Universiteit Antwerpen
Faculteit Wetenschappen
Departement Fysica

Towards accurate image reconstruction from truncated X-ray CT projections

---

Methoden voor accuratere beeldreconstructie vanuit getrunceerde X-stralen CT projecties

Proefschrift voorgelegd tot het behalen van de graad van

Doctor in de Wetenschappen

aan de Universiteit Antwerpen, te verdedigen door

Gert VAN GOMPEL

Promotor
Prof. Dr. Jan Sijbers

Copromotoren
Prof. Dr. Michel Defrise
Dr. K. Joost Batenburg

Antwerpen, 2009

Cover illustration

Truncated sinogram of a diamond slice. The back cover image represents a slice-by-slice reconstruction of a complete diamond from transaxially truncated projections.

Predefense
Monday, August 31 2009 at 4.00PM, CDE N1.08
Universiteit Antwerpen

Public defense
Monday, September 28 2009 at 3.00PM, CMI G.010
Universiteit Antwerpen

Members of the jury

Chairman
Prof. Dr. Wim Decraemer, Universiteit Antwerpen

Supervisors
Prof. Dr. Jan Sijbers, Universiteit Antwerpen
Prof. Dr. Michel Defrise, Vrije Universiteit Brussel
Dr. K. Joost Batenburg, Universiteit Antwerpen

Members
Prof. Dr. Marleen Verhoye, Universiteit Antwerpen
Prof. Dr. Willy Malfliet, Universiteit Antwerpen
Prof. Dr. Franz Pfeiffer, Technische Universität München
Prof. Dr. Johan Nuyts, Katholieke Universiteit Leuven

Contact information

✉ Gert Van Gompel
Vision Lab, Dept. of Physics
University of Antwerp (CDE)
Universiteitsplein 1, Building N1.19
B-2610 Wilrijk – Antwerpen, Belgium

☎ +32 (0)3 265 24 72

+32 (0)3 265 22 45

`gert.vangompel@ua.ac.be`

`http://visielab.ua.ac.be/staff/gevgo`

Contents

List of Figures

List of abbreviations

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
AIBHC	Accelerated Iterative Beam Hardening Correction
BH	Beam hardening
ConSiR	Consistent Sinogram Recovery
CT	Computed Tomography
DART	Discrete Algebraic Reconstruction Technique
DBP	Differentiated Backprojection
DBPS	Differentiated Backprojection method for uniform Star objects
FBP	Filtered Backprojection
FDK	Feldkamp-David-Kress
FOV	Field Of View
FWHM	Full Width at Half Maximum
IBHC	Iterative Beam Hardening Correction
IPR	Iterative Post Reconstruction
MAE	Mean Absolute Difference
PET	Positron Emission Tomography
ROI	Region Of Interest
RT	Radon Transform
SEM	Simple Extrapolation Method
SIRT	Simultaneous Iterative Reconstruction Technique
SIRT-t	Simultaneous Iterative Reconstruction Technique - threshold
SPECT	Single Photon Emission Computed Tomography
TVmin	Total Variation minimization

Summary

Context

X-ray transmission computed tomography (CT) is a non-destructive imaging technique providing 3D structural information of an object under examination. The desired 3D image is computed from a set of X-ray projections of the object, that are recorded by the X-ray CT scanner at different orientations. CT has many important applications such as medical diagnostics and small animal imaging, which supports pre-clinical drug testing. Other applications are found in the industry (diamond, packaging, construction), where CT scanners are employed for non-destructive testing.

Accurate image reconstruction requires that the detector measures the projection data of the complete object in all directions such that the recorded X-ray projections are not truncated. However, in a practical setup, this requirement is often not met. As a consequence, significant cupping artifacts are observed in the reconstructed images, hindering accurate segmentation and quantitative analysis of the reconstructed attenuation values. In addition, the reconstruction of the object regions outside the detector field-of-view provides only vague structural information. This thesis concerns the problem of the accurate recovery of a two-dimensional object slice from one-dimensional projections that are transaxially truncated.

Truncation of the projections is often present in medical X-ray CT. The diameter of the gantry is relatively small (typically 70 cm), and the scanning field-of-view diameter even smaller ($\pm$50 cm). These small dimensions inevitably lead to truncated data if the examined patient is obese, incorrectly placed on the scanning bed, or if a scan at shoulder level is aimed at.
Alternatively, adequate truncation artifact reduction would allow X-ray beam collimation such that only the desired body volume is exposed to the radiation and the dose can be reduced significantly.

For biomedical and industrial applications, often μCT systems are used, which provide high resolution images of small objects. The resolution of a small *region-of-interest* (ROI) in the object can be increased by translating the object towards the source of the X-ray cone beam such that the ROI covers a larger part of the detector. However, this inevitably causes the detector to miss transmission data from the other parts of the object in certain directions.
In μCT, one also frequently faces the problem where the object is too large to be covered by the detector in each projection direction. Some applications, such as diamond optimization, require accurate reconstruction of the complete object, rather than only in the detector field-of-view.

Outline

This thesis is subdivided into four parts:

Part I. Background: The first chapter provides a general introduction to X-ray transmission computed tomography, covering the production of X-rays, X-ray imaging, and the most commonly used reconstruction methods. Also, some artifacts are discussed that are encountered in practice. Two groups or artifacts are considered: artifacts caused by degraded data, such as beam hardening, and artifacts caused by incomplete data.

Part II. Reduction of truncation artifacts in standard 2D X-ray CT: This part concerns the accuracy enhancement in images reconstructed from truncated projections. In particular, the focus is put on accurate reconstruction in the field-of-view (FOV), which is the object region that is covered by the detector in all projection directions. Chapter 2 shows that traditional reconstruction methods fail at obtaining an accurate image from truncated projections for any level of truncation, and describes how this problem can be partially alleviated by using an alternative analytical Radon inversion formulation. Subsequently, the most recent data sufficiency conditions are summarized that determine the object regions for which a unique solution exists. The corresponding methods that recover this unique solution are briefly described. Finally, an overview is provided of the available data completion methods, and of empirical, approximating reconstruction techniques that are currently available.
In Chapter 3, a new approximating method (ConSiR) for the reduction of truncation artifacts is presented. This technique exploits the sinogram consistency to extrapolate the missing sinogram data. The algorithm is applied to simulated and real X-ray data and its performance is compared to that of several other methods proposed in the literature.

Part III. Reconstruction of piecewise uniform objects from truncated data: In recent years, interesting reconstruction results have been obtained for a variety of limited data problems by incorporating certain prior knowledge of the object. Examples of such prior knowledge are sparseness, such as in angiography where blood vessels are imaged, and piecewise uniformity of the object, which is frequently the case in non-destructive testing. Part III investigates the possibility of recovering accurate images of piecewise uniform objects from truncated projections, not only in the FOV such as in Part II, but in the complete object support.

In Chapter 4, it is proved that a binary star shaped phantom is uniquely determined from its *interior* data, which represents the most severe type of truncation. The stability is discussed and a numerical algorithm is proposed, which transforms the 2D inverse problem into a set of 1D problems along radial lines. This reconstruction method is quantitatively evaluated in simulation experiments considering noise-free and noisy data. Chapter 5 investigates whether algorithms that treat a 2D image as a whole are valuable compared to the line-per-line technique proposed in Chapter 4. In addition, the aim of Chapter 5 is to experimentally indicate the extent of object complexity for which the inverse Radon transform from truncated data can be recovered. To this end, we investigated the new application of DART (*discrete algebraic reconstruction technique*) on truncated data of non-star shaped piecewise uniform objects containing one or multiple densities.

Part IV. Reconstruction of piecewise uniform objects in practice: For real X-ray data, the techniques from Chapters 4 and 5 frequently do not lead to accurate reconstructions. The main reason is that non-linear physical effects such as beam hardening invalidate the assumption of constant densities. Therefore, Part IV concerns beam hardening correction for piecewise uniform objects, which is a necessary step before applying the previous methods designed for truncated data.

In Chapter 6, the literature on beam hardening reduction is summarized, and a novel beam hardening correction algorithm for piecewise uniform objects is presented that does not require information on the source spectrum or on the energy dependent attenuation coefficients of the present materials. The method is applied on real non-truncated polychromatic X-ray data of several physical phantoms, and the resulting reconstruction images are quantitatively evaluated. The ultimate goal is to provide accurate images from truncated data, or other types of limited data, even if the data are degraded by beam hardening artifacts. To this end, the beam hardening correction method is combined with iterative approaches that are specifically developed for limited data. The resulting combination methods are then applied on real X-ray CT data for several cases of limited data problems (Chapter 7).

Nederlandse samenvatting

Context

X-stralen transmissie computer tomografie (CT) is een niet-destructieve techniek voor de 3D beeldvorming van objectstructuren. De beelden worden berekend vanuit een set X-stralen projecties van het object, die opgenomen worden door de CT scanner onder verschillende orientaties. CT heeft een hele waaier aan belangrijke toepassingen gaande van medische diagnostiek en beeldvorming van muizen en ratjes voor het preklinisch testen van geneesmiddelen, tot het niet-destructief testen in, bv., de diamant-, verpakkings-, en constructie-industrie.

Accurate beeldreconstructie vereist, voor de klassieke CT acquisitie geometrieën, dat de volledige projectie van het object gekend is voor elke bron-detector oriëntatie. In de praktijk zijn de projecties echter vaak getrunceerd. De daaruit voortvloeiende reconstructies zijn typisch gecontamineerd met een additieve, niet-constante, object-afhankelijke bias, die accurate segmentatie en kwantitatieve analyse van de beelden verhindert. Meer nog, in het gebied rondom het gezichtsveld van de detector ('field of view' (FOV)) kan met de klassieke methodes slechts beperkte structurele informatie verkregen worden. Dit proefschrift heeft tot doel een accuratere beeldreconstructie te bekomen van een tweedimensionale (2D) object snede vanuit 1D projecties die transaxiaal getrunceerd zijn.

Truncatie van projecties komt niet zelden voor in medische X-stralen CT. De diameter van de opening rond dewelke de bron en de detector draaien, is relatief klein (typisch 70 cm), en de radius van het gezichtsveld van de detector is zelfs nog kleiner (ongeveer 50 cm). Deze kleine dimensies veroorzaken onvermijdelijk getrunceerde data wanneer de patient obees is, slecht gecentreerd ligt op het CT bed, of bij scans op schouderhoogte.
In het kader van dosisreductie kan truncatie ook intentioneel plaatsvinden door de belichting te reduceren tot het gebied van interesse (bv. enkel het hart), in plaats van de volledige doorsnede van het lichaam.

In μCT, waarbij hoofdzakelijk hoge-resolutie beelden worden nagestreefd, kan de resolutie van een kleine regio in het object verbeterd worden door het object dichter bij de bron te plaatsen, zodat de projectie van deze kleine regio een groter gedeelte van de detector bedekt. Dit zorgt echter onvermijdelijk voor truncatie van de projecties van de omliggende regio's in het object.
Anderzijds past in μCT, net zoals in medische CT, het object vaak niet voor elke projectierichting in het gezichtsveld van de detector. Sommige toepassingen vereisen ook in deze situatie een accurate reconstructie van het volledige object, eerder dan enkel van het gebied in het gezichtsveld van de detector. Diamantoptimalisatie, een toepassing die aan bod komt in Hoofdstuk 5 van dit proefschrift, is hier een voorbeeld van.

Overzicht

Deze thesis is onderverdeeld in vier delen.

Achtergrond: Het eerste hoofdstuk geeft een algemene inleiding tot X-stralen transmissie tomografie, waarbij de productie van X-stralen, X-stralen beeldvorming en de standaard reconstructiemethodes aan bod komen. Verder worden enkele artifacten besproken die in de praktijk vaak voorkomen. Twee klassen van artefacten worden daarbij beschouwd: degene afkomstig van verstoorde data, of veroorzaakt door een tekort aan data.

Reductie van truncatie artefacten in standaard 2D X-stralen CT: Deel II streeft accuratere reconstructie na van beelden die gereconstrueerd worden uit getrunceerde projecties. De focus wordt in het bijzonder gelegd op accurate reconstructie van het gebied dat zich voor alle projectierichtingen in het gezichtsveld van de detector bevindt. Hoofdstuk 2 toont dat de traditionele standaard methodes falen in het verkrijgen van accurate beelden voor elke graad van projectie truncatie. Verder wordt ook beschreven hoe dit probleem gedeeltelijk opgelost wordt door het gebruik van een alternatieve analytische formulering van de inverse Radon transformatie. Vervolgens worden de meest recente uniciteitsvoorwaarden voor de projectiedata beschreven. Tenslotte volgt een overzicht van de beschikbare datavervolledigingsmethodes en van de empirische benaderende reconstructiemethodes die momenteel beschikbaar zijn.
In hoofdstuk 3 wordt een nieuwe methode (ConSiR) gepresenteerd voor de reductie van truncatie-artefacten. Deze techniek vervolledigt de ontbrekende data door het optimaliseren van de sinogram consistentie. Het algoritme wordt toegepast op gesimuleerde en reële data en de resulterende reconstruc-

tiebeelden worden kwantitatief vergeleken met die van verscheidene methodes uit de literatuur.

Reconstructie van stuksgewijs uniforme objecten vanuit getrunceerde data: Recent werden nieuwe reconstructie resultaten bekomen voor diverse gelimiteerde-data problemen door het inbrengen van specifieke voorkennis. Zulke voorkennis is in verscheidene toepassingen voorhanden. In angiografie bijvoorbeeld, zijn de af te beelden bloedvaten typisch spaars, en in industriële toepassingen kan vaak aangenomen worden dat het object stuksgewijs uniform is. Deel III onderzoekt de mogelijkheid om accurate beelden van stuksgewijs uniforme objecten te reconstrueren vanuit een set getrunceerde projecties, waarbij reconstructie beoogd wordt in het volledige domein van het object, en dus niet enkel in de FOV, zoals in Deel II.
In Hoofdstuk 4 wordt aangetoond dat een binair stervormig object uniek bepaald wordt door getrunceerde data, zelfs wanneer de FOV zich volledig in het inwendige van het object bevindt, wat overeenkomt met het meest ernstige type van projectietruncatie. De stabiliteit wordt onderzocht en een numeriek algoritme wordt voorgesteld dat het 2D inverse problem transformeert naar een set van 1D problemen langsheen radiale lijnen. Deze reconstructiemethode wordt geëvalueerd in simulatie-experimenten voor data zonder en met ruis. Hoofdstuk 5 onderzoekt of algoritmes die een 2D beeld als geheel reconstrueren, te verkiezen zijn boven de lijn-per-lijn techniek die voorgesteld werd in Hoofdstuk 4. Verder heeft Hoofdstuk 5 tot doel om experimenteel af te tasten voor welke graad van objectcomplexiteit reconstructie nog mogelijk is in de praktijk. Hiervoor onderzoeken we de toepassing van DART (discrete algebraische reconstructie techniek) op getrunceerde data van niet-stervomige stuksgewijs-constante objecten bestaande uit een of meerdere densiteiten.

Reconstructie van stuksgewijs uniforme objecten in de praktijk: Voor reële X-stralen data leiden de technieken die gebruikt werden in Hoofdstukken 4 en 5 vaak niet tot accurate reconstructies. De belangrijkste reden is dat niet lineaire fysische effecten zoals bundelverharding ('beam hardening' (BH)) de aanname van constante grijswaarden ongeldig maken.
Hoofdstuk 6 vat de literatuur over de reductie van BH-artefacten samen, en presenteert een nieuw BH-correctie-algorithme voor stuksgewijs uniforme objecten dat geen informatie vereist over het bronspectrum of over de energieafhankelijke attenuatiecoefficienten van de aanwezige materialen. De methode wordt toegepast op reële niet-getrunceerde polychromatische X-stralen data van verscheidene fysische fantomen, en de resulterende reconstructiebeelden worden quantitatief geëvalueerd. Het uiteindelijke streefdoel is om accurate beelden van stuksgewijs uniforme objecten te bekomen vanuit getrunceerde projecties of andere types van gelimiteerde data, zelfs als de data

verstoord is door bundelverharding. In Hoofdstuk 7 wordt het BH-correctie-algoritme gecombineerd met enkele iteratieve methoden die de weinige beschikbare data optimaal benutten. De resulterende combinatiemethodes worden vervolgens toegepast op reële X-stralen CT data voor verscheidene gelimiteerde-data problemen.

Part I

Background

Chapter 1

Introduction to Computed Tomography

This chapter provides a general introduction to computed tomographic imaging. The first section explains how X-rays are produced, how they interact with matter, and how the X-ray intensity decays when traversing matter. Section 1.2 describes how the properties of X-rays can be exploited for 3D tomographic imaging. The Radon transform, which is a linear mathematical model for X-ray attenuation, is introduced in Section 1.3, and a few standard methods are described for the recovery of the object function from its Radon transform. Section 1.4 consists of examples of artifacts that are encountered in practice.

For more introductory information, we refer to the work of Kak and Slaney [1], and the more recent books of Hsieh [2], Kalender [3] and Buzug [4]. A thorough theoretical study on tomographic image reconstruction is found in Natterer [5] and Natterer and Wübbeling [6].

1.1 Fundamentals of X-rays

1.1.1 Production of X-rays

X-rays are electromagnetic waves with a wavelength λ varying between 10 nm and 10^{-3} nm. These wavelengths correspond to energies from $E = 1.24 \times 10^{-1}$ keV to 1.24×10^{3} keV conform following expression:

$$E = \frac{hc}{\lambda}, \tag{1.1}$$

where h represents Planck's constant, and c the speed of light.

X-rays are generated in a vacuum tube containing a cathode and an anode. By thermal excitation, electrons are freed from the cathode. Subject to the electric voltage between the cathode and the anode, these electrons accelerate and eventually hit the anode at very high speed. Three possible types of interaction take place (see Fig. 1.1):

- A high-speed electron collides with an outer-shell electron, while transferring part of its energy towards the second electron. The energy of both electrons is then dissipated into heat. Collisions with outer-shell electrons (not shown in Fig. 1.1) represent the majority of the interactions, which explains why anode cooling is very important.

- A high-speed electron interacts with an inner-shell electron, which is then ejected from the atom, leaving an unoccupied lower energy level. This gap is filled by an outer-shell electron striving to the lowest energy position, and simultaneously a photon is released with an energy corresponding to the difference of energy levels of the two shells. Such photons take up only discrete energy values and are therefore called "characteristic radiation".

- A high-speed electron is suddenly decelerated when it passes near the nucleus of the atom, and emits a photon containing the excess energy. The amount of deceleration (and consequently the energy of the emitted photon) is determined by the distance of the electron to the nucleus. If the distance is large, the electron path undergoes a small deflection and deceleration, causing a low-energy photon to be emitted. When the electron collides directly onto the nucleus, it decelerates completely, and a high energy photon is released. The radiation emitted by this type of interaction, called "Bremsstrahlung", has a continuous spectrum.

1.1.2 Interaction with matter

X-rays interact with both electrons and atomic nuclei, and can be absorbed or (in)elastically scattered. The most important processes in the diagnostic energy range are the photo-electric effect, Compton scattering.

- The photo-electric effect (Fig. 1.2a) occurs when a bounded electron is hit by an X-ray photon with an energy larger than the electron's binding energy, and entirely absorbs the photon energy. Consequently, the electron is ejected from the atom and the excess photon energy is converted into kinetic energy. When an electron from a higher energy shell fills the gap on the lower energy

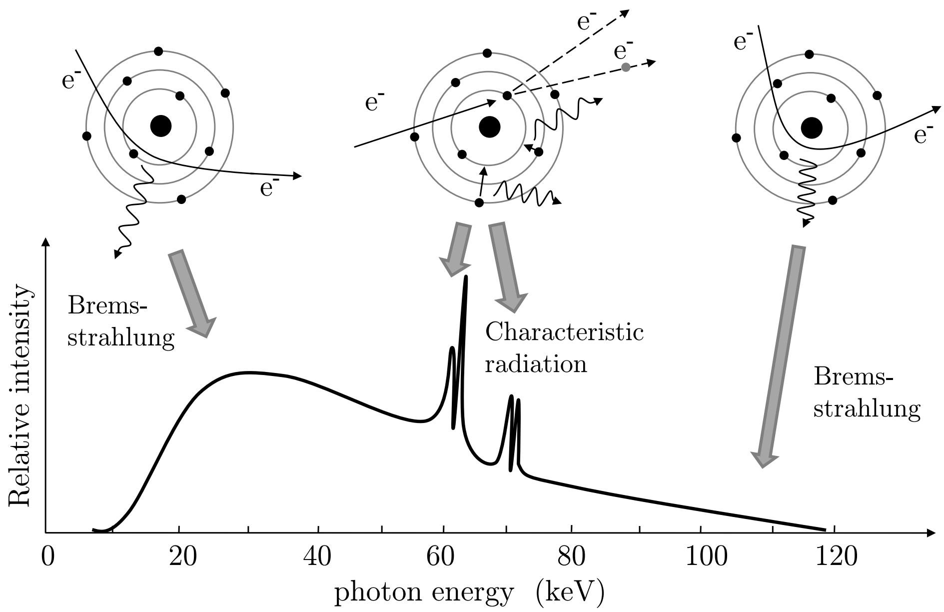
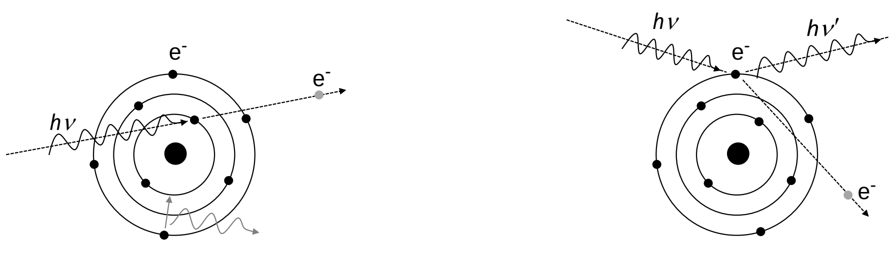


Figure 1.1: *X-ray spectrum of a tungsten tube. The peaks correspond to the characteristic radiation; the continuous part of the spectrum represents the Bremsstrahlung.*



(a) Photo-electric effect

(b) Compton scattering

Figure 1.2: *X-ray-matter interactions in the diagnostic range.*

shell, a so called "characteristic photon" is released. This X-ray photon is usually soft and has very small penetration capacity.
The probability of the photo-electric interaction is inversely proportional to the third power of the excess photon energy. For materials with a low atomic number Z, the binding energy of the K-shell electrons is small (e.g. roughly 500 eV for soft biological tissue) compared to the energy of the diagnostic X-rays, causing it to be almost transparent to the X-rays. Calcium has a slightly higher K-shell binding energy (approx. 4 keV), but the cubic probability relation provides a much higher attenuation so that calcium is significantly less transparent. Hence, the photo-electric effect yields a large contrast between materials with only slightly different atomic numbers.

- Compton scattering (Fig. 1.2b) occurs when an X-ray photon interacts with a bounded electron at a much higher energy than the electron's binding energy. During the collision, the incident photon gives up part of its energy to eject the electron from the atom. The incident photon then continues with reduced energy on a path deflecting from its original path conform the law of momentum conservation. Low energy photons are preferentially backscattered (angle from 90 to 180 degrees) and thus not detected, while the high frequency photons have a high probability to be scattered in the forward direction (angle 0 to 90 degrees). These high energy photons, which might undergo several of such collisions, lose the spatial information of the interaction with the electrons, and therefore lead to artifacts (see Section 1.4.1.1) when they finally are detected. The occurance rate of the Compton scatter is proportional to the density of electrons in the material and not on the atomic number.

X-rays with an energy lower then 10 keV are considered as soft and of less value for diagnostic imaging since their penetration length is small, which results in low contrast ratios on the imaging device. Hard X-rays ($E > 140$ keV) have large penetration lengths in bone and tissue, and since that results in low contrast rates, they are also less valuable for medical diagnostics as well. For industrial imaging, however, X-rays of 100 keV to 300 keV are frequently used, e.g. for metal objects.

1.1.3 X-ray attenuation

The interactions described in Section 1.1.2 result in a gradual intensity loss or *attenuation* of an X-ray beam that enters a material. Consider a monochromatic X-ray beam with intensity I propagating through a homogeneous material. The absorption of the beam with respect to the traversed distance is described by the law of Beer-Lambert, stating that each layer of equal thickness dt absorbs an equal fraction $\frac{dI}{I}$ of the intensity of the beam that traverses it:

$$\frac{dI}{I} = -\mu dt, \tag{1.2}$$

with constant coefficient μ. Integration yields

$$I(t) = I_0 e^{-\mu t} \tag{1.3}$$

where I_0 is the incident intensity. Let T be the total intersection length between the object and the straight ray path. The measured intensity of the beam after passing through the object is then

$$I = I_0 e^{-\mu T} \tag{1.4}$$

Defining the *attenuation* A as

$$A = -\ln\left(\frac{I}{I_0}\right), \tag{1.5}$$

then the combination with Eq. (1.4)

$$A = \mu T \tag{1.6}$$

shows that the 'measured' attenuation A is linear with respect to the thickness of the object. The material dependent coefficient μ is called the *linear attenuation coefficient*. It mainly consists of contributions from the Compton effect μ_σ and the photo-electric effect μ_τ so that $\mu \simeq \mu_\tau + \mu_\sigma$.

When X-rays are traveling through inhomogeneous matter, μ depends on the position t along the ray, and the expressions for the intensity and attenuation are generalized as:

$$I = I_0 e^{-\int_0^T \mu(t)dt} \tag{1.7}$$

$$A = -\ln\left(\frac{I}{I_0}\right) = \int_0^T \mu(t)dt. \tag{1.8}$$

In general, X-ray beams are not monochromatic and the linear attenuation coefficients depend on the energy of the rays. Consequently, formula Eq. (1.7) has to be adapted for polychromatic X-rays:

$$I = \int_0^{E_0} I_0(E) e^{-\int_0^T \mu(t)dt} dE \tag{1.9}$$

This non-linear relation will be neglected until Chapter 6, which concerns the reduction of beam hardening artifacts.

Figure 1.3: *X-ray photo of the hand of Wilhelm Conrad Röntgen's wife [7].*

1.2 X-ray imaging

As explained in the previous section, the output intensity that is measured after an X-ray beam passed through an object, depends on the material distribution of the object along the ray. Consequently, the complete illumination of an object by an X-ray beam yields a 2D intensity image in which the contrast is induced by the varying structures and attenuation properties in the object. Such an intensity image, which basically represents a projection of a 3D object onto a 2D plane, is an *X-ray photo* or *projection*. Note that the term "X-ray projection", depending on context, also refers to the corresponding attenuation image that is acquired using Eq. (1.5). Fig. 1.3 shows the X-ray photo of the hand of Wilhelm Conrad Röntgen's wife, recorded shortly after he discovered X-rays in 1896.

X-ray projections are used widely, especially in medical applications because they offer high-contrast images of important substances in the body such as air, bone, tissue, fat, etc.

One limitation of X-ray photos is that they do not provide depth information, since the measured intensity of the beam that traversed the object is independent of the material order along the ray (Eq. (1.7)). However, when X-ray projections of the object are acquired at many different orientations, a complete 3D attenuation map of the object can be reconstructed. This technique, called *Tomographic imaging*, is widely used as a medical diagnostic tool, but has also various industrial applications.

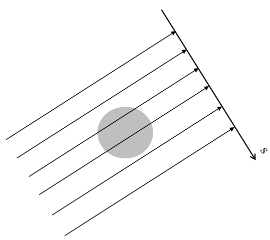
Fig. 1.4 shows a typical setup of a clinical CT scanner: a source-detector pair rotates around a patient that is lying on a bed. Medical setups typically generate 3D reconstructions with 1 mm^3 resolution. Alternative X-ray CT systems exist that

(a)

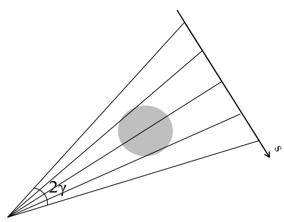
(b)

Figure 1.4: *(a) Philips clinical CT scanner [`http://www.healthcare.philips.com`]; (b) SkyScan 1172 micro CT scanner [`http://www.skyscan.be`]*

generate images of small objects at micrometer resolution. Such μCT systems are used for research on small animals (e.g. mice, rats) or in industrial applications (e.g. for quality study of metal foam). μ-CT systems for industrial applications often have a different scanning setup: the source and detector are fixed, while the sample is placed on a rotating sample holder (Fig. 1.4(b)).



(a)



(b)

Figure 1.5: *2D tomographic setup: (a) fan beam; (b) parallel beam.*

Several scanning geometries can be considered such as parallel beam, fan beam and cone beam geometry. The parallel (Fig. 1.5(a)) and fan-beam (Fig. 1.5(b)) geometry were typically used in early generations of commercial X-ray CT systems. Both geometries record the attenuation data slice-by slice. In a parallel beam geometry, the X-rays travel along parallel paths. In a fan beam geometry, characterized by opening angle 2γ, the X-rays are emitted from a single focal point for each projection.

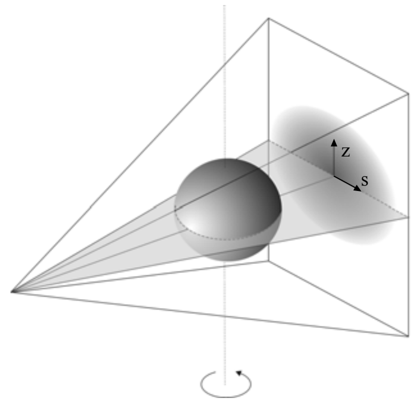


Figure 1.6: *3D cone beam tomographic setup.*

Opposed to parallel and fan beam, cone beam CT or multislice CT (Fig. 1.6) is essentially three dimensional. The X-ray point source, emitting a cone beam, illuminates the entire object or a stack of slices at each orientation, which results in 2D projections on the detector. Note that the signal on the detector line for $z = 0$ corresponds to a fan beam geometry. The corresponding illuminated object slice is referred to as the *central slice*.

Although the majority of current X-ray tomography systems uses a cone beam geometry, research for fan and parallel beam remains highly relevant since the standard 3D reconstruction algorithms basically perform a set of weighted 2D reconstructions along tilted planes. Throughout this thesis, we will consider a 2D parallel geometry, unless indicated otherwise.

1.3 Image reconstruction

In this section, standard methods are described for the reconstruction of a 2D object slice from its set of one-dimensional projections.

1.3.1 Radon transform

Consider a 2-dimensional function $f(x, y)$, representing the position dependent attenuation coefficient in an object slice, in a fixed coordinate system (x, y), see Fig. 1.7. Define a second coordinate system (s, t), which is a rotated version of (x, y) with orientation angle θ. The coordinate transforms between the two systems are

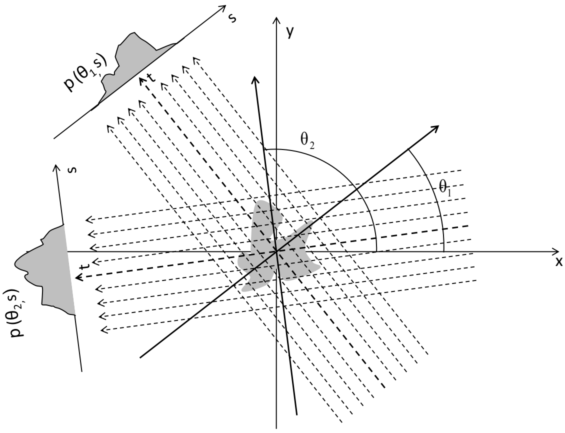


Figure 1.7: *Schematic figure for 2 projections with $\theta = \theta_1$ and $\theta = \theta_2$ of an object function $f(x, y)$. The variables (x, y) and (s, t) represent the fixed and the rotating coordinate system, respectively.*

given by:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} \tag{1.10}$$

and

$$\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{1.11}$$

Consider a straight path $L(\theta, s)$ at angle θ and signed distance to the center of rotation s. The line integral of f along the ray $L(\theta, s)$ is given by

$$p(\theta, s) = \int_{L(\theta,s)} f(x(s,t), y(s,t))dt, \tag{1.12}$$

or, alternatively,

$$p(\theta, s) = \int_{-\infty}^{\infty} f(s\boldsymbol{\alpha} + t\boldsymbol{\alpha}^{\perp})dt, \tag{1.13}$$

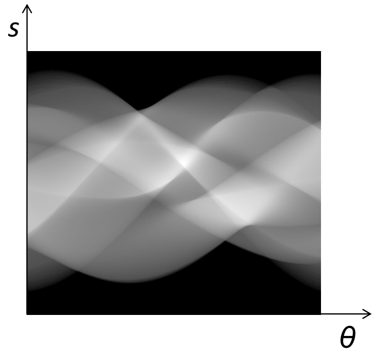
with $\boldsymbol{\alpha} = (\cos\theta, \sin\theta)$ and $\boldsymbol{\alpha}^{\perp} = (-\sin\theta, \cos\theta)$.

The transform $\mathcal{R}$ that maps the object function $f(x, y)$ onto its complete set of line integrals

$$(\mathcal{R}f)(\theta, s) = p(\theta, s) \tag{1.14}$$

is known as the *Radon Transform* (RT), named after Johann Radon who published an inversion formula in 1917. For instance, Fig. 1.8(b) depicts the Radon transform of a certain function $f(x, y)$, shown in Fig. 1.8(a). From now on, the indication of

(a)



(b)

Figure 1.8: *Illustration of the sinogram (b) of a certain object function $f(x,y)$ depicted in (a).*

the (θ, s) axes in images of the Radon Transform are omitted. The RT of a 2D slice is often called *sinogram* since the support [1] of the RT of a Dirac delta function corresponds to a sine. The Radon Transform has following properties :

- $p(\theta, s)$ is periodic in θ with period 2π

$$p(\theta, s) = p(\theta + 2\pi, s), \tag{1.15}$$

- $p(\theta, s)$ is symmetric in θ around π

$$p(\theta, s) = p(\theta \pm \pi, -s). \tag{1.16}$$

Assuming a monochromatic X-ray beam, note that the attenuation of an X-ray beam traversing an object along a straight line corresponds to the Radon transform. Hence, the problem of reconstructing the attenuation coefficient $\mu(x, y)$ is equivalent to the problem of finding an object function $f(x, y)$ from its Radon transform. In this thesis, except in Chapter 6, we will use the more general term *object function* $f(x, y)$ to represent the attenuation coefficient distribution $\mu(x, y)$.

The set of line integrals for $s \in [-\infty, \infty]$ at a fixed angle θ, illustrated in Fig. 1.7 is called a *parallel projection*, or briefly *projection*, and corresponds to a projection in a parallel beam imaging setup as shown in Fig. 1.5 (a). Note that the Radon transform is completely determined by the set of parallel projections $p(\theta, s)$ covering the angular range $\theta \in [0, \pi)$. In a fan beam setup (Fig. 1.5(b)) with fan angle

[1]The support of a function $f(x, y)$ is defined as the set of coordinates (x,y) for which the function $f(x, y)$ is nonzero.

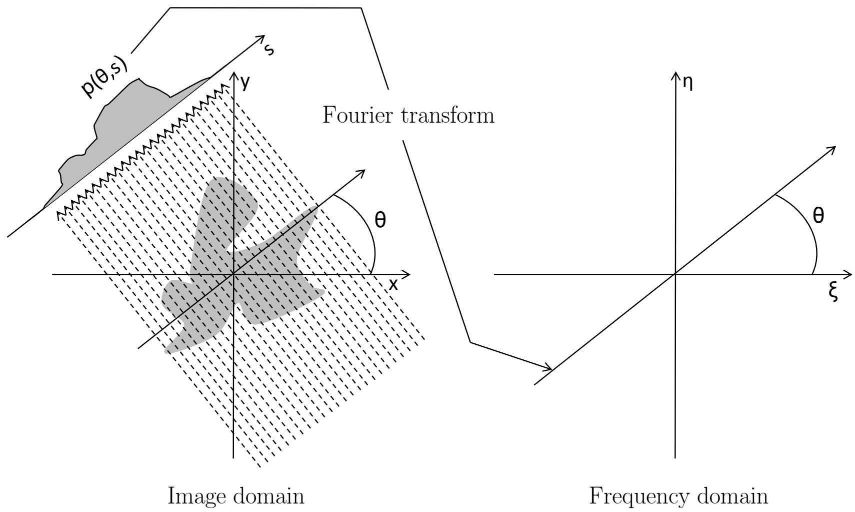


Figure 1.9: *Fourier slice theorem*

2γ, two fan beam projections recorded in the opposite directions do not represent the same line integrals. For fan beam, minimally a projection range of $\pi + 2\gamma$, called a 'short scan' is required to represent the complete Radon transform. Note that by re-arranging the ray paths, fan beam data can be converted into a parallel sinogram, which is called *rebinning*.

1.3.2 Analytical reconstruction

1.3.2.1 Fourier Slice Theorem

The Fourier Slice theorem plays a fundamental role in image reconstruction, since it relates the Fourier transform of the projections to the 2D Fourier transform of the object function.

To derive the Fourier Slice Theorem, define the two-dimensional Fourier transform $F(\xi, \eta)$ of the object function $f(x, y)$ as

$$F(\xi, \eta) = \mathcal{F}_{2\mathrm{D}} f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-i2\pi(x\xi + y\eta)} dx dy. \qquad (1.17)$$

On the other hand, working in the (s, t) coordinate system (Eq. (1.10) and Eq. (1.11)), consider a projection $p(\theta, s)$ with fixed θ

$$p(\theta, s) = \int_{-\infty}^{\infty} f(s, t) dt, \qquad (1.18)$$

and its one-dimensional Fourier transform $P(\theta, \sigma)$ with respect to s

$$P(\theta, \sigma) = \mathcal{F}_{1\mathrm{D}} p(\theta, s) = \int_{-\infty}^{\infty} p(\theta, s) e^{-i2\pi\sigma s} ds. \tag{1.19}$$

Subsituting Eq. (1.18) in Eq. (1.19) yields

$$P(\theta, \sigma) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s, t) e^{-i2\pi\sigma s} dsdt \tag{1.20}$$

which, using a coordinate transform to the (x, y) system (Eq. (1.10)), is equivalent to

$$P(\theta, \sigma) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-i2\pi\sigma(x\cos\theta + y\sin\theta)} dxdy. \tag{1.21}$$

The right-hand-side of this equation corresponds to the Fourier transform of the object function along one radial line $(\xi, \eta) = (\sigma\cos\theta, \sigma\sin\theta)$ (Eq. (1.26)):

$$P(\theta, \sigma) = F(\sigma\cos\theta, \sigma\sin\theta). \tag{1.22}$$

This expression is known as the Fourier Slice theorem. It states that the Fourier transform of a parallel projection with orientation angle θ coincides with the 2D Fourier transform of the object function along a radial line with orientation angle θ (see Fig. 1.9). Hence, each radial line of the 2D Fourier transform of the object is known by Fourier transforming the corresponding measured projection. In case of an infinite number of projections, the 2D frequency space of the object function can be filled completely. A 2D inverse Fourier transform then yields the reconstructed object function.

In practice, only a finite number of projections is available from the measurements, yielding a sampling of the frequency space as shown in Fig. 1.10. Suppose N projections $p(\theta, s)$ with equal angular spacing were measured. The following reconstruction strategy could then be applied:

1. Calculate the 1D Fourier transform of the measured projections

$$P(\theta, \sigma) = \mathcal{F}_{1\mathrm{D}} p(\theta, s) \tag{1.23}$$

   with respect to the second variable.

2. Arrange the Fourier transformed projections onto a 2D radial grid. The 2D frequency space of the object function is now radially sampled:

$$F(\sigma\cos\theta, \sigma\sin\theta) = P(\theta, \sigma). \tag{1.24}$$

3. Resample the data points to a rectangular grid (ξ, η) by interpolation.

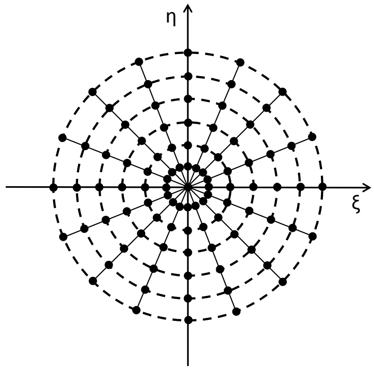


Figure 1.10: *Sampling of the Fourier space of the 2D object function.*

4. Perform a 2D inverse Fourier transform of $F(\xi, \eta)$ to recover the object function

$$f(x, y) = \mathcal{F}_{\mathrm{2D}}^{-1} F(\xi, \eta) \tag{1.25}$$

The main disadvantage of the above method is located in step 3. For higher values of σ, $F(\sigma \cos\theta, \sigma \sin\theta)$ is only coarsely sampled, which causes a regular interpolationto be less accurate for higher frequencies. Therefore, complicated interpolation procedures (cfr. gridding [8]) are necessary to avoid high-frequency artifacts.

1.3.2.2 Filtered Back Projection

A reformulation of the Fourier Slice theorem in polar coordinates yields a 2-step reconstruction method consisting of a projection filtering and a backprojection onto the image domain. Advantageous in this *Filtered Back Projection* (FBP) method, derived below, is that the difficult interpolation problem that was encountered in the Fourier Slice reconstruction method, is confined to a linear interpolation in the image domain.

