perfusion map in Fig. 5.9(c). The large difference in image quality of the CBV map for LEST with and without the intermediate optimization in the artery region (Fig. 5.9(c)) is mainly due to the difference in quality of the AIF, which is based on the TCCs in the artery region (Fig. 5.9(a)).

Finally, it can be observed that the stationary region update has a positive influence on the reconstruction quality in both the stationary and the dynamic region, which is quantified in the plot of Fig. 5.10.

5.5.2 Brain phantom

Since this phantom is no longer spatially sparse (in the sense of having a lower total variation), the PICCS algorithm performs significantly worse in comparison to the LEST algorithm on all possible validation measures (see Table 5.2). Besides the lack of sparsity, also the quality of the estimated AIFs (displayed in Fig. 5.11) have an influence on the final perfusion map quality (see Fig. 5.12). Again, the same observation can be made for different number of projections per time frame or for different noise levels, of which the results are plotted in Fig. 5.13 and Fig. 5.14. The plots in Fig. 5.14 illustrate the ability of LEST to detect regions of reduced perfusion within the brain. Also, notice that the quality of the LEST-gt and the LEST-est reconstruction is again highly similar.

5.5.3 Clinical perfusion data

As can be seen from the peak attenuation values and the perfusion maps (Fig. 5.16c, Fig. 5.17c and Fig. 5.17h), the PICCS reconstruction algorithm results in many artifacts, even though its parameters were optimized to have a minimum RRMSE with respect to the ground truth. The SIRT reconstruction (Fig. 5.16b) contains a high level of noise, which is partially alleviated in the LEST reconstruction (Fig. 5.16d). The perfusion maps derived from the LEST reconstruction have the highest similarity to the ground truth, which is also numerically confirmed by the RRMSE values in Table 5.3. The AIF's estimate, displayed in Fig. 5.18, has different properties depending on the reconstruction method it was derived from. The SIRT reconstruction results in a slight underestimation of the AIF, while the LEST reconstruction shows a slight overestimation. Finally, the PICCS reconstruction accurately captures the peak concentration of the AIF, but is slightly elongated in the temporal direction.

5.6 Conclusion

In general, the radiation dose in PCT experiments can be decreased by undersampling the projection data. Standard approaches that reconstruct the object independently at different time frames, however, result in images containing artifacts that are introduced by the lack of projection data.

In this chapter, the LEST algorithm was presented. It is an iterative reconstruction algorithm tailored specifically to cerebral PCT problems. The algorithm exploits prior knowledge in a twofold manner. On the one hand, the reconstruction quality in the stationary regions (bone and void space) is improved by assuming these regions to have the same attenuation at all time frames. On the other hand, image quality in the dynamic region is further improved by the assumption that the regions with brain tissue and vessels are correlated over time. An intermediate gamma variate basis function optimization approach was introduced, resulting in increased accuracy of the derived AIF, and consequently also in increased quality of the estimated perfusion maps.

The LEST algorithm was validated with simulation experiments and clinical data (based on a high quality reconstruction of a Discovery CT750 HD scan). These experiments illustrate that, in comparison to standard approaches, the LEST algorithm significantly improves image quality for the same radiation dose and has similar image quality for a lower radiation dose.

References

[1] L. Axel, "Cerebral blood flow determination by rapid-sequence computed tomography: theoretical analysis.," *Radiology*, vol. 137, pp. 679–86, 1980.

[2] K. A. Miles and M. R. Griffiths, "Perfusion CT: a worthwhile enhancement?," *The British Journal of Radiology*, vol. 76, no. 904, pp. 220–231, 2003.

[3] M. Wintermark, R. Sincic, D. Sridhar, and J. Chien, "Cerebral perfusion CT: Technique and clinical applications," *Journal of Neuroradiology*, vol. 35, no. 5, pp. 253 – 260, 2008.

[4] E. G. Hoeffner, I. Case, R. Jain, S. K. Gujar, G. V. Shah, J. P. Deveikis, R. C. Carlos, B. G. Thompson, M. R. Harrigan, and S. K. Mukherji, "Cerebral perfusion CT: Technique and clinical applications 1," *Radiology*, vol. 231, no. 3, pp. 632–644, 2004.

