UNIVERSITY OF NOVI SAD
FACULTY OF TECHNICAL SCIENCES

Regularized Models for Tomographic Image Reconstruction

DOCTORAL DISSERTATION

Modeli sa regularizacijom za rekonstrukciju slika u tomografiji

DOKTORSKA DISERTACIJA

Supervisor:

Prof. Dr. Tibor Lukić

Candidate:

Marina Šulc

Novi Sad, 2023

УНИВЕРЗИТЕТ У НОВОМ САДУ **ОБРАЗАЦ – 5а**
ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА

КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА[1]

Врста рада:	Докторска дисертација
Име и презиме аутора:	Марина Шулц
Ментор (титула, име, презиме, звање, институција)	Др Тибор Лукић, редовни професор,Факултет техничких наука, Универзитет у Новом Саду
Наслов рада:	Модели са регуларизацијом за реконструкцију слика у томографији
Језик публикације (писмо):	Енглески (латиница)
Физички опис рада:	Унети број: Страница: 212 Поглавља: 6 Референци: 100 Табела: 5 Слика: 34 Графикона: 0 Прилога: 3
Научна област:	Примењена математика
Ужа научна област (научна дисциплина):	Теоријска и примењена математика
Кључне речи / предметна одредница:	Рачунарска обрада слика, реконструкција слике, дискретна томографија, бинарна томографија, дескриптори облика, минимизација енергије
Резиме на језику рада:	Фокус ове тезе је проблем реконструкције слике у дискретној и бинарној томографији. Теза даје формулацију проблема реконструкције слике у томографији, као и преглед постојећих алгоритама који се баве овим проблемом. У тези је уведена нова метода за реконструкцију слика заснована на минимизацији функције енергије у дискретној томографији која комбинује градијентну методу са методом сечења графова. Овај метод показује добре перформансе у поређењу са постојећим методама које се баве сличним проблемима. У тези је показано да се перформансе метода за реконструкцију слика у дискретној томографији са малим бројем пројекционих података могу значајно побољшати уколико се у њих укључи и а приори информација о објекту. У тези је посебан акценат стављен на оријентацију и циркуларност облика.

[1] Аутор докторске дисертације потписао је и приложио следеће Обрасце:
5б – Изјава о ауторству;
5в – Изјава о истоветности штампане и електронске верзије и о личним подацима;
5г – Изјава о коришћењу.
Ове Изјаве се чувају на факултету у штампаном и електронском облику и не кориче се са тезом.

	У тези је представљена нова метода, која интегрише методу сечења графа са техником минимизације засноване на градијенту, користећу оријентацију као познату информацију о облику. Метода показује завидне перформансе приликом реконструкције бинарних слика на основу пројекционих података из једне димензије. Додатно, описан је метод који користи циркуларност облика као а приори информацију о објекту, као и резултати експеримената који показују велику ефикасност овог метода у реконструкцији слика са смањеним пројекционим подацима.
Датум прихватања теме од стране надлежног већа:	25.03.2021.
Датум одбране: (Попуњава одговарајућа служба)	
Чланови комисије: (титула, име, презиме, звање, институција)	Председник: Др Биљана Михаиловић, редовни професор, Факултет техничких наука, Универзитет у Новом Саду Члан: Др Татјана Давидовић, научни саветник, Математички институт САНУ, Београд Члан: Др Лидија Чомић, ванредни професор, Факултет техничких наука, Универзитет у Новом Саду Члан: Др Буда Бајић Папуга, доцент, Факултет техничких наука, Универзитет у Новом Саду, Ментор: Др Тибор Лукић, редовни професор, Факултет техничких наука, Универзитет у Новом Саду
Напомена:	

UNIVERSITY OF NOVI SAD
FACULTY OF TECHNICAL SCIENCES

KEY WORD DOCUMENTATION[2]

Document type:	Doctoral dissertation
Author:	Marina Šulc
Supervisor (title, first name, last name, position, institution)	Prof. Dr. Tibor Lukić, full professor at Faculty of Technical Sciences, University of Novi Sad
Thesis title:	Regularized Models for Tomographic Image Reconstruction
Language of text (script):	English language
Physical description:	Number of: Pages: 212 Chapters: 6 References: 100 Tables: 5 Illustrations: 34 Graphs: 0 Appendices: 3
Scientific field:	Applied Mathematics
Scientific subfield (scientific discipline):	Theoretical and Applied Mathematics
Subject, Key words:	Digital image processing, image reconstruction, discrete tomography, binary tomography, shape descriptors, energy minimization methods
Abstract in English language:	The focus of these theses is on the problem of image reconstruction in discrete and binary tomography. The thesis provides a formulation of the image reconstruction problem in tomography, as well as an overview of existing algorithms addressing this issue. A new method for image reconstruction based on energy function minimization in discrete tomography is introduced in the thesis. This method combines gradient descent with graph-cutting techniques and demonstrates superior performance compared to existing methods addressing similar problems. The thesis shows that the performance of image reconstruction methods in discrete tomography with a small number of projection data can be significantly improved by incorporating a priori information about the object. The thesis emphasizes the orientation and circularity of shapes. A new method is presented in the thesis, integrating graph-cutting techniques

[2] The author of doctoral dissertation has signed the following Statements:
5б – Statement on the authority,
5в – Statement that the printed and e-version of doctoral dissertation are identical and about personal data,
5г – Statement on copyright licenses.
The paper and e-versions of Statements are held at he faculty and are not included into the printed thesis.

	with gradient-based minimization using orientation as known shape information. This method exhibits remarkable performance in reconstructing binary images based on one-dimensional projection data. Additionally, a method using the circularity of shapes as a priori information about the object is described, along with experimental results demonstrating the high efficiency of this method in reconstructing images with reduced projection data.
Accepted on Scientific Board on:	25. 03. 2021
Defended: (Filled by the faculty service)	
Thesis Defend Board: (title, first name, last name, position, institution)	President: Dr. Biljana Mihailović, full professor at Faculty of Technical Sciences, University of Novi Sad Member: Dr. Tatjana Davidović, research professor at Mathematical Institute, Belgrade Member: Dr. Lidija Čomić, associate professor at Faculty of Technical Sciences, University of Novi Sad Member: Dr. Buda Bajić Papuga, assistant professor at Faculty of Technical Sciences, University of Novi Sad Member, Supervisor: Dr. Tibor Lukić, full professor at Faculty of Technical Sciences, University of Novi Sad
Note:	

DEDICATION

To my two Mirjanas, for one has shown me what it means to be unconditionally loved, and the other what it means to love unconditionally...

Abstract

Tomography refers to various imaging methods that use penetrating waves to gather data about an inaccessible or obscured object of interest. In general, the waves are systematically propagated through the object from multiple directions in order to gather the necessary data for a good reconstruction. The object under reconstruction is perceived as a function characterized by a domain that can either be discrete or continuous, and its output range comprises real numbers. The primary objective is to restore this function by leveraging available data, encompassing integrals or summations over portions of its domain. In the case of *Discrete Tomography* (DT), the function's range is typically a finite collection, commonly employed for reconstructing digital images featuring multiple gray levels, referred to as multi-level discrete tomography. Another specific case is *Binary Tomography* (BT), which focuses on reconstructing binary images.

In DT, very few projections are in ideal situations used for recovering the object. Consequently, DT finds extensive application in domains where the material composition of the object being examined is already known, such as industrial non-destructive testing or electron tomography[41, 42].

So far, only a handful of algorithms have been proposed to address the DT problem, particularly when confronted with the multi-level tomography image reconstruction challenge. These algorithms encompass the Discrete Algebraic Reconstruction Technique (DART) introduced by Batenburg and Sijbers in their work [5], the Multi-Well Potential based method (MWPDT) outlined in the publication by Lukić [57], a hybrid approach that combines non-local projection constraints with continuous convex relaxation for multilabeling problem resolution, and the Non-Linear Discretization function-based reconstruction algorithm (NLD) as described by Chen and coauthors [23]. Nevertheless, it is important to note that certain limitations are associated

with some of these methods, such as the tendency to produce radical solutions, decreased accuracy in reconstructions when confronted with limited projection data, or susceptibility to becoming trapped in local minimum solutions.

In recent years several machine learning algorithms and techniques have been applied in discrete tomography to enhance reconstruction accuracy, speed up the process, or handle specific challenges. These algorithms can be based on Convolutional Neural Networks (CNNs) [48] or Reinforcement Learning [86]. Combining traditional optimization techniques with machine learning approaches can lead to hybrid methods that leverage the strengths of both. For example, using machine learning for initialization or as a post-processing step in combination with optimization algorithms. The downsides of using machine learning in discrete tomography include challenges related to the need for large labeled datasets, interpretability issues with complex models, potential overfitting, computational complexity, and the requirement for expertise in both machine learning and domain-specific knowledge. Additionally, concerns about robustness to noise, limited explainability, and dependency on training distribution may impact the applicability of machine learning in discrete tomography applications. It's important to note that the application of machine learning in discrete tomography is an active area of research, and new methods and improvements continue to emerge.

Graph cuts have emerged as a prominent strategy in the fields of image processing and computer vision for addressing various problem-solving tasks. This methodology revolves around the creation of a dedicated graph to represent the energy function, with the objective of achieving either global or local energy minimization by the identification of the minimal cut on the graph. Efficient computation of the minimum cut can be achieved by utilizing max-flow algorithms, which often yield a solution that comes with theoretical guarantees on its quality.

In DT, the reconstruction problem is usually under-determined with only two projections, and it becomes NP-hard with more than two projection directions. Prior information is crucial in tackling this issue and making the reconstruction process more feasible. This prior knowledge can encompass a wide range of details about the objects in the image. For instance, it might involve knowing that the objects in the image are expected to have specific shapes or adhere to particular geometrical properties, such as being circular or rectangular. By incorporating this prior information into the reconstruction process, we aim to reduce the ambiguity in the problem and improve the accuracy of the final reconstructed image.

This raises an intriguing question: Can shape descriptors be used as prior information? In specific situations, there might be a requirement for reconstructions utilizing only one projection direction. Image reconstruction from a single projection, a challenging task in fields like medical imaging, astronomy, and security, is applied when obtaining multiple measurements is not feasible. For instance, emergency medical X-rays use a single image for rapid injury diagnosis, and in astronomy, occultation events rely on changes in starlight to unveil celestial object characteristics. Security scanners minimize radiation exposure by reconstructing baggage contents from a single scan. In industrial non-destructive testing, single projections are employed to inspect welds or structures. In such instances, harnessing a known shape descriptor of the observed object can be highly beneficial in obtaining satisfactory images.

This thesis provides an overview of discrete tomography, beginning with the definition of digital images, followed by explanations of tomography and discrete tomography, as well as some of the shape descriptors that can be used as a priori information in reconstruction. Existing reconstruction methods in discrete tomography are reviewed. The thesis outlines three innovative approaches to address image reconstruction challenges in discrete tomography. Initially, it introduces a novel method that integrates the gradient approach with the graph cuts method. Subsequently, it enhances this method by incorporating two additional pieces of a priori information—specifically, shape orientation and shape circularity. These pieces of information serve as substitutes in situations where there is a limited number of projection directions.

Rezime

Digitalna slika

Digitalnu sliku matematički možemo definisati kao dvodimenzionalni niz diskretnih vrednosti, gde svaka vrednost predstavlja određenu karakteristiku slike, kao što su intenzitet, boja ili tekstura. Formalno, digitalnu sliku možemo predstaviti kao funkciju $f(x,y)$, gde x i y predstavljaju prostorne koordinate slike, a $f(x,y)$ je vrednost slike na toj određenoj lokaciji.

U matematičkoj notaciji, digitalnu sliku možemo definisati kao:

$$f(x,y) = \begin{cases} I(x,y), & \text{ako } (x,y) \in D \\ 0, & \text{inače,} \end{cases}$$

gde je $I(x,y)$ intenzitet piksela na poziciji (x,y), D je domen slike, i $f(x,y)$ je digitalna slika.

Domen slike D obično se definiše kao oblast u Dekartovom koordinatnom sistemu, u kojoj su vrednosti piksela definisane, dok se pikseli izvan ovog regiona smatraju nulama. Vrednosti intenziteta $I(x,y)$ obično se predstavljaju kao skup diskretnih vrednosti, kao što su celobrojni ili binarni brojevi, u zavisnosti od formata slike i broja bitova koji se koriste za predstavljanje svakog piksela.

Digitalna slika često se predstavlja kao matrica, gde svaki element matrice odgovara vrednosti intenziteta piksela na slici. Na ovaj način, vrednosti digitalne slike mogu se matematički predstaviti i lako obraditi, analizirati i manipulisati.

U matematičkim razmatranjima često se pretpostavlja da je funkcija slike $f(x,y)$ neprekidna i dovoljno puta diferencijabilna. To omogućava primenu različitih metoda matematičke analize, kao što su višestruki integrali, diferencijalni operatori i diferencijalne jednačine. Na ovaj način, dobijeni rezultati (jednačine/operatori) zatim se diskretizuju i prilagođavaju za upotrebu u digitalnom okruženju.

Tomografija

Reč tomografija se odnosi na različite metode snimanja u kojima se koristi talasna energija koja prolazi kroz nepoznati objekat od interesa da bi se prikupili podaci o njemu; obično je taj objekat teško dostupan ili nevidljiv. U većini slučajeva, talasi se moraju slati kroz objekat iz velikog broja različitih pravaca kako bi se prikupilo dovoljno podataka za uspešnu rekonstrukciju. Objekat koji se pokušava rekonstruisati posmatra se kao funkcija sa domenom koji može biti diskretan ili neprekidan, i sa skupom slika koji čine dati skup (obično) realnih brojeva. Zadatak tomografije jeste rekonstrukcija ove funkcije na osnovu poznatih podataka (integrala ili suma podskupova njenog domena).

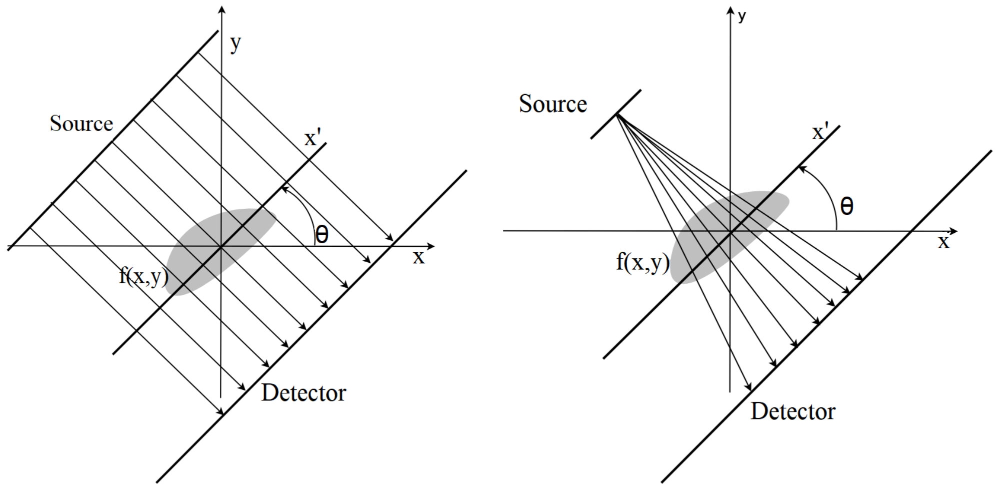
Projektivna geometrija u tomografiji

Glavni problem postavljen u tomografiji je rekonstrukcija informacija o objektu na osnovu posmatranih podataka. Posmatrani podaci se dobijaju merenjem intenziteta talasa koji prodiru kroz objekat iz različitih pravaca. Dok talas iz izvora prolazi kroz objekat, njegova snaga oslabi, i nova vrednost intenziteta talasa se beleži na detektoru. Tu vrednost nazivamo projekcijom.

Projektivna geometrija bavi se odnosima između objekta i projekcije objekta na neko drugo područje (detektor). Najčešći tipovi projekcija u tomografiji su projekcije paralelnog zraka i projekcije lepezastog zraka (Slika 1).

U projekciji paralelnog zraka, kroz objekat prolaze paralelni zraci koji idu od izvora (izvor se kreće u paralelnom smeru) do detektora. U projekciji lepezastog zraka, izvor je fiksan, rotira se oko objekta i šalje radijalne zrake prema detektoru.

U ovoj disertaciji, fokusiramo se na projekcije paralelnog zraka. Geometrija lepezastog zraka može se konvertovati u geometriju paralelnog zraka, tako da se sva razmatranja iz ovog poglavlja mogu primeniti i na projekcije lepezastog zraka [6].



(a) Paralelni zrak (b) Lepezasti zrak

Slika 1: Geometrije slikanja paralelnog zraka i lepezastog zraka

Diskretna tomografija

Diskretna tomografija (DT) [41, 42] može se definisati kao grana tomografije koja se bavi rekonstrukcijom diskretnih objekata ili slika iz ograničenog broja projekcija. Za razliku od kontinuirane tomografije, koja se bavi rekonstrukcijom slika sa kontinuiranim vrednostima intenziteta, diskretna tomografija se usmerava na rekonstrukciju slika sa diskretnim ili binarnim vrednostima. Diskretna priroda objekata ili slika uvodi dodatne izazove i ograničenja u proces rekonstrukcije. Pojam diskretne tomografije je skovao Lari Šep (Larry Shepp), koji je organizovao prvu konferenciju na ovu temu 1994. godine.

U idealnom slučaju, za tačno rekonstruisanje objekta u DT-i, često je potreban mali broj projekcija. Zbog toga DT ima širok spektar primena u oblastima gde su materijali objekta pod istraživanjem poznati unapred, kao što su industrijsko ispitivanje materijala ili elektronska tomografija [41, 42].

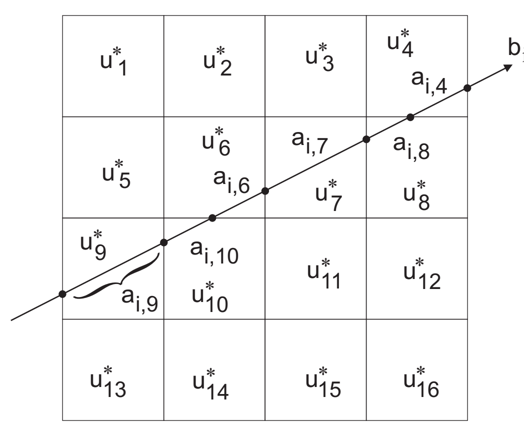
Problemi rekonstrukcije u diskretnoj tomografiji obično se formulišu kao optimizacioni problemi, gde se definiše funkcija cilja koja meri usklađenost između objekta i projekcionih podataka. Potom se traži minimum funkcije cilja, koristeći različite numeričke metode.

Standardna metoda za objašnjavanje procesa prikupljanja projekcionih podataka u diskretnoj tomografiji koristi koncept linija projekcija. Kako bismo to ilustrovali, razmotrimo 2D rešetku koja predstavlja objekat koji treba rekonstruisati. Svaka ćelija u ovoj rešetki može imati vrednost iz diskretnog skupa. Proces prikupljanja projekcionih podataka simulira šta bi se desilo ako bi linije bile projektovane kroz objekat iz različitih pravaca i gde bi se te linije sekle sa objektom.

Kako se prikupljaju projekcioni podaci, sakuplja se informacija o presecima između linija projekcije i ćelija. Sakupljeni projekcioni podaci iz različitih pravaca mogu se organizovati u matricu. Svaki red u matrici odgovara jednom uglu projekcije, a svaka kolona odgovara određenom broju preseka popunjenih ćelija.

Primarni cilj diskretne tomografije je rekonstruisati originalnu rešetku objekta koristeći matricu prikupljenih projekcionih podataka. Ovaj proces uključuje rešavanje inverznog problema: pronalaženje konfiguracije rešetke koja bi dovela do posmatranih projekcionih podataka.

Posmatrajmo sliku u^* dimenzija $N = 4 \times 4 = 16$. Slika 2 prikazuje primer izračunavanja vrednosti projekcije na u^*. Linija projekcije prodire kroz piksele slike. Vrednost projekcije b_i računa se kao $b_i = a_{i,4}u_4^* + a_{i,6}u_6^* + a_{i,7}u_7^* + a_{i,8}u_8^* + a_{i,9}u_9^* + a_{i,10}u_{10}^*$.



Slika 2: Primer izračunavanja vrednosti projekcije na slici

Diskretna tomografija se suočava sa izazovima povezanim sa diskretnom prirodom reprezentacije objekta, što rezultira specijalizovanim metodama dizajniranim posebno za rekonstrukcije sa diskretnim ili binarnim rešetkama.

Jedan od izazova u problemima rekonstrukcije u diskretnoj tomografiji je suočavanje sa šumom i nesigurnošću u projekcionim podacima. Projekcioni podaci mogu biti pogođeni šumom, kao što su elektronski šum ili rasipanje svetlosti, što može otežati preciznu rekonstrukciju objekta. Pored toga, projekcioni podaci mogu sadržati slučajnu odnosno stohastičku komponentu, što znači da se merenja ne znaju tačno. Stoga, metode rekonstrukcije često uključuju regularizacioni član koji pomaže stabilizaciji rešenja i smanjenju uticaja šuma i nesigurnosti.

Još jedan izazov u diskretnoj tomografiji je kombinatorna priroda problema. Skup mogućih rešenja je neodređen, što znači da postoji vrlo veliki broj mogućih objekata koji bi mogli proizvesti dati skup merenja. Stoga se metode rekonstrukcije često oslanjaju na neki oblik a priori znanja o objektu kako bi se problem učinio manje neodređenim.

Problem rekonstrukcije u diskretnoj tomografiji može se prikazati sledećim linearnim sistemom jednačina:

$$a_{11}u_1 + a_{12}u_2 + a_{13}u_3 + \ldots + a_{1N}u_N = b_1$$

$$a_{21}u_1 + a_{22}u_2 + a_{23}u_3 + \ldots + a_{2N}u_N = b_2$$

$$a_{31}u_1 + a_{32}u_2 + a_{33}u_3 + \ldots + a_{3N}u_N = b_3$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$a_{M1}u_1 + a_{M2}u_2 + a_{M3}u_3 + \ldots + a_{MN}u_N = b_M,$$

koji se može posmatrati u matričnoj formi kao:

$$A\,u = b, \tag{1}$$

gde su: $A \in \mathbb{R}^{M\times N}$, $u \in \Lambda^N$, $b \in \mathbb{R}^M$ i $\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_k\}$ za $k \geq 2$.

Zadatak je rekonstruisati sliku predstavljenu nepoznatim vektorom u. Opseg mogućih vrednosti za sliku, predstavljen skupom Λ, definisan je od strane korisnika i može biti binaran ili sa više nivoa sive boje. Projekcioni podaci su sadržani u projekcionoj matrici A, gde svaki red određuje dužinu preseka između piksela i projekcionih zraka koji prolaze kroz njih. Elementi matrice određeni su dužinom ovih preseka. Projekcioni vektor b računa se kao suma proizvoda intenziteta piksela i dužine projekcionih zraka koji prolaze kroz njih.

Proces projekcije koristi različite pravce i koristi paralelnu metodu projekcije zraka, gde se uzima više paralelnih projekcionih zraka za svaki pravac. Ugao α odgovoran je za određivanje pravca projekcije. Kako bi se osigurala pokrivenost cele rešetke slike, udaljenost između susednih paralelnih projekcionih zraka jednaka je dužini stranice piksela i ravnomerno su raspoređeni. Broj paralelnih projekcionih zraka je specificiran kako bi se pokrila cela rešetka slike.

Rekonstrukcioni problem podrazumeva pronalaženje slike u predstavljene linearnim sistemom jednačina (1), koristeći projekcionu matricu A i vektor b. Ovaj sistem često nema jednistveno rešenje, sa $N > M$. Cilj je ne samo pronaći rešenje koje se podudara sa datim projekcijama, već i rešenje koje blisko odražava originalnu sliku. Da bi se dobilo visokokvalitetno i zadovoljavajuće rešenje, potrebno je koristiti sva dostupna znanja (a priori informacije) o objektu u pitanju.

Metodi minimizacije energije su moćne tehnike koje se koriste u obradi slika za rešavanje različitih problema, kao što su uklanjanje šuma sa slike, segmentacija slika, popunjavanje slika i obnova slika. Ovi metodi imaju za cilj da pronađu optimalnu konfiguraciju energetske funkcije koja predstavlja grešku ili neslaganje između obrađene slike i željenog rezultata.

Fundamentalni koncept uključuje formulisanje problema diskretne tomografije kao modela minimizacije, gde je cilj pronaći vrednost funkcije u koja minimizuje funkcionelu $E(u)$. Tipično, u ovom kontekstu, u predstavlja sliku. Pojam energija potiče iz analogije u fizici, gde stabilan sistem karakteriše minimalna ukupna energija.

U svakom pristupu minimizaciji energije, moraju se zadovoljiti dva ključna kriterijuma. Prvo, dizajn energetske funkcije ili modela mora usko odražavati stvarni problem koji se rešava, a njen minimum, po mogućstvu globalni, treba predstavljati optimalno rešenje problema. Drugo, optimizacioni algoritam koji se koristi za minimizaciju energije treba da bude brz i precizan, omogućavajući dobro aproksimiranje minimalne vrednosti uz efikasnu upotrebu dostupnih računarskih resursa. Nepoštovanje bilo kog od ovih kriterijuma može značajno smanjiti efikasnost metoda ili ih potpuno učiniti neprikladnim za praktične primene.

U obradi slika, energetska funkcija (poznata i kao funkcija cilja) kvantifikuje kvalitet ili prikladnost date slike. Energetska funkcija obično se sastoji od dve glavne komponente: funkcije koja meri sličnost rekonstrikcije prikupljenim projekcionim podacima i regularizacione funkcije.

Prilikom primene minimizacije energije u rekonstrukciji slika, u najopštijem kontekstu, pokušava se dobiti rekonstruisana forma posmatrane slike u minimizacijom sledeće funkcije:

$$E(u) = F(Lu, b) + \lambda R(u). \tag{2}$$

Argument u^r koji minimizuje energetsku funkciju,

$$u^r = \arg\min_u E(u) \tag{3}$$

smatra se procenom originalne slike.

Funkcija F, koja meri udaljenost između projekcionih podataka b i rekonstrukcije u nakon primene operatera L, procenjuje koliko dobro rekonstruisana slika odgovara posmatranim podacima. U kontekstu uklanjanja šuma sa slike, član za vernost podacima kažnjava razlike između slike degradirane šumom i slike bez šuma.

Regularizaciona funkcija R nameće a priori znanje o rešenju u. Očekuje se da niske vrednosti R doprinose, do određene mere, eliminaciji neželjenih karakteristika. Regularizacija takođe obezbeđuje numeričku stabilnost problema. Regularizacioni parametar λ reguliše balans između ova dva člana, odnosno nivo uklanjanja šuma u odnosu na precizno rekonstruisanje detalja slike. Regularizacioni član podstiče određene osobine ili karakteristike u rekonstruisanoj slici i sprečava pojavu prekomplikovanih rešenja ili rešenja sa dosta šuma. Regularizacija pomaže u sprečavanju preprilagođavanja i proizvodi vizuelno privlačnije rezultate.

Nalaženje analitičkog rešenja za problem (2) obično nije izvodljivo zbog njegove dimenzionalnosti. Stoga je potreban odgovarajući optimizacioni pristup za njegovo rešavanje.

Član za podudaranje podataka u E može se oblikovati kao konveksna kvadratna funkcija u obliku sume kvadrata grešaka, što čini njegovo numeričko rešavanje relativno jednostavnim. Metoda konjugovanih gradijenata pokazala se kao jedan od najefikasnijih pristupa za minimizaciju ove funkcije, pružajući rešenje u najviše N (dimenzije u) iteracija. Iako se ovaj čllan često predstavlja kao konveksna kvadratna funkcija, važno je napomenuti da može odstupati od ovog oblika. Nekvadratni članovi za podudaranje podataka javljaju se u složenim sistemima ili modelima, stvarajući probleme za numeričku obradu. U takvim slučajevima, tradicionalne metode optimizacije razvijene za kvadratne probleme mogu biti manje efikasne [20].

Međutim, regularizacioni član R može imati potpuno drugačiji oblik. Može pokazivati visoku nelinearnost, negativnost, pa čak i nediferencijabilnost u određenim tačkama. Ove karakteristike čine minimizaciju E izazovnom. Visoka nelinearnost povećava računarsku složenost tokom numeričkih evaluacija, dok negativnost može rezultirati time da energetska funkcija E nije konveksna, što otežava određivanje globalnog minimuma. Osim toga, nediferencijabilnost funkcije R implicira da je i E nediferencijabilna, čime mnoge metode minimizacije zasnovane na gradijentu ili izvodima višeg reda postaju neprikladne. Ovo je veliko ograničenje, jer se mnoge efikasne determinističke metode oslanjaju na računanje gradijenata.

Analiza gore navedenog pokazuje da problemi obrade slika sa regularizacijom nisu uvek *dobro postavljeni* (well-possed). Na primer, kada energetska funkcija sadrži negativan regularizacioni član, to može dovesti do negativnosti problema, što može rezultirati višestrukim lokalnim minimumima bez jedinstvenog globalnog minimuma. U takvim slučajevima, glavni doprinos regularizacije je sužavanje prvobitnog skupa rešenja, ali ne i nužno dovođenje do jedinstvenog rešenja.

Problem (2) predstavlja problem neprekidne optimizacije. Međutim, određene primene, kao što su diskretna tomografija ili defazifikacija, ograničavaju prostor pretrage na diskretan skup. Ograničeni problem sa regularizacijom formulisan je kao:

$$\min_{u \in \Omega} E_Q(u), \tag{4}$$

gde Ω predstavlja dopustiv skup. Ispunjavanje uslova ograničenja predstavlja dodatni izazov, uz prethodno analizirane probleme, koji treba rešiti. Jedan mogući pristup je transformisati ograničeni problem u neprekidni putem reformulacije uslova ograničenja kao novog regularizacionog sabirka. Primer takvog pristupa je konveksno-konkavna regularizacija u diskretnoj tomografiji [84]. Alternativno, drugi način za rešavanje ovog izazova je direktna primena odgovarajuće metode optimizacije posebno dizajnirane za probleme sa ograničenjima.

U rekonstrukciji slika, regularizacija je tehnika koja se koristi za ograničavanje prostora rešenja inverznog problema kako bi se dobilo jedinstveno i stabilno rešenje. Inverzni problem u rekonstrukciji slika odnosi se na zadatak procene nepoznate slike na osnovu datog skupa merenja ili posmatranja. Merenja ili posmatranja mogu biti degradirana šumom ili mogu biti nepotpuna, što inverzni problem čini *slabo postavljenim* (ill-posed problem).

Regularizacija se može definisati kao dodatan sabirak koji se dodaje funkciji cilja koja se minimizuje kako bi se pronašlo rešenje inverznog problema. Regularizaciona funkcija nameće određene osobine rešenju, kao što su glatkost, orijentacija i slično, što rešenje čini smislenijim i manje osetljivim na šum ili nepotpunost podataka.

Regularizacioni član često je funkcija slike sama po sebi i bira se na osnovu karakteristika slike i vrste inverznog problema. Regularizacija se može matematički formulisati kao član kazne, uslov ili a priori verovatnoća raspodele slike. Cilj regularizacije je da se dobije stabilno rešenje inverznog problema.

Postoji nekoliko funkcija za regularizaciju koje se često koriste u rekonstrukciji slika, svaka sa svojim prednostima i manama.

Do sada je samo nekoliko algoritama predloženo za rešavanje problema DT-a. Ti algoritmi uključuju Diskretnu Algebarsku Rekonstrukcionu Tehniku (DART) [5], metodu zasnovanu na Višestrukim Potencijalima (MWPDT) [57], kombinaciju nepostojanih projekcionih ograničenja sa kontinuiranim konveksnim relaksiranjem problema višestrukog obeležavanja [99] i algoritam rekonstrukcije zasnovan na nelinearnoj diskretizaciji funkcije (NLD) [92]. Međutim, neki od ovih metoda imaju ograničenja, kao što su pružanje radikalnih rešenja, manje tačne rekonstrukcije, posebno kada se smanji broj projekcija, ili ostajanje u lokalnim minimumima, tj. polu-kontinualna rešenja.

U prethodnih par godina metode mašinskog učenja su počele da imaju primenu i u rekonstrukcija slika u tomografiji [48, 86]. Kombinovanje tradicionalnih tehnika optimizacije sa pristupima mašinskog učenja može dovesti do hibridnih metoda koje koriste prednosti oba pristupa. Na primer, upotreba mašinskog učenja za inicijalizaciju ili kao postupak postprocesiranja u kombinaciji sa tradicionalnim algoritmima optimizacije. Za adekvatnu primenu metoda mašinskog učenja, neophodno je imati veliku bazu podataka na osnovu kojih bi se mogao trenirati algoritam. Dodatno, mana ovakvog pristupa leži u nemogućnosti poptpune interpretacije rešenja, kao i u velikoj računarskoj kompleksnosti.

Jedan od često korišćenih pristupa za rešavanje problema obrade slika i računarske vizije je zasnovan na presecanju grafova. Ovaj pristup podrazumeva izgradnju specijalizovanog grafa za datu funkciju energije, tako da minimum preseka na grafu minimizuje energiju, bilo globalno ili lokalno. Minimum preseka se može efikasno izračunati pomoću algoritama maksimalnog protoka, a rezultat obično pruža rešenje sa teoretskim garancijama kvaliteta.

U DT-u, problem rekonstrukcije je obično neodređen sa samo dve projekcije, dok postaje NP-težak sa više od dva projekciona pravca [38]. Da bi se smanjila neodređenost, često se u rekonstrukciju uključuju unapred poznate informacije o objektu, kao što su konveksnost, povezanost, homogenost, sličnost sa modelnom slikom, obim i orijentacija. To postavlja pitanje da li se i koji deskriptori oblika mogu koristiti kao a priori informacije. Potreba za rekonstrukcijama koristeći samo jedan pravac projekcije može se javiti kada su i projekcioni sistem i posmatrani objekat fiksirani. Primer takve situacije može biti ispitivanje materijala u građevinarstvu, kao što je ispitivanje strukture zida. Primene u medicini mogu zahtevati izuzetno nisku radijaciju za pacijenta. U slučajevima kada je određeni deskriptor oblika posmatranog objekta (npr.

određeni unutrašnji organ) poznat, metoda rekonstrukcije koja koristi ove informacije može obezbediti prihvatljive slike koristeći samo jedan pravac projekcije, čime se smanjuje doza radijacije za pacijenta.

Ova teza pruža pregled diskretne tomografije, počevši od definicije digitalnih slika, zatim definisanja tomografije i diskretne tomografije, kao i pregled postojećih metoda rekonstrukcije u diskretnoj tomografiji i postojeće deskriptore slike koji mogu da se koriste kao a priori informacija prilikom rekonstrukcije slike. Na kraju, teza pruža definiciju novih metoda rekonstrukcije u diskretnoj tomografiji koje pokazuju bolje performanse u poređenju sa postojećim metodama.

Metode za rekonstrukciju slika u digitalnoj tomografiji zasnove na sečenju grafova

Graf G je uređen par $G = (X, \rho)$, gde je X konačan neprazan skup elemenata koji se zovu čvorovi, a ρ je konačan skup uređenih, ili neuređenih parova, različitih elemenata skupa X koji se zovu grane. Ukoliko je ρ skup uređenih parova, reč je o usmerenom grafu, u suprotnom imamo neusmereni graf. Ukoliko se svakoj grani grafa dodeli neka vredonst, dobijamo ponderisani (težinski) graf.

Tehnika sečenja grafa je moćna tehnika koja se koristi u obradi slike za različite zadatke, poput segmentacije slike, prepoznavanja objekata i matiranja slike. Ovaj koncept uključuje particionisanje slike na dva ili više segmenata pronalaženjem optimalnog reza kroz graf, kojim je predstavljena slika.

U tehnikama rezanja grafa, slika se predstavlja kao graf, gde je svaki piksel čvor, a grane predstavljaju odnose između piksela. Cilj tehnike rezanja grafa je pronaći particiju grafa koja minimizuje funkciju cilja, u ovu svrhu mogu se koristiti različiti algoritmi, kao što su algoritam maksimalnog protoka-minimalnog reza ili metoda sečenja grafa zasnovana na dinamičkom programiranju.

Optimizacija uz upotrebu tehnike sečenja grafa pruža praktičan pristup rešavanju raznovrsnih izazova u obradi slike koji se mogu izraziti u kontekstu minimizacije energije, kako je dokumentovano u različitim istraživanjima [8, 12, 13, 15, 16, 49, 52, 55].

Potsov model

U teoriji sečenja grafa, Potsov model [95] se koristi kao način formulisanja i rešavanja optimizacionih problema koji uključuju grafove. Cilj je podeliti čvorove grafa u grupe tako da su grane koje povezuju čvorove unutar grupe minimizovane, dok su grane koje povezuju čvorove između grupa maksimizovane. Ovo može biti korisno za zadatke poput segmentacije slike, gde je cilj podeliti sliku na različite regione ili objekte. U našoj primeni, Potsov model se zasniva na minimizaciji sledeće energije:

$$E(d) = \sum_{p \in \mathcal{P}} D(p, d_p) + \sum_{(p,q) \in \mathcal{N}} K_{(p,q)} \cdot (1 - \delta_{d_p, d_q}), \tag{5}$$

gde $d = \{d_p \,|\, p \in \mathcal{P}\}$ označava intenzitet piksela slike $p \in \mathcal{P}$. Izraz $D(p, d_p)$ označava trošak dodeljivanja oznake d_p pikselu p. Potencijalna interakcija između parova susednih piksela p i q označava se kao $K_{(p,q)}$, a $\mathcal{N}$ predstavlja skup susednih parova piksela. Funkcija δ_{d_p, d_q} je Kronekerova delta funkcija.

Primena metoda sečenja grafova u diskretnoj tomografiji

Tehnika za rekonstrukciju slike opisana u ovom odeljku je prvi put predstavljena od strane Šulc i Lukića u [59] i naziva se metodom rekonstrukcije Diskretne Tomografije Tehnikom Sečenja Grafa (GCDT).

Prvi korak u procesu rekosntrukcije je pronalazak neprekidnog rešenja problema (6).

$$\min_{u \in [0,1]^N} E_Q(u) := \|Au - b\|^2. \tag{6}$$

Gde je, E_Q kvadratna funkcija nad skupom $\Omega = [0, 1]^N$. Za rešavanje ovog problema minimizacije koristimo algoritam Spektralnog Projektovanog Gradijenta (SPG) [10].

Sledeći korak u rekonstrukciji je diskretizacija neprekidnog rešenja jednačine (6), dobijenog primenom SPG algoritma. U ovu svrhu, koristimo algoritam predložen u

[85, 93] i [58, 61], i biramo Potsov model interakcije zbog njegove sposobnosti da promoviše kompaktnost u rešenjima, kako je primećeno u [14, 36, 89]. Izraz D, u (5.2) je formulisan na osnovu intenziteta piksela, $u(p)$, i dizajniran je tako da bude mali ili jeftin u blizini određenih sivih vrednosti.

$$
\begin{aligned}
&D(p,0) = |u(p) - \lambda_1|,\\
&D(p,1) = |u(p) - \lambda_2|,\\
&D(p,2) = |u(p) - \lambda_3|,\\
&\vdots\\
&D(p,k-1) = |u(p) - \lambda_k|.
\end{aligned}
$$

Potencijal interakcije, $K_{(p,q)}$, između susednih piksela je postavljen kao konstantna vrednost 1. Energetska funkcija (5) se zatim minimizuje koristeći algoritam za Optimizaciju Sečenjem Grafa(GCO) [14, 16, 27, 53]. GCO algoritam dodeljuje vrednost oznake, d_p, svakom pikselu, koja odgovara unapred definisanom nivou sive boje i određuje intenzitete piksela u konačnom diskretnom rešenju.

Metoda zasnovana na orijentaciji oblika

Naš pristup rekonstrukciji tomografije sjedinjuje metodu sečenja grafa sa tehnikom minimizacije baziranom na gradijentu, koristeći orijentaciju oblika kao ključnu a priori informaciju.

U početnom koraku naše metode, izračunavamo vrednosti troškova podataka za svaki piksel unutar slike. Ove vrednosti proizlaze iz intenziteta kontinualno aproksimirane konačne rekonstruisane slike, postignute minimizacijom energetske funkcije:

$$
\begin{aligned}
\min_{u\in[0,1]^N} E_Q(u) = &w_P\|Au-b\|^2 + w_H \sum_{i=1}^{N} \sum_{j\in\Upsilon(i)} (u_i - u_j)^2 \\
&+ w_{\mathcal{O}}(\Phi(u) - \alpha^*)^2 + \mu\langle u, \tau - u\rangle.
\end{aligned} \tag{7}
$$

Ovde uvodimo ključne elemente:

- $\tau = [1, 1, \ldots, 1]^T$ kao N-dimenzionalni vektor,
- w_P koeficijent za kontrolu prilagođavanja podacima,
- w_H koeficijent za regulisanje kompaktnosti rešenja,
- $\Upsilon(i)$ predstavljanja indekse susednih piksela piksela i,
- $\Phi(u)$ za orijentaciju trenutnog rešenja,
- α^* kao stvarnu orijentaciju (a priori informacija),
- $w_{\mathcal{O}}$ koeficijent za određivanje uticaja regularizacije orijentacije,
- $\langle u, \tau - u \rangle$ za podsticanje intenziteta piksela prema binarnim vrednostima,
- μ za kontrolu uticaja binarizacije.

Za svako fiksirano μ, koristimo iterativni optimizacioni algoritam SPG za rešavanje problema (7). U sledećem koraku, vršimo sveobuhvatnu binarizaciju neprekidnog rešenja dobijenog algoritmom SPG. Ova binarizacija koristi metodu sečenja grafa baziranu na Potsovom modelu. Funkcija troška podataka D u (5) oblikovan je na osnovu informacija dobijenih iz glatkog rešenja u:

$$D(p, 0) = u(p),$$

$$D(p, 1) = 1 - u(p).$$

Takođe, definišemo skup susednih parova $\mathcal{N}$, sa $(p, q) \in \mathcal{N}$. Dva piksela su susedna kade se njihove koordinate na slici razlikuju za samo jednu vrednost. Potencijal interakcije $K_{(p,q)}$ je 1.