The 2D inverse Fourier transform of $F(\xi, \eta)$ is

$$f(x, y) = \mathcal{F}_{\mathrm{2D}}^{-1} F(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\xi, \eta) e^{2\pi i (x\xi + y\eta)} d\xi \, d\eta. \tag{1.26}$$

Expressing this equation in a polar coordinate system (θ, σ) with

$$\xi = \sigma \cos\theta \tag{1.27}$$

$$\eta = \sigma \sin\theta \tag{1.28}$$

$$d\xi \, d\eta = \sigma \, d\sigma \, d\theta \tag{1.29}$$

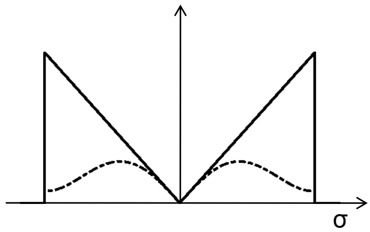


Figure 1.11: *Solid line: the Ram-Lak filter frequency response. Dashed line: resulting frequency response of the Ram-Lak filter multiplied by the Hamming function*

yields

$$f(x,y) = \int_0^{2\pi} \int_0^{\infty} F(\theta,\sigma) e^{i2\pi\sigma s} \sigma \, d\sigma \, d\theta. \tag{1.30}$$

with $s = x\cos\theta + y\sin\theta$ (Eq. (1.11)). Splitting the outer integral into $\theta \in [0,\pi)$ and $\theta \in [\pi, 2\pi)$, and using $F(\theta,\sigma) = F(\theta+\pi, -\sigma)$ yields

$$f(x,y) = \int_0^{\pi} \left[\int_{-\infty}^{\infty} F(\theta,\sigma)|\sigma| e^{i2\pi\sigma s} \, d\sigma \right] d\theta, \tag{1.31}$$

Following the Fourier Slice theorem, the 2D Fourier transform $F(\theta,\sigma)$ along the radial line with orientation θ is given by the Fourier transform of the projection $P(\theta,\sigma)$, which yields

$$f(x,y) = \int_0^{\pi} \left[\int_{-\infty}^{\infty} P(\theta,\sigma)|\sigma| e^{i2\pi\sigma s} d\sigma \right] d\theta. \tag{1.32}$$

In this integral, two operations can be distinguished: a projection filtering

$$q(\theta,s) = \int_{-\infty}^{\infty} P(\theta,\sigma)|\sigma| e^{i2\pi\sigma s} d\sigma = \mathcal{F}_{1\mathrm{D}}^{-1} \left[\mathcal{F}_{1\mathrm{D}} \left(p(\theta,s) \right) |\sigma| \right], \tag{1.33}$$

followed by the backprojection of the filtered projections

$$f(x,y) = \int_0^{\pi} q(\theta, x\cos\theta + y\sin\theta) d\theta. \tag{1.34}$$

The filtering procedure is basically a multiplication of the projections with the kernel $|\sigma|$ in Fourier space. This high-pass filter compensates for the inhomogeneous sampling density in the frequency space (Fig. 1.10). For a discrete implementation, the filter, also called the 'Ram-Lak' filter or the 'ramp' filter, is cut off at the Nyquist frequency, as shown in Fig. 1.11 (solid line). Often, a smoothing window function such as the Hamming window (dashed line in Fig. 1.11) is used that suppresses the highest spatial frequencies, to reduce the high-frequency noise and aliasing artifacts in the reconstruction image.

The reconstructed image is obtained after backprojection of the filtered projections. During the backprojection, illustrated in Fig. 1.12, the filtered projections are 'smeared out' onto the image plane. For each angle θ and each coordinate value s, the filtered projection value $q(\theta, s)$ is added to all image pixels on the line of constant s.

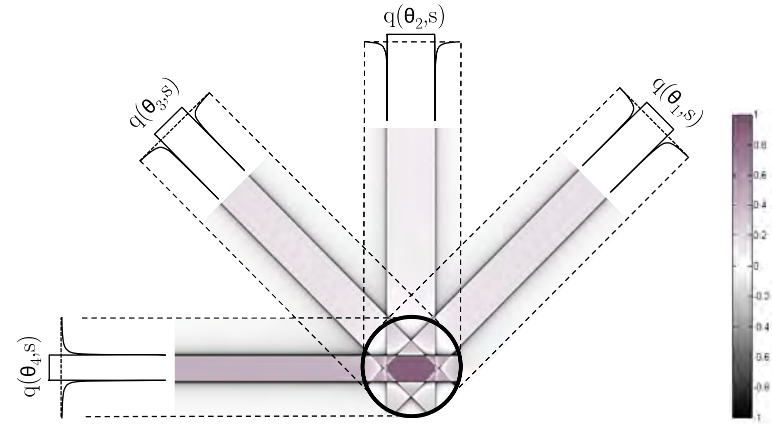


Figure 1.12: *Illustration of the backprojection procedure using 4 filtered projections $q(\theta, s)$ of an ellipse.*

In subsection 1.3.3 iterative methods are discussed. Compared to these methods, the analytical FBP reconstruction method is very fast. Mainly for this reason, the use of the FBP method and its variants such as, e.g., the Feldkamp-Davis-Kress (FDK) algorithm for cone beam (see [1]), is standard practice for tomographic image reconstruction. The major drawback of the analytical reconstruction methods is the difficulty to incorporate prior knowledge such as a positivity constraint, or a sophisticated imaging model.

1.3.3 Iterative reconstruction

An entirely different approach to image reconstruction is offered by iterative reconstruction methods. Using a discretized representation of the image space, the unknown object function is found by iteratively solving a system of linear equations, each representing one measured value as a linear combination of the unknowns.

Consider the object function $f(x, y)$ superimposed on the image grid (Fig. 1.13). The unknown function values in the 2D grid can be ordered in a one-dimensional

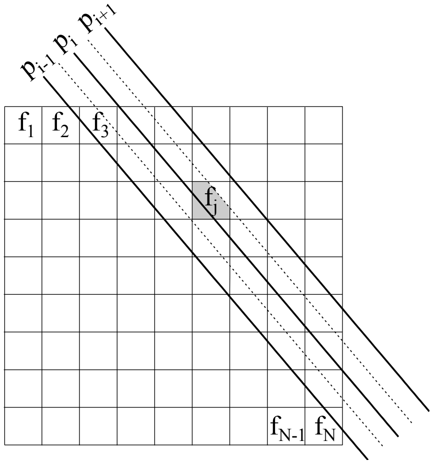


Figure 1.13: *Image grid for iterative reconstruction methods. The grey area denotes the area of pixel j that contributes to projection value p_i*

array $\mathbf{f} = (f_1, f_2, ..., f_j, f_{j+1}, ...f_N)$ with N the number of pixels on the reconstruction grid. Consider measurements $\mathbf{p} = (p_1, p_2, ..., p_i, p_{i+1}, ...p_M)$ with M the total number of projection values, i.e. the number of detector pixels multiplied by the number of projections. The system of linear equations is then given by

$$p_i = \sum_{j=1}^{N} a_{ij} f_j, \tag{1.35}$$

where element a_{ij} of the $M \times N$ system matrix A represents the contribution of pixel j to the i^{th} ray sum p_i. Many system models can be used, each yielding different noise properties in the reconstruction image; for an overview, see [9]. In the example of Fig. 1.13, the matrix element a_{ij} is represented by the intersection surface of the pixel j and the i^{th} ray strip.
Typically, the system matrix is huge (in the order of 10^{12} elements for a 1000×1000 pixel image), which turns direct matrix inversion infeasible. Alternatively, a wide range of algebraic and statistical methods exist that calculate the object function iteratively. An overview can be found in [10]. For example, the *Simultaneous Iterative Reconstruction Technique* (SIRT) iteratively computes an image update $f_j^{(k+1)}$ from a previous guess $f_j^{(k)}$ using the following recurrence equation:

$$f_j^{(k+1)} = f_j^{(k)} + \frac{\sum_i \left[a_{ij} \left(p_i - \sum_h a_{ih} f_h^{(k)} \right) / \sum_h a_{ih} \right]}{\sum_i a_{ij}}. \tag{1.36}$$

This formula basically represents a backprojection of the projection difference

$$p_i - \sum_h a_{ih} f_h^{(k)}, \tag{1.37}$$

which is the difference between the measured projection value p_i and the forward projection along the i^{th} ray. Let R be the diagonal matrix of which the diagonal elements are defined as $r_{ii} = 1/\sum_j a_{ij}$. Define the residual error E of an image $\mathbf{f}$ as

$$E(\mathbf{f}) = \|\mathbf{p} - A\mathbf{f}\|_R^2, \tag{1.38}$$

where $\|\mathbf{p} - A\mathbf{f}\|_R^2 = (A\mathbf{f} - \mathbf{p})^T R(A\mathbf{f} - \mathbf{p})$ (see, e.g., [11]). It is well known that the iterative SIRT method converges to the solution of the least squares problem

$$\mathbf{f}^{\text{final}} = \arg\min_{\mathbf{f}}(E(\mathbf{f})). \tag{1.39}$$

Note that, if the number of equations M is smaller than the number of unknowns N, the solution is not uniquely determined. In that case, SIRT converges to the solution that is the closest to the initial guess $\mathbf{f}^0$ [1], i.e.

$$\mathbf{f}^{\text{final}} = \arg\min_{\mathbf{f}} |\mathbf{f}^0 - \mathbf{f}|. \tag{1.40}$$

Compared to analytical reconstruction methods, iterative techniques have the major advantage that prior knowledge is relatively easily incorporated using penalty functions. Moreover, iterative reconstructions easily deal with various acquisition geometries, as long as the ray path corresponding to each measured value is known. In addition, in case of incomplete or irregularly sampled data, the iterative reconstruction methods are able to find the least squares solution of Eq. (1.39), as opposed to the FBP method.

However, analytical methods such as FBP strongly outperform iterative reconstruction methods when it comes to computation time. Consequently, iterative methods are found infeasible for most practical applications. Exceptions are PET (Positron Emission Tomography) and SPECT (Single Photon Emission Computed Tomography), where typically low-resolution images are reconstructed, and where the low photon count requires a statistical approach.

1.4 Artifacts in tomography

This section discusses some of the most important artifacts encountered when reconstructing CT images. An artifact can be defined as any discrepancy between the reconstructed image and the true attenuation coefficient, that affects the quantitative or qualitative analysis of the image. We consider two groups of artifacts: those caused by degraded CT data and those caused by limited CT data.

1.4.1 Artifacts caused by degraded data

Standard tomographic reconstruction techniques use the Radon transform to model the logarithm of the intensity decay when an X-ray beam traverses an object. This simplified model, however, is not adapted to address artifacts caused by scattering, beam hardening, detector deficiency, afterglow, object motion, etc. [2]. This subsection discusses two of these artifacts, namely scatter and beam hardening, which both have a potentially large impact on the image quality in both medical CT and the area of non-destructive testing. Beam hardening artifact reduction will be discussed in Part IV of this dissertation.

1.4.1.1 Scatter

CT reconstruction methods assume that the measured intensity is exclusively related to the attenuation along the incident ray path. This assumption is violated by high energy photons that are inelastically scattered (Compton scattering) in the forward direction along deflected paths (see section 1.1.2). Alternatively, a (typically low-energy) photon can be scattered elastically during an interaction with a bounded electron, which causes a change in the direction of the photon without a change in its wavelength or energy.
If a scattered photon penetrates through the entire object, it contributes to a detector pixel that is not related to the original ray path. As a result, scattering induces an additional slowly varying object-dependent background intensity onto the intensity signal of the primary photons, which is illustrated in Fig. 1.14(a). Scattering has a larger impact on lower intensity measures, which correspond to rays through highly absorbing materials. The resulting reconstructions suffer from a reduced contrast and signal to noise ratio, and from shading between highly attenuating structures. Fig. 1.14(b) shows a phantom (upper) and its reconstruction (lower) from data contaminated by scattering. The scattering in the latter image is simulated by adding a constant value to the intensity sinogram $I(\theta, s)$. Scattering artifacts can be reduced by placing a collimator or *scatter grid* before detecting the signal, to prevent the detection of scattered photons with a large deflection angle.

1.4.1.2 Beam hardening

When a monochromatic X-ray beam traverses a homogeneous object, the attenuation is linearly related to the thickness of the object along that ray (Beers law). In general, however, CT X-ray sources are polychromatic (see Eq. (1.9)). Since the low energy (soft) X-rays are more easily absorbed than high energy (hard) X-rays, the beam hardens when propagating through matter. Consequently, the attenuation-thickness relation becomes a non-linear curve, deflecting from the expected linear line. This is illustrated for a homogeneous material in Fig. 1.15(a),

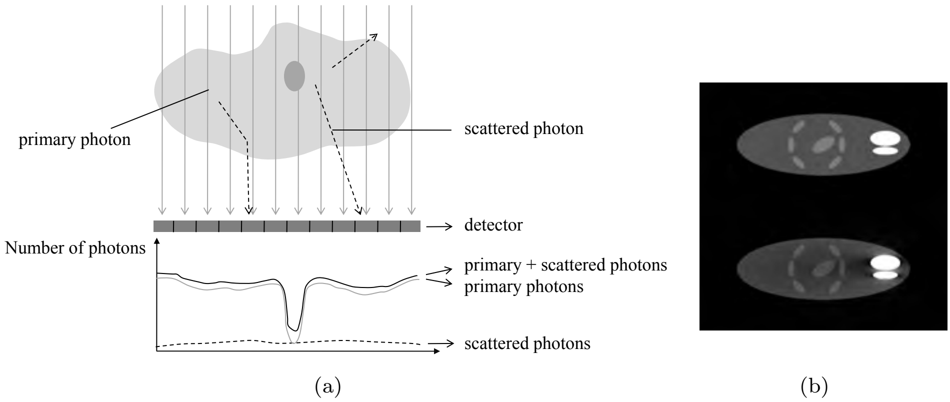
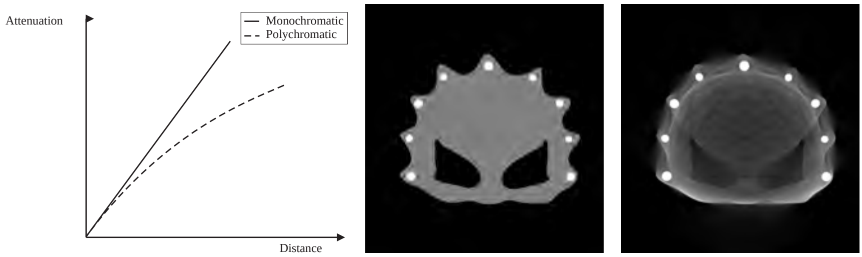


Figure 1.14: *(a) Simulated effect of scattering on the measured intensity values (picture based on a picture of Hsieh [2]). (b) the lower image represents a reconstruction of the phantom that is shown in the upper image, from data contaminated by simulated scattering: a first order simulation is performed by adding a constant value to the intensity sinogram $I(\theta, s)$. Notice the shadowing artifacts close to the highly absorbing structures, and the decreased contrast between the structures.*

where the shaded line, representing the nonlinear beam hardening curve, deflects from the monochromatic straight line (solid). If the energy dependence of the absorption is not taken into account, reconstructions are contaminated by cupping and streak artifacts [1], as shown in Fig. 1.15(c), which depicts a reconstruction of the phantom in (b) from polychromatic data. In Chapter 6, the beam hardening problem is described in more detail, a literature overview is given and a new beam hardening reduction method is proposed for piecewise uniform objects.

1.4.2 Artifacts caused by incomplete projection data

The artifacts discussed above are inherent to the nature of X-rays. Other artifacts are encountered in case of incomplete data. For example in a medical setup, patients do not always fit in the FOV of the detector, and consequently, line integrals are missing in some directions. In transmission electron tomography, which is a technique that uses a transmission electron microscope and an electron beam to measure projections at a resolution of $5 - 20$ nm, practical limitations restrict the tilt angle θ of the sample holder typically to a range of 120 degrees.
To determine whether or not to classify an inverse problem as incomplete or insufficient, the well-posedness of the problem needs to be analysed. A problem of finding an object function f given a Radon transform g so that $Ag = f$ is called "well-posed" [12] if it has a unique solution and if the solution continuously depends on the input data.



(a) (b) (c)

Figure 1.15: *(a) Attenuation-thickness relation in case of a monochromatic beam (solid line) and a polychromatic beam (dashed line). (b) the Mega Mindy phantom. (c) Filtered Backprojection reconstruction of the Mega Mindy phantom from data with simulated beam hardening using the model proposed in Chapter 6. Notice the large cupping artifact, i.e. the decrease of the grey value towards the center, and the streaks between the highly attenuating structures.*

We consider three types of incomplete data problems :

- Small number of projections. Here, projections are known for only a small set of angles θ distributed over $[0, \pi)$. In this case, the solution is not unique and the problem is severely ill-posed. Note that the problem of reconstructing an object from a large but finite number of projections is also essentially underdetermined (see Theorem II.3.7 in [5]), but for well behaved functions and a sufficient number of projections, this indeterminacy is dissolved (see Theorem VI.2.2 in [5]).

- Limited angular range. Here, the projections are only given for a limited angular interval $\theta \in [0, \Theta]$, where $\Theta < \pi$. This problem is uniquely determined, but reconstruction is highly unstable (see Chapter 6 in [5]).

- *Interior problem.* In this special category of truncation problems, only the set of line integrals through a *region of interest* (ROI) that is embedded in the interior of the object, is measured. This is the case when the Radon transform $p(\theta, s)$ is only known for $|s| \leq w$ with w the radius of a centered circle in the interior of the object.
  It is well-known that the desired ROI of the object function is not uniquely determined by these data, which means that one cannot exactly reconstruct the attenuation map of the object function even when assuming that the data is given over a continuous set of lines without measurement errors (see also Section 2.1).

In Fig. 1.16, FBP reconstructions of the Shepp-Logan phantom [2] are shown for each of the discussed limited data problems. Fig. 1.16(a) represents the FBP reconstruction from a (complete) sinogram, consisting of 360 views equally sampled in $[0, \pi)$ and 512 radial samples per view. Fig. 1.16b illustrates the FBP reconstruction from a sinogram containing only 18 (equally sampled in $[0, \pi/2)$ views; (c) is reconstructed from a Radon transform with limited angular range $\theta \in [0, \pi/2]$; (d) represents the reconstruction from interior data with $|s| \leq 128$.

The uniqueness or well-posedness of the reconstruction problem from limited data can be restored, provided adequate prior knowledge is included. For example in digital subtraction angiography, where blood vessels are imaged by using contrast agents, one can assume that the images are sparse, i.e. that they contain only a small number of non-zero pixels.
Reconstruction methods for limited data problems can be subdivided into two classes: iterative reconstruction methods and data completion for analytic reconstruction. Iterative methods have the advantage that prior knowledge is relatively easily incorporated in the algorithm by adding a penalty function. For example, for the blood-vessel example, an ℓ_1-norm penalty function is often added to the cost function in Eq. (1.39):

$$\mathbf{f}^{\text{final}} = \arg\min_{\mathbf{f}} \left\{ \|\mathbf{p} - A\mathbf{f}\|_R^2 + \|\mathbf{f}\|^1 \right\}, \tag{1.41}$$

with $\|\mathbf{f}\|^1 = \sum_h |f_h|$, which typically results in sparse images. The drawback of iterative methods is, however, their reconstruction speed. Analytical methods, on the other hand, are much faster, but have large difficulties with incorporating prior knowledge. Therefore, the data are often completed in a preprocessing step by using prior knowledge such as the object support, consistency conditions, previous scans, etc.
Alternatively, for the limited angular range and the truncation problem, recent results have shown that some parts of the object can accurately be reconstructed using only knowledge of the object support. For a more thorough overview of correction methods for the truncation problem, see Chapter 2.

[2]The Shepp-Logan phantom, proposed by L.A. Shepp and B.F. Logan [13] represents a simulation of a human head section, and is standardly used for the comparison of the quality of reconstruction techniques.

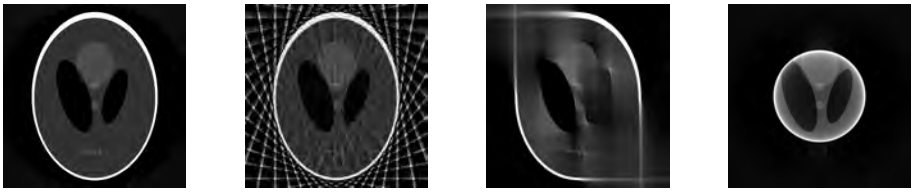


Figure 1.16: *(a) Shepp-Logan (SL) phantom [13]. (b) Reconstruction of the SL-phantom from 18 projections. (c) Reconstruction of the SL-phantom from a limited angular range ($[0, \pi/2)$. (d) Reconstruction of the SL-phantom from interior data.*

Bibliography

[1] A. Kak and M. Slaney, *Principles of computerized tomographic imaging*, SIAM, 2001.

[2] J. Hsieh, *Computed tomography, principles, design, artifacts and recent advances*, SPIE Press, 2003.

[3] W. A. Kalender, *Computed Tomography*, Publicis, Erlangen, 2005.

[4] T. M. Buzug, *Computed Tomography*, Springer, 2008.

[5] F. Natterer, *The mathematics of computerized tomography*, SIAM, 2001.

[6] F. Natterer and F. Wübbeling, *Mathematical methods in image reconstruction*, Monographs on Mathematical Modeling and Computation. SIAM, 2001.

[7] "`http://www.worldsfamousphotos.com`," .

[8] J.D. O' Sullivan, "A fast sinc function gridding algorithm for fourier inversion in Computer Tomography," *IEEE Transactions on Medical Imaging*, vol. 4, no. 4, pp. 200–207, 1985.

[9] F. Xu and K. Mueller, "A comparative study of popular interpolation and integration methods for use in computed tomography," in *Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on*, April 2006, pp. 1252–1255.

[10] Fessler, J. A., "Statistical Methods for Image Reconstruction," `http://www.eecs.umich.edu/~fessler/papers/files/talk/06/isbi,p1,note.pdf`, 2004.

[11] J. Gregor and T. Benson, "Computational analysis and improvement of SIRT," *IEEE Transactions on Medical Imaging*, vol. 27, no. 7, pp. 918–924, 2008.

[12] J. Hadamard, *Le problème de Cauchy et les equations aux derivées partielles linéaires hyperboliques*, Herman, Paris, 1932.

[13] L.A. Shepp and B.F. Logan, "Fourier reconstruction of a head section," *IEEE Transactions on Nuclear Science*, vol. 21, no. 3, pp. 21–43, 1974.

Part II

Reduction of truncation artifacts in standard 2D X-ray CT

Chapter 2

Region-of-interest tomography: state of the art

2.1 Introduction

Region of interest (ROI) tomography considers the problem of reconstructing a certain area in the object from a set of truncated projections, i.e. data that consists solely of the integrals along the complete set of lines that intersect this area. A schematic example of ROI tomography for a parallel beam geometry is shown in Fig. 2.1. In this example, the circular field of view (FOV) is covered by the detector in all directions; each point of the remaining area in the object support is not covered by the detector in at least one projection direction.

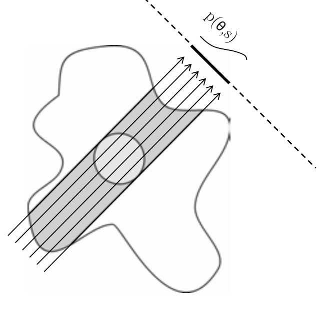
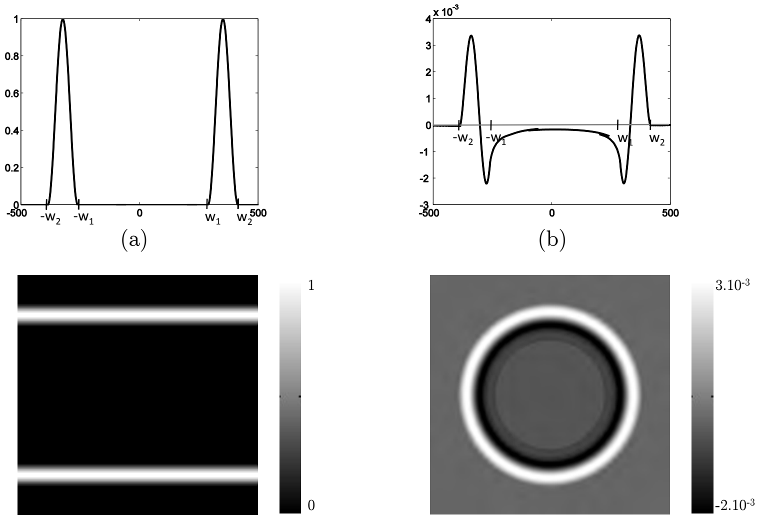


Figure 2.1: *Illustration of a truncation problem. The circle represents the field of view (FOV), covered by the detector in each projection direction.*

Since the attenuation integrals are measured along all lines through the FOV,



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

Figure 2.2: *Illustration of the non-uniqueness of the solution from interior data. (c) sinogram composed of identical 1D projections $h(s)$ (shown in (a)). (d) unique solution from sinogram (c); a radial cross section is plotted in (b).*

intuitively one expects that the FOV can be completely recovered. Note, however, that the measured projection values are contaminated by the attenuation of the surrounding object region, sometimes with important consequences. For example, in case of an interior problem, i.e. if the FOV is completely embedded in the object support, this contamination destroys the uniqueness of the solution. This non-uniqueness is illustrated in Fig. 2.2. Fig. 2.2(a) shows an even one-dimensional infinitely differentiable function $h(s)$ that is zero for all $|s| \notin (w_1, w_2)$. It is proven (see Section VI.4 in Natterer [1]) that the 2D function $q(\theta, s)$, constructed by $q(\theta, s) = h(s)$ for all $\theta \in [0, 2\pi)$, represents a consistent sinogram, and has a unique solution, which is depicted in Fig. 2.2 (d). A plot of the reconstruction value along a radial line (red) through the reconstructed image, shows that the solution is non-zero in the centered circle with radius w_1. Note, however, that the truncated sinogram, acquired by measuring $q(\theta, s)$ only for $|s| < w_1$, is zero everywhere. Hence, the truncated sinogram also represents the truncated Radon transform of a second function: $f(x, y) = 0$ for $(x, y) \in \mathbb{R}^2$. This illustrates the non-uniqueness of the solution from interior Radon data.

After a short introduction to data truncation in Section 2.2, Section 2.3 considers the use of FBP for the reconstruction of an image from truncated data, and discusses the problems associated with this approach. In Section 2.4, a recently obtained alternative formulation for the Radon inverse is described, and it is shown how this new formulation partly alleviates these non-locality problems. Section 2.5 consists of recent results describing the area in which the solution to the reconstruction problem is uniquely determined, for various types of truncation problems. Also, reconstruction methods to obtain this unique solution are mentioned. An overview of empirical truncation artifact reduction methods is provided in Section 2.6.

2.2 Data truncation

A schematic example of a scanning acquisition is shown in Fig. 2.3. In this figure, the dashed line describes the object support Ω; the FOV is assumed to be circular. Define r as the minimal radius of the FOV that is necessary to cover the complete object in all projection directions. The dataset is called truncated if the FOV has a radius w with $w < r$. The corresponding truncated sinogram $(\mathcal{R}_w f)(\theta, s)$ is then measured only for $s \in [-w, w]$.
Two disjoint regions $A \subset \Omega$ and $B = \Omega \backslash A$ in the object support Ω are distinguished (see Fig. 2.3). The region A (light grey) describes the intersection of the object support with the FOV. The region B (dark grey) is the complement of region A with respect to the object support. For all points in B, line integrals are missing at least in some directions. This corresponds to a limited angle problem and the measured angular range becomes smaller for points at larger distance from region A. Therefore it can be expected that no part of region B can be recovered accurately [2].

2.3 FBP reconstruction from truncated projections

For a long period of time, the mainstream idea was that accurate and stable reconstruction of any region of interest in the object requires a complete set of line integrals of the whole object. This idea partly stems from the non-uniqueness result for the interior problem, and partly from the formulation of the standard FBP reconstruction method (see Eq. (1.34) and Eq. (1.33)):

$$f(x, y) = \int_0^\pi q(\theta, x\cos\theta + y\sin\theta)d\theta. \tag{2.1}$$

with

$$q(\theta, s) = \int_{-\infty}^{\infty} P(\theta, \sigma)|\sigma|e^{i2\pi\sigma s}d\sigma = \mathcal{F}_{\mathrm{1D}}^{-1}(P(\theta, \sigma)|\sigma|). \tag{2.2}$$

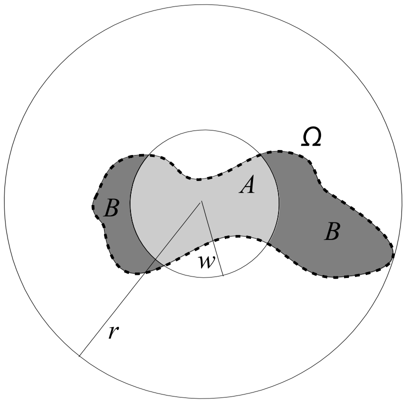
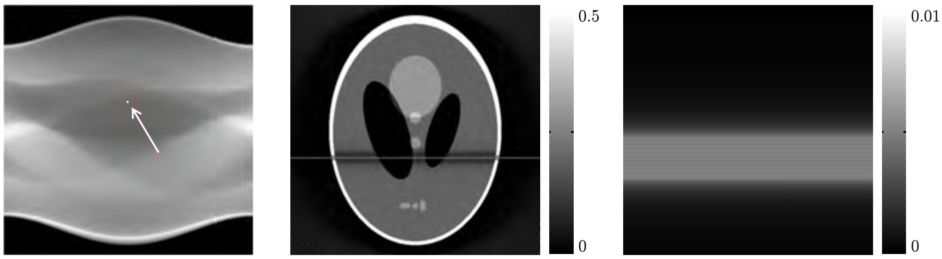


Figure 2.3: *Schematic example of a truncated scanning acquisition. The dashed line represents the object support Ω; the circle with radius w represents the field of view (FOV) of the detector in case of truncation. The dark grey area reflects region B, the lighter grey area denotes region A. The large circle with radius r represents the FOV for a non-truncated case.*

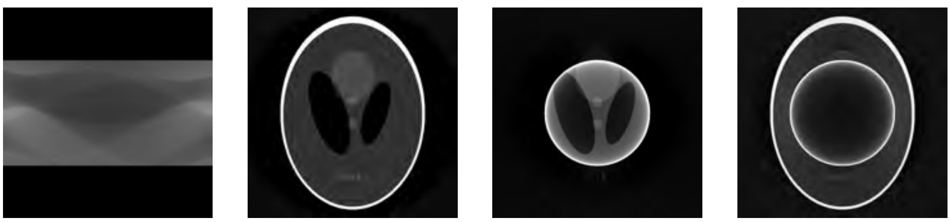
This FBP formula shows that a local measurement distortion in a single projection contaminates the complete projection after filtering. During the subsequent back-projection (Eq. (1.34)), every filtered projection contributes to the reconstruction of a single point (x, y). Hence, even a tiny data contamination prohibits accurate FBP recovery of the object function $f(x, y)$ in any point (x, y) of the object support. This is illustrated in Fig. 2.4 for the reconstruction of the Shepp-Logan phantom from a complete sinogram in which a single pixel is severely distorted; the difference image is everywhere non-zero. Note, however, that the distortion mainly affects the neighbourhood of the line along which the distortion is backprojected. At larger distance from this line, the reconstruction quality improves.

Hence, the FBP reconstruction from a truncated sinogram (Fig. 2.5 (a)), in which basically the unknown sinogram values are replaced by zeros , is never exact. It typically suffers from cupping artifacts as observed in Fig. 2.5(c). The FBP ramp filtering of projections with an abrupt transition between measured values and zeros, caused by the truncation, induces large artifacts at the boundary of the FOV, where a bright rim is noticed. The cupping artifact decreases towards the center of the FOV, but the bias is nowhere zero. For this reason, quantitative analysis is prohibited, but structural information can still be recovered in the FOV. Note also that the FBP from truncated projections offers only severely distorted density and structural information in the region surrounding the FOV.



(a) (b) (c)

Figure 2.4: *Illustration of the non-local behaviour of the FBP algorithm. (a) Sinogram of Shepp-Logan phantom (see Fig. 1.16(a)) with one distorted pixel (indicated with white arrow). (b) reconstruction from the distorted sinogam shown in (a). (c) absolute value of the difference image with respect to the true Shepp-Logan phantom. This figure illustrates how the complete 2D reconstructed image is contaminated by a one-pixel distortion in the sinogram.*



(a) Truncated sinogram

(b) FBP reconstruction from a complete sinogram

(c) FBP reconstruction from a truncated sinogram (a)

(d) Absolute value of (b)-(c)

Figure 2.5: *Illustration of the FBP reconstruction from truncated data. (a) truncated sinogram; (b) reconstruction of a ROI from complete data; (c) reconstruction of the same ROI from sinogram (a); (d) difference image*

2.4 DBP reconstruction from truncated projections

For decades, FBP and its variants were the only closed-form analytic expressions to recover a 2D object function from its parallel projections. In 2004, a new analytical reconstruction method, relating the differentiated backprojection of the projections to the Hilbert transform of the object function, was proposed for a parallel beam geometry by Noo et al. [3], for a fan beam geometry by Noo et al. [3] and Zou et al. [4], and by Zhuang et al. [5] for a cone beam geometry. The 2-step inverse Hilbert method of Noo et al. [3] directly leads to new insights and sufficiency conditions for 2D parallel beam tomography. We refer to this method as the differentiated backprojection method (DBP). In this section, the Hilbert transform and the DBP method are introduced, and it is shortly discussed how the DBP method partially circumvents the problem of non-locality encountered by the FBP method.

2.4.1 Hilbert transform

The Hilbert transform of a one-dimensional function $f(t)$ is defined as its convolution with kernel $\frac{1}{\pi t}$:

$$\mathcal{H}f(t) = p.v. \int_{-\infty}^{\infty} \frac{f(t-t')}{\pi t'} dt'. \tag{2.3}$$

where "p.v." denotes that the singularity at $t' = 0$ is handled in the Cauchy principal value sense (see, e.g, p. 191 in [6]). In the derivation below, the notation p.v. is omitted. The Fourier representation of the Hilbert transform is given by:

$$\mathcal{H}f(t) = \int_{-\infty}^{\infty} -i\,\mathrm{sgn}(\tau) F(\tau) e^{i2\pi t\tau} d\tau, \tag{2.4}$$

where $F(\tau) = \mathcal{F}_{1\mathrm{D}} f(t) = \int_{-\infty}^{\infty} f(t) e^{-i2\pi\tau t} dt$ denotes the Fourier transform of function f, and where $-i\,\mathrm{sgn}(\tau)$ is the well known Fourier response of the Hilbert kernel. Using the Hilbert transform, the expression of the FBP filtering in Eq. (1.33) can be rewritten as

$$q(\theta, s) = \int_{-\infty}^{\infty} -i\,\mathrm{sgn}(\sigma) \left[P(\theta,\sigma) i\sigma\right] e^{i2\pi\sigma s} d\sigma = \frac{1}{2\pi} \mathcal{H} \frac{\partial p(\theta, s)}{\partial s}, \tag{2.5}$$

since multiplication with $2\pi i\sigma$ in the Fourier domain corresponds to differentiation in real space.
Now consider a 2-dimensional function $f(\mathbf{x})$ with $\mathbf{x} = (x, y)$. The one-dimensional Hilbert transform of f along the x-axis, denoted as $\mathcal{H}_0 f$, is given by

$$\mathcal{H}_0 f(\mathbf{x}) = \int_{-\infty}^{\infty} \frac{f(x-t', y)}{\pi t'} dt' = \int_{-\infty}^{\infty} \frac{f(t', y)}{\pi(x-t')} dt'. \tag{2.6}$$

Let $F(\boldsymbol{\xi}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\mathbf{x}) e^{-i2\pi\boldsymbol{\xi}\cdot\mathbf{x}} d\xi d\eta$ represent the 2D Fourier transform of $f(\mathbf{x})$, with $\boldsymbol{\xi} = (\xi, \eta)$. The Fourier representation of Eq. (2.6) is then:

$$\mathcal{H}_0 f(\mathbf{x}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} -i\,\mathrm{sgn}(\xi) F(\boldsymbol{\xi}) e^{i2\pi\boldsymbol{\xi}\cdot\mathbf{x}} d\xi d\eta. \tag{2.7}$$

More generally, the one-dimensional Hilbert transform $H_\phi(f)$ [1] along lines with direction $\boldsymbol{\beta} = (\cos\phi, \sin\phi)$ and a fixed angle ϕ with respect to the x-axis is given by

$$\mathcal{H}_\phi f(\mathbf{x}) = \int_{-\infty}^{\infty} \frac{f(\mathbf{x} - t'\boldsymbol{\beta})}{\pi t'} dt', \tag{2.8}$$

[1] Notice the difference with the definition of Noo et al., where $\mathcal{H}_\phi$ denotes the Hilbert transform along lines with the perpendicular direction $(-\sin\phi, \cos\phi)$.

where $\boldsymbol{\beta} = (\cos\phi, \sin\phi)$.
The corresponding Fourier representation of this formula is written

$$\mathcal{H}_\phi f(\mathbf{x}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} -i\,\mathrm{sgn}(\boldsymbol{\xi}\cdot\boldsymbol{\beta})F(\boldsymbol{\xi})e^{i2\pi\boldsymbol{\xi}\cdot\mathbf{x}}d\xi d\eta. \tag{2.9}$$

Consider the coordinate system (s, t), rotated with respect to the system (x, y), such that $s = \mathbf{x}\cdot\boldsymbol{\beta}^\perp$, $t = \mathbf{x}\cdot\boldsymbol{\beta}$, $\boldsymbol{\beta}^\perp = (\sin\phi, -\cos\phi)$, and $\mathbf{x} = s\boldsymbol{\beta}^\perp + t\boldsymbol{\beta}$. The Hilbert transform along angle ϕ can then be rewritten as

$$\mathcal{H}_\phi f(s\boldsymbol{\beta}^\perp + t\boldsymbol{\beta}) = \int_{-\infty}^{\infty}\frac{f(s\boldsymbol{\beta}^\perp + (t-t')\boldsymbol{\beta})}{\pi t'}dt' = \int_{-\infty}^{\infty}\frac{f(s\boldsymbol{\beta}^\perp + t'\boldsymbol{\beta})}{\pi(t-t')}dt'. \tag{2.10}$$

The following characteristics of the Hilbert transform are relevant for the remainder of the derivation:

- Convolution with the Hilbert kernel $1/(\pi t)$ of a function with compact support results in a function of unbounded support. This is particularly true for object functions, which are positive within their support, and zero elsewhere. The 1D Hilbert transform of $f(\mathbf{x})$ has therefore infinite extent in the direction of the transform ϕ. In the perpendicular direction, the extent is unchanged and thus finite. As an illustration, the Hilbert transform of a uniform ellipse along vertical lines is shown in Fig. 2.6.

- From Eq. (2.4) it is seen that

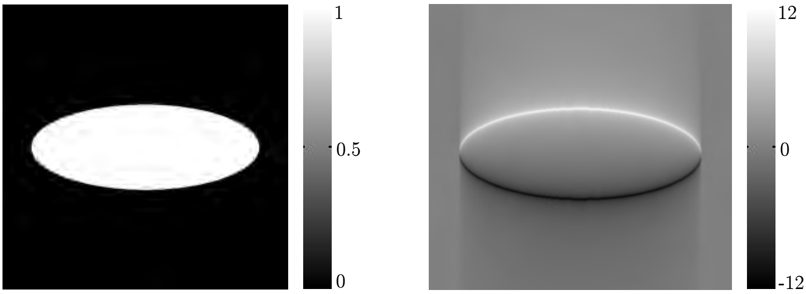
$$\mathcal{H}_\phi^2 f = -f. \tag{2.11}$$

  Consequently, by switching the sign, Eq. (2.3)-(2.10) can also be used for the inverse Hilbert transform.