[5] A. Fieselmann, M. Kowarschik, A. Ganguly, J. Hornegger, and R. Fahrig, "Deconvolution-based CT and MR brain perfusion measurement: Theoretical model revisited and practical implementation details.," *International Journal of Biomedical Imaging*, vol. 2011, p. 20, 2011.

[6] A. A. Konstas, G. V. Goldmakher, T.-Y. Lee, and M. H. Lev, "Theoretic basis and technical implementations of CT perfusion in acute ischemic stroke, part 1: Theoretic basis.," *American Journal of Neuroradiology*, vol. 30, no. 4, pp. 662–8, 2009.

[7] A. A. Konstas, G. V. Goldmakher, T.-Y. Lee, and M. H. Lev, "Theoretic basis and technical implementations of CT perfusion in acute ischemic stroke, part 2: technical implementations.," *American Journal of Neuroradiology*, vol. 30, no. 5, pp. 885–92, 2009.

[8] S. L. Murphy, J. Xu, K. D. Kochanek, and V. Statistics, "Deaths: Final data for 2010," *National Vital Statistics Reports*, vol. 61, no. 4, 2013.

[9] B. D. Murphy, A. J. Fox, D. H. Lee, D. J. Sahlas, S. E. Black, M. J. Hogan, S. B. Coutts, A. M. Demchuk, M. Goyal, R. I. Aviv, S. Symons, I. B. Gulka, V. Beletsky, D. Pelz, V. Hachinski, R. Chan, and T. Y. Lee, "Identification of penumbra and infarct in acute ischemic stroke using computed tomography perfusion-derived blood flow and blood volume measurements," *Stroke*, vol. 37, no. 7, pp. 1771–1777, 2006.

[10] A. Mnyusiwalla, R. I. Aviv, and S. P. Symons, "Radiation dose from multidetector row CT imaging for acute stroke.," *Neuroradiology*, vol. 51, no. 10, pp. 635–40, 2009.

[11] M. Wintermark, W. S. Smith, N. U. Ko, M. Quist, P. Schnyder, and W. P. Dillon, "Dynamic perfusion CT: optimizing the temporal resolution and contrast volume for calculation of perfusion CT parameters in stroke patients.," *American Journal of Neuroradiology*, vol. 25, no. 5, pp. 720–9, 2004.

[12] M. Wiesmann, S. Berg, G. Bohner, R. Klingebiel, V. Schöpf, B. M. Stoeckelhuber, I. Yousry, J. Linn, and U. Missler, "Dose reduction in dynamic perfusion CT of the brain: effects of the scan frequency on measurements of cerebral blood flow, cerebral blood volume, and mean transit time.," *European Radiology*, vol. 18, no. 12, pp. 2967–74, 2008.

[13] J. Ma, J. Huang, Q. Feng, H. Zhang, H. Lu, Z. Liang, and W. Chen, "Low-dose computed tomography image restoration using previous normal-dose scan.," *Medical Physics*, vol. 38, no. 10, pp. 5713–31, 2011.

[14] J. Ma, H. Zhang, Y. Gao, and J. Huang, "Iterative image reconstruction for cerebral perfusion CT using a pre-contrast scan induced edge-preserving prior," *Physics in Medicine and Biology*, vol. 57, no. 22, pp. 7519–7542, 2012.

[15] J. C. Ramirez-Giraldo, J. Trzasko, S. Leng, L. Yu, a. Manduca, and C. H. McCollough, "Nonconvex prior image constrained compressed sensing (NCPICCS): Theory and simulations on perfusion CT," *Medical Physics*, vol. 38, no. 4, p. 2157, 2011.

[16] B. E. Nett, R. Brauweiler, W. Kalender, H. Rowley, and G.-H. Chen, "Perfusion measurements by micro-CT using prior image constrained compressed sensing (PICCS): initial phantom results.," *Physics in Medicine and Biology*, vol. 55, no. 8, pp. 2333–50, 2010.

[17] B. Nett, J. Tang, S. Leng, and G. Chen, "Tomosynthesis via total variation minimization reconstruction and prior image constrained compressed sensing (PICCS) on a C-arm system," *Proceedings of SPIE*, pp. 1–14, 2008.