S ovim definicijama, spremni smo da minimiziramo energetsku funkciju u (5) koristeći algoritam za optimizaciju rezanja grafa (Graph Cuts Optimization - GCO)

[16]. Algoritam GCO dodeljuje vrednosti oznaka $d_p \in 0, 1$ svakom pikselu p. Ovde opisana metoda je predložena u [66].

Metoda zasnovana na cirkularnosti oblika

Metoda rekonstrukcije za rešavanje problema diskretne tomografije predložena u [67] i ukratko opisana u ovom odeljku sastoji se od dva dela:

1. Pronalaženje kontinualnog rešenja problema minimizacije energije koristeći metod minimizacije baziran na gradijentu. Energetska funkcija uključuje informacije o cirkularnosti originalnog objekta.

2. Diskretizacija dobijenog neprekidnog rešenja primenom algoritma baziranog na sečenju grafa. Vrednosti piksela neprekidne slike koriste se za definisanje funkcije troška podataka za graf.

Energetska funkcija koja se koristi za izračunavanje neprekidnog rešenja data je sledećom jednačinom:

$$\begin{aligned}\min_{u \in [0,1]^N} E_Q(u) := w_P \|Au - b\|_2^2 + w_H \sum_{i=1}^{N} \sum_{j \in \Upsilon(i)} (u_i - u_j)^2 + \\ + w_C \left(\mathcal{C}(u) - \mathcal{C}^*\right)^2 + \mu \left\langle u, \tau - u \right\rangle ,\end{aligned} \tag{8}$$

i sastoji se od sledeća četiri člana:

1. Član prilagođavanja podacima, $\|Au - b\|_2^2$, regulisan parametrom $w_P > 0$, koji osigurava sličnost projekcionim podacima.

2. Funkcija homogenosti, $\sum_{i=1}^{N} \sum_{j \in \Upsilon(i)} (u_i - u_j)^2$, regulisana parametrom $w_H > 0$; Ovde, $\Upsilon(i)$ predstavlja skup indeksa susednih piksela pikselu i. Ovaj član osigurava glatkoću rešenja;

3. Član, $(\mathcal{C}(u) - \mathcal{C}^*)^2$, meri udaljenost između cirkularnosti trenutnog rešenja ($\mathcal{C}(u)$) i poznate cirkularnosti originalne slike ($\mathcal{C}^*$). Parametar $w_C > 0$ određuje uticaj regularizacije cirkularnosti.

4. Član regularizacije konkavnosti, $\langle u, \tau - u \rangle$, gde je $\tau = [1, 1, ..., 1]^T$ vektor veličine N. Ovaj član pomaže u pomeranju intenziteta piksela prema binarnim vrednostima, a njegov uticaj postepeno raste tokom procesa rekonstrukcije, regulisan parametrom $\mu > 0$.

Optimizacioni problem (8) je ograničen, kvadratni problem minimizacije energije koji se može rešiti različitim metodama optimizacije. Mi smo izabrali optimizacioni algoritam SPG [9] zbog dobrih performansi u sličnim problemima.

Gradijent regularizacionog člana $(\mathcal{C}(u) - \mathcal{C}^*)^2$ u energetskoj funkciji (8) određen je analitički, što omogućava brzu minimizaciju i određivanje neprekidnog rešenja pomoću SPG algoritma.

Kriterijum zaustavljanja prilikom traženja neprekidnog rešenja je $\langle u, \tau - u \rangle < E_{bin}$, gde E_{bin} reguliše stepen binarizacije rešenja u, i u našim eksperimentima ima vrednost 100.

Nakon izračunavanja neprekidnog rešenja, sledeći korak je njegova potpuna binarizacija. Ovo se postiže primenom Potsove metode.

Navedeni algoritam daje vrednosti oznaka d_p za svaki piksel p, gde je d_p predefinisan kao $d_p = 0 \rightarrow 0$ i $d_p = 1 \rightarrow 1$. Ove vrednosti oznaka određuju intenzitete piksela u konačnom (binarnom) rešenju, označavajući kraj procesa rekonstrukcije.

Naučni doprinosi i originalni rezultati

U tezi je uvedena nova metoda za rekonstrukciju slika zasnovana na minimizaciji funkcije energije u diskretnoj tomografiji koja kombinuje gradijentnu metodu sa metodom sečenja grafova. Ova metoda je objavljena u publikaciji [59] i pokazuje dobre performanse u poređenju sa postojećim metodama koje se bave sličnim problemima.

U tezi je pokazano da se performanse metoda za rekonstrukciju slika u diskretnoj tomografiji sa malim brojem projekcionih podataka mogu značajno poboljšati ukoliko se u njih uključi i a priori informacija o objektu. U tezi je poseban akcenat stavljen na orijentaciju i cirkularnost oblika.

U tezi je predstavljena nova metoda, obljavljena u publikaciji [66], koja integriše metodu sečenja grafa sa tehnikom minimizacije zasnovane na gradijentu, koristeći orijentaciju kao poznatu informaciju o obliku. Metoda pokazuje zavidne performanse prilikom rekonstrukcije binarnih slika na osnovu projekcionih podataka iz jedne dimenzije. Dodatno, opisan je novi metod predstavljen u publikaciji [67] koji koristi cirkularnost oblika kao a priori informaciju o objektu, kao i rezultati eksperimenata koji pokazuju veliku efikasnost ovih metoda u rekonstrukciji slika sa smanjenim projekcionim podacima.

Struktura teze

Ova doktorska disertacija je podeljena u šest poglavlja. U uvodnom poglavlju predstavljeni su istraživački problem i njegov značaj. Poglavlje pruža pregled cele teze, ističući ključne ciljeve i doprinose sprovedenog istraživanja. Takođe sadrži listu originalnih radova koji čine osnovu teze. Do kraja ovog poglavlja, čitaoci će imati jasno razumevanje opsega istraživanja i njegove relevantnosti u oblasti digitalne obrade slika i tomografije.

Poglavlje 2 služi kao osnova dalje diskusije. Počinje sa definicijom digitalne slike, opisujući kako se ona predstavlja i anlizira. Zatim se definiše tomografska projektivna geometrija, ističući njen značaj. Uvedena je Radonova transformacija, jedan od osnovnih matematičkih alata u tomografiji, a zatim se detaljno istražuju njena svojstva i primene. Sledeća je predstavljena Furijeova transformacija, naglašavajući njenu važnost u analizi frekvencijskog sadržaja. Poglavlje se završava objašnjenjem Furijeove teoreme i njene primene kroz algoritam filtrirane rekonstrukcije.

Poglavlje 3 prebacuje fokus na diskretnu tomografiju, značajan pristup rekonstrukciji slika sa diskretnim nivoima intenziteta. Počinje formulisanjem problema diskretne tomografije i razmatranjem povezanih izazova, kao i definicijom binarne slike. Poglavlje zatim istražuje problem rekonstrukcije slike u digitalnoj tomografiji. Objašnjen je koncept regularizacije koji se koristi u situacijama gde znamo neke informacije o objektu na slici koji se rekonstruiše, ali se rekonstrukcija vrši na osnovu malog broja projektivnih podataka. Predstavljen je Rajserov algoritam, zajedno sa drugim algebarskim tehnikama rekonstrukcije kao što su Algebraic Reconstruction Technique (ART), Simultaneous Iterative Reconstruction Technique (SIRT), Simultaneous Algebraic Reconstruction Technique (SART), Discrete Algebraic Reconstruction Technique (DART) i Simulated Annealing Algorithm. Takođe, razmatrane su metode rekonstrukcije zasnovane na gradijentu kako bi se pružio širi uvid u različite pristupe tomografskoj rekonstrukciji.

U poglavlju 4 predstavljeni su desktiptori oblika, kao što su geometrijski momenti i momenti invarijantnosti, orijentacija i cirkularnost oblika. Pomenuti deskriptori oblika su temeljno opisani, naglašavajući njihov potencijal u analizi i rekonstrukciji slika.

Poglavlje 5 prikazuje originalne doprinose disertacije. Prvi glavni doprinos je primena optimizacije sečenjem grafa. Koncept sečenja grafa je detaljno objašnjen, kao i originalna ideja njegove primene u rekonstrukciji slike u diskretnoj tomografiji. Predstavljeni su eksperimentalni rezultati za metodu zasnovanu na sečenju grafa, koji pokazuju njenu efikasnost i performanse u poređenju sa postojećim pristupima. Poglavlje zatim integriše informaciju o cirkularnosti i orijentaciji oblika u metode rekonstrukcije zasnovane na sečenju grafa, što dovodi do razvoja dve nove strategije za rekonstrukciju. Detaljna objašnjenja ovih metoda, zajedno sa eksperimentalnim rezultatima, pružaju ocenu njihovih performansi i mogućih prednosti.

U zaključnom poglavlju 6, sumirani su glavni rezultati i doprinosi disertacije. Poglavlje potvrđuje značaj istraživanja i njegov potencijalni uticaj na oblast diskretne tomografije i rekonstrukcije slika i daje predlog mogućih budućih pravaca istraživanja.

Disertacija se zatvara sveobuhvatnim popisom svih referenci citiranih u celom dokumentu. Ovaj deo prepoznaje izvore i prethodna istraživanja koja su podržala i uticala na disertaciju.

Acknowledgements

I would like to express my utmost gratitude to my supervisor, Dr. Tibor Lukić, whose tremendous guidance and support were vital in the successful completion of my work. As a consequence of his feedback, I gained a fresh viewpoint and adopted a more critical approach towards my studies. Tibor was motivating, understanding, supportive, and always willing to share his knowledge and provide me with new opportunities, and for this, I will forever be grateful.

I extend my heartfelt appreciation to the members of my thesis defense committee for their invaluable contributions to the completion of this research. I am sincerely grateful to Dr Biljana Mihailović, Dr Tatjana Davidović, Dr Lidija Čomić, and Dr Buda Bajić Papuga for dedicating their time to review and assess my work. Their thoughtful questions, insightful suggestions, and constructive criticism have significantly enriched the content of this thesis.

Additionally, I want to thank to my family and closest friends for their constant and unwavering encouragement and support. Their belief in my abilities, even during the most challenging moments of this academic journey, has been a source of strength and motivation. Their understanding and patience have allowed me to fully immerse myself in the research process, and for that, I am truly grateful.

Finally, I extend my gratitude to my two little beans, whose presence provided the ultimate motivation and inspiration for completing this thesis. I hope you can forgive me for keeping you awake during so many nights while working on this endeavor.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ART	Algebraic Reconstruction Technique
BP	Basis Pursuit (algorithm)
BT	Binary Tomography
CMYK	Cyan, Magenta, Yellow, Black
CNN	Convolutional Neural Networks
CSP	Constraint Satisfaction Problem
CT	Computed Tomography
DART	Discrete Algebraic Reconstruction Technique
DT	Discrete Tomography
DVT	Digital Volume Tomography
EIT	Electrical Impedance Tomography
FBP	Filtered Back Projection
GCCIRCBT	Graph Cuts Binary Tomography Assisted by the Circularity prior
GCDT	Graph Cuts Discrete Tomography
GCO	Graph-Cut based optimization
GCORIENTBT	Graph Cuts Tomography Assisted by the Orientation prior
LAB	Lightness, a, b
MCMC	Markov Chain Monte Carlo
MRI	Magnetic Resonance Imaging
MWPDT	Multi Well Potential based method
m.r.	Misclassification Rate

NLM	Non-local Means
NLD	Non-Linear Discretization
OCT	Optical Coherence Tomography
PE	Pixel Error
PET	Positron Emission Tomography
PRE	Projection Error
RGB	Red, Green, Blue
ROF	Rudin-Osher-Fatemi (model)
SA	Simulated Annealing
SART	Simultaneous Algebraic Reconstruction Technique
SIRT	Simultaneous Iterative Reconstruction Technique
SPECT	Single-Photon Emission Computed Tomography
SPG	Spectral Projected Gradient
TRDT	Method based on classical threshold
TV	Total Variation

CHAPTER 1

Introduction

1.1 Thesis motivation and contribution

Tomography is an umbrella term for a broad range of imaging techniques that rely on the process of image slicing. Tomographic imaging can be divided into two categories:

- Transmission tomography
- Emission tomography

In transmission tomography, the source of the radiation is outside the object. A source is transmitting low energy ray through the object, a wave penetrates the object, and part of the energy is attenuated. The detector then registers the attenuated intensity of the wave. From the intensity of these projections, a reconstruction of the image of the object is being made.

On the other hand, in emission tomography, radioactive substances are injected and redistributed into the object. The unstable radionuclide decays by generating γ-rays, detected by a detector array encircling the target. The acquired multiview projection data set is then used to reconstruct the image.

In this thesis, we are focused only on transmission tomography.

In recent decades, we have seen enormous growth in the development and usage of tomography. The main reason for the development of this field is its wide application in many spheres of modern life.

In 1930 radiologist Alessandro Vallebona developed the basics of tomography in medicine in the form of "classic" X-ray tomography. Subsequently, a multitude of tomographic techniques have been developed and applied within the medical domain to facilitate the early detection of diseases. These methods include ultrasound diagnostics (sonography), computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), single-photon emission computed tomography (SPECT), optical coherence tomography (OCT), electrical impedance tomography (EIT), digital volume tomography (DVT), and "classic" X-ray tomography. The key distinguishing factor among these imaging techniques is their energy sources.

Angiography serves as an illustration of a medical diagnostic technique employing tomography to provide visualizations of the interior of organs and blood vessels, with a primary focus on arteries, veins, and the chambers of the heart. During the procedure, a contrast agent with radio-opaque properties is introduced into the blood vessel, rendering it visible through X-ray-based methods like fluoroscopy.

During the COVID-19 pandemic, Computed Tomography (CT) imaging has played a crucial role in various aspects related to the virus. CT scans have been widely used as a supplementary diagnostic tool to detect and assess lung involvement caused by the SARS-CoV-2 virus. When the initial Polymerase Chain Reaction (PCR) tests were inconclusive or had false-negative results, CT imaging provided valuable information for diagnosing COVID-19 cases. It can reveal characteristic lung abnormalities, such as ground-glass opacities and consolidations, which are indicative of COVID-19 pneumonia.

Apart from diagnosis, CT scans have been instrumental in evaluating the severity and extent of lung involvement in COVID-19 patients, helping healthcare professionals determine the appropriate level of care and treatment. Serial CT scans have also been used to monitor disease progression, allowing clinicians to assess whether the disease is improving or worsening and adjust treatment plans accordingly.

In situations of overwhelmed healthcare systems, CT imaging has assisted in triaging patients based on the severity of lung involvement, enabling more efficient resource allocation. Additionally, CT scans have been a valuable tool in advancing the scientific understanding of COVID-19. By studying CT scans from large cohorts of patients, researchers have gained insights into the patterns of lung involvement, disease progression, and the impact of various treatments.

Furthermore, CT imaging has been used to evaluate lung damage and fibrosis in patients who have recovered from COVID-19. Some individuals may experience persistent lung issues even after recovery, and CT scans aid in assessing the extent of long-term lung damage.

Throughout the pandemic, the responsible and efficient use of CT imaging in COVID-19 cases has been subject to ongoing discussions. Concerns about radiation exposure, resource allocation, and potential virus spread in radiology facilities have been considered, leading to an increased need for image reconstruction methods that require fewer projections for a valid reconstruction.

Next to medicine, tomography has its application in security and cargo inspection since it can be used in materials categorization and detection, thus allowing the exposure of dangerous and prohibited goods [88].

Computed Tomography has emerged as a highly valuable and non-destructive tool in the field of archaeology. Using advanced imaging techniques, CT scans have revolutionized how researchers study artifacts, human remains, and archaeological sites, providing unparalleled insights into the past [45].

In artifact analysis, CT scanning enables archaeologists to delve into the inner compositions of items like pottery, metalwork, and statuettes, revealing hidden features, inscriptions, and construction details that may not be visible to the naked eye. This non-invasive approach allows for thorough examination without the risk of damage.

Human remains examination has also significantly benefited from CT scanning. By producing detailed images of skeletal remains, researchers can study ancient diseases, injuries, and even the mummification processes used in various societies. Additionally, CT scans aid in facial reconstruction and the analysis of burial practices, shedding light on the lives of past civilizations.

Beyond archaeology, CT scanning extends its utility to paleontology, enabling the examination of fossils and other ancient remains. The technology facilitates the investigation of internal structures, offering valuable insights into extinct species' biology and evolutionary history.

Moreover, CT scanning plays a crucial role in exploring monuments and archaeological sites without causing harm to delicate structures. By scanning subsurface layers, archaeologists can discover hidden chambers, tunnels, or artifacts, enhancing our understanding of ancient civilizations.

Preservation of cultural heritage is another crucial aspect of CT's application in archaeology. Fragile or sensitive artifacts can be digitally preserved through CT scans, reducing physical handling and the risk of damage during conservation efforts. In figure 1.1.3, taken from [79], we can see how CT assisted in discovering and restoring objects hidden within soil block.

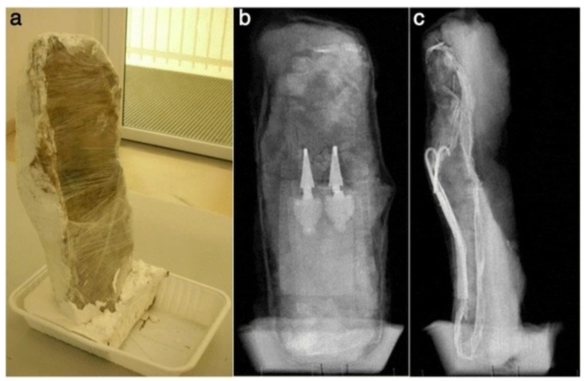


Fig. 1.1.3: Radiographs. A soil block (a) and two radiographs (b-c) of it viewed from different angles (images are taken from [79]).

Additionally, CT scanning has been employed to read and virtually "unroll" ancient scrolls and manuscripts that are too fragile to open physically, enabling access to their contents without causing harm.

Additionally, tomography is used for non-destructive material testing, in the food industry, geophysics, oceanography, and other areas of science.

Tomography is a cost-effective, non-destructive, non-invasive imaging technique that has the potential to undertake many future challenges in various areas of application.

Tomography is inherently interdisciplinary, drawing insights from mathematics, computer science, image processing, and other related fields. This cross-disciplinary nature allows researchers to collaborate and explore innovative approaches to problem-solving. This, together with the diverse application of tomography, was the biggest motivation to start the research presented in this thesis.

The primary objective of here presented research is to address the challenge of improving image reconstruction in discrete tomography using projections obtained from a very limited number of angles. While there exist reconstruction methods that produce relatively good smooth solutions, this thesis goes a step further by concentrating on finding an optimal way to discretize these smooth solutions. By achieving a more refined discretization, we aim to enhance the quality and accuracy of the final image reconstruction.

Within this thesis, we introduce an innovative technique for obtaining image reconstructions of grayscale images using a sparse set of projections. Our method incorporates the Spectral Projection Gradient (SPG) algorithm, which aids in obtaining a smooth solution. Additionally, we employ a method based on graph cuts to further refine the solution and achieve a discrete representation of the reconstructed image.

In many discrete tomography applications, we possess prior knowledge about the object within the image. Taking advantage of this valuable knowledge, we introduce a regularization term into the algorithm, which acts as a constraint during the reconstruction process. In particular, we incorporate information about the shape circularity and shape orientation into the regularization term. By doing so, we guide the reconstruction algorithm to generate solutions that align with the known properties of the object being imaged.

To evaluate the effectiveness of our proposed method, we conducted numerous experiments and comparative analyses. The results conclusively demonstrate that our approach, utilizing the SPG algorithm and graph cuts, outperforms existing methods in terms of both reconstruction accuracy and computational efficiency. By leveraging the available a priori information, we have successfully achieved substantial improvements in image reconstruction from a sparse set of projections.

Overall, this research represents an advancement in the field of discrete tomography, as it tackles the challenge of limited projections by proposing an innovative and effective reconstruction method. The incorporation of a priori knowledge further enhances the reconstruction quality, making it well-suited for various real-world applications where obtaining a sufficient number of projections is challenging or time-consuming.

1.2 Thesis outline

This PhD thesis is divided into six chapters. In the introductory chapter, the research problem and its significance are presented. The chapter provides an overview of the entire thesis, outlining the key objectives and contributions. It also includes a list of the original papers that constitute the core of the thesis. By the end of this chapter, readers will have a clear understanding of the research scope and its relevance in the field of digital image processing and tomography.

Chapter 2 provides a fundamental basis for future analyses and discussions. It starts by introducing the basics of digital images, explaining how they are represented and processed. Tomographic projective geometry is then explored, emphasizing its role in capturing projections from different angles. The Radon Transform, a crucial mathematical tool in tomography, is introduced, followed by an in-depth exploration of its properties and applications. Next, the Fourier Transform is presented, highlighting its relevance in analyzing frequency content. The chapter concludes with an explanation of the Fourier Slice Theorem and its application through the Filtered Back Projection algorithm.

Chapter 3 shifts the focus to discrete tomography, a significant approach in reconstructing images with discrete intensity levels. It begins by formulating the discrete tomography problem and discussing the challenges associated with it. The binary image representation and the 0-1 intensity assumption are introduced. The chapter then delves into the reconstruction problem, exploring how to infer missing information from the acquired projections. Concept of regularization terms, used to enforce smoothness or constraints, is examined in detail. The Ryser algorithm, a popular method for discrete tomography, is presented, along with a comprehensive exploration of algebraic reconstruction techniques such as Algebraic Reconstruction Technique (ART), Simultaneous Iterative Reconstruction Technique (SIRT), Simultaneous Algebraic Reconstruction Technique (SART), Discrete Algebraic Reconstruction Technique (DART), and the Simulated Annealing Algorithm. Additionally, gradient-based reconstruction methods are discussed to provide a broader perspective on different tomographic reconstruction approaches.

Chapter 4 focuses on Area Based Shape Descriptors, such as geometric moments and moment invariants, shape orientation, and shape circularity. Each shape descriptor is explained thoroughly, emphasizing its potential in image analysis and reconstruction.

Chapter 5 showcases the original contributions of the thesis. The first major contribution is the use of Graph Cut Optimization. The concept of graph cuts is explained in detail, and its application to discrete tomography reconstruction is elucidated. Experimental results for the Graph Cut-based method are presented, demonstrating its effectiveness and performance compared to existing approaches. The chapter then integrates shape circularity and orientation into Graph Cuts Reconstruction Methods, leading to the development of two novel approaches. Detailed explanations of these methods, along with their experimental results and quality metrics, are provided to assess their performance and potential advantages.

In the concluding chapter 6, the main findings and contributions of the thesis are summarized. The chapter reaffirms the significance of the research and its potential impact on the field of discrete tomography and image reconstruction. Limitations and potential challenges faced during the research are acknowledged, leading to future research directions and possible improvements to the proposed methods. The chapter concludes with a thought-provoking outlook on the potential applications and extensions of the work presented in the thesis.

The thesis concludes with a comprehensive list of all the references cited throughout the document. This section acknowledges the sources and prior research that have influenced and supported the thesis.

1.3 List of original papers

- **Marčeta, Marina,** and Tibor Lukić. "Regularized graph cuts based discrete tomography reconstruction methods." Journal of Combinatorial Optimization 44.4 (2022): 2324-2346. (M22)
- Ćebić, Dejan, Nebojša M. Ralević Ralević, and **Marina Marčeta**. "An optimal sixteenth order family of methods for solving nonlinear equations and their basins of attraction." Mathematical Communications 25.2 (2020): 269-288. (M22)
- **Marčeta, Marina**, and Tibor Lukić. "Graph cuts based tomography enhanced by shape orientation." International Workshop on Combinatorial Image Analysis. Springer, Cham, 2020. (M33)
- Lukić, Tibor, and **Marina Marčeta**. "Gradient and graph cuts based method for multi-level discrete tomography." International Workshop on Combinatorial Image Analysis. Springer, Cham, 2017. (M33)
- **Marčeta, Marina,** and Tibor Lukić, "Area based shape descriptors for image analysis", Proceedings of 3rd Conference on Mathematics in Engineering: Theory and Applications (META 2018), pp. 52–58, Novi Sad, Serbia, May 12-13, 2018. (M63)
- **Marčeta, Marina,** and Tibor Lukić, "Discrete Tomography Reconstruction Approaches", Proceedings of 2nd Conference on Mathematics in Engineering: Theory and Applications (META), pp. 134-140, Novi Sad, Serbia, June 23-24, 2017. (M63)
- Mihailović Biljana, Ljubo Nedović, and **Marina Marčeta** "Dualnost realnih fazi mera", In Proceedings of the First Conference on Mathematics in Engineering: Theory and Applications (META), Novi Sad, March 4-6, 2016. (M63)

CHAPTER 2

Background

In this chapter, we give a short introduction to tomography and image reconstruction.

2.1 Digital image

A digital image can be mathematically described as a two-dimensional grid of discrete values. Within this grid, each value corresponds to a specific attribute of the image, such as its intensity, color, or texture. To be more formal, we can represent a digital image as a function $f(x, y)$, where x and y denote the spatial coordinates within the image, and $f(x, y)$ signifies the value associated with that particular position.

In mathematical notation, a digital image can be defined as:

$$f(x, y) = \begin{cases} I(x, y), & \text{if } (x, y) \in D \\ 0, & \text{otherwise.} \end{cases}$$

$I(x, y)$ is the intensity value at the location (x, y), D is the domain of the image, and $f(x, y)$ is the digital image.

The domain of the image D is typically defined as a rectangular region in the Cartesian plane, where the pixel values are defined, and the pixels outside of this region are considered to be zero. The intensity values $I(x, y)$ are typically represented by a set of discrete values, such as integers or binary values, depending on the image format and the number of bits used to represent each pixel.

A digital image is often represented as a matrix, where each element of the matrix corresponds to the intensity value of a pixel in the image. In this way, a digital image's values can be represented mathematically, and it is easy to process, analyze, and manipulate.

In mathematical considerations, it is often assumed that the function of an image $f(x, y)$ is continuous and sufficiently many times differentiable. This enables the application of various methods of mathematical analysis, such as multiple integrals, differential operators, and differential equations. In this way, the obtained results (equations/operators) are then discretized and thus adapted for use in a digital environment.

There are several important properties of digital images that are commonly used to describe and analyze them:

Resolution: This property refers to the number of pixels in the image and is typically measured in terms of the number of pixels per inch (ppi) or the number of pixels per centimeter (ppc). The higher the resolution, the more detailed the image will be.

Bit depth: This characteristic denotes the quantity of bits allocated to represent each pixel within the image. Standard bit depths include 8-bit (256 levels of gray or 24-bit color), 16-bit (65536 levels of gray or 48-bit color), and 32-bit (4.3 billion levels of gray or 96-bit color). The higher the bit depth, the more color or tonal variations can be represented in the image.

Color space: This property refers to the color model used to represent the colors in the image. Common color spaces include RGB (red, green, blue), CMYK (cyan, magenta, yellow, black), and LAB (lightness, a, b). Each color space possesses its unique merits and drawbacks, and the selection of a specific color space can impact the ultimate visual presentation of the image.

Compression: This property refers to the method employed for diminishing the size of an image file. Typical compression methods encompass lossless techniques like PNG, which retain all the original information in the image, and lossy methods such as JPEG, which sacrifice some information to achieve a smaller file size.

File format: This property refers to the file type used to save the image. Common file formats include JPEG, PNG, GIF, and TIFF. Each file format has its advantages and disadvantages, and the choice of file format can affect the image's quality, size, and compatibility.

Geometry: This property refers to the shape of the image, including its size, orientation, and distortion. The geometry of an image can be affected by factors such as camera position, lens distortion, and image processing techniques.

Spectral content: This property refers to the distribution of colors or tones in the image. The spectral content of an image can be affected by factors such as lighting conditions, camera settings, and post-processing techniques.

Noise: This property refers to the random variations in the intensity values of the image that are not part of the true signal. Noise can be caused by factors such as camera sensor noise, electronic noise, insufficient light conditions, and image compression.

Contrast: This property refers to the variation in brightness levels between the brightest and darkest areas within the image. The image's contrast can be influenced by factors such as lighting circumstances, camera configurations, and post-processing methods.

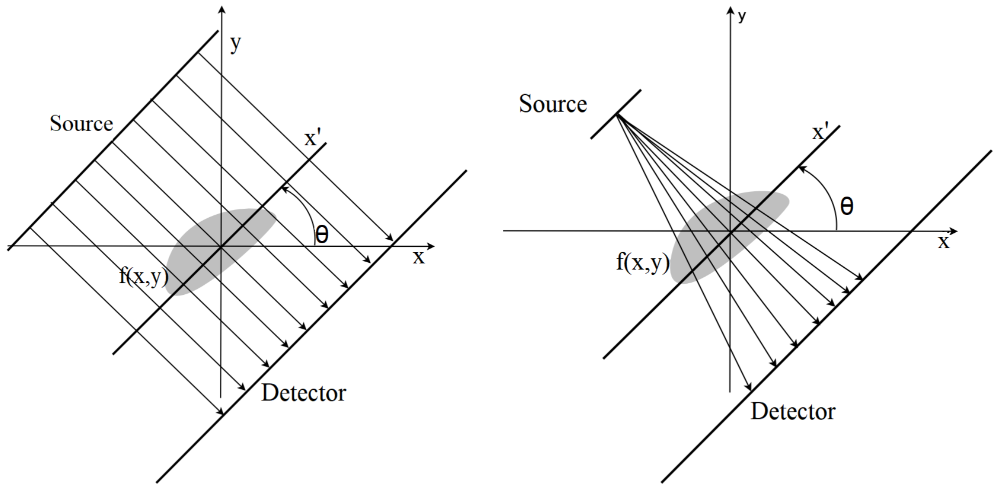
Sharpness: This property refers to the degree of detail and clarity in the image. The sharpness of an image can be affected by factors such as lens quality, camera settings, and post-processing techniques.

2.2 Tomographic projective geometry

The main problem posed in tomography is reconstructing the information about an object based on the observed data. Observed data is obtained by measuring the intensity of the waves penetrating the object from different angles. While a wave from the source passes through the object, its power weakens, and the new value is recorded on the detector. This value we call a projection.

Projective geometry deals with relationships between an object and its projections to some other area. The primary categories of projections in tomography include parallel-beam projections and fan-beam projections (Figure 2.2.4).

In the parallel beam projection, an object is penetrated by parallel rays that



(a) Parallel-beam (b) Fan-beam

Fig. 2.2.4: The imaging geometries of parallel-beam and fan-beam imaging using flat detector

go from a source (the source is moving in a parallel direction) to the detector. In fan-beam projection, the source is fixed, rotates around the object, and sends radial beams towards the detector. Two common types of detectors used in tomography are flat detectors and curved detectors. Flat detectors are essentially planar arrays of detectors arranged in a grid. Unlike flat detectors, curved detectors have a curved shape, which can match the geometry of certain optical systems or imaging setups.

This thesis primarily focuses on two-dimensional $(2-D)$ projections. However, it's important to note that in three-dimensional $(3-D)$ scenarios, there are additional projection geometries. For instance, cone-beam geometry involves the X-ray source emitting rays in a cone shape, covering a larger volume of the object with each projection. Another example is helical or spiral geometry, where the X-ray source and detector continuously rotate around the object, while the object is simultaneously moved along the axis.

In this dissertation, the primary focus centers on parallel beam projections. It is important to note that fan-beam geometry can be transformed into parallel beam geometry [6]. Converting fan beam geometry to parallel beam geometry involves mathematical transformations to reinterpret acquired data. In fan beam imaging, diverging rays are used, and projections are obtained along curved paths. The process typically includes parameterizing the fan beam geometry and applying rebinning techniques (such as Fourier rebinning [26]) to organize data into parallel beam-like projections. Consequently, all the considerations presented in this chapter are equally applicable to fan-beam projections as well.

2.3 Radon transform

When rays are sent from the source to the object, the energy of those waves that do not hit the object stays the same, while the power of those rays that hit the object gets attenuated since some part of it is absorbed by the object. We can think of this process as taking the intensity of every pixel of the object image and adding them up (integrating them) to get an image profile from a specific direction (Figure 2.3.5). By changing the direction of the penetrating wave, we obtain new different profiles of the same object (Figure 2.3.6).

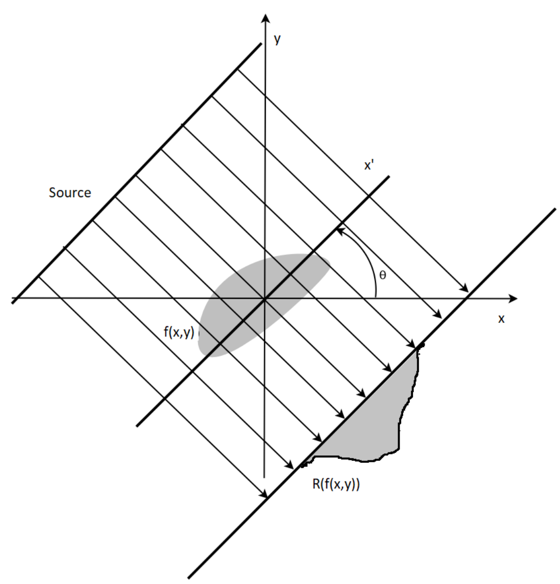
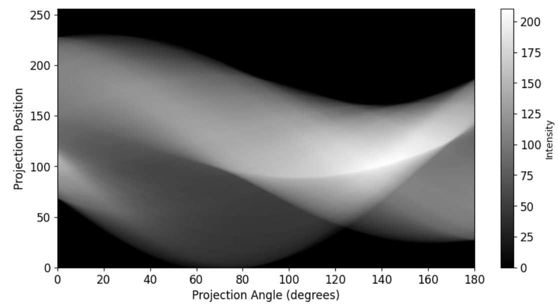


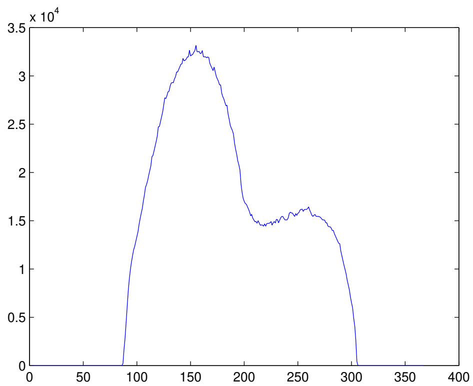
Fig. 2.3.5: Geometric illustration of the Radon transform of a 2-D function

Let us now observe the projection of an image $f(x, y)$ under an angle θ. For each wave a_i and fixed angle θ, there is precisely one intensity detected on the detector.

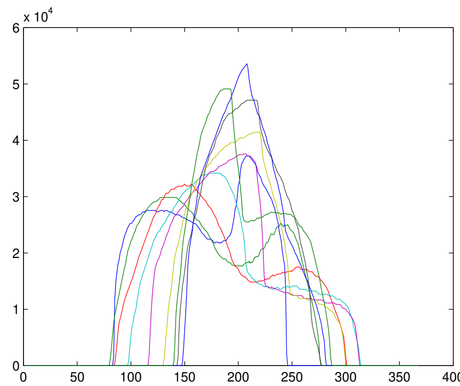
(a) Original image



(b) Sinogram



(c) $\theta = 45°$



(d) $\theta = 0° : 20° : 179°$

Fig. 2.3.6: Original image, its sinogram and Radon transforms from different positions

We can mark that value as $p(\rho, \theta)$. The value $p(\rho, \theta)$ is actually a line integral along the a_i line, after representing a_i in parametric form:

$$a_i : x \cos\theta + y \sin\theta = \rho\,.$$

For every ρ and θ we obtain the value $p(\rho, \theta)$ of the projection wave ρ through the image $f(x, y)$ under angle θ in the following way:

$$p(\rho, \theta) = \mathcal{R}(f(\rho, \theta)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\delta(x \cos\theta + y \sin\theta - \rho) dx dy, \tag{2.1}$$

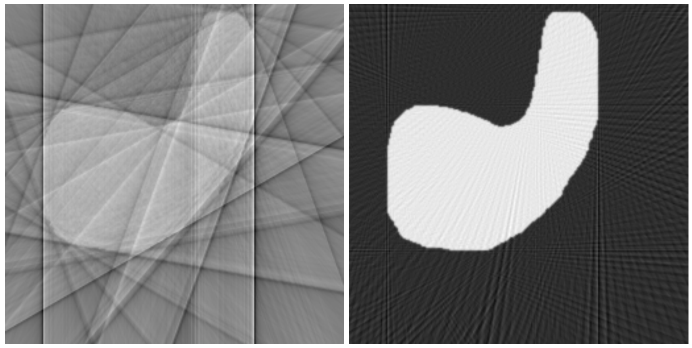
where δ is a Dirac δ function [28]. Transform (2.1) is named Radon transform [76]. By visualizing the function $\mathcal{R}(f(\rho, \theta))$, we get a sinogram. The Radon transform evaluates the integral of the function $f(x, y)$ along the line defined by the angle θ and the distance parameter ρ. Radon transform is the mathematical basis for connecting spatial coordinates (x, y) and projection coordinates (ρ, θ).

We can obtain an image reconstruction if we take the projections from each angle and inverse them (Figure 2.3.7).

The inverse Radon transform $\mathcal{R}^{-1}[\mathcal{R}(f(\rho, \theta))]$ is defined as:

$$\mathcal{R}^{-1}[\mathcal{R}(f(\rho, \theta))] = f(x, y) = \frac{1}{2\pi} \int_{0}^{\pi} \int_{-\infty}^{\infty} \mathcal{R}(\rho, \theta) \cdot \delta(x \cos\theta + y \sin\theta - \rho)\, d\rho\, d\theta. \tag{2.2}$$

Within this equation, the integration is performed over the entire parameter space. The fundamental objective of this integration process is to reconstruct the original spatial distribution of the function $f(x, y)$ in the plane. The process is essentially a transition from the projection domain to the spatial domain. The projection data are methodically merged and re-projected to reconstitute the original function's spatial properties, which were first recorded as line integrals of the function along various angles and distances.



(a) 9 directions (b) 90 directions

Fig. 2.3.7: Inverse Radon transform

Properties of Radon transform

Although the properties of the Radon transform presented here are applicable for more dimensions, we limit ourselves to 2-D situations because it is the most relevant for the application. Proofs of these properties [30] follow directly from the definition of the Radon transform.

Let us denote with $\mathcal{R}(f(\rho,\theta))$ as Radon transform of a piecewise continuous function with bounded support $f(x,y)$ on the line $x\cos\theta + y\sin\theta = \rho$. The following properties are in place:

- **Linearity:**

  $\mathcal{R}(\alpha_1 f_1(x,y) + \alpha_2 f_2(x,y)) = \alpha_1\mathcal{R}(f_1(\rho,\theta)) + \alpha_2\mathcal{R}(f_2(\rho,\theta)),\ \ \alpha_1,\alpha_2 \in \mathbb{R}.$

- **Symmetry:**

  The parameter set of $\rho \in [0,\infty)$ and $\theta \in [0,\pi]$ denotes every element of Radon transform, since $\mathcal{R}(f(\rho,\theta)) = \mathcal{R}(f(-\rho,\theta \pm \pi))$.

- **Periodicity:**

  $\mathcal{R}(f(\rho,\theta)) = \mathcal{R}(f(\rho,\theta + 2k\pi)),\ \forall\, k \in \mathbb{Z}.$

- **Scaling by factor α:**

  $\mathcal{R}(f(\alpha x, \alpha y)) = \frac{1}{|\alpha|}\mathcal{R}(f(\alpha\rho,\theta)),\ \alpha \neq 0.$

- **Rotation by an angle θ_0:**

  If we write function $f(x, y)$ in polar coordinates $f(r, \Phi)$ and rotate it for an angle θ_0 we obtain following:

  $\mathcal{R}(f(r, \Phi + \theta_0)) = \mathcal{R}(f(\rho, \theta + \theta_0))$.

- **Shifting by a vector (x_0, y_0):**

  $\mathcal{R}(f(x - x_0, y - y_0)) = \mathcal{R}(f(\rho - x_0 \cos\theta - y_o \sin\theta, \theta))$.

By utilizing Radon's direct and inverse transforms, we can establish a mathematical link between the images and measurement data obtained from the detector. The Radon transform captures this connection.

2.4 Fourier transform

The fundamental ideas behind the Fourier transform can be traced back to Joseph Fourier (1768 - 1830), a French mathematician and physicist who introduced the concept in the early 19th century. Fourier's work was published in his book [33] in 1822, which is where most of his related work can be found.

The Fourier transform is a mathematical operation that converts a function of space (or time) into a function of frequency. It decomposes a complex signal into its constituent sine and cosine waves of various frequencies.

As we transition to the next chapter on the Fourier transform, it is noteworthy to recognize the interplay between these transforms in imaging and signal processing. After obtaining Radon projections, the Fourier transform can be applied to analyze the frequency components inherent in these projections. This combined approach enhances our ability to interpret and process the captured data, providing a comprehensive foundation for understanding and manipulating images in applications like medical imaging and beyond.

The sinusoid function can be written as

$$f(x) = A \sin(2\pi\omega x + \theta), \tag{2.3}$$

where A represents the amplitude, θ is the phase or the shift of the sinusoid. Sinusoid function is a periodic function with a period T and frequency $\omega = \frac{1}{T}$.

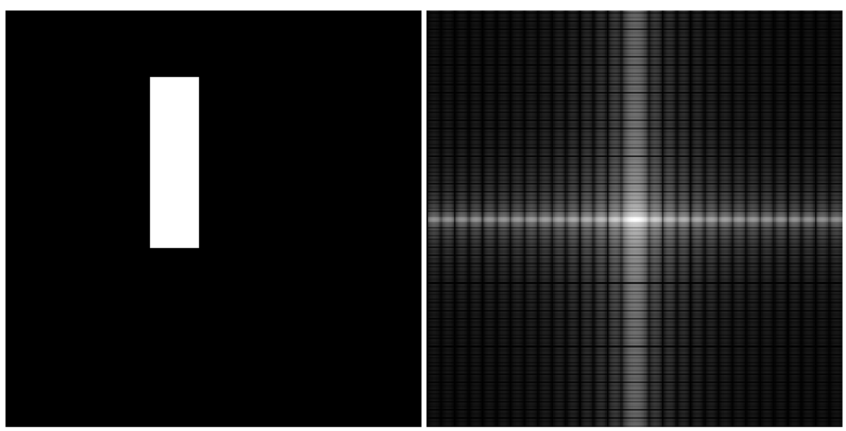
The $L^1(\mathbb{R}^n)$ is the vector space of equivalence classes of integrable functions on $\mathbb{R}^n$, where a function f is equivalent to a function g if $f = g$ almost everywhere. The L^1 norm of a function f is defined as $\|f\|_{L^1} = \int |f| dx$, where integral is a Lebesgue integral. A function f is in the L^1 space if this integral converges, meaning that the function is absolutely integrable. This makes $L^1(\mathbb{R}^n)$ into a normed vector space.