2.4.2 Differentiated Backprojection and the Hilbert transform

In this section, we essentially follow the derivation in the paper of Noo et al. [3] to show the relation between the differentiated backprojection and the 1D Hilbert transform of the object function.
Consider a 2D image $g_\phi(\mathbf{x})$ that is the result of differentiating the projections $p(\theta, s)$ along direction s, followed by a backprojection, with $\theta \in [\phi - \frac{\pi}{2}, \phi + \frac{\pi}{2}]$, i.e.

$$g_\phi(\mathbf{x}) = -\frac{1}{2}\int_{\phi-\frac{\pi}{2}}^{\phi+\frac{\pi}{2}}\frac{\partial p(\theta, s)}{\partial s}d\theta, \tag{2.12}$$



(a) (b)

Figure 2.6: *Illustration of 1D Hilbert transform of a 2D object function. (a) uniform ellipse phantom. (b) Hilbert transform of the ellipse along the vertical direction.*

where $s = \mathbf{x} \cdot \boldsymbol{\alpha}$ and $\boldsymbol{\alpha} = (\cos\theta, \sin\theta)$. For now, ϕ denotes a constant angle. The choice of ϕ is important and will be discussed in Section 2.4.5. Recall the Radon transform property that $p(\theta, s) = p(\theta+\pi, -s)$, and thus that $\frac{\partial p(\theta,s)}{\partial s} = -\frac{\partial p(\theta+\pi,-s)}{\partial s}$. Eq. (2.12) can then be rewritten as:

$$g_\phi(\mathbf{x}) = -\frac{1}{2}\int_0^\pi \operatorname{sgn}(\cos(\theta-\phi))\frac{\partial p(\theta,s)}{\partial s}d\theta. \tag{2.13}$$

We will show that the object function is related to image g_ϕ through the Hilbert transform in direction ϕ.
The Fourier representation of the differentiated projection $p'(\theta, s)$ is

$$\frac{\partial p(\theta,s)}{\partial s} = 2\pi\int_{-\infty}^{\infty} i\sigma P(\theta,\sigma)e^{-i2\pi s\sigma}d\sigma, \tag{2.14}$$

where $P(\theta, \sigma) = \mathcal{F}_{1\text{D}}p(\theta, s)$ is the 1D Fourier transform of $p(\theta, s)$ with respect to the second variable. Substituting Eq. (2.14) in Eq. (2.13) yields

$$\begin{aligned} g_\phi(\mathbf{x}) &= -\pi\int_0^\pi\int_{-\infty}^{\infty} i\sigma \operatorname{sgn}(\cos(\theta-\phi))P(\theta,\sigma)e^{-i2\pi\sigma(\mathbf{x}\cdot\boldsymbol{\alpha})}d\sigma d\theta && (2.15)\\ &= -\pi\int_0^\pi\int_{-\infty}^{\infty} i \operatorname{sgn}(\cos(\theta-\phi)\sigma)P(\theta,\sigma)e^{-i2\pi\sigma(\mathbf{x}\cdot\boldsymbol{\alpha})}|\sigma|d\sigma d\theta. && (2.16)\end{aligned}$$

Note that $\cos(\theta - \phi) = \boldsymbol{\alpha}\cdot\boldsymbol{\beta}$.
In the next step, the Fourier Slice theorem $P(\theta, \sigma) = F(\sigma\boldsymbol{\alpha})$ is applied, which relates the 1D Fourier transform of the projections to the 2D Fourier transform of the object function (see Eq. (1.22)):

$$g_\phi(\mathbf{x}) = -\pi\int_0^\pi\int_{-\infty}^{\infty} i \operatorname{sgn}(\sigma\boldsymbol{\alpha}\cdot\boldsymbol{\beta})F(\sigma\boldsymbol{\alpha})e^{-i2\pi\sigma(\mathbf{x}\cdot\boldsymbol{\alpha})}|\sigma|d\sigma d\theta. \tag{2.17}$$

Finally, the polar coordinates are transformed into cartesian coordinates $\boldsymbol{\xi} = (\xi, \eta) = \sigma\boldsymbol{\alpha}$, yielding

$$g_\phi(\mathbf{x}) = -\pi \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i \,\mathrm{sgn}(\boldsymbol{\xi} \cdot \boldsymbol{\beta}) F(\xi) e^{-i2\pi(\boldsymbol{\xi}\cdot\mathbf{x})} d\xi d\eta. \tag{2.18}$$

The comparison of Eq. (2.18) with Eq. (2.9) shows that the differentiated backprojection from the measured projection data corresponds to the Hilbert image $\mathcal{H}_\phi f(\mathbf{x})$ scaled by π:

$$g_\phi(\mathbf{x}) = \pi \mathcal{H}_\phi f(\mathbf{x}). \tag{2.19}$$

Note that the Hilbert transform angle ϕ can be selected arbitrarily by rearranging the projections $p(\theta, s)$ using the Radon transform symmetry ($p(\theta, s) = p(\theta+\pi, -s)$) such that $\theta \in [\phi - \pi/2, \phi + \theta/2)$ (see Eq. (2.12)).

Fig. 2.7 illustrates the differentiated backprojection procedure. Fig. 2.7(a) depicts the used phantom, and (b) represents its corresponding Radon data. Fig. Fig. 2.7(c) shows the differentiated sinogram for a preselected Hilbert transform angle ϕ. The Hilbert image that is obtained after backprojection of this differentiated sinogram, is depicted in (d).

2.4.3 Finite Hilbert inverse

Following Eq. (2.19), the object function can be recovered from the differentiated backprojection $g_\phi(\mathbf{x})$ by a 1D inverse Hilbert transform along parallel lines with direction $\boldsymbol{\beta}$ for fixed s. Along each line $\mathcal{L}$ with direction $\boldsymbol{\beta}$, this problem is equivalent with recovering a 1D function $u(t)$ from $\mathcal{H}_\phi u(t)$, where $\mathcal{H}_\phi u(t)$ corresponds to the slice of $g_\phi(\mathbf{x})$ along the line $\mathcal{L}$.
Recovering $u(t)$ using $u(t) = \mathcal{H}_\phi^{-1}(\mathcal{H}_\phi u(t)) = -\mathcal{H}_\phi^2 u(t)$ (see Eq. (2.11)), requires $\mathcal{H}_\phi u(t)$ to be known in its complete support. This condition is not met, since $\mathcal{H}_\phi u(t)$ has infinite extent and is only known from the measurements in a finite interval $t \in [L, U]$ (see Fig. 2.8), where L and U are determined by the intersection of the ray path with the FOV.

An alternative formula, called *finite inverse Hilbert transform* (Tricomi [7]), requires $\mathcal{H}_\phi u(t)$ only to be known in a finite interval that covers the support of $u(t)$ with an excess of $\epsilon > 0$ at each side. This formula basically represents a weighted version of the Hilbert transform of $\mathcal{H}_\phi u(t)$. Suppose $u(t)$ is zero outside some interval $[L + \epsilon, U - \epsilon]$ and $\mathcal{H}_\phi u(t)$ is known in interval $[L, U]$. Then, for all $t \in [L + \epsilon, U - \epsilon]$, $u(t)$ can be recovered using:

$$u(t) = -\frac{1}{\pi\sqrt{(t-L)(U-t)}} \left(\int_L^U \sqrt{(t'-L)(U-t')} \frac{\mathcal{H}u(t')}{(t-t')} dt' + C \right), \tag{2.20}$$

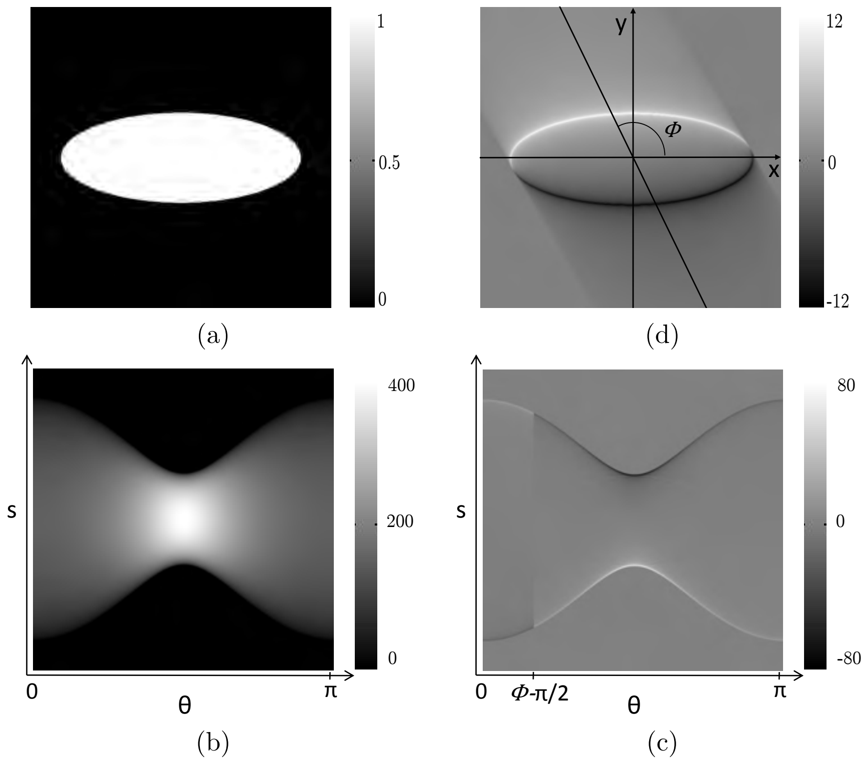


Figure 2.7: *Illustration of the differentiated backprojection. (a) uniform ellipse phantom. (b) attenuation sinogram of the ellipse phantom in (a). (c) shows the differentiated sinogram for a certain preselected Hilbert transform angle ϕ. The Hilbert image that is obtained after backprojection of this differentiated sinogram, is depicted in (d).*

with

$$C = \int_{-\infty}^{\infty} u(t)dt. \tag{2.21}$$

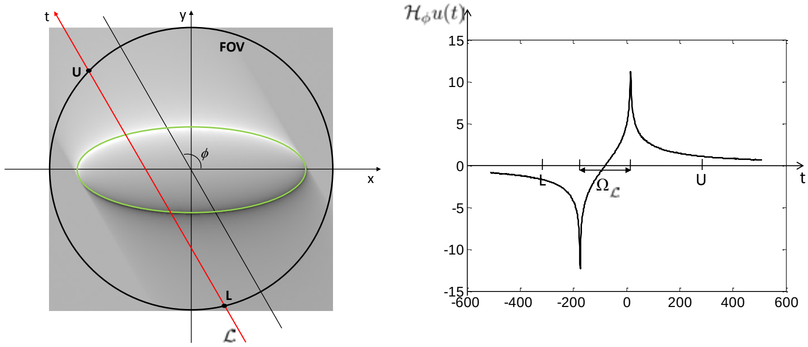
The 2D reconstruction of the object function is thus obtained by an inverse Hilbert transform using Eq. (2.20) along separate lines with direction $\boldsymbol{\beta}$:

$$f(\mathbf{x}) \quad = \quad f(s\boldsymbol{\beta}^{\perp} + t\boldsymbol{\beta}) = -\frac{1}{\pi\sqrt{(t-L_s)(U_s-t)}} \tag{2.22}$$

$$\times \left(\int_{L_s}^{U_s} \sqrt{(t'-L_s)(U_s-t')} \frac{\mathcal{H}f(s\boldsymbol{\beta}^{\perp} + t'\boldsymbol{\beta})}{(t-t')} dt' + C_s \right) \tag{2.23}$$

where $t = \mathbf{x} \cdot \boldsymbol{\beta}$ and $s = \mathbf{x} \cdot \boldsymbol{\beta}^{\perp}$. Since the required line integrals in Eq. (2.20) are known from the projection data, the constant C_s is given by

$$C_s = p(\phi, s). \tag{2.24}$$



(a) (b)

Figure 2.8: *(a) Differentiated backprojection $g_\phi(\mathbf{x})$ from a complete dataset, such that the complete object support (green ellipse) is covered by the FOV. $\mathcal{H}_\phi f(\mathbf{x})$ is known in the complete FOV. (b) plot of $\mathcal{H}_\phi u(t)$, which is a one dimensional section of $\mathcal{H}_\phi f(\mathbf{x})$ along the line $\mathcal{L}$, which has direction β. Along this line, $\mathcal{H}_\phi u(t)$ is known from the measurements in the interval $t \in [L, U]$.*

2.4.4 DBP reconstruction method: summary

The DBP reconstruction technique is initialized by selecting the direction ϕ. This can be done by rearranging the projections using the Radon symmetry ($p(\theta, s) = p(\theta + \pi, -s)$) such that the angular range is situated within the interval $\theta \in [\phi - \pi/2, \phi + \pi/2]$. The importance of a good selection of ϕ is discussed in Section 2.4.5. The DBP algorithm then basically consists of two steps:

1. Differentiated backprojection: the projections are first differentiated and then backprojected using Eq. (2.13), yielding an image $g_\phi(\mathbf{x}) = \pi\mathcal{H}_\phi f(\mathbf{x})$.

2. Inverse Hilbert transform: perform a finite inverse Hilbert transform using Eq. (2.23) and Eq. (2.24).

In a practical implementation, it is convenient to align the direction ϕ to the orientation of the pixels, whenever possible. The singularity in Eq. (2.23) is circumvented by sampling the differentiated backprojection $\mathcal{H}f$ in step 1 at half-pixel shifted samples.

2.4.5 DBP and truncation

The DBP method described above represents an alternative for the standard FBP method. A major advantage of the DBP method is that the DBP formulation is less restrictive in the amount of data that is needed to reconstruct the object function in a single point (x, y).

The calculation of $\left.\frac{\partial p(\theta,s)}{\partial s}\right|_{s=s_0}$ requires $p(\theta, s)$ to be known in a small neighborhood surrounding s_0, i.e. for $s \in [s_0 - \epsilon, s_0 + \epsilon]$ with ϵ infinitesimally small. Consequently, the differentiated backprojection $g_\phi(\mathbf{x})$ of a truncated sinogram $\mathcal{R}_a f(\theta, s)$ can be calculated in the complete FOV, except for a strip of width ϵ along its boundary.
The inverse Hilbert transform $\mathcal{H}_\phi^{-1}$, on the other hand, is a non-local 1D operation along the line $s = \mathbf{x} \cdot \boldsymbol{\beta}^\perp$ for constant s. The inversion formula 2.23 requires the DBP to be known in an open interval embedding the object support along that line. If this condition is met, the object function can be reconstructed along the whole line. Otherwise, the reconstruction is not exact along the complete line. Therefore, if a certain area in the sinogram is distorted, it is important to choose ϕ in such a way that the distortion affects the DBP $g_\phi(z)$ in a minimal number of lines with direction $\boldsymbol{\beta}$. In case of truncated projections, ϕ should be chosen such that the interval $[L_s, U_s]$ covers the object support for as many lines with direction $\boldsymbol{\beta}$ as possible. After all, along these lines, the object function can accurately be recovered. A new data sufficiency condition for truncated projections can therefore be derived, which is discussed in Section 2.5.

As an example, Fig. 2.9 depicts the DBP reconstruction from a complete sinogram in which a single pixel is distorted. When the direction of the inverse Hilbert transform β is optimally chosen, i.e. parallel to the backprojected line of the distortion, the differentiated sinogram is given by Fig. 2.9(e), and the resulting DBP reconstruction in Fig. 2.9(f) is obtained. As opposed to the FBP method (see Fig. 2.4), for this choice of β, the DBP method localizes the distortion along a strip with width 2ϵ. This can clearly bee seen in the difference image Fig. 2.9(g), that is calculated with respect to the DBP reconstruction from the true sinogram.
If the choice of β is suboptimal, larger parts of the image are distorted. Fig. 2.9(b) shows the differentiated sinogram for the extreme case in which the direction of the inverse Hilbert transform is perpendicular to the backprojected line of distortion. The resulting DBP reconstruction is shown in Fig. 2.9(c). Its difference image with respect to the reconstruction from the true sinogram, is nowhere zero. However, the image accuracy of the reconstruction improves further away from the backprojected line of the distortion.

In Fig. 2.10, the reconstruction (c) from a truncated sinogram (a) is compared with the reconstruction (f) from complete data (c), both using the DBP reconstruction method. Fig. 2.10 (b) and (e) depict the area in which the DBP $b_\phi(\mathbf{x})$ is accurately obtained from the sinograms in (a) and (c), respectively.

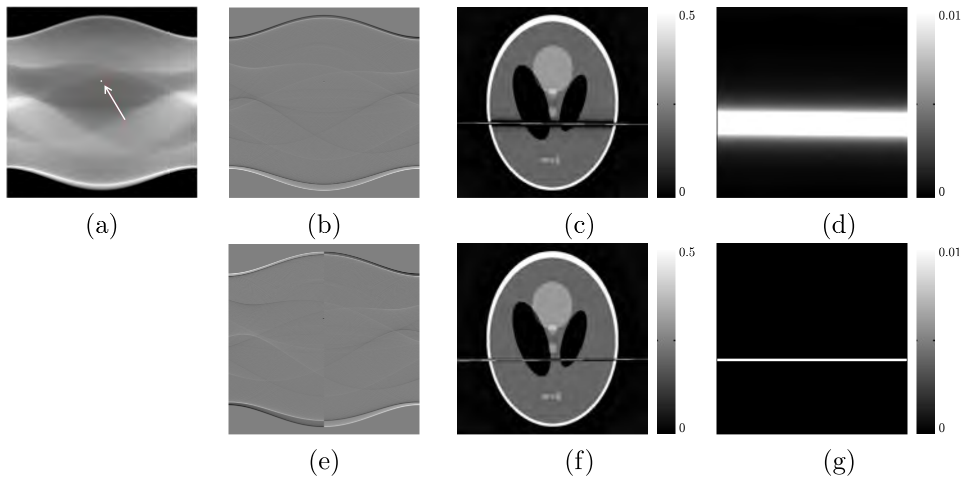


Figure 2.9: *Illustration of the non-local behaviour of the DBP algorithm. (a) Sinogram of Shepp-Logan phantom (see Fig. 1.16(a)) with one distorted pixel (indicated with white arrow). (b) and (e) represent the differentiated sinograms with preselected Hilbert angles that are chosen perpendicular (b) and parallel (e) to the backprojected line of the distortion. (c) and (f) denote the DBP reconstructions based on the differentiated sinograms in (b) and (e), respectively. (d) and (g): absolute value of the difference image with respect to the true Shepp-Logan phantom for images (c) and (f), respectively. Note that the difference image (d) is nowhere zero, while difference image (g), which corresponds to the optimal Hilbert angle, is zero everywhere except along the white area.*

2.5 Uniqueness results and exact reconstruction

Before 2002, it was generally believed that exact 2D reconstruction requires complete sinogram data. This idea changed drastically from the moment that whole series of reconstruction techniques and sufficiency conditions were proposed, triggered by new results on the reconstruction problem for 3D helical cone beam CT. In this section, an overview of the recent uniqueness results for reconstruction from truncated projections is given.

The first breakthrough for reconstruction of 2D tomographic images from incomplete data was achieved for a set of non-truncated projections, covering an incomplete angular range. Fig. 2.11 depicts a circular object and two fan beam source trajectories. The short scan trajectory $\lambda_1 \rightarrow \lambda_2$ covers an angular range $\pi+$ fan angle 2γ, which is the necessary range to determine the complete Radon transform of the object. The second trajectory $\lambda_1' \rightarrow \lambda_2$ is a so called 'super short scan' trajectory covering an angular range smaller than that of the short scan. The projections are considered to be non-truncated, i.e. the complete object support

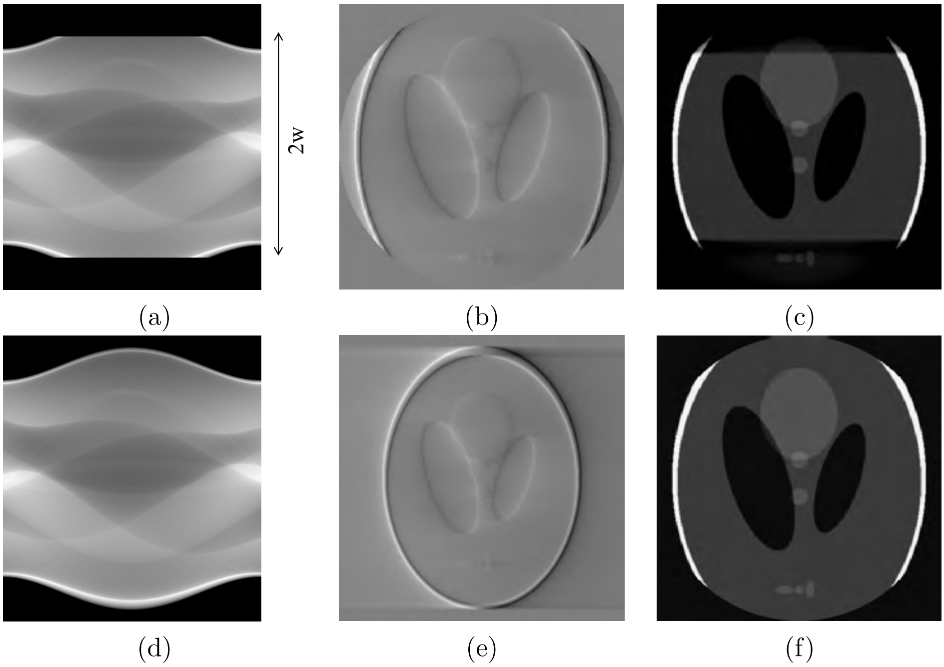


Figure 2.10: *DBP reconstruction from truncated (top row) and complete (bottom row) sinograms. Left column: sinograms; center column: the area in which the DBP $b_\phi(\mathbf{x})$ is accurately obtained from the corresponding sinogram. Right: reconstruction in a ROI with radius w, obtained after inverse Hilbert transform of the DBP.*

is covered in each projection. Inspired by the algorithms for cone-beam spiral CT proposed by Katsevich [8], Noo et al. [9] found that in case of a super short scan, some object regions can still be reconstructed accurately. They proved that *complete (non-truncated) fan beam projections provide sufficient information for the reconstruction of an ROI when every line passing through the ROI intersects the vertex path in a non-tangential way.* Furthermore, they proposed a filtered backprojection formula for reconstruction from super short scan data. In Fig. 2.11, the shaded area corresponds to the region that can be reconstructed according to this uniqueness theorem.

Clackdoyle et al. [2] extended the result of Noo et al. for truncated parallel projections. They noticed that an ROI in the object can be recovered accurately, provided that the truncated parallel data can be converted into non-truncated fan beam data corresponding to a virtual trajectory that obeys the 2D fan beam data consistency condition of Noo et al [9]. In Fig. 2.12, it is illustrated for a non convex object and four truncation scenarios, which object area (shaded) can be recon-

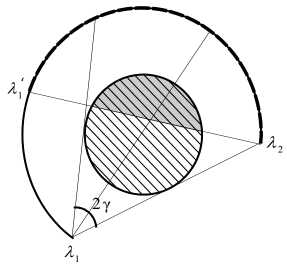


Figure 2.11: *Example of a limited angle fan beam geometry. The inner circle represents the object support. The trajectory from λ_1 to λ_2 (full line) represents a 'short scan', which enables accurate reconstruction in the complete hatched region. The trajectory $\lambda_1' \to \lambda_2$ (dashed line) represents a 'super short scan'; its corresponding fan beam data uniquely determines the object function in the shaded area.*

structed according to this new data sufficiency condition of Clackdoyle et al. [2]. In each scenario, the circle denotes the FOV. Fig. 2.12(d) represents an interior problem in which a small part (hatched region) is known a priori.

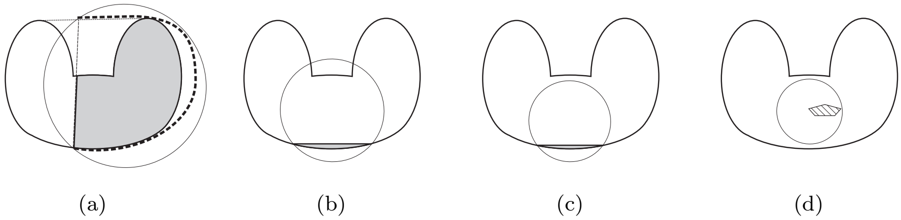
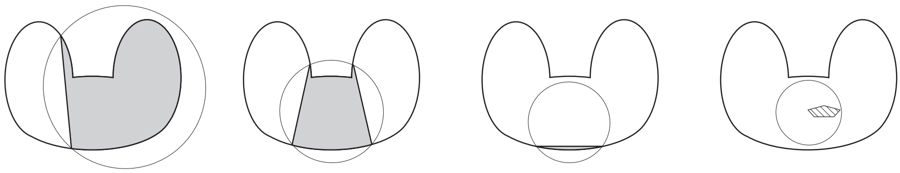


Figure 2.12: *Different scenarios of truncation, depicted for a non-convex object support (solid line) and a circular FOV. In each scenario, the grey area shows the reconstructable region, based on the data sufficiency condition obtained by Clackdoyle et al. [2]. In (a), the optimal source trajectory is drawn that leads to the maximal reconstructable region that is shown.*

The following important step was achieved by Noo et al. [3] and Zou et al. [4], based on the relationship between the differentiated backprojection (DBP) and the one-dimensional Hilbert transform along certain lines of the object function. This relationship is a particular case of the results of Gelfand and Graev [10], and combined with a finite untruncated Hilbert transform inversion formula [7], it leads to the new inversion formula discussed in Section 2.4, and a new data sufficiency condition for 2D tomographic reconstructions.



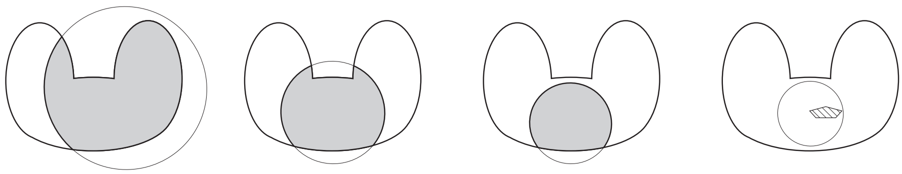
(a) (b) (c) (d)

Figure 2.13: *Different scenarios of truncation, depicted for a non-convex object support (solid line) and a circular FOV. In each scenario, the grey area shows the reconstructable region, based on the data sufficiency condition obtained by Noo et al. [3] and Zou et al. [4].*

The data sufficiency condition can be summarized as follows. Let Ω' be the support of the object that is enlarged with a margin of thickness ϵ with ϵ infinitesimally small, and B' its corresponding B-region. The sufficiency condition for 2D tomography based on the DBP reconstruction formula, can then be summarized as follows : *A region of interest (ROI) inside region A can be reconstructed if the ROI lies within the union of infinite lines that do not contact region B' [3].*
Fig. 2.13 depicts the same non-convex object and truncation scenarios as in Fig. 2.12. For each scenario, the grey area depicts the region that can be recovered based on this sufficiency condition. Notice the enlarged reconstructable region for the cases in Fig. 2.13(a) and (b) compared to Fig. 2.12; in Fig. 2.13(c) and Fig. 2.13(d), the reconstructable areas remain the same as in Fig. 2.12.

Several other one-step FBP-type methods were proposed that accurately recover parts of the object function for specific data truncation problems, such as [11, 12]. They have the advantage that the reconstruction of a certain ROI from complete data requires the computationally intensive backprojection only to be calculated in the ROI.

Building on the results of Noo et al. [3], Defrise et al. [13] analysed the problem of recovering a 1D real function from its Hilbert transform that is known on a finite segment covering the object boundary only at one side (instead of the required two sides in previous sufficiency condition), which lead to a new sufficiency condition. Recall that Ω' represents the object support that is enlarged with a margin of thickness ϵ, with $\epsilon > 0$, and B' is its corresponding B-region. In analogy with Noo's sufficiency condition, the sufficiency condition of Defrise et al. [13] can then be summarized as follows : *A region of interest (ROI) inside region A can be reconstructed if the ROI lies within the union of infinite segments that contact region B' at most at one side [3].* As an illustration, Fig. 2.14 depicts the non-convex object support and truncation scenarios of Fig. 2.12. For each scenario, the grey area represents the region that can be recovered based on Defrise's sufficiency condition. Notice the significant enlargement of the reconstructable area in Fig. 2.14(a), (b) and (c) compared to Fig. 2.12 and Fig. 2.13. For Fig. 2.14(d), again, no part can be reconstructed.



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

Figure 2.14: *Different scenarios of truncation, depicted for a non-convex object support (solid line) and a circular FOV. In each scenario, the grey area shows the reconstructable region, based on the data sufficiency condition obtained by Defrise et al. [13].*

Very recently, Ye et al. [14] and Kudo et al. [15] independently extended the previous uniqueness results. By following a similar approach as Defrise et al. [13] and using less restrictive assumptions on the knowledge of the Hilbert transform, they showed that *the interior problem can be solved if a tiny a priori knowledge on the object $f(x, y)$ is available in the form that $f(x, y)$ is known on a small region located inside the region of interest* [15]. In [16], the uniqueness results are extended to the whole space, and a rigourous stability estimate is performed. As a consequence of these recent uniqueness results, the interior problem in Fig. 2.15 (d) with known object function in the hatched pentagon, that remained unsolved in the previous sufficiency conditions, can now completely be solved. Recall that Fig. 2.14(a), (b) and (c) were already solved using the uniqueness theorem of Defrise et al. [13].

No closed form analytic inversion formula is yet derived for reconstruction of the area that can be accurately determined based on the uniqueness results in [13–16]. Instead, a DBP-POCS method was introduced by Defrise et al. [13] and extended by Kudo et al. [15]. Alternatively, a *Maximum a posteriori* (MAP) method [13] and *Maximum Likelihood Expectation Maximization*(ML-EM) method [17] are successfully applied for the reconstruction of a ROI from truncated data. Such algorithms have the advantage that they are applicable to various types of truncation problems in a unified way.

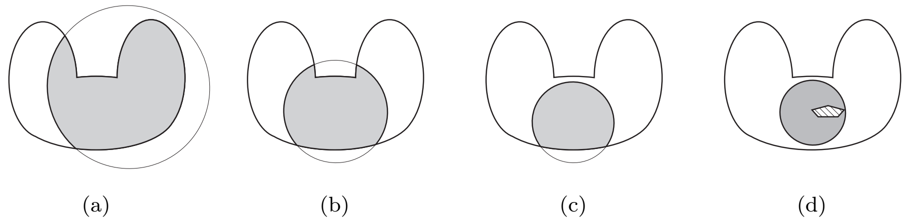


Figure 2.15: *Different scenarios of truncation, depicted for a non-convex object support (solid line) and a circular FOV. In each scenario, the grey area shows the reconstructable region, based on the data sufficiency condition obtained by Ye et al. [14] and Kudo et al. [15].*

2.6 Empirical truncation artifact reduction methods

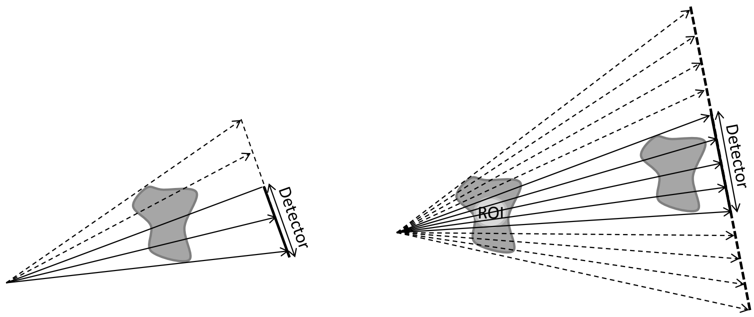
Section 2.5 discussed the uniqueness results that were obtained recently to recover an ROI from truncated data, and summarized the reconstruction techniques that were proposed to obtain accurate reconstruction of the ROI. This section consists of a literature overview of data completion methods and empirical approaches for truncation artifact reduction.

2.6.1 Data completion methods using extra scans

Several techniques have been proposed that complete the missing data by performing additional measurements in an adapted acquisition geometry.

Cho et al. [18] shift the insufficiently large detector along the detector plane (see Fig. 2.16a)) to reduce the problem to one-sided truncation for fan beam and cone beam, respectively. The missing attenuation values are then estimated using the measured lines in the opposite direction, obtained from a full 360° scan. The sinograms are merged using a so-called *iterative reconstruction-reprojection* (IRR) method (see Section 2.6.2). Hooper et al. [19] use a rebinning method for the fusion of multiple sinograms acquired from two or more scans with a repositioned patient. By blocking redundant X-rays, the patient dose can be maximally limited.

In microtomography, one often strives to optimal resolution of a small part of the object. An optimal resolution is acquired by scanning the object positioned at a small distance from the source so that the area of interest covers a larger part of the detector (see Fig. 2.16b). As a result, some line integrals from other parts of the object are not measured. Several authors [21, 22] worked on a technique



(a) (b)

Figure 2.16: *Acquisition schemes for 2nd scan sinogram completion methods. (a) the completion technique of Nassi et al. [20] and Cho et al. [18], who shifted the insufficiently large detector along the detector plane (see Fig. 2.16a)) to reduce the problem to one-sided truncation. The missing attenuation values are estimated using the measured lines in the opposite direction, obtained in a full* $360°$ *scan. (b) the second scan method of [21, 22] sometimes used to enhance resolution in microtomography [23, 24].*

involving a second full scan at lower resolution after translating the object towards the detector until it completely fits in the field of view of the detector. From these data, a complete low resolution reconstruction can be computed, and a forward projection of this image estimates the line integrals that were missing in the high resolution scan. This technique provides relatively accurate reconstructions of the ROI, and is applied in practical experiments, for example in small-animal imaging [25]. Several authors [22, 24] reported that an additional coarse sampling of the truncated part of the sinogram suffices for accurate reconstruction of the ROI.

In another approach, the truncated sinogram is decomposed using appropriate wavelets or related multiresolution models of which the support remains more or less local after ramp filtering; see, among others [23, 26, 27]. Since locality is not completely preserved, additional low resolution information of the truncated part of the sinogram is used. This approach was criticized by Tisson [28] stating that the major part of the improvement is provided by the additional data and not by the local behaviour of the wavelet technique.
Alternative multiresolution methods that do not use prior knowledge have been proposed for the reconstruction of the object discontinuities, for the case when accurate edge detection is aimed at rather than the accurate reconstruction of the attenuation map [29, 30].

2.6.2 Sinogram extrapolation

In many situations, it is not possible to complete the missing data via extra scans, due to practical or dose limitations. The missing line integrals are then typically estimated using a mathematical extrapolation function that incorporates any available prior knowledge. The main aim of the extrapolation function is the restoration of smoothness at the truncation transition in the sinogram to reduce the high frequency artifacts that appear as a bright rim in the reconstructed image. Moreover, the hope is that by using an appropriate model and by exploiting prior knowledge, the unknown sinogram values can be approximated sufficiently accurately to yield a significant reduction of the cupping artifacts, which is important for accurate segmentation of the reconstructed images.

Suppose $(\mathcal{R}_w f)(\theta, s)$ is a truncated sinogram, measured for $s \in [-w, w]$ and let $[s_{\text{ext1}}(\theta), s_{\text{ext2}}(\theta)]$ represent the estimated support of the projection identified with angle θ. If the sinogram support is unknown, a constant interval defined by $s_{\text{ext1}} = -s_{\text{ext}}$ and $s_{\text{ext2}} = s_{\text{ext}}$ is used, where s_{ext} is the radius of the estimated object support circle.
For convenience, the extrapolated sinograms below are described for $s \in [-\infty, 0]$ only; the extrapolation for $s \in [0, \infty]$ is similar. The completed sinogram $p_c(\theta, s)$ is then obtained as follows:

$$p_c(\theta, s) = \begin{cases} 0 & s < s_{\text{ext1}} \\ p_{\text{ext}}(s) & s_{\text{ext1}} \leq s < -w \\ p(\theta, s) & -w \leq s \leq 0 \end{cases}, \tag{2.25}$$

with p_{ext} a smooth function such that $p_{\text{ext}}(w) = p(\theta, w)$ and $p_{\text{ext}}(s_{\text{ext1}}) = 0$. Various extrapolation functions p_{ext} are used in literature, such as a polynomial [31], or $\cos^2$ [28]. The *Simple Extrapolation Method* (SEM) of Lewitt and Bates [32] uses projection data of a physical shape, in this case a circle, as extrapolation function. This function, multiplied by a polynomial, is fitted to the outer extremities of the sinogram. Explicitly, this extrapolation function is written as:

$$p_{\text{ext}}(s) = \sqrt{1 - \left(\frac{s}{s_{\text{ext1}}}\right)^2} \left(c_0 + c_1 \frac{s}{s_{\text{ext1}}}\right) \qquad w < s \leq s_{\text{ext1}}, \tag{2.26}$$

where the coefficients c_0 and c_1 are found by fitting the above function to at least two outer samples of the measured data. In the method of Ohnesorge et al. [31], the unknown part of the sinogram is estimated by a weighted mirroring of the measured data:

$$p_{\text{ext}}(s) = (2p(\theta, w) - p(\theta, 2w - s)) \left(\cos(\frac{s - w}{s_{\text{ext1}} - s})\right)^{0.75} \qquad w < s \leq s_{\text{ext1}}, \tag{2.27}$$

Fig. 2.17 plots a single truncated projection together with various projection extension curves, corresponding to the above described extrapolation methods.

The above extrapolation methods fit a smoothing function to each projection

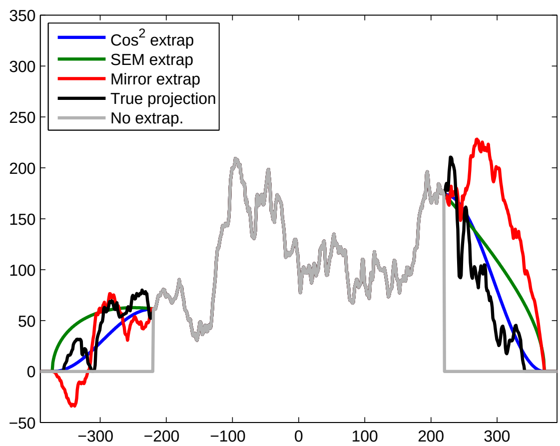


Figure 2.17: *Plot of different extrapolation functions for projection completion.*

separately. Another approach is to perform a global fit for all projections simultaneously by taking into account that a complete sinogram obeys a set of consistency conditions. This idea was introduced for limited angle problems by Lewitt and Bates [32], Peres [33], Louis [34], Natterer[1], and Prince and Willsky [35].

For truncation problems, however, the prior knowledge of the consistency conditions alone is not sufficient to find a unique sinogram extrapolation [32]. Therefore extra assumptions of the object shape are often taken into account. For example, Sourbelle et al. [36], proposed a method where 2 consistency conditions were used to estimate the parameters of a uniform ellipse of which the projection data is used as sinogram extrapolation. Hsieh et al. [37] extrapolate each projection separately using the projection data of a water cilinder, but stretch the extrapolation interval to satisfy the first consistency condition. In Chapter 3, a new sinogram extrapolation method is developed that uses sinogram consistency methods to estimate the missing sinogram data.

Several approaches are proposed to recover the sinogram iteratively. One such method is *iterative reconstruction reprojection* (IRR), which estimates missing line

integrals by iteratively switching between sinogram and object space and each time imposing prior knowledge. This approach was originally developed for the reduction of artifacts in soft tissue caused by neighboring highly attenuating structures [38, 39], and is frequently applied for the reduction of truncation artifacts [18, 20] and other types of incomplete data, e.g. in [40, 41].
Related to this IRR is the method of *projection on convex sets* (POCS) [42]. If the unknown function can be considered as an element of an appropriate Hilbert space, POCS can be used to find a solution that satisfies a set of constraints, provided that each constraints restricts the possible solution onto a convex set. By projecting a guess sinogram consecutively onto the convex sets, a solution in the intersection of the convex sets is found, provided the intersection is not empty. Several authors [43, 44] explored this technique for sinogram restoration in general. Kudo and Saito [45] proposed a POCS method for the recovery of truncated sinograms, using constraints such as the measured data, nonnegativity, the support of the sinogram, a predescribed reference image, and consistency conditions.

2.7 Conclusions

This chapter discussed the reconstruction of a ROI from truncated data in 2D CT. Firstly, it is shown that the standard FBP reconstruction method distributes sinogram distortions over the complete reconstruction domain. Hence, the FBP reconstruction from truncated data systematically yields images contaminated by cupping artifacts. Together with non-uniqueness results for the interior problem, these observations contributed to the general belief that complete data are required for the reconstruction of any ROI. Recently, however, a new formulation of the inverse radon transform using the differentiated backprojection of the projections (see Section 3) was proposed, showing that accurate reconstruction is possible in some areas of the ROI. The most recent uniqueness results and corresponding reconstruction methods were discussed in Section 2.5.

Section 2.6 provided an overview of existing data completion methods and empirical extrapolation methods. This latter class can be subdivided in very simple approaches, mainly aiming at a reduction of the high-frequency artifacts, and more advanced methods, aiming at a reduction of the low-frequency bias, by, for example, incorporating information on the consistency of sinograms.
In the next chapter, a new sinogram extrapolation method is developed that uses sinogram consistency methods to estimate the missing sinogram data.

Bibliography

[1] F. Natterer, *The mathematics of computerized tomography*, SIAM, 2001.

[2] R. Clackdoyle, F. Noo, J. Guo, and J.A. Roberts, "A quantitative reconstruction from truncated projections in classical tomography," *IEEE Transactions on Nuclear Science*, vol. 51, no. 5, pp. 2570–2578, October 2004.

[3] F. Noo, R. Clackdoyle, and J. D. Pack, "A two-step Hilbert transform method for 2D image reconstruction," *Physics in Medicine and Biology*, vol. 49, pp. 3903, August 2004.

[4] Y. Zou, X. Pan, and E.Y. Sidky, "Image reconstruction in regions-of-interest from truncated projections in a reduced fan-beam scan," *Physics in Medicine and Biology*, vol. 50, pp. 13–28, January 2005.