[18] M. T. Manhart, M. Kowarschik, A. Fieselmann, Y. Deuerling-Zheng, K. Royalty, A. K. Maier, and J. Hornegger, "Dynamic iterative reconstruction for interventional 4-D C-arm CT perfusion imaging," *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1336–1348, 2013.

[19] C. Neukirchen, M. Giordano, and S. Wiesner, “An iterative method for tomographic x-ray perfusion estimation in a decomposition model-based approach,” *Medical Physics*, vol. 37, no. 12, p. 6125, 2010.

[20] M. Wagner, Y. Deuerling-Zheng, M. Möhlenbruch, M. Bendszus, J. Boese, and S. Heiland, “A model based algorithm for perfusion estimation in interventional C-arm CT systems.,” *Medical Physics*, vol. 40, no. 3, p. 031916, 2013.

[21] M. D. Harpen and M. L. Lecklitner, “Derivation of gamma variate indicator dilution function from simple convective dispersion model of blood flow,” *Medical Physics*, vol. 11, p. 690, 1984.

[22] A. J. Riordan, M. Prokop, M. A. Viergever, J. W. Dankbaar, E. J. Smit, and H. W. A. M. de Jong, “Validation of CT brain perfusion methods using a realistic dynamic head phantom,” *Medical Physics*, vol. 38, no. 6, pp. 3212–21, 2011.

[23] M. Manhart, “Digital brain perfusion phantom,” http://www5.cs.fau.de/data/, 2013.

[24] A. Kaestner, B. Munch, T. Pavel, and L. Butler, “Spatiotemporal computed tomography of dynamic processes,” *Optical Engineering*, vol. 50, no. 12, p. 123201, 2011.

Part III

Conclusions and Appendices

6

Conclusions

Limited data problems arise in various CT applications, ranging from pore quantification in electron tomography to fast fluid flow and cerebral perfusion CT. Solving these problems with classical tomography algorithms typically results in reconstructions that are deteriorated by various limited data artifacts.

In this thesis, local model-based reconstruction methods for improved reconstruction quality in a limited data tomography setting were introduced. There are several design principles that were taken into account during the development of the proposed algorithms:

- Prior knowledge about the scanned object is exploited by imposing model-based constraints on the reconstruction during iterative reconstruction.

- The model assumptions are applied locally, i.e., in a specific region only, thereby differentiating adequately between specific properties of mutually different regions.

- The validity of the local model assumptions can be checked during reconstruction, in order to increase the robustness of the algorithm.

By more accurately modelling the sample, the reconstruction problem becomes more determined, which generally results in improved reconstruction quality. These design principles were applied during the development of the iterative algorithms in Chapter 2-5. In what follows, the main conclusions drawn in these chapters are discussed.

Part I: Stationary CT

Chapter 2 – Reconstruction method for improved porosity quantification of porous materials

In this chapter, the PORES algorithm (“*POre REconstruction and Segmentation*”) was introduced. It is a tailor-made, integral approach, for the reconstruction, segmentation, and quantification of porous (nano)materials. In the reconstruction step, prior knowledge about the sample is introduced into the reconstruction algorithm by simultaneously reconstructing the sample and classifying the interior region to the pores. The resulting reconstruction can be directly plugged into the remaining processing chain of the PORES algorithm, resulting in accurate individual pore quantification and full sample pore statistics. The proposed approach was extensively validated on both simulated and experimental data, indicating its ability to generate accurate statistics of porous materials, even if only few projection data is available.

Part II: Dynamic CT

Chapter 3 – Region-based iterative reconstruction

In this chapter, a reconstruction method was introduced for locally time-varying objects, that is, objects that change dynamically only in a local region inside the object. The reconstruction approach of this chapter, i.e., the region-based SIRT (rSIRT) algorithm, forms the basis of the algorithms that were introduced in Chapter 4 and 5.