Definition 1 *For $f \in L^1(\mathbb{R})$ the Fourier transform $\mathcal{F}[f(x)] = F(\omega)$ and its inverse Fourier transform $\mathcal{F}^{-1}[F(\omega)] = f(x)$ are defined by*

$$\mathcal{F}[f(x)] = F(\omega) = \int_{-\infty}^{\infty} f(x) e^{-i2\pi\omega x} dx, \tag{2.4}$$

$$\mathcal{F}^{-1}[F(\omega)] = f(x) = \int_{-\infty}^{\infty} F(\omega) e^{i2\pi\omega x} d\omega, \ \sqrt{i} = -1. \tag{2.5}$$

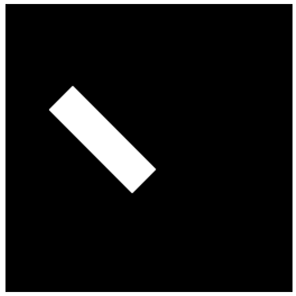
Observing the equation (2.5), we conclude that the original function (signal) is a sum of frequencies. $F(\omega)$ is complex and holds the amplitude and the phase of the sinusoid of the frequency ω.



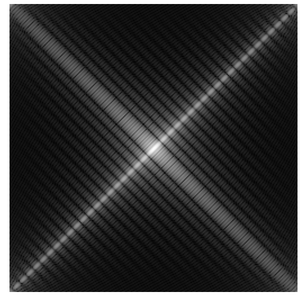
(a) Original image (b) Fourier transform

Fig. 2.4.8: Original image and its Fourier transform

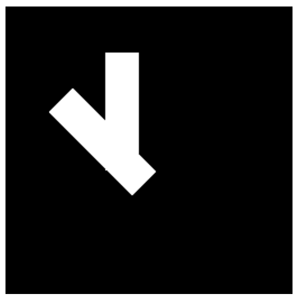
Examining the image 2.4.8, one can discern several distinct characteristics in the



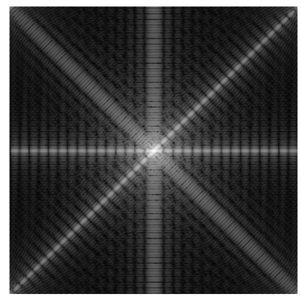
(a) Rotated original



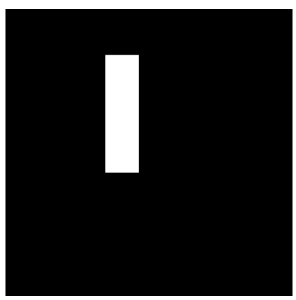
(b) FT of the rotated original



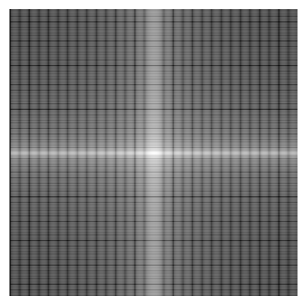
(c) Linear combination of 2 images



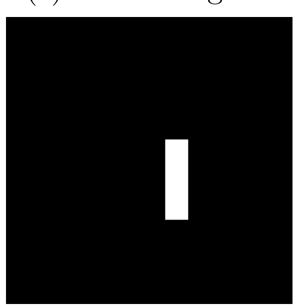
(d) FT Linear combination



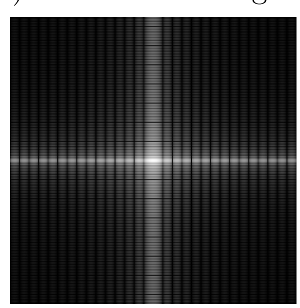
(e) Scaled original



(f) FT of scaled original



(g) Translated original



(h) FT of translated original

Fig. 2.4.9: Properties Fourier transform

resulting Fourier transform. The central portion of the Fourier transform image mirrors the presence of low-frequency elements within the original image, encompassing its smooth and gradually changing areas. In contrast, the outer regions of the Fourier transform image correspond to the higher-frequency components that capture the finer details, such as edges and textures, present in the original image. When the original image contains recurring patterns or lines, these patterns manifest as corresponding lines in the Fourier transform image, highlighting the presence of these repetitions. Additionally, bright spots in the Fourier transform image indicate strong frequency components in the corresponding positions of the original image, often indicating significant features or objects. It is important to note that the image's central region specifically represents the zero-frequency component.

Fourier transform $\mathcal{F}[f] = F(\omega)$ of function $f(x)$ holds following properties [70] illustrated in Figure 2.4.9:

- **Linearity:**

  $\mathcal{F}[\alpha_1 f_1(x) + \alpha_2 f_2(x)] = \alpha_1 \mathcal{F}[f_1(x)] + \alpha_2 \mathcal{F}[f_2(x)],\ \alpha_1, \alpha_2 \in \mathbb{R}.$

- **Scaling:**

  $\mathcal{F}[f(\alpha x)] = \frac{1}{|\alpha|} F(\frac{\omega}{\alpha})\,, \alpha \neq 0.$

- **Shifting:**

  $\mathcal{F}[f(x - \alpha)] = e^{-i2\omega\alpha} F(\omega).$

- **Differentiation:**

  $\mathcal{F}[\frac{d^n}{dx^n}(f(x))] = (i2\pi\omega)^n F(\omega).$

2-D Fourier transform $\mathcal{F}[f(x,y)] = F(u,v)$ of the function $f(x,y)$ we define as:

$$\mathcal{F}[f(x,y)] = F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) e^{-i2\pi(ux+vy)} dxdy. \tag{2.6}$$

The inverse Fourier transform is

$$\mathcal{F}^{-1}[F(u,v)] = f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v) e^{i2\pi(ux+vy)} dudv, \tag{2.7}$$

where u and v are frequencies along x and y respectively.

Convolution theorem

Definition 2 *For $f, g \in L^1(\mathbb{R})$ convolution of functions f and g is:*

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-y)g(y)dy,\ x \in \mathbb{R}.$$

Theorem 1 *If $f, g \in L^1(\mathbb{R})$ are functions in spacial domain and $\mathcal{F}[f] = F(\omega)$ and $\mathcal{F}[g] = G(\omega)$ are Fourier transforms of functions f and g, then:*

$$(f * g)(x) = \mathcal{F}^{-1}(F(\omega) \cdot G(\omega)).$$

The convolution theorem, in other words, states that performing convolution in the spatial domain is essentially the same as executing point-wise multiplication in the frequency domain. This theorem also holds in 2-D.

The convolution theorem [33] holds significant importance as it enables complex and costly linear filtering processes in the spatial domain to be replaced with straightforward and computationally efficient multiplications in the frequency domain.

Fourier slice theorem

The Fourier Slice Theorem [17] gives a connection between the Fourier transform of each of the projections and the original image.

Theorem 2 *Fourier Slice: The Fourier transform $P(\omega, \theta)$ of a projection $p(\rho, \theta)$ of an image $I(x, y)$ satisfies:*

$$P(\omega, \theta) = F(\omega \cos\theta, \omega \sin\theta), \quad \forall\theta \in [0, \pi), \forall\omega \in \mathbb{R},$$

where $F(\omega \cos\theta, \omega \sin\theta)$ represents the Fourier transform of the original image $I(x, y)$ along the line $u = \omega\cos\theta; v = \omega\sin\theta$ in the frequency domain (u, v).

The claim made here can be proved using well-known facts. The original image has a 2-D Fourier transform, while each projection has a 1-D Fourier transform.

Let us consider a projection under fixed angle $\theta_k = \theta$ and take a 1-D Fourier transform of that projection with respect to ρ:

$$P(\omega) = \int_{-\infty}^{\infty} p(\rho, \theta) e^{-i2\pi\omega\rho} d\rho. \tag{2.8}$$

By including equation (2.1) in the previous equation, we get:

$$\begin{aligned} P(\omega) &= \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y)\delta(x\cos\theta + y\sin\theta - \rho) dxdy \Big) e^{-i2\pi\omega\rho} d\rho \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y) \Big(\int_{-\infty}^{\infty} \delta(x\cos\theta + y\sin\theta - \rho) e^{-i2\pi\omega\rho} d\rho \Big) dxdy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y) e^{-2\pi i\omega(x\cos\theta + y\sin\theta)} dxdy. \end{aligned}$$

If in the equation above we say that $u = \omega\cos\theta$ and $v = \omega\sin\theta$ we get:

$$P(\omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y) e^{-2\pi i(ux+vy)} dxdy, \ \forall\theta \in [0, \pi). \tag{2.9}$$

On the right-hand side of the equation (2.9), it is depicted the 2-D Fourier transform of the initial image $I(x,y)$ along the trajectory defined by $u = \omega\cos\theta$ and $v = \omega\sin\theta$.

Equation (2.9) elucidates that the 1-D Fourier transform of an image projection at an angle θ is equivalent to a cross-section taken through the 2-D Fourier transform of the original image at the same angle, as visually represented in Figure 2.4.10.

2.5 Filtered back projection

To reconstruct an image, we need to get the sum of Fourier transforms of the image from as many as possible projections and then apply the inverse Fourier transform.

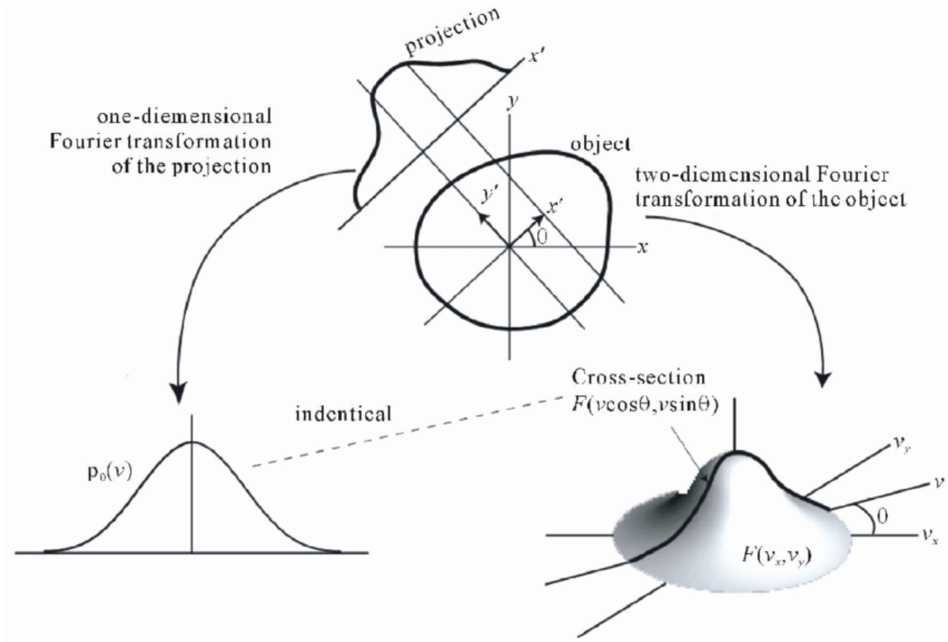


Fig. 2.4.10: Illustration of Fourier slice theorem (image is taken from [72])

While summing up the Fourier transforms, the middle part of the object is overly representative, which causes a blurry image (as shown on Figure 2.5.11). To avoid this, we can use a filtered back projection.

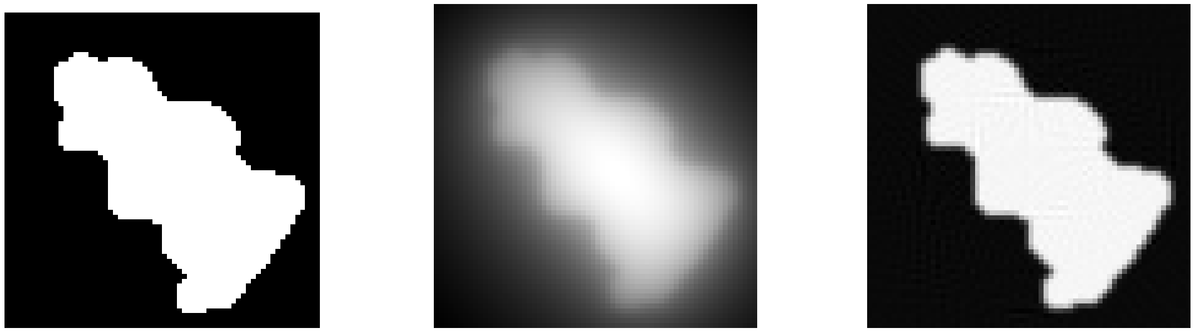
Frequency domain filters are tools used to refine images by manipulating their high- and low-frequency components, leading to effects like smoothing and sharpening. They differ from spatial domain filters because they primarily deal with the frequency characteristics of images. Essentially, these filters focus on modifying the frequency content to achieve two primary outcomes: smoothing and sharpening.

The original image can be represented using an inverse Fourier transform.

$$I(x,y) = \mathcal{F}^{-1}[F(u,v)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)e^{i2\pi(ux+vy)}dudv. \tag{2.10}$$

If we use the polar coordinates in equation (2.10) $u = \omega\cos\theta$ and $v = \omega\sin\theta$, $dudv = \omega d\omega d\theta$, we get:

$$I(x,y) = \int_{0}^{2\pi}\int_{-\infty}^{\infty} F(\omega\cos\theta,\omega\sin\theta)e^{i2\pi((\omega\cos\theta)x+(\omega\sin\theta)y)}\omega d\omega d\theta. \tag{2.11}$$



Original image Non-Filtered back projection Filtered back projection

Fig. 2.5.11: Comparison of simple and filtered back projection image reconstruction

By the Fourier slice theorem we have $P(\omega, \theta) = F(\omega \cos\theta, \omega \sin\theta)$, this yields:

$$I(x, y) = \int_0^{2\pi} \int_{-\infty}^{\infty} P(\omega, \theta) e^{i2\pi(x\cos\theta + y\sin\theta y)\omega} \omega d\omega d\theta.$$

If we apply shifting property that $P(\omega, \theta + \pi) = P(-\omega, \theta)$, we get following:

$$\begin{aligned} I(x, y) &= \int_0^{\pi} \int_{-\infty}^{\infty} |\omega| P(\omega, \theta) e^{i2\pi(x\cos\theta + y\sin\theta y)\omega} d\omega d\theta \\ &= \int_0^{\pi} \left[\int_{-\infty}^{\infty} |\omega| P(\omega, \theta) e^{i2\pi\omega\rho} d\omega \Bigg|_{\rho = x\cos\theta + y\sin\theta} \right] d\theta. \end{aligned} \tag{2.12}$$

Within the given equation (2.12) the inner part represents the inverse Fourier transform of the projection, which is then multiplied by a filter function denoted as $|\omega|$. In this way, we obtain filtered back projections [77].

The application of filters in the Fourier domain generally involves these steps:

- Compute the Fourier transform of the input image, converting it into the frequency domain.
- Multiply the transformed image by the frequency response of the chosen filter.
- Perform an inverse Fourier transform to convert the modified frequency-domain image back into the spatial domain.

Filtered back projection reconstruction

Filtered back projection (FBP) is a widely used reconstruction algorithm in computed tomography (CT) imaging, converting raw projection data into a 2-D or 3-D image. The choice of filter in FBP is crucial, influencing the quality and characteristics of the reconstructed image.

Common filters in FBP

1. **Ram-Lak (Ramp) Filter [97]:**
    - **Frequency Response:** Constant amplitude and linearly increasing frequency.
    - **Characteristics:** Simple and commonly used. Emphasizes high frequencies, contributing to good spatial resolution but may introduce more noise.
2. **Shepp-Logan Filter [87]:**
    - **Frequency Response:** Similar to Ram-Lak but with a more complex shape.
    - **Characteristics:** Designed to reduce artifacts and noise, providing improved image quality compared to Ram-Lak.
3. **Butterworth Filter [19]:**
    - **Frequency Response:** Adjustable based on filter order and cutoff frequency.
    - **Characteristics:** Offers flexibility; lower orders provide smoother images, while higher orders can enhance edges but may amplify noise.
4. **Hann (Hanning) Filter [65]:**
    - **Frequency Response:** Bell-shaped curve with smoother roll-off.
    - **Characteristics:** Provides smoother images by suppressing higher frequencies, balancing noise reduction with spatial resolution.

5. **Hamming Filter [40]:**

    - **Frequency Response:** Similar to Hann but with a wider main lobe.
    - **Characteristics:** Offers smoother images, reducing high-frequency noise at the expense of some spatial resolution.

Considerations in filter selection

- **Spatial Resolution vs. Noise Trade-off:** Filters like Ram-Lak emphasize spatial resolution but may introduce more noise. Filters like Hanning and Hamming prioritize noise reduction but may sacrifice spatial resolution.
- **Clinical Application:** Filters may be chosen based on specific clinical needs, such as emphasizing fine structures in vascular imaging.
- **Artifact Reduction:** Filters like Shepp-Logan and Butterworth are designed to reduce artifacts, contributing to overall image quality.
- **User Preferences:** The choice of filter may be influenced by user preferences and institutional protocols.

High pass filter

High-pass filters in the Fourier domain enhance high-frequency components, emphasizing edges and fine details. Types include:

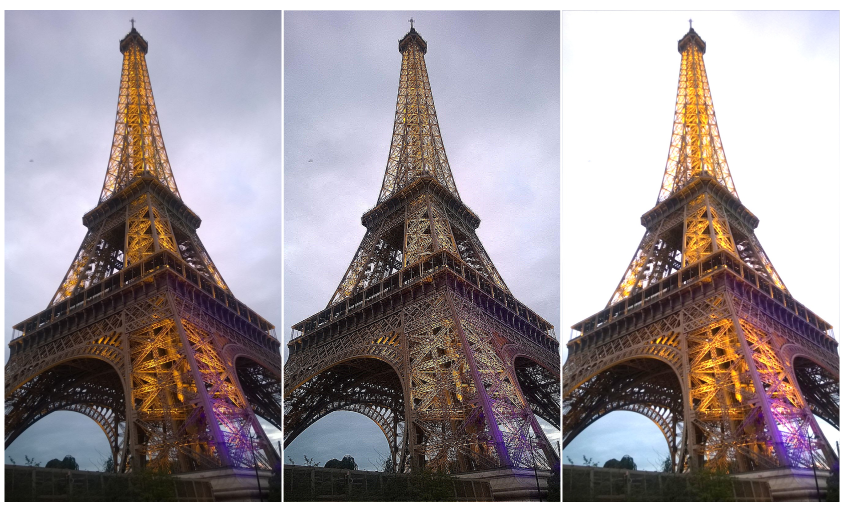
- **Ideal High-Pass Filter:** Abrupt cutoff, introducing spatial artifacts.
- **Butterworth High-Pass Filter:** Smoother transition controlled by order, mitigating artifacts.
- **Gaussian High-Pass Filter:** Uses Gaussian function, with standard deviation controlling the transition smoothness.

Low pass filter

Low-pass filters in the Fourier domain emphasize low-frequency components, leading to smoother variations in color and intensity. Types include:

- **Ideal Low-Pass Filter:** Abrupt cutoff with potential ringing artifacts.
- **Butterworth Low-Pass Filter:** Gradual transition controlled by order, reducing ringing artifacts.
- **Gaussian Low-Pass Filter:** Uses Gaussian function, with standard deviation controlling the transition sharpness.

While high-pass and low-pass filters (Figure 2.5.12) excel in enhancing specific image features, their application requires careful consideration of potential noise amplification and artifacts. Additionally, these filters can be used in conjunction with FBP to optimize the trade-off between spatial resolution and noise suppression based on the imaging requirements.



Original image High-pass filter Low-pass filter

Fig. 2.5.12: Comparison of high and low pass filters

CHAPTER 3

Discrete Tomography

This chapter delves into discrete tomography (DT) [41, 42], a specialized field dedicated to reconstructing discrete objects from a limited set of projections. In contrast to continuous tomography, which deals with images of continuous intensity values, discrete tomography focuses specifically on binary or discrete values, introducing unique challenges in terms of reconstruction algorithms, model formulations, and solution uniqueness. Coined by Larry Shepp in 1994, discrete tomography has garnered attention for its diverse applications, ranging from medical imaging to material science and security.

DT is positioned as a specialized form of CT, given that discrete functions can be considered a subset of general functions. While it logically extends findings from CT to discrete functions, DT requires its own set of principles to address inquiries about coherence, existence, and uniqueness. An additional incentive to explore dedicated discrete reconstruction methods is the hope that, due to the discrete nature of the unknown image, determining it might require less data compared to the requirements for general functions. As a result, DT commonly employs a modest number of projections.

Discrete tomography finds applications in various fields due to its ability to reconstruct objects or images with limited information.

In the field of medical imaging, discrete tomography can be utilized to reconstruct images of various body structures, including bones, teeth, and other organs and tissues. For instance, in dental computed tomography (CT), discrete tomography is used to reconstruct images of teeth and jaw bones. This can help in the diagnosis and treatment planning of disorders that are related to the teeth and jaws. Discrete tomography can also be used to reconstruct images of the inner ear, which is helpful in the diagnosis and treatment of hearing impairments.

In materials science, discrete tomography can be used to reconstruct images of materials at the micro and nanoscale. For example, in electron tomography, discrete tomography is used to reconstruct images of materials at the nanoscale, which can aid in understanding their mechanical, electrical, and optical properties.

Pipes, metal castings, and printed circuit boards are just some of the industrial structures that might benefit from discrete tomography examination. For instance, discrete tomography can be used in non-destructive testing to reconstruct images of the interior structure of pipes and metal castings, which can then be used to help spot flaws like fractures and voids.

Discrete tomography is also used in a wide range of other fields, such as archaeology, computer vision, and cryptography.

The significance of DT lies in bridging the gap between continuous tomography and the discrete world, allowing for the reconstruction of objects with discrete intensity values prevalent in real-world applications. By tailoring algorithms and models for discrete tomography, unique challenges associated with discrete objects or images can be addressed, leading to improved reconstruction accuracy, computational efficiency, and a more profound understanding of discrete structures.

In the subsequent sections of this chapter, we will explore the formulation of discrete tomography models, including binary tomography models, grayscale tomography models, hybrid models, and various reconstruction algorithms employed in discrete tomography.

3.1 Discrete tomography formulation

In discrete tomography, a reconstruction problem involves the challenge of recreating the inner structure of an object based on a collection of projection data. Typically, this projection data is acquired by illuminating the object with X-rays, electrons, or other forms of radiation through the object and measuring the intensity of the radiation that emerges on the other side. The goal of the reconstruction problem is to use this projection data to construct an image of the internal structure of the object.

In discrete tomography, the object being imaged can only take on a discrete set of values, such as black and white or multiple grayscale levels. Therefore, the reconstruction problem is finding the object's specific discrete values that are most consistent with the projection data.

Discrete tomography requires a suitable representation of objects or images in the discrete domain. Depending on the application and desired level of detail, different types of discrete representations can be employed, such as pixel-based representations or higher-level discrete structures.

Pixel-based Representations: Pixel-based representations are commonly used in discrete tomography, where the objects or images are discretized into a grid of pixels. Each pixel can have a binary value (0 or 1) or represent different intensity levels in grayscale tomography. The arrangement and relationships between the pixels play a crucial role in the reconstruction process.

Higher-Level Discrete Structures: In certain cases, the discrete representation may involve higher-level structures beyond individual pixels. For example, the objects or images may be represented using line segments, rectangles, squares, or other combinatorial structures, which in this context are seen as arrangements of discrete elements that follow certain rules or constraints. These higher-level structures introduce additional constraints and considerations in the reconstruction models.

Discretization of the Tomographic Projection Process: The tomographic projection process, which captures the interactions between the objects or images and the projection angles, also needs to be discretized to fit the discrete tomography framework. This involves defining discrete projection angles and discretizing the measurement process.

Discrete Projection Angles: In discrete tomography, projection angles are discretized to a finite set of discrete values. The choice of discrete projection angles depends on factors such as the desired resolution, the number of projections available, and the specific reconstruction algorithm employed. Common discretization schemes include equidistant angles, evenly spaced angles, or predefined sets of angles based on specific requirements.

Discretization of Measurement Process: The measurement process in tomography involves computing the projection values or measurements by quantifying the interactions between the objects or images and the projection angles. In discrete tomography, this measurement process must be adapted to the discrete nature of the objects or images. It may involve quantizing the continuous measurements or applying specific discretization techniques tailored to the discrete representation used.

By properly formulating the discrete representation of objects/images and discretizing the tomographic projection process, discrete tomography models can capture the essential characteristics and constraints needed for accurate reconstruction.

3.2 Reconstruction problem

Reconstruction problems in discrete tomography are typically formulated as optimization problems, in which an objective function is defined to measure the consistency between the object and the projection data. The objective function is then minimized with respect to the object using methods such as gradient descent or linear programming.

The standard method to explain the process of projection data collection in discrete tomography is by using the concept of line projections. To illustrate, consider a 2-D grid or lattice that represents the object to be reconstructed. Each cell in this grid can have a value from some discrete set. The process of collecting projection data simulates what would happen if lines were projected through the object from different angles and where these lines intersect with the object.

The object's representation is established using a 2-D grid, where each cell's status (empty or filled) captures the discrete nature of the object. For each projection angle, a projection line is conceptualized through the grid. These lines can be described

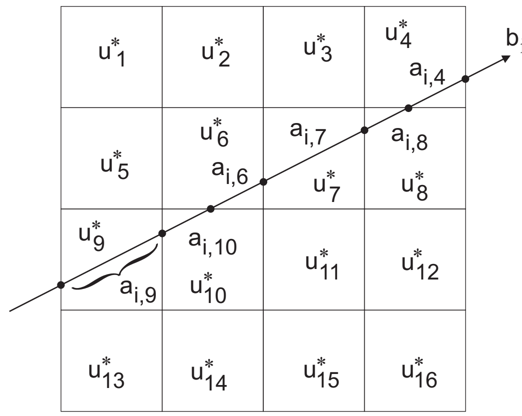


Fig. 3.2.13: Example of a projection value calculation on an image

parametrically, using equations like the slope-intercept form or by specifying an angle and distance from a reference point. As projection data is collected, information about the intersections between projection lines and cells is gathered. The collected projection data from different angles can be organized into a matrix. Each row in the matrix corresponds to a projection angle, and each column corresponds to a particular count of intersected filled cells. The primary objective of discrete tomography is to reconstruct the original object grid by utilizing the projection data matrix gathered earlier. This process involves solving an inverse problem: finding a configuration of the grid that would lead to the observed projection data.

Figure 3.2.13 shows an example of a projection value calculation on an image u^* of size $N = 4 \times 4 = 16$. A projection ray penetrates through the image pixels. The projection value b_i is calculated by $b_i = a_{i,4}u_4^* + a_{i,6}u_6^* + a_{i,7}u_7^* + a_{i,8}u_8^* + a_{i,9}u_9^* + a_{i,10}u_{10}^*$.

The data collected in continuous tomography is typically in the form of continuous projections, which represent the integral of the object's properties along the path of the imaging beam. Data collected in discrete tomography consists of projections that involve counting the number of discrete elements (e.g., pixels) along certain directions. Discrete tomography addresses the challenges associated with the discrete nature of object representation, resulting in specialized computational methods designed specifically for reconstructions involving discrete or binary grids. One of the challenges in reconstruction problems in tomography is dealing with noise and uncertainty in the projection data. The projection data may be affected by noise, such as electronic noise or scatter, which can make it difficult to reconstruct the object accurately. Additionally, the projection data may be uncertain, meaning that the measurements are not known exactly. Therefore, reconstruction methods often include regularization terms that help to stabilize the solution and reduce the effects of noise and uncertainty.

Another challenge in discrete tomography is the combinatorial nature of the problem. The set of possible solutions is combinatorially large, meaning that there are a very large number of possible objects that could have produced a given set of measurements. Therefore, reconstruction methods often rely on some form of prior knowledge about the object, such as smoothness [44], sparsity [83] or non-negativity constraints [22] to make the problem more tractable.

The problem of DT reconstruction can be explained through the linear set of equations provided below.

$$
\begin{aligned}
&a_{11}u_1 + a_{12}u_2 + a_{13}u_3 + \ldots + a_{1N}u_N = b_1 \\
&a_{21}u_1 + a_{22}u_2 + a_{23}u_3 + \ldots + a_{2N}u_N = b_2 \\
&a_{31}u_1 + a_{32}u_2 + a_{33}u_3 + \ldots + a_{3N}u_N = b_3 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&a_{M1}u_1 + a_{M2}u_2 + a_{M3}u_3 + \ldots + a_{MN}u_N = b_M,
\end{aligned}
$$

which we examine in its matrix form:

$$A\,u = b, \tag{3.1}$$

where $A \in \mathbb{R}^{M\times N},\ \ u \in \Lambda^N,\ b \in \mathbb{R}^M$ and $\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_k\}$ for $k \geq 2$.

The task is reconstructing an image represented by the unknown column vector u. The range of possible values for the image, represented by set Λ, is defined by the user and can be binary or multi-colored. The projection data is captured in the projection matrix A, where each row corresponds to the intersection length between pixels and the projection rays traversing them. These matrix elements are determined based on the extent of these intersections. The projection vector b is computed as the summation of the products of pixel intensities and the lengths of the projection rays passing through them.

In the process of projection, various directions are employed, using a parallel beam projection method where multiple parallel projection rays are taken for each direction. The angle α plays a key role in determining the direction of the projection. To ensure that the entire image grid is covered, the spacing between adjacent parallel projection beams is set equal to the size of a pixel, with even distribution. The number of parallel projection rays is carefully chosen to guarantee comprehensive coverage of the entire image grid.

Now, the reconstruction task at hand involves finding the image solution, denoted as u, which is represented by a linear system of equations (3.1). This system is often underdetermined, meaning there are more unknowns (N) than equations (M). The objective goes beyond merely finding a solution that aligns with the provided projections; it also aims to create an image that closely resembles the original one. To achieve a high-quality and satisfactory solution, it is crucial to make use of all available knowledge, including any prior information, about the object being reconstructed.

3.3 Binary image representation and the 0-1 intensity assumption

Binary tomography involves the reconstruction of binary images, where each pixel or element in the image is represented by a binary value, typically 0 or 1. This binary representation simplifies the reconstruction problem by reducing the complexity of intensity values. The *0-1 intensity assumption* assumes that the objects of interest are perfectly opaque (1) or completely transparent (0), allowing for a straightforward binary representation.

The binary image is represented as a matrix, where each entry corresponds to a pixel in the image. A value of 1 indicates that the pixel belongs to the object or region of interest, while a value of 0 indicates that the pixel is part of the background or is transparent. The matrix can be denoted as M, where $M_{i,j}$ represents the value of the pixel at row i and column j.

The binary representation simplifies the reconstruction problem by reducing the number of possible intensity values, making it more amenable to discrete tomography techniques. Additionally, the binary nature of the representation allows for the utilization of various combinatorial structures and algorithms tailored to binary objects.

By assuming the 0-1 intensity values, binary tomography focuses on the presence or absence of objects rather than their varying intensity levels. This simplification is particularly suitable for scenarios where the primary interest lies in detecting the presence or absence of specific objects or features.

A potential use of the 0-1 intensity assumption can be found in human X-ray angiography. This involves creating images of blood vessels and heart chambers using X-ray tomography. By introducing a high-contrast agent into the body area of interest, the problem can be addressed through BT. This entails detecting the contrast agent's presence in specific positions [21, 74].

The main challenges in binary tomography arise from the fact that the reconstructed image is binary, which introduces non-linearity and combinatorial complexity into the reconstruction problem. The goal is to find the binary image that best matches the given projections, subject to certain constraints and regularization techniques.

Binary tomography reconstruction methods can be classified into four main classes: algebraic methods, stochastic sampling methods, heuristic combinatorial, and relaxation methods approaches. Each of these methods has its own strengths and weaknesses in solving the binary tomography problem [47].

Algebraic Methods: Algebraic approaches leverage the inherent algebraic properties of binary tomography problems to offer valuable insights into solution uniqueness and the necessary number of projections. These methods often rely on mathematical equations and principles to reconstruct binary images from limited projection data.

Although these approaches are theoretically elegant, they can encounter difficulties when extended to more practical projection models and when confronted with noisy data. They are commonly used for theoretical analysis and small-scale problems where the algebraic structure can be adequately captured.

Heuristic Combinatorial Approaches: Heuristic combinatorial approaches combine ideas from combinatorial optimization and iterative methods. These methods offer practical efficiency and are known to perform well in practice. They often rely on heuristics, which are rules of thumb or approximations, to guide the reconstruction process. Some of the heuristic approaches are simulated annealing [1], tabu search [35], and evolutionary algorithms [73]. While these algorithms demonstrate effectiveness in handling diverse binary tomography scenarios, it is essential to recognize that they do not guarantee optimal solutions. The quality of the reconstruction heavily relies on the specific heuristics employed, necessitating careful consideration and tuning of hyperparameters. Achieving optimal performance often involves balancing trade-offs, and the effectiveness of these approaches is closely tied to the selection and fine-tuning of heuristic parameters.

Stochastic Sampling Methods: Stochastic sampling methods take a probabilistic approach to binary tomography reconstruction. They construct probability density functions on the space of discrete images, enabling the sampling of potential solutions. Markov Chain Monte Carlo (MCMC) techniques are often employed to explore the solution space and find suitable reconstructions [69, 34, 7, 71, 29]. These methods offer a high degree of flexibility, making them well-suited for addressing intricate reconstruction scenarios. However, it's worth noting that the computational cost can be significant, especially for large-scale datasets. Like heuristic combinatorial approaches, the successful implementation of stochastic sampling methods demands careful consideration and tuning of model internal parameters to strike a balance between computational efficiency and reconstruction accuracy. The effectiveness of stochastic sampling methods lies in their ability to provide probabilistic reconstructions, accommodating uncertainty in the reconstruction process. Despite their computational demands, stochastic sampling methods remain a valuable tool in binary tomography reconstruction, offering a probabilistic framework to capture the inherent uncertainty in the imaging process.

Relaxation Methods: Relaxation methods involve relaxing the constraints of the binary tomography problem to enable a more tractable solution. These methods often use convex or non-convex relaxation techniques, leading to natural extensions of variational formulations and iterative algorithms. They are known for their efficiency in solving large-scale binary tomography problems. However, ensuring convergence to the correct binary solution can be challenging, and some relaxation methods may not perform well when the data is noisy or when dealing with complex structures. Linear-programming based methods, a variant of relaxation techniques, can work well on small-scale images and noise-free data [94].

In the field of binary tomography, researchers continuously explore and develop new algorithms and variants that suit different problem settings.

3.4 Ryser algorithm

In this subsection, we present a brief summary of Ryser's [54, 82] theoretical solutions that answer the existence question in the binary tomography problem.

The reconstruction of a binary matrix from its row and column sum vectors is studied in circumstances when some elements of the matrix may be specified and the matrix can be identified from these values. This allows for the reconstruction of binary pictures using just two projection angles.

Let us consider two non negative vectors $R = (r_1, r_2, ..., r_m)$ and $S = (s_1, s_2, ..., s_m)$. The tomographic equivalent class of all binary matrices $A = (a_{ij}) \in \{0, 1\}^{m \times n}$ for which stands:

$$r_i = \sum_{j=1}^{n} a_{i,j}, \;\; s_j = \sum_{i=1}^{m} a_{i,j} \tag{3.2}$$

is denoted by $\mathcal{U}(R, S)$.

Ryser has demonstrated the condition for the existence of such matrices, that is, whether $\mathcal{U}(R, S)$ is empty or not.

Consider the matrix C, where the $i-th$ row is composed of r_i ones followed by $n - r_i$ zeros. Such a matrix is referred to as maximal. It is clear that row sums determine a uniquely maximal matrix. Let vector $\bar{S}$ denote the column sum of C.

Theorem 3 *(Ryser, 1957) Let S and R be a pair of two non negative vectors. The class $\mathcal{U}(R, S)$ is nonempty if and only if*

$$\sum_{j=l}^{n} s'_j \geq \sum_{j=l}^{n} \bar{s}_j, \quad 2 \leq l \leq n, \tag{3.3}$$

where $\bar{s}_j \in \bar{S}$, $s'_j \in S'$, and vector S' is the non increasing permutation of the elements of vector S.

A pair (R, S) is considered compatible if it is feasible to create a binary matrix that adheres to the given row sums R and column sums S.

Assume that the set $\mathcal{U}(R, S)$ contains a binary matrix A. Consequently, the class $\mathcal{U}(R, S')$ accommodates a binary matrix A', derived from A through a suitable permutation of its columns. If there exists any disparity between $\bar{A}$ and A', the former can be derived from the latter by left-shifting the 1's in the rows of A'. This process aligns with the relationship described in equation 3.3. Assuming equation 3.3 holds for vectors R and S, our goal is to generate a binary matrix A using the Ryser algorithm described in Algorithm 3.4.1. The proof that Algorithm 3.4.1 produces a matrix A with a row sum vector R and a column sum vector S can be found in [54].

Ryser demonstrated that if two matrices consisting of 0's and 1's have identical sums for both their rows and columns, then it is possible to transform the first matrix into the second using a series of simple operations. Each operation involves changing 1's to 0's and 0's to 1's, while keeping the sum of each row and column the same. This finding is akin to a principle in binary tomography, where matrices composed of 0's and 1's can be seen as binary images. When two matrices with matching row and column sums are considered, they correspond to two binary images that share the same horizontal and vertical projections. Such images are referred to as being

Algorithm 3.4.1 Ryser Algorithm

input : A compatible pair of vectors (R, S) satisfying 3.3;
output: A binary matrix A;
begin
  Step 1. Construct S' from S by permutation π;
  Step 2. Let $B = \bar{A}$ and $k = n$;
  Step 3.
  while $k > 1$ **do**
    while $s'_k > \sum_{i=1}^{m} b_{ik}$ **do**
      let $j_0 = \max_{i=1}^{m} j < k | b_{ij} = 1, b_{i,j+1} = \cdots = bik = 0$;
      let row i_0 be where such a j_0 was found;
      set $b_{i_0 j_0} = 0$ and $b_{i_0 k} = 0$ (i.e., shift the 1 to the right);
    reduce k by 1;
  Step 4. Construct the matrix A from B by permutation π^{-1} of the columns;
  return Matrix A

"tomographically equivalent."

To illustrate the Ryser algorithm, let us consider a simple example of a 5×6 binary image that should be reconstructed from its projections. This binary image can be represented as a matrix $A = (a_{i,j})_{5\times 6}$, $a_{i,j} \in \{0, 1\}$, where 0 represents an empty pixel, and 1 represents an object pixel.

Let us assume we have two projection directions of an unknown image, represented by matrix A, one horizontal and one vertical. Horizontal projection: $R = [2, 4, 3, 4, 1]$ (sum of the pixels in each row), vertical projection: $S = [3, 4, 3, 2, 1, 1]$ (sum of the pixels in each column). In order to generate matrix A, we firstly construct a maximal matrix based on the row sums and we calculate column sums of that matrix, $\bar{S}$ = [5,4,3,2,0,0], then we construct a non-increasing permutation of vector S, S'=[4,3,3,2,1,1], vectors $\bar{S}$ and S' satisfy (3) which means that there is a class of binary matrices whose row sum is vector R and column sum is vector S. Thus, it is possible to reconstruct a binary image from these projection data. In the following lines, we describe finding one solution for this problem using the Ryser algorithm.

1. **Step 1:** Arrange the column sums S in a non-increasing order to create a new sequence, denoted as S'.

2. **Step 2:** Create a maximum matrix $\bar{A}$ defined by vector R.

$$\bar{A} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

3. **Step 3:** Move elements from the rightmost columns of $\bar{A}$ to the columns where the sum of elements in the corresponding column is smaller than the value in S'. This step adjusts the matrix while maintaining the specified column sums.

$$\Rightarrow \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

4. **Step 4:** Reverse the permutation used in step 1. This step ensures that the resulting matrix matches the original order of the column sums and generates the following matrix as one solution to the given problem:

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The Ryser Algorithm provides a method to reconstruct the binary matrix A based on the given row sums R and column sums S and its time complexity is $O(n(m + \log n))$.

3.5 Discrete algebraic reconstruction technique

The Discrete Algebraic Reconstruction Technique (DART) is one of the widely used iterative reconstruction algorithms in discrete tomography. It is based on Algebraic Reconstruction Methods (ARM) and was introduced by Batenburg and Sijbers in [5].

Algebraic Reconstruction Methods (ARM) are iterative algorithms used in computed tomography for image reconstruction from angular projections. Unlike analytical methods such as filtered back-projection, ARM approaches the problem iteratively. It updates the image iteratively to converge to the best approximation of the true image. ARM aims to minimize differences between measured and estimated projections by solving a system of linear equations. Each equation corresponds to projection data and its corresponding ray path through the image. ARM gradually improves image reconstruction through iterative solving. However, ARM faces challenges like computational complexity and sensitivity to noise, as mentioned in [5].

3.5.1 Algebraic Reconstruction Technique (ART)

The Algebraic Reconstruction Technique (ART) models the reconstruction problem as a system of linear equations solved iteratively. Pixels' values are variables in vector u, and the imaging process is described by matrix A. The angular projections are in vector b. For a matrix A of dimensions $m \times n$ and vector b, ART iteratively approximates the solution as follows:

$$u^{k+1} = u^k + \lambda_k \frac{b_i - \langle a_i, x^k \rangle}{\|a_i\|^2} a_i, \tag{3.4}$$

where $i = k \mod m + 1$, a_i is the i-th row of matrix A, b_i is the i-th component of vector b, and λ_k is a relaxation parameter. ART handles unconventional scanning geometry (non-uniform angular sampling and limited-angle tomography) but requires more computational resources.