[5] T. Zhuang, S. Leng, B. E. Nett, and G-H. Chen, "Fan-beam and cone-beam image reconstruction via filtering the backprojection image of differentiated projection data," *Physics in Medicine and Biology*, vol. 49, pp. 5489–5503, 2004.

[6] R.P. Kanwal, *Linear integral equations: theory and techniques (2nd edition)*, Boston: Birkhäuser, 1996.

[7] F.G. Tricomi, *Integral Equations, vol. 5*, Dover 1985, 1957.

[8] A. Katsevitch, "Theoretically exact filtered backprojection-type inversion algorithm for spiral CT," *SIAM Journal of Applied Mathematics*, vol. 62, no. 6, pp. 2012–2026, 2002.

[9] F. Noo, M. Defrise, R. Clackdoyle, and H. Kudo, "Image reconstruction from fan-beam projections on less than a short scan," *Physics in Medicine and Biology*, vol. 47, pp. 2525–2546, July 2002.

[10] I. M. Gelfand and M. I. Graev, "Crofton's function and inversion formulas in real integral geometry," *Functional analysis and applications*, pp. 1–5, 1991.

[11] R. Clackdoyle and F. Noo, "A large class of inversion formulae for the 2D Radon transform of functions of compac support," *Inverse Problems*, vol. 20, 2004.

[12] R. Clackdoyle, F. Noo, M. S. Ould Mohammed, and C. Mennessier, "Filtered-backprojection reconstruction formula for 2D tomography with bilateral truncation," in *IEEE Nuclear Science Symposium Conference Record*, 2006, pp. M12–2.

[13] M. Defrise, F. Noo, R. Clackdoyle, and H. Kudo, "Truncated Hilbert transform and image reconstruction from limited data," *Inverse Problems*, vol. 22, pp. 1037–1053, 2006.

[14] Y. Ye, H. Yu, Y. Wei, and G. Wang, "A general local reconstruction approach based on a truncated Hilbert transform," *International Journal of Biomedical Imaging*, p. ID 63634, 2007.

[15] H. Kudo, M. Courdurier, F. Noo, and M. Defrise, "Tiny a priori knowledge solves the interior problem in computed tomography," *Physics in Medicine and Biology*, vol. 53, pp. 2207–2231, 2008.

[16] M. Courdurier, F. Noo, M. Defrise, and H. Kudo, "Solving the interior problem of computed tomography using a priori knowledge," *Inverse Problems*, vol. 24, pp. 065001, 2008.

[17] B. Zhang and G.L. Zeng, "Two-dimensional iterative region-of-interest (ROI) reconstruction from truncated projection data," *Medical Physics*, vol. 34, pp. 935–944, 2007.

[18] P.S. Cho, A.D. Rudd, and R.H. Johnson, "Cone-beam CT from width-truncated projections," *Computerized Medical Imaging and Graphics*, vol. 20, no. 1, pp. 49–57, 1996.

[19] H.R. Hooper and B.G. Fallone, "Technical note: Sinogram merging to compensate for truncation of projection data in tomotherapy imaging," *Medical Physics*, vol. 29, no. 11, pp. 2548–2551, November 2002.

[20] M. Nassi, W.R. Brody, B.P. Medoff, and A. Macovski, "Iterative reconstruction-reprojection: An algorithm for limited data cardiac-computed tomography," *IEEE Transactions on Biomedical Engineering*, vol. BME-295, no. 2, pp. 330–340, 1982.

[21] S. Azevedo, Ph. Rizo, and P. Grangeat, "Region-of-interest cone-beam computed tomography," in *Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, 1995, pp. 381–384.

[22] G. Tisson, P. Scheunders, and D. Van Dyck, "ROI cone-beam CT on a circular orbit for geometric magnification using reprojection," in *IEEE Medical Imaging Conference*, October 2004.

[23] B Sahiner and A.E. Yagle, "Region-of-interest tomography using exponential radial sampling," *IEEE Transactions on Image Processing*, vol. 4, no. 8, pp. 1120–1125, 1995.

[24] R.N. Chityala, K.R. Hoffmann, D.R. Bednarek, and S. Rudin, "Region of interest (ROI) computed tomography," in *Proceedings of SPIE Medical Imaging: Physics of Medical Imaging*, 2004, vol. 5368, pp. 534–541.

[25] I.K. Chun, Myung H. Cho, S.C. Lee, Min H. Cho, and S.Y Lee, "X-ray micro-tomography system for small-animal imaging with zoom-in imaging capability," *Physics in Medicine and Biology*, vol. 49, pp. 3889–3902, August 2004.

[26] T. Olson and J. DeStefano, "Wavelet localization of the radon transform," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2055–2066, August 1994.

[27] A.H. Delany and Y. Bresler, "Multiresolution tomographic reconstruction using wavelets," *IEEE Transactions on Image Processing*, vol. 4, no. 6, pp.

799–813, June 1995.

[28] G. Tisson, *Reconstruction from transversely truncated projections in cone beam microtomography*, PhD thesis, University of Antwerp, 2006.

[29] F. Rashid-Farrokhi, K.J.R. Liu, C.A. Berenstein, and D. Walnut, "Wavelet-based multiresolution local tomography," *IEEE Transactions on Image Processing*, vol. 6, pp. 141230, 1995.

[30] E.J. Candés and F. Guo, "New multiscale transforms, minimum total variation synthesis: applications to edgepreserving image reconstruction," *Signal Processing*, vol. 82, pp. 15191543, 2002.

[31] B. Ohnesorge, T. Flohr, and K. Schwarz, "Efficient correction for CT image artifacts caused by objects extending outside the scan field of view," *Medical Physics*, vol. 27, no. 1, pp. 39–46, 2000.

[32] R.M. Lewitt and R.H.T. Bates, "Image reconstruction from projections: III: Projection completion methods (theory)," *Optik*, vol. 50, no. 3, pp. 189–204, October 1977.

[33] A. Peres, "Tomographic reconstruction from limited angular data," *Journal of Computer Assisted Tomography*, vol. 3, pp. 800–803, 1979.

[34] A.K. Louis, "Picture reconstruction from limited angular data," *Mathematical Methods in Applied Science*, vol. 2, pp. 209–220, 1980.

[35] J.L. Prince and A.S. Willsky, "Constrained sinogram restoration from limited-angle tomography," *Optical Engineering*, vol. 29, pp. 535–543, 1990.

[36] K. Sourbelle, M. Kachelriess, and W.A. Kalender, "Reconstruction from truncated projections in CT using adaptive detruncation," *European Radiology*, vol. 15, no. 5, pp. 1008–1014, 2005.

[37] J. Hsieh, E. Chao, J. Thibault, B. Grekowicz, A. Horst, S. McOlash, and T.J. Myers, "A novel reconstruction algorithm to extend the CT scan field-of-view," *Medical Physics*, vol. 31, no. 9, pp. 2385–2391, 2004.

[38] I.S. Reed, W.V.Jr. Glenn, T.K. Truong, Y.S. Kwoh, and C.M. Chang, "X-ray reconstruction of the spinal cord, using bone suppression," *IEEE Transactions on Biomedical Engineering*, vol. BME-27, pp. 293–298, June 1980.

[39] G. C. McKinnon and R. H. T. Bates, "Towards imaging the beating heart usefully with a conventional CT scanner," *IEEE Transactions on Biomedical Engineering*, vol. BME-28, pp. 123–127, February 1981.

[40] J.H. Kim, K.Y. Kwak, S.B. Park, and Z.H. Cho, "Projection Space Iteration Reconstruction-Reprojection," *IEEE Transactions on Medical Imaging*, vol. 4, no. 3, pp. 139–143, 1985.

[41] K.C. Tam, J.W. Eberhard, and K.W. Mitchell, "Incomplete-data CT image reconstructions in industrial applications," *IEEE Transactions on Nuclear Science*, vol. 37, no. 3, pp. 1490–1499, 1990.

[42] D.C. Youla and H. Webb, "Image restoration by the method of convex pro-

jections: Part 1 - Theory," *IEEE Transactions on Medical Imaging*, vol. MI-1, pp. 81–94, 1980.

[43] M.I. Sezan and H. Stark, "Tomographic image reconstruction from incomplete view data by convex projections and direct Fourier inversion," *IEEE Transactions on Medical Imaging*, vol. MI-3, pp. 91–98, 1984.

[44] P. Oskoui and H. Stark, "Tomographic image reconstruction using the theory of convex projections," *IEEE Transactions on Medical Imaging*, vol. 7, pp. 45–58, 1988.

[45] H. Kudo and T. Saito, "Sinogram recovery with the method of convex projection for limited-data reconstruction in computed tomography," *Journal of Optical Society of America*, vol. 8, pp. 1148–1160, 1991.

Chapter 3

Sinogram extrapolation based on consistency conditions

3.1 Introduction

In this chapter, a sinogram recovery method is presented that uses a set of consistency conditions for the data extrapolation. The approach is similar to the methods in [1–4] that are used in SPECT/PET to estimate the unknown attenuation map for accurate reconstruction of the emission map. We estimate the object function in the region outside the FOV using a set of uniform ellipses. The missing sinogram data can be completed using the corresponding data of the ellipses. The unknown ellipse parameters are obtained by optimizing a cost function based on the Helgason-Ludwig consistency conditions for transmission Radon data. Our method differs from the methods in [1–4] by the definition of the cost function which favours continuity of the extrapolation, by the complexity of the extrapolation function, and by the optimization approach.

This chapter is outlined as follows. In Section 3.2, the Helgason-Ludwig consistency conditions are introduced. Section 3.3 describes the proposed ConSiR method. In Section 3.4, reconstructions of ConSiR extrapolated sinograms are shown for simulated and real μCT data. The reconstruction images are compared to reconstructions using the SEM and the DBP method (see Chapter 2). Finally, conclusions are drawn in Section 3.5.

3.2 Consistency Conditions

Consider an object function $f(\mathbf{x})$ with compact support Ω. Theoretically, the two-dimensional Radon transform (RT) obeys a set of consistency conditions, known as the Helgason-Ludwig (HL) conditions. These conditions state that the function $H_n(\theta)$, described by

$$H_n(\theta) = \int_{-\infty}^{\infty} s^n (\mathcal{R}f)(\theta, s) ds \tag{3.1}$$

is a homogeneous polynomial in $\sin(\theta)$ and $\cos(\theta)$ of degree n for any integer $n \geq 0$.

For example, the first condition (n=0) describes that the total attenuation should be equal for all directions. The second condition describes the consistency of the 'center of mass', which, in Radon space, corresponds to the first moment of the projections. The location of the center of mass, which is a point in image space, describes a sinusoid in the radon space and therefore the result of the integral also needs to be sinusoidal. For a truncated sinogram, these consistency conditions are in general not fulfilled.

An alternative formulation of the HL-conditions [5] is given by:

$$H_{n,k}(\mathcal{R}f) = \int_0^{\pi} \int_{-\infty}^{\infty} s^n e^{ik\theta} (\mathcal{R}f)(\theta, s) ds d\theta = 0 \tag{3.2}$$

for integers $k > n \geq 0$ and $k - n$ even. If the sinogram $(\mathcal{R}f)(\theta, s)$ is inconsistent, the $H_{n,k}(\mathcal{R}f)$ values differ from zero and can therefore be used as a quantitative measure of inconsistency.
Define r as the minimal FOV radius that covers the entire object support Ω such that $(\mathcal{R}f)(\theta, s) = 0$ for all $\theta \in [0, \pi]$ and $|s| > r$. The $H_{n,k}(\mathcal{R}f)$ values can then be written as:

$$H_{n,k}(\mathcal{R}f) \quad = \quad \int_0^{\pi} \int_{-r}^{r} s^n e^{ik\theta} (\mathcal{R}f)(\theta, s) ds d\theta \tag{3.3}$$

3.3 Consistent sinogram recovery method

Suppose a truncated sinogram $(\mathcal{R}_w)f(\theta, s)$ of $f(\mathbf{x})$ that is measured in the limited interval $s \in [-w, w]$ with $w < r$. The corresponding region B (see Chapter 2) is given by $B = \Omega \backslash FOV$, where the FOV is a circle with diameter $2w$.

In the proposed sinogram completion approach, region B is modeled by a set of superposed uniform ellipses. Each of the m_e uniform ellipses $e(\mathbf{x}, \mathbf{p}_m)$ with $m = 1...m_e$, is characterized by 6 parameters $\mathbf{p}_m = (\mu_m, a_m, b_m, x_{m,0}, y_{m,0}, \phi_m)$ representing the density, the semi major and minor axis, the position of the center and the angle between the major axis and the x-axis, respectively [1]. The object function in region B is then approximated by

$$f_B(\mathbf{x}) \approx \sum_{m}^{m_e} e(\mathbf{x}, \mathbf{p}_m). \tag{3.4}$$

In the next step, the missing data are estimated by the corresponding Radon data of the ellipses $(\mathcal{R} f_B)(\theta, s)$, yielding an extrapolated sinogram $(\mathcal{R}_e f)(\theta, s)$.

$$(\mathcal{R}_e f)(\theta, s) = \begin{cases} (\mathcal{R}_w f)(\theta, s) & \text{for } |s| \leq w \\ (\mathcal{R} f_B)(\theta, s) & \text{for } |s| \geq w \end{cases} \tag{3.5}$$

A schematic overview of this procedure is depicted in Fig. 3.1.

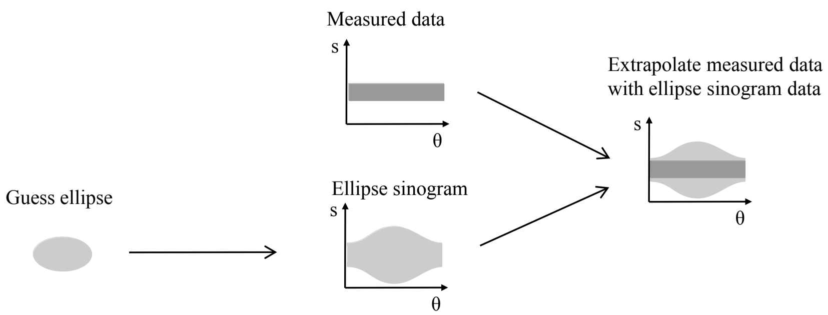


Figure 3.1: *Schematic overview of the sinogram extrapolation.*

The aim is to find an extrapolated sinogram with optimized consistency and with a smooth transition between measured data and extrapolated data, to avoid high-frequency artifacts induced by the FBP ramp filter. Therefore, we determine the initially unknown parameters $\mathbf{p} = (\mathbf{p}_1, ..., \mathbf{p}_N)$ by minimizing a cost function $\Phi(\mathbf{p})$ defined as the weighted sum of an inconsistency penalty I and a discontinuity penalty D

$$\mathbf{p} = \arg\min_{\mathbf{p}} \{\Phi(\mathbf{p})\}, \tag{3.6}$$

with

$$\Phi(\mathbf{p}) = I(\mathbf{p}) + \alpha D(\mathbf{p}), \tag{3.7}$$

where the weight parameter α is determined empirically at $\alpha = \frac{1}{100}$ (see subsection 'Inconsistency', below). The penalties I and D are defined below.

[1] Note that the formulation of the algorithm is not restricted to uniform ellipses, which is specifically interesting when some information on the shape of the object is available. The expected object boundary is then fitted to the truncated sinogram using affine operations, represented by parameters $\mathbf{p} = (\mu_m, a_m, b_m, x_{m,0}, y_{m,0}, \phi)$.

3.3.0.1 Inconsistency

The inconsistency penalty I is constructed using the Helgason-Ludwig values $H_{n,k}(\mathcal{R}_e f)$ Eq. (3.3). Each consistency condition is considered equally important. To this end, the different $H_{n,k}$ are weighted with a factor $\left(\frac{\beta}{r}\right)^n$, where β is a constant. The values of β and α (see Eq. (3.7)) were simultaneously determined by maximizing, for a few examples of truncated sinograms, the correlation between the cost $\Phi(\mathbf{p})$ of the extrapolated sinogram and the inaccuracy of the resulting reconstruction image, which is quantified by the Mean Absolute Difference (see Section 3.4.2). The weight factor β is set at $\beta = 4$, and remains fixed throughout all the experiments in this chapter. The inconsistency penalty I is then defined as the weighted squared L_2 norm of the HL-consistency integrals :

$$I(\mathbf{p}) = \frac{1}{n_c} \sum_{(k,n)\in\mathcal{I}} \left[\left(\frac{\beta}{r}\right)^n H_{n,k}(\mathcal{R}_e f) \right]^2 \tag{3.8}$$

with n_c the number of used consistency conditions, and $\mathcal{I}$ the set of (k, n) representing the different consistency conditions. We use only a small set $\mathcal{I}$ of conditions, namely

$$\mathcal{I} = \{(k, n) : n = 0, 1, 2; k = n + 2, n + 4, n + 6\}. \tag{3.9}$$

3.3.0.2 Discontinuity

The discontinuity penalty D reflects the magnitude of the sinogram discontinuity at $|s| = w$, i.e. at the transition between the measured and the extrapolated part of the sinogram. Therefore, two discontinuity measures $d^-(\theta)$ and $d^+(\theta)$ are calculated for each projection angle θ, at $s = -w$ and $s = w$, respectively. To determine $d^+(\theta)$, two parabolas $q_f^+(\theta, s)$ and $q_e^+(\theta, s)$ are fitted over a small number of pixels [2], respectively at the measured side and at the extrapolated side of the transition (see Fig. 3.2). The shift between values of the two parabolas in $s = w$ then yields $d^+(\theta)$. The same procedure is used to determine $d^-(\theta)$ such that

$$d^+(\theta) = \left| q_f^+(\theta, w) - q_e^+(\theta, w) \right| \tag{3.10}$$

$$d^-(\theta) = \left| q_f^-(\theta, -w) - q_e^-(\theta, -w) \right| \tag{3.11}$$

The discontinuity penalty D is then given by

$$D = \frac{1}{n_\theta} \sum_{n=1}^{n_\theta} (d^-(\theta)^2 + d^+(\theta)^2), \tag{3.12}$$

with n_θ the number of projections.

[2] In this work we fitted the parabola on 3 pixels, but in case of very noisy data, it may be preferable to use more.

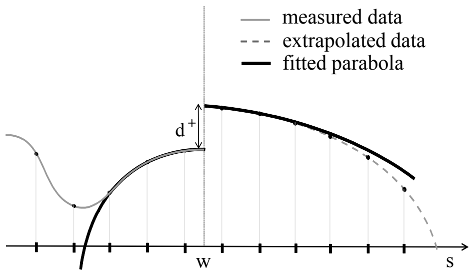


Figure 3.2: *Definition of d^+, the shift at the transition between a parabola fitted to the measured side and a parabola fitted to the extrapolated side for $s = w$.*

3.3.0.3 Optimization

In our experiments we found that the cost function contains many local minima, and consequently, the optimum found by a gradient based optimization algorithms will be very dependent on the choice of the initialization parameters for the optimization. Instead, we used a Differential Evolution (DE) [6] optimization approach, which is an iterative population-based global minimization algorithm for continuous functions. This optimization technique is initialized by generating a population of k random vectors or 'parents'; the number of components equals the number of parameters to optimize, which is $6 \times m_e$. At each iteration (called 'generation'), k child vectors are created by combining two random parents and a third specific strategy vector through a weighted vector sum. A child replaces the 'worst' parent when its cost $\Phi(\mathbf{p})$ is lower than that of the worst parent.

Note that, using this method, there is no guarantee that the global optimum is found. In particular for the case of a large number of ellipses, a high-dimensional cost function is to be optimized with an ill-conditioned optimum. This problem is partly alleviated by limiting the number of ellipses.
Because of the slow convergence and the large number of function evaluations per iteration, DE methods are very time consuming. Therefore, we perform only a relatively small number of iterations, enough to roughly determine the area containing an interesting minimum. An additional speed-up is obtained by performing the optimization using a downsampled sinogram. Afterwards, the minimum value found by the DE method is used as start value for a gradient based method that rapidly converges to the corresponding local minimum.

3.3.0.4 Sinogram smoothing

Using the optimized parameters $\mathbf{p}$, the extrapolated sinogram $\mathcal{R}_e f(\theta, s)$ can be calculated using (3.5). Although the cost function that is minimized contains a discontinuity penalty, the extrapolated sinogram is not necessarily continuous at transition $|s| = w$, since the extrapolation consisting of multiple ellipses is in general not perfectly adapted to the truncated sinogram. We smooth the remaining discontinuities, reflected by the shifts $d^-(\theta)$ and $d^+(\theta)$, by 'smearing them out' over the extrapolation width using a $\cos^2$ function:

$$(\mathcal{R}_e f)(\theta, s) = \begin{cases} (\mathcal{R}_w f)(\theta, s) & \text{for } s \in [-w, w] \\ (\mathcal{R}_e f)(\theta, s) + d^-(\theta)\cos^2(\frac{s-w}{r^-(\theta)-s}) & \text{for } s \in [r^-(\theta), w] \\ (\mathcal{R}_e f)(\theta, s) + d^+(\theta)\cos^2(\frac{s-w}{r^+(\theta)-s}) & \text{for } s \in [w, r^+(\theta)] \end{cases} \tag{3.13}$$

where $s = r^-(\theta)$ and $s = r^+(\theta)$ delimit the projection support of the extrapolated projections. Note that the smoothing results in a slightly decreased consistency of the extrapolated sinogram.
In a final step, the image is computed by performing an FBP reconstruction from the extrapolated sinogram.

3.3.0.5 Summary

In summary the ConSiR method contains the following steps:

1. Initialise the ConSiR optimization by considering guess parameters $\mathbf{p}_m = (\mu_m, a_m b_m, x_{m,0}, y_{m,0}, \phi_m)$ for each of the m_e uniform ellipses. Complete the truncated sinogram (using Eq. (3.5)) with the Radon transform of the ellipses. This yields an extrapolation sinogram $(\mathcal{R}_e f)(\theta, s)$.

2. Find the ellipse parameters that minimize the cost function $\Phi(\mathbf{p})$ (Eq. (3.7)). To this end, compute the Radon transform $(\mathcal{R} f_B)(\theta, s)$ of the current-guess ellipses, and complete the measured truncated sinogram with $(\mathcal{R} f_B)(\theta, s)$ for $|s| > w$ (see Eq. (3.5)). One optimization iteration basically consists of evaluating the cost function for this extrapolated sinogram, proposing a new guess for the ellipse parameters, and computing the updated extrapolated sinogram.
   The optimization strategy is subdivided in two steps. Firstly, a small number of Differential Evolution iterations is performed to obtain a rough estimate of the optimal ellipse parameters. Secondly, the parameters are finetuned by optimizing the cost function using a gradient based optimization method.

3. Smooth the optimally extrapolated sinogram at the transitions $|s| = w$ using a $\cos^2$ function (see Eq. (3.13)).

4. Reconstruct the completed sinogram using FBP.

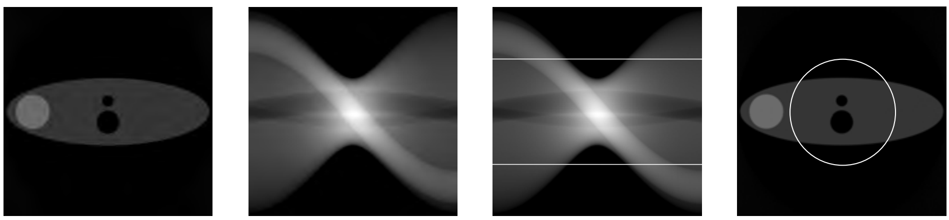
3.4 Results and discussion

3.4.1 Objects consisting of uniform ellipses

We consider a phantom consisting of a few uniform ellipses (see Fig. 3.3(a)). The truncated Radon transform of such an object function has a consistent solution for the extrapolation using a small number of uniform ellipses. We aim to verify whether the ConSiR method finds this consistent extrapolated sinogram.

A complete parallel sinogram, consisting of 1000 radial samples and 300 equally spaced angular views and depicted in Fig. 3.3(b), is simulated for the phantom in Fig. 3.3(a). A truncation problem is simulated by setting the data outside an interval $s \in [-w, w]$ to zero, with $w = 250$ the radius of the FOV. The object and sinogram support are considered unknown.

The truncated sinogram, extrapolated using the ConSiR method with 2 ellipses, is shown in Fig. 3.3 (c); the white lines indicate the transition between the measured and the extrapolated data. It can be observed that the extrapolation is quasi identical to the missing data, which is confirmed by the FBP reconstruction from this extrapolated sinogram, depicted in Fig. 3.3(d).



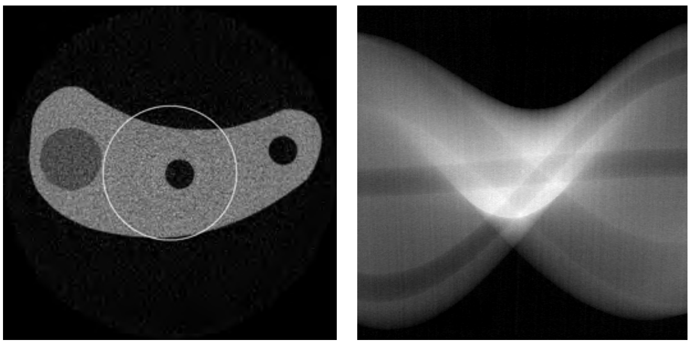
(a) (b) (c) (d)

Figure 3.3: *ConSiR extrapolation for an object consisting of ellipses: (a) phantom, (b) complete sinogram, (c) sinogram extrapolated with ConSiR, the 2 white lines delimit the measured data. (d) the reconstruction from the ConSiR extrapolated sinogram. The white circle represents the FOV.*

3.4.2 The Bean Phantom

In this section, the performance of the ConSiR method is discussed for a non-elliptical phantom of which the object region surrounding any FOV cannot be described in an exact manner using a small set of ellipses. We consider the Bean phantom, depicted in Fig. 3.4(a), which is asymmetrical and has a non-convex

and elongated support. This allows the study of multiple types of truncation by decreasing the radius of a circular FOV.
The Beam phantom (3×1cm) consists of plexiglass containing two small air holes and one larger hole filled with white spirit, which has an intermediate density between that of plexiglass and air. Real X-ray CT data for this phantom is acquired at 100 kV and 100 μA in a SkyScan 1172 scanner with a circular cone beam geometry. The X-ray CT scan is complete in the sense that the FOV of the source-detector pair covers the complete object support in all views. To reduce artifacts from beam hardening, the X-rays are pre-filtered through a thin aluminium and copper plate. We consider only the fan beam data from the central slice, which is rebinned to a parallel beam sinogram ([7]). The resulting sinogram, depicted in Fig. 3.4(b) has 1000 radial samples at 300 equally spaced angular views. The reconstruction from this non-truncated sinogram will be used as ground truth.



(a) (b)

Figure 3.4: *(a) The Bean phantom. The white circle superposed on the phantom denotes the FOV of the experiments in Fig. 3.6. (b) The complete sinogram of the Beam phantom.*

We consider the truncation problem corresponding to a FOV with a diameter of 400 pixel units; the corresponding truncated sinogram is shown in Fig. 3.5(a). The object and sinogram support are considered to be unknown. Fig. 3.5(b) depicts the FBP reconstruction from the truncated sinogram in Fig. 3.5(a). The sinogram that is extrapolated using the ConSiR method with 3 ellipses, is depicted in Fig. 3.5(c). It can be observed that the ConSiR extrapolation provides a good estimation of the sinogram support. The corresponding reconstruction Fig. 3.5(d) shows that the image is well-restored in the FOV. Moreover, it can be observed that the ConSiR extrapolation provides much more structural information than the FBP reconstruction in the region surrounding the FOV.

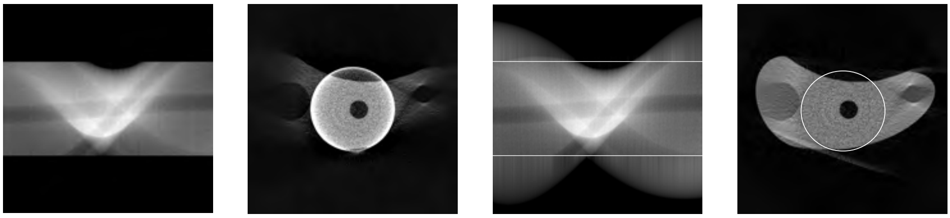


Figure 3.5: *ConSiR extrapolation for a non-elliptical object: (a) truncated sinogram, (b) FBP reconstruction from the truncated sinogram, (c) sinogram extrapolated with ConSiR, the 2 white lines delimit the measured data. (d) the reconstruction from the ConSiR extrapolated sinogram. The white circle represents the FOV with a diameter of 500 pixel units.*

To quantify the accuracy of the ROI reconstructions, we use the mean absolute error (MAE):

$$MAE = \frac{1}{M}\sum_{i=1}^{M} |I_{\text{ref}}(i) - I_{\text{Rec}}(i)| \tag{3.14}$$

with M the number of pixels inside the FOV, I_{ref} the reference reconstruction and I_{rec} the reconstructions from the (extrapolated) truncated sinogram.

In figures 3.6, the reconstruction accuracy of the ConSiR method from truncated projections is compared with several other methods. A FOV with a diameter of 400 pixel units is considered. The left column displays the sinograms that are extrapolated using various methods. In the middle column, the corresponding reconstructions are depicted. The right column shows the difference images of the reconstructions with respect to the reconstruction from complete data. The first row shows the FBP reconstruction from the truncated data. The second row depicts the results obtained using the *Differentiated Back Projection* (DBP) method of Noo et al. (see Section 2.4 and [8])(second row). Although the sinogram support is unknown, we consider the optimal direction for the inverse Hilbert transform to be known (in this case along vertical lines), because it can roughly be estimated if at least some projections are not truncated, which is the case for this example. The DBP difference image in the right column is obtained with respect to the DBP reconstruction from complete data.
The third row presents results for the *Simple Extrapolation Method* (SEM) of Lewitt et al. [9], which basically fits weighted projection data of a circle to each truncated projection separately. In the presented examples, it is assumed that the range of

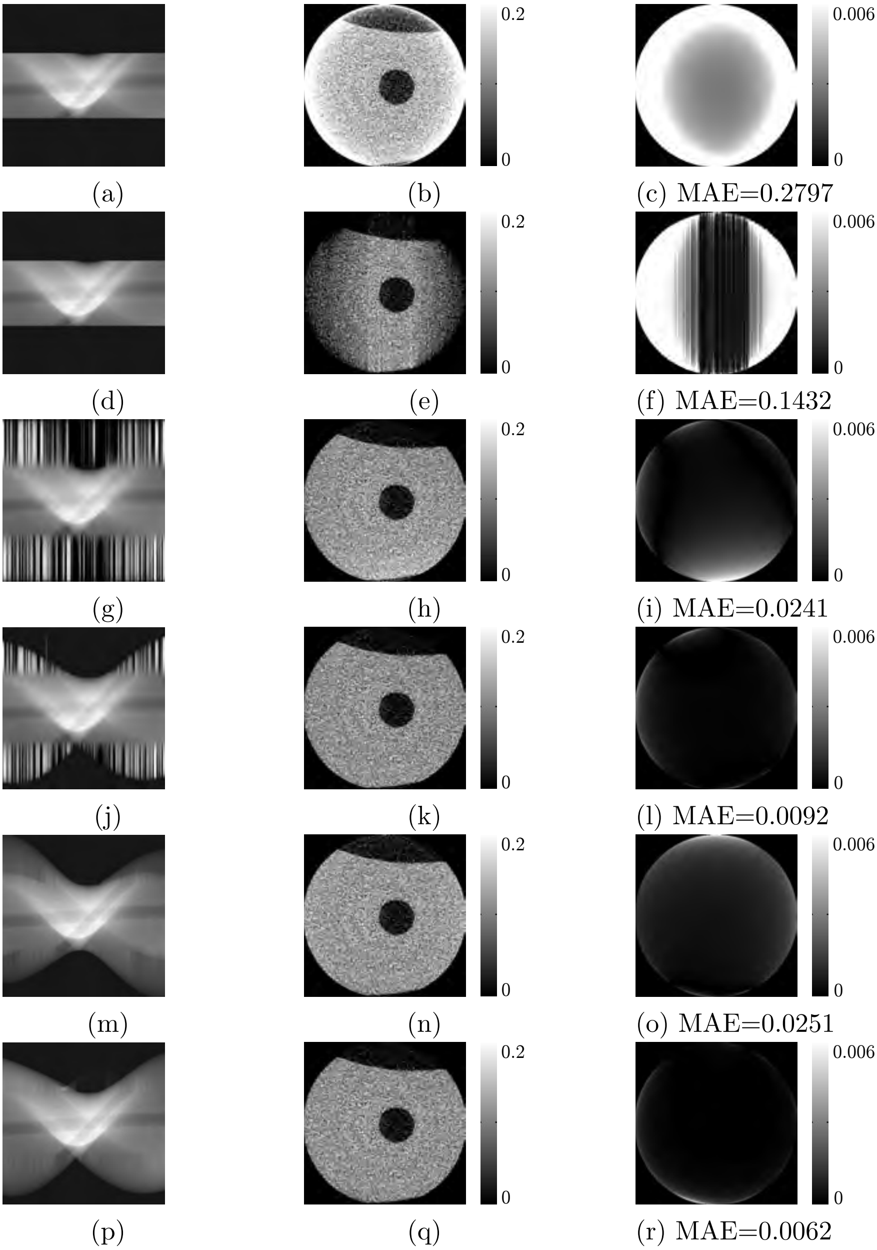


Figure 3.6: *Results of various reconstruction methods for the Bean phantom and truncated data with FOV diameter $2w = 400$ pixel units. Left column: extrapolated sinogram; middle column: reconstruction image; right column: difference image with reference reconstruction. From top to bottom: FBP and DBP without extrapolation (first and second row), SEM without and with sinogram support knowledge (third and fourth row), and ConSiR extrapolation using 2 and 3 ellipses (fifth and sixth row).*

the extrapolated projection corresponds to the detector range of a full scan (1000 pixels). In addition, the SEM reconstruction is shown for the case when the object support is known (SEMS) (fourth row); the extrapolation range is then projection dependent. Finally, the truncated sinograms are extrapolated with ConSiR using 2 ellipses (fifth row) and 3 ellipses (sixth row).
Note that the displayed gray scales are arbitrary and do not correspond to Hounsfield units, which is the commonly used quantitative scale to express the reconstructed attenuation coefficients.

Figure 3.7 plots the MAE for each of the considered methods as a function of the cross-section of the field of view.

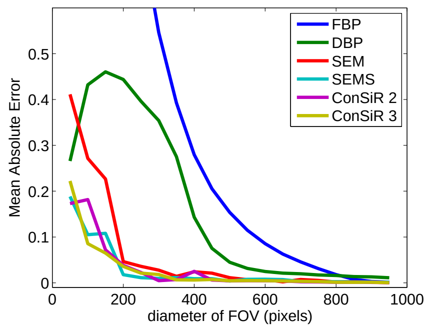


Figure 3.7: *Mean absolute error of the reconstructions performed with various truncation artifact reduction methods, as a function of the FOV diameter for the Bean phantom.*

It can be observed in Fig. 3.7 that the three-ellipse ConSiR extrapolation typically results in more accurate reconstruction images than the two-ellipse extrapolation. This is intuitively expected although it cannot be guaranteed by the ConSiR method. After all, the ConSiR method optimizes the consistency, which gives an indication of, but is not directly related to the image accuracy.

Fig. 3.6 shows that the ConSiR method provides a good sinogram support estimation, which can be seen by comparing the ConSiR extrapolated sinograms (in subfigures (m) and(p)) with the SEMS extrapolated sinogram (displayed in subfigures (j)), which uses the real sinogram support .

The graphs in Fig. 3.7 suggest that the ConSiR method improves the image accuracy compared to the SEM method, but the difference in reconstruction accuracy between both methods is relatively small compared to the magnitude of the cupping reduction already achieved by the SEM method. In case the object support is known a priori, no significant accuracy difference is observed between the ConSiR and SEMS reconstructions.

Note that, since the ConSiR method and SEM extrapolation methods use FBP reconstruction, accurate reconstruction of the attenuation coefficient is not guaranteed, which prohibits its quantitative interpretation. In contrast with the FBP method, the formulation of the DBP method does allow for accurate reconstruction in specific parts of the FOV for certain types of truncation (see Chapter 2 and [8]), which is confirmed in our experiments (second row in Fig. 3.6). The accurate image recovery in this part of the FOV, however, does not necessarily result in a small MAE-value since the remaining part of the FOV has a poor reconstruction quality. In this remaining part, the ConSiR and SEM extrapolation methods seem to outperform the DBP method with respect to image accuracy.

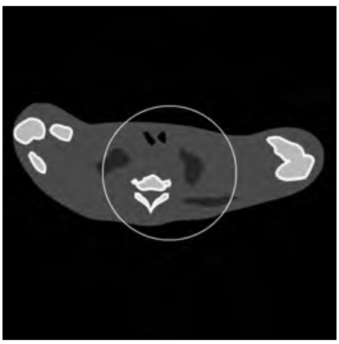
	Comp. time (s)
FBP reconstruction	13
DBP reconstruction	21
SEM (incl. FBP)	22
SEMS (incl. FBP)	22
ConSiR 2 ell. (incl. FBP)	115
ConSIR 3 ell. (incl. FBP)	169

Table 3.1: *Computation times for our implementation of the various methods that were applied for the example in Fig. 3.6.*

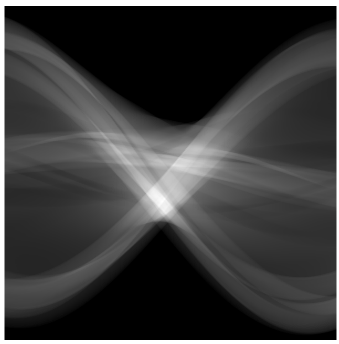
Table 3.1 presents the computation times on a single CPU for our implementation of the various methods that were applied for the example in Fig. 3.6. The most computationally expensive step in the FBP method, is the backprojection, which has time complexity $N^2 n_\theta$, where N is the number of pixels per projection, and n_θ is the number of projections. In this particular example, $N = 400$ and $n_\theta = 300$. The computation times of the SEM and DBP method are similar in this example, although, apart from the backprojection which they share, the underlying processes are completely different. The computation time of the SEM extrapolation, which basically fits a 3-parameter function to a few measured samples at the outer extremities of the truncated sinogram, scales with the number of projections n_θ. The DBP method consists of a backprojection and a convolution. The time complexity in our explicit implementation of the convolution is ($O(N^3)$), but, if the convolution is executed in the Fourier domain using the Fast Fourier Transform (FFT), it can be reduced to $O(N^2 \log(N))$, which corresponds to the complexity of the

Ramp filter in the FBP method.
The ConSiR method optimizes a certain cost function to find the optimal ellipse parameters for the sinogram extrapolation. Recall that m_e denotes the number of ellipses, and define N_r as an estimation of the number of detector pixels that is required to cover the complete object (in this example $N_r = 1000$). Each evaluation of this cost function consists of three steps: analytical calculation of the ellipse sinogram ($O(N_r n_\theta m_e)$), the calculation of the discontinuity ($O(n_\theta)$) and that of the consistency ($O(N_r^2)$). The computation time of one function evaluation is actually very small, but the number of function evaluations that is required for the global and local optimization critically depends on the behavior of the cost function. In this particular example, the number of required function evaluations is large, which causes the total processing time for the ConSiR method to become significantly larger than that of the SEM method. However, the computation time of ConSiR remains in the order of minutes, which seems reasonable for the additional image accuracy that is obtained.

3.4.3 The Thorax phantom



(a)



(b)

Figure 3.8: *(a)The Thorax phantom. The white circle superposed on the phantom denotes the FOV of the experiments in Fig. 3.5. (b) the complete sinogram of the Thorax phantom.*

In this section, the experiments of Section 3.4.2 are repeated for the computer simulated Thorax phantom, which is depicted in figure Fig. 3.8(a). The corresponding complete sinogram is shown in Fig. 3.8(b). The Thorax phantom represents a slice through the human body at shoulder height, and contains various low and high-contrast structures. Truncated data for this phantom represents the frequently posed problem in medical CT systems where the detector is too small to detect

X-rays through the full width of the patient in all directions. Similar to the Bean phantom, the Thorax phantom is asymmetrical and has a non-convex and elongated support, which allows the study of multiply types of truncation by decreasing the radius of a circular FOV. A complete parallel sinogram with 1000 radial samples and 300 equally spaced angular views is simulated. As for the Bean phantom, the reconstruction from the complete parallel sinogram is used as reference.

We consider a truncation problem corresponding to a FOV with a diameter of 400 pixel units. Fig. 3.9, similarly organized as Fig. 3.6, depicts the reconstruction images using the FBP (first row), DBP (second row), SEM (third and fourth row for unknown and known support, respectively), and the ConSiR method (fifth and sixth row for the 2- and 3-ellipse optimization).

In addition, in Fig. 3.10, the MAE for the various methods is plotted with respect to the diameter of the FOV (expressed in pixel units).