The rSIRT algorithm assumes the existence of stationary regions, i.e., regions that remain unaltered over time, within the scanned object. This assumption is frequently exploited by calculating updates for the time-varying and the stationary region independently. The update for the time-varying region is based on projection data that corresponds solely to each time frame, thereby accurately representing the state-of-knowledge about each time frame, while the update for the stationary region is based on all projection data. Furthermore, we introduced an iterative optimization routine that can automatically determine these regions. The proposed algorithm was validated on simulation data and experimental μCT data, illustrating its capability to reconstruct locally time-varying objects more accurately in comparison to current techniques.

Chapter 4 – An iterative CT reconstruction algorithm for fast fluid flow imaging

In this chapter, an iterative CT reconstruction algorithm for improved temporal/spatial resolution in the imaging of fluid flow through solid matter was introduced. Analogously to Chapter 3, the time-varying object is assumed to consist of stationary (the solid matter) and dynamic regions (the fluid flow). Secondly, the attenuation curve of a particular voxel in the dynamic region is modeled by a piecewise constant function over time, which is in accordance with the actual advancing fluid/air boundary. These model assumptions make our proposed reconstruction algorithm more robust to limited data. This was illustrated by quantitative and qualitative experiments with different simulation phantoms and a real neutron tomography dataset. These experiments show that in comparison to state-of-the-art algorithms, the proposed algorithm allows reconstruction from substantially fewer projections per rotation without image quality loss. This makes it possible to further increase the temporal resolution in fluid flow experiments, thereby providing a much-needed method for probing high speed fluid dynamics.

Chapter 5 – Iterative reconstruction for low-dose cerebral perfusion CT

In this chapter, we introduced the Local Enhancement Steered Tomography (LEST) method. The brain volume is assumed to have stationary regions over time (bone and void space) and the temporal relation between the different time frames in the dynamic region (brain vessels and tissue) is exploited. Furthermore, the shape of the arterial input function (AIF) is independently optimized based on the projection data at intermediate iterations. The LEST algorithm was extensively validated with simulation and real clinical experiments and its performance was compared to different basic and state-of-the-art methods. Quantitative and qualitative results show that LEST is able to substantially reduce the radiation dose while maintaining image quality in comparison to these methods.

The progress made in this work is a step forward for accurate reconstruction in limited data CT problems arising in a variety of applications, including pore quantification in electron microscopy, the reconstruction of locally time-varying objects, fluid flow reconstruction problems and cerebral perfusion CT.

B

List of common abbreviations and symbols

Common symbols

$\boldsymbol{C}$	$\in \mathbb{R}^{N\times N}$	Diagonal matrix with inverse column sums of the projection matrix $\boldsymbol{W}$.
$\tilde{\boldsymbol{C}}_r$	$\in \mathbb{R}^{N\times N}$	Diagonal matrix with inverse column sums of $\tilde{\boldsymbol{W}}_r$.
$\boldsymbol{C}_r$	$\in \mathbb{R}^{N\times N}$	Diagonal matrix with inverse column sums of $\boldsymbol{W}_r$.
f	$: \mathbb{R}\times\mathbb{R}\rightarrow\mathbb{R}$	Function that describes the object's attenuation coefficient μ at each two-dimensional coordinate.
i	$\in \mathbb{N}_0$	Index associated with the different projection values in the projection data.
I_0	$\in \mathbb{R}^+$	The X-ray beam intensity at the source position.
$\boldsymbol{I}_S$	$\in \{0,1\}^{N\times N}$	Binary diagonal matrix representing the operator that sets all pixels belonging to the time-varying region inside the reconstruction domain to 0.
$\boldsymbol{I}_V$	$\in \{0,1\}^{N\times N}$	Binary diagonal matrix representing the operator that sets all pixels belonging to the stationary region inside the reconstruction domain to 0.
j	$\in \mathbb{N}_0$	Index associated with the different attenuation values in a reconstruction.
k	$\in \mathbb{N}$	Iteration number.
K	$\in \mathbb{N}$	The total number of iterations.