3.5.2 Simultaneous Iterative Reconstruction Technique (SIRT)

The Simultaneous Iterative Reconstruction Technique (SIRT) is a variant of ART that provides slightly improved images at a slower pace [3]. In SIRT, matrix A represents the scanner's action, and matrix A^T back-projects the projection images onto the reconstruction region. SIRT alternates forward and back projections using the update equation:

$$u^{k+1} = u^k + CA^T R(b - Au^k), \tag{3.5}$$

where C and R are diagonal matrices containing inverses of column and row sums of the system matrix, respectively. These matrices adjust for ray-pixel interactions.

The Simultaneous Algebraic Reconstruction Technique (SART) [2] combines ART and SIRT. It updates u_k per projection angle, striking a balance between ART and SIRT. SART computes updates for all rays per projection angle in blocks, making it a block iterative method.

3.5.3 DART algorithm

The Discrete Algebraic Reconstruction Technique (DART) algorithm starts by generating a continuous reconstruction through a fixed number of SIRT iterations. This initial reconstruction serves as a starting point for the DART process. The reconstructed image is then subjected to a segmentation step, which aims to obtain an image containing only permissible grey values. This segmentation is achieved by rounding the pixel values to the nearest admissible grey values, thereby simplifying the representation of the image.

Subsequently, the segmented image is partitioned into two distinct groups of pixels: free pixels and fixed pixels. Free pixels are those that are adjacent to at least one pixel with a different grey value, effectively delineating the edges of the object within the image.

The core of the DART algorithm involves performing a predetermined number of SIRT iterations specifically on the free pixels, while keeping the fixed pixels at their respective grey values. By maintaining certain pixels at fixed values, the problem is simplified, resulting in a system with fewer variables but the same number of equations. However, the presence of noise in the data, combined with the fixed pixel values, can lead to fluctuations in the values of the free pixels after each round of SIRT iterations.

To mitigate the influence of noise and stabilize the reconstruction, a smoothing operation is applied to the free pixels. This operation involves using a Gaussian smoothing filter with a radius of 1. The filter works to even out the values of neighboring pixels and reduce the impact of noise, contributing to a smoother and more accurate reconstruction.

The DART algorithm can be terminated based on specific criteria. It may halt when the total projection error falls below a predefined threshold or after a certain fixed number of iterations have been executed. Once the algorithm concludes, the final reconstruction is obtained by performing segmentation on the resulting image. This final reconstruction contains only pixels with permissible grey values, providing an accurate representation of the object under study.

DART is particularly well-suited for scenarios where objects consist of distinct compositions, each corresponding to a constant grey value in the reconstructed image. It employs a fixed threshold function for discretization without regularization, which can sometimes result in more radical solutions, especially when dealing with limited projection data. Despite this, DART finds extensive use across various applications [57, 59, 93].

3.6 Energy minimization in image reconstruction

Energy-minimization methods are powerful techniques used in image processing to solve various problems, such as image denoising, image segmentation, image inpainting, and image restoration. These methods aim to find the optimal configuration of an energy function, which represents the cost or discrepancy between the processed image and the desired result.

The fundamental concept involves formulating the task as a minimization model, wherein the objective is to find the value of a function u that minimizes the functional $E(u)$. Typically, u represents an image in this context. The term "energy" is drawn from a physical analogy, where a stable system is characterized by having minimal total energy.

In any energy-minimization approach, two crucial criteria must be satisfied. Firstly, the design of the energy function, or model, must closely mimic the real problem being addressed, and its minimum point, preferably global, should represent the optimal solution to the problem. Secondly, the optimization algorithm used for energy minimization should be both fast and accurate, allowing for a good approximation of the minimum value while making efficient use of available computational resources. Failure to meet either of these criteria could significantly reduce the method's effectiveness or render it entirely unsuitable for practical applications.

In image processing, an energy function (also known as an objective function or cost function) quantifies the quality or appropriateness of a given image. The energy function typically consists of two main components: the data fidelity term and the regularization term.

When applying energy minimization in image reconstruction, in its most generic context, one attempts to recover a reconstructed form of the observed image u by minimizing the following energy function:

$$E(u) = F(Lu, b) + \lambda R(u). \tag{3.6}$$

An argument u^r that minimizes this energy function,

$$u^r = arg \min_u E(u) \tag{3.7}$$

is regarded as an estimate of the original image.

The function F measures how dissimilar the reconstructed image u is from the original data b after applying the operator L (where L typically represents a linear transformation or measurement operator). It is generally referred to as the "data fidelity term," and it essentially assesses how well the processed image matches the observed data. In the context of image denoising, the data term penalizes differences between the noisy and denoised images.

On the other hand, the "regularization term," denoted as R, introduces prior knowledge or expectations about the solution u. Essentially, it is a tool to encourage smoother and more desirable characteristics in the processed image while discouraging undesired features. Lower values of the regularization term, R, are expected to contribute to the removal of unwanted details. Additionally, regularization plays a crucial role in enhancing the numerical stability of the image reconstruction process. The regularization parameter, λ, acts as a control knob that dictates the trade-off between the degree of smoothing applied to the image and the accuracy of recovering fine image details. The regularization term encourages certain properties or characteristics in the processed image. It serves as a smoothness constraint and discourages overly complex or noisy solutions. Regularization helps prevent overfitting and produces more visually appealing results.

Finding an analytical solution for the problem (3.6) is generally not feasible due to its large-scale nature. Consequently, an appropriate optimization approach is required for solving it.

The data fitting term in E can be formed as a convex quadratic function in the form of a sum of the squares of the errors, making its numerical treatment relatively straightforward. The conjugate gradient method proves to be one of the most efficient approaches for minimizing this term, providing a solution in at most N (size of u) iteration steps. While the data fitting term in the energy function E is often represented as a convex quadratic function, it's important to recognize that it can deviate from this form. Non-quadratic data fitting terms arise in complex systems or models, introducing challenges for numerical treatment. In such cases, traditional optimization methods designed for quadratic problems may be less effective [20].

However, the regularization term R can take a completely different form compared to the data fitting term. It may exhibit high non-linearity, non-convexity, and even non-differentiability at certain points. These characteristics make the minimization of E challenging. High non-linearity increases the computational complexity during numerical evaluations, while non-convexity can result in the energy function E being non-convex, making it difficult to determine the global minimum. Furthermore, the non-differentiability of R implies that E is also non-differentiable, rendering many minimization methods based on gradient or higher-order differential information unsuitable. This is a significant restriction, as many efficient deterministic methods rely on gradient calculations.

The analysis above demonstrates that the regularized image processing problems are not always well-posed. For instance, when the energy function contains a non-convex regularization term, it can lead to non-convexity in the problem, resulting in multiple local minima without a unique global minimum. In such cases, the main contribution of the regularization is to significantly restrict the originally vast set of solutions without necessarily leading to a unique solution.

The problem (3.6) represents an unconstrained optimization problem. However, certain applications, such as discrete tomography or defuzzification, restrict the search space to a discrete set. The constrained regularized problem is formulated as follows:

$$\min_{u \in \Omega} E_Q(u), \tag{3.8}$$

where Ω represents the feasible set. Addressing the constraint condition poses an additional challenge on top of the previously analyzed issues, which needS to be resolved. One possible approach is to transform the constrained problem into an unconstrained one by reformulating the constraint condition as a new regularization term. An example of such an approach is convex-concave regularization in discrete tomography [84]. Alternatively, another way to tackle this challenge is to directly apply an appropriate optimization method specifically designed for constrained problems.

3.6.1 Regularization terms

In image reconstruction, regularization is a technique used to constrain the solution space of an inverse problem in order to obtain a unique and stable solution. The inverse problem in image reconstruction refers to the task of estimating an unknown image from a given set of measurements or observations. The measurements or observations may be corrupted by noise, or they may be incomplete, which makes the inverse problem ill-posed.

Regularization can be defined as an additional term added to the objective function that is being minimized in order to find the solution to the inverse problem. The regularization term imposes certain properties on the solution, such as smoothness, sparsity, or piecewise constancy, which makes the solution more meaningful and less sensitive to noise or incompleteness of data.

The regularization term is often a function of the image itself, and it is chosen based on the characteristics of the image and the type of the inverse problem. The regularization term can be mathematically formulated as a penalty term, a constraint, or a prior probability distribution on the image. The goal of regularization is to balance the goodness of fit of the model to the data and the regularization term in order to obtain a stable and interpretable solution to the inverse problem.

There are several regularization terms that are commonly used in image reconstruction, each with its own advantages and disadvantages. Some of the most widely used regularization terms include:

L2-norm regularization: This term is also known as Tikhonov regularization, and it promotes smooth solutions by minimizing the L2-norm of the image. A popular L2-norm regularization method is the Tikhonov regularization algorithm, which was first introduced in [91]. Building upon the foundation laid by L2-norm regularization, Total Variation (TV) regularization emerged as another tool in image processing. This regularization technique is based on the idea that natural images often exhibit piecewise constant or piecewise smooth structures. By minimizing the total variation of an image, TV regularization promotes solutions that capture these characteristics. The versatility of TV regularization is evident in its applications, spanning image denoising, deblurring, and inpainting. A popular TV regularization method is the Rudin-Osher-Fatemi (ROF) model, introduced in [81].

L1-norm regularization: This term promotes sparse solutions by minimizing the L1-norm of the image, which is the sum of the absolute values of the image's pixels. This method is widely used in compressed sensing and sparse representation. A popular L1-norm regularization method is the basis pursuit (BP) algorithm, which was first introduced in the paper [24].

Non-local Means (NLM) regularization: This term is based on the idea that similar patches in an image should have similar intensities. NLM regularization promotes solutions that are consistent with this idea by using non-local means to estimate the image's pixels. This method is widely used in image denoising, and many papers have been published on this topic. A popular NLM regularized method is the non-local means denoising algorithm, which was first introduced in [18]. Deep learning based regularization: This term is based on the idea of using deep neural networks to extract features and to make predictions [51]. Many papers have been published on this topic, and it has been used in various image reconstruction tasks such as image inpainting, deblurring, and denoising.

Regularization using geometric moments: In certain scenarios, it is feasible to have prior knowledge about aspects such as the orientation of shapes, the location of shape centroids, or the circularity of shapes within discrete tomography problems. These descriptors of shape characteristics can be effectively quantified using geometric moments. Consequently, they can be seamlessly integrated into the energy function as a regularization component, as demonstrated in the works by Lukić and Balazs [58, 62, 63]. By incorporating geometric moments into the regularization term of the energy function, the reconstruction process benefits from improved shape fidelity and structural consistency, ultimately leading to more reliable results.

These are some of the most popular regularization terms used in image reconstruction. However, there are many other techniques, and many papers have been published on this topic [62, 64, 67].

3.6.2 Simulated annealing algorithm

One of the image reconstructing methods that is based on energy minimization is the Simulated Annealing (SA) algorithm. SA is a stochastic optimization technique. The fundamental idea of SA originated in 1953 [68], when scientists applied the concept of slow cooling of material in a heat bath to solve a physical problem. In a real annealing process, the observed system starts at a high temperature and high energy, gradually cooling down until it reaches approximate thermodynamic equilibrium, converging to a steady, frozen ground state. This idea was extended to optimization problems in 1982 [50] and introduced as the general SA optimization algorithm. SA has found applications in tomography reconstruction problems [60, 61, 93].

The SA algorithm relies solely on objective function values during the reconstruction process, which offers great flexibility in incorporating different types of regularization terms into the energy function. However, SA is non-deterministic, meaning different runs of the same problem may yield different solutions. Additionally, SA can have relatively high running times and require careful tuning of its parameters.

3.6.3 Gradient based reconstruction methods

Gradient based reconstruction methods in image processing are algorithms that use the gradient information of an image to reconstruct or restore the image. These methods are often used in image restoration and reconstruction tasks, such as deblurring, denoising, and inpainting, which aim to remove noise, blur, or missing information from an image.

One of the main advantages of gradient-based reconstruction methods is that they can effectively preserve fine details and edges in the image, which are often lost or distorted in other types of reconstruction methods. These methods also often have fast convergence rates and can be implemented relatively easily.

There are several different types of gradient-based reconstruction methods, including total variation (TV) methods [81], which minimize the total variation of the image gradient, and wavelet-based methods, which use wavelet transforms to decompose the image into different frequency bands and reconstruct the image using the gradient information in these bands.

Gradient-based reconstruction methods are widely used in image processing because they can effectively restore and reconstruct images that have been degraded by noise, blur, or missing information while preserving important image features such as edges and fine details. These methods can be used in a variety of applications, including medical imaging, satellite imaging, and microscopy.

The gradient method used within this research to obtain the smooth solution of an image is the Spectral Projection Gradient (SPG) method introduced in [9]. SPG is a deterministic optimization algorithm that is used to solve optimization problems of the form:

$$\min_{x \in \Omega} E_Q(x),$$

where the feasible region Ω is a closed convex set in $\mathbb{R}^n$ and E_Q is a smooth function. The method involves projecting the current iterate x onto the set Ω, $P_\Omega(x)$, using a spectral decomposition of the projection operator and then using the projected iterate as the next iterate in the optimization process.

SPG algorithm is outlined in Algorithm 3.6.1. To begin the reconstruction process, we initiate an arbitrary initial solution $u_0 \in \Omega$. The tolerance of the final stopping criterium is controlled by parameter Err. The prerequisites for the application of the SPG algorithm are as follows:

i) The projection P_Ω of any point $x \in \mathbb{R}^n$ onto the set Ω is defined.

ii) The function E_Q is defined and possesses continuous partial derivatives on an open set that encompasses Ω.

The SPG algorithm combines the non-monotone line search technique [37] and the spectral gradient step-length selection method [4, 11, 78]. If conditions i) and ii) are satisified, the algorithm converges to a constrained stationary point. See [9] for detailed analysis.

Algorithm 3.6.1 SPG optimization algorithm.

Initialize u^0;
Initialize $d^0 = P_\Omega(u^0 - \nabla E_Q(u^0)) - u^0$; Set $k = 0$;
repeat
  Determine the current step-length $\lambda^k > 0$ using a non-monotone line search approach;
  $u^{k+1} = u^k + \lambda^k d^k$;
  Calculate the gradient spectral step-length $\theta_{k+1} > 0$;
  $d^{k+1} = P_\Omega(u^{k+1} - \theta_{k+1} \nabla E_Q(u^{k+1})) - u^{k+1}$; $k = k + 1$;
until $\|u^k - u^{k-1}\|_\infty < Err$;

The SPG method utilizes the spectral properties of the operator that relate the image to the measurements or observations. It projects the gradient of the cost function onto the eigenspace of the operator and updates the image in the direction of the projected gradient. This allows the method to take into account the spectral properties of the operator and the regularization term and to converge faster to a solution.

This method has several attractive properties, including global convergence, fast convergence rates, and the ability to handle large-scale optimization problems. It is also relatively simple to implement and can be used with a wide range of optimization problems.

One of the main advantages of the SPG method is that it can handle optimization problems with complex constraints, such as those involving inequality or equality constraints or those involving multiple sets or subspaces.

CHAPTER 4

Area Based Shape Descriptors

As mentioned in previous chapters, in today's data-rich landscape, images are indispensable across various domains, including medicine, security, industry, geology, and archaeology. These images often contain numerous objects that need to be accurately recognized, categorized, and labeled. Initially, the traditional approach involved directly working and comparing the objects themselves, which often resulted in computationally intensive and inaccurate comparisons between objects. However, a more effective strategy involves converting these objects of interest into numerical representations, typically vectors in a mathematical space (often denoted as $\mathbb{R}^d$). This transformation allows for more advanced operations and analysis of the objects.

To make this transformation possible, we need to identify and quantify specific properties of the objects efficiently.

Shape represents an attribute that allows for numerical characterization and holds substantial potential for discriminating between objects. Over time, numerous shape descriptor techniques have been developed [90]. These descriptors encompass those tailored to specific shapes and those capturing shared characteristics across multiple shapes, including circularity [75], ellipticity, rectangularity, triangularity [80], symmetry [96], and more. Even within a single shape attribute, a variety of alternative measurements often exist.

Shapes can be analyzed based on information derived solely from boundary points (boundary-based) or from all points within the shape (area-based). Area-based methods are known for their robustness, especially in the presence of noise, and computational efficiency. Moreover, these techniques are well-suited for discrete domains such as digital images.

4.1 Geometric moments and moment invariants

Geometric moments are numerical characteristics of an image that are used for digital image processing and computer vision applications. They are used to describe the shape, size, orientation and other characteristics of an object or group of objects. Geometric moments provide a more accurate description of an object than other methods such as Fourier transform or the Hough transform.

Moment invariants are numerical characteristics of an image that remain unchanged under certain transformations such as scaling, rotation and translation. These invariants are used to recognize objects in an image regardless of orientation, scale or position. They can also be used to detect and track objects in an image. Moment invariants are computed using geometric moments.

The calculation of geometric moments involves the integration of intensity values across all pixels within an image. The moments are calculated by multiplying the pixel intensity by its coordinates, then summing over the entire image. Moment invariants are calculated by combining two or more moments with different orders.

The most commonly used moment invariants are the Hu invariants, which are a set of seven moment invariants developed by Hu in 1962. These invariants can be used to recognize objects in an image regardless of orientation, scale, or position. Other moment invariants include the Zernike moments, describing the shape of an object, and the Legendre moments, which are used to recognize objects in an image regardless of orientation.

Geometric moments and moment invariants are powerful tools that can be used to recognize, track, and classify objects in an image. Additionally, they can facilitate the identification of similarities between different objects.

The geometric (p, q)-moment of a given planar shape S is defined as:

$$m_{p,q}(S) = \iint_S x^p y^q \, dx \, dy. \tag{4.1}$$

Order of moment $m_{p,q}$ is $p + q$. When in discrete spaces, $m_{p,q}$ is approximated in the following way:

$$m_{p,q}(S) = \sum_{(i,j) \in dig(S)} i^p j^q, \tag{4.2}$$

where $dig(S)$ is the digitization of the real shape S, (i, j) are spatial coordinates of each pixel in the image, pixels are considered to be of size 1×1. Moments are used in image processing and computer vision to describe very common features. For example, the position of a shape, which is one of its basic features, is generally described in terms of moments. Specifically, the shape centroid $(x_c(S), y_c(S))$ tells us the position of a given shape S is. This is specified as:

$$(x_c(S), y_c(S)) = \left(\frac{m_{1,0}(S)}{m_{0,0}(S)}, \frac{m_{0,1}(S)}{m_{0,0}(S)} \right). \tag{4.3}$$

Next, we present central moments which are translation invariant

$$\overline{m}_{p,q}(S) = \iint_S (x - x_c(S))^p (y - y_c(S))^q \, dx \, dy. \tag{4.4}$$

Following we introduce the normalized moment. Normalized geometric moments are a valuable tool in image processing because they offer scale and rotation invariance, noise robustness, and the ability to capture essential shape characteristics of objects in an image. These properties make them suitable for a wide range of applications

$$\mu_{p,q}(S) = \frac{\overline{m}_{p,q}(S)}{m_{0,0}(S)^{1+\frac{p+q}{2}}}. \tag{4.5}$$

It is easy to see that $\mu_{p,q}(S) = \mu_{p,q}(rS)$, where r is a scaling factor.

In medical imaging, accurately identifying and characterizing tumors in CT scans is critical for diagnosis and treatment planning. However, tumors can vary in size, shape, and position within the body, and CT scans may capture them at different orientations. This variability makes it challenging to develop a consistent and automated method for tumor detection.

Hu moments introduced by Ming-Kuei Hu in [43] can help address this problem by providing a way to describe and compare tumor shapes while being invariant to rotation and scale. The following are the Hu moments.

$$
\begin{aligned}
I_1 &= \mu_{2,0} + \mu_{0,2} \\
I_2 &= (\mu_{2,0} - \mu_{0,2})^2 + 4(\mu_{1,1})^2 \\
I_3 &= (\mu_{3,0} - 3\mu_{1,2})^2 + (3\mu_{2,1} - \mu_{0,3})^2 \\
I_4 &= (\mu_{3,0} + \mu_{1,2})^2 + (\mu_{2,1} + \mu_{0,3})^2 \\
I_5 &= (\mu_{3,0} - 3\mu_{1,2})(\mu_{3,0} + \mu_{1,2})\left[(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{2,1} + \mu_{0,3})^2\right] \\
&\quad + (3\mu_{2,1} - \mu_{0,3})(\mu_{2,1} + \mu_{0,3})\left[3(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2\right] \\
I_6 &= (\mu_{2,0} - \mu_{0,2})\left[(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2\right] \\
&\quad + 4\mu 1,1(\mu_{3,0} + \mu_{1,2})(\mu_{2,1} + \mu_{0,3}) \\
I_7 &= (3\mu_{2,1} - \mu_{0,3})(\mu_{3,0} + \mu_{1,2})\left[(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{2,1} + \mu_{0,3})^2\right] \\
&\quad + (\mu_{3,0} - 3\mu_{1,2})(\mu_{2,1} + \mu_{0,3})\left[3(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2\right].
\end{aligned}
$$

A reader can notice that Hu moments use normalized moments hence exhibiting invariance to translation and scaling. Moreover, it can be demonstrated that these moments demonstrate rotational invariance [31].

Zernike moments [98] are a powerful set of orthogonal moments widely employed in image processing and pattern recognition for shape representation. Developed by Frits Zernike in 1934, these moments have gained popularity due to their desirable properties, such as rotation invariance and compactness in capturing shape information. Zernike moments are particularly useful for characterizing the boundary shape of objects or regions in images, making them well-suited for shape analysis tasks, such as object recognition and shape matching.

Zernike moments are based on the Zernike polynomials, which are a set of orthogonal polynomials defined over the unit disk. Each Zernike polynomial is associated with two integer indices, n and m, where n represents the radial order, and m denotes the azimuthal order. The radial order determines the number of nodes in the radial direction, while the azimuthal order governs the number of times the shape rotates around its center within 360 degrees.

The general form of the Zernike polynomial $Z_{n,m}(\rho, \phi)$ is given by:

$$Z_{n,m}(\rho, \phi) = R_{n,m}(\rho) \cdot e^{im\phi},$$

where ρ represents the radial distance from the center of the unit disk to a point on the boundary of the shape (normalized to lie within the unit circle, $0 \leq \rho \leq 1$), ϕ is the azimuthal angle of the point measured from the reference axis, and $R_{n,m}(\rho)$ is the radial polynomial defined as:

$$R_{n,m}(\rho) = \sum_{k=0}^{(n-m)/2} \frac{(-1)^k \cdot (n-k)!}{k! \cdot ((n+m)/2-k)! \cdot ((n-m)/2-k)!} \cdot \rho^{n-2k}.$$

To compute the Zernike moments for a given shape, the region of interest (ROI) containing the shape is first converted to a binary image, with the shape of interest represented as white pixels on a black background. The image is then mapped to the unit disk using polar coordinates, and the Zernike moments are calculated by integrating the product of the binary intensity values and the corresponding Zernike polynomials over the shape's boundary.

The calculation of Zernike moments can be computationally efficient due to their orthogonality properties, which allow moments of higher orders to be expressed as linear combinations of lower-order moments. This property leads to a reduction in the number of required calculations, making Zernike moments practical for real-time applications.

One of the key advantages of Zernike moments is their ability to capture shape features in a rotation-invariant manner. Since the Zernike polynomials are orthogonal, the moments computed from a shape and its rotated version remain the same, albeit with possible sign changes. This property is particularly beneficial for shape matching and recognition tasks, where the orientation of objects may vary.

Once the Zernike moments are computed, they serve as a compact and meaningful representation of the shape. These moments can be used as feature vectors for various shape analysis tasks, such as object recognition, shape classification, and shape matching. By comparing the Zernike moments of different shapes using appropriate distance metrics, shape similarities or dissimilarities can be quantified, enabling effective shape recognition and matching algorithms.

4.2 Shape orientation

Shape orientation is a crucial aspect of characterizing objects in image analysis, computer vision, and various other fields. One effective method for determining the orientation of a shape involves the use of geometric moments. In this section, we will explore how geometric moments can be employed to calculate the orientation of a shape, specifically by finding the axis of the least second moment of inertia. This method is widely used in the field of image processing for shape analysis [46, 90].

The primary objective in shape orientation determination is to minimize the integral of squared distances from points within a shape to a specific line. Mathematically, this can be formulated as follows:

$$I(\alpha, S, \rho) = \iint_S r^2(x, y, \alpha, \rho) dxdy. \tag{4.6}$$

Here, $I(\alpha, S, \rho)$ represents the integral of squared distances, α is the angle of the line with respect to a reference axis, and S is the shape of interest.

The distance function $r(x, y, \alpha, \rho)$ is defined as the perpendicular distance from a point $(x, y) \in S$ to a line given by:

$$X \sin \alpha - Y \cos \alpha = \rho. \tag{4.7}$$

The axis of the least second moment of inertia corresponds to the line that minimizes the integral defined in Equation (4.6). It is worth noting that this axis passes through the centroid of the shape. Consequently, we can set $\rho = 0$ and focus on minimizing $F(\alpha, S) = I(\alpha, S', \rho = 0)$, where S' is the translated shape such that its centroid coincides with the origin.

The squared distance function $r^2(x, y, \alpha, \rho = 0)$ simplifies to $(x \sin \alpha - y \cos \alpha)^2$, leading us to the minimizing function:

$$F(\alpha, S) = \sin^2 \alpha \, \overline{m}_{2,0}(S) + \cos^2 \alpha \, \overline{m}_{0,2}(S) - \sin \alpha \, \overline{m}_{1,1}(S). \tag{4.8}$$

In Equation (4.8), $\overline{m}_{2,0}(S)$, $\overline{m}_{0,2}(S)$, and $\overline{m}_{1,1}(S)$ represent the central moments of the shape S. These moments capture important information about the shape's geometry.

The orientation of a given shape S is determined by finding the angle α at which the function $F(\alpha, S)$ reaches its minimum. To identify this angle, we can look for points where the first derivative of $F(\alpha, S)$ equals zero.

Taking the derivative of $F(\alpha, S)$ with respect to α and setting it equal to zero, we obtain the orientation angle α:

$$\frac{\sin(2\alpha)}{\cos(2\alpha)} = \frac{2\overline{m}1, 1(S)}{\overline{m}2, 0(S) - \overline{m}_{0,2}(S)}. \tag{4.9}$$

Equation (4.9) provides a straightforward and computationally efficient method for determining the orientation of a shape based on its geometric moments. By solving this equation, we can find the angle at which the shape is oriented.

While the method described here is a fundamental and widely-used approach to shape orientation determination, it may have limitations, especially for highly symmetrical shapes. Researchers have developed additional methods to address these limitations [25, 39].

Geometric moments provide a robust and computationally efficient method for determining the orientation of shapes in image analysis. By minimizing the integral of squared distances to a line, based on central moments, the orientation angle of a shape can be readily calculated.

Shape orientation has practical applications in fields like binary tomography, as demonstrated in a paper by Lukić and Balazs [58]. In this context, shape orientation serves as a regularization condition during image reconstruction. The reconstruction process involves minimizing an energy or objective function, denoted as $E(u)$, where u represents the binary image to be reconstructed. The energy function for binary image reconstruction, proposed by the authors, encompasses three integral components: firstly, the Data Fitting Term assesses the correspondence between the reconstructed image and the observed projection data, ensuring their alignment. Secondly, the Smooth Regularization Term enforces uniformity in pixel values among neighboring regions in the reconstructed image, a characteristic commonly found in real images. Lastly, the Orientation Term evaluates how well the reconstructed image aligns with a predetermined orientation angle, denoted as α^*, representing the expected orientation of the original object within the image. Researchers adjust the impact of these terms using parameters, with the option to exclude the Orientation Term, simplifying the energy function into a form suitable for discrete tomography. As shown by the authors, by incorporating shape orientation as an additional constraint in the reconstruction process, it becomes possible to improve the quality of reconstructions, especially in scenarios with limited projection data.

4.3 Shape circularity

Circularity serves as a valuable metric for quantifying the roundness or resemblance to a circle of a given shape.

One of the most common ways to quantify circularity is through the standard circularity measure, denoted as $C_{st}(S)$. This measure exploits a fundamental geometric principle – the circle has the largest area among all shapes with the same perimeter. Accordingly, $C_{st}(S)$ is defined as follows:

$$C_{st}(S) = \frac{4\pi A(S)}{(P(S))^2}, \tag{4.10}$$

where $A(S)$ represents the area of the shape S, while $P(S)$ represents the perimeter of the shape S. Notably, $C_{st}(S)$ takes into account both interior and boundary information of the shape, making it a holistic descriptor of circularity.

Another approach to assess circularity involves utilizing Hu invariants, which are geometric invariants associated with a shape. The circularity measure $C(S)$, introduced by [100], incorporates the Hu invariant I_1:

$$C(S) = \frac{1}{2\pi(\mu_{2,0} + \mu_{0,2})} = \frac{1}{2\pi I_1(S)}, \tag{4.11}$$

where $\mu_{2,0}$ and $\mu_{0,2}$ are the central moments of the shape S. Unlike $C_{st}(S)$, $C(S)$ is primarily area-based and does not penalize deep intrusions into the shape. This property makes it more robust to noise, as it focuses on the overall shape characteristics rather than the shape's perimeter. In contrast, $C_{st}(S)$ is sensitive to shape irregularities that lead to significant perimeter increases.

CHAPTER 5

Thesis Contributions

5.1 Graph cut optimization

Graphs and directed graphs (also known as digraphs) are fundamental mathematical structures used in various fields, including image processing and image reconstruction [56].

Definition 3 *The graph G is an ordered pair $G = (X, \rho)$, where X is a finite non-empty set of elements called nodes (vertices), and ρ is a finite set of ordered or un-ordered pairs, with distinct elements from the set X, called edges.*

The edges of the graph represent relationships or connections between the nodes. Graphs can be used to model and represent various types of relationships and data. They are typically categorized into two main types, undirected and directed graph.

Definition 4 *An undirected graph $G = (X, \rho)$ is an ordered pair that satisfies the following conditions:*

1. *X is a finite non-empty set of elements called nodes,*
2. *ρ is a finite set of unordered pairs of distinct elements from X, representing edges.*

In an undirected graph, the edges in G have no direction; they connect two vertices without specifying a starting or ending point. Undirected graphs are used to represent relationships where the order of connection between nodes does not matter, such as social networks or road networks.

Definition 5 *A directed graph (or digraph) $G = (X, \rho)$ is an ordered pair that satisfies the following conditions:*

1. *X is a finite non-empty set of elements called nodes.*

2. *ρ is a finite set of ordered pairs of distinct elements from X, representing directed edges.*

Directed graphs are used to represent relationships with a clear direction, such as flow networks, dependencies, or sequential processes.

Definition 6 *A weighted graph, denoted as $G = (X, \rho, w)$, is an ordered triple where X is a finite non-empty set of nodes, ρ is a finite set of edges (ordered or unordered pairs of distinct elements from X), and $w : \rho \to \mathbb{R}$ is a function associating each edge in ρ with a real number, referred to as its weight.*

The weights in a weighted graph represent some measure or cost associated with traveling from one node to another along the edge. In other words, a weighted graph assigns a numerical value to each connection in the graph to represent the distance, cost, capacity, or any other relevant quantity between the vertices connected by that edge. Graphs play a crucial role in image processing and reconstruction by providing a structured way to represent and analyze image data. Graphs can be used to represent images where each pixel or region is a node, and edges connect neighboring pixels or regions. Techniques like graph-based segmentation use graph properties to partition an image into meaningful segments. Graph-based methods, such as graph cuts and spectral graph theory, are used to compress images efficiently while preserving important features.

5.1.1 Graph cuts

Graph cut is a powerful technique used in image processing and computer vision for a variety of tasks, such as image segmentation, object recognition, and image matting.

Definition 7 *Let $G = (X, \rho, w)$ be a directed weighted graph, with non negative weights, and let $a, b \in X$. An $a - b$-cut of a graph G is a partition of nodes X into two disjoint sets, A and B, so that $a \in A$ and $b \in B$. The cost of the cut, denoted by cut(A, B), is the sum of the costs of all edges that go from A to B:*

$$cut(A, B) = \sum_{\substack{u \in A \\ v \in B}} w(u, v),$$

where A and B form a partition of the set of nodes X, i.e., $A \cup B = X$ and $A \cap B = \emptyset$.

For an $a - b$ - cut, node a is often referenced as a source and node b as a sink. Source and sink nodes are refereed as terminals. The graph cuts algorithm aims to find the optimal cut that minimizes the cost of partitioning set X. Algorithms for this purpose are detailed in [14]. If the weights in the graph are derived from a certain energy function, the graph cuts algorithm becomes a powerful tool for energy minimization. In this scenario, finding the optimal graph cut is equivalent to minimizing the energy function. The algorithm excels in solving binary or multi-label segmentation problems, efficiently determining an optimal partition that balances data fidelity and smoothness.

In the graph cuts application in image processing, the image is represented as a directed weighted graph, where pixels are nodes, and edges encode relationships between pixels (e.g. nodes representing pixels that are spatially adjacent are connected by edges). In general, the direction of the edge between neighbouring pixels is arbitrary and can be chosen based on the specific implementation of the graph cuts algorithm . The energy function $E(X)$ drives the labeling configuration of pixels and consists of a data term and a smoothness term:

$$E(A, B) = \sum_{(i,j) \in \delta(A,B)} w_{ij} + \lambda \sum_{i \in A} d_i.$$

Here, $\delta(A, B)$ represents edges crossing the cut, w_{ij} is the weight associated with edge (i, j), and d_i is a data term capturing fidelity to observed data. The algorithm involves creating a specialized graph corresponding to the energy function. Regular image pixels become nodes, and edges are of two types, $n - links$ for neighboring pixels and $t - links$ for connecting pixels to terminals. Edge costs derive from the energy function, and the minimum cut on this graph minimizes the energy. Each graph cut provides a different image segmentation.

The choice of source and sink nodes depends on the specific requirements of image processing task. The source node represents the starting set, often associated with the foreground or the region of interest in the image. In the context of image segmentation, the source node could be related to the pixels or regions that should be identified or highlighted. The sink node represents the destination set, often associated with the background or the region we want to distinguish from the foreground.

The cornerstone of Graph Cuts in image reconstruction is the Min-Cut/Max-Flow algorithm [32]. This algorithm seeks to find the minimum cut in a flow network, which is modeled using the constructed graph G. The final cut yields the optimal segmentation of the image into disjoint regions, balancing the trade-off between data fidelity and smoothness.

Graph cut optimization offers a convenient approach for tackling a diverse range of image processing challenges that can be expressed in the context of energy minimization, as documented in various studies [8, 12, 13, 15, 16, 49, 52, 55].

The graph cuts method has several advantages over other image reconstruction techniques. For example, it can handle large amounts of missing data and can produce good results even when the quality of the original image is poor. However, it can be computationally expensive and may not always produce the best results in all cases.

5.1.2 Potts model

Graph cuts serve as a potent tool for energy minimization, particularly in addressing image reconstruction challenges. The key lies in formulating an energy function that is amenable to graph cuts and can effectively tackle image reconstruction problems. One such formulation is provided by the Potts model [95].

To apply the Potts model for image reconstruction, a graph is created to depict pixel relationships. Nodes in this graph correspond to image pixels, and edges con-

0	0.3	0.9
0.2	0.4	0.8
0.1	0.7	1

Fig. 5.1.14: Given gray scale image for graph cut optimisation

nect neighboring pixels. Labels denote potential intensity values or categories for each pixel. The objective is to group nodes, maximizing edges within groups while minimizing edges between groups. This proves valuable in image segmentation tasks, where the aim is to partition an image into distinct regions or objects.

In our application, the Potts model is governed by the minimization of the following energy equation:

$$E(d) = \sum_{p \in \mathcal{P}} D(p, d_p) + \sum_{(p,q) \in \mathcal{N}} K_{(p,q)} \cdot (1 - \delta_{d_p, d_q}), \tag{5.1}$$

where $d = \{d_p \,|\, p \in \mathcal{P}\}$ represents the labeling of image pixels $p \in \mathcal{P}$. The term $D(p, d_p)$ denotes the data cost, which is a penalty or cost associated with assigning a label d_p to a pixel p. The interaction potential between neighboring pixel pairs p and q is denoted as $K_{(p,q)}$, where $\mathcal{N}$ represents the set of neighboring pixel pairs. The function δ_{d_p, d_q} is the Kronecker delta function defined as:

$$\delta_{i,j} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} .$$

The second term in 5.1 promotes spatial coherence by penalizing inconsistencies between adjacent pixels. The goal of image reconstruction in the Potts model is to find the labeling d that minimizes the total energy of the system.

Now, let's bridge these theoretical concepts with a practical example. Consider a grayscale image, as illustrated in Figure 5.1.14. To achieve the objective of image binarization, the image is processed using the Potts model and graph cuts algorithm. We start by constructing a directed graph associated with the image, as shown in

Figure 5.1.15a. In this graph, each pixel corresponds to a node, and edges connect neighboring pixels, reflecting the defined energy function. In our specific case (Potts model), interactions between neighboring pixels are implemented using a Kronecker delta function, which registers only whether the pixels are adjacent or not. To maintain coherence with the general definition, in our example two directed edges (p, q) and (q, p) between neighbouring pixels p and q are created, both of the edges have the constant weight of 1. Edges connecting nodes to the labels (0) and (1) are weighted based on the terminal label assignment cost. The function $D(p, d_p)$ represents the cost of assigning the label (source or sink) d_p to pixel p, and it depends on how close the intensity of the pixel $u(p)$ is to the value d_p. The values of the function $D(p, d_p)$ are specified at the beginning of the process. Subsequently, based on the values of $D(p, d_p)$, the weights $w(p, d_p)$ of edges (p, d_p) are calculated in the corresponding directed graph.

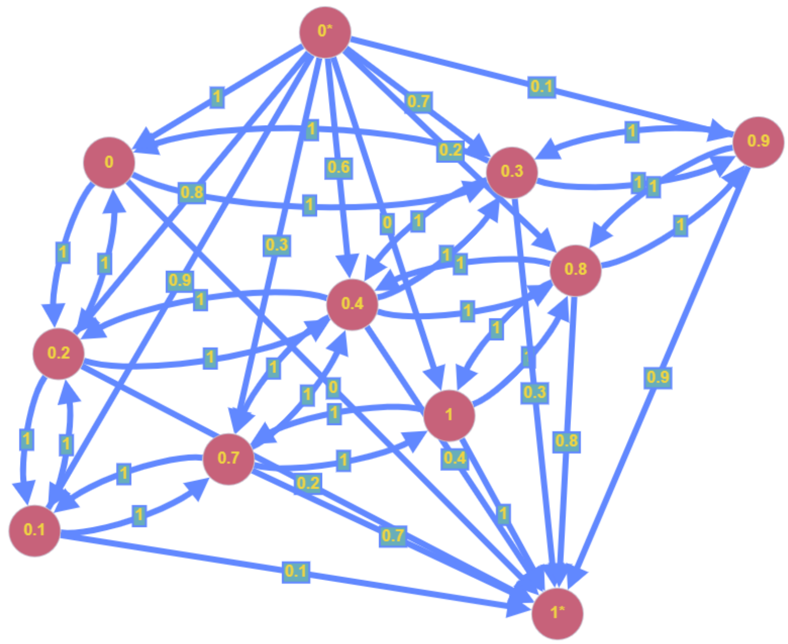
In our example we define $w(p, q)$ in the following manner:

$$w(p,q) = \begin{cases} 1, & \text{p and q are neighbouring pixels} \\ |u(p) - 1|, & q = 0^* \\ |u(p)|, & q = 1^* \end{cases},$$

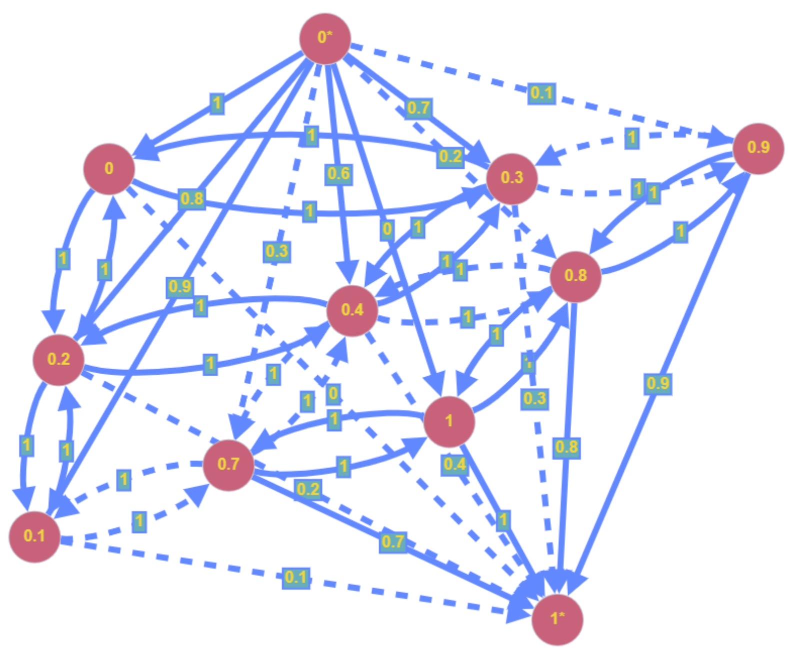
where $u(p)$ is the intensity of the pixel p, 0^* and 1^* are source and sink respectively. This selection for the weights function was made in order to ensure that the minimum cut separates less similar pixels.

Figure 5.1.15b illustrates one cut on the graph. The cut represents one segmentation of the image. The cost of the cut is calculated as the sum of weights of the removed edges. In our case, $cut(0^*, 1^*) = 4 \cdot 2 \cdot 1 + 0.3 + 0.4 + 0.2 + 0.1 + 0 + 0.1 + 0.2 + 0.3 + 0 = 9.6$.