The reconstructions for the Thorax phantom confirm the results obtained for the Bean phantom. Again, it can be observed that the 3-ellipse ConSiR optimization is favourable above the 2-ellips ConSiR reconstruction. The results suggest also that the ConSiR method typically leads to more accurate images compared to the SEM method if the object support is unknown. An important limitation to the ConSiR method is that, although it often yields images with small MAE, the method cannot guarantee a good image accuracy in the parts that can be recovered in an exact manner using the DBP method. However, in the remaining parts of the FOV, the ConSiR reconstruction generally yields significantly better image accuracy than the DBP method.

3.5 Conclusions

In this chapter, a new sinogram extrapolation method (ConSiR) is proposed that uses sinogram consistency methods to estimate the missing sinogram data. This sinogram recovery method extrapolates the truncated sinogram with data of one or more ellipses of which the parameters are determined by optimizing the consistency and the continuity of the extrapolated sinogram. The ConSiR method is compared to a variety of ROI reconstruction algorithms. The experimental results suggest that this approach often yields more accurate reconstructions than the *Simple Extrapolation Method*(SEM) [9], but that the difference of image accuracy resulting from both methods is relatively small compared to the magnitude of the cupping reduction already achieved by the SEM method.
Since the ConSiR and SEM methods are essentially based on FBP reconstruction,

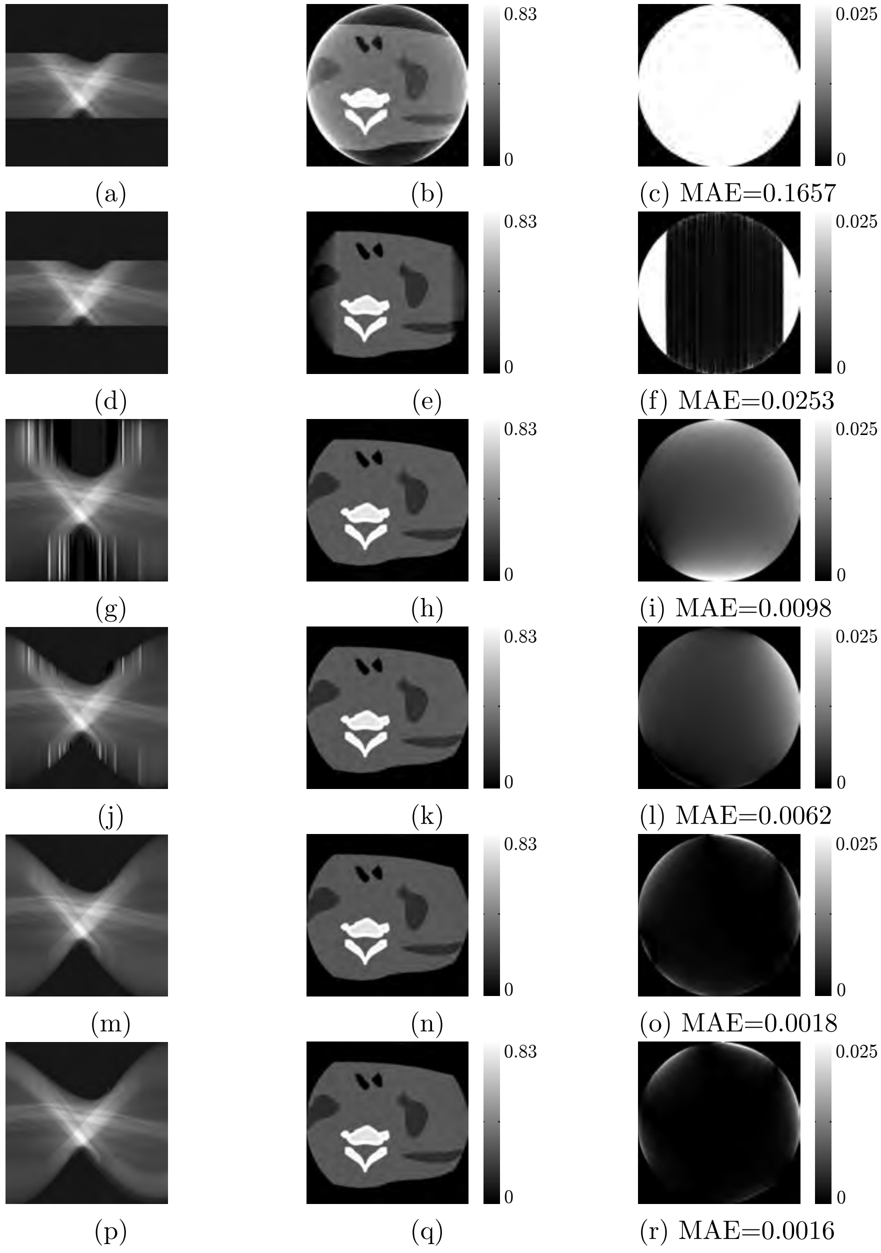


Figure 3.9: *Results of various reconstruction methods for the Thorax phantom and truncated data with FOV diameter $2w = 400$ pixel units. Left column: extrapolated sinogram; middle column: reconstruction image; right column: difference image with reference reconstruction. From top to bottom: FBP and DBP without extrapolation (first and second row), SEM without and with sinogram support knowledge (third and fourth row), and ConSiR extrapolation using 2 and 3 ellipses (fifth and sixth row).*

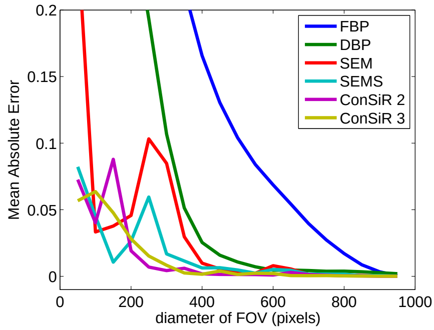


Figure 3.10: *Mean absolute error of the reconstructions performed with various truncation artifact reduction methods, as a function of the FOV diameter for the Thorax phantom.*

the accuracy of the reconstructed attenuation coefficient is not guaranteed, which prohibits its quantitative interpretation. As opposed to the FBP method, the formulation of DBP reconstruction does allow accurate reconstructions in a certain region of the FOV, described by the uniqueness theorem of Noo et al. (see 2.4 and [8]), which was confirmed in our experiments. However, we found that the extrapolation methods ConSiR and SEM outperform the DBP method with respect to image accuracy in the remaining parts of the FOV. We conclude that data extrapolation methods are still relevant since they can be applied regardless of the type of truncation and they can be used in combination with the existing scanning software.

Bibliography

[1] A. Welch and R. Clack, "Toward accurate attenuation correction in SPECT without transmission measurements," *IEEE Transactions on Medical Imaging*, vol. 16, no. 5, pp. 532–541, October 1997.

[2] A. Welch, C. Campbell, and R. Clackdoyle, "Attenuation correction in PET using consistency information," *IEEE Transactions on Nuclear Science*, vol. 45, no. 6, pp. 3134–3141, December 1998.

[3] I. Laurette, R. Clackdoyle, A. Welch, F. Natterer, and G. T. Gullberg, "Comparison of Three Applications of ConTraSPECT," *IEEE Transactions on Nuclear Science*, vol. 46, pp. 6, 1999.

[4] D. Gourion, D. Noll, P. Gantet, A. Celler, and J.P. Esquerré, "Attenuation Correction Using SPECT Emission Data Only," 2001, pp. 1522–1526.

[5] F. Natterer and F. Wübbeling, *Mathematical methods in image reconstruction*,

Monographs on Mathematical Modeling and Computation. SIAM, 2001.

[6] R. Storn and K. Price, "Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, pp. 341–359, 1997.

[7] A. Kak and M. Slaney, *Principles of computerized tomographic imaging*, SIAM, 2001.

[8] F. Noo, R. Clackdoyle, and J. D. Pack, "A two-step Hilbert transform method for 2D image reconstruction," *Physics in Medicine and Biology*, vol. 49, pp. 3903, August 2004.

[9] R.M. Lewitt and R.H.T. Bates, "Image reconstruction from projections: III: Projection completion methods (theory)," *Optik*, vol. 50, no. 3, pp. 189–204, October 1977.

Part III

Reconstruction of piecewise uniform objects from truncated projections

Chapter 4

Reconstruction of a uniform star object from interior x-ray data: uniqueness, stability, algorithm.

4.1 Introduction

This chapter concerns the problem of reconstructing an object with uniform density from x-ray projections. In particular, we consider the reconstruction of star-shaped objects from limited projection data, where the detector only covers an interior-field-of-view. Figure 4.1a shows an example of such a star-shaped object, along with a single projection collected by a detector that is significantly smaller than the diameter of the object. We will show that a 2D star-shaped object of uniform but unknown density is determined by its parallel projections sampled over a full π angular range with a detector that only covers an interior field-of-view.

In a slightly different form, this chapter has been published as:

G. Van Gompel, M. Defrise and K.J. Batenburg, *Reconstruction of a uniform star object from interior x-ray data: uniqueness, stability, algorithm*, Inverse Problems (25), ID 65010, 2009

4.2 Context

We consider the reconstruction of star-shaped objects from limited projection data, where the detector only covers an interior-field-of-view. A broad class of similar problems have been analyzed in the past, all aiming at characterizing objects with constant or piecewise constant densities from tomographic data, with applications mainly for non-destructive testing. These problems differ by the nature of the data, by the data sampling and by the class of objects in which a solution is sought. In particular, the problem considered in this chapter belongs to the field of *Geometric Tomography*, which focuses on the reconstruction of geometric objects from their sections, orthogonal projections or both [1, 2].

Consider the case where the data consist of x-ray projections: an x-ray projection is the set of integrals of the object along a fan (in 2D) or a cone (in 3D) of lines diverging from a vertex, which corresponds physically to the anode of the x-ray source in CT. Volčič [3] showed for instance that a convex 2D object with known uniform density is determined by x-ray projections measured from any set of three non-collinear vertices not contained within the object. For this result, and most other results in the literature on Geometric Tomography, it is assumed that the density of the object is known beforehand and that the projections are not truncated in the sense that the integral of the object is measured (or known to be zero) for *all* lines diverging from the vertex. In CT however the density represents the linear attenuation coefficient for x-rays, and an accurate estimate of that quantity can be obtained only if both the nature of the material and the incident x-ray spectrum are known. In addition, especially with micro-CT scanners, the detector is sometimes too small to cover the whole sample and in that case the projections are necessarily truncated. This observation motivates the present study of the *interior problem* for an object with uniform but unknown density. In the interior problem, the integral of the object is measured only for those lines that intersect a circular *field-of-view* (FOV) contained within the support of the object. This corresponds to CT data acquired with a short detector and a 2π rotation of the assembly x-ray source-detector. We assume that these data are parameterized as *parallel projections*. This choice of the parametrization does not restrict the generality of the uniqueness theorem in Section 4.5.

In medical tomography, the density to be reconstructed is an arbitrary function and in that case the solution of the interior problem is not unique (see Theorem 6.5 in [4] and the singular value analysis in [5]). Uniqueness however can be restored if strong prior knowledge is available. For instance, it was shown recently (see Chapter 2 and [6, 7]) that interior data determine the density function in a unique and stable way within the measured FOV, as soon as this function is known a priori in a subset of that FOV. This chapter follows a similar approach based on the DBP to show in Section 4.5 that the interior problem also allows unique reconstruction provided the object is known to be uniform and star-shaped (in the sense that each half-line diverging from the center of the FOV has one and only one intersection with the boundary of the object support). A stability estimate is obtained in Section 4.7 using the Cramer-Rao bound.

The major goal of this chapter is to prove these new uniqueness and stability results for the reconstruction of a star object from interior data. Interestingly however, the proof leads directly to a reconstruction algorithm, denoted as the DBP algorithm and described in Section 4.8. As discussed above, this algorithm requires that the object has a uniform density. This assumption is restrictive, not only because it prevents the application to piecewise uniform samples containing different materials but also because of confounding physical effects such as beam-hardening.

4.3 Notation and concepts

Let $\mathbb{R}_{>0} = \{x \in \mathbb{R} : x > 0\}$. Let $\mathcal{J}$ be the class of finite and integrable functions $f : \mathbb{R}^2 \to \mathbb{R}$ with a compact support for which the Radon transform

$$p(s,\theta) = (\mathcal{R}f)(s,\theta) = \int_{-\infty}^{\infty} f(s\cos\theta - t\sin\theta, s\sin\theta + t\cos\theta)\, dt, \tag{4.1}$$

is defined almost everywhere in $s \in \mathbb{R} \times 0 \leq \theta < \pi$, with the symmetry $p(s,\theta) = p(-s,\theta+\pi)$. The variables s and θ are the usual sinogram variables: s is the signed distance between the line and the origin of the coordinate system, and θ defines the line orientation as shown in figure 4.1.

Let $O = (0,0)$. A set $S \subset \mathbb{R}^2$ is called *star-shaped* at O if every line through O that meets S does so in a line segment. By $\mathcal{S}$, we denote the collection of nonempty, compact sets that are star-shaped at O and for which O is an interior point. Let $S \in \mathcal{S}$. The *radial function* $\rho_S : \mathbb{R}^2 \to \mathbb{R}_{>0}$ of S is defined by

$$\rho_S(x,y) = \max\{h : (hx, hy) \in S\}. \tag{4.2}$$

The radial function is often restricted to the unit circle. Switching to polar coordinates, it can then be represented by a periodic function $u_S : \mathbb{R} \to \mathbb{R}_{>0}$ with period 2π, such that for $r > 0$ and $0 \leq \phi < 2\pi$:

$$(r\cos\phi, r\sin\phi) \in S \iff r \leq u_S(\phi). \tag{4.3}$$

We call a star-shaped set S a *star object* if u_S is a continuous function. Let S be a star object and let $c \in \mathbb{R}$. Define a *star object of density* c by

$$f_{S,c}(r\cos\phi, r\sin\phi) = \tilde{f}_{S,c}(r,\phi) = \begin{cases} c & r \leq u_S(\phi) \\ 0 & r > u_S(\phi) \end{cases}, \tag{4.4}$$

where $\tilde{f}_{S,c}$ corresponds to the polar coordinate representation of $f_{S,c}$. Note that the term *star object* will be used to refer to either the star-shaped set or its representation as a function, depending on context. The set of all functions $f_{S,c}$, where S is a star object and $c \in \mathbb{R}$ is denoted by $\mathcal{F}$. From now on, we omit the index S, c and S for the functions f and u.
It is well-known that under very general assumptions (theorem 1.7 in [8]), any function in $\mathcal{J}$ is uniquely determined by its Radon transform. This chapter concerns the problem of reconstructing a star object from only part of its Radon transform, corresponding to a detector for which the field of view covers only part of the interior of the object. The following uniqueness theorem is proved:

UNIQUENESS THEOREM *Let $w > 0$. Every function $f \in \mathcal{F}$ is uniquely determined, among all $g \in \mathcal{F}$, by the value $p(s,\theta)$ of its Radon transform on the set $|s| \leq w$, $0 \leq \theta < \pi$.*

Note that no central symmetry is assumed, i.e. the radial function u is not assumed to satisfy $u(\phi) = u(\phi + \pi)$. This uniqueness theorem will be proved in Section 4.5. As the proof is constructive, it also leads to an algorithm for solving the following reconstruction problem:

PROBLEM 1. (Reconstruction problem for unknown density). *Let $w > 0$. Suppose that the Radon transform $p(s,\theta)$ of $f \in \mathcal{F}$ is given for $|s| \leq w$ and $0 \leq \theta < \pi$. Compute f from this partial Radon transform.*

In Section 4.7, it is shown that the solution of Problem 1 can be quite unstable. Stability can be significantly improved if the density of f is known a priori. This leads to the following reconstruction problem, which fits perfectly in the domain of geometric tomography:

PROBLEM 2. (Reconstruction problem for known density). *Let S be a star object, let f be the indicator function of S and let $w > 0$. Suppose that the Radon transform $p(s,\theta)$ of f is given for $|s| \leq w$ and $0 \leq \theta < \pi$. Compute f from this partial Radon transform.*

4.4 The differential backprojection and the Hilbert transform

We recall in this section the differential backprojection (DBP), see Chapter 2. Let $f \in \mathcal{J}$ be a finite and integrable function with a compact support, and let $p(s,\theta)$

be its Radon transform. Consider a fixed direction $0 \leq \phi < \pi$ and the central line $\mathcal{L}_\phi = \{(z\cos\phi, z\sin\phi),\, z \in \mathbb{R}\}$. For each point z along this line, define the backprojection of the derivative of the Radon transform Eq. (2.13):

$$g_\phi(z) = \frac{-1}{2}\int_0^\pi \operatorname{sgn}(\cos(\theta-\phi))\, \frac{\partial p(s,\theta)}{\partial s}\Big|_{s=z\cos(\theta-\phi)} d\theta\,. \tag{4.5}$$

This function is defined almost everywhere in $z \in \mathbb{R}$ because the derivative $\partial p/\partial s$ of the Radon transform of $f \in \mathcal{J}$ can be defined almost everywhere. If necessary, it can be defined as a distribution (see Chapter 10 in [9] for the definition of the Radon transform of distributions and section 4.1 in [10] for the definition of the DBP and Hilbert transform of functions in $L_0^2(\mathbb{R}^2)$).
Noo et al [11] and Zou et al [12] have shown that the DBP function in Eq. (4.5) is related to the Hilbert transform of f along $\mathcal{L}_\phi$ by

$$g_\phi(z) = (\mathcal{H}f_\phi)(z) = p.v.\int_{-\infty}^{\infty} \frac{1}{z-z'}\, f_\phi(z')\, dz' \quad z \in \mathbb{R}, \tag{4.6}$$

where $p.v.$ denotes the Cauchy principal value and f_ϕ is the restriction of f to $\mathcal{L}_\phi$:

$$f_\phi(z) = f(z\cos\phi, z\sin\phi) = \begin{cases} \tilde{f}(|z|,\phi) & z \geq 0 \\ \tilde{f}(|z|,\phi+\pi) & z < 0 \end{cases}\,. \tag{4.7}$$

with $\tilde{f}$ the polar coordinate representation of f. To simplify notations, this definition (Eq. (4.6)) of the Hilbert transform $\mathcal{H}$ differs from the standard definition (Eq. (2.3)) by a factor π. Note that the DBP separates the 2D problem into a set of independent 1D problems along a family of central lines $\mathcal{L}_\phi$. The usefulness of the DBP stems from the existence of a closed form expression (equation 12 p. 175 in [13]) for the *inverse finite Hilbert transform*, which allows recovering a function $f_\phi(z)$ that vanishes outside the interval $(-1, 1)$ from its Hilbert transform $g_\phi(z)$ on $z \in [-1, 1]$.

4.5 Uniqueness for star objects

In this section we prove Theorem 1 in a constructive manner. The proof will also provide the basis of a reconstruction algorithm, described in Section 4.8.
We will show that for any $w > 0$ the star object (4.4) can be reconstructed in a unique way from its Radon transform $p(s,\theta)$, measured over the region $|s| \leq w$, $0 \leq \theta < \pi$. The line integrals of f in (4.1) are measured for all lines that intersect a circular field-of-view of radius w centered at $x = y = 0$, where the parameter w is determined by the size of the detector. Note that the field-of-view can be centered at any point such that each half line diverging from this point has only one intersection with the boundary of the support of f. Without loss of

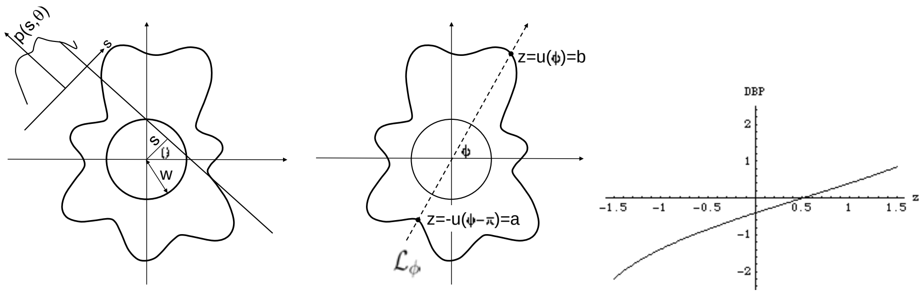


Figure 4.1: *Left: A star object with the interior field-of-view of radius w. The line integrals $p(s,\theta)$ of this object are measured for all lines intersecting the field-of-view, i.e. all lines with $|s| \leq w$ and $0 \leq \theta < \pi$. Middle: The data derivative $\partial p/\partial s$ is backprojected along a central line $\mathcal{L}_\phi$ through the origin to obtain the DBP $g_\phi(z)$ on the segment $-w \leq z \leq w$. Right: the radius $b = u(\phi)$ and $-a = u(\phi+\pi)$ of the star object along $\mathcal{L}_\phi$ and its density c are determined by fitting Eq. (4.8) to the DBP $g_\phi(z)$, here illustrated for $a = -2, b = 3$ and $w = 1.5$.*

generality we assume that $u(\phi) > w$ for $0 \leq \phi < 2\pi$, the problem is then referred to as an *interior problem.* This assumption is not restrictive since w can always be decreased until an interior problem is obtained.
To prove Theorem 1, we use the differential backprojection (DBP). For the specific case of the star object, inserting (4.7) into (4.6) and using (4.4) yields

$$\begin{aligned} g_\phi(z) &= -c\,\log(u(\phi) - z) + c\,\log(z + u(\phi+\pi)) \qquad -w < z < w \\ &= -c\,\log(b - z) + c\,\log(z - a)\,, \end{aligned} \tag{4.8}$$

where we have defined the end points $a = -u(\phi+\pi)$ and $b = u(\phi)$ of the support of f_ϕ (i.e. the support of f along the line $\mathcal{L}_\phi$) and we omit the dependence of a and b on ϕ to simplify notations. Note that $a < -w$ and $b > w$ because the interior field-of-view of radius w is contained within the support of f.
In practice one would estimate the three parameters a, b, c using typically a least-square fit of the RHS of (4.8) to the DBP $g_\phi(z)$ on the segment $z \in [-w, w]$ where it can be recovered from the interior data. To prove uniqueness however, we are not concerned by stability, and we simply consider the first and second derivatives of the Hilbert transform,

$$\begin{aligned} g'_\phi(z) &= \frac{dg_\phi(z)}{dz} = \frac{c}{b-z} + \frac{c}{z-a}, \\ g''_\phi(z) &= \frac{d^2 g_\phi(z)}{dz^2} = \frac{c}{(b-z)^2} - \frac{c}{(z-a)^2}, \end{aligned} \tag{4.9}$$

which are continuous on $-a < -w \leq z \leq w < b$, with $g'_\phi(z) > 0$. An additional equation is given by the integral of f along $\mathcal{L}_\phi$, which is measured because that central line intersects the field-of-view. This integral is equal to $p(0, \phi - \pi/2) = c(b - a)$. Using (4.9), one obtains

$$
\begin{aligned}
a + b &= \frac{-p(0, \phi - \pi/2)\, g''_\phi(0)}{(g'_\phi(0))^2}, && (4.10)\\
ab &= \frac{-p(0, \phi - \pi/2)}{g'_\phi(0)}. && (4.11)
\end{aligned}
$$

If the data are consistent the system (4.10,4.11) has a unique solution such that $b > 0$ and $a < 0$. The object density is then recovered as

$$c = p(0, \phi - \pi/2)/(b - a). \tag{4.12}$$

This concludes the proof of Theorem 1.

4.6 Generalizations

Theorem 1 has been obtained for the parallel-beam parametrization of the 2D Radon transform, Eq. (4.1). Extension to the 2D fan-beam parametrization is straightforward by resampling the fan-beam data into parallel-beam data, as is often done in CT. Alternatively, it is possible to avoid this resampling by exploiting instead of Eq. (4.5) a similar equation that directly relates the Hilbert transform of f to the backprojection of its differentiated fan-beam data (see [14] and Eq. (24) in [15]). That relation also holds for the 3D x-ray transform and could therefore be applied for a spiral data acquisition with a multi-row CT scanner.

The uniqueness theorem for the interior problem with star objects can also be extended to more general classes of objects. Note first that Eq. (4.5) and (4.6) are valid for arbitrary density functions in $\mathcal{J}$. Applying the same approach as in the previous section to a class of objects which can be described along each line $\mathcal{L}_\phi$ by J parameters $a_{j,\phi}, j = 1, \ldots, J$, the interior problem is reduced to estimating these parameters by fitting the function $g_\phi(z)$ that has been recovered on $z \in [-w, w]$ using the DBP. Uniqueness should then be verified for each type of parametrization. More generally, we conjecture that a uniqueness theorem might be obtained for general binary objects by using the analyticity lemma 2.1 of [16]. It is likely however that the stability with respect to measurement noise will rapidly degrade with objects of increased complexity. Chapter 5 aims to indicate the extent of object complexity for which the inverse problem from truncated data can be recovered in practice. In this chapter, the attention is restricted to star objects.

4.7 Stability

The stability of the inverse problem in the previous section can be analyzed by calculating for each radial line $\mathcal{L}_\phi$ the Cramer-Rao lower bound (see e.g. section 13.3.5 in [17]) for the variance of an unbiased estimator of the parameters $a = -u(\phi+\pi)$ and $b = u(\phi)$. Recall that an estimator $\hat{a}$ of the parameter a is unbiased if its expectation $E(\hat{a})$ is equal to a. We derive this Cramer-Rao lower bound under the following assumptions:

- The DBP $g_\phi(z)$ calculated using (4.5) is a white Gaussian stochastic process on $-w \leq z \leq w$, with mean value given by (4.8) and with uniform variance σ^2,
- The ray-sum $p(0, \phi-\pi/2)$ along $\mathcal{L}_\phi$ is a Gaussian random variable with mean value $c(b-a)$ and with variance σ_p^2.
- There is no correlation between the noise on $p(0, \phi - \pi/2)$ and on $g_\phi(z)$.

The logarithm of the likelihood function is then

$$L(g, p|a, b, c) = \frac{-1}{2\sigma_p^2}(p(0, \phi - \pi/2) - c(b-a))^2 \tag{4.13}$$

$$-\frac{1}{2\sigma^2}\int_{-w}^{w} (g_\phi(z) + c\,\log(b-z) - c\,\log(z-a))^2\, dz\,. \tag{4.14}$$

The Fisher matrix for the three parameters a, b, and c is

$$F = \begin{pmatrix} F_{a,a} & F_{a,b} & F_{c,a} \\ F_{a,b} & F_{b,b} & F_{c,b} \\ F_{c,a} & F_{c,b} & F_{c,c} \end{pmatrix}, \tag{4.15}$$

with

$$\begin{aligned}
F_{a,a} &= -E\left(\frac{\partial^2 L}{\partial a^2}\right) = \frac{c^2}{\sigma_p^2} + \frac{1}{\sigma^2}\int_{-w}^{w} \frac{c^2}{(z-a)^2}\, dz = \frac{c^2}{\sigma_p^2} + \frac{2c^2 w}{\sigma^2(a^2-w^2)}, \\
F_{a,b} &= -E\left(\frac{\partial^2 L}{\partial a \partial b}\right) = \frac{-c^2}{\sigma_p^2} + \frac{1}{\sigma^2}\int_{-w}^{w} \frac{c^2}{(z-a)(b-z)}\, dz \\
&= \frac{-c^2}{\sigma_p^2} + \frac{c^2}{\sigma^2(a-b)} \log\left\{\frac{(a+w)(w-b)}{(w-a)(b+w)}\right\}, \\
F_{b,b} &= -E\left(\frac{\partial^2 L}{\partial b^2}\right) = \frac{c^2}{\sigma_p^2} + \frac{1}{\sigma^2}\int_{-w}^{w} \frac{c^2}{(b-z)^2}\, dz = \frac{c^2}{\sigma_p^2} + \frac{2c^2 w}{\sigma^2(b^2-w^2)},
\end{aligned}$$

$$
\begin{aligned}
F_{c,c} &= -E\left(\frac{\partial^2 L}{\partial c^2}\right) = \frac{(b-a)^2}{\sigma_p^2} + \frac{1}{\sigma^2}\int_{-w}^{w} \left(\log(b-z) - \log(z-a)\right)^2 dz, \\
F_{c,a} &= -E\left(\frac{\partial^2 L}{\partial c \partial a}\right) = \frac{c(a-b)}{\sigma_p^2} + \frac{1}{\sigma^2}\int_{-w}^{w} \left(\log(b-z) - \log(z-a)\right) \frac{c}{z-a}\, dz, \\
F_{c,b} &= -E\left(\frac{\partial^2 L}{\partial c \partial b}\right) = \frac{c(b-a)}{\sigma_p^2} + \frac{1}{\sigma^2}\int_{-w}^{w} \left(\log(b-z) - \log(z-a)\right) \frac{c}{b-z}\, dz,
\end{aligned}
\tag{4.16}
$$

where $E()$ denotes the expectation value. The variance of any unbiased estimator $\hat{a}, \hat{b}, \hat{c}$ of the three parameters is then bounded below by the diagonal elements of the inverse Fisher matrix,

$$\text{Var } \hat{a} \geq (F^{-1})_{a,a}, \quad \text{Var } \hat{b} \geq (F^{-1})_{b,b}, \quad \text{Var } \hat{c} \geq (F^{-1})_{c,c}. \tag{4.17}$$

When the density c is known beforehand, the lower bound on the variance of $\hat{a}$ and $\hat{b}$ is obtained by calculating the inverse of the 2×2 Fisher matrix

$$F = \begin{pmatrix} F_{a,a} & F_{a,b} \\ F_{a,b} & F_{b,b} \end{pmatrix}, \tag{4.18}$$

with the same matrix elements as in Eq. (4.16).

As an illustration, figure 4.2 shows the lower bound for the variance of $\hat{a}$ as a function of the radius w of the field-of-view, when the true values of the parameters are $a = -2$, $b = 3$ and $c = 1$, and the variances are $\sigma^2 = \sigma_p^2 = 1$. For this example, the variance increases dramatically when the radius of the field-of-view, w, is small, which could be expected. In contrast, the variance bound is small when $w \to |a| = 2$: in that limit the field-of-view approaches the corresponding boundary of the support of f along the line $\mathcal{L}_\phi$. The localization of that boundary (a in this example) is easy to determine in this limit because the Hilbert transform is singular at $w = -a$. This can also be seen by noting from Eq. (4.16) that $F_{a,a} \to \infty$ when $w \to |a|$. Another expected observation in figure 4.2 is that the variance bound is much better when the density c is known a priori.

Figure 4.3 shows the value of $F^{-1}_{b,c} / \sqrt{F^{-1}_{b,b}\, F^{-1}_{c,c}}$ as a function of σ_p. This quantity is equal to the asymptotic value of the correlation coefficient between the maximum-likelihood estimators of b and of c. Intuitively one expects this correlation coefficient to be negative, i.e. one expects that the estimated density tends to increase when the estimated object shrinks. However, when the variance σ_p^2 on the measured value of the ray-sum is large, figure 4.3 reveals the counter-intuitive result that the correlation coefficient is positive.

To conclude this section, we stress that the stability estimates above have been obtained by considering separately each radial line $\mathcal{L}_\phi$. Intrinsically the problem is two-dimensional and it is therefore likely that better variance bounds could

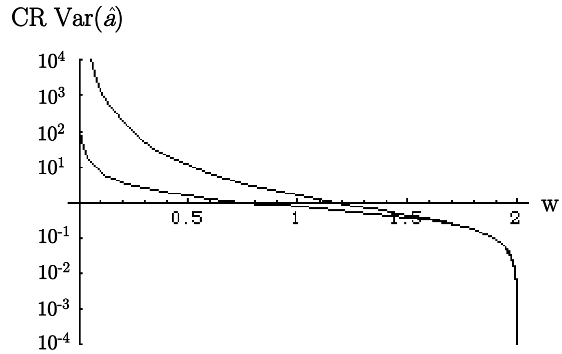


Figure 4.2: Cramer-Rao lower bound for the variance of an unbiased estimator of the boundary $\hat{a}$, assuming a Gaussian distribution with variance 1 for the Hilbert transform data $g_\phi(z)$ and for the ray sum $p(0, \phi - \pi/2)$. Logarithmic vertical scale for the Cramer-Rao variance bound. Horizontal axis: the radius of the field-of-view w. The true values of the parameters are $a = -2$, $b = 3$ and $c = 1$. The lower curve corresponds to the case where the density c is known. The upper curve corresponds to the case where c is unknown.

be obtained by handling the full interior data set $\{p(s,\theta), |s| \leq w, 0 \leq \theta < \pi\}$ simultaneously. Another limitation is our assumption of white noise on the DBP $g_\phi(z)$: this is at best an approximation because noise correlations are introduced when calculating the DBP (4.5) from the measured data $p(s, \phi)$.

4.8 The DBPS algorithm for Problem 1 and Problem 2

The Cramer-Rao bound suggests that the reconstruction of a star object from interior data can be rather unstable. To improve the stability, we propose a two-step reconstruction, which is not optimized but at least partially alleviates the sub-optimality due to the separate handling of each radial line: in a first step an estimate of the density c is obtained for each radial line $\mathcal{L}_\phi$ by using Eq. (4.10), (4.11) and (4.12), with values of $g'_\phi(0)$ and $g''_\phi(0)$ estimated by fitting a polynomial to the DBP $g_\phi(z)$ on $-w \leq z \leq w$. The estimated density is then averaged over all radial lines, and that average value is used as a known density during a second step in which only the boundaries a and b are to be determined. If the Gaussian noise assumptions made when deriving the Cramer-Rao bound are valid, and if the variances σ^2 and σ_p^2 are known, maximum-likelihood estimates can be obtained by

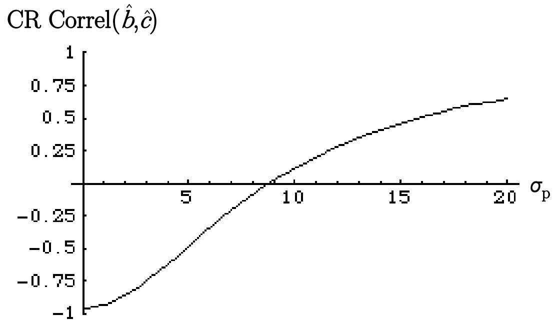


Figure 4.3: *Asymptotic Cramer-Rao value for the correlation coefficient between the estimator of $\hat{b}$ and of the density $\hat{c}$, assuming a Gaussian distribution with variance 1 for the Hilbert transform data $g_\phi(z)$ and with variance σ_p^2 for the ray sum $p(0, \phi - \pi/2)$. Horizontal axis: the value of σ_p. The true values of the parameters are $a = -2$, $b = 3$ and $c = 1$, and the width of the FOV is $w = 1$.*

maximizing the log-likelihood function (4.14):

$$(\hat{a}, \hat{b}) = \arg \max_{a<0, b>0} L(g, p|a, b, c) \,. \tag{4.19}$$

Noting that the statistical properties of $g_\phi(z)$ are unknown and probably complex, we have implemented instead the following weighted least square method, which is simpler and avoids the non-linear optimization of $L(g, p|a, b, c)$. Using the known value of the density c, we define the function

$$h_\phi(z) = e^{-g_\phi(z)/c}, \tag{4.20}$$

with g_ϕ the DBP computed from the measured projections $p(\theta, s)$ using Eq. 4.5, and estimate the parameters a and b by minimizing

$$\begin{aligned}\Psi_\phi(a, b) \;&=\; \int_{-w}^{w} \left(h_\phi(z) - \frac{b-z}{z-a}\right)^2 (z-a)^2 \, dz \\ &+\; \int_{-w}^{w} \left(h_\phi^{-1}(z) - \frac{z-a}{b-z}\right)^2 (b-z)^2 \, dz \\ &+\; \frac{2w}{c^2} \, \beta \left(p(0, \phi - \pi/2) - c(b-a)\right)^2 .\end{aligned} \tag{4.21}$$

This is a least-square fit with weighting factors $(z-a)^2$ and $(b-z)^2$. These weights are not optimal in terms of noise but are chosen for simplicity because they lead to a quadratic cost function (4.21), and therefore to a closed form solution. The parameter $\beta \geq 0$ in Eq. (4.21) determines the weight given to the ray-sum data $p(0, \phi - \pi/2)$. Typically one would select a small value of β when the uncertainty on $p(0,\phi-\pi/2)$ is large compared to the uncertainty on the DBP $g_\phi(z)$ (i.e. when the parameter σ_p in the stability study of section 6 is large). The parameters $\hat{a}$ and $\hat{b}$ that minimize $\Psi_\phi(a,b)$ are the solutions of the 2×2 linear system

$$\begin{pmatrix} \int_{-w}^{w}(h_\phi^2(z)+1)\,dz+2w\,\beta & \int_{-w}^{w}(h_\phi(z)+h_\phi^{-1}(z))\,dz-2w\,\beta \\ \int_{-w}^{w}(h_\phi(z)+h_\phi^{-1}(z))\,dz-2w\,\beta & \int_{-w}^{w}(h_\phi^{-2}(z)+1)\,dz+2w\,\beta \end{pmatrix}\begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix}=\begin{pmatrix}q_a\\ q_b\end{pmatrix}, \tag{4.22}$$

with

$$\begin{pmatrix}q_a\\ q_b\end{pmatrix}=\begin{pmatrix} \int_{-w}^{w} z\,(h_\phi^2(z)+1+h_\phi(z)+h_\phi^{-1}(z))\,dz-\frac{2w}{c}\,\beta\,p(0,\phi-\pi/2) \\ \int_{-w}^{w} z\,(h_\phi^{-2}(z)+1+h_\phi(z)+h_\phi^{-1}(z))\,dz+\frac{2w}{c}\,\beta\,p(0,\phi-\pi/2) \end{pmatrix}. \tag{4.23}$$

In summary, the DBP based algorithm for star shaped objects (DBPS method), applied in Section 4.9 for Problem 1 consists of the following steps:

1. Estimate the derivative $\partial p(s,\theta)/\partial s$ (we use a two-point difference estimate with a half-sample shift).
2. Using (4.5), backproject on a family of radial lines to obtain the DBP $g_\phi(z), |z|\leq w, 0\leq\phi<\pi$.
3. For noisy data, smooth $g_\phi(z)$ by applying a Gaussian filter along the angular variable ϕ.
4. For each radial line $\mathcal{L}_\phi$:
    - Fit a polynomial to $g_\phi(z)$ to estimate $g'_\phi(0)$ and $g''_\phi(0)$ (we empirically found that a polynomial of degree 5 is sufficient to accurately describe $g_\phi(z)$ for $|z|\leq w$).
    - Apply Eq. (4.11) and (4.12) to obtain an estimate $\hat{c}_\phi$ of the object density.
5. Calculate the average density estimate $\hat{c}=1/\pi\int\hat{c}_\phi\,d\phi$.
6. Using $\hat{c}$ as the true density, calculate the function h_ϕ in (4.20) and solve for each radial line the system (4.22) to obtain an estimate of $a=-u(\phi+\pi)$ and $b=u(\phi)$.

For Problem 2, the density is known beforehand and the same algorithm is used, skipping steps 4 and 5.

4.9 Numerical examples with simulated data

Figure 4.4: *The simulated star object.*

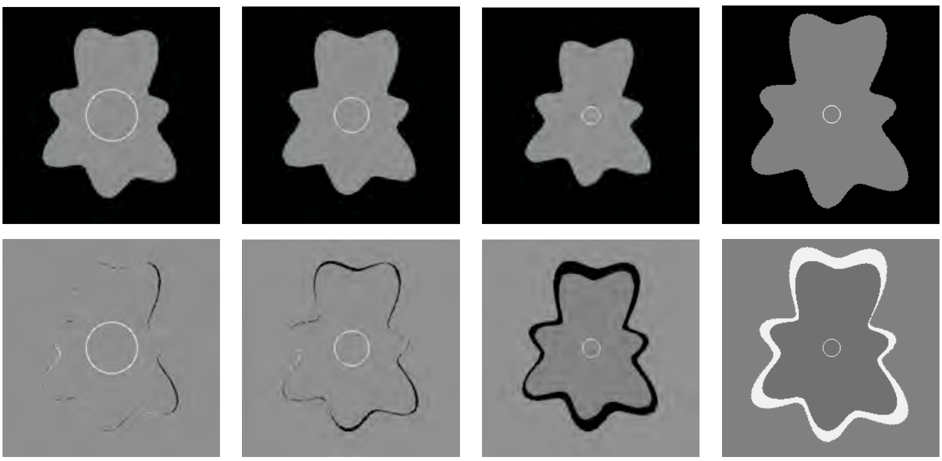


Figure 4.5: *DBPS reconstructions from noise-free data using the algorithm of Section 4.8. The object density is unknown. The first three columns correspond from left to right to FOV diameters $N_z = 60, 40, 20$, with estimated densities equal to $\hat{c} = 1.006, 0.997$ and 0.892 respectively, using $\beta = 0$. In the fourth column, $N_z = 20$ and $\hat{c} = 0.892$ are the same as in the third column, but $\beta = 0.2$. The interior FOV used for reconstruction is shown as a superimposed white circle. Upper row: the reconstructed star object. Bottom row: difference between the reconstruction and the true object of figure 4.4.*

The star object in figure 4.4 has a density $c = 1$ and a radial function

$$u(\phi) = 40\,(2 + 0.4\cos(2\phi) + 0.3\sin(3\phi + \pi/3) - 0.33\cos(7\phi - \pi/6))\,. \qquad (4.24)$$

We generated a digital image of this object on a 1024×1024 matrix with pixel size 0.25. Parallel projections $p(s, \theta)$ were calculated for 256 uniformly spaced angular samples over $[0, \pi)$ and with radial sampling $\Delta s = 1.0$, by forward projecting the 1024×1024 image using Joseph's algorithm with linear interpolation [18]. The object was reconstructed on a 256×256 matrix with pixel size 1.0. We considered both the case of unknown and of known density. The accuracy of the reconstructions was quantified by the ratio of the area of the symmetric difference between the estimated object and true object, and the area of the true object

$$\epsilon = \frac{\text{Area } \{\text{supp}(\hat{f}) \triangle \text{supp}(f)\}}{\text{Area } \{\text{supp}(f)\}}, \tag{4.25}$$

where $\text{supp}(f)$ denotes the support of object function f. The error ϵ varies in the range $[0, \infty)$, where $\epsilon = 0$ represents the ideal case and $\epsilon = 1$ the case when $\hat{f}$ is zero everywhere.