M	$\in \mathbb{N}_0$	The number of measured projection values ($M = M_t M_d$).
M_t	$\in \mathbb{N}_0$	The number of acquired projections per time frame.
M_d	$\in \mathbb{N}_0$	The number of detector elements in a single projection.
N	$\in \mathbb{N}_0$	The number of pixels in a reconstruction image.
n_w	$\in \mathbb{N}_0$	The window size in the sliding window approach for projection data subset selection.
$\boldsymbol{p}$	$\in \mathbb{R}^M$	The log-corrected projection data associated with a stationary tomography problem.
$\boldsymbol{p}_r$	$\in \mathbb{R}^M$	The log-corrected projection data associated with the rth time frame of the reconstruction in a dynamic tomography problem ($r = 1, \ldots, R$).
$\tilde{\boldsymbol{p}}$	$\in \mathbb{R}^{RM}$	The log-corrected projection data associated with a dynamic tomography problem in which R time frames are considered.
r	$\in \{1, \ldots, R\}$	Index referring to a specific time frame.
R	$\in \mathbb{N}_0$	The number of time frames.
$\boldsymbol{R}$	$\in \mathbb{R}^{RM \times RM}$	Diagonal matrix with inverse row sums of the projection matrix $\boldsymbol{W}$.
$\tilde{\boldsymbol{R}}_r$	$\in \mathbb{R}^{n_w M_d \times n_w M_d}$	Diagonal matrix with inverse row sums of $\tilde{\boldsymbol{W}}_r$.
$\boldsymbol{R}_r$	$\in \mathbb{R}^{M \times M}$	Diagonal matrix with inverse row sums of $\boldsymbol{W}_r$.
$\boldsymbol{S}_K$	$\in \mathbb{R}^{RN \times RM}$	The linear operator that represents K iterations of the rSIRT algorithm with initial estimate $\tilde{\boldsymbol{x}}^{(0)} = \boldsymbol{0}$.
S	$\subset \{1, \ldots, N\}$	Set of pixel indices that correspond to the stationary regions within the reconstruction.
t_r	$\in \mathbb{R}^+$	The time point associated with the rth time frame.
V	$\subset \{1, \ldots, N\}$	Set of pixel indices that correspond to the time-varying regions within the reconstruction.
$\boldsymbol{W}$	$\in \mathbb{R}^{RM \times N}$	Sparse matrix that encodes the forward projection of a reconstruction $\boldsymbol{x}$ along all acquisition angles ($R = 1$ in stationary CT and $R > 1$ in dynamic CT).
$\tilde{\boldsymbol{W}}$	$\in \mathbb{R}^{RM \times RN}$	Sparse matrix that simulates the forward projection of a dynamic reconstruction to its corresponding simulated projection data.

$\boldsymbol{W}_r$	$\in \mathbb{R}^{M \times N}$	Forward projection matrix that maps the reconstruction's rth time frame to its corresponding projection data.
$\boldsymbol{x}$	$\in \mathbb{R}^{N}$	The column vector representing the attenuation values of the reconstructed image in a stationary tomography problem.
$\boldsymbol{x}_r$	$\in \mathbb{R}^{N}$	The column vector representing the attenuation values of the rth time frame of the reconstructed image in a dynamic tomography problem.
$\tilde{\boldsymbol{x}}$	$\in \mathbb{R}^{RN}$	The column vector representing the attenuation values of the reconstructed time-evolving image in a dynamic tomography problem.
μ	$\in \mathbb{R}^{+}$	The attenuation coefficient.
ω_l	$\in [0, 2\pi]$	The acquisition angle of the lth projection ($l = 1, \ldots, M_t R$).

C

Scientific contribution

Journal articles

- *G. Van Eyndhoven*, M. Kurttepeli, C. J. Van Oers, P. Cool, S. Bals, K. J. Batenburg, and J. Sijbers, *"Pore REconstruction and Segmentation (PORES) method for improved porosity quantification of nanoporous materials"*, Ultramicroscopy, vol. 148, pp. 10–19, 2015.

- *G. Van Eyndhoven*, K. J. Batenburg, and J. Sijbers, *"Region-based iterative reconstruction of structurally changing objects in CT"*, IEEE Transactions on Image Processing, vol. 23, no. 2, pp. 909–919, 2014.

- *G. Van Eyndhoven*, K. J. Batenburg, D. Kazantsev, V. Van Nieuwenhove, P. D. Lee, K. J. Dobson, and J. Sijbers, *"An iterative CT reconstruction algorithm for fast fluid flow imaging"*, IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 4446-4458, 2015.