To achieve image binarization, the algorithm seeks the minimum cut in the graph Figure 5.1.15a. Various algorithms, such as those developed by Kolmogorov and Zabih [53], are employed to find the optimal cut. Once optimal graph is obtained, the image graph is partitioned into two sets as shown on Figure 5.1.16a. Asigning 0

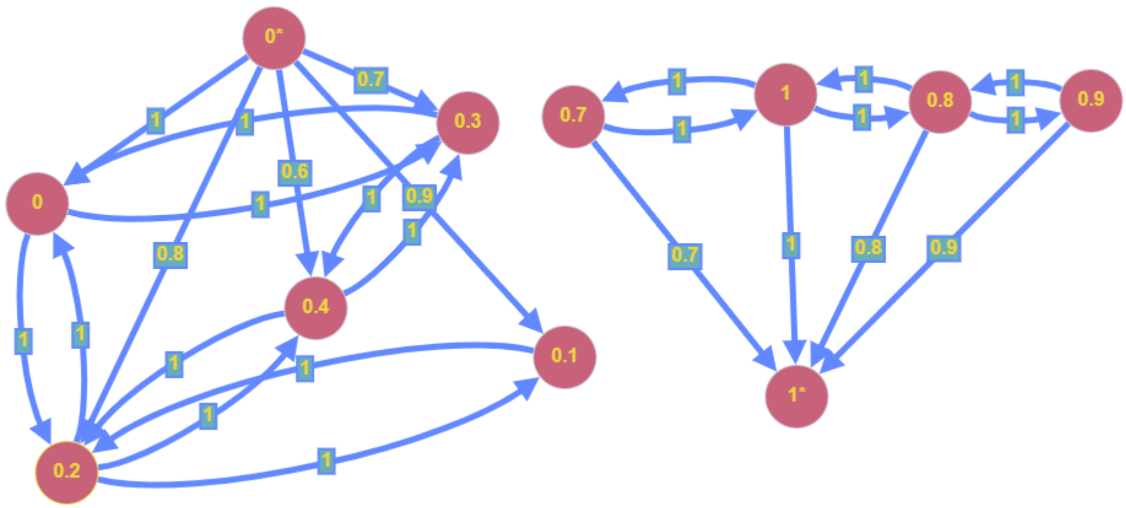


(a) Directed graph created for image 5.1.14 associated to the energy E, numbers on edges reflect their weights



(b) Cut on the graph 5.1.15a, node 0^* is the source, node 1^* is the sink

Fig. 5.1.15: Specially designed graph for image 5.1.14 and its cut



(a) Optimal cut of the graph 5.1.15a

0	0	1
0	0	1
0	1	1

(b) Binary reconstruction of the image 5.1.14

Fig. 5.1.16: Graph cut reconstruction

label to all the nodes connected to the source and label 1 to the nodes connected to the sink yields the binary reconstruction (Figure 5.1.16b). This final result effectively separates foreground and background pixels, minimizing the overall energy of the system.

The Potts model tackles a discrete optimization problem, where the image is represented by a set of discrete labels. The regularization term accounts for the cost of assigning different labels to neighboring pixels, with the cost defined as a function of the labels.

5.1.3 Discrete tomography reconstruction based on graph cuts method

In this chapter, we delve into a novel approach to discrete tomography reconstruction, fusing the powerful graph cuts method with the quadratic iterative minimization technique. Our journey begins with the computation of data cost values for each pixel within the image. These values are extracted from the intensity of the reconstructed image, which is the solution to the energy-minimization problem articulated as follows:

$$\min_{u\in[0,1]^N} E_Q(u) := \|Au - b\|^2, \tag{5.2}$$

$$A \in \mathbb{R}^{M\times N},\ \ u \in \Lambda^N,\ b \in \mathbb{R}^M, \quad \Lambda = \{\lambda_1, \lambda_2, ..., \lambda_k\}, \quad k \geq 2.$$

.

E_Q takes the form of a quadratic function, $\Omega = [0,1]^N$ represents the feasible set, k is the number of different gray level values and the set Λ is given by the user and represents pixel labels. We solve this minimization problem through a selection of optimization algorithms, with our preferred choice being the Spectral Projected Gradient (SPG) optimization algorithm.

For the SPG algorithm to be effective, following prerequisites must be met:

For the SPG algorithm to operate effectively, certain conditions must be satisfied:

1. The objective function should possess continuous partial derivatives within an open set that encompasses the Ω.

2. An availability of the projection function P_Ω for any given vector onto the set Ω.

The objective function in (5.2) is differentiable in $\mathbb{R}^N$, and the projection P_Ω is defined as:

$$[P_\Omega(u)]_i = \begin{cases} 0, & u_i \leq 0 \\ 1, & u_i \geq 1 \\ u_i, & \text{elsewhere} \end{cases}, \qquad \text{where } \ i = 1, \ldots, N\,.$$

Thus, we fulfill the necessary conditions for the application of the SPG algorithm. The outcome of the SPG optimization is a smooth solution to (5.2).

Our subsequent step entails discretizing this smooth solution obtained from Equation 5.2, which we achieved using the SPG algorithm. We draw upon discrete tomography reconstruction algorithms proposed by Schüle et al. [85, 93] and Lukić et al. [58, 61], and opt for the Potts interaction model for its capacity to promote compactness in solutions, as observed in [14, 36, 89]. The data cost term, denoted as D in (5.2), and defined in (5.3) is formulated based on the pixel intensities, $u(p)$, and is designed to be small or inexpensive near specific gray values.

$$D(p,i) = |u(p) - \lambda_i|, \quad \text{for } i = 0, 1, 2, \ldots, k-1. \tag{5.3}$$

The interaction potential, $K_{(p,q)}$, between adjacent pixels, in our method, is set as a constant value of 1. The energy function in (5.1) is then minimized utilizing the Graph Cuts Optimization (GCO) algorithm [14, 16, 27, 53]. The GCO algorithm assigns a label value, d_p, to each pixel, which corresponds to a predefined gray level and determines the pixel intensities in the final discrete solution.

The technique described in this section is firstly introduced by Šulc and Lukić in [59] and is referred to as the Graph Cuts Discrete Tomography (GCDT) reconstruction method.

5.1.4 Experimental results for GCDT method

In this section, we provide a brief showcase of the effectiveness of the suggested reconstruction technique based on graph cuts (referred to as GCDT) to justify its application in the later introduced methods. We use the Shepp-Logan image for evaluation, and additional experimental results can be found in Section 5.4. The Shepp-Logan image is a well-known phantom in discrete tomography, containing 6 gray levels. We perform experiments employing multiple projection directions, capturing 128 parallel rays for each direction. These rays are uniformly sampled within the range of 0 to 180 degrees.

We compare the results obtained from the GCDT with DART method [5]. Furthermore, as a control method, we employ a basic yet less potent approach called TRDT (Thresholding and Discretization Technique). Within this approach, we apply a thresholding function on the continuous solution and assign predetermined gray levels based on pixel intensities.

PH1 PH2 PH3 Shepp-Logan

Fig. 5.1.17: Original test images (128×128). Phantoms PH1, PH2, and PH3 are composed of three distinct gray levels, specifically 0, 0.5, and 1. Shepp-Logan phantom comprises six different gray levels, namely $0, 0.1, 0.2, 0.3, 0.4$, and 1.

Figure 5.1.17 shows the original test images used in the experiments, including Shepp-Logan.

Table 5.1.1 presents the summary of the experimental results. We assess the outcomes achieved through the application of three distinct reconstruction techniques: TRDT, DART, and GCDT. The results pertaining to the DART technique are ex-

smooth rec.

12 proj. 15 proj. 18 proj.

Fig. 5.1.18: Reconstructions of the Shepp-Logan test images by the proposed GCDT method.

Table 5.1.1: Results from experiments conducted on the Shepp-Logan image employing three distinct reconstruction methods. The term "m.r." denotes the misclassification rate, where a lower value signifies superior reconstruction, and "d" represents the quantity of projections.

	d	TRDT (m.r. %)	DART (m.r. %)	GCDT (m.r. %)
Shepp-Logan	12	12.74	14.21	5.72
	15	10.44	8.44	3.17
	18	10.03	2.56	2.14

tracted from the work of Batenburg and Sijbers [5]. The experimental setup involves the acquisition of projection data from three different configurations, namely 12, 15, and 18 projection directions. Notably, across all these configurations, the GCDT method consistently emerges as the most efficient, yielding the lowest mean reconstruction error (m.r.) values, where m.r. is the pixel error measure relative to the total number of image pixels.

In summary, our experimental findings underscore the effectiveness of the proposed GCDT method in the context of Shepp-Logan image reconstruction. This method surpasses alternative techniques in terms of both reconstruction quality and computational efficiency. For a comprehensive exploration of additional experimental results and detailed performance analyses, readers are encouraged to consult the original research paper [59].

5.2 A priori information in image reconstruction

A priori information, also known as prior information or prior knowledge, is information that is known or assumed before considering new data. In the context of image reconstruction algorithms, a priori information about the shape or characteristics of the object being reconstructed can be valuable in improving the accuracy and efficiency of the reconstruction process.

There are various ways to obtain a priori information, and the method often depends on the specific application and the nature of the imaging problem. In some cases, a priori information can be obtained through direct physical measurements of the object. For example, in medical imaging, certain characteristics of tissues may be known from previous experiments or measurements using different imaging modalities. Mathematical models and simulations can provide a priori information about the expected characteristics of the object. Researchers may use computational models based on physics or other relevant principles to simulate how an object should appear in the given imaging system. Machine learning techniques can be employed to learn a priori information from a large dataset of representative examples. Convolutional neural networks (CNNs) and other deep learning approaches can learn patterns and features from training data, capturing the inherent structures in images. Knowledge from domain experts can be considered as a valuable source of a priori information. Expert input can help define constraints, assumptions, or features relevant to the specific application.

The accuracy of image reconstruction algorithms can be sensitive to the correctness of a priori information. If the prior information is inaccurate or does not align well with the actual characteristics of the object, it may lead to artifacts or errors in the reconstructed image.

Some algorithms are designed to be more robust to uncertainties in the a priori information. For example, Bayesian frameworks allow for the incorporation of uncertainties, and regularization techniques can help mitigate the effects of inaccurate prior information.

There is often a trade-off between relying on a priori information and adapting to the data. Balancing the weight given to prior information versus the observed data is a critical aspect of algorithm design. In summary, obtaining a priori information can involve various methods, and the impact of inaccuracies in this information depends on the specific algorithm and application. Robust algorithms are designed to handle uncertainties and balance the influence of prior information with the observed data.

5.3 Graph cuts reconstruction methods assisted by shape circularity and shape orientation

The idea of incorporating a priori information into the reconstruction process is rooted in the belief that supplementing the algorithm with relevant knowledge about the imaged object can improve accuracy and robustness. In this regard, the question arises: Can the shape descriptors described in chapter 4 serve as a valuable a priori information, potentially replacing the need for extensive projection data in the reconstruction process?

To further enhance the performance of graph cuts reconstruction methods, we explore the possibility of introducing an additional regularization term representing shape circularity or shape orientation.

The hypothesis suggests that even with limited projection data from a single direction, knowing an object's circularity could lead to promising reconstruction outcomes. By integrating circularity as prior knowledge, the reconstruction algorithm can be regularized, steering it toward more plausible and physically meaningful solutions.

In addition to circularity, shape orientation is another important aspect of a shape's geometry. It can provide crucial information about the object's alignment, which is often valuable in various imaging applications.

These novel approaches aim to capitalize on the benefits of circularity and orientation as a priori information. By integrating both descriptors into the graph cuts reconstruction algorithm, we seek to achieve superior results compared to traditional methods that rely solely on projection data.

In the subsequent sections of this chapter, we will delve deeper into the practical implementation of shape circularity and orientation as regularization terms in graph cuts reconstruction methods. We will explore experimental results and case studies to validate the hypothesis that leveraging these geometric descriptors can indeed lead to more accurate and robust reconstructions in scenarios with limited projection data.

5.3.1 The new method based on shape orientation

Our novel tomography reconstruction approach seamlessly integrates the graph cuts method with a gradient-based minimization technique, all while leveraging shape orientation as crucial a priori information.

In the initial step of our method, we calculate data cost values for each pixel within the image. These values are derived from the intensities of a smoothly approximated final reconstructed image, achieved through the minimization of an energy function:

$$\begin{aligned}\min_{u\in[0,1]^N} E_Q(u) =& w_P\|Au-b\|^2 + w_H \sum_{i=1}^{N}\sum_{j\in\Upsilon(i)}(u_i-u_j)^2 \\ &+ w_{\mathcal{O}}(\Phi(u)-\alpha^*)^2 + \mu\langle u, \tau-u\rangle.\end{aligned} \tag{5.4}$$

Here, we introduce key elements:

1. Data fitting term, $\|Au-b\|_2^2$, regularized by the parameter $w_P > 0$. This term ensures adherence to the projection data. The regularization parameter w_P controls the impact of this term. A higher value of w_P penalizes deviations from the data more strongly, leading to a solution that closely fits the observed data. However, setting it too high may result in overfitting.

2. Homogeneity term, $\sum_{i=1}^{N}\sum_{j\in\Upsilon(i)}(u_i-u_j)^2$, regularized by the parameter $w_H > 0$. Here, $\Upsilon(i)$ represents the set of indices of neighboring pixels (in the x and y axis directions) of pixel i. This term encourages neighboring pixels to have similar intensities, promoting smoothness in the solution. A higher w_H strengthens the smoothness constraint. It helps in reducing noise in the solution. Too high values, though, may oversmooth the solution, potentially causing a loss of important details.

3. Term, $(\Phi(u)-\alpha^*)^2$, which quantifies the disparity in orientation between the current solution $\Phi(u)$ and the known orientation of the original image α^*. The parameter $w_{\mathcal{O}} > 0$ determines the impact of the orientation regularization.

4. Concave regularization term, $\langle u, \tau-u\rangle$, where $\tau = [1,1,...,1]^T$ is a vector of size N. This term helps in moving pixel intensities toward binary values, and its influence gradually increases during the reconstruction process, regulated by the parameter $\mu > 0$. The parameter μ controls the strength of this regularization. Higher values of μ increase the influence of this term, which can be useful for obtaining binary solutions. However, an excessively high μ might lead to a binary solution that does not capture the underlying structure.

For each fixed μ, we employ the Spectral Projected Gradient (SPG) iterative optimization algorithm to solve the problem (5.4). In the subsequent step, we perform a comprehensive binarization of the smoothed solution obtained from the SPG algorithm. This binarization utilizes the graph cuts method based on the Potts model (details in Section 5.1.3). The data cost term D in (5.1) is crafted from the information derived from the smooth solution u:

$$D(p,0) = u(p),$$

$$D(p,1) = 1 - u(p).$$

We also define a set of neighboring pairs, denoted as $\mathcal{N}$. For two different pixels p and q, we have that $(p,q) \in \mathcal{N}$ if the image coordinates of p and q differ by at most 1 in either the horizontal or vertical direction. The interaction potential $K_{(p,q)}$ is set as a constant with a value of 1.

With these definitions, we are ready to minimize the energy function in (5.1) using the GCO (Graph Cuts Optimization) algorithm [16]. The GCO algorithm assigns label values d_p to each pixel p, where each label value corresponds to either 0 or 1.

We henceforth refer to this method as the Graph Cuts Binary Tomography Assisted by Orientation (GCORIENTBT).

5.3.2 The new method based on shape circularity

The approach suggested in this section for addressing the discrete tomography problem is divided into two components:

1. Identification of a continuous (smooth) solution to the energy minimization problem through the utilization of a gradient-based minimization technique. The energy function incorporates information about the circularity of the original object.

2. Making the acquired smooth solution discrete by employing a graph cuts-based

algorithm. The pixel values from the smooth image are employed to establish the data cost term for the graph.

The energy function used for calculating the smooth solution is given by the following equation:

$$\begin{aligned}\min_{u\in[0,1]^N} E_Q(u) := w_P\|Au-b\|_2^2 + w_H \sum_{i=1}^{N} \sum_{j\in\Upsilon(i)} (u_i-u_j)^2 + \\ + w_C\left(\mathcal{C}(u)-\mathcal{C}^*\right)^2 + \mu\left\langle u, \tau-u\right\rangle,\end{aligned} \tag{5.5}$$

and is constructed of following terms:

- $\tau = [1, 1, \ldots, 1]^T$ as an N-dimensional vector,
- w_P term for controlling data fitting,
- w_H term for regulating homogeneity of the solution,
- $\Upsilon(i)$ to represent neighboring pixel indexes of pixel i,
- $\mathcal{C}(u)$ for the circularity of the solution,
- $\mathcal{C}^*$ as the true circularity (a priori information),
- $w_C > 0$ to determine circularity regularization impact,
- $\langle u, \tau - u\rangle$ to encourage pixel intensities toward binary values,
- μ to control the influence of the binarization term.

The problem of minimizing energy, as expressed in equation (5.5) through constrained quadratic optimization, can be tackled using various optimization methods. We specifically chose the Spectral Projected Gradient (SPG) algorithm [9] due to its proven effectiveness in similar problem domains.

Efficient minimization and extraction of the smooth solution in the energy function (5.5) are made possible by analytically determining the gradient for the regularization term $(\mathcal{C}(u) - \mathcal{C}^*)^2$. The termination condition for the smooth solution is defined as $\langle u, \tau - u \rangle < E_{\text{bin}}$, with E_{bin} determining the level of binarization for the solution u and set to 100 in our experiments.

After computing the smooth solution, the next step involves complete binarization, achieved through the graph cuts method based on the Potts model. The resulting label values d_p for each pixel p (where $d_p = 0 \rightarrow 0$ and $d_p = 1 \rightarrow 1$) determine pixel intensities in the final binary solution, indicating the conclusion of the reconstruction procedure. This approach is referred to as Graph Cuts Binary Tomography Assisted by Circularity (GCCIRCBT).

GCCIRCBT provides a notable advantage over GCORIENTBT, primarily due to its analytical determination of the regularization gradient. This enables a swift determination of the smooth solution using the SPG algorithm, leading to a significant reduction in algorithm runtime compared to existing similar techniques.

5.4 Experimental results

In this section, we assess the effectiveness of different algorithms in reconstructing discrete tomography images, with particular emphasis on our proposed methods (GCDT, GCORIENTBT, GCCIRCBT). To gauge the effectiveness of our approaches, we conducted comprehensive experiments using a diverse set of test images (Figure 5.4.19). PH1-3 encompass 3 shades of gray, PH4-6 encompass 6 shades of gray, and PH7-12 depict binary images. PH1-PH11 consist of synthetic images, while PH12 is a binary-segmented fluorescence image of Calcein-stained Chinese hamster ovary cells. Each projection direction for multi-gray-level images involves a total of 128 parallel rays, whereas binary images utilize 64 projection rays. In every case, the projection directions are evenly distributed across the range of 0 to 180 degrees. This set of projection data is utilized as input for the reconstruction algorithms.

Our evaluation includes a comparison of several established reconstruction algorithms:

- Graph Cuts Discrete Tomography Algorithm (GCDT) [59]

- Discrete Algebraic Reconstruction Technique (DART) [5]
- Method based on classical threshold (TRDT)
- Multi Well Potential based method (MWPDT) [57]
- Graph Cuts Tomography Assisted by the Orientation prior (GCORIENTBT) [66]
- Graph Cuts Binary Tomography Assisted by the Circularity prior (GCCIRCBT) [67]

All the reconstruction methods (GCDT, DART, TRDT, MWPDT, GCORIENTBT, GCCIRCBT) are fully implemented in the Matlab programming language.

In our experiments, we employed the GCORIENTBT and GCCIRCBT reconstruction process with the parameter values given in table (5.4.2).

Parameter	**Value**
Initial vector	$u_0 = [0.5, 0.5, \ldots, 0.5]^T$
Data fitting weight	$w_P = 0.1$
Homogeneity weight (GCORIENTBT)	$w_H = 0.5$
Homogeneity weight (GCCIRCBT)	$w_H = 0.1$
Orientation preservation weight	$w_{\mathcal{O}} = 0.1$
Circularity regularization weight	$w_C = 3000$
Concave regularization weight	$\mu = 0.0001$

Table 5.4.2: Model parameter values used in the experiments

These values were determined through an iterative experimental process on our test set. To obtain these values, we performed multiple runs, adjusting the parameters based on the observed reconstruction quality. The selection aimed to strike a balance between model adherence to the data, smoothness of the solution, preservation of circular patterns or aimed orientation, and promotion of binary-like intensity values. Future improvements may involve exploring Adaptive Regularization Weighting techniques, allowing the algorithm to dynamically adjust parameters during the optimization process based on the characteristics of the data. Additionally, conducting a systematic sensitivity analysis and employing cross-validation techniques could further enhance the robustness and generalization of the reconstruction algorithm.

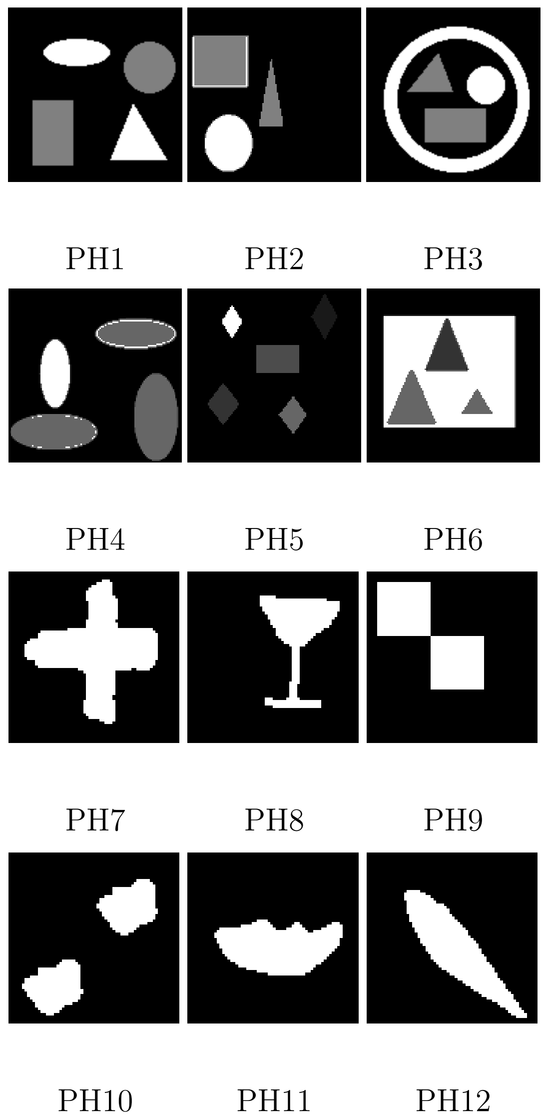


Fig. 5.4.19: Original test images. Phantoms PH1, PH2, and PH3 contain three distinct gray levels, while PH4, PH5, and PH6 contain 6 different gray levels. Phantoms PH7 through PH12 feature binary images.

5.4.1 Quality metrics

To provide a comprehensive evaluation of reconstruction quality, we employ a set of fundamental metrics. For all these metrics, a lower value indicates better reconstruction.

- Pixel Error (PE): This metric quantifies the absolute number of misclassified pixels, offering insights into the accuracy of reconstructions.
- Misclassification Rate (m.r.): It provides a normalized view of pixel error relative to the total number of image pixels, helping us understand the overall fidelity of the reconstructions.
- Projection Error (PRE): PRE assesses the alignment of the reconstructed images with the given projection data, which is critical for accurate reconstruction.

The direction of projection is determined by the angle α, and we denote the number of different projection angles used as d. When using horizontal and vertical projection data, we can effectively determine circularity and orientation shape descriptors, as discussed in Lukić et al. [62]. Therefore, when using three or more projection angles, including circularity and orientation as prior information becomes redundant, as these attributes are already present in the projection values. Consequently, we omit results for GCORIENTBT and GCCIRCBBT when using a higher number of projections since they would be identical to those obtained by GCDT.

5.4.2 Comparison of algorithms

The outcomes pertaining to the effectiveness of various algorithms on test images PH1, PH2, and PH3 have been summarized in Table 5.4.3 and Table 5.4.4, with corresponding visual representations in Figure 5.4.20 and Figure 5.4.22. In terms of the metrics PE and m.r., the GCDT method demonstrated superior performance in 10 out of 12 instances, while for the PRE metric, GCDT excelled in 8 cases. However, when considering execution time, the MWPDT method emerged as the leader, as GCDT necessitated a notably higher number of iterations to achieve a smooth solution.

The reconstruction outcomes for phantoms featuring 6 distinct gray levels are detailed in Table 5.4.5. In this context, the GCDT method surpassed TRDT and DART in 10 out of 12 scenarios, with DART exhibiting superior performance in the remaining 2 cases. Visual representations of the reconstructions from 6 and 15 projection directions are provided in Figure 5.4.21 and Figure 5.4.23, respectively.

The analysis conducted thus far underscores the competitive efficacy of a model combining graph cuts and a gradient-based approach (GCDT). This promising outcome has spurred further testing and refinement of the algorithm.

Table 5.4.3: Results from experiments conducted on images PH1, PH2, and PH3 are presented, employing three distinct reconstruction methods. The symbol d denotes the number of projections, and the most outstanding performance is highlighted in bold font.

	PH1				PH2				PH3			
d	6	9	12	15	6	9	12	15	6	9	12	15
MWP (PE)	**255**	159	59	35	**143**	138	20	18	655	456	275	174
MWP (m.r. %)	**1.55**	0.97	0.36	0.21	**0.87**	0.84	0.12	0.11	3.99	2.78	1.67	1.06
TRDT (PE)	412	175	48	28	209	141	17	17	412	301	101	41
TRDT (m.r. %)	2.51	1.06	0.29	0.17	1.28	0.86	0.10	0.10	2.51	1.83	0.61	0.25
GCDT (PE)	272	**69**	**8**	**5**	225	**124**	**12**	**12**	**272**	**116**	**20**	**9**
GCDT (m.r. %)	1.66	**0.42**	**0.04**	**0.03**	1.37	**0.76**	**0.07**	**0.07**	**1.66**	**0.70**	**0.12**	**0.05**

Table 5.4.4: The experimental findings for images PH1, PH2, and PH3, employing three diverse reconstruction methods. In this context, "e.t." represents elapsed time in minutes, while d signifies the number of projections. The optimal performance is denoted in bold font.

		PH1				PH2				PH3			
d		6	9	12	15	6	9	12	15	6	9	12	15
MWPDT	(PRE)	**14.70**	12.19	9.96	9.08	**14.11**	**18.94**	6.08	7.71	**19.83**	18.77	18.80	16.43
	(e.t.)	1.76	2.63	3.17	4.06	5.34	8.17	6.36	11.62	2.19	2.87	4.30	4.66
TRDT	(PRE)	18.66	14.72	10.61	8.87	17.98	17.30	7.09	7.09	23.64	17.87	13.66	10.61
	(e.t.)	7.73	12.58	14.55	17.77	6.24	10.82	16.01	17.74	7.28	11.07	13.39	16.00
GCDT	(PRE)	23.24	**11.12**	**6.52**	**4.39**	26.77	21.04	**6.01**	**6.00**	25.87	**14.96**	**7.59**	**5.60**
	(e.t.)	7.73	12.58	14.55	17.77	6.25	10.82	16.01	17.74	7.29	11.07	13.40	16.01

Table 5.4.5: Results of the experiments for PH4, PH5 and PH6 images, employing three distinct reconstruction methods. The symbol d denotes the number of projections, and the most outstanding performance is highlighted in bold font.

		PH4				PH5				PH6			
d		6	9	12	15	6	9	12	15	6	9	12	15
GCDT	(PE)	1976	**804**	**551**	**399**	**219**	**134**	**42**	**28**	727	**473**	**251**	**192**
	(m.r. %)	12.06	**4.91**	**3.36**	**2.44**	**1.34**	**0.82**	**0.26**	**0.17**	4.44	**2.89**	**1.53**	**1.17**
TRDT	(PE)	2435	1415	1188	998	1364	1330	1286	1274	889	807	587	552
	(m.r. %)	14.86	8.64	7.25	6.09	8.32	8.12	7.85	7.78	5.43	4.92	3.58	3.37
DART	(PE)	**1695**	1242	1177	1089	488	379	288	319	**649**	836	596	707
	(m.r. %)	**10.34**	7.58	7.18	6.65	2.98	2.31	1.76	1.95	**3.96**	5.10	3.64	4.32

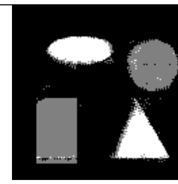
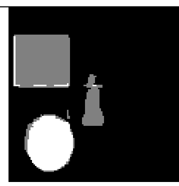
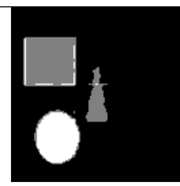
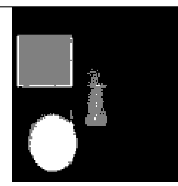
 PE=272 (1.66%)	 PE=255 (1.55%)	 PE=412 (2.51%)
PH1		
 PE=225 (1.37%)	 PE=209 (1.28%)	 PE=143 (0.87%)
PH2		
 PE=367 (2.24%)	 PE=655 (3.99%)	 PE=519 (3.16%)
PH3		
GCDT	MWPDT	TRDT

Fig. 5.4.20: Reconstructions of the 3 gray level test images using data from 6 projection directions.

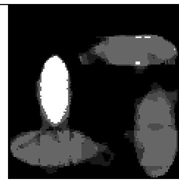
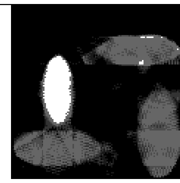
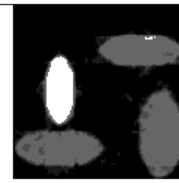
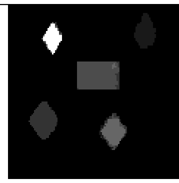
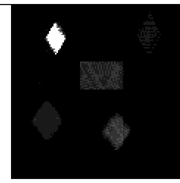
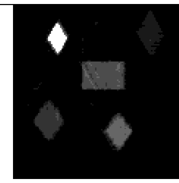
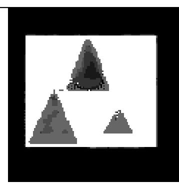
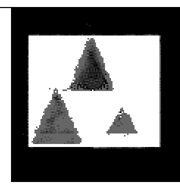
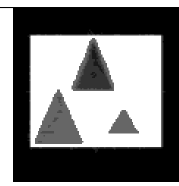
<table>
<tr><td>
PE=1976 (12.06%)</td><td>
PE=2435 (14.86%)</td><td>
PE=1695 (10.34%)</td></tr>
<tr><td colspan="3">PH4</td></tr>
<tr><td>
PE=219 (1.34%)</td><td>
PE=1364 (8.33%)</td><td>
PE=488 (2.98%)</td></tr>
<tr><td colspan="3">PH5</td></tr>
<tr><td>
PE=727 (4.44%)</td><td>
PE=889 (5.43%)</td><td>
PE=649 (3.96%)</td></tr>
<tr><td colspan="3">PH6</td></tr>
<tr><td>GCDT</td><td>TRDT</td><td>DART</td></tr>
</table>

Fig. 5.4.21: Reconstructions of the 6 gray level test images using data from 6 projection directions.

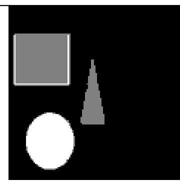
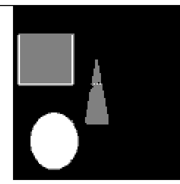
		
PE=5 (0.03%)	PE=35 (0.21%)	PE=28 (0.17%)
PH1		
		
PE=12 (0.07%)	PE=17 (0.10%)	PE=18 (0.11%)
PH2		
		
PE=9 (0.05%)	PE=174 (1.06%)	PE=41 (0.25%)
PH3		
GCDT	MWPDT	TRDT

Fig. 5.4.22: Reconstructions of the 3 gray level test images using data from 15 projection directions.

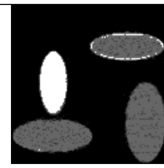
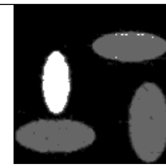
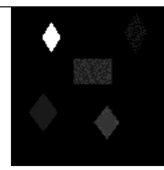
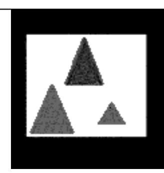
 PE=399 (2.44%)	 PE=998 (6.09%)	 PE=1089 (6.65%)
PH4		
 PE=28 (0.17%)	 PE=1274 (7.78%)	 PE=319 (1.95%)
PH5		
 PE=192 (1.17%)	 PE=552 (3.37%)	 PE=707 (4.32%)
PH6		
GCDT	TRDT	DART

Fig. 5.4.23: Reconstructions of the 6 gray levels test images using data from 15 projection directions.

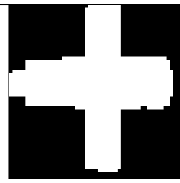
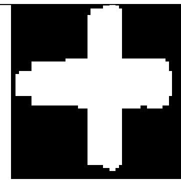
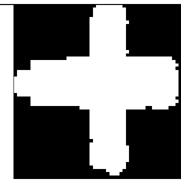
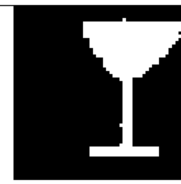
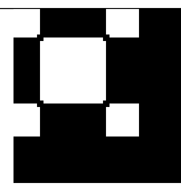
		
PE=49 (1.20%)	PE=44 (1.07%)	PE=45 (1.10%)
PH7		
		
PE=66 (1.61%)	PE= 67 (1.64%)	PE=45 (1.10%)
PH8		
		
PE=800 (19.53%)	PE=800 (19.53%)	PE=604 (14.75%)
PH9		
GCDT	TRDT	DART

Fig. 5.4.24: Reconstructions of the binary test images (PH7, PH8, PH9) using data from 2 projection directions (vertical and horizontal).

Our investigations on binary images, as illustrated in Figure 5.4.24 and Figure 5.4.25, unveiled that the GCDT method yields unsatisfactory outcomes, producing in some cases an entirely black image, particularly when reconstructing from only two projections. To overcome this limitation, we propose enhancing the GCDT method by incorporating orientation and circularity as prior information, resulting in the formulation of the GCORIENTBT and GCCIRCBT algorithms.

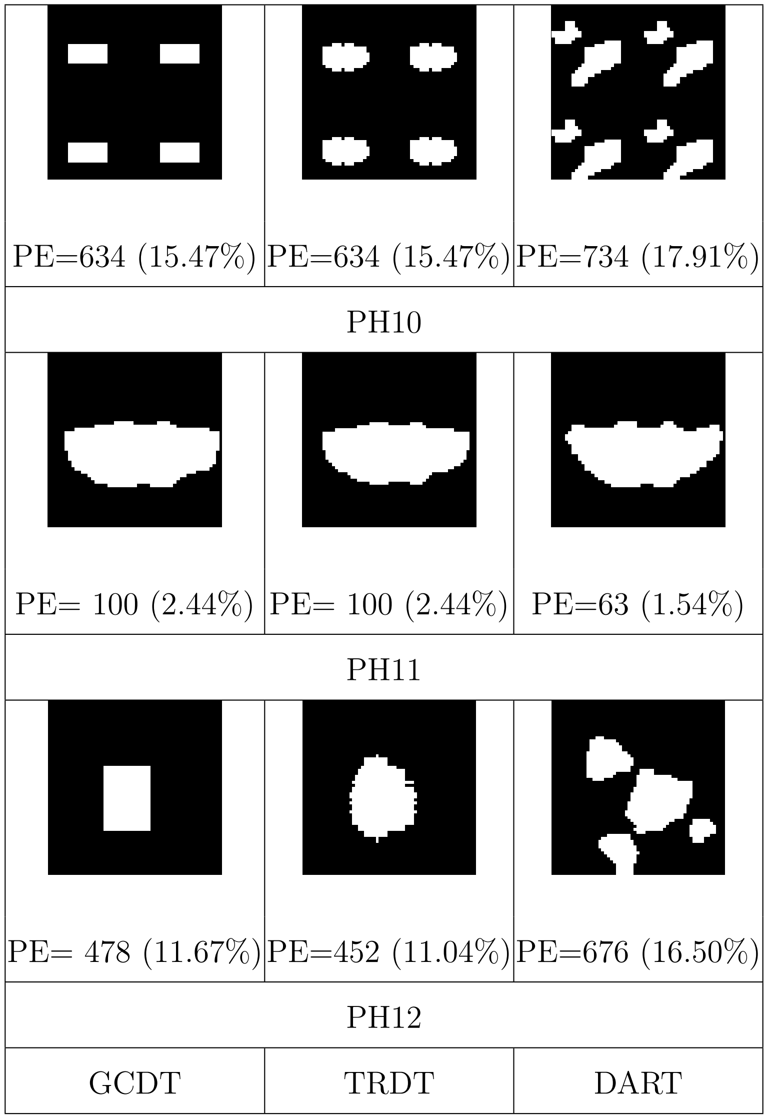


Fig. 5.4.25: Reconstructions of the binary test images (PH10, PH11, PH12) using data from 2 projection directions (vertical and horizontal).

We proceeded to compare these algorithms with two other reconstruction methods, namely DART and GCDT, as depicted in Figure 5.4.32 and Figure 5.4.33. For our analysis, we utilized 6 binary images with data obtained from a single projection and tested the models across 6 different projection angles.

In our prior study [66], GCORIENTBT demonstrated outstanding results in scenarios with limited projection view availability. In a subsequent investigation [67], we explored whether circularity is as effective or potentially superior as a regularization term. The results indicate that in 17 out of 36 cases, GCCIRCBT produced the best reconstruction (smallest PE/m.r.), while GCORIENTBT prevailed in 13 cases. As anticipated, incorporating the prior information into the GCDT method led to significantly improved results for binary images, particularly when dealing with limited projection data. Moreover, GCCIRCBT exhibited a noteworthy advantage in terms of running time, being, on average, 2.12 times faster than its top competitor, GCORIENTBT (Figure 5.4.34).

Summarizing the outcomes from the analysis of reconstruction tasks involving multiple gray levels, as detailed in Tables 5.4.3 and 5.4.5, the GCDT method showcased superior reconstruction quality in 20 out of 24 instances, accounting for 83% of the cases. Additionally, the combined performance of GCORIENTBT and GCCIRCBT proved to be superior in 83% of the cases, with GCCIRCBT holding a slight advantage. This underscores the commendable efficacy of the graph cuts-based reconstruction approach in DT, emphasizing the notable benefits of incorporating shape circularity and orientation as a priori information.

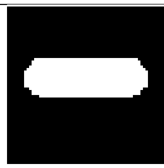
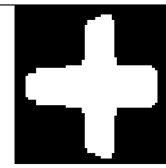
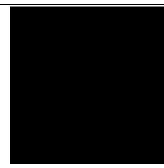
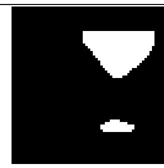
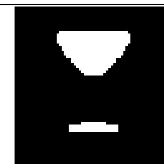
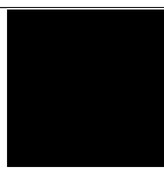
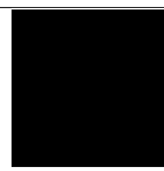
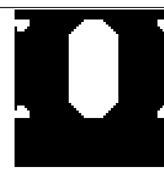
			
PE=444 (10.83%)	PE=682 (16.655%)	PE=682 (16.65%)	PE=252 (6.15%)
PH7			
			
PE=509(12.43%)	PE=509 (12.43%)	PE=178 (4.35%)	PE=469(11.45%)
PH8			
			
PE=800 (19.53%)	PE=800 (19.53%)	PE=800 (19.53%)	PE=886 (21.63%)
PH9			
DART	GCDT	GCORIENTBT	GCCIRCBT

Fig. 5.4.26: Reconstructions of the binary test images (PH7, PH8, PH9) using data from 1 projection direction, $\alpha = 0°$.

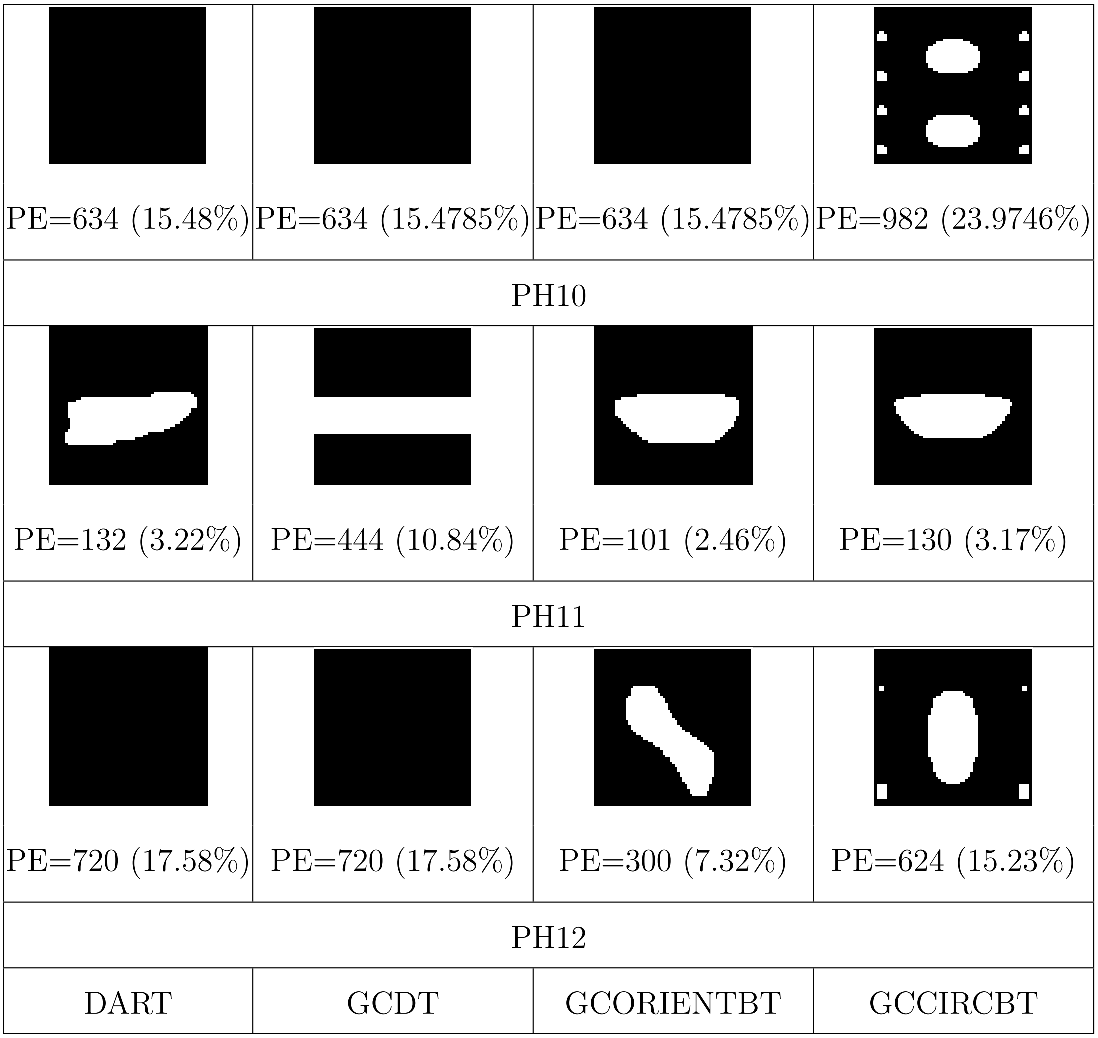


Fig. 5.4.27: Reconstructions of the binary test images (PH10, PH11, PH12) using data from 1 projection direction, $\alpha = 0°$.