For the DBPS algorithm, $g_\phi(z)$ was calculated for the same 256 uniformly spaced angular samples over $\phi \in [0, \pi)$, and for N_z values of z uniformly sampled over $[-w, w]$, with sampling distance $\Delta z = 1.0$ so that $w = N_z/2$.

4.9.1 Problem 1: unknown density

The object with unknown density is reconstructed using the 2-step DBPS algorithm. The first three columns in figure 4.5 show the reconstructed images for different values of the FOV diameter, $N_z = 60, 40, 20$, using $\beta = 0$ and without angular filtering. The FOV circular boundary has been superimposed on the reconstruction. The difference with the true object illustrates the expected degradation of the algorithm accuracy with a decreasing FOV. Even though the simulated data were noise-free for this example, discretization as well as numerical round off errors have an effect similar to that of random noise, and the results in figure 4.5 illustrate the poor stability of the DBPS algorithm when the FOV is small. The density c estimated by the algorithm was respectively $\hat{c} = 1.006$, $\hat{c} = 0.997$ and $\hat{c} = 0.892$ for $N_z = 60, 40$ and 20. The ratio ϵ was respectively equal to 0.019, 0.047 and 0.233.

From the images in third column of figure 4.5, one notices that the values $-\hat{a}$ and $\hat{b}$ are underestimated if $\hat{c}$ is underestimated. For an explanation we refer to figure 4.3 and the corresponding discussion in Section 4.7, which show that there is a positive correlation between $-\hat{a}$ and $\hat{b}$ on one hand, and $\hat{c}$ on the other hand when σ_p is large, which, as noticed above, corresponds to small values of β. To illustrate the impact of the parameter β, the images in the fourth column of figure 4.5 show the reconstruction for $N_z = 20$ with $\beta = 0.2$. In this case, the underestimation of $\hat{c} = 0.892$ is paired with an overestimation of $-\hat{a}$ and $\hat{b}$, as expected from the fact that the correlation in figure 4.3 is negative for small σ_p.

To illustrate the stability of the algorithm, pseudo-random Poisson noise was added

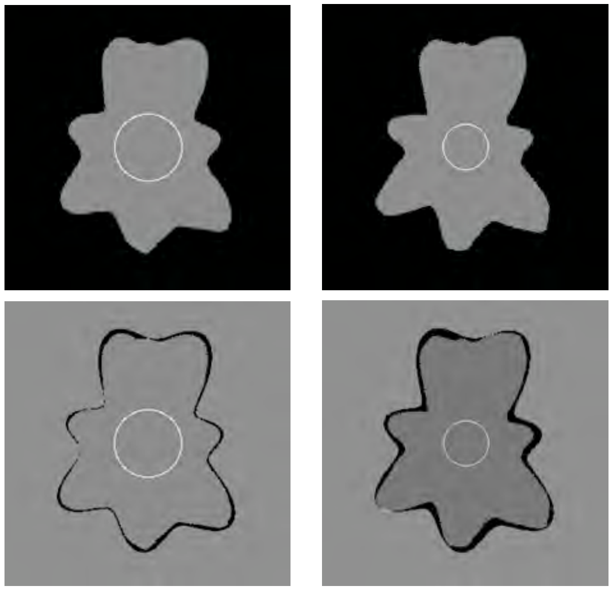


Figure 4.6: DBPS reconstructions from noisy data using the algorithm of Section 4.8. The density is unknown. The columns correspond to FOV diameters $N_z = 60, 40$ with estimated densities equal to $\hat{c} = 0.973$ and 0.812 respectively. Upper row: the reconstructed star object (FOV superimposed). Bottom row: difference between the reconstruction and the true object of figure 4.4.

to the sinogram, corresponding to a total of one million photons; this resulted in a relative standard deviation of 0.005 for the largest sinogram sample. Figure 4.6 shows the reconstructions for $N_z = 60$ and $N_z = 40$, done with a filtering of the DBP data with an angular Gaussian filter of FWHM equal to 10 samples and with $\beta = 0.05$. The error ratio is $\epsilon = 0.076$ and 0.120 for $N_z = 60$ and $N_z = 40$ respectively. For the small FOV ($N_z = 20$) the DBPS reconstruction with noisy data and unknown density failed.

With the examples in figures 4.5 and 4.6, the density c was underestimated by the DBPS algorithm when the field-of-view was small or when the data were noisy. The bias is due to the non-linearity of the system of Eq. (4.10,4.11,4.12). The sign of the bias is determined by the curvature of the solution $\hat{c} = C(g'_\phi(0), g''_\phi(0), p(0, \phi-\pi/2))$ to Eq. (4.10,4.11,4.12). In our example the dominant eigenvalues of the Hessian matrix of the function C are negative for most lines $\mathcal{L}_\phi$, leading to the observed negative bias.

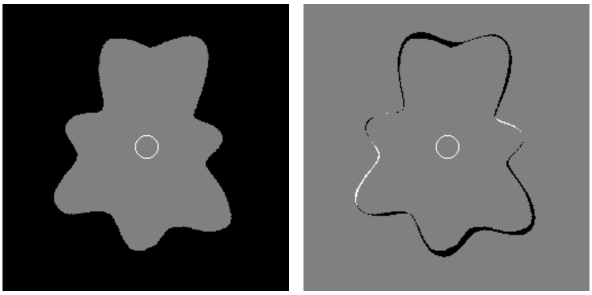


Figure 4.7: DBPS reconstructions from noise-free data using the algorithm of Section 4.8 with prior knowledge of the density, for the FOV diameter $N_z = 20$. Left: the reconstructed star object (FOV superimposed). Right : difference between the reconstruction and the true object of figure 4.4.

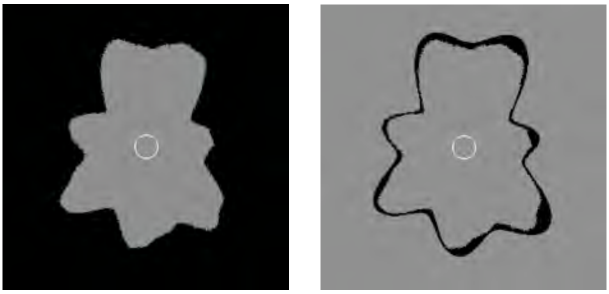


Figure 4.8: DBPS reconstructions from noisy data using the algorithm of Section 4.8 with prior knowledge of the density, for the FOV $N_z = 20$. Left: the reconstructed star object (FOV superimposed). Right : difference between the reconstruction and the true object of figure 4.4.

4.9.2 Problem 2: known density

We reconstructed the same data sets as above, now assuming that the density c is known beforehand. The knowledge of c significantly improves the stability in the presence of noise, as expected from the analysis in Section 4.7. Results are only shown for the smallest FOV ($N_z = 20$). The DBPS reconstruction from noise-free data with $\beta = 0$ is shown in figure 4.7. The error ratio $\epsilon = 0.064$ is much smaller than the ratio 0.233 found when the density is unknown. The DBPS reconstruction from noisy data with $\beta = 0.05$ is shown in figure 4.8, with error ratio $\epsilon = 0.145$.

4.10 Conclusion

We have shown that a star-shaped 2D object with uniform but unknown density is determined by its integrals along all lines intersecting an interior field-of-view. To the best of our knowledge this uniqueness theorem is new. The uniqueness proof is based on the relation between the backprojection of the derivative of the data and the Hilbert transform of the object along a family of central lines. This proof is constructive and leads to a numerical algorithm, which handles each radial line independently.

This algorithm was applied to simulated and measured x-ray projections to illustrate how the stability of the reconstruction depends on the size of the interior field-of-view and on the presence of noise. One limitation of this work is that only a single noisy data set was studied, but this case-study shows a dramatic degradation of the stability as the radius of the interior field-of-view decreases. This degradation is predicted by the Cramer-Rao lower bound for the variance of the estimated density and shape of the star-object, but a more systematic study will be needed to verify and quantify this property. Stability was improved to some extent by using a two-step algorithm and by filtering the noisy data in the angular variable. Additional regularization might be achieved, for instance by improving the calculation of the data derivative in the first step of the algorithm. However an optimal stability requires a global two-dimensional approach, which avoids separating the reconstruction into a set of one-dimensional reconstructions along central lines. This is illustrated in chapter 5 using the *Discrete Algebraic Reconstruction Technique* (DART).

The uniqueness result in Section 4.5 is derived and illustrated for a 2D star-shaped object, but the approach could be extended by noting that the DBP reduces the reconstruction problem to the inversion of the truncated Hilbert transform along a family of central lines. This reduction is valid for arbitrary objects and also for the 3D x-ray transform, and we therefore conjecture that uniqueness might also hold for uniform objects with more complex shapes. Stability, however, is likely to worsen with increasing complexity, which will be illustrated in Chapter 5. A thorough investigation of these generalizations will be the subject of future work.

Bibliography

[1] R. J. Gardner, *Geometric tomography, 2nd edition*, Cambridge University Press, 2006.

[2] R. J. Gardner, “Geometric tomography,” *Notices of the AMS*, vol. 42, pp. 422–429, 1995.

[3] A. Volčič, “A three-point solution to Hammer’s x-ray problem,” *Journal of the London Mathematical Society*, vol. 34, pp. 349–359, 1986.

[4] F. Natterer and F. Wübbeling, *Mathematical methods in image reconstruction*, Monographs on Mathematical Modeling and Computation. SIAM, 2001.

[5] P. Maass, “The Interior Radon Transform,” *SIAM Journal of Applied Mathematics*, vol. 52, pp. 710–724, 1992.

[6] Y. Ye, H. Yu, Y. Wei, and G. Wang, “A general local reconstruction approach based on a truncated Hilbert transform,” *International Journal of Biomedical Imaging*, p. ID 63634, 2007.

[7] H. Kudo, M. Courdurier, F. Noo, and M. Defrise, “Tiny a priori knowledge solves the interior problem in computed tomography,” *Physics in Medicine and Biology*, vol. 53, pp. 2207–2231, 2008.

[8] D.C. Solmon, “The X-ray transform,” *Journal of Mathematical Analysis and Applications*, vol. 56, pp. 61–83, 1976.

[9] A.G. Ramm and A.I. Katsevich, *The radon transform and local tomography*, Boca Raton, FL: CRC Press, 1996.

[10] M. Courdurier, *Restricted measurements for the x-ray transform*, Ph.D. thesis, University of Washington, 2007.

[11] F. Noo, R. Clackdoyle, and J. D. Pack, “A two-step Hilbert transform method for 2D image reconstruction,” *Physics in Medicine and Biology*, vol. 49, pp. 3903, August 2004.

[12] Y. Zou, X. Pan, and E.Y. Sidky, “Image reconstruction in regions-of-interest from truncated projections in a reduced fan-beam scan,” *Physics in Medicine and Biology*, vol. 50, pp. 13–28, January 2005.

[13] F.G. Tricomi, *Integral Equations, vol. 5*, Dover 1985, 1957.

[14] Y. Zou and X Pan, “Exact image reconstruction on PI-lines from minimum data in helical cone-beam CT,” *Physics in Medicine and Biology*, vol. 49, pp. 941–959, February 2004.

[15] J. D. Pack, F. Noo, and R. Clackdoyle, “Cone-beam reconstruction using the backprojection of locally filtered projections,” *IEEE Transactions on Medical Imaging*, vol. 24, pp. 70–85, 2005.

[16] M. Courdurier, F. Noo, M. Defrise, and H. Kudo, “Solving the interior problem of computed tomography using a priori knowledge,” *Inverse Problems*, vol. 24, pp. 065001, 2008.

[17] H. H. Barrett and K. J. Myers, *Foundation of Image Science*, Wiley, 2004.

[18] P.M. Joseph, “An improved algorithm for reprojecting rays through pixel images,” *IEEE Transactions on Medical Imaging*, vol. 1, no. 3, pp. 192–196, 1982.

Chapter 5

Reconstruction of piecewise uniform objects from truncated data

5.1 Introduction

The differential backprojection method for truncated data proposed in Chapter 4 transforms the 2D inverse problem into a set of 1D problems along radial lines $\mathcal{L}_\phi$. This separation is numerically efficient but does not optimally exploit the data, especially in the presence of noise. This limitation is only partially overcome by the two-step approach in the DBPS algorithm of Section 4.8.

Algorithms that treat the 2D inverse problem as a whole, instead of transforming it into a set of 1D lines are expected to yield a better stability in the presence of noise. An example of such a method that exploits the 2D information is the iterative discrete algebraic reconstruction technique (DART) [1]. In each iteration, DART thresholds the image and then updates solely the pixels at the boundaries of the piecewise constant areas. This approach requires the attenuation values or the 'densities' to be known in advance.
Alternatively, some approaches encourage rather than strictly enforce piecewise uniformity. This can be done using penalized maximum-likelihood algorithms for image reconstruction, with various types of penalties that favor sparse solutions in some appropriate basis. For instance Candes et al [2], Sidky et al [3] and Herman and Davidi [4] have shown that the total-variation penalty allows accurate reconstructions of fairly complex piecewise objects from a very small number of projections.

In the first part of this chapter, the stability of the reconstruction from star-shaped objects using such an approach that handles the whole dataset simultaneously, is experimentally investigated. The experiments are performed using the DART reconstruction technique, and it is assumed that the object density is known beforehand.

In Chapter 4, it was suggested that the uniqueness result for binary star shaped objects from interior data can be generalized to more complex objects. The DBPS method, proposed in Chapter 4, cannot be applied for such objects since it explicitly assumes star-shaped object functions. The 2D approaches described above, however, do not make any specific assumption about the shape of the reconstructed object, although the obtained accuracy is expected to depend on the type of object and the noise. The second part of this chapter aims to experimentally indicate the extent of object complexity and number of densities for which the inverse truncated Radon transform from truncated data can be recovered in practice. The study is performed using DART as reconstruction method, considering the densities as prior knowledge.

This chapter is organized as follows. In Section 5.2, the DART algorithm is described and discussed. In Section 5.3, DART is applied and compared with the proposed DBPS method for star-shaped phantom data and real X-ray data of a diamond. Section 5.4 contains experiments in which DART is applied for non-star shaped objects containing one or multiple densities, and discusses the results. Finally, conclusions are drawn in Section 5.5.

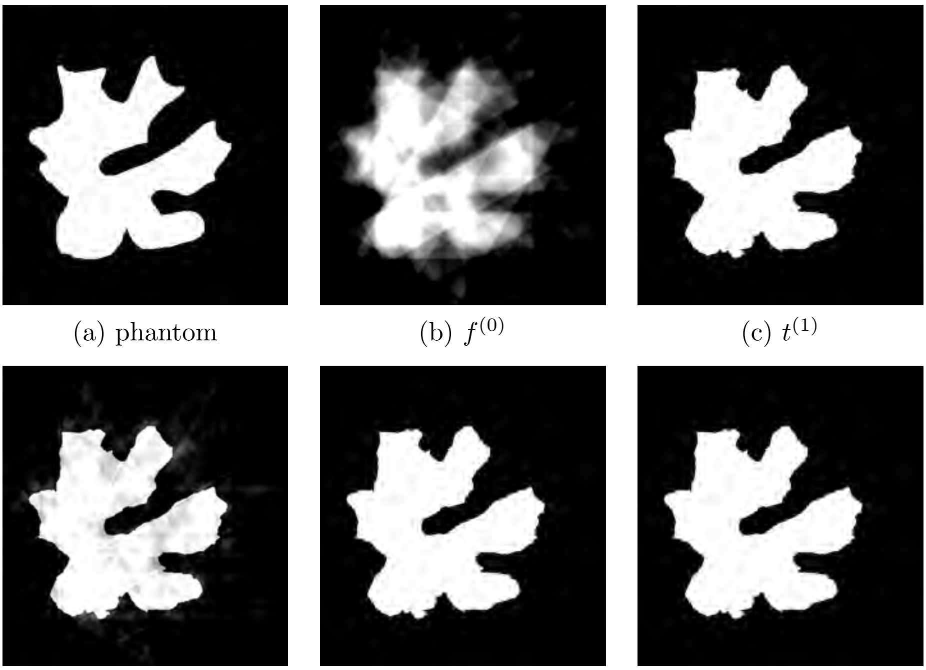
5.2 The discrete algebraic reconstruction technique (DART)

DART is an iterative algebraic reconstruction method for the reconstruction of piecewise uniform objects. It was originally designed for the reconstruction of piecewise uniform objects from data consisting of a very small number of projections and/or a limited angular projection range [1], and has been successfully applied for several reconstruction problems, for example in material science [5–7]. In this chapter, we investigate the new application of DART for the reconstruction of piecewise uniform objects from interior data.

5.2.1 Approach

Consider the problem of reconstructing a binary 2D object function $f(\mathbf{x})$ such as the phantom in Fig. 5.1(a) from a very small number (six) of projections. This corresponds to solving a system of linear equations with a large number of unknowns (the pixel values) and an insufficient number of equations. Since the solution from this set of linear equations is not uniquely determined, SIRT converges to the solution closest to the initial guess (see Eq. (1.40)), which most likely does not correspond to the true object function. This SIRT reconstruction is shown in Fig. 5.1(b).

Assume that the density c of the binary object is known beforehand such that $f(\mathbf{x}) \in \{0, c\}$ for all $\mathbf{x}$. An intuitive approach to enforce a piecewise uniform reconstruction could be the successive application of a SIRT reconstruction alternated with an image thresholding operation, which we call the *SIRT-t* method. The SIRT-t method is initialized with a SIRT reconstruction $f^{(1)}$ from the limited



(a) phantom (b) $f^{(0)}$ (c) $t^{(1)}$

(d) $f^{(1)}$ (e) $t^{(2)}$ (f) $t^{(3)}$

Figure 5.1: *Overview of successive SIRT-t steps for the reconstruction of the binary phantom in (a) from 6 complete projections.*

data. In each iteration $n = 1, 2, \ldots$ of the SIRT-t method, the following steps are executed:

1. The current reconstruction $f^{(n)}$ is segmented using a simple threshold method, forming the image $t^{(n)}$:

$$t^{(n)}(x,y) = \begin{cases} c & \forall (x,y) \in \Omega : f^{(n)}(x,y) \geq \frac{c}{2} \\ 0 & \forall (x,y) \in \Omega : f^{(n)}(x,y) < \frac{c}{2} \end{cases} . \tag{5.1}$$

2. The thresholded image is used as the input for a new SIRT reconstruction, yielding an updated image $f^{(n+1)}(x,y)$.

Fig. 5.1 shows successive steps in the SIRT-t method: (b)-(f) denote the functions $f^{(0)}$, $t^{(1)}$, $f^{(1)}$, $t^{(2)}$ and $t^{(3)}$, respectively. Consider the residual error $E(t^{(1)})$ (Eq. (1.38)) of the thresholded image $t^{(1)}$. If the majority of the pixels in the image $t^{(1)}$ is assigned the correct value, as is the case in Fig. 5.1(c), this residual error is typically small. During the subsequent SIRT update, this small residual error is then redistributed over all pixels in the image domain. Consequently, the pixel values in the resulting image $f^{(2)}$ (Fig. 5.1(d)) have only slightly changed compared to the thresholded image $t^{(1)}$ (Fig. 5.1(c)), typically insufficiently to induce a switch of the assigned density value by the following thresholding operation. Indeed, by comparing Fig. 5.1 (c), (e) and (f) it can be observed that the thresholded image remains equal throughout successive iterations, and that iterative alternation between a SIRT update and a threshold operation is not useful.

(a) (b) (c)

Figure 5.2: *Subsequent steps in the DART method. (a) SIRT reconstruction on boundary pixels using the thresholded SIRT image in Fig. 5.1 (c) as input image; (b) image acquired after merging image (a) with the non-boundary pixel image, and thresholding. (c) final image after 30 DART iterations*

Instead of redistributing the small residual error of the thresholded image $t^{(n)}$ over the complete image, the key idea of DART is to redistribute this error only on the pixels at the boundary of a segment. Using this approach, the number of unknowns for the linear system with small number of equations, is reduced. Intuitively, this means that the residual error is concentrated at the boundary pixels, rather than being diffused over all pixels (such as in the SIRT-t method). Consequently, the subsequent thresholded image typically differs at the boundary pixels from the thresholded image in the previous iteration. In the subsequent DART iteration, all pixels are re-classified in boundary and non-boundary pixels, and the whole procedure is repeated again. Fig. 5.2 illustrates one iteration of the DART procedure. Fig. 5.2 (a) depicts the SIRT update of the boundary pixels, using Fig. 5.1 (c) as start image. This updated boundary is then merged with the image of the non-boundary pixels, and thresholded afterwards, yielding the image Fig. 5.2 (b). After 30 iterations, the reconstructed image, shown in Fig. 5.2 (c), is identical to the true object.

Note that if the estimation of the non-boundary pixels is incorrect, which is often the case (also in the shown example), the system of linear equations, where only the boundary pixels are assumed unknown, does not necessarily have a solution, and convergence of the algorithm is not guaranteed. Nevertheless, the heuristic algorithm proves to be very valuable in practical situations.

5.2.2 Method

We describe our implementation of DART. For other variants of this algorithm and more details on the underlying algorithmic ideas we refer to [1].

A high level flow chart of DART is shown in figure 5.3. DART relies on an underlying reconstruction algorithm for continuous tomography, which is repeatedly used as a subroutine. In our implementation, SIRT (see e.g. section 5.3 in [8]) is used as the continuous method. For simplicity's sake, the DART method is described below for binary objects, but it is trivially extensible to general piecewise uniform objects containing multiple densities.

First, an initial reconstruction $f^{(1)}$ is computed using SIRT. Subsequently, several DART-iterations are performed. In each iteration $n = 1, 2, \ldots$, following steps are executed:

1. The current reconstruction $f^{(n)}$ is segmented using a simple threshold method, forming the image $t^{(n)}$:

$$t^{(n)}(x,y) = \begin{cases} c & \forall (x,y) \in \Omega : f^{(n)}(x,y) \geq \frac{c}{2} \\ 0 & \forall (x,y) \in \Omega : f^{(n)}(x,y) < \frac{c}{2} \end{cases} . \tag{5.2}$$

2. The imaging region Ω is divided into two disjoint subregions: boundary and non-boundary pixels. Let $M(x, y)$ denote the set of pixels that are completely

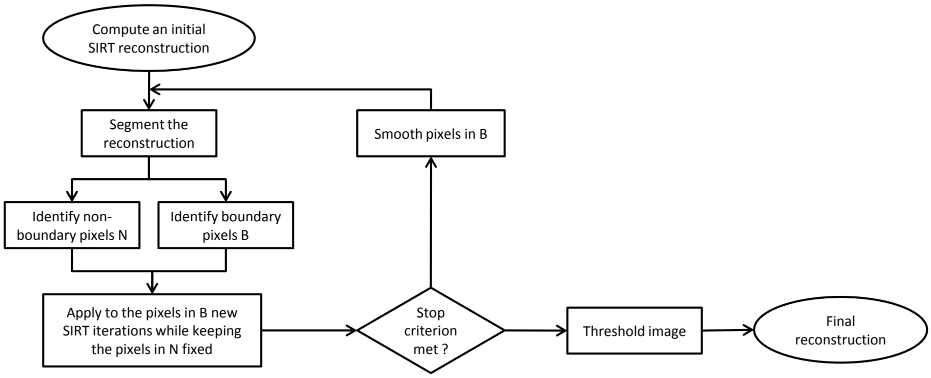


Figure 5.3: *Flow chart of the DART algorithm.*

contained within a circle of radius r_n centered at $(x, y) \in \Omega$ [1]. The set $B_1^{(n)}$ of boundary pixels is defined as the set of all pixels (x, y) such that the segmented image $t^{(n)}$ is not constant within the neighbourhood $M(x, y)$ of (x, y):

$$B_1^{(n)} = \Big\{(x, y) \in \Omega \,|\, \exists (x', y') \in M(x, y) : t^{(n)}(x', y') \neq t^{(n)}(x, y)\Big\}. \qquad (5.3)$$

To increase robustness for noise data inconsistencies, and to ensure that potential holes in the object function are found, the set of boundary pixels is extended by the set $B_2^{(n)}$ which consists of a certain percentage ς of randomly selected pixels in the area $\Omega \backslash B_1^{(n)}$. The eventual set of boundary pixels $B^{(n)}$ is then defined as $B^{(n)} = B_1^{(n)} \bigcup B_2^{(n)}$. The set of non-boundary pixels is defined as $N^{(n)} = \Omega \backslash B^{(n)}$.

3. The boundary and non-boundary pixels are now processed separately:

   - The non-boundary pixels are kept fixed at their thresholded values, yielding an image $t_{N^{(n)}}^{(n)}(x, y) = t^{(n)}(x, y)\, \mathbf{1}_{N^{(n)}}(x, y)$, where $\mathbf{1}_{N^{(n)}}(x, y)$ denotes the indicator function of $N^{(n)}$.
   - Exclusively on the set of boundary pixels, several SIRT iterations are performed. To this end, the 2D Radon transform of $t_{N^{(n)}}^{(n)}(x, y)$ is calculated and subtracted from the projection data p which yields modified projection data $p^{(n)}(s, \theta) = p(s, \theta) - (\mathcal{R}t_{N^{(n)}}^{(n)})(s, \theta)$ for $|s| \leq w, 0 \leq \theta < \pi$. The boundary pixel values are then updated by applying SIRT to the data $p^{(n)}$, yielding an image $h_{B^{(n)}}^{(n)}$ with support in the region $B^{(n)}$.

The image $h_{B^{(n)}}^{(n)}$ is merged with the image $t_{N^{(n)}}^{(n)}$, forming the image $h^{(n)}$.

[1] Typically, this radius is set $r_n = 3\sqrt{2}/2$ pixel units such that $M(x, y)$ consists of the eight neighbouring pixels to pixel (x, y)

4. As a means of regularization, the image $h^{(n)}$ is blurred using a simple weighted sum, forming $f^{(n+1)}$:

$$f^{(n+1)}(x,y) = b\,h^{(n)}(x,y) + (1-b) \sum_{(x',y')\in M(x,y)} \frac{h^{(n)}(x',y')}{|M(x,y)|}\,, \tag{5.4}$$

   with weight factor b such that $0 < b < 1$, and with $|M(x,y)|$ the cardinality of $M(x,y)$. The accuracy is not very dependent on the choice of b. It mainly affects the smoothness of the boundary. In our work, typically $b = 0.7$ is selected.

After $T = 1000$ DART-iterations, the algorithm terminates and the current reconstruction $f^{(1001)}$ is thresholded, forming the binary constant reconstruction $t^{(1001)}$.

The importance of an additional random selection of boundary pixels is illustrated in Fig. 5.4 for the DART reconstruction from 10 projections of a phantom containing small holes. Two DART reconstructions were performed, with $\varsigma = 0\%$ (Fig. 5.4(b)) and $\varsigma = 2\%$ (Fig. 5.4 (c)), respectively. The images show that the small holes, which are not detected if $\varsigma = 0\%$, are recovered if random pixels are selected. Note that DART ultimately becomes equivalent to the SIRT-t method if $\varsigma = 100\%$. In this chapter, ς is systematically set at $\varsigma = 2\%$.

(a) phantom (b) $\varsigma = 0\%$ (c) $\varsigma = 2\%$

Figure 5.4: *DART reconstructions from 10 projections of a phantom containing small holes.*

5.3 Results for star-shaped objects

5.3.1 Reconstruction from simulated data

We re-use the digital image and the simulated Radon data introduced in Section 4.9. We consider only the case of known density, since this knowledge is required by the DART algorithm. The accuracy ϵ of the reconstructions is given by Eq. (4.25). The truncated data consist of N_z values of z uniformly sampled over $[-w, w]$, with sampling distance $\Delta z = 1.0$ such that $w = N_z/2$.

The DART reconstructions from noise-free and noisy data with $N_z = 20$ are shown in figures 5.5 and 5.6. The accuracy ϵ of the DART and DBPS reconstructions is listed in Table 5.1. Note the considerable improvement compared to the DBPS algorithm in the presence of noise.

ϵ	DBPS	DART
No noise	0.014	0.013
Noise	0.026	0.020

Table 5.1: *Comparison of the image error ϵ for reconstructions using the DBPS and DART, from noise-free and noisy data.*

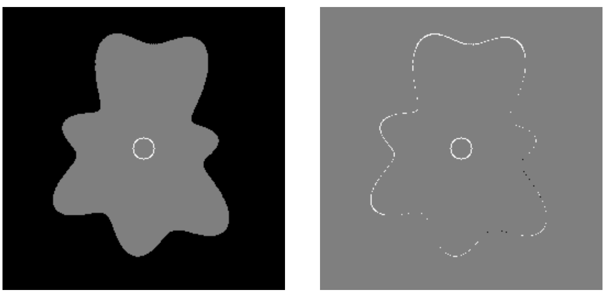


Figure 5.5: *DART reconstruction from noise-free data (see 4.9) of the star-shaped phantom shown in Fig. 4.4, with prior knowledge of the density and a FOV $N_z = 20$. Left: the reconstructed star object (FOV superimposed). Right : difference between the reconstruction and the true object.*

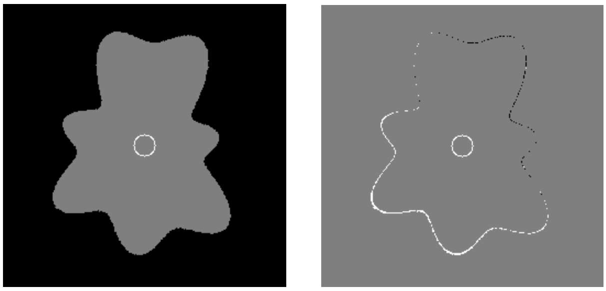
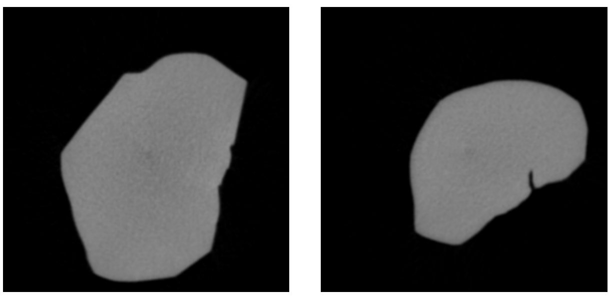


Figure 5.6: *DART reconstruction from noisy data (see 4.9) of the star-shaped phantom shown in Fig. 4.4, with prior knowledge of the density and a FOV* $N_z = 20$*. Left: the reconstructed star object (FOV superimposed). Right : difference between the reconstruction and the true object.*

5.3.2 Application to real X-ray CT data of diamonds

DiamCad (Antwerp, Belgium), a diamond processing company that performs a detailed study of rough stones, scans rough diamonds to retrieve detailed information on their shapes. Recently, DiamCad encountered the problem that one of the diamonds was too large to be covered by the field of view of the detector, which resulted in truncated projection data for some of the slices. Since diamonds consist of only one material (apart from the impurities) and their shape is fairly simple, the truncated data problem forms a nice application for the proposed uniqueness theorem and DBPS algorithm.



(a) Slice A (b) Slice B

Figure 5.7: *FBP reconstructions from non-truncated X-ray data of a diamond.*

A diamond was scanned at 70 kVp in a Scanco μCT 40 (Scanco Medical, Brüttisellen, Switzerland) with a circular cone beam geometry. The data were

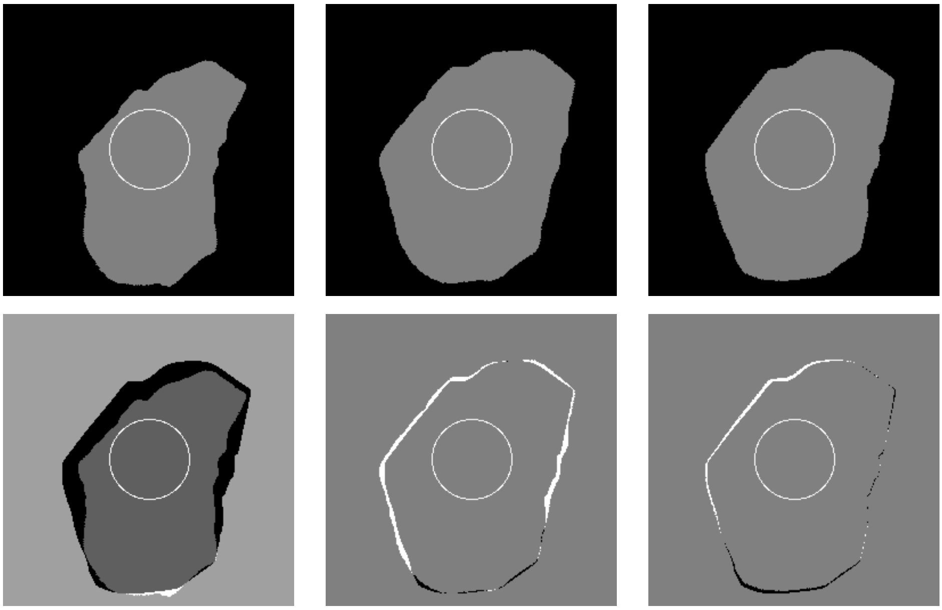


Figure 5.8: *Reconstructions from truncated X-ray data of diamond Slice A with a FOV diameter of $N_z = 70$. The columns correspond from left to right to the DBPS reconstruction with unknown density, and the DBPS and DART reconstructions with known density. The estimated density in the left reconstruction is $\hat{c} = 0.2048$. Upper row: the reconstructed diamond slice (FOV superimposed). Bottom row: difference between the reconstruction and our "ground truth" FBP reconstruction image from complete data in figure 5.7.*

ϵ	Slice A	Slice B
DBPS without density knowledge	0.198	0.126
DBPS with known $c = 0.346$	0.057	0.032
DART with known $c = 0.346$	0.033	0.032

Table 5.2: *Error ratio values ϵ of slices A and B using the DBPS method without known density, and using the DBPS and DART method with known $c = 0.346$.*

recorded at 256 angles in $[0, \pi)$ using a 1024×56 (transaxial$\times$ axial) pixel detector. To cover the full axial length of the object, 500 circular cone beam scans were performed at equally spaced axial positions. The data were linearized using the manufacturer's software to avoid data inconsistencies due to beam hardening. Afterwards, the data were rebinned to parallel beam, yielding a 1024×256 sized sinogram per slice, and then downsampled to 256×256 sinograms. For this results

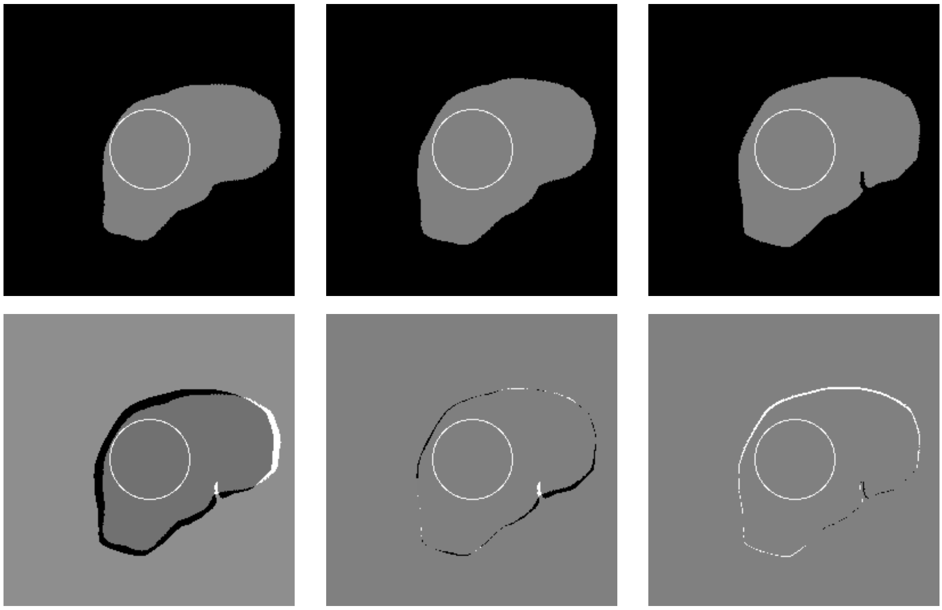


Figure 5.9: *Reconstructions from truncated X-ray data of Slice B with a FOV diameter of $N_z = 70$. The columns correspond from left to right to the DBPS reconstruction with unknown density, and the DBPS and DART reconstructions with known density. The estimated density in the left reconstruction is $\hat{c} = 0.276$. Upper row: the reconstructed diamond slice (FOV superimposed). Bottom row: difference between the reconstruction and our "ground truth" FBP reconstruction image from complete data in Fig. 5.7.*

section, a star shaped slice A and a nearly star shaped slice B of the diamond are selected of which the full sinograms are available. The FBP reconstructions of these slices are shown in figure 5.7. The accuracy of the interior reconstructions is evaluated with respect to ground truth uniform images obtained by performing the histogram based segmentation procedure of Otsu [9] of these FBP reconstructions. This segmentation also yields the density value $c = 0.346$.

The measured sinograms of slices A and B were artificially truncated so as to obtain an interior FOV with diameter $N_z = 70$. Both truncated datasets are reconstructed using the DBPS method without prior knowledge of the density c, and using the DBPS and DART methods assuming that the density c is known beforehand. Figure 5.8 shows the respective reconstructions and their difference images for diamond slice A. Figure 5.9 shows the corresponding reconstructions for slice B. The DBPS reconstructions are performed using $\beta = 0.05$. The error ratios of the reconstructions are given in table 5.2 for both slices A and B.

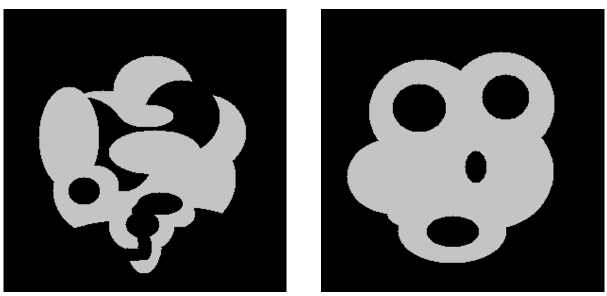
Similar to the observations for the simulated data, the DBPS reconstruction where the density c is known outperforms the DBPS reconstruction method with unknown c for both slices. For slice A, the DART reconstruction with known density provides the smallest error ratio, while for Slice B, the error ratios of the DBPS and DART methods with known density are similar.

5.4 Non-star shaped objects

In Section 4.6, it was suggested that uniqueness is still valid for non-star shaped objects, but the stability will worsen with increasing level of object complexity. In this section, non-star shaped object functions are reconstructed from truncated data. As opposed to the proposed DBPS method, which exclusively reconstructs star-shaped objects, DART does not restrict the class of object shapes on which it can be applied, though the obtained accuracy is likely to depend on the noise and the type of object. To enable proper comparison between the various DART reconstructions, the parameters of DART are kept fixed throughout the whole section.

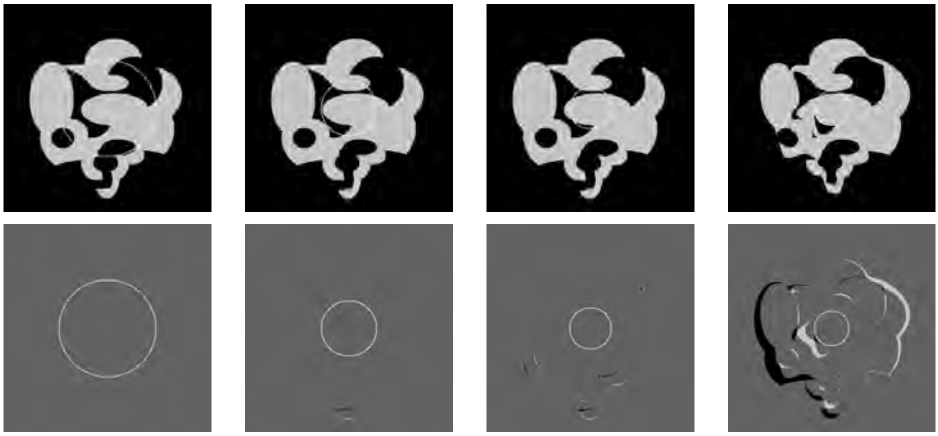
5.4.1 Uniform non-star shaped objects

Fig. 5.10 (a) and (b) depict two uniform but non-star shaped phantoms, Phantom 1 and Phantom 2, respectively. The phantoms are defined on a 300×300 grid. Sinograms are simulated using 300 radial and 180 angular samples. Note that this 'full' dataset actually already contains an insufficient number of equations to determine a unique solution.