- D. Kazantsev, W. M. Thompson, W. R. B. Lionheart, *G. Van Eyndhoven*, A. P. Kaestner, K. J. Dobson, P. J. Withers, and P. D. Lee, *"4D-CT reconstruction with unified spatial-temporal patch-based regularization"*, Inverse Problems and Imaging, vol. 9, no. 2, pp. 447–467, 2015.

- D. Kazantsev, *G. Van Eyndhoven*, W. R. B. Lionheart, P. J. Withers, K. J. Dobson, S. A. McDonald, R. Atwood, and P. D. Lee, *"Employing temporal self-similarity across the entire time domain in computed tomography reconstruction"*, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 373, no. 2043, The Royal Society, 2015.

- *G. Van Eyndhoven*, K.J. Batenburg, N. Buls, J. Vandemeulebroucke, J. De Beenhouwer, and J. Sijbers, *"A local enhancement steered tomography (LEST)* *algorithm for low-dose cerebral perfusion CT"*, IEEE Transactions on Computational Imaging, in Submission.

Conference proceedings (full paper)

- G. Van Eyndhoven, J. Sijbers, and K. J. Batenburg, *"Combined Motion Estimation and Reconstruction in Tomography"*, ECCV'12, Ws/Demos, vol. 7583, Firenze, LNCS, pp. 12-21, 2012.
- G. Van Eyndhoven, K. J. Batenburg, and J. Sijbers, *"Region based 4D tomographic image reconstruction: application to cardiac X-ray CT"*, IEEE International Conference on Image Processing, in Press, 2015.

Conference proceedings (extended abstract)

- G. Van Eyndhoven, K. J. Batenburg, and J. Sijbers, *"An Algebraic Reconstruction Technique for the Study of Local Structural Changes during CT"*, 1st International Conference on Tomography of Materials and Structures, 2013.
- G. Van Eyndhoven, K. J. Batenburg, and J. Sijbers, *"Region-Based SIRT algorithm for the reconstruction of phase bins in dynamic micro-CT"*, Bruker Micro-CT User Meeting, 2013.
- G. Van Eyndhoven, K. J. Batenburg, and J. Sijbers, *"Accurate reconstruction of porous samples in CT"*, International Conference on Industrial Computed Tomography (ICT), Wels, Austria, 2014.
- V. Van Nieuwenhove, G. Van Eyndhoven, J. De Beenhouwer, and J. Sijbers, *"Compensation of affine deformations in fan and cone beam projections"*, Micro-CT User Meeting, Ostend, Belgium, pp. 187-189, 2014.

Conference abstracts

- G. Van Eyndhoven, K. J. Batenburg, D. Kazantsev, V. Van Nieuwenhove, P. D. Lee, K. J. Dobson, and J. Sijbers, *"A 4D CT reconstruction algorithm for fast liquid flow imaging"*, Applied Inverse Problems Conference, Helsinki, Finland, 2015.
- G. Van Eyndhoven, K. J. Batenburg, D. Kazantsev, and J. Sijbers, *"An iterative prior knowledge based reconstruction algorithm for increased temporal/spatial resolution in the CT imaging of fluid flow through solid matter"*, Interpore 2015, Padova, Italy, 2015.

- V. Van Nieuwenhove, *G. Van Eyndhoven*, J. De Beenhouwer, and J. Sijbers, *"Combined Estimation of Affine Movement and Reconstruction in Tomography"*, International Congress on 3D Materials Science, Annecy, France, 2014.
- *G. Van Eyndhoven*, K. J. Batenburg, and J. Sijbers, *"A local enhancement based tomographic reconstruction technique for radiation exposure reduction in cerebral perfusion computed tomography"*, European Congress of Radiology 2014 (ECR 2014), Vienna, Austria, 2014.
- *G. Van Eyndhoven*, K. J. Batenburg, C V. Oers, M. Kurttepeli, S. Bals, P. Cool, and J. Sijbers, *"Reliable Pore-size Measurements Based on a Procedure Specifically Designed for Electron Tomography Measurements of Nanoporous Samples"*, International Congress on 3D Materials Science 2014, Annecy, France, 2014.