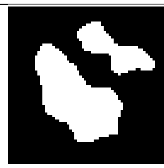
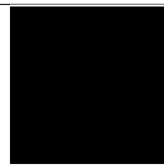
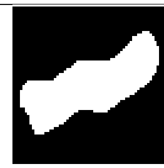
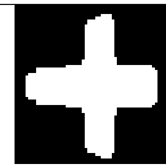
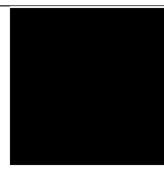
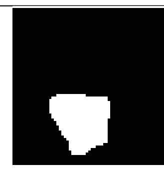
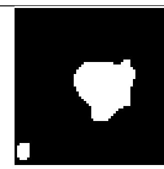
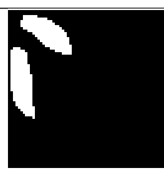
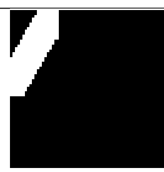
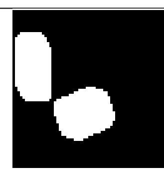
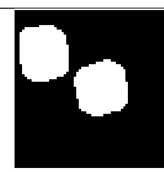
			
PE=846 (20.65%)	PE=1110 (27.10%)	PE=954 (23.29%)	PE=647 (15.8%)
PH7			
			
PE=509(12.43%)	PE=509 (12.43%)	PE=891 (21.75%)	PE=757 (18.48%)
PH8			
			
PE=810 (19.78%)	PE=724 (17.68%)	PE=870 (21.24%)	PE=334 (8.15%)
PH9			
DART	GCDT	GCORIENTBT	GCCIRCBT

Fig. 5.4.28: Reconstructions of the binary test images (PH7, PH8, PH9) using data from 1 projection direction, direction angle is $\alpha = 60°$.

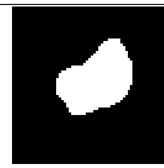
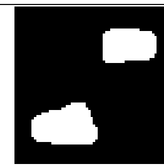
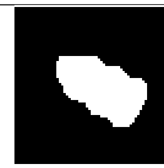
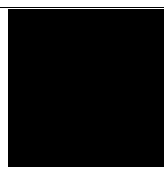
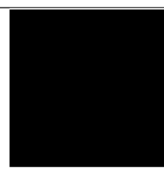
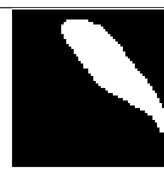
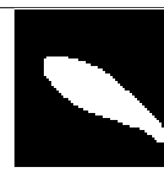
			
PE=634 (15.48%)	PE=634 (15.48%)	PE=870 (21.24%)	PE=372 (9.08%)
PH10			
			
PE=710 (17.33%)	PE= 710 (17.33%)	PE= 710 (17.33%)	PE=423 (10.32%)
PH11			
			
PE=720 (17.58%)	PE=720 (17.58%)	PE=1353 (33.03%)	PE=519 (12.67%)
PH12			
DART	GCDT	GCORIENTBT	GCCIRCBT

Fig. 5.4.29: Reconstructions of the binary test images (PH10, PH11, PH12) using data from 1 projection direction, direction angle is $\alpha = 60°$.

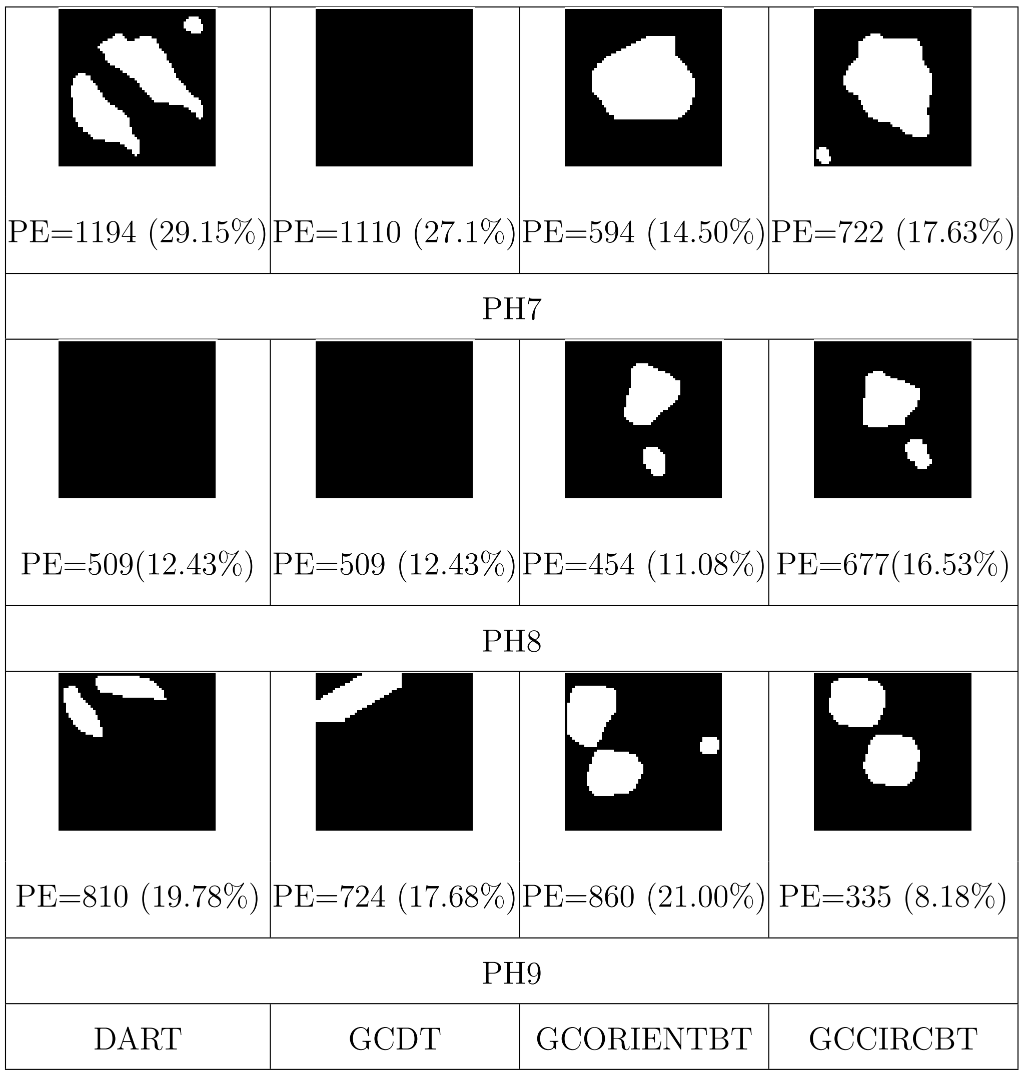


Fig. 5.4.30: Reconstructions of the binary test images (PH7, PH8, PH9) using data from 1 projection direction, direction angle is $\alpha = 30°$.

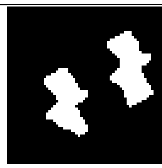
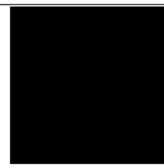
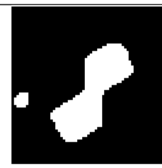
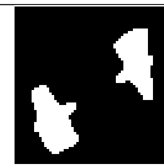
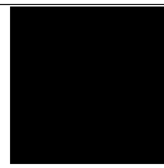
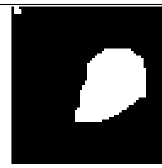
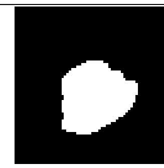
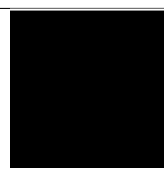
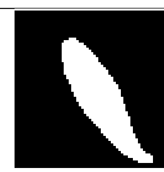
 PE=744 (18.16%)	 PE=634 (15.48%)	 PE=661 (16.14%)	 PE=389 (9.50%)
PH10			
 PE=710 (17.33%)	 PE=710 (17.33%)	 PE=552 (13.48%)	 PE=377 (9.20%)
PH11			
 PE=720 (17.58%)	 PE=720 (17.58%)	 PE=1426 (34.81%)	 PE=365 (8.91%)
PH12			
DART	GCDT	GCORIENTBT	GCCIRCBT

Fig. 5.4.31: Reconstructions of the binary test images (PH10, PH11, PH12) using data from 1 projection direction, direction angle is $\alpha = 30°$.

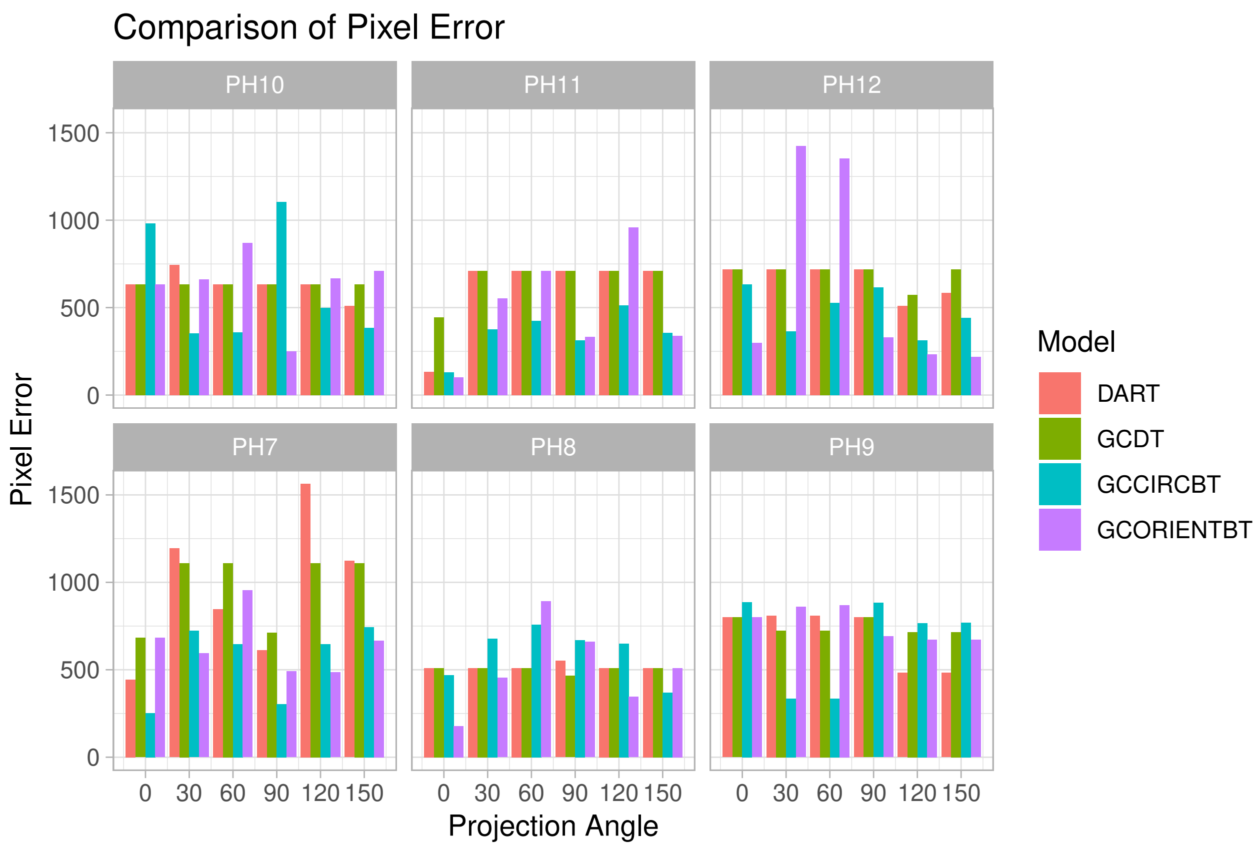


Fig. 5.4.32: Experimental outcomes for BT employing four distinct reconstruction methods, focusing on Pixel Error.

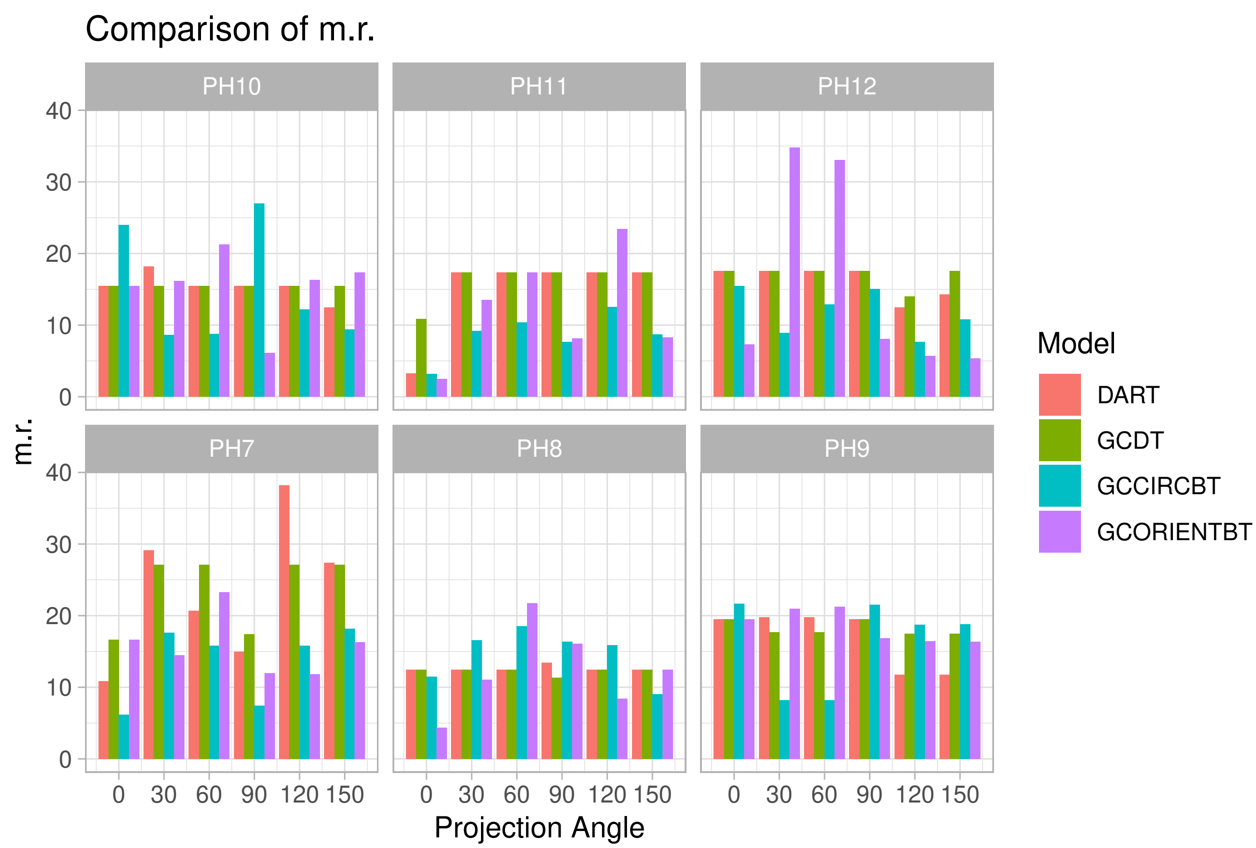


Fig. 5.4.33: Experimental outcomes for BT with an emphasis on misclassification rate (m.r)., employing four diverse reconstruction methods.

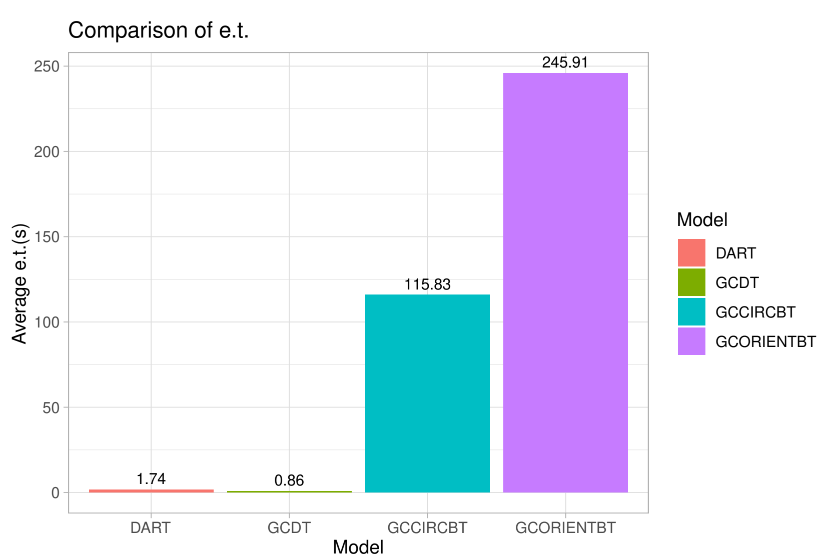


Fig. 5.4.34: Experimental outcomes for BT with an emphasis on average elapsed time (e.t.)., employing four diverse reconstruction methods.

CHAPTER 6

Conclusion and Future Work

In this thesis, we have conducted an in-depth investigation into the field of discrete tomography, a discipline primarily concerned with reconstructing discrete objects or images from a limited set of projections. Our exploration has encompassed various aspects of discrete tomography, including its foundational formulation, binary tomography models, and a range of reconstruction algorithms. Throughout our research, our primary objective has been the development and evaluation of regularized models tailored for tomographic image reconstruction, specifically addressing the challenges associated with generating high-quality images from sparse projection data.

Within the thesis, we emphasize the significant performance of an algorithm that combines gradient-based techniques with graph cuts optimization for solving discrete tomography problems. In scenarios where the availability of projections is severely constrained, we have adapted this method by incorporating shape descriptors as prior knowledge about the objects being reconstructed. Our experimental results conclusively demonstrate that the proposed approach outperforms previously published reconstruction methods regarding reconstruction quality. These findings establish that the integration of a gradient-based method with graph cuts optimization, enhanced by the inclusion of area-based shape descriptors as prior information, represents an effective strategy for achieving high-quality reconstructions in the context of discrete tomography.

While this thesis has made significant progress in exploring regularized models for tomographic image reconstruction, several avenues for future research remain open. We identify the following areas as promising directions for extending and refining our work:

Multi-Modal Shape Regularization: Explore the integration of multiple area-based shape descriptors as regularization terms simultaneously. Combining shape orientation and circulation with other descriptors like eccentricity, roundness, or elongation could provide a more comprehensive and robust regularization framework, capturing different aspects of the object's shape and improving the overall reconstruction accuracy.

Adaptive Regularization Weighting: Investigate adaptive approaches for adjusting the weights or importance of the shape descriptors during the reconstruction process. Some regions of the tomographic data might benefit more from specific shape descriptors than others. Developing adaptive regularization schemes that automatically adjust the weights based on the local image characteristics can lead to more effective regularization and enhanced reconstruction results.

Parameter Tuning for Shape Regularization: Conduct a thorough study on the impact of the parameters associated with shape orientation and circulation regularization. Fine-tuning these parameters can significantly influence the regularization's effectiveness in preserving shapes and structures of the object. Utilize techniques like cross-validation or optimization algorithms to determine the optimal parameter values for specific imaging scenarios.

Deep Learning-Based Regularization: As deep learning approaches have shown remarkable success in various imaging tasks, integrating deep neural networks into our regularization framework holds great promise. Investigating the use of convolutional neural networks (CNNs) and generative models for regularization can potentially improve the accuracy and efficiency of tomographic image reconstruction.

By pursuing these future research directions, we believe that the field of tomographic image reconstruction can benefit from more robust, accurate, and efficient methods, enabling a wide range of applications in various scientific and industrial domains.

In conclusion, this thesis has provided a comprehensive overview of discrete tomography, its formulation, binary tomography models, and reconstruction algorithms.

The research conducted here contributes to the growing body of knowledge in discrete tomography and offers a foundation for future advancements in this field. The implications and limitations discussed in this chapter highlight promising directions for future research, paving the way for improved reconstruction techniques and broader applications of discrete tomography in various domains.

REFERENCES

[1] Aarts, E., Korst, J.: Simulated annealing and Boltzmann machines: a stochastic approach to combinatorial optimization and neural computing. John Wiley & Sons, Inc. (1989)

[2] Andersen, A.H., Kak, A.C.: Simultaneous algebraic reconstruction technique (sart): a superior implementation of the art algorithm. Ultrasonic imaging 6(1), 81–94 (1984)

[3] Aster, R.C., Borchers, B., Thurber, C.H.: Parameter estimation and inverse problems, vol. 90. Academic Press (2011)

[4] Barzilai, J., Borwein, J.M.: Two point step size gradient methods. IMA Journal of Numerical Analysis 8, 141–148 (1988)

[5] Batenburg, K.J., Sijbers, J.: DART: A Fast Heuristic Algebraic Reconstruction Algorithm for Discrete Tomography. In: Proceedings of International Conference on Image Processing (ICIP). pp. 133–136 (2007)

[6] Besson, G.: CT image reconstruction from fan-parallel data. Medical Physics 26(3), 415–426 (1999)

[7] Betancourt, M.: A conceptual introduction to hamiltonian monte carlo. arXiv preprint arXiv:1701.02434 (2017)

[8] Birchfield, S., Tomasi, C.: Multiway cut for stereo and motion with slanted surfaces. Proc. International Conf. Computer Vision pp. 489–495 (1999)

[9] Birgin, E.G., Martínez, J.M., Raydan, M.: Algorithm: 813: SPG - software for convex-constrained optimization. ACM Transactions on Mathematical Software 27, 340–349 (2001)

[10] Birgin, E., Chambouleyron, I., Martínez, J.: Estimation of Optical Constants of Thin Films Using Unconstrained Optimization. J. of Computational Physics 151, 862–880 (1999)

[11] Birgin, E., Martínez, J.: Spectral Conjugate Gradient Method for Unconstrained Optimization. Appl. Math. Optimization 43, 117–128 (2001)

[12] Boykov, Y., Jolly, M.P.: Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images. In: Proc. International Conf. Computer Vision. pp. 105–112 (2001)

[13] Boykov, Y., Kolmogorov, V.: Computing Geodesics and Minimal Surfaces via Graph Cuts. In: Proc. International Conf. Computer Vision. pp. 26–33 (2003)

[14] Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Transactions on PAMI 26(9), 1124–1137 (2004)

[15] Boykov, Y., Veksler, O., Zabih, R.: Markov random fields with efficient approximations. In: Proceedings of IEEE Conf. Computer Vision and Pattern Recognition. pp. 648–655 (1998)

[16] Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Transactions on PAMI 23(11), 1222–1239 (2001)

[17] Bracewell, R.N.: Strip integration in radioastronomy. Australian Journal of Physics 9, 198–217 (1956)

[18] Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). vol. 2, pp. 60–65. Ieee (2005)

[19] Butterworth, S.: On the theory of filter amplifiers. Wireless Engineer 7, 536–541 (1930)

[20] Chambolle, A., Pock, T.: An introduction to continuous optimization for imaging. Acta Numerica 25, 161–319 (2016)

[21] Chang, S.K., Chow, C.: The reconstruction of three-dimensional objects from two orthogonal projections and its application to cardiac cineangiography. IEEE Transactions on Computers 100(1), 18–28 (1973)

[22] Chen, B., Liu, Y., Zhang, Z., Yu, Z., Thompson, R., Sidky, E., Pan, X.: Investigation of non-negativity constraint on basis images in half-rotation data reconstruction in spectral ct. In: 2016 IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD). pp. 1–3 (2016)

[23] Chen, G.H., Tang, J., Leng, S.: Prior image constrained compressed sensing (piccs): a method to accurately reconstruct dynamic ct images from highly undersampled projection data sets. Medical physics 35(2), 660–663 (2008)

[24] Chen, S., Donoho, D.: Basis pursuit. In: Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers. vol. 1, pp. 41–44. IEEE (1994)

[25] Cortadellas, J., Amat, J., De la Torre, F.: Robust normalization of silhouettes for recognition applications. Pattern Recognition Letters 25(5), 591–601 (2004)

[26] Defrise, M., Kinahan, P., Townsend, D., Michel, C., Sibomana, M., Newport, D.: Exact and approximate rebinning algorithms for 3-D pet data. IEEE Transactions on Medical Imaging 16(2), 145–158 (1997)

[27] Delong, A., Osokin, A., Isack, H.N., Boykov, Y.: Fast approximate energy minimization with label costs. Proc. IEEE Conf. Computer Vision and Pattern Recognition 96(1), 1–27 (2010)

[28] Dirac, P.A.M.: The Principles of Quantum Mechanics. Clarendon Press (1950)

[29] Durmus, A., Moulines, E., Pereyra, M.: Efficient bayesian computation by proximal markov chain monte carlo: when langevin meets moreau. SIAM Journal on Imaging Sciences 11(1), 473–506 (2018)

[30] Fiddy, M.: The radon transform and some of its applications. Optica Acta: International Journal of Optics 32(1), 3–4 (1985)

[31] Flusser, J., Suk, T.: Rotation moment invariants for recognition of symmetric objects. IEEE Transactions on Image Processing 15(12), 3784–3790 (2006)

[32] Ford Jr, L.R., Fulkerson, D.R.: Maximal flow through a network. Canadian Journal of Mathematics 8(3), 399–404 (1956)

[33] Fourier, J.B.J.: Théorie Analytique de la Chaleur. Firmin Didot, Père et Fils (1822)

[34] Girolami, M., Calderhead, B.: Riemann manifold langevin and hamiltonian monte carlo methods. Journal of the Royal Statistical Society Series B: Statistical Methodology 73(2), 123–214 (2011)

[35] Glover, F., Laguna, M.: Tabu search. Springer (1998)

[36] Greig, D., Porteous, B., Seheult, A.: Exact maximum a posteriori estimation for binary images. Journal of the Royal Statistical Society 51(2), 271–279 (1989)

[37] Grippo, L., Lampariello, F., Lucidi, S.: A Nonmonotone Line Search Technique for Newton's Method. SIAM J. Numer. Anal. 23, 707–716 (1986)

[38] Gritzmann, P., Prangenberg, D., de Vries, S., Wiegelmann, M.: Success and failure of certain reconstruction and uniqueness algorithms in discrete tomography. International journal of imaging systems and technology 9(2-3), 101–109 (1998)

[39] Ha, V.H., Moura, J.M.: Affine-permutation invariance of 2-d shapes. IEEE Transactions on Image Processing 14(11), 1687–1700 (2005)

[40] Hamming, R.W.: Digital filters. Courier Corporation (1998)

[41] Herman, G.T., Kuba, A.: Discrete Tomography: Foundations, Algorithms and Applications. Birkhäuser (1999)

[42] Herman, G.T., Kuba, A.: Advances in Discrete Tomography and Its Applications. Birkhäuser (2006)

[43] Hu, M.K.: Visual pattern recognition by moment invariants. IRE transactions on information theory 8(2), 179–187 (1962)

[44] Huang, K., Gao, Z., Yang, F., Zhang, H., Zhang, D.: An improved discrete algebraic reconstruction technique for limited-view based on gray mean value guidance. Journal of Nondestructive Evaluation 42(1), 6 (2023)

[45] Hughes, S.: Ct scanning in archaeology. Computed Tomography-Special Applications/Ed. L. Saba. InTech Europe pp. 57–70 (2011)

[46] Jain, A.K., Vailaya, A.: Shape-based retrieval: A case study with trademark image databases. Pattern recognition 31(9), 1369–1390 (1998)

[47] Kadu, A., van Leeuwen, T.: A convex formulation for binary tomography. IEEE Transactions on Computational Imaging 6, 1–11 (2020)

[48] Kalare, K.W., Bajpai, M.K.: Recdnn: deep neural network for image reconstruction from limited view projection data. Soft Computing 24(22), 17205–17220 (2020)

[49] Kim, J., Zabih, R.: Automatic Segmentation of Contrast- Enhanced Image Sequences. In: Proc. International Conf. Computer Vision. pp. 502–509 (2003)

[50] Kirkpatrick, S., Gelatt, C., Vecchi, M.: Optimization by Simulated Annealing. Science 220(4598), 671–680 (1983)

[51] Kofler, A., Haltmeier, M., Schaeffter, T., Kachelrieß, M., Dewey, M., Wald, C., Kolbitsch, C.: Neural networks-based regularization for large-scale medical image reconstruction. Physics in Medicine & Biology 65(13), 135003 (2020)

[52] Kolmogorov, V., Zabih, R.: Visual Correspondence with Occlusions Using Graph Cuts. In: Proc. International Conf. Computer Vision. pp. 508–515 (2001)

[53] Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts? IEEE Transactions on PAMI 26(2), 147–159 (2004)

[54] Kuba, A., Herman, G.T.: Discrete tomography: A historical overview. In: Discrete tomography, pp. 3–34. Springer (1999)

[55] Kwatra, V., Schoedl, A., Essa, I., Turk, G., Bobick, A.: Graphcut Textures: Image and Video Synthesis Using Graph Cuts. In: Proc. SIGGRAPH 2003. pp. 277–286. ACM Trans. Graphics (2003)

[56] Lézoray, O., Grady, L.: Image processing and analysis with graphs: theory and practice. CRC Press (2017)

[57] Lukić, T.: Discrete Tomography Reconstruction Based on the Multi-well Potential. In: Proceedings of Combinatorial Image Analysis - 14th International Workshop (IWCIA). LNCS, vol. 6636, pp. 335–345. Springer-Verlag, Madrid, Spain (2011)

[58] Lukić, T., Balázs, P.: Binary tomography reconstruction based on shape orientation. Pattern Recognition Letters 79, 18–24 (2016)

[59] Lukić, T., Marčeta, M.: Gradient and Graph Cuts Based Method for Multi-level Discrete Tomography. In: Proceedings of Combinatorial Image Analysis - 18th International Workshop (IWCIA). pp. 322–333. LNCS, Springer-Verlag, Plovdiv, Bulgaria (2017)

[60] Lukić, T., Nagy, B.: Energy-minimization based Discrete Tomography Reconstruction Method for Images on Triangular Grid. In: Proceedings of Combinatorial Image Analysis - 15th International Workshop (IWCIA). LNCS, vol. 7655, pp. 274–284. Springer-Verlag, Austin (TX), USA (2012)

[61] Lukić, T., Nagy, B.: Deterministic discrete tomography reconstruction method for images on triangular grid. Pattern Recognition Letters 49, 11–16 (2014)

[62] Lukić, T., Balázs, P.: Shape circularity assisted tomography reconstruction. Physica Scripta 95(10), 105211 (2020)

[63] Lukić, T., Balázs, P.: Limited-view binary tomography reconstruction assisted by shape centroid. The Visual Computer 38(2), 695–705 (2022)

[64] Lukić, T., Nagy, B.: Regularized binary tomography on the hexagonal grid. Physica Scripta 94(2), 025201 (2019)

[65] Lyra, M., Ploussi, A.: Filtering in spect image reconstruction. Journal of Biomedical Imaging 2011, 1–14 (2011)

[66] Marčeta, M., Lukić, T.: Graph cuts based tomography enhanced by shape orientation. In: Lukić, T., Barneva, R.P., Brimkov, V.E., Comic, L., Sladoje, N. (eds.) Combinatorial Image Analysis - 20th International Workshop, IWCIA 2020, Novi Sad, Serbia, July 16-18, 2020, Proceedings. Lecture Notes in Computer Science, vol. 12148, pp. 219–235. Springer (2020), `https://doi.org/10.1007/978-3-030-51002-2\_16`

[67] Marčeta, M., Lukić, T.: Regularized graph cuts based discrete tomography reconstruction methods. Journal of Combinatorial Optimization 44(4), 2324–2346 (2022)

[68] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E.: Equation of State Calculations by Fast Computing Machines. Journal of Chemical Physics 21(6), 1087–1092 (1953)

[69] Minh, D.D., Minh, D.L.: Understanding the hastings algorithm. Communications in Statistics-Simulation and Computation 44(2), 332–349 (2015)

[70] Natterer, F., Wübbeling, F.: Mathematical methods in image reconstruction. SIAM (2001)

[71] Neal, R.M.: Slice sampling. The annals of statistics 31(3), 705–767 (2003)

[72] Noori-Asl, M., Sadremomtaz, A.: Analytical image reconstruction methods in emission tomography. Journal of Biomedical Science and Engineering 06, 100–107 (01 2013)

[73] Paton, R.C., Back, T., Fogel, D., Michalewicz, T.: Principles of genetics. In: Evolutionary Computation 1: basic algorithms and operators, pp. 27–38. Institute of Physics Publishing, Bristol and Philadelphia (2000)

[74] Prause, G.P., Onnasch, D.G.: Binary reconstruction of the heart chambers from biplane angiographic image sequences. IEEE Transactions on Medical Imaging 15(4), 532–546 (1996)

[75] Proffitt, D.: The measurement of circularity and ellipticity on a digital grid. Pattern Recognition 15(5), 383–387 (1982)

[76] Radon, J.: Über die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten. Berichte über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse 69, 262–277 (1917)

[77] Ramachandran, G., Lakshminarayanan, A.: Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of fourier transforms. Proceedings of the National Academy of Sciences 68(9), 2236–2240 (1971)

[78] Raydan, M.: The Barzilai and Browein Gradient Method for the Large Scale Unconstrained Minimization Problem. SIAM J. of Optimization 7, 26–33 (1997)

[79] Re, A., Corsi, J., Demmelbauer, M., Martini, M., Mila, G., Ricci, C.: X-ray tomography of a soil block: a useful tool for the restoration of archaeological finds. Heritage Science 3(1), 1–7 (2015)

[80] Rosin, P.L.: Measuring shape: ellipticity, rectangularity, and triangularity. Machine Vision and Applications 14(3), 172–184 (2003)

[81] Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D: nonlinear phenomena 60(1-4), 259–268 (1992)

[82] Ryser, H.J.: Combinatorial properties of matrices of zeros and ones. Canadian Journal of Mathematics 9, 371–377 (1957)

[83] Saha, S.K., Tahtali, M., Lambert, A., Pickering, M.: Sparsity prior computed tomography reconstruction using a nonstandard simultaneous x-ray acquisition model. Journal of Medical Imaging and Radiation Sciences 47(3), 251–266 (2016)

[84] Schüle, T., Schnörr, C., Weber, S., Hornegger, J.: Discrete tomography by convex-concave regularization and D.C. Programming. Discrete Appl. Math. 151, 229–243 (2005)

[85] Schüle, T., Schnörr, C., Weber, S., Hornegger, J.: Discrete Tomography by Convex-concave Regularization and D.C. Programming. Discrete Appl. Math. 151, 229–243 (2005)

[86] Shen, C., Gonzalez, Y., Chen, L., Jiang, S.B., Jia, X.: Intelligent parameter tuning in optimization-based iterative ct reconstruction via deep reinforcement learning. IEEE transactions on medical imaging 37(6), 1430–1439 (2018)

[87] Shepp, L.A., Logan, B.F.: The fourier reconstruction of a head section. IEEE Transactions on Nuclear Science 21(3), 21–43 (1974)

[88] Shikhaliev, P.M.: Computed tomography imaging of large cargos. Applied Physics Initiative, LLC, Baton Rouge, LA pp. 1–13 (2017)

[89] Snow, D., Viola, P., Zabih, R.: Exact voxel occupacy with graph cuts. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. vol. 1, pp. 345–352 (2000)

[90] Sonka, M., Hlavac, V., Boyle, R.: Image processing, analysis, and machine vision. Thomosn-Engineering (2007)

[91] Tikhonov, A.N.: On the solution of ill-posed problems and the method of regularization. In: Doklady akademii nauk. vol. 151, pp. 501–504. Russian Academy of Sciences (1963)

[92] Varga, L., Balázs, P., Nagy, A.: An energy minimization reconstruction algorithm for multivalued discrete tomography. In: Proceedings of 3rd International Symposium on Computational Modeling of Objects Represented in Images. pp. 179–185. Taylor & Francis, Rome, Italy (2012)

[93] Weber, S., Nagy, A., Schüle, T., Schnörr, C., Kuba, A.: A benchmark evaluation of large-scale optimization approaches to binary tomography. In: Proc. of 13th International Conference on Discrete Geometry for Computer Imagery (DGCI). LNCS, vol. 4245, pp. 146–156. Springer-Verlag, Szeged, Hungary (2006)

[94] Weber, S., Schnorr, C., Hornegger, J.: A linear programming relaxation for binary tomography with smoothness priors. Electronic Notes in Discrete Mathematics 12, 243–254 (2003)

[95] Yuan, J., Bae, E., Tai, X.C., Boykov, Y.: A continuous max-flow approach to potts model. In: Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part VI 11. pp. 379–392. Springer (2010)

[96] Zabrodsky, H., Peleg, S., Avnir, D.: Symmetry as a continuous feature. IEEE Transactions on Pattern Analysis and Machine Intelligence 17(12), 1154–1166 (1995)

[97] Zeng, G.L.: Revisit of the ramp filter. In: 2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). pp. 1–6. IEEE (2014)

[98] Zernike, F.: Beugungstheorie des schneidenver-fahrens und seiner verbesserten

form, der phasenkontrastmethode. Physica 1(7), 689–704 (1934), `https://www.sciencedirect.com/science/article/pii/S0031891434802595`

[99] Zisler, M., Petra, S., Schnörr, C., Schnörr, C.: Discrete Tomography by Continuous Multilabeling Subject to Projection Constraints. In: Proc. of 38th German Conference on Pattern Recognition. pp. 261–272. LNCS, Springer-Verlag (2016)

[100] Žunić, J., Hirota, K., Rosin, P.L.: A hu moment invariant as a shape circularity measure. Pattern Recognition 43(1), 47–57 (2010)

Biografija

Marina Šulc (Marčeta) rođena je 09.09.1990. u Bihaću. Osnovnu školu „Miroslav Antić“završila je u Futogu, 2005. godine kao đak generacije. Iste godine je upisala Gimnaziju "Svetozar Marković" u Novom Sadu, koju završava 2009. godine kao nosilac Vukove diplome. Po završetku gimnazije, 2009. godine, upisala je osnovne akademske studije na Prirodno - matematičkom fakultetu u Novom Sadu, smer primenjena matematika, matematika finansija, koje je završila u junu 2012. godine, sa prosečnom ocenom 9.96, a zatim i master studije na istom smeru sa prosečom ocenom 9.87. Godine 2016. upisuje doktorske studije na Fakultetu tehničkih nauka Univerziteta u Novom Sadu, smer matematika u tehnici. Položila je sve ispite predviđene nastavnim planom i programom doktorskih studija septembra 2018. godine, sa prosečnom ocenom 10.00.

Bila je stipendista nemačke službe za internacionalnu razmenu DAAD. Kao stipendista nemačke vlade obavljala je praksu u sedištu Deutsche Bank u Frankfurtu. Biši je stipendista Fonda za mlade talente Republike Srbije, kao i pitomac Srpskog privrednog društva „Privrednik“. U periodu od septembra 2014. do oktobra 2019. bila je zaposlena kao saradnik u nastavi, odnosno asistent, na katedri za matematiku Fakulteta tehničkih nauka u Novom Sadu gde je 2018. godine bila dobitnik nagrade iz fonda „Momčilo Momo Novković“ za najboljeg asistenta. 2019. godine se seli u Beč gde i danas živi i radi na poziciji Senior Quantitative Analyst i vođe tima u kompaniji Entain Austria GmbH.

Biography

Marina Šulc (née Marčeta) was born on September 9, 1990, in Bihać. She completed her primary school education at *Miroslav Antić* School in Futog in 2005, graduating as the top student of her class. In the same year, she enrolled in *Svetozar Marković* Gymnasium in Novi Sad, where she graduated in 2009 as the recipient of the *Vuk Karadžić* Diploma, an award for high-achieving students. Following her gymnasium education in 2009, she pursued her undergraduate studies at the Faculty of Sciences, University of Novi Sad, specializing in Applied Mathematics with a focus on Financial Mathematics. She completed her undergraduate studies in June 2012 with an average grade of 9.96. Subsequently, she pursued her master's studies in the same field, achieving an average grade of 9.87.

In 2016, she enrolled in the doctoral program at the Faculty of Technical Sciences, University of Novi Sad, specializing in Mathematics in Engineering. Marina successfully completed all required courses for her doctoral studies in September 2018, achieving an average grade of 10.00.

She was a recipient of the DAAD (German Academic Exchange Service) scholarship for international exchange. She conducted an internship at the headquarters of Deutsche Bank in Frankfurt during her time as a German government scholarship holder. Marina was also a scholarship recipient of the Fund for Young Talents of the Republic of Serbia and a member of the Serbian Economic Society "Privrednik."

From September 2014 to October 2019, she worked as a teaching assistant in the Department of Mathematics at the Faculty of Technical Sciences in Novi Sad. In 2018, she was awarded the *Momčilo Momo Novković* fund prize for the best teaching assistant. In 2019, she relocated to Vienna, where she currently resides, and works as a Senior Quantitative Analyst and team lead at Entain Austria GmbH.

APPENDIX A

Publication I

Marčeta, Marina, and Tibor Lukić. ”Regularized graph cuts based discrete tomography reconstruction methods.” Journal of Combinatorial Optimization 44.4 (2022): 2324-2346.