(a) Phantom1 (b) Phantom 2

Figure 5.10: *Presentation of two uniform non-star shaped phantoms.*



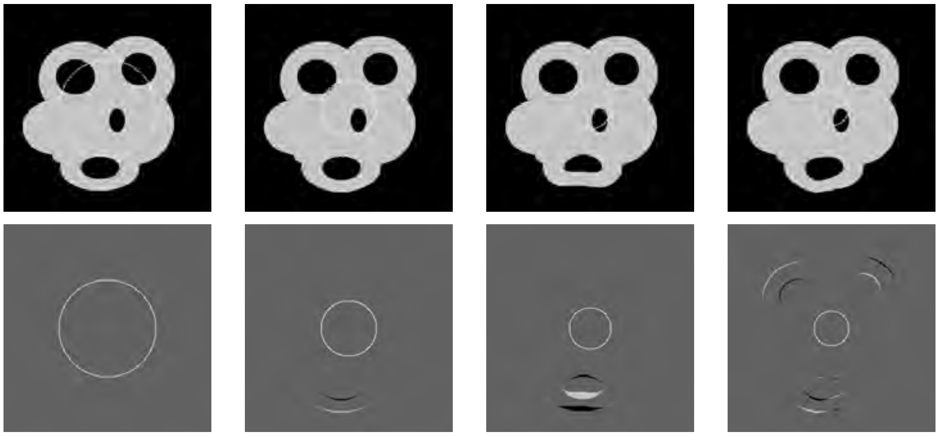
$N_z = 140$	$N_z = 80$	$N_z = 60$	$N_z = 50$
$\epsilon = 0.0012$	$\epsilon = 0.0037$	$\epsilon = 0.0093$	$\epsilon = 0.2019$

Figure 5.11: *DART reconstructions from simulated data of Phantom 1, for varying degrees of truncation. The white circle, imposed on each image, represents the FOV; N_z denotes the diameter of the FOV.*

Truncation is simulated by using exclusively N_z radial samples at the center of each projection. The upper row in Fig. 5.11 and Fig. 5.12 depicts DART reconstructions from truncated data for varying FOV's, for Phantom 1 and Phantom 2, respectively. The second row in Fig. 5.12 and Fig. 5.11 represents the difference images of the DART reconstructions with the ground truth phantom images (Fig. 5.10). Fig. 5.11 and Fig. 5.12 illustrate that accurate reconstructions of uniform non-star objects can be obtained from interior data.

To illustrate the robustness for noise, the previous experiments are repeated using truncated noisy data as the input for DART. These data were obtained by applying Poisson noise (assuming a total of 50000 photon counts) on the intensity image that corresponds to the attenuation sinogram. This represents a significant amount of noise, as demonstrated in Fig. 5.13, where the FBP reconstructions from the noisy data are shown for Phantom 1 and Phantom 2, respectively.

Fig. 5.14 depicts graphs representing the DART reconstruction error ϵ as a function of the FOV radius w for Phantom 1 and Phantom 2, for noise-free and noisy data, respectively. The error curves appear to behave very similar for noise-free and noisy data. Apparently, the considered noise level, which is of the same order of magnitude as the noise encountered in X-ray CT images, does not significantly affect the accuracy of the reconstruction image, although the image accuracy will



$N_z = 140$	$N_z = 80$	$N_z = 60$	$N_z = 50$
$\epsilon = 0.0018$	$\epsilon = 0.0075$	$\epsilon = 0.0274$	$\epsilon = 0.0191$

Figure 5.12: *DART reconstructions from simulated data of Phantom 2, for varying degrees of truncation. The white circle, imposed on each image, represents the FOV;* N_z *denotes the diameter of the FOV.*

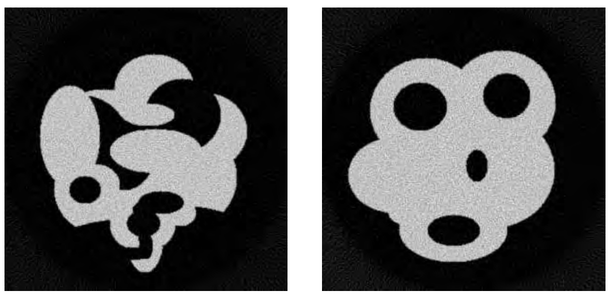


Figure 5.13: *FBP reconstructions of Phantom 1 and Phantom 2 from a non-truncated dataset to which Poisson noise was added.*

inevitably decrease above a certain level of noise.

5.4.2 Objects consisting of two materials

In this subsection, DART reconstructions from truncated data of objects containing two densities are discussed.

The error ϵ as defined in Eq. (4.25) cannot directly be applied for objects with multiple densities. Say $d(x, y) = \hat{f}(x, y) - f(x, y)$, with $\hat{f}$ the DART estimation of

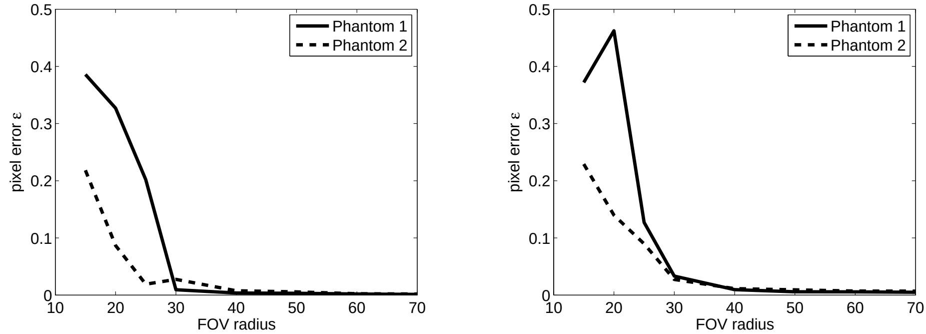
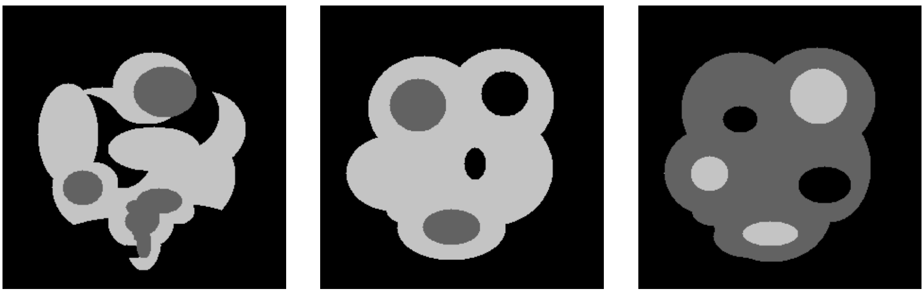


Figure 5.14: Graphs representing the image pixel error of the DART reconstruction with respect to the FOV radius for both phantoms, for noise-free (left) and noisy data (right).



Phantom 3 Phantom 4 Phantom 5

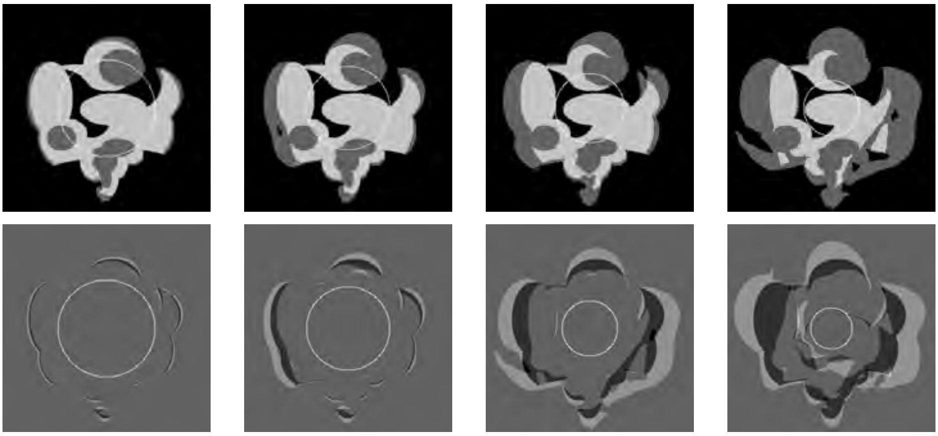
Figure 5.15: Two-density phantoms with densities $t = 0, 0.5, 1$ ($t = 0$ corresponds to the background).

the object function. The accuracy for the reconstruction of objects consisting of multiple known densities is quantified as:

$$\epsilon_d = \frac{\text{Area}\ \{\text{supp}(d)\}}{\text{Area}\ \{\text{supp}(f)\}}, \tag{5.5}$$

where supp(f) denotes the support of object function f. Note that the definition of ϵ_d becomes identical to the definition of ϵ in Eq. (4.25) when considering uniform objects with known density.

Three phantoms are considered: Phantom 3, Phantom 4 and Phantom 5, shown in Fig. 5.15. The DART reconstructions of these three phantoms for varying levels of truncation are shown in Fig. 5.16 for Phantom 3, Fig. 5.17 for Phantom 4, and in Fig. 5.18 for Phantom 5. For Phantoms 3 and 4, it can be seen that the



$N_z = 140$	$N_z = 120$	$N_z = 100$	$N_z = 80$
$\epsilon_d = 0.0580$	$\epsilon_d = 0.1344$	$\epsilon_d = 0.1778$	$\epsilon_d = 0.3749$

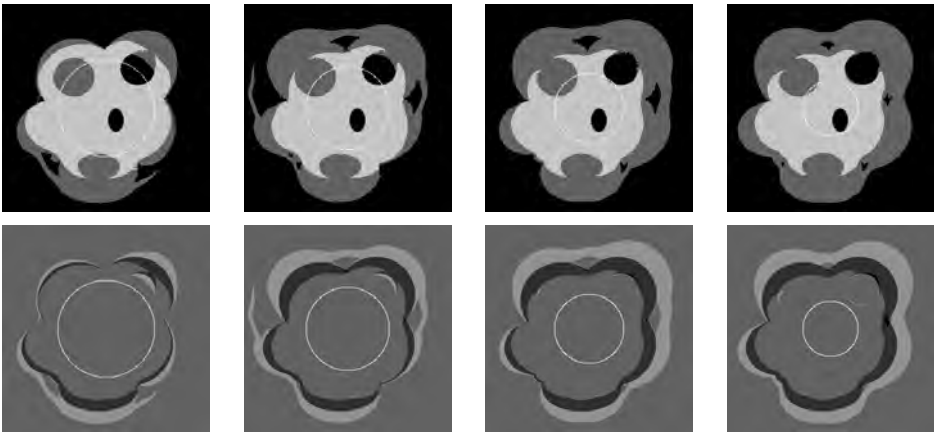
Figure 5.16: *DART reconstructions from simulated data of Phantom 3, for varying degrees of truncation. The white circle, imposed on each image, represents the FOV; N_z denotes the diameter of the FOV.*

reconstruction quality quickly degrades for a decreasing FOV diameter. Clearly, the presence of the second material has a large impact on the image accuracy. On the other hand, although the difference with the one-material is still significant, the reconstruction quality seems reasonably well preserved for Phantom 5.

Below, an intuitive explanation for this different behaviour is given.
Consider an object pixel (x, y) that is classified as a boundary pixel in iteration n such that $(x, y) \in B^{(n)}$. Assume that this pixel has true value $f(x, y) = 1$, but was assigned value $f^{(n-1)}(x, y) = 0$ in the previous iteration. The next step in the DART method is a SIRT reconstruction that is performed exclusively on the set of boundary pixels. Suppose that the image is binary, with two densities $(c_0, c_1) = (0, 1)$. Hence, it is sufficient that the value of the SIRT reconstruction $h^{(n)}(x, y)$ is larger than 0.5 to obtain the correct thresholded value $t^{(n)}(x, y) = 1$. However, suppose that an additional intermediate density $c_1 = 0.5$ is present in the object slice. In that case, it is necessary that $h^{(n)}(x, y) > 0.75$ to obtain the correct value after thresholding. Recall that SIRT, when the system of equations is underdetermined, finds the solution which is the closest to the initial guess (see Eq. (1.40)). The norm in Eq. (1.40) favours smooth solutions with a large number of pixels having small values. Hence, in case of an additional intermediate density, more stringent conditions (and thus more data) are required. This problem is less prominent if the additional density is non-intermediate. In Phantoms 3 and 4, the

additional density is intermediate for the majority of the pixels. The opposite is valid for Phantom 5, which explains the difference in image accuracy.

The experiments were repeated for noisy data of the 2-density phantoms. Poisson noise corresponding to a total number of 5×10^4 photon counts was added to the data. Fig. 5.19 plots the relation of the image pixel error ϵ versus the FOV radius for the three phantoms. The left graph represents reconstructions from the clean sinogram data, while the right image corresponds to the noisy data. These graphs confirm the observations in the previous section, namely that the DART reconstructions are not significantly affected by the noise level that is considered here, although the image accuracy will inevitably decrease if a certain noise level is exceeded.

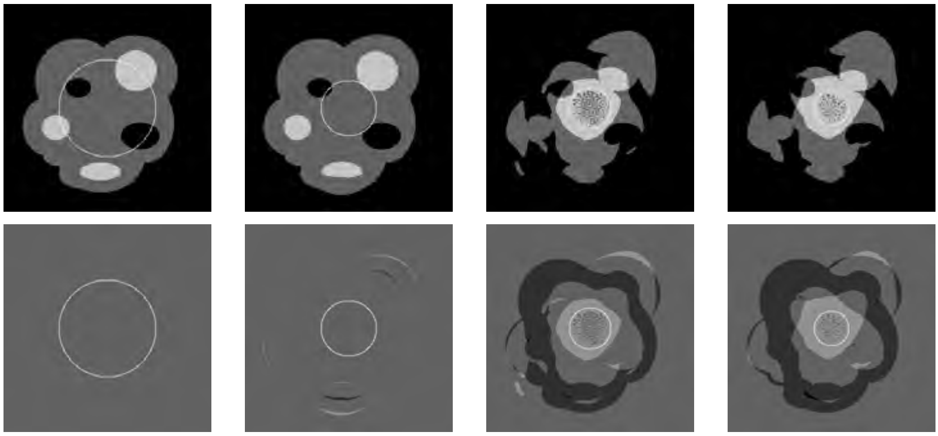


$N_z = 140$	$N_z = 120$	$N_z = 100$	$N_z = 80$
$\epsilon_d = 0.1689$	$\epsilon_d = 0.2837$	$\epsilon_d = 0.3395$	$\epsilon_d = 0.3656$

Figure 5.17: *DART reconstructions from simulated data of Phantom 4, for varying degrees of truncation. The white circle, imposed on each image, represents the FOV; N_z denotes the diameter of the FOV.*

5.5 Conclusion

This chapter concerns the problem of reconstructing general piecewise uniform object functions from truncated data. In the first part of the chapter, we suggested that a reconstruction method that handles the whole truncated data set simultaneously, is preferred above approaches that subdivide a 2D reconstruction in a set of 1D problems as in the DBPS method for star-shaped objects proposed in Chapter



$N_z = 140$ $\epsilon_d = 0.0031$ | $N_z = 80$ $\epsilon_d = 0.0182$ | $N_z = 60$ $\epsilon_d = 0.5662$ | $N_z = 50$ $\epsilon_d = 0.6120$

Figure 5.18: *DART reconstructions from simulated data of Phantom 5, for varying degrees of truncation. The white circle, imposed on each image, represents the FOV; N_z denotes the diameter of the FOV.*

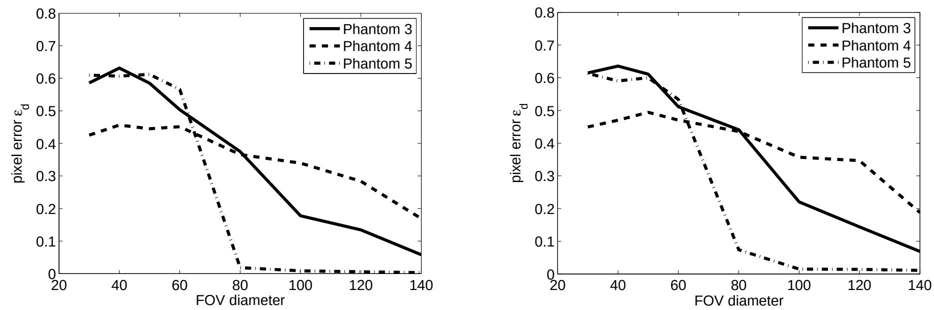


Figure 5.19: *Graphs representing the image pixel error of the DART reconstruction with respect to the FOV radius for both phantoms, for noise-free(left) and noisy data (right).*

4. As an illustration, the discrete algebraic reconstruction method DART was implemented, and the experiments for the star-shaped phantom from noise-free and noisy data from Chapter 4 were repeated using DART. The results show that the stability improves for noisy data if DART is used. Moreover, it was demonstrated for real X-ray data of a diamond that both with the DBPS and the DART method, accurate reconstructions were obtained from significantly truncated sinograms.

The second part of this chapter aims at accurate reconstruction of objects with more complex shapes and multiple densities. In Chapter 4 it was suggested that the uniqueness result for binary star shaped objects from interior data could be generalized to more complex objects. Since the DBPS method that was proposed in Chapter 4 explicitly assumes star-shaped object functions, it cannot be applied for the more complex objects. However, several approaches that treat the 2D inverse problem as a whole, such as DART, do not make any specific assumptions on the shape of the reconstructed object. Experiments were performed using the DART reconstruction technique, aiming at indicating the extent of object complexity for which the inverse problem from truncated data can be recovered in practice. The experiments suggest that also binary objects with relatively complex shapes can still stably be reconstructed from interior data. However, for objects consisting of multiple materials, the stability of the DART reconstruction from interior data is reduced drastically. An interesting direction for future work is to repeat these experiments for other iterative approaches for the reconstruction of piecewise uniform objects from limited data, such as ℓ_1-minimization [2, 3]. In particular, with these methods, the stability can be investigated for the case the densities are unknown beforehand.

Bibliography

[1] K.J. Batenburg and J. Sijbers, "DART: a fast heuristic algebraic reconstruction algorithm for discrete tomography," *Proceedings of ICIP*, vol. 4, pp. 133–136, 2007.

[2] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, pp. 489–509, 2006.

[3] E.Y. Sidky, C.M. Kao, and X Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *Journal of X-ray Science and Technology*, vol. 14, pp. 119–139, 2006.

[4] G.T. Herman and R. Davidi, "Image reconstruction from a small number of projections," *Inverse Problems*, vol. 24, pp. 045011, 2008.

[5] S. S. Bals, K.J. Batenburg, D. Liang, O. Lebedev, G. Van Tendeloo, A. Aerts, J.A. Martens, and C.E.A. Kirschhock, "Quantitative three-dimensional modeling of Zeotile through discrete electron tomography," *Journal of the American Chemical Society*, vol. 131, no. 13, pp. 4769–4773, 2009.

[6] 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.

[7] S. Turner, S.M.F. Tavernier, G. Huyberechts, E. Biermans, S. Bals, K. J. Baten-

burg, and G. Van Tendeloo, "Assisted spray pyrolysis production and characterisation of ZnO nanoparticles with narrow size distribution," *to appear in Journal of Nanoparticle research*, 2009.

[8] F. Natterer and F. Wübbeling, *Mathematical methods in image reconstruction*, Monographs on Mathematical Modeling and Computation. SIAM, 2001.

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

Part IV

Reconstruction of piecewise uniform objects: practice

Chapter 6

Beam hardening artifact reduction for piecewise uniform objects

6.1 Introduction

The algorithms that were used in Chapters 4 and 5 for reconstruction of piecewise uniform objects require that the object has one (or few) uniform densities. This assumption is restrictive because of confounding physical effects such as beam-hardening. This motivates the study in this chapter, in which a new method is presented for the reduction of beam hardening artifacts for piecewise uniform objects. In Chapter 7, the proposed beam hardening method will be applied for limited data problems such as truncation.

This chapter is organized as follows. Section 6.2 consists of a short literature overview on beam hardening reduction methods, and puts the proposed method into its context. A detailed description of the proposed iterative correction procedure is presented in Section 6.3. Section 6.4 contains results after applying the beam hardening correction method on real X-ray CT data of 2- and 3-material hardware phantoms. This section also includes a comparison with the method of Krumm et al. [1], which was simultaneously and independently developed, and plays a similar approach. In Section 6.6, the results are discussed, and conclusions are drawn.

6.2 Context

When a monochromatic X-ray beam traverses a homogeneous object, the attenuation is linearly related to the thickness of the object along that ray (Beer's law). In general, however, CT X-ray sources are polychromatic. The linear relation does not hold for polychromatic beams, since lower energy photons are more easily absorbed than higher energy photons, which causes the beam to 'harden' as it propagates through the object. This non-linear effect is referred to as beam-hardening (BH). If the energy dependence of the absorption is not taken into account, reconstructions are contaminated by cupping and streak artifacts [2]. Beam hardening correction is important in both medical and industrial CT applications to improve the visual quality of the images and to obtain more accurate segmentations, which is necessary for quantitative image analysis.

Beam hardening artifacts have been a subject of research for decades, resulting in a broad variety of artifact reduction strategies. Beam hardening correction methods can roughly be subdivided into four classes : hardware filtering, dual energy, statistical polychromatic reconstruction, and linearization.

- Hardware filtering is a common method to narrow the broad source spectrum. Thin metal plates that are placed between the source and the object absorb the lower energy photons of the beam before the beam enters the object. Although hardware filters reduce the beam hardening artifacts in the resulting image, the lower photon count also results in a decrease of the signal to noise ratio.

- In dual energy methods [2–4], the energy-dependence of the attenuation coefficients is described as a linear combination of two basis functions representing the separate contributions of the photo-electric effect and the scattering. The coefficients of the two basis functions are needed for each image pixel. Therefore, intensity measures from two scans at different source voltages are required. After determining the coefficients, reconstructions of the linear attenuation coefficient can be obtained at any energy within the diagnostic range. The main drawback of the method is that it requires CT-scans at two different, preferably non-overlapping [4], energy spectra. Another drawback is that the spectra are often unknown and therefore accurate calibration experiments are necessary to avoid artifacts in the reconstructions. For these reasons, dual energy is limited to specific applications e.g. for measuring the bone mineral density in the lumbar spine [5].

- Statistical reconstruction of polychromatic data is an approach explored by several authors [6–8]. The statistical beam hardening reduction methods basically incorporate the polychromatic nature of the beam in a maximum likelihood expectation maximization (ML-EM) algorithm, yielding reconstruction images with significantly reduced beam hardening artifacts. This approach assumes that the object consists of N known base substances, and that the energy dependence of the attenuation coefficient for each pixel can be described as a linear combination of the known energy dependencies of the base substances. Statistical methods are generally very flexible with respect to various geometries, prior knowledge, noise statistics, etc. However, such methods are computationally very intensive.

- Linearization methods, which are used standardly for beam hardening suppression, aim at transforming the measured polychromatic attenuation data into monochromatic attenuation data. For homogeneous objects, the beam hardening curve, describing the attenuation-thickness relation of the material, is acquired from a calibration scan or by estimating for each sinogram pixel the corresponding object thickness based on a preliminary reconstruction. In a next step, a parametric model (e.g., polynomial, bimodal energy [9]) is fitted to the beam hardening curve. Using this model, the measured attenuation values can then be corrected for beam hardening.
  For objects containing more than one material, linearization is performed iteratively using a so called *iterative post reconstruction* (IPR) method [2, 10–17]. In IPR methods, a preliminary reconstruction is performed. A segmentation of this image then allows to determine the intersection length of each material along each ray. Given these material thicknesses, a poly- and monochromatic sinogram can be calculated, and the difference between these simulated sinograms is then used as an additive correction to the measured sinogram. Finally, a new image is reconstructed from the corrected sinogram. This procedure is performed iteratively, resulting in an improved segmentation and consequently in improved beam hardening correction. Similarly to the statistical methods, the linearization techniques assume that the object consists of a known number of materials with known energy dependence of the attenuation coefficients. Some methods require uniform materials (e.g. [2, 12, 14]), others allow for mixtures of these base materials (e.g. [15–17]).

An important drawback of the statistical and linearization methods is the large amount of prior information that is required, which often consists of the knowledge of the source spectrum and the material specific energy-dependent attenuation coefficients. In many industrial and non-standard medical applications, the exact material composition is unknown, hence, the required prior knowledge is not available.

Recently, several linearization methods were developed that do not require prior knowledge such as the spectrum and the attenuation coefficients of the materials.

These methods only require the number of materials N to be known and assume that the object consists of *uniform* materials. The methods of Van de Casteele et al. [18], Gao et al. [19] and Mou et al. [20] offer nice results but have some limitations. The first two methods are restricted to a very small class of objects, while the third method is restricted to complete data and it uses a cost function based on adapted Helgason-Ludwig consistency conditions, which are specific for the used fan beam geometry.

In this chapter, an alternative IPR method is proposed that does not require prior spectrum or material knowledge. By using a physical model with a small number of parameters, the polychromatic sinogram is simulated. The source spectrum is parameterized by discretizing the energy range using a small number of energy bins. The unknown parameters, representing the fractional source spectrum intensity for each energy bin and the corresponding attenuation values, are found by minimizing the difference between the measured and the simulated polychromatic sinogram.

6.3 Method

6.3.1 Notations and concepts

Assume a piecewise uniform object that consists of N materials with attenuation coefficients $\mu_n(E)$ that depend on the energy E. Consider a ray path $L(\theta, s)$, which is the straight line defined by the angle θ and the distance s to the center of rotation. To enhance readability, we omit (θ, s) in the equations below. When a monochromatic X-ray beam with intensity I_0 and energy E_0 passes through uniform segments along a ray path L, the monochromatic exit intensity for that specific ray is expressed by the Beer-Lambert law as

$$I_M(L) \quad = \quad I_0 e^{-\sum_{n=1}^{N} \mu_n(E_0) t_n(L)}, \tag{6.1}$$

where $t_n(L)$ denotes the intersection length between path L and material n. The monochromatic attenuation A_M is then defined as the logarithm of the ratio of the input and output intensities, and behaves linearly with respect to the traversed thickness:

$$A_M(L) = \sum_{n=1}^{N} \mu_n(E_0) t_n(L). \tag{6.2}$$

Note that $A_M(L)$ is the desired quantity used by analytic reconstruction algorithms such as Filtered Back Projection (FBP).
In practice, the emitted X-ray photons have different energies E, where $E \in [0, E_{\max}]$ where $E_{\max}$ denotes the maximal energy of the photons emitted by the source. The measured intensity of such a polychromatic beam along a path L that traversed a piecewise uniform object can then be expressed as the sum of the monochromatic contributions for each energy E:

$$I_P^{\text{meas}}(L) = \int_0^{E_{\max}} I_0(E) e^{-\sum_{n=1}^{N} \mu_n(E) t_n(L)} dE, \tag{6.3}$$

where $I_0(E)$ represents the intensity of the emitted source spectrum. The polychromatic attenuation $A_P(L)$ along a path L is given by

$$A_P^{\text{meas}}(L) = \log\left(\frac{I_0}{I_P^{\text{meas}}(L)}\right), \tag{6.4}$$

with I_0 the total incident beam intensity $I_0 = \int_0^{E_{\max}} I_0(E) dE$.

6.3.2 Linearization of the sinogram

The beam hardening problem concerns the estimation of the monochromatic attenuation $A_M(L)$ from the measured polychromatic sinogram $A_P^{\text{meas}}(L)$. We construct a correction formula by noting that the required amount of correction is given by the difference between $A_M(L)$ and $A_P(L)$. Hence, a corrected sinogram $A_M^{\text{cor}}(L)$ is obtained using

$$A_M^{\text{cor}}(L) = A_P^{\text{meas}}(L) + (A_M^{\text{sim}}(L) - A_P^{\text{sim}}(L)), \tag{6.5}$$

where $A_M^{\text{sim}}(L)$ and $A_P^{\text{sim}}(L)$ denote a simulated mono- and polychromatic sinogram, respectively. These mono- and polychromatic sinograms can be simulated using Eq. (6.2) and Eq. (6.4), given an estimate of $t_n(L)$ and provided that $I_0(E)$ and $\mu_n(E)$ are known a priori. In IPR methods, the path lengths $t_n(L)$ are usually estimated based on a segmentation of the preliminary reconstruction image that is obtained from the uncorrected sinogram. Note that, in case of a poor segmentation image and consequently a poor estimation of $t_n(L)$, the corrected sinogram $A_M^{\text{cor}}(L)$ in Eq. (6.5) is not completely linearized. However, the the hope is that the corrected sinogram will yield an improved reconstruction image and segmentation. Hence, an improved estimate of $t_n(L)$ can be used for a new update of the corrected sinogram $A_M^{\text{cor}}(L)$.

In this chapter, the material composition $\mu_n(E)$ and the source spectrum $I_0(E)$ are considered to be unknown, which implies that Eq. (6.2) and Eq. (6.4) cannot be used directly to determine the correction term in Eq. (6.5). We model the X-ray energy spectrum as a discrete set of J energy bins with corresponding intensities $\{I_j\}$. As will be seen below, the corresponding energies $\{E_j\}$ need not to be specified. Using this model, the output polychromatic intensity I_P^{sim} for a path L is given by

$$I_P^{\text{sim}}(L) = \sum_{j=1}^{J} I_j e^{-\sum_{n=1}^{N} \mu_{n,j} t_n(L)}, \tag{6.6}$$

with unknown attenuation coefficients $\mu_{n,j} = \mu_n(E_j)$. The corresponding polychromatic output attenuation A_P^{sim} is

$$A_P^{\text{sim}}(L) = -\log\left(\sum_{j=1}^{J} I_j^F e^{-\sum_{n=1}^{N} \mu_{n,j} t_n(L)}\right), \tag{6.7}$$

where $I_j^F = I_j/I_0$ is the fraction of the total spectrum intensity corresponding to the j^{th} energy bin, with $\sum_j I_j^F = 1$. Given an estimate of the lengths $\mathbf{t} = \{t_n\}$ for each path L, and given the measured intensities $I_P^{\text{meas}}(L)$, the unknown parameters $\mathbf{I}^F = \{I_j^F\}$ and $\boldsymbol{\mu} = \{\mu_{n,j}\}$ can be estimated by minimizing the following cost function:

$$\Phi(\boldsymbol{\mu}, \mathbf{I}^F, \mathbf{t}) = \sum_{L(s,\theta)} \left(\log\left(\frac{I_P^{\text{meas}}(L)}{I_0}\right) - \log\left(\sum_{j=1}^{J} I_j^F e^{-\sum_{n=1}^{N} \mu_{n,j} t_n(L)}\right)\right)^2. \tag{6.8}$$

The number of parameters K to be estimated is $K = (N+1) \times J - 1$. In our experiments, we model the polychromatic behavior of the attenuation data $A_P^{\text{sim}}(L)$ using $J = 3$ energy bins. We observed that a larger number of energy bins does not significantly improve the fit.

The monochromatic linear attenuation $A_M^{\text{sim}}(L)$ can be written as

$$A_M^{\text{sim}}(L) = \sum_{n=1}^{N} \overline{\mu}_n t_n(L), \tag{6.9}$$

where $\{\overline{\mu}_n\}$ is a set of *reference* attenuation coefficients which can be chosen, for instance, using an arbitrary weighting of the attenuation coefficients $\mu_{n,j}$. We expect that the robustness of the method is optimized if the reference attenuation coefficients $\overline{\boldsymbol{\mu}} = \{\overline{\mu}_n\}$ are selected by minimizing the magnitude of the correction term in Eq. (6.5), i.e. by minimizing the following quadratic functional with respect to $\overline{\boldsymbol{\mu}}$:

$$\begin{aligned}\Psi(\overline{\boldsymbol{\mu}}, \mathbf{t}) &= \sum_{L(s,\theta)} \left(A_M^{\text{sim}}(L) - A_P^{\text{sim}}(L)\right)^2 \\ &= \sum_{L(s,\theta)} \left(\sum_{n=1}^{N} \overline{\mu}_n t_n(L) + \log\left(\sum_{j=1}^{J} I_j^F e^{-\sum_{n=1}^{N} \mu_{n,j} t_n(L)}\right)\right)^2 . \end{aligned} \tag{6.10}$$

To solve this minimization problem, define a matrix $\mathbf{B}$ with elements

$$b_{n,n'} = \sum_{L(s,\theta)} t_n(L) t_{n'}(L) \qquad n, n' = 1, ..., N, \tag{6.11}$$

and a vector $\mathbf{v}$ as

$$v_n = \sum_{L(s,\theta)} t_n(L) \log\left(\sum_{j=1}^{J} I_j^F e^{-\sum_{n=1}^{N} \mu_{n,j} t_n(L)}\right) \qquad n = 1, ..., N. \tag{6.12}$$

The minimum of the functional in Eq. (6.10) is then given by

$$\overline{\boldsymbol{\mu}} = \mathbf{B}^{+} \mathbf{v}, \tag{6.13}$$

where $\mathbf{B}^{+}$ denotes the Moore-Penrose generalized inverse of matrix $\mathbf{B}$.

Characteristic to this method, shared with the method of Krumm et al., is that the optimized reference attenuation coefficients do not provide accurate information on the Hounsfield numbers, which is the price to pay if $I_0(E)$ and $\mu_n(E)$ are not known a priori.

6.3.3 Iterative beam hardening correction

The proposed iterative beam hardening correction (IBHC) method has a similar outline as a standard postreconstruction method. An overview is depicted in Fig. 6.1. In each iteration $w = 1, 2, \ldots$, the following steps are executed:

(i) Using FBP, reconstruct the image from the sinogram $A_M^{\text{cor},w-1}(L)$, where $A_M^{\text{cor},0}(L)$ is given by the measured sinogram.

(ii) Segment the reconstruction. In our implementation, the simple histogram based thresholding method of Otsu [21] was used.

(ii) Given this segmented image, estimate thickness sinogram $t_n^w(L)$ for each material by forward projecting the indicator function of the corresponding material area.

(iv) Estimate the unknown parameters $\boldsymbol{\mu}^w$ and $\mathbf{I}^{F,w}$ by minimizing the cost function in Eq. (6.8):

$$(\boldsymbol{\mu}^w, \mathbf{I}^{F,w}) = \underset{(\boldsymbol{\mu}, \mathbf{I}^F)}{\arg\min}\ \Phi(\boldsymbol{\mu}, \mathbf{I}^F, \mathbf{t}^w), \tag{6.14}$$

while imposing the extra constraint that the attenuation values decrease for increasing energy E_j. This constraint is justified by the observation that the attenuation coefficient decreases with increasing energy for most materials in the typical experimental energy range.
We handle this non-linear optimization problem using a gradient based optimization algorithm. The cost function $\Phi(\boldsymbol{\mu}, \mathbf{I}^F, \mathbf{t}^w)$ is non convex and there is no guarantee that the solution to this optimization problem is unique. However, in our experiments we did not encounter convergence problems, and the optima found in successive IBHC iterations gradually resulted in an improved reconstruction quality.

(v) Simulate the polychromatic attenuation sinogram using Eq. (6.7).

(vi) Simulate the monochromatic attenuation sinogram using Eq. (6.9). The reference attenuation coefficients $\overline{\boldsymbol{\mu}}^w$ are determined by calculating the least squares solution (Eq. (6.13)) of the cost function in Eq. (6.10).

(vii) Calculate the updated sinogram $A^{\mathrm{cor},w}(L)$ by adding the difference between the simulated mono- and polychromatic sinogram to the measured data, as formulated in Eq. (6.5).

Iterating this loop results in an improved estimate of the thickness sinograms and the corresponding BH correction.

We define the polychromatic model error ϵ^w in iteration w as the minimum of the cost function $\Phi(\boldsymbol{\mu}^w, \mathbf{I}^{F,w})$ in Eq. (6.8) with respect to the beam hardening parameters

$$\epsilon^w = \min_{\boldsymbol{\mu}^w, \mathbf{I}^{F,w}} \left(\Phi(\boldsymbol{\mu}^w, \mathbf{I}^{F,w})\right). \tag{6.15}$$

The algorithm, described by steps (i) to (vii) is highly non-linear and there is no guarantee that it converges. However, the convergence of the algorithm can be monitored by following the value of ϵ^w. We observe a decreasing trend for ϵ^w, and we stop the iterations when

$$\frac{\epsilon^w + \epsilon^{w-1}}{\epsilon^{w-2} + \epsilon^{w-3}} > 0.99, \tag{6.16}$$

where the averaging between pairs of successive iterations aims at improving the robustness to the oscillations that often occur.

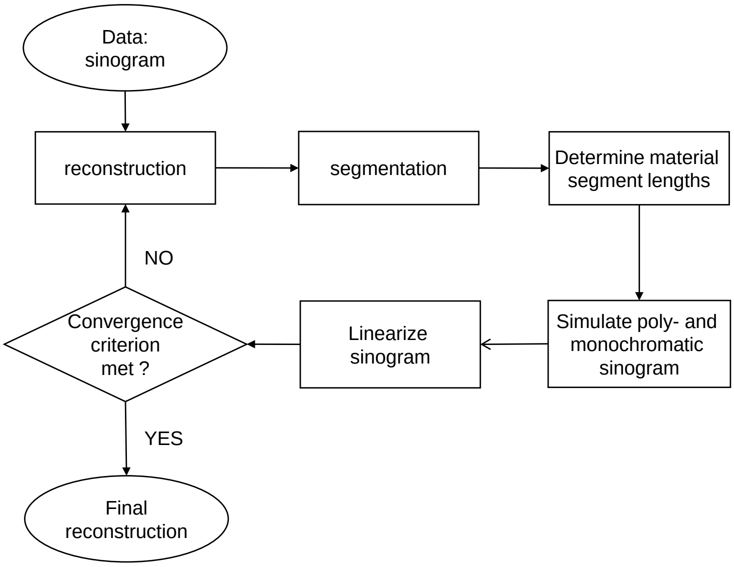


Figure 6.1: *Flowchart of the proposed beam hardening correction method.*

6.3.4 Accelerated IBHC method

Several papers in the literature [1, 2, 10–14] present results for mild beam hardening problems, and state that, although the post reconstruction approach is iterative, the number of iterations is very small in practice. However, in practical applications, severe streaks often hinder proper segmentation and therefore the linearization procedure requires many iterations. In Section 6.4, examples of challenging beam hardening artifacts are presented for which the number of iterations that is required to meet condition (6.16), exceeded 30. Performing a large number of reconstructions on a complete 3D data set is not desirable. We address this problem by performing a preliminary fast iterative beam hardening correction on a downsampled sinogram and image grid. In addition, we found that the number of iterations can be reduced by smoothing the reconstructed image prior to the segmentation, in each iteration.

In practice, we use the following automated algorithm, which we call the *accelerated IBHC* (AIBHC) method.

(1) Downsample the measured sinogram. Perform IBHC iterations for the downsampled sinogram, until condition (6.16) is met. In each iteration, smooth the reconstructed image using a Gaussian smoothing filter, preceding the segmentation.

(2) Use the resulting sinogram of step (1) to initialize a new series of IBHC

iterations without Gaussian smoothing, until condition (6.16) is met.

(3) Upsample the corrected sinogram resulting from step (2) to the size of the original measured sinogram using bilinear interpolation. Finalize the algorithm by performing two full-resolution iterations.

6.4 Experiments and results

6.4.1 Quality criterion

In Section 6.4, reconstructions with and without beam hardening correction are shown. One way to compare the quality of two reconstructed images is visual observation, but preferably, these observations are supported by a quantitative measure. For beam hardening correction methods, the restored uniformity in the corrected images can be quantified by measuring the variance in each material area. Define the neighborhood $\mathcal{N}_m(x_0, y_0)$ as the set of pixels within a radius m from pixel (x_0, y_0):

$$\mathcal{N}_m(x_0, y_0) = \{(x, y)|(x - x_0)^2 + (y - y_0)^2 < m^2\}, \tag{6.17}$$

and define $\mathcal{S}_n$ as the set of pixels that belong to material n. The objects used in this chapter are known, which allows us to obtain the ground truth sets $\mathcal{S}_n$. Consider $\Omega_n \subset \mathcal{S}_n$ to be the set of pixels (x, y) in $\mathcal{S}_n$ that are not in the neighborhood of the material boundary:

$$\Omega_n = \{(x, y)|\mathcal{N}_m(x, y) \subset \mathcal{S}_n\}. \tag{6.18}$$

The attenuation uniformity of the material classes is now determined by calculating the variance σ_n^2 of the attenuation in each material class separately:

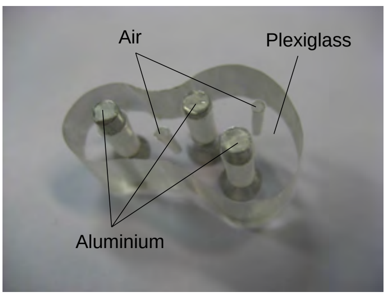
$$\sigma_n^2 = \frac{1}{|\Omega_n|} \sum_{(x,y)\in\Omega_n} \left(\mu(x, y) - \langle\mu_n\rangle\right)^2, \tag{6.19}$$

where $\langle\mu_n\rangle$ denotes the mean attenuation value for the set Ω_n. Since the images have varying noise properties, the variance measures are performed on smoothed reconstruction images, to make sure that the variance mainly reflects slow varying non-uniformities instead of noise. A uniform circular smoothing kernel was used, with radius r. For noisy images (Barbapapa and Bean phantom), the parameters are set at $r = 10$ and $m = 12$, which is large enough to avoid contributions of the boundaries to the variance measures. For the low-noise images of Section 6.5, we set $r = 1$ and $m = 2$.

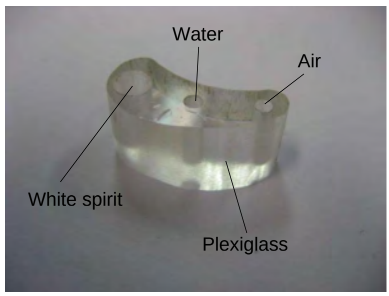
6.4.2 Beam hardening correction

In this subsection, beam hardening problems are considered in which the preliminary reconstruction yields a poor segmentation and therefore a poor estimate of the material thicknesses t_n. For our experiments, X-ray data were acquired at 60kV using a SkyScan 1172 μCT scanner, which has a circular cone beam geometry. Hardware filtering and the software beam hardening correction option are turned off. Our implementation uses a 2D parallel back- and forward projector. Hence, we selected only the fan beam data for the central slice, and rebinned it to a parallel sinogram. The parallel beam sinogram contains 300 equally spaced views with angular range $[0, \pi)$, and 1000 radial samples. The images were reconstructed on a 1000×1000 grid. The results are shown for two phantoms: the *Barbapapa* and the *Bean* phantom, which are depicted in Figures (6.2a) and (6.2b), respectively. The Barbapapa phantom consists of plexiglass and three aluminum cylinders; the Bean phantom consists of three materials: plexiglass, white spirit, and water.

To correct beam hardening artifacts for the Barbabapa and Bean phantoms,



(a) Barbapapa phantom



(b) Bean phantom

Figure 6.2: *Pictures of the used hardware phantoms.*

we followed the AIBHC strategy described in Section 6.3.4. The AIBHC method was initialized using a downsampled sinogram of 250 radial by 150 angular samples, and a Gaussian filter was selected with a standard deviation of 0.5 pixels. Table 6.1 lists the numbers of iterations that were executed to meet the stop condition Eq. (6.16) in steps (1) and (2) of the AIBHC method, for the Barbapapa and the Bean phantom.

The left image in Figure 6.3 shows the FBP reconstruction of the Barbapapa phantom from the uncorrected measured sinogram. This image suffers from severe cupping and streak artifacts which hinder an accurate segmentation. As an illustration, Figure 6.4 (left) depicts the Otsu segmentation of this FBP reconstruc-

AIBHC	Barbapapa	Bean
Step (1)	19	28
Step (2)	11	8

Table 6.1: *Number of iterations that is executed in steps (1) and (2) of the AIBHC method to meet condition (6.16).*

tion. The center and right images of Figure 6.4 show the gradual improvement of the segmentation image through successive iterations of the AIBHC method. The right image in Figure 6.3 represents the reconstruction obtained with the AIBHC method. In this image, the cupping and streak artifacts are strongly suppressed, which is confirmed by the variance measures in Table 6.2.

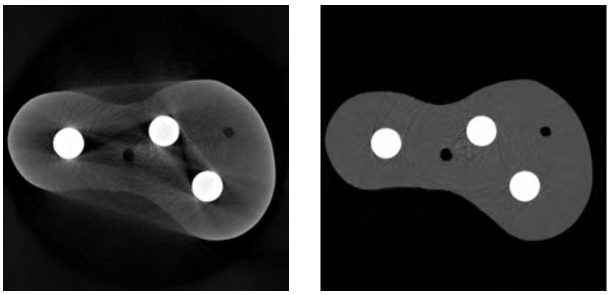


Figure 6.3: *Reconstructions from experimental data of the Barbapapa phantom. Left: FBP reconstruction. Right: AIBHC reconstruction.*

Variance ($\times 10^{-2}$)	Plexiglass	Aluminum
No correction	8.9	0.34
AIBHC	0.08	0.014

Table 6.2: *Variance per material segment in the reconstructed images for the Barbapapa phantom.*

Figure 6.5 shows the FBP and AIBHC reconstructions of the Bean phantom, respectively. Similarly, these images and the variance measures that are listed in Table 6.3, suggest that the AIBHC method is very effective in the reduction of both cupping and streak artifacts. Although the Bean phantom consists of three materials, only two material segments were considered during segmentation. The reason is given by the fact that water and plexiglass have very similar attenuation

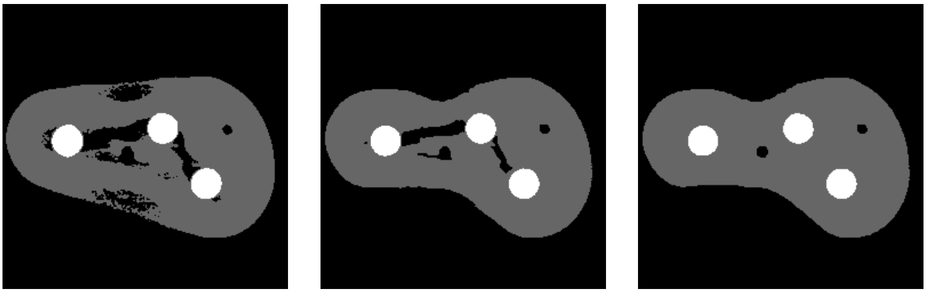


Figure 6.4: *Illustration of the the improved segmentation accuracy in successive iterations of the beam hardening correction procedure. Left to right : iteration 1,6,20.*

properties. As a result, the Otsu [21] segmentation classifies the two materials into one segment and a false additional material class would be detected when using 3 material classes. Note that, despite the fact that the water is not segmented as a separate material, it can still be distinguished from the plexiglass in the final reconstruction. This is because the final image is reconstructed from the corrected sinogram defined by Eq. (6.5), rather than directly from the simulated monochromatic sinogram $A_M^{\mathrm{sim}}(L)$. The latter, being based on only 2 materials, would not elicit the water region.

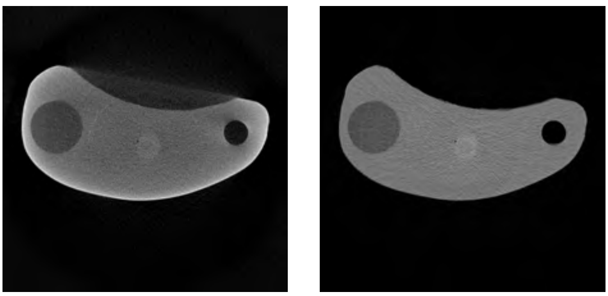


Figure 6.5: *Reconstructions from real X-ray data of the Bean phantom. Left: FBP reconstruction. Right: AIBHC reconstruction.*

6.5 Comparison with the method of Krumm et al.

Simultaneously and independently, Krumm et al. [1] proposed an alternative IPR method, in which the poly- and monochromatic sinograms are simulated by fitting respectively a hypersurface and hyperplane to the data. Similar to the IBHC

Variance ($\times 10^{-2}$)	Plexiglass	White Spirit	Water
No correction	3.1	0.53	0.15
AIBHC	0.072	0.12	0.098

Table 6.3: *Variance per material segment in the reconstructed images for the Bean phantom.*

method, this approach does not require the spectrum or material properties to be known a priori.

In this section, we compare the IBHC method with the beam hardening correction method proposed by Krumm et al. [1] These methods differ mainly in the way the polychromatic sinogram is simulated. In the IBHC method, the polychromatic sinogram is simulated by evaluating Eq. (6.7) after estimating the unknown coefficients by optimizing the cost function Φ in Eq. (6.8). Krumm et al. obtain the simulated polychromatic sinogram by fitting a smooth N-dimensional hypersurface to $A_P^{\mathrm{meas}}(\mathbf{t})$, which represents the measured attenuation as a function of estimated thicknesses, where $\mathbf{t} = (t_1, ..., t_N)$. For this fitting step, radial basis functions are used, which allow to perform smooth fits. The radial basis approximation to $A_P^{\mathrm{meas}}(\mathbf{t})$ is a function $A_P^{\mathrm{sim}}(\mathbf{t})$ defined by

$$A_P^{\mathrm{sim}}(\mathbf{t}) = \sum_{k=1}^{K} \lambda_k \varphi\big(|\mathbf{t} - \mathbf{t}_k|\big), \tag{6.20}$$

with φ the one-dimensional radial basis function [1], and where $\mathbf{t}_k$ represents a set of well-chosen thickness coordinates [1]. The unknown coefficients $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_K)$ can then be estimated by minimizing the cost function $\tilde{\Phi}(\boldsymbol{\lambda})$ given by

$$\tilde{\Phi}(\boldsymbol{\lambda}) = \sum_{L(s,\theta)} \left(\log\left(\frac{I_P^{\mathrm{meas}}(\mathbf{t}(L))}{I_0} \right) - \sum_{k=1}^{K} \lambda_k \varphi\big(|\mathbf{t}(L) - \mathbf{t}_k|\big) \right)^2 . \tag{6.21}$$

This represents a linear optimization problem in which K parameters are to be optimized. To maintain equal sampling density, the number of parameters K increases exponentially with the number of materials N.

Due to the fact that the method of Krumm et al. is part of a commercial product, its implementation details are not available to us. Therefore, we use the same experimental data as in [1], and compare our IBHC reconstructions with reconstructions provided by Krumm et al. As an illustration, we show results for real X-ray CT data of an aluminium cylinder in which five steel pins were drilled. Detailed scan settings can be found in [1]. We used the central slice of the cone beam scan, and rebinned it to parallel beam. The beam hardening problem in this dataset is rather mild in the sense that the first segmentation already largely corresponds to the expected object shape. Hence, both algorithms found the final reconstruction after 2 iterations.

Figure 6.6 shows the FBP reconstructions without BH correction, with the BH correction of Krumm et al. and with IBHC correction, respectively. We observe that both methods largely restore uniformity, which is confirmed by the variance values in table 6.4. These results illustrate that the IBHC method achieves accurate reconstructions by using a physical model using only $K = 3(N+1) - 1$ parameters (in this case K=8). This is an important advantage compared to the method of Krumm et al., where the number of parameters exponentially increases with the number of materials N (for a 2 material object, Krumm et al. used $K = 400$).

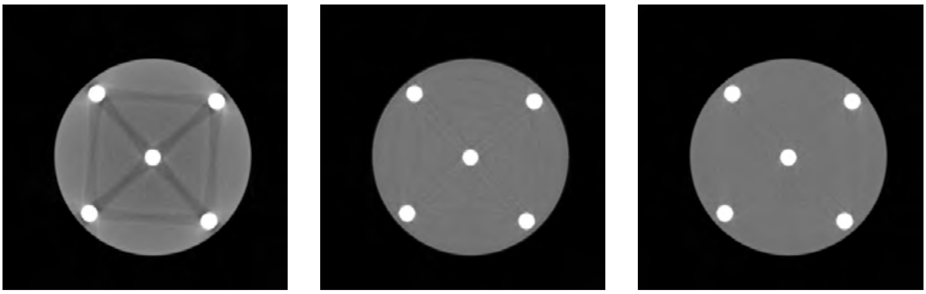


Figure 6.6: *Reconstructions of the central slice of the aluminum-steel phantom of Krumm et al. Left: FBP reconstruction from non-linear data. Center : FBP reconstruction after beam hardening correction with the method of Krumm et al. Right: FBP reconstruction after beam hardening correction using IBHC.*

Variance ($\times 10^{-2}$)	Aluminum	Steel
No correction	0.85	0.74
method of Krumm et al.	0.040	0.0168
AIBHC	0.0361	0.0182

Table 6.4: *Variance measures of the attenuation coefficients in each material segment of the reconstructions shown in Figure 6.6.*

Note that the optimization problem in Krumm's method is linear, hence it can be solved by direct matrix inversion, which has a very small computational cost. The cost function in the IBHC method, on the other hand, is non-quadratic, and therefore a gradient based optimization method is used to determine the unknown parameters, which is computationally more expensive. In our implementation, the optimization in one iteration of the example in Fig. 6.6, required approximately 6 s. This represents a relatively small contribution to the computation time of a complete IBHC iteration (involving FBP, thresholding, optimization, etc) which used 30 s for this example.

6.6 Discussion and Conclusion

In this chapter, an new iterative method is described for beam hardening correction for objects consisting of multiple, uniform materials. The number of materials is considered to be known beforehand, but no information on the source spectrum or on the energy dependent attenuation coefficients of the materials is used. This is a significant practical advantage compared to most other beam hardening correction methods, which often require additional calibration experiments. The price to pay is that the obtained reconstructions do not provide accurate quantitative information on the Hounsfield numbers. The method, however, successfully reduces the cupping artifacts caused by the beam polychromaticity in such a way that the reconstruction of each homogeneous region has a low variance. The method is described for a 2D parallel geometry, but it is readily extensible to any acquisition geometry because the beam hardening correction is performed ray per ray, and the segmentation step can be applied independently of the 2D or 3D algorithm used to generate the image. Moreover, the IBHC method can be combined with any reconstruction algorithm that yields adequate segmentation images, which will be illustrated in Chapter 7.

Bibliography

[1] M. Krumm, S. Kasperl, and M. Franz, "Reducing non-linear artifacts of multi-material objects in industrial 3D computed tomography," *Nondestructive Testing and Evaluation*, vol. 41, pp. 242–251, 2008.

[2] J.P. Stonestrom, R. E. Alvarez, and A.M. Mackowski, "A framework for spectral artifact corrections in X-ray CT," *IEEE Transactions on Biomedical Engineering*, vol. 28, no. 2, pp. 128–141, February 1981.

[3] R.E. Alvarez and A. Macowski, "Energy-selective reconstructions in X-ray computerized tomography," *Physics in Medicine and Biology*, vol. 21, no. 5, pp. 733–744, 1976.

[4] 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.

[5] W. A. Kalender, D. Felsenberg, H. K. Genant, M. Fischer, J. Dequeker, and F. Reeve, "The European Spine Phantom a tool for standardization and quality control in spinal bone mineral measurements by DXA and QCT," *European Journal of Radiology*, vol. 20, pp. 83–92, April 1995.

[6] I.A. Elbakri and J.A. Fessler, "Segmentation-free statistical image reconstruction for polyenergetic x-ray computed tomography with experimental validation," *Physics in Medicine and Biology*, vol. 48, pp. 2453–2477, 2003.

[7] N. Menvielle, Y. Goussard, D. Orban, and G. Soulez, "Reduction of beam-hardening artifacts in x-ray ct," *Engineering in Medicine and Biology Society, EMBS, 27th Annual International Conference of the IEEE*, pp. 1865–1868, 2005.

[8] B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens, "An iterative maximum-likelihood polychromatic algorithm for CT," *IEEE Trans. Med. Imaging*, vol. 20, no. 10, pp. 999–1008, 2001.

[9] E. Van de Casteele, D. Van Dyck, J. Sijbers, and E. Raman, "An energy-based beam hardening model in tomography," *Physics in Medicine and Biology*, vol. 47, no. 23, pp. 4181–4190, 2002.

[10] P.M. Joseph and R.D. Spital, "A method for correcting bone induced artifacts in computed tomograpy scanners," *Journal of Computer Assisted Tomography*, vol. 2, no. 1, pp. 100–109, January 1978.

[11] G.T. Herman, "Correction for beam hardening in computed tomography," *Physics in Medicine and Biology*, vol. 24, no. 1, pp. 81–106, 1979.

[12] G.T. Herman and S.S. Trivedi, "A comparative study of two postreconstruction beam hardening correction methods," *IEEE Transactions on Medical Imaging*, vol. 2, no. 3, pp. 128–135, September 1983.

[13] P.K. Kijewski and B.E. Bjärngard, "Correction for beam hardening in computed tomography," *Medical Physics*, vol. 5, no. 3, pp. 209–214, May 1978.

[14] P.M. Joseph and C. Ruth, "A method for simultaneous correction of spectrum hardening artifacts in CT images containing both bone and iodine," *Medical Physics*, vol. 24, no. 10, pp. 1629–1634, October 1997.

[15] J. Hsieh, R.C. Molthen, C.A. Dawson, and R.H. Johnson, "An iterative approach to the beam hardening correction in cone beam CT," *Medical Physics*, vol. 27, no. 1, pp. 23–29, January 2000.

[16] V.S.V.M. Vedula and P. Munshi, "An improved algorithm for beam-hardening corrections in experimental X-ray tomography," *NDT&E International*, vol. 41, pp. 25–31, 2008.

[17] J.M. Meagher, C.D. Mote, and H.B. Skinner, "CT image correction for beam hardening using simulated projection data," *IEEE Transactions on Nuclear Science*, vol. 37, no. 4, pp. 1520–1524, August 1990.

[18] E. Van de Casteele, D. Van Dyck, J. Sijbers, and E. Raman, "A model-based

correction method for beam hardening artefacts in X-ray microtomography," *Journal of X-ray Science and Technology*, vol. 12, no. 1, pp. 43–57, 2004.

[19] H. Gao, L. Zhang, Z. Chen, Y. Xing, and L. Shuanglei, "Beam hardening correction for middle-energy industrial computerized tomography," *IEEE Transactions on Nuclear Science*, vol. 53, no. 5, pp. 2796–2807, 2006.

[20] X. Mou, S. Tang, and H. Yu, "Comparison on beam hardening correction of CT based on HL consistency and normal water phantom experiment," in *SPIE Medical Imaging: Image Processing*, 2006, vol. 6318, p. 63181V.

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

Chapter 7

Beam hardening correction for insufficient data problems

7.1 Introduction

The IBHC beam hardening correction method, proposed in Chapter 6, essentially uses a *ray per ray* approach in the sense that each data point in the sinogram is corrected independently. In addition, the segmentation step can be applied independently of the 2D or 3D algorithm that is used to generate the image. An important advantage of such an approach is that it can trivially be extended for any scanning geometry, provided that the corresponding ray paths are known. Moreover, it can be combined with any reconstruction algorithm that yields accurate segmentation images. This chapter aims to demonstrate that such a combined approach can result in an advantageous synergy between the IBHC method and the reconstruction algorithm in terms of reconstruction accuracy.

7.2 Beam hardening correction for insufficient data problems

Recall that FBP reconstruction from incomplete data results in severely degraded reconstructed images (see Section 2.3). Because segmentation is an essential step in the IBHC method (and iterative post reconstruction methods in general), beam hardening correction fails when FBP is used as reconstruction method.

Data insufficiency problems require specific reconstruction methods that exploit the available data and incorporate certain assumptions about the objects to restore uniqueness. For example, algorithms that minimize the *total variation* such as the *TVmin* method of Sidky et al [1], assume that objects are relatively constant over larger volumes. The *Discrete Algebraic Reconstruction Technique* (DART) [2], on the other hand, assumes that the object is piecewise uniform and that the attenuation values are known. Such algorithms can be very powerful and yield accurate reconstructions from very small amounts of data. However, like FBP, these algorithms are based on the assumption that the data are line integrals of the attenuation coefficients. A violation of this assumption, caused by beam hardening, may have a negative impact on the accuracy of the reconstructions. In this section, three data insufficiency problems are considered in combination with beam hardening: a small number of projections, a limited angular range, and truncated projections. We demonstrate that the IBHC method can be applied on incomplete data, and that the combined approach of the beam hardening correction procedure and the specific limited data reconstruction methods (DART and TVmin) results in an improved object recovery, notwithstanding the poor quality of the preliminary segmentation.

The experiments are performed on the real X-ray CT data of the Barbapapa phantom (see Section 6.4.2), which is significantly affected by beam hardening. This dataset, consisting of 300 projections with 1000 radial samples, is downsampled to a sinogram of 250 radial and 150 angular samples. The corresponding reconstructions are computed on a 250×250 grid.

7.2.1 Small number of projections.

Recall from Section 5.1 that methods minimizing the total variation (TV) allow for accurate reconstructions of fairly complex piecewise uniform objects from a very small number of projections (see Candes et al [3], Sidky et al [1] and Herman and Davidi [4]). We implemented the heuristic iterative TVmin algorithm of Sidky et al. [1], in which each iteration basically consists of a general iterative reconstruction (e.g. SIRT), followed by a steepest descent TV minimization. This TV minimization method is incorporated into the IBHC algorithm by replacing the FBP reconstruction step. A schematic overview of the resulting method, further referred to as the IBHC-TVmin method, is shown in Fig. 7.1. The grey rectangle indicates the step that was altered, compared to the IBHC method (Fig. 6.1).

Consider the dataset of the Barbapapa phantom, of which only 10 projections, uniformly sampled in $[0, \pi)$, are used. Figure 7.2 depicts several reconstructed images from this limited dataset. From left to right, the three images represent the FBP, TVmin and the IBHC-TVmin reconstructions, respectively. The left and central image yield poor structural information of the phantom in the sense that in larger areas, the object cannot be distinguished from the background, and the

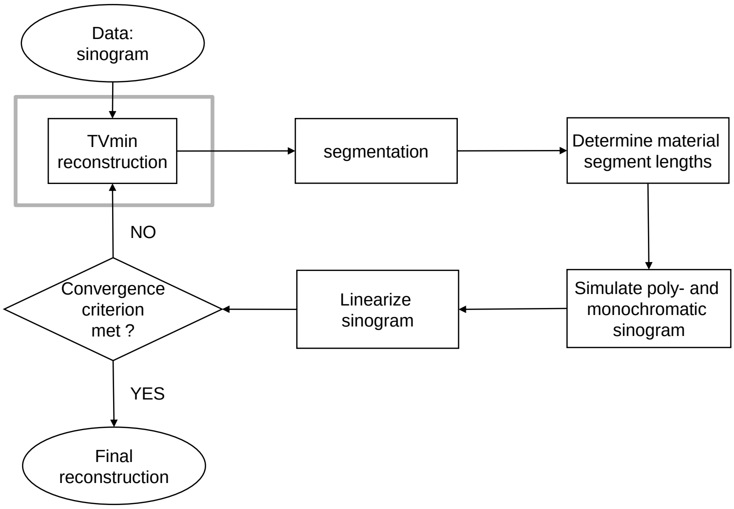
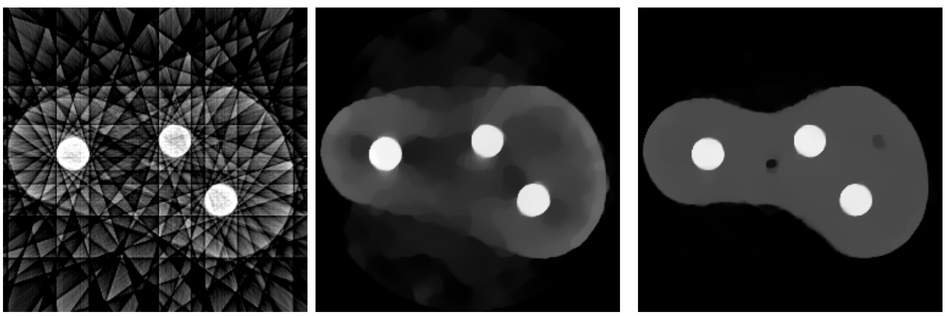


Figure 7.1: *Flowchart of the IBHC-TVmin method.*

air holes are not detected. As opposed to these images, the IBHC-TVmin reconstruction, obtained after 60 iterations, shows that the object is largely recovered and beam hardening artifacts are well suppressed.

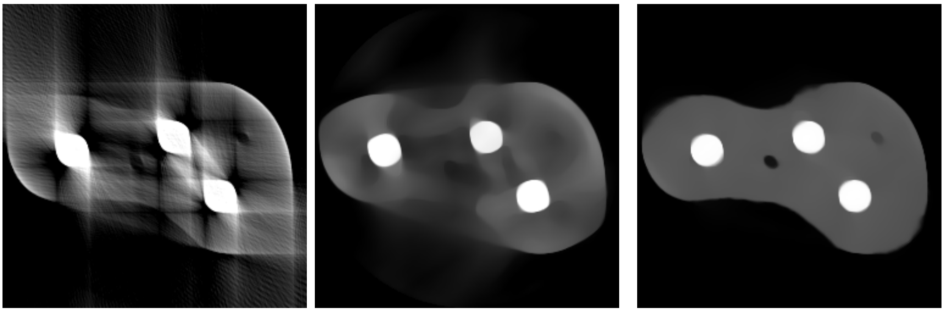


(a) FBP reconstruction (b) TVmin reconstruction (c) IBHC-TVmin reconstruction

Figure 7.2: *Reconstructions from 10 X-ray projections of the Barbapapa phantom.*

7.2.2 The limited angular range problem.

A limited angular range dataset is created by using only the 75 projections in the angular range $\theta \in [0, \pi/2)$. Figure 7.3 shows the resulting reconstructions for the limited angular range problem. The left and center figures represent the FBP and TVmin reconstruction, respectively. The right image results from applying 18 IBHC-TVmin iterations. Only one of the two air-holes is recovered. However, it can be observed that the object boundary is much better defined in the right image.



(a) FBP reconstruction (b) TVmin reconstruction (c) IBHC-TVmin reconstruction

Figure 7.3: *Reconstructions from limited angle X-ray data of the Barbapapa phantom. The angular range for this example is* $\theta \in [0, \pi/2)$.

7.2.3 The truncated projection problem.

Consider a piecewise uniform object, consisting of K materials with unknown density, and corresponding projections, which are affected by beam hardening and truncation. For this problem, to maintain the analogy with Chapter 5, the IBHC method is combined with the iterative DART reconstruction method (see Section 5.2).

The DART method is incorporated in the IBHC method by replacing the FBP reconstruction and the segmentation step. Preceding each DART reconstruction, the unknown densities are estimated in two steps:

1. Perform a SIRT reconstruction from the updated truncated sinogram. The reconstructed image obtained in the previous IBHC iteration serves as the input image for SIRT. An exception is made for the first iteration, in which the measured data and a blanc input image are used.

2. Determine K threshold values u_k with $k = 1, ..., K$ by applying the method of Otsu [5] exclusively on the FOV area of the SIRT reconstruction. Let u_0 and u_{K+1} be the minimum and maximum value in the FOV area of the

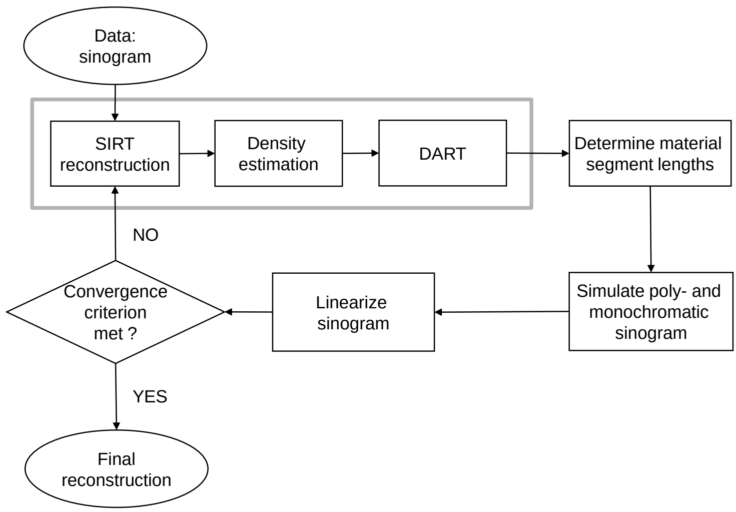


Figure 7.4: *Flowchart of the IBHC-DART method.*

SIRT image, respectively. The unknown densities (c_k) are estimated using $c_k = \frac{u_{k+1}}{2}$ for $k = 1...K$.

A schematic overview of the IBHC-DART method is depicted in Fig. 7.4, where the grey rectangle indicates the steps that are altered with respect to the IBHC method.

Recall that the complete dataset consists of 250 radial samples in each of the 150 projections. Projection truncation of this data is simulated by restricting the projection data to the $N_z = 124$ radial samples at the center of each projection. Fig. 7.5 depicts reconstructions from these truncated projections. The superimposed white circle indicates the FOV. From these images, it can be observed that the FBP reconstruction (Fig. 7.5 (a)) is clearly distorted by truncation and beam hardening artifacts. The DART reconstruction in Fig. 7.5(b) also offers very little structural information of the object. The reconstructed image (Fig. 7.5(c)) that is obtained after 95 iterations of the IBHC-DART method, however, nicely illustrates the advantageous synergy between the IBHC and the DART method in terms of image accuracy.

7.3 Conclusion

This chapter focusses on the correction of beam hardening artifacts for limited data problems, which require specific reconstruction methods that incorporate prior knowledge about the objects. Since the non-linearity caused by beam harden-

(a) FBP reconstruction (b) DART reconstruction (c) IBHC-DART reconstruction

Figure 7.5: *Reconstructions from truncated non-linear data of the Barbapapa phantom with with $N_z = 124$.*

ing violates the assumptions made by those algorithms, we proposed to combine these iterative reconstruction approaches with the IBHC beam hardening correction method. Three incomplete data problems (few projections, limited angular range, and truncation) and two specific iterative reconstruction methods (TVmin [1] and DART [2]) were considered. The results illustrate that beam hardening artifacts can be significantly reduced, or in some cases be removed completely, by using a combined iterative approach of a beam hardening reduction and a limited data reconstruction method.

Bibliography

[1] E.Y. Sidky, C.M. Kao, and X Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *Journal of X-ray Science and Technology*, vol. 14, pp. 119–139, 2006.

[2] K.J. Batenburg and J. Sijbers, "DART: a fast heuristic algebraic reconstruction algorithm for discrete tomography," *Proceedings of ICIP*, vol. 4, pp. 133–136, 2007.

[3] E. Candes, J. Romberg, and Tao T., "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, pp. 489–509, 2006.

[4] G.T. Herman and R. Davidi, "Image reconstruction from a small number of projections," *Inverse Problems*, vol. 24, pp. 045011, 2008.

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

Part V

Conclusions

Conclusions and future work

Conclusions and original contributions

An important issue faced in computer tomography is the accurate reconstruction of an object slice from incomplete data. We considered the problem in which the projections are transaxially truncated, for instance when the detector is too small to cover the projection image of the object in all projection directions.

Traditional analytical reconstruction methods, such as Filtered Backprojection (FBP), distribute a sinogram distortion over the complete reconstruction. In case of truncated projections, the FBP method typically provides an image with accurate structural information in the area inside the Field Of View (FOV) of the detector, but the grey values are contaminated by a slowly varying bias, which hinders accurate segmentation and quantitative gray value analysis. In the area outside the FOV, only severely distorted structural information is available.

In this thesis, we proposed and discussed several methods for the reconstruction of (parts of) an object from its truncated data, each accounting for a different sort of prior knowledge about the object.

Part II. In this part, we considered truncation problems for which no prior knowledge about the object is available. In this case it can be expected that no part of the region surrounding the FOV can be reconstructed accurately, since each point in the area outside the FOV is not covered by the detector at least in one projection direction. Therefore, we focussed on the accuracy enhancement of the reconstruction in the FOV-region only. Note, however, that even within this region the solution is not always uniquely determined (cfr. the interior problem).

We proposed a new empirical algorithm that extrapolates the missing data by exploiting the consistency between projections. This consistent sinogram extrapolation method (ConSiR) is used as a preprocessing step for the FBP

reconstruction method. Our experimental results suggest that this algorithm often yields more accurate reconstructions than the empirical Simple Extrapolation Method (SEM) [1] which is commonly used in literature. Note, however, that ConSiR is an approximating method for the estimation of the missing data, which hinders quantitative analysis on the grey values.
Alternatively, for some types of truncation, algorithms were proposed in literature, that accurately recover specific parts within the FOV from truncated projections. One example is the analytical differentiated backprojection method (DBP)[2]. In our experiments, we indeed found that the DBP method provides very accurate reconstructions of certain areas in the FOV. However, in the remaining parts of the FOV, the empirical methods ConSiR and SEM seem to outperform the DBP method with respect to image accuracy.
We conclude that in case accurate segmentations are aimed at, data extrapolation methods such as the ConSiR method are still relevant for truncation artifact reduction since they can be applied regardless of the type of truncation.

Part III. Methods such as SEM, ConSiR and DBP focus on the reconstruction of (parts of) the FOV from truncated projection data, and not on the recovery of the object region outside the FOV, since this reconstruction problem is severely underdetermined. In recent years, it was shown for several limited data problems that such underdeterminacy can be alleviated or solved by incorporating certain prior knowledge of the object. In many applications, some assumptions on the object can be made, for instance that an object is piecewise uniform. In part III, we investigated the possibility of recovering a complete object from its truncated data in case the object can be assumed to be piecewise uniform.

Initially, we considered a small subset of piecewise uniform objects: uniform star-shaped objects. We proved that a star-shaped object with uniform but unknown density is uniquely determined by its truncated projections, even in case the truncated data corresponds to an interior problem. To the best of our knowledge this uniqueness theorem is new. The proof resulted in a new reconstruction algorithm, the differentiated backprojection method for star-shaped objects (DBPS), that handles each radial line of the star object independently. In experiments with simulated data, a severe degradation of the stability was observed as the radius of the FOV is decreased in the presence of noise. However, if the density is known beforehand, accurate image

reconstructions can still be obtained for significantly reduced FOV's.
Optimal stability, however, requires a global two-dimensional approach, which avoids separating the reconstruction into a set of one-dimensional reconstructions along central lines. An example of such an approach is the discrete algebraic reconstruction technique (DART) [3], which is an iterative technique assuming piecewise uniformity of the objects with known densities. We adopted the DART algorithm, which was originally proposed for few-projections and limited-angular-range problems, for the application on truncated data of uniform star-shaped objects and we observed an increased stability compared to the DBPS method.

The proof of the uniqueness theorem for star-shaped objects suggested that the uniqueness result for binary star shaped objects from interior data could be generalized to more complex objects. Since the DART algorithm does not make any explicit assumptions on the shape of the considered objects, we applied it for the reconstruction of more general piecewise uniform objects to indicate the extent of object complexity for which the inverse problem from truncated data can be recovered in practice. The experiments suggest that binary objects with relatively complex shapes can still be reconstructed stably from interior data. However, for objects consisting of multiple materials, the reconstruction accuracy of the DART reconstruction from interior data is reduced drastically.
This research for piecewise constant objects was motivated by a specific question from DiamCad, a diamond processing company that scans rough diamonds to retrieve detailed information on their shapes. We received a truncated dataset of a diamond that was too big to fit in the field of view of the detector. Based on the above mentioned study, we were able to accurately recover the complete diamond from the truncated data. Alternative applications can potentially be found in industrial CT, where often very large but uniform objects (such as engines) are scanned.

Part IV. In the study of Part III, we found that the underdeterminacy caused by the truncation of projections, can be alleviated or dissolved by exploiting prior knowledge such as the piecewise uniformity of the object. It is very important that the assumptions are valid to a good approximation. In practical situations, the assumption of piecewise uniformity is often invalidated because of confounding physical effects such as beam-hardening. Beam hardening is the phenomenon that takes place when the incident X-ray beam is polychromatic. Since the absorption properties of materials vary with the energy of the X-ray photons, the spectrum of the beam changes as the beam penetrates deeper into the object, which violates the assumption of the standard reconstruction methods that the attenuation is linearly related to the thickness of the traversed material. The resulting reconstructions typically suffer from cupping and shadowing artifacts, which hinder an accurate image segmentation.

In Part IV, we presented a new iterative beam hardening correction (IBHC) method for piecewise uniform objects from complete data. The IBHC method does not require additional prior knowledge such as the material dependent attenuation coefficients or the source spectrum. This is a significant practical advantage compared to most other beam hardening correction methods, which often require additional calibration experiments. Our experiments for several physical phantoms show that the IBHC method successfully reduces the cupping artifacts caused by the polychromatic beam, in such a way that the reconstruction of each uniform region is to good accuracy homogeneous.

An important advantage of the IBHC method is that it can be combined with any reconstruction technique, provided that this reconstruction algorithm leads to an accurate segmentation in the monochromatic case. Therefore, the IBHC method can also be applied for limited data problems provided that the reconstruction is done using an appropriate algorithm, for instance DART. The results for a physical phantom illustrate that the combination of the IBHC method with several limited data reconstruction method leads to a satisfactory recovery of the object function that is virtually free of beam hardening artifacts.

Future work

The work presented in this thesis focussed on the reconstruction of a 2D object slice from a set of 1D parallel X-ray projections. Since modern X-ray scanners typically use a 3D cone beam geometry, the next step is to extend the methods and experiments to this geometry. Whereas this extension is rather straightforward for DART and the beam hardening correction method, it is much harder for the ConSiR method. The consistency conditions for 3D tomography essentially require four-dimensional data [4], which are not provided by a circular 3D cone beam scan. Note, however, that the standard 3D reconstruction method for a circular cone beam acquisition such as the Feldkamp-David-Kress (FDK) algorithm, actually performs a weighted 2D reconstruction along tilted planes, assuming that the vertical cone angle is sufficiently small. A similar approach could be considered for the extension of the consistency conditions, which would enable the use of ConSiR for a cone beam geometry.

Chapter 5 consisted of experiments of DART reconstruction for piecewise uniform object functions from truncated data. We found that the reconstruction accuracy quickly degrades as the number of densities increases. Although a decreasing relation is expected, it is unclear whether the obtained results are specifically related to the properties of DART or whether they indicate that the prior knowledge is optimally exploited. Therefore, similar experiments could be performed using iterative approaches that encourage, rather than strictly enforce piecewise uniformity of the solution, such as ℓ_1-minimization [5, 6]. This is particularly interesting since these methods do not require the densities of the objects to be known in advance.

The beam hardening method presented in Chapter 6 proved to recover the homogeneity of regions up to a good accuracy. The phantoms used in this study, however, contained large uniform regions. For objects with more detailed structures, partial volume effects will be more prominently present. To alleviate artifacts caused by this partial volume, one could allow mixtures of materials in the pixels at boundaries of segments, instead of classifying pixels in strictly separate material classes. In addition, the adaption of the IBHC to take other physical phenomena such as scattering into account, is a subject for future work.

Bibliography

[1] R.M. Lewitt and R.H.T. Bates, "Image reconstruction from projections: III: Projection completion methods (theory)," *Optik*, vol. 50, no. 3, pp. 189–204, October 1977.

[2] F. Noo, R. Clackdoyle, and J. D. Pack, "A two-step Hilbert transform method for 2D image reconstruction," *Physics in Medicine and Biology*, vol. 49, pp. 3903, August 2004.

[3] K.J. Batenburg and J. Sijbers, "DART: a fast heuristic algebraic reconstruction algorithm for discrete tomography," *Proceedings of ICIP*, vol. 4, pp. 133–136, 2007.

[4] S.K. Patch, "Consistency conditions upon 3D CT data and the wave equation," *Physics in Medicine and Biology*, vol. 47, pp. 26372650, 2002.

[5] E.Y. Sidky, C.M. Kao, and X Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *Journal of X-ray Science and Technology*, vol. 14, pp. 119–139, 2006.

[6] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, pp. 489–509, 2006.

List of publications

Journal articles

- G. Van Gompel, M. Defrise, and J. Batenburg, *Reconstruction of a uniform star object from interior x-ray data: uniqueness, stability and algorithm*, Inverse Problems, Vol. 25, ID 65010, June 2009.

- G. Van Gompel, M. Defrise, J. Batenburg and J. Sijbers, *Beam hardening correction for piecewise uniform objects and insufficient data in CT*, in preparation.

Conference proceedings (full paper)

- G. Van Gompel, M. Defrise, and J. Batenburg, *Reconstruction of a uniform star object from interior X-ray data*, Nuclear Science Symposium Conference Record, NSS08, IEEE, Dresden, Germany, October 2008.

- G. Van Gompel, M. Defrise and D. Van Dyck *Elliptical Extrapolation of Truncated 2D CT Projections using Helgason-Ludwig consistency conditions*, SPIE Medical Imaging: Physics of Medical Imaging, Vol. 6142, p. 1408-1417, San Diego, California, USA, February 2006.

- G. Van Gompel, G. Tisson, D. Van Dyck and J. Sijbers, *A new algorithm for Region Of Interest Tomography*, SPIE Medical Imaging: Physics of Medical Imaging, Vol. 127, p. 2105-2113, San Diego, California, USA, February 2004.

Conference proceedings (abstracts)

- G. Van Gompel, J. Batenburg, M. Defrise, and J. Sijbers, *An iterative post-reconstruction method for beam hardening reduction without prior information*, Liege Image Days 2008: Medical Imaging, March 2008.

- G. Van Gompel, M. Defrise, and D. Van Dyck, *Consistent Extrapolation of Truncated 2D CT Projections*, Micro tomography 3D, Gasthuisberg, Leuven, Belgium, September 2006.
- G. Van Gompel, D. Van Dyck, G. Tisson, and J. Sijbers, *A new algorithm for Region Of Interest Tomography*, 19th Annual Symposium of the Belgian Hospital Physicists Association, Brussels, Belgium, January 2004.