Journal of Combinatorial Optimization
https://doi.org/10.1007/s10878-021-00730-4

Regularized graph cuts based discrete tomography reconstruction methods

Marina Marčeta[1] **· Tibor Lukić**[1]

Accepted: 31 March 2021

Abstract
The topic of this paper includes graph cuts based computed tomography reconstruction methods in binary and multi-gray level cases. This approach combines the graph cuts and a gradient based method. The present paper introduces and analyses the shape circularity as a new regularization and incorporates it in a graph cuts based computed tomography reconstruction approach, thus introducing a new energy-minimization based reconstruction algorithm for binary tomography. Proposed method is capable for reconstructions in cases of limited projection view availability. Results of experimental evaluation of the considered graph cuts type reconstruction methods for both binary and multi-level tomography are presented.

Keywords Discrete tomography · Binary tomography · Shape circularity · Graph cuts optimization · Energy minimization methods

1 Introduction

Image reconstruction represents a collection of methods used to enhance and improve the quality of the image or to extract additional information from the image. Very often we need to obtain information about an object which is not visible or easily accessible. An area of image processing whose scope are these type of problems is named tomography.

Authors acknowledge the project "Innovative scientific and artistic research from the domain of the Faculty of Technical Sciences", grant no. $451-03-68/2020-14/200156$.

Marina Marčeta
marceta.marina@gmail.com

Tibor Lukić
tibor@uns.ac.rs

[1] Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia

Published online: 10 April 2021

Tomography deals with reconstructing images from the given projection data. Projection data is obtained by a wave penetrating through an unknown object. This wave is then detected on the opposite side of an object. A set of projection data is obtained at a particular angle. The source and detector are then rotated at a small angle, and a new projection is obtained. The object aimed to be restored is seen as a function with a domain that can be discrete or continuous and a range that is a given set of (usually) real numbers. *Discrete Tomography* (DT) (Herman and Kuba 1999, 2006) is a field of tomography that focuses on reconstruction of discrete images (finite number of pixel values) using much fewer number of projections. We distinguish binary tomography (BT) concentrating on binary images and multi-level discrete tomography concentrating on digital images that consist of numerous gray levels. DT has a wide range of applications in areas where the materials of the object under investigation are known before, such as industrial non-destructive testing or electron tomography, as well in many diagnostic approaches in medicine (Herman and Kuba 1999, 2006).

In the related conference publication (Marčeta and Lukić 2020), we proposed a new graph cuts based binary tomography reconstruction algorithm (GCORIENTBT) for limited projection availability. This approach incorporates the shape orientation (Žunić et al. 2006) as an a priori information about the solution into the reconstruction process.

This paper brings a new regularization approach based on the shape circularity (Žunić et al. 2010). We use the same graph cuts based optimization approach as in the case of the GCORIENTBT method, but, instead of the shape orientation, the shape circularity is reviewed and applied as an a priori information in the reconstruction process. We found the motivation for this choice in recently published paper (Lukić and Balázs 2020), where the circularity prior is successfully applied in a combination with convex-concave based regularization (Schüle et al. 2005). The proposed circularity based method (GCCIRCBT) has an important advantage compared to GCORIENTBT, since the gradient of the regularization is determined in an analytical way which makes the determination of the smooth solution fast by the SPG algorithm. Running time of the algorithm is significantly decreased compared to existing similar techniques. We demonstrate by experiments that the prior information can boost the performance of reconstruction in cases of very low number of projections. Additionally, this paper gives an overview and experimental evaluation of the most often used algorithms for multi-level tomography reconstruction problem, which, to the best of our knowledge was addressed by only few researches.

This paper is organized as follows. Section 2 gives a brief overview of the basic reconstruction problem. Section 3 begins by examining the approach that uses graph cuts for energy minimization, followed by the introduction of shape orientation and circularity as shape descriptors and finishing with describing and analyzing the new reconstruction method. Experimental results are provided in Sect. 4. Our conclusions are drawn in the final section.

2 Reconstruction problem

In this chapter we introduce some notations and define the DT problem mathematically.

DT reconstruction problem can be outlined by the following linear system of equations

$$
\begin{aligned}
&a_{11}u_1 + a_{12}u_2 + a_{13}u_3 + \ldots + a_{1N}u_N = b_1\\
&a_{21}u_1 + a_{22}u_2 + a_{23}u_3 + \ldots + a_{2N}x_N = b_2\\
&a_{31}u_1 + a_{32}u_2 + a_{33}u_3 + \ldots + a_{3N}x_N = b_3\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{M1}u_1 + a_{M2}u_2 + a_{M3}u_3 + \ldots + a_{MN}x_N = b_M,
\end{aligned}
$$

which we examine in its matrix form

$$A\,u = b, \tag{1}$$

where $A \in \mathbb{R}^{M\times N},\ \ u \in \Lambda^N,\ b \in \mathbb{R}^M$ and $\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_k\}$ for $k \geq 2$.

Unknown image to be reconstructed is denoted by u and is represented in a column vector form. Set Λ is given by the user and denotes the gray levels of the image, if $k = 2$, problem becomes problem of binary tomography. Matrix A is named *projection matrix*. Each row of this matrix is determined by one projection ray and its entries are calculated as the length of the intersection of the pixels and projection ray passing through them. The corresponding components of the vector b consist of the detected projection values calculated as a sum of products of the pixel's intensity and the corresponding length of the projection ray through that pixel.

Projection process is performed from different directions. For each projection direction a number of parallel projection rays is taken (parallel beam projection). The projection direction is determined by the angle α. Every two adjacent parallel projection rays are equidistant with their distance being equal to the side length of pixels. Number of parallel projection rays is specified in a way to cover the whole image grid.

The reconstruction problem means finding the solution image u of the linear system of equations (1), where the projection matrix A and the projection vector b are given. This system is often undetermined ($N > M$). We are interested in finding a reconstruction which resembles the original image as closely as possible, not just one that corresponds to the given projections. Therefore, it is necessary to use all available information (a priori information) about the object of interest in order to determine quality and acceptable solution.

3 Graph cuts reconstruction methods assisted by shape circularity and shape orientation

Image reconstruction is commonly performed by regularized energy minimization, due to its simplicity and generally good performance. In its most general setting one tries to recover a reconstructed version of the observed image y by minimizing an energy function which has the following form:

$$E(u) = F(Lu, b) + \lambda R(u). \tag{2}$$

An argument u^r which minimizes the energy function,

$$u^r = arg\min_u E(u) \tag{3}$$

is considered to be an estimate of the original image. The function F is called the data fidelity term and measures the distance between the data b and the reconstruction u after the forward operator L has acted on it. The function R is called the regularization term and it imposes a priori knowledge on the solution u. It is expected that small values of R will lead, up to a certain extent, to the elimination of the undesirable features. Regularization also provides numerical stabilization of image reconstruction problem. The regularization parameter λ controls the trade-off between the two terms, i.e. the level of smoothing vs. faithful recovery of the image detail.

3.1 Graph cut optimization

Graph cut optimization can be conveniently utilized to solve a wide variety image processing problems that can be formulated in terms of energy minimization (Boykov et al. 1998, 2001; Birchfield and Tomasi 1999; Kolmogorov and Zabih 2001; Kwatra et al. 2003; Boykov and Kolmogorov 2003; Boykov and Jolly 2001; Kim and Zabih 2003).

A directed, weighted graph $G = (X, \rho)$, is determined by a set of nodes X, that are connected together through edges ρ. All the edges are directed from one node to another and appointed some weight or cost. A cut of a graph G is a partition of set X into two disjoint subsets A named source and B named sink. Any cut determines a unique cut-set consisting of a set of edges that have one endpoint in each subset of the partition. Cost of a cut is calculated as a sum of weights of all edges going from A to B. The minimum cut problem consists of finding a cut with minimum cost among all possible cuts. Algorithms to solve this problem can be found in (Boykov and Kolmogorov 2004).

The main idea behind application of graph cuts method in energy minimization is construction of a graph specially designed for the energy function so that the solution of minimum cut problem also minimizes the energy function. The solution of the minimum cut problem, in turn, can be computed very efficiently by max-flow algorithms.

The Potts model in graph cuts theory, on which min-cut/max-flow algorithm is applied, is based on the minimization of the following energy

$$E(d) = \sum_{p \in \mathcal{P}} D(p, d_p) + \sum_{(p,q) \in \mathcal{N}} K_{(p,q)} \cdot (1 - \delta_{d_p, d_q}), \tag{4}$$

where $d = \{d_p | p \in \mathcal{P}\}$ represents the labelling of the image pixels $p \in \mathcal{P}$. By $D(p, d_p)$ we denote the data cost term, where $D(p, d_p)$ is a penalty or cost for assigning a label d_p to a pixel p. $K_{(p,q)}$ is an interaction potential between neighboring pairs p and q, $\mathcal{N}$ is a set of neighboring pairs. Function δ_{d_p, d_q} is a Kronecker delta function.

3.2 Geometric moments

The geometric moment of a digitized image u is defined by

$$m_{p,q}(u) = \sum_{(i,j)\in\Omega} u(i,j) i^p j^q,$$

where $\Omega \subseteq \mathbb{R}^2$ denotes the image domain.

The center of gravity (or centroid) of an image (or a shape) u is defined by

$$(C_x(u), C_y(u)) = \left(\frac{m_{1,0}(u)}{m_{0,0}(u)}, \frac{m_{0,1}(u)}{m_{0,0}(u)} \right).$$

The centroid enables the definition of the centralized moment which is translation invariant. The centralized moment of an image u of order $p+q$ is given by

$$\overline{m}_{p,q}(u) = \sum_{(i,j)\in\Omega} u(i,j)(i - C_x(u))^p (j - C_y(u))^q.$$

The shape is a characteristic of an object which allows numerical characterizations and, in addition, has high object discrimination capacity. Many approaches regarding shape descriptors have been developed (Sonka et al. 2007). There are shape descriptors that accurately describe specific shapes and the ones that describe single characteristics that are present over a variety of shapes. In this paper we will focus on two shape descriptors, namely orientation and circularity, and we will measure them using geometric moments.

3.3 Shape orientation

Shape orientation is determined by the angle α, which represents the slope of the axis of the second moment of inertia (orientation axis) of the considered shape (Sonka et al. 2007). The orientation (angle α) for the the given image u can be calculated by the following equation:

$$\frac{\sin(2\alpha)}{\cos(2\alpha)} = \frac{2 \cdot \bar{m}_{1,1}(u)}{\bar{m}_{2,0}(u) - \bar{m}_{0,2}(u)}. \tag{5}$$

Moments in (5) are translation invariant, making the orientation invariant to translation transformations, for more details see (Lukić and Balázs 2016; Žunić et al. 2006).

The graph cuts reconstruction method which applies the shape orientation in the binary tomography reconstruction process (GCORIENTBT) is proposed and analyzed in our recently published paper (Marčeta and Lukić 2020), therefore we omit its detailed description.

3.4 Shape circularity

Shape circularity is a familiar shape descriptor. Exploiting the fact that the circle has the largest area among all the shapes with the same perimeter, the most standard method defines the shape circularity $C_{st}(S)$ in the following way

$$C_{st}(S) = \frac{4\pi A(S)}{(P(S))^2}, \tag{6}$$

where $A(S)$ is the area of the shape S and $P(S)$ is the perimeter of S. It is easy to notice that $C_{st}(S)$ is not area based nor boundary based, since it uses the information both from the interior and boundary points.
For the given shape, represented by the image u, the circularity can also be rated or measured by the following formula

$$\mathcal{C}(u) = \frac{1}{2\pi(\mu_{2,0}(u) + \mu_{0,2}(u))}, \tag{7}$$

where $\mu_{p,q}(u)$ is the normalized moment of u of order $p + q$. It is defined by the following formula

$$\mu_{p,q}(u) = \frac{\overline{m}_{p,q}(u)}{m_{0,0}(u)^{1+\frac{p+q}{2}}}.$$

It is trivial to show, that the normalized moment, besides being translation invariant, is invariant to uniform scaling as well. The circularity measure $\mathcal{C}(u)$ is proposed and thoroughly analyzed in (Žunić et al. 2010). In addition, in the same paper it has been proven that circularity is highly performant in shape classification problems. The following Theorem summarizes the most important properties of $\mathcal{C}(u)$.

Theorem 1 *(Žunić et al. 2010) The circularity measure $\mathcal{C}(u)$, for a compact shape u (closed and bounded), satisfies:*

(a) $\mathcal{C}(u) \in (0, 1]$, *for all shapes defined by u;*
(b) $\mathcal{C}(u) = 1 \Leftrightarrow u$ *represents a circle;*
(c) $\mathcal{C}(u)$ *is invariant w.r.t. similarity transformations (translation, rotation and scaling);*
(d) For each $\delta > 0$ *there is a shape u such that* $0 < \mathcal{C}(u) < \delta$.

The standard circularity measure $C_{st}(S)$ penalizes deep intrusions into the shape, because such intrusions lead to an essential perimeter increase, which, by the definition, decreases $C_{st}(S)$. The measure $C(S)$ is area based and does not penalize such intrusions. On the other hand, $C(S)$ is robust to noise, as area based descriptor, whereas $C_{st}(S)$ can only cope with small levels of noise because it uses the shape perimeter for the computation. For our model we use $C(S)$ as a measure of circularity.

3.5 The new method based on shape circularity

Reconstruction method for solving discrete tomography problem we propose in this paper consists of two parts:

- Finding a continuous (smooth) solution of the energy minimization problem using gradient based minimization method. Information about the circularity of the original object is incorporated in the energy function.
- Discretization of the obtained smooth solution applying graph cuts based algorithm. Pixel values of the smooth image are used to define data cost term for the graph.

$$\min_{u\in[0,1]^N} E_Q(u) := w_P\|Au-b\|_2^2 + w_H \sum_{i=1}^{N}\sum_{j\in\Upsilon(i)}(u_i-u_j)^2 + \\ + w_C\left(\mathcal{C}(u)-\mathcal{C}^*\right)^2 + \mu\left\langle u,\tau-u\right\rangle, \tag{8}$$

Energy function we use for calculation of the smooth solution is given in the equation (8) and is constructed of four terms:

1. Data fitting term, $\|Au-b\|_2^2$, regularized by parameter $w_P > 0$. Data fitting term ensures adherence to the projection data.
2. Homogeneity term, $\sum_{i=1}^{N}\sum_{j\in\Upsilon(i)}(u_i-u_j)^2$, regularized by parameter $w_H > 0$. $\Upsilon(i)$ is set of indices of two neighbouring pixels (in x and y axis directions) of pixel i. This term ensures the smoothness of the solution.
3. Term, $(\mathcal{C}(u)-\mathcal{C}^*)^2$, measures the distance between the circularity of current solution($\mathcal{C}(u)$) and known circularity of the original image ($\mathcal{C}^*$). Parameter $w_C > 0$ determines the impact of the circularity regularization.
4. Concave regularization term, $\langle u,\tau-u\rangle$, where $\tau=[1,1,\ldots,1]^T$ is a vector of size N, has the role to move pixels intensities toward binary values. Influence of this term is gradually increased during the reconstruction and it is regulated by parameter $\mu > 0$.

Problem (8) is a constrained and quadratic type energy-minimization problem that can be solved by several optimization methods. We have selected Spectral Projected Gradient (SPG) optimization algorithm (Birgin et al. 2001) for this task, since it has shown good performance in successful application in similar problems (Lukić and Balázs 2016; Lukić and Nagy 2014; Nagy and Lukić 2016; Birgin et al. 2000). The SPG algorithm combines the non-monotone line search algorithm (Grippo et al. 1986) and the spectral gradient step-length selection (Barzilai and Borwein 1988; Raydan 1997; Birgin and Martínez 2001), its pseudo-code is presented in Alg. 1.

The gradient of the regularization term

$$\left(\mathcal{C}(u)-\mathcal{C}^*\right)^2$$

in the energy function (8) is determined in a fully analytical manner, for its exact expression see (Lukić and Balázs 2020). This allows a fast minimization process and

Algorithm 1: SPG optimization algorithm.

$u^0 = [0.5, 0.5, ..., 0.5]^T$;
$d^0 = P_{\Omega}(u^0 - \nabla E_Q(u^0)) - u^0$; $k = 0$;
repeat
  Determine the current step-length $\lambda^k > 0$ by a line search approach, see (Birgin et al. 2001);
  $u^{k+1} = u^k + \lambda^k d^k$;
  Calculate the gradient spectral step-length $\theta_{k+1} > 0$, see (Birgin et al. 2001);
  $d^{k+1} = P_{\Omega}(u^{k+1} - \theta_{k+1}\nabla E_Q(u^{k+1})) - u^{k+1}$; $k = k + 1$;
until $\|u^k - u^{k-1}\|_\infty < 10^{-2}$;
$u^{new} = u^k$;

determination of the smooth solution by SPG algorithm, in contrast to the GCORI-ENTBT method when the gradient of the shape orientation based regularization is calculated in numerical manner.

The stopping criterion for smooth solution is given by

$$\langle u, \tau - u \rangle < E_{bin},$$

where E_{bin} regulates the degree of binarization of the solution u and it is set to 100 in our experiments. Partial binarization of the continuous solution boosts the determination of data cost terms for graph cuts method.

Next action after calculation of the smooth solution is its full binarization. This is done by applying the graph cuts method based on the Potts model, described in Sect. 3.1. We construct energy function according to the one used in the Potts model (4) in the following way:

- Data cost term, D,

$$D(p, 0) = u(p),$$
$$D(p, 1) = 1 - u(p),$$

  where $u(p)$ represents the intensity of a pixel p.
- Set of neighboring pairs, $\mathcal{N}$,
  $(p, q) \in \mathcal{N}$ if the image coordinates of p and q differ for one value only.
- Interaction potential, K,

$$K_{(p,q)} = 1.$$

After successful construction of the energy function (4), the next task is solving a problem of finding a minimum of this function. That is achieved by implementing GCO graph cuts based optimization algorithm, introduced in (Boykov et al. 2001) and further analyzed in (Boykov and Kolmogorov 2004; Delong et al. 2010; Kolmogorov and Zabih 2004). The output of GCO algorithm are label values d_p for each pixel p, where d_p is predefined as $d_p = 0 \rightarrow 0$ and $d_p = 1 \rightarrow 1$. As a result of the adequate construction of the function (4), obtained label values determine intensities of pixels in the final (binary) solution, marking the end of the reconstruction process.

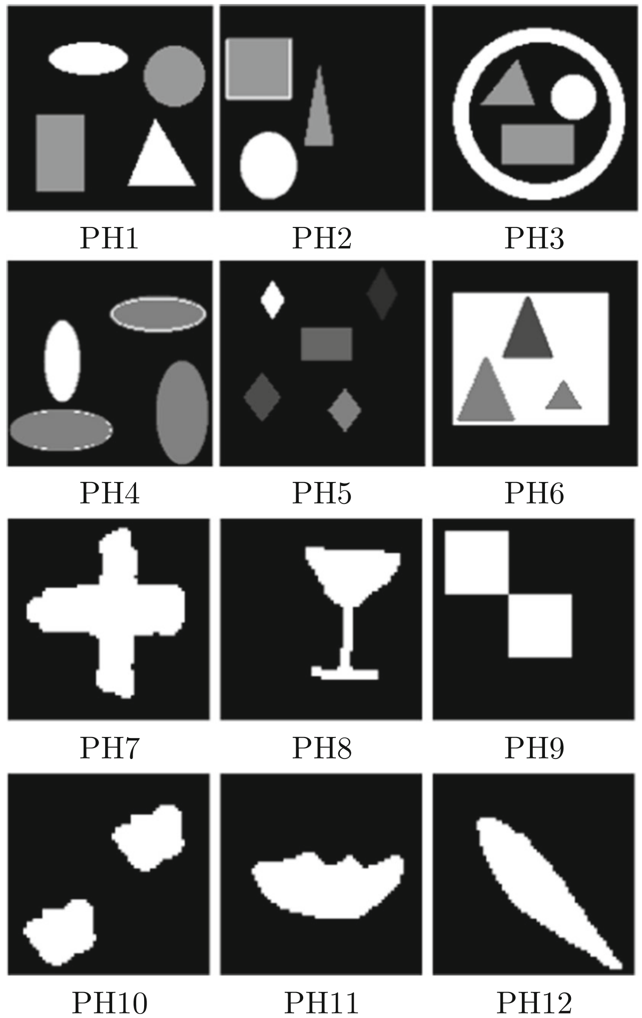


Fig. 1 Original test images. Phantoms PH1, PH2 and PH3 contain 3 different gray levels, PH4, PH5 and PH6 contain 6 different gray levels, while PH7, PH8, PH9, PH10, PH11, PH12 present binary images

We denote this method by Graph Cuts Binary Tomography Assisted by the Circularity prior (GCCIRCBT) reconstruction method.

4 Experimental results

In this section we aim to evaluate the performance of the algorithms that are, to the best of our knowledge, most commonly used for solving similar reconstruction problems in DT and to experimentally confirm if the new circularity prior improves the reconstruction quality. In order to achieve above mentioned goal, we use test set containing 12 test images (phantoms) presented in Fig. 1. PH1-3 contain 3 gray levels, PH4-6 contain 6 gray levels, while PH7-12 represent binary images. Images PH1-PH11 are synthetic, whereas PH12 is a binary segmented florescence image of Calcein stained Chinese hamster ovary cell. A total of 128 parallel rays is taken for each projection direction for multi gray level images and 64 projection rays for binary images. In all cases, the projection directions are uniformly selected between 0 and 180 degrees. This projection information is used as input in reconstruction algorithms:

Table 1 Experimental results for PH1, PH2 and PH3 images, using three different reconstruction methods. The abbreviation d indicates the number of projections.

d		PH1				PH2				PH3			
		6	9	12	15	6	9	12	15	6	9	12	15
MWP	(PE)	**255**	159	59	35	**143**	138	20	18	655	456	275	174
	(m.r. %)	**1.55**	0.97	0.36	0.21	**0.87**	0.84	0.12	0.11	3.99	2.78	1.67	1.06
TRDT	(PE)	412	175	48	28	209	141	17	17	412	301	101	41
	(m.r. %)	2.51	1.06	0.29	0.17	1.28	0.86	0.10	0.10	2.51	1.83	0.61	0.25
GCDT	(PE)	272	**69**	**8**	**5**	225	**124**	**12**	**12**	**272**	**116**	**20**	**9**
	(m.r. %)	1.66	**0.42**	**0.04**	**0.03**	1.37	**0.76**	**0.07**	**0.07**	**1.66**	**0.70**	**0.12**	**0.05**

The values representing the best performance are highlighted by bold font.

Table 2 Experimental results for PH1, PH2 and PH3 images, using three different reconstruction methods. The abbreviation e.t. means elapsed time in minutes and *d* indicates the number of projections.

d		PH1				PH2				PH3			
		6	9	12	15	6	9	12	15	6	9	12	15
MWPDT	(PRE)	**14.70**	12.19	9.96	9.08	**14.11**	**18.94**	6.08	7.71	**19.83**	18.77	18.80	16.43
	(e.t.)	1.76	2.63	3.17	4.06	5.34	8.17	6.36	11.62	2.19	2.87	4.30	4.66
TRDT	(PRE)	18.66	14.72	10.61	8.87	17.98	17.30	7.09	7.09	23.64	17.87	13.66	10.61
	(e.t.)	7.73	12.58	14.55	17.77	6.24	10.82	16.01	17.74	7.28	11.07	13.39	16.00
GCDT	(PRE)	23.24	**11.12**	**6.52**	**4.39**	26.77	21.04	**6.01**	**6.00**	25.87	**14.96**	**7.59**	**5.60**
	(e.t.)	7.73	12.58	14.55	17.77	6.25	10.82	16.01	17.74	7.29	11.07	13.40	16.01

The values representing the best performance are highlighted by bold font.

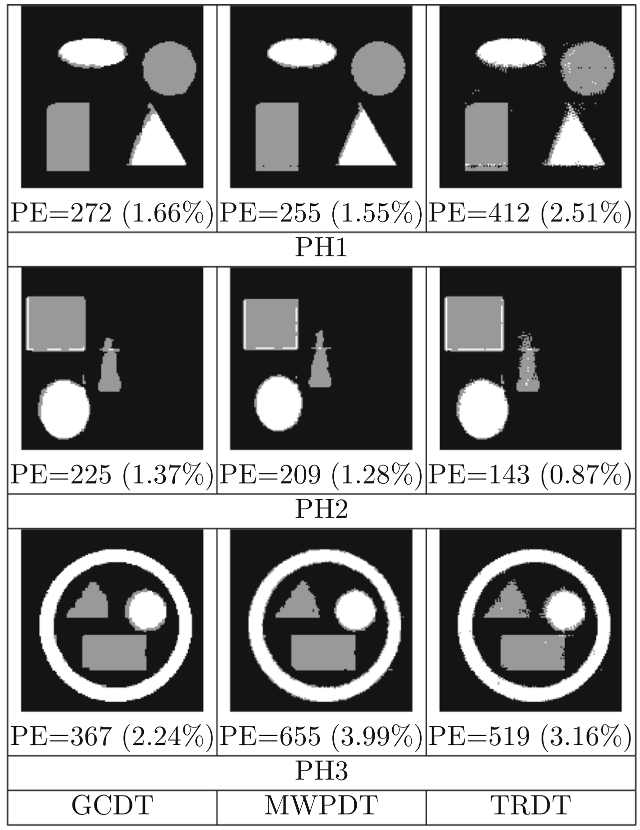


Fig. 2 Reconstructions of the 3 gray level test images using data from 6 projection directions

- Graph Cuts Discrete Tomography Algorithm (GCDT) (Lukić and Marčeta 2017)
- Discrete Algebraic Reconstruction Technique (DART) (Batenburg and Sijbers 2007)
- Method based on classical threshold (TRDT)
- Multi Well Potential based method (MWPDT) (Lukić 2011)
- Graph Cuts Tomography Assisted by the Orientation prior (GCORIENTBT) (Marčeta and Lukić 2020)
- Graph Cuts Binary Tomography Assisted by the Circularity prior (GCCIRCBT), introduced in this paper

MWPDT method is developed and used only for phantoms with 3 gray levels, GCORIENTBT and GCCIRCBT only for binary images, while the rest of the algorithms mentioned in this section can be used for reconstruction of images with arbitrary number of gray levels. In our experiments, each reconstruction method (GCDT, DART, TRDT, MWPDT, GCORIENTBT, GCCIRCBT) is completely implemented in programming language Matlab.

In the evaluation process, we analyze the quality of the reconstructions. The quality of the reconstructions is expressed by the pixel error (PE), i.e. the absolute number of the misclassified pixels, and by the misclassification rate ($m.r.$), i.e. the pixel error measure relative to the total number of image pixels. Additionally, as a qualitative error measure, we consider the projection error, defined by $PRE = \|Au^r - b\|$, where u^r represents the reconstructed image. This error indicates the accordance of the reconstruction with the given projection data.

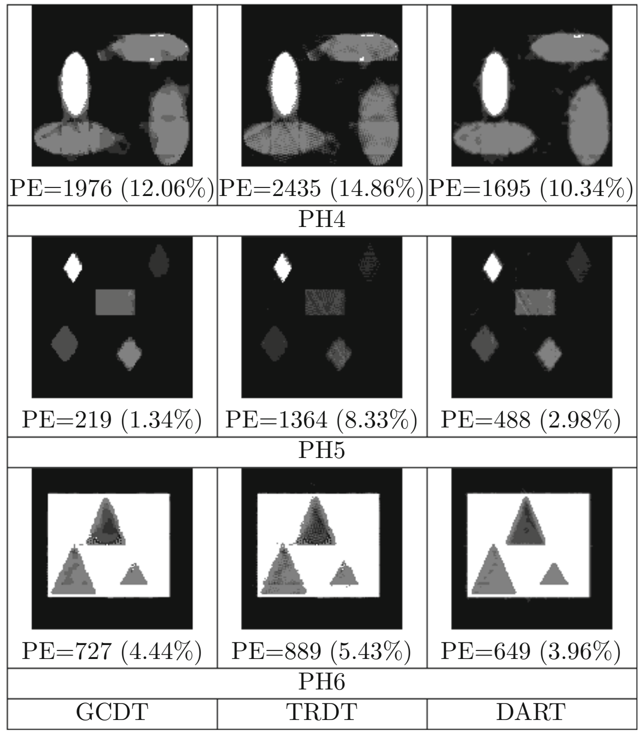


Fig. 3 Reconstructions of the 6 gray level test images using data from 6 projection directions

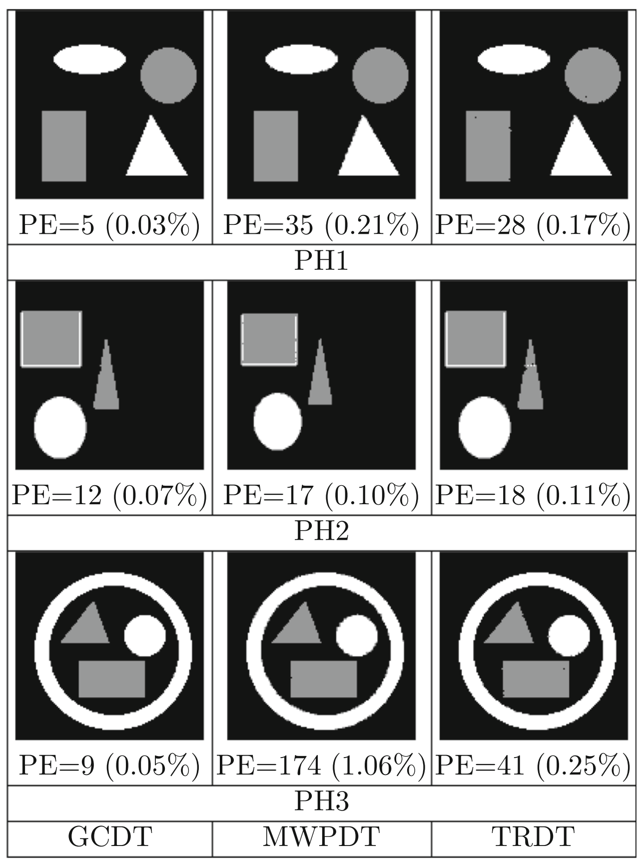


Fig. 4 Reconstructions of the 3 gray level test images using data from 15 projection directions

Table 3 Experimental results for PH4, PH5 and PH6 images, using three different reconstruction methods. The abbreviation *d* indicates the number of projections.

d		PH4				PH5				PH6			
		6	9	12	15	6	9	12	15	6	9	12	15
GCDT	(PE)	1976	**804**	**551**	**399**	**219**	**134**	**42**	**28**	727	**473**	**251**	**192**
	(m.r. %)	12.06	**4.91**	**3.36**	**2.44**	**1.34**	**0.82**	**0.26**	**0.17**	4.44	**2.89**	**1.53**	**1.17**
TRDT	(PE)	2435	1415	1188	998	1364	1330	1286	1274	889	807	587	552
	(m.r. %)	14.86	8.64	7.25	6.09	8.32	8.12	7.85	7.78	5.43	4.92	3.58	3.37
DART	(PE)	**1695**	1242	1177	1089	488	379	288	319	**649**	836	596	707
	(m.r. %)	**10.34**	7.58	7.18	6.65	2.98	2.31	1.76	1.95	**3.96**	5.10	3.64	4.32

The values representing the best performance are highlighted by bold font.

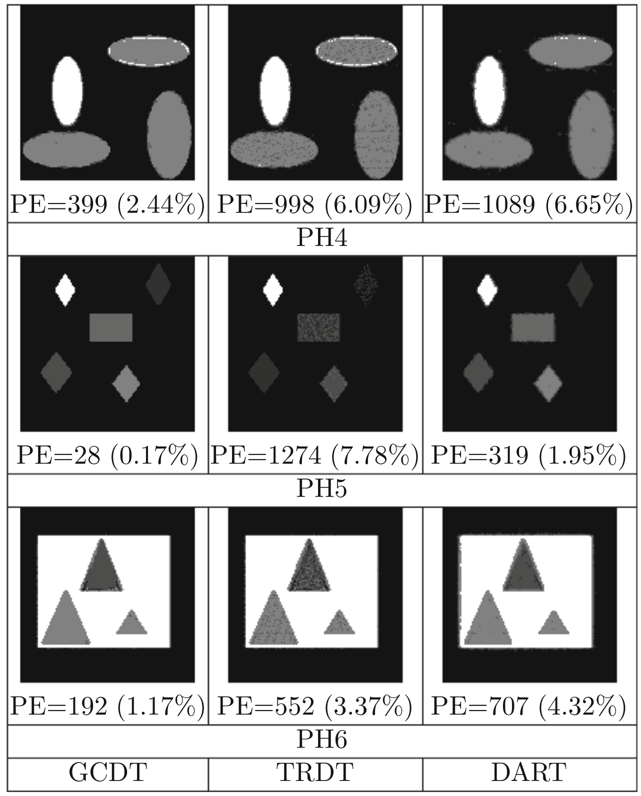


Fig. 5 Reconstructions of the 6 gray levels test images using data from 15 projection directions

We compare performance of the observed algorithms on the various projection data of the test images. The projection direction is determined by the angle α and, in this section, number of different projection angles used for obtaining projection data is denoted by d. Horizontal and vertical projection data provide enough information for determination of circularity and orientation shape descriptors (see Lukić and Balázs 2020). Thus, orientation and circularity of a shape do not present any additional information about an object if horizontal and vertical projection data is known. Therefore, in cases of 3 or more projection angles circulation and orientation as a priori information are redundant, as they are already present in projection values. This is the reason why for reconstructions that use higher number of projections, we do not show results for GCORIENTBT and GCCIRCBBT (they would be identical to those obtained by GCDT).

Results regarding the performance of different algorithms on test images PH1, PH2 and PH3 (Tables 1, 2 and Figs. 2, 4), show that for PE and $m.r.$ as metrics, method GCDT provided the best results in 10 out of 12 cases, for PRE metric GCDT method dominates in 8 cases, while in terms of the execution time MWPDT method prevails. GCDT uses significantly higher number of iterations for obtaining the smooth solution compared to MWPDT method in total, thus resulting in greater consumption of time.

Reconstruction results of phantoms with 6 different gray levels (Table 3) show that, compared to TRDT and DART, GCDT method prevails in 10 out of 12 cases, whilst DART performs the best in 2 cases. In Figs. 3, 5 reconstructions from 6 and 15 projection directions respectively are presented.

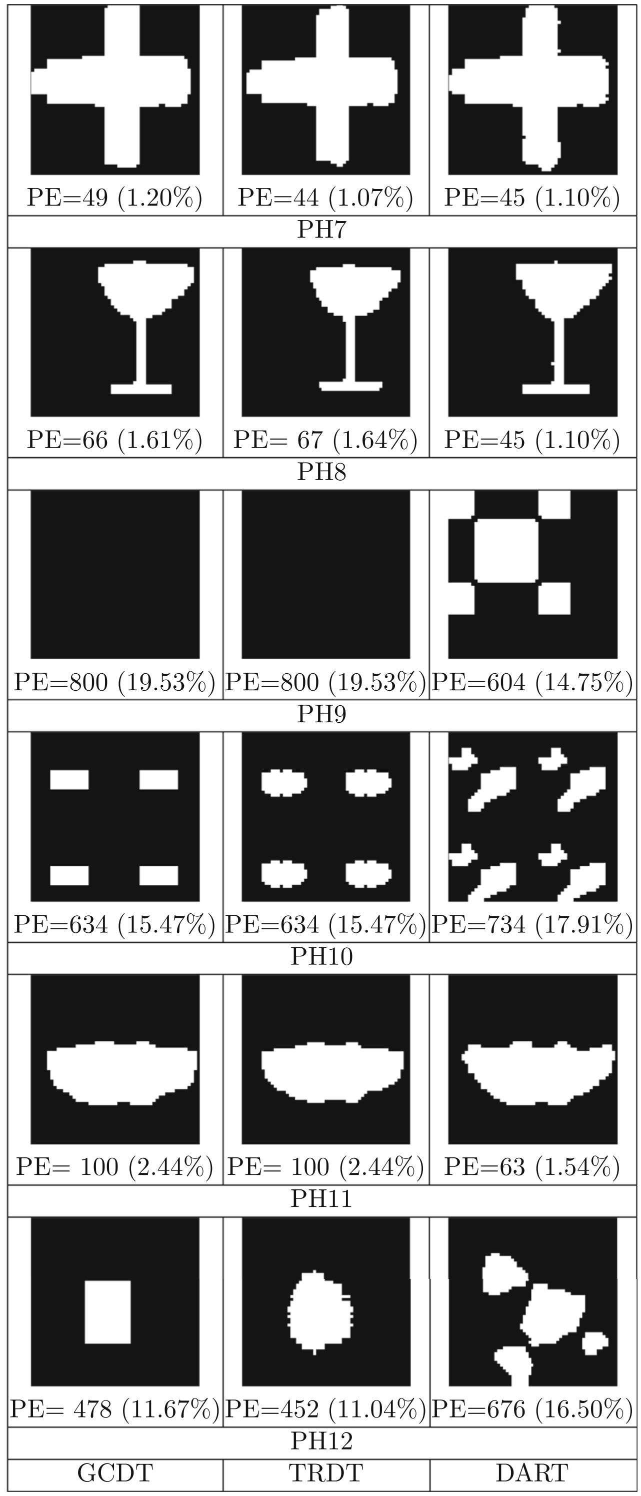


Fig. 6 Reconstructions of the binary test images using data from 2 projection directions (vertical and horizontal)

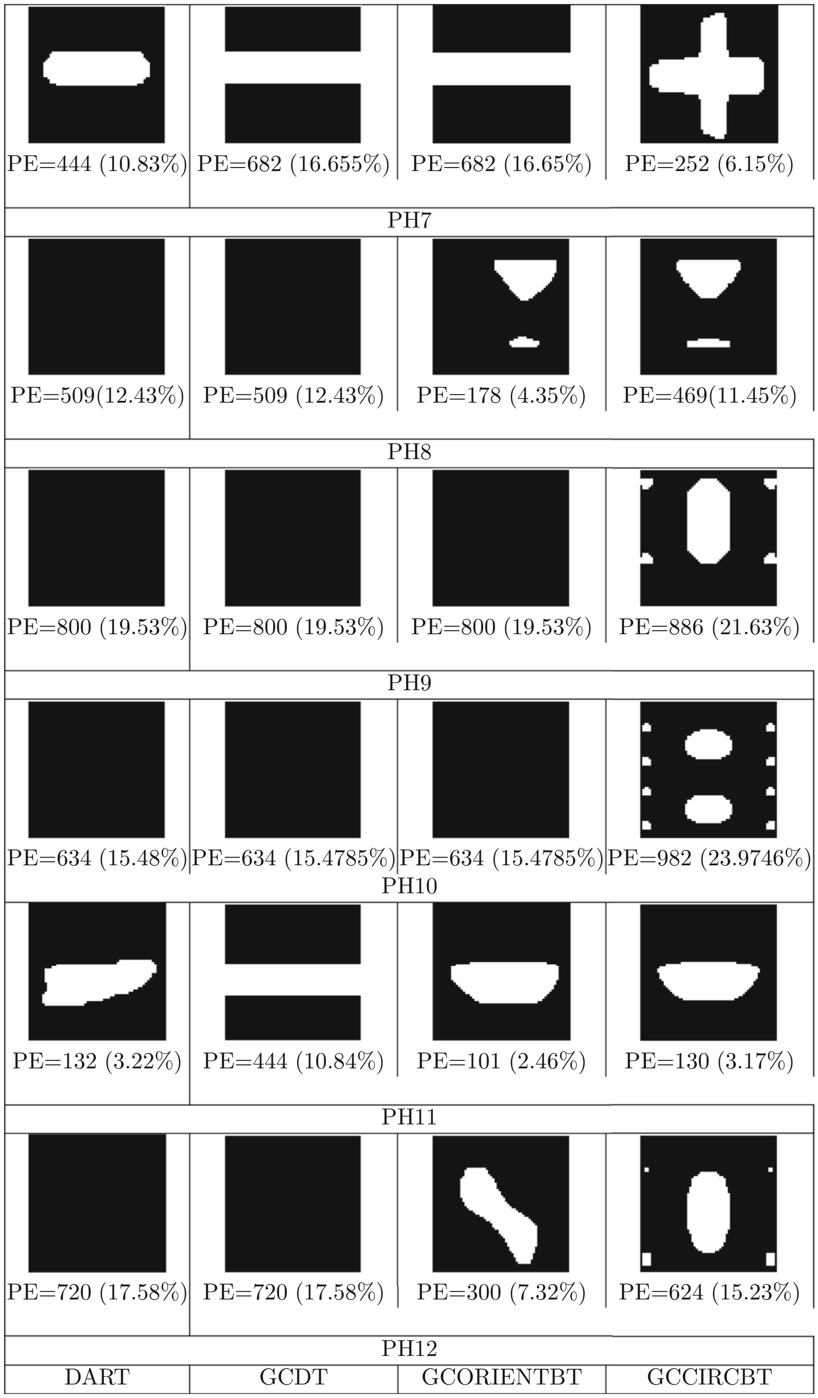


Fig. 7 Reconstructions of the binary test images using data from 1 projection direction, $\alpha = 0°$

Results down to this point of the analysis show competitive performance of a model based on the combination of graph cuts and a gradient based method (GCDT). This encourages us to test and develop this algorithm further.

Our experiments on binary images (Fig. 6) demonstrate that GCDT method gives poor results in cases of the reconstruction from two projections. In order to avoid this drawback we can add orientation and circularity as a priori information to GCDT

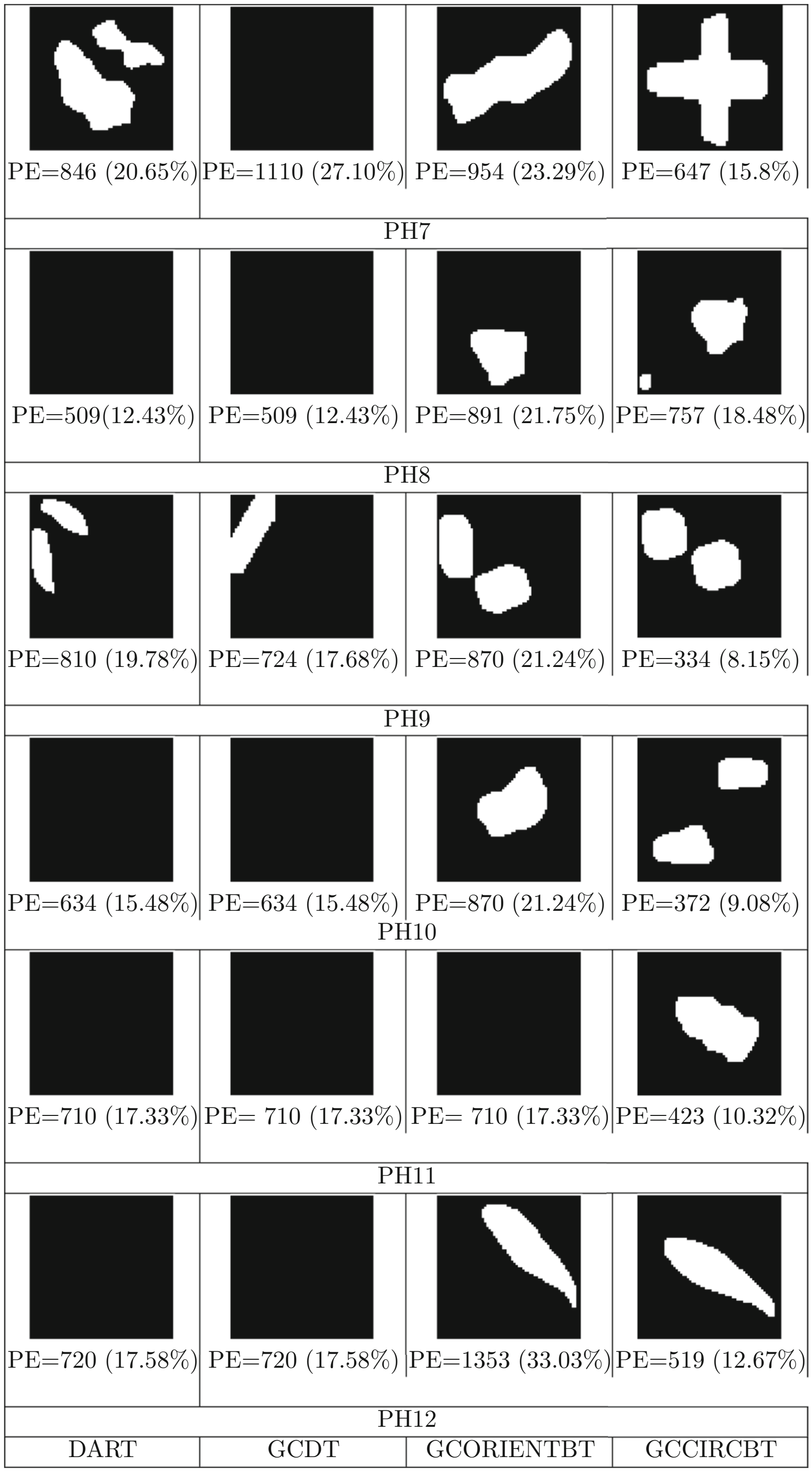


Fig. 8 Reconstructions of the binary test images using data from 1 projection direction, direction angle is $\alpha = 60^\circ$

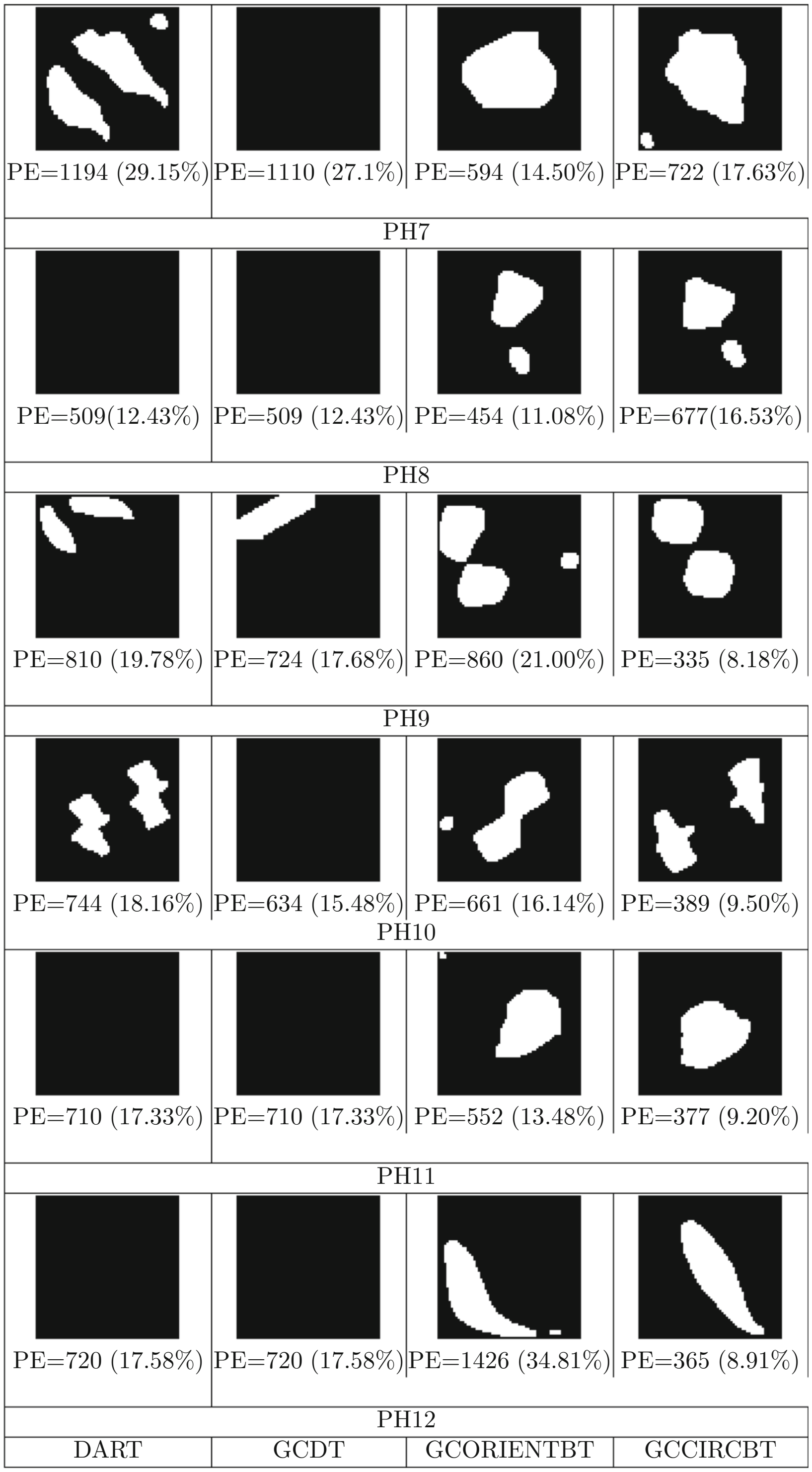


Fig. 9 Reconstructions of the binary test images using data from 1 projection direction, direction angle is $\alpha = 30°$

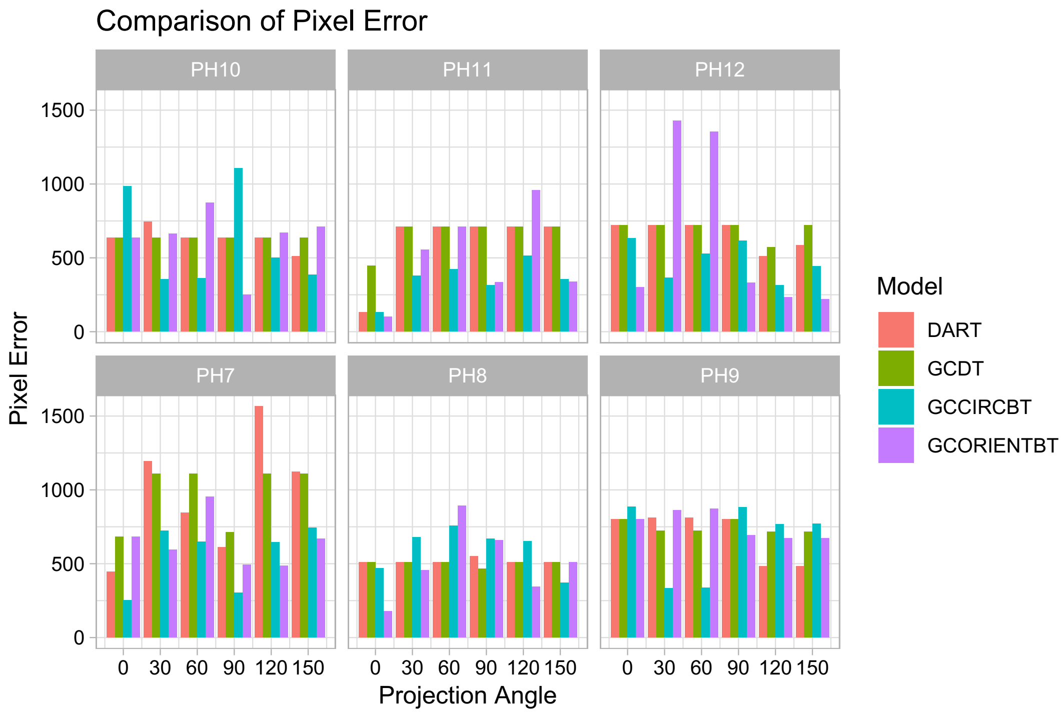


Fig. 10 Experimental results for BT using four different reconstruction methods (Pixel Error)

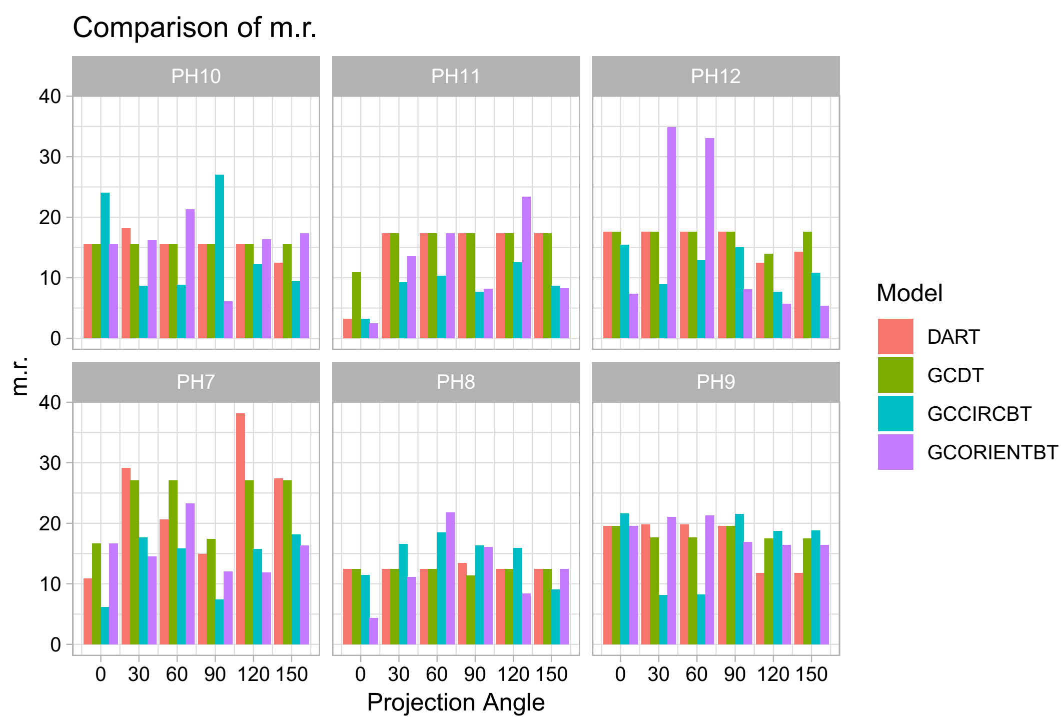


Fig. 11 Experimental results for BT using four different reconstruction methods (m.r.)

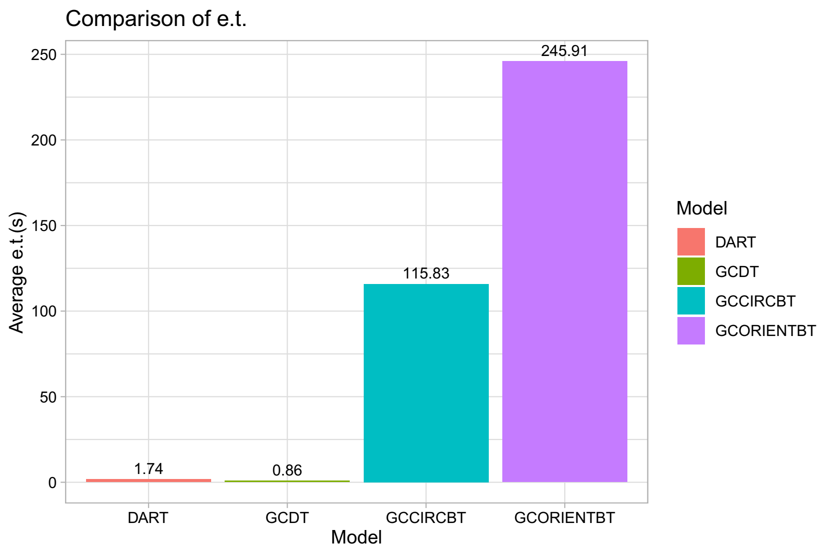


Fig. 12 Experimental results for BT using four different reconstruction methods (average e.t.)

method, thus building the GCORIENTBT and GCCIRCBT algorithm. Later, we have compared these algorithms with two other reconstruction methods (DART, GCDT) (Figs. 10, 11). For analyzis we have used 6 binary images, data was given from one projection, and we have tested the model by using 6 different projection angles.

In (Marčeta and Lukić 2020) it was shown that GCORIENTBT gives very good results in cases of limited projection view availability. We have now been inquiring if the circularity is equally good or even better to use as a regularization term. In 17 out of 36 cases GCCIRCBT gives the best reconstruction (smallest PE/m.r.) and GCORIENTBT wins in 13 cases. It can be noticed that, as expected, by adding the prior to GCDT method, significantly better results for BT are obtained in cases of limited projection data availability. The noticeable advantage of GCCIRCBT is in running time, execution time of GCCIRCBT is in most of the cases significantly shorter compared to its best competitor GCORIENTBT (Fig. 12).

Summarizing the results obtained by the total of 24 multi-gray level analyzed reconstruction tasks, see Tables 1 and 3, the quality of the reconstruction, indicated by m.r. for the proposed GCDT method was the best in 20 cases, i.e. in 83% of the analyzed cases. Further, GCORIENTBT and GCCIRCBT together performed better in 83% of the cases, with GCCIRCBT being slightly superior, thus indicating excellent performance of graph cuts based reconstruction approach in DT as well as prevailing advantages of using shape circularity as an a priori information.

5 Conclusions

This paper has highlighted the noteworthy performance of an algorithm based on the combination of gradient based method and graph cuts optimization method for solving problems in Discrete Tomography. In cases of very limited projection accessibility we modified the method using shape descriptor circularity as an a priori information about the object. Conducted experiments gave priority in reconstruction quality to the proposed method compared to the formerly published reconstruction methods. In conclusion, our results show that it is suitable to use the combination of a gradient based method with the graph cuts optimization method, which can be successfully enhanced by circularity as an a priori information, for providing high quality reconstructions in discrete tomography.

Data availability The data (digital test images) that support the findings of this study are available from the corresponding author, upon reasonable request.

Declarations

Conflict of interest The authors declare that there is no conflict of interest.

Funding The research is not financially supported.

Code availability The Matlab code used in this paper is available on request per email from the corresponding author.

References

Barzilai J, Borwein JM (1988) Two point step size gradient methods. IMA J Num Anal 8:141–148

Batenburg KJ, Sijbers J (2007) DART: A Fast heuristic algebraic reconstruction algorithm for discrete tomography. In Proceedings of International conference on image processing (ICIP), pages 133–136

Birchfield S, Tomasi C (1999) Multiway cut for stereo and motion with slanted surfaces. Proc. International Conf. Computer Vision, pages 489–495

Birgin E, Martínez J (2001) Spectral conjugate gradient method for unconstrained optimization. Appl Math Optim 43:117–128

Birgin E, Martínez J, Raydan M (2000) Nonmonotone spectral projected gradient methods on convex sets. SIAM J Optim 10:1196–1211

Birgin EG, Martínez JM, Raydan M (2001) Algorithm: 813: SPG - software for convex-constrained optimization. ACM Trans Math Softw 27:340–349

Boykov Y, Jolly MP (2001) Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In Proc. International Conf. Computer Vision, pages 105–112

Boykov Y, Kolmogorov V (2003) Computing geodesics and minimal surfaces via graph cuts. In Proc. International Conf. Computer Vision, pages 26–33

Boykov Y, Kolmogorov V (2004) An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Trans. PAMI 26(9):1124–1137

Boykov Y, Veksler O, Zabih R (1998) Markov random fields with efficient approximations. In Proceedings of IEEE Conf. Computer Vision and Pattern Recognition, pages 648–655

Boykov Y, Veksler O, Zabih R (2001) Fast approximate energy minimization via graph cuts. IEEE Trans PAMI 23(11):1222–1239

Delong A, Osokin A, Isack H. N, Boykov Y (2010) Fast approximate energy minimization with label costs. Proc. IEEE Conf. computer vision and pattern recognition 96(1):1–27

Grippo L, Lampariello F, Lucidi S (1986) A nonmonotone line search technique for Newton's method. SIAM J Numer Anal 23:707–716

Herman GT, Kuba A (1999) Discrete tomography: foundations, algorithms and applications. Birkhäuser

Herman GT, Kuba A (2006) Advances in discrete tomography and its applications. Birkhäuser

Kim J, Zabih R (2003) Automatic segmentation of contrast- enhanced image sequences. In Proc. International Conf. Computer Vision, pages 502–509

Kolmogorov V, Zabih R (2001) Visual correspondence with occlusions using graph cuts. In Proc. International Conf. Computer Vision, pages 508–515

Kolmogorov V, Zabih R (2004) What energy functions can be minimized via graph cuts? IEEE Trans PAMI 26(2):147–159

Kwatra V, Schoedl A, Essa I, Turk G, Bobick A (2003) Graphcut textures: image and video synthesis using graph cuts. In Proc. SIGGRAPH 2003, pages 277–286. ACM Trans. Graphics

Lukić T (2011) Discrete tomography reconstruction based on the multi-well potential. In Proceedings of combinatorial image analysis - 14th international workshop (IWCIA), volume 6636 of *LNCS*, pages 335–345, Madrid, Spain. Springer-Verlag

Lukić T, Balázs P (2016) Binary tomography reconstruction based on shape orientation. Pattern Recognit Lett 79:18–24

Lukić T, Balázs P (2020) Shape circularity assisted tomography reconstruction. Physica Scripta 95(10):105211

Lukić T, Marčeta M (2017) Gradient and graph cuts based method for multi-level discrete tomography. In Proceedings of combinatorial image analysis - 18th International workshop (IWCIA), LNCS, pages 322–333, Plovdiv, Bulgaria, Springer-Verlag

Lukić T, Nagy B (2014) Deterministic discrete tomography reconstruction method for images on triangular grid. Pattern Recognit Lett 49:11–16

Marčeta M, Lukić T (2020) Graph cuts based tomography enhanced by shape orientation. In T. Lukic, R. P. Barneva, V. E. Brimkov, L. Comic, and N. Sladoje, editors, Combinatorial image analysis - 20th International workshop, IWCIA 2020, Novi Sad, Serbia, July 16-18, 2020, Proceedings, volume 12148 of lecture notes in computer science, pages 219–235. Springer. https://doi.org/10.1007/978-3-030-51002-2_16

Nagy B, Lukić T (2016) Dense projection tomography on the triangular tiling. Fundamenta Informaticae 145:125–141

Raydan M (1997) The Barzilai and Browein gradient method for the large scale unconstrained minimization problem. SIAM J Optim 7:26–33

Schüle T, Schnörr C, Weber S, Hornegger J (2005) Discrete tomography by convex-concave regularization and D.C. programming. Discrete Appl Math 151:229–243

Sonka M, Hlavac V, Boyle R (2007) Image processing, analysis, and machine vision. Thomosn-Engineering

Žunić J, Rosin PL, Kopanja L (2006) On the orientability of shapes. IEEE Trans Image Process 15:3478–3487

Žunić J, Hirota K, Rosin PL (2010) A hu invariant as a shape circularity measure. Pattern Recognit 43:47–57

APPENDIX B

Publication II

Marčeta, Marina, and Tibor Lukić. "Graph cuts based tomography enhanced by shape orientation." International Workshop on Combinatorial Image Analysis. Springer, Cham, 2020

Graph Cuts Based Tomography Enhanced by Shape Orientation

Marina Marčeta(✉) and Tibor Lukić

Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia
marceta.marina@gmail.com, tibor@uns.ac.rs

Abstract. The topic of this paper includes graph cuts based tomography reconstruction methods in binary and multi-gray level cases. A energy-minimization based reconstruction method for binary tomography is introduced. This approach combines the graph cuts and a gradient based method, and applies a shape orientation as an a priori information. The new method is capable for reconstructions in cases of limited projection view availability. Results of experimental evaluation of the considered graph cuts type reconstruction methods for both binary and multi level tomography are presented.

Keywords: Discrete tomography · Shape orientation · Energy minimization methods

1 Introduction

The word tomography comes from Greek words *tomos* which means slice and *graphein* which means to write and it denotes an area in image processing that deals with reconstructing images from the given projection data. Its main focus usually are the objects which are not easily accessible or visible.

A wave penetrates through an unknown object and collects the projection data from the object. In order to gather enough data for a successful reconstruction, waves usually need to penetrate the object from large number of directions. The unknown object that tomography pursues to restore is identified as a function with a domain that can be discrete or continuous and a range that is a given set of (usually) real numbers. Therefore, in order to obtain the image of the unknown object, it is needed to reconstruct this function based on the known data (integrals or sums over subsets of its domain). In *Discrete Tomography* (DT) [11,12] the range of the function is a finite set. DT that deals with the problem of reconstruction of binary images is named binary tomography (BT). If DT deals with the reconstruction of digital images which consist of numerous gray levels, it is referenced as multi-level discrete tomography.

Although problems of multi-level discrete tomography image reconstruction can occur frequently in the application, to the best of our knowledge, there are only few reconstruction algorithms that deal with such problems. Discrete Algebraic Reconstruction Technique (DART) [1], Multi-Well Potential based method

T. Lukić et al. (Eds.): IWCIA 2020, LNCS 12148, pp. 219–235, 2020.
https://doi.org/10.1007/978-3-030-51002-2_16

(MWPDT) [17], a combination of non-local projection constraints with a continuous convex relaxation of the multilabeling problem [25] and the Non-Linear Discretization function based reconstruction algorithm (NLD) [24] are among them. A recently introduced method (GCDT) [20], which combines a gradient based algorithm with a graph cuts type optimization method, showed good performance for this type of problem. This paper gives an overview and experimental evaluation of most often used algorithms for multi-level tomography reconstruction problem.

The good performance of the GCDT algorithm for multi-level case motivate us to make a step further and apply and adjust this approach to an other interesting problem: binary tomography for limited projection availability. We propose a new method which incorporates an a prior knowledge about the solution into the reconstruction process. The smooth solution is determined by the regularized gradient based SPG algorithm [4] which is subsequently binarized applying a max-flow type graph cut algorithm, introduced in [9] and further analyzed in [7,10,15]. The added prior information is the shape orientation descriptor [26]. Our motivation for the selection of this type of prior information is lies in the fact that it shows excellent performance in combination with convex-concave and gradient based approaches, see the reconstruction method (BTO) [18].

The paper has the following structure. Section 2 gives the description of the basic reconstruction problem. In Sect. 3 the new reconstruction method is presented. Experimental results are provided in Sect. 4. Finally, the conclusion is given in the Sect. 5.

2 Reconstruction Problem

In this paper we consider the DT reconstruction problem, represented by a linear system of equations

$$Au = b, \quad A \in \mathbb{R}^{M \times N}, \ \ u \in \Lambda^N, \ b \in \mathbb{R}^M, \quad \Lambda = \{\lambda_1, \lambda_2, ..., \lambda_k\}, \quad k \geq 2 \quad (1)$$

where k is the number of different gray level values. The set Λ is given by the user. The matrix A is a so-called *projection matrix*, whose each row corresponds to one projection ray. The corresponding components of the vector b contain the detected projection values, while the vector u represents the unknown image to be reconstructed. The i-th row entries $a_{i,\cdot}$ of A represent the length of the intersection of the pixels and the i-th projection ray passing through them. The projection value measured by a projection ray is calculated as a sum of products of the pixel's intensity and the corresponding length of the projection ray through that pixel. Projections are taken from different directions. For each projection direction a number of parallel projection rays are taken (parallel beam projection). The projection direction is determined by the angle β. The distance between two adjacent parallel projection rays can vary depending on the reconstruction problem, we set this distance to be equal to the side length of pixels. The reconstruction problem means finding the solution image u of the linear system of Eq. (1), where the projection matrix A and the projection vector b are given. This system is often undetermined ($N > M$).

3 Graph Cuts Reconstruction Method Assisted by Shape Orientation

A directed, weighted graph $G = (X, \rho)$, consists of a set of nodes X and a set of directed edges ρ that connect them. The nodes,in image processing interpretations, mostly correspond to pixels or voxels in 3D. All edges of graph are assigned some weight or cost.

Let $G = (X, \rho)$ be a directed graph with non-negative edge weights that has two special nodes or terminals, the source A and the sink B. An $a-b$-cut (which is referred informally as a cut) $C = A, B$ is a partition of the terminals in X into two disjoint sets A and B so that $a \in A$ and $b \in B$. The cost of the cut is the sum of costs of all edges that go from A to B:

$$c(A, B) = \sum_{x \in A, y \in B, (x,y) \in \rho} c(x, y).$$

The minimum $a-b$-cut problem is to find a cut C with the minimum cost among all cuts. Algorithms to solve this problem can be found in [7].

The approach that uses graph cuts for energy minimization has, as basic technique, construction of a specialized graph for the energy function to be minimized such that the minimum cut on the graph also minimizes the energy. The form of the graph depends on the exact form of X and on the number of labels. The minimum cut, in turn, can be computed very efficiently by max flow algorithms.

These methods have been successfully used in the last 20 years for a wide variety of problems, naming image restoration [8,9], stereo and motion [2,14], image synthesis [16], image segmentation [6] and medical imaging [5,13].

The Potts model in graph cuts theory is based on the minimization of the following energy

$$E(d) = \sum_{p \in \mathcal{P}} D(p, d_p) + \sum_{(p,q) \in \mathcal{N}} K_{(p,q)} \cdot (1 - \delta_{d_p, d_q}), \tag{2}$$

where $d = \{d_p | p \in \mathcal{P}\}$ represents the labelling of the image pixels $p \in \mathcal{P}$. By $D(p, d_p)$ we denote the data cost term, where $D(p, d_p)$ is a penalty or cost for assigning a label d_p to a pixel p. $K_{(p,q)}$ is an interaction potential between neighboring pairs p and q, $\mathcal{N}$ is a set of neighboring pairs. Function δ_{d_p, d_q} is Kronecker delta function.

3.1 Shape Orientation

In this section we give a short description and a calculation method of the shape orientation.

The geometrical moment of a digitize image u is defined by

$$m_{p,q}(u) = \sum_{(i,j) \in \Omega} u(i, j) i^p j^q,$$

where $\Omega \subseteq \mathbb{R}^2$ denotes the image domain.

The center of gravity (or centroid) of an image (or a shape) u is defined by

$$(C_x(u), C_y(u)) = \left(\frac{m_{1,0}(u)}{m_{0,0}(u)}, \frac{m_{0,1}(u)}{m_{0,0}(u)} \right).$$

The centroid enables the definition of the centralized moment which is translation invariant. The centralized moment of an image u of order $p+q$ is given by

$$\overline{m}_{p,q}(u) = \sum_{(i,j)\in\Omega} u(i,j)(i - C_x(u))^p (j - C_y(u))^q.$$

The shape orientation is an often used and well known shape descriptor [26]. The orientation is determined by the angle α, which represents the slope of the axis of the second moment of inertia (orientation axis) of the considered shape [23]. The orientation (angle α) for the the given image u can be calculated by the following equation:

$$\frac{\sin(2\alpha)}{\cos(2\alpha)} = \frac{2 \cdot \bar{m}_{1,1}(u)}{\bar{m}_{2,0}(u) - \bar{m}_{0,2}(u)}. \tag{3}$$

Moments in (3) are translation invariant, making the orientation invariant to translation transformations, for more details see [18,26].

3.2 The Proposed Method

Our tomography reconstruction approach is a combination of the graph cuts method and a gradient based minimization method. In the first step, we determine the data cost values for each image pixels. The data cost values are determined as intensity values of the continuous or smooth approximation (solution) of the final reconstruction image. The smooth solution is obtained by the following energy-minimization problem

$$\begin{aligned} \min_{u\in[0,1]^N} E_Q(u) := w_P\|Au - b\|_2^2 + w_H \sum_{i=1}^{N} \sum_{j\in\Upsilon(i)} (u_i - u_j)^2 \\ + w_{\mathcal{O}} \left(\phi(u) - \alpha^*\right)^2 + \mu \left\langle u, \tau - u \right\rangle, \end{aligned} \tag{4}$$

where $\tau = [1, 1, \ldots, 1]^T$ is a vector of size N. Parameter w_P regulates the influence of the data fitting term, w_H controls the second term, whose role is to enhance the homogeneity or compactness of the solution. By $\Upsilon(i)$ two neighbor pixel indexes (in x and y axis directions) of pixel i is denoted. The orientation of the solution u is denoted by $\phi(u)$, while α^* is given true orientation. Parameter w_O determines the impact of the orientation regularization. The task of the concave regularization term $\langle u, \tau - u\rangle$ (inner product of vectors u and $1 - u$) is to push pixel intensities toward binary values. Parameter μ regulates the influence of this binarization term and its value is gradually increased during the reconstruction process. The problem (4), for each fixed μ, is solved by the Spectral

Projected Gradient (SPG) iterative optimization algorithm, originally proposed by Birgin et al. [3]. Motivation for application of this algorithm is supported by its successful application in similar problems, see [19,21,22]. The reconstruction process (4) is terminated before the fully binary solution is achieved. The termination criterion is given by

$$\langle u, \tau - u \rangle < E_{bin},$$

where E_{bin} regulates the degree of binarization of the solution u. In our experiments its value is set as 100. This partially binarization of the smooth solution is applied in order to improve the determination of data cost terms for graph cuts method.

In the next step we have to fully binarize the smooth solution of the problem (4), obtained by the SPG algorithm. For this task we apply the graph cuts method based on the Potts model, described in Sect. 3. The data cost term D in (2) is determined using information provided by the smooth solution u. More precisely, we define it in the following way

$$\begin{aligned} D(p, 0) &= u(p), \\ D(p, 1) &= 1 - u(p), \end{aligned}$$

where $u(p)$ represents the intensity of a pixel p. The idea is to make data cost small or cheap in the vicinity of the given gray values. The neighbor pairs are defined based on 1-neighboring system, i.e., $(p, q) \in \mathcal{N}$ if the image coordinates of p and q differ for one value only. The interaction potential $K_{(p,q)}$ (see (2)) in our experiments is set as a constant and its value is 1. Now, the energy function in (2) is determined and ready to be minimized. For this task we use the GCO graph cuts based optimization algorithm, introduced in [9] and further analyzed in [7,10,15]. The GCO algorithm determines the label values d_p for each pixel p. Each label value is assigned to one predefined gray level in the following way: $d_p = 0 \rightarrow 0$ and $d_p = 1 \rightarrow 1$. Therefore, the obtained label values also determine intensities of pixels in the final (binary) solution, therefore the reconstruction process is terminated. We denote this method by Graph Cuts Tomography Assisted by the Orientation prior (GCORIENTBT) reconstruction method. If in the Eq. 4 we set values of parameters w_o, w_H and μ as 0, and solve it using the combination of SPG and GCO algorithms, we are getting algorithm introduced in [20] and denoted by GCDT.

4 Experimental Results

In this section we experimentally evaluate the proposed graph cuts based reconstruction methods, denoted by GCDT and GCORIENTBT. In the experiments for multi gray level tomography we use 6 test images (phantoms), as originals in reconstructions, presented in Fig. 1. Phantoms PH1, PH2 and PH3 contain 3 gray levels, while phantoms PH4, PH5 and PH6 contain 6 gray levels. In addition, GCDT was also tested on binary images using phantoms PH7, PH8, PH9 as well as on two medical images: binary segmented CT image of a bone implant, inserted in a leg of a rabbit (PH10) and a Binary segmented florescence image of Calcein stained Chinese hamster ovary cell (PH11). We consider reconstructions of these images obtained from different projection directions. A total of 128 parallel rays are taken for each projection direction for multi gray level images and 64 projection rays for binary images. In all cases, the projection directions are uniformly selected between 0 and 180°. The obtained results are compared with two reconstruction methods suggested for multi level discrete tomography, so far: 1) Multi Well Potential based method (MWPDT) proposed in [17] (this method is developed and used only for phantoms with 3 gray levels); and 2) DART method, proposed in [1]. We also include into the evaluation process a simple method based on classical threshold, denoted by TRDT. Additionally, binary images are compared with similar reconstruction method for binary tomography introduced in [18] and denoted by BTO. All reconstruction methods (BTO, DART, GCORIENTBT, MWPDT, GCDT and TRDT) are implemented completely in Matlab.

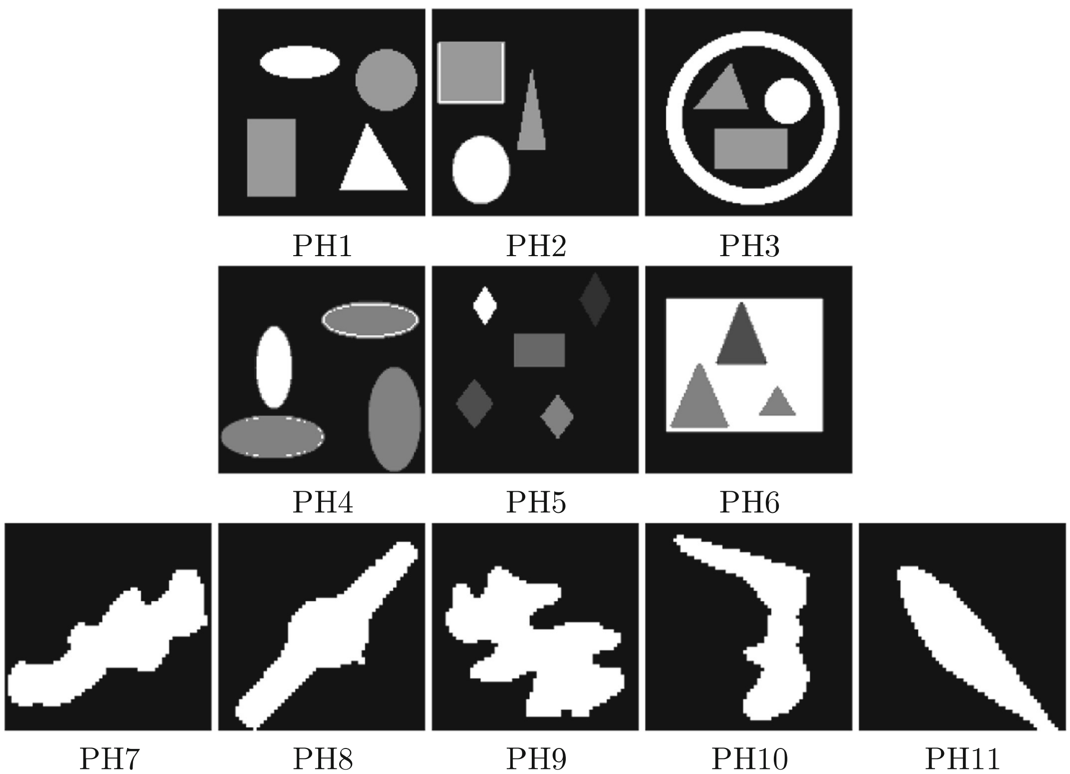


Fig. 1. Original test images (128x128). Phantoms PH1, PH2 and PH3 contain 3 different gray levels (0,0.5,1), PH4, PH5 and PH6 contain 6, while PH7, PH8, PH9, PH10, PH11 present binary images.

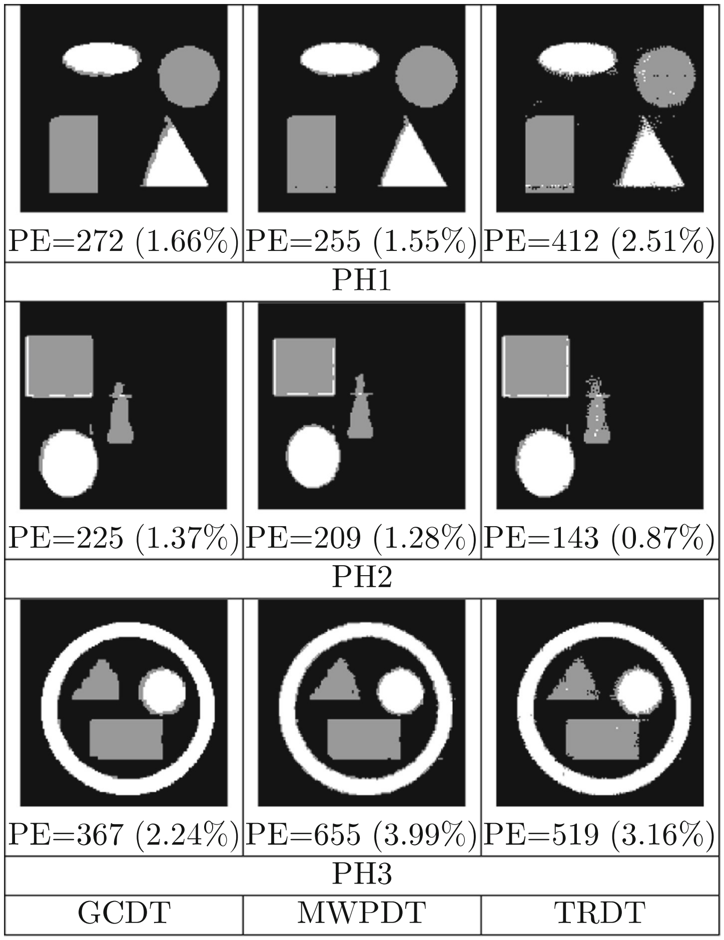


Fig. 2. Reconstructions of the test images using data from 6 projection directions.

In the evaluation process, we analyze the quality of the reconstructions and required running times. The quality of the reconstructions are expressed by the pixel error (PE), i.e., the absolute number of the misclassified pixels, and by the misclassification rate ($m.r.$), i.e., the pixel error measure relative to the total number of image pixels. Also, as a qualitative error measure, we consider the projection error, defined by $PRE = \|Au^r - b\|$, where u^r represents the reconstructed image. This error express the accordance of the reconstruction with the given projection data.

In Table 1 we present pixel errors for reconstructions of three phantom images (PH1, PH2 and PH3) obtained from different number of projections by three different methods (MWPDT, GCDT and TRDT). From total of 12 different reconstruction problems, GCDT method provided best results in 10 cases, while in 2 cases the dominant was the MWP method. Regrading the PRE values, see Table 2, the proposed GCDT method dominated in 8 cases, while MWP in 4 cases.

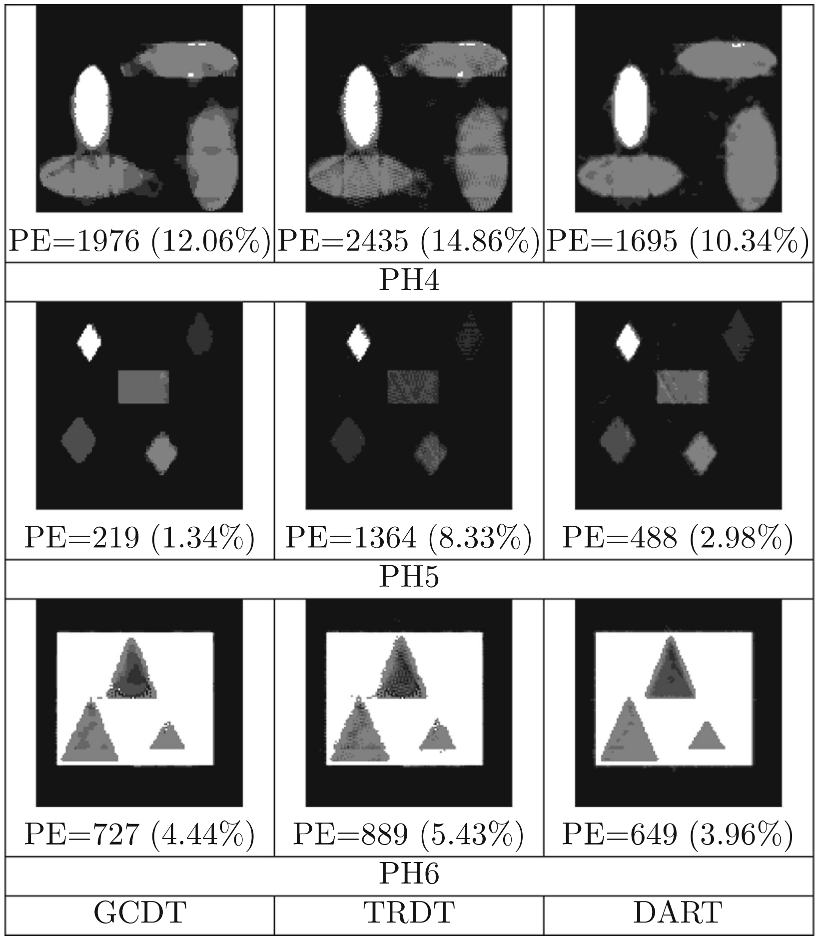


Fig. 3. Reconstructions of the test images using data from 6 projection directions.

Best running times in all experiments were achieved by the MWP method. GCDT and TRDT methods use the smooth solution/reconstruction as a first step, before the "discretization process" starts.

The smooth solution is achieved as a final termination, with high precision, which requires significantly higher number of iterations compared to MWPDT method in total, thus resulting in greater consumption of time. In Figs. 2 and 4 reconstructions from 6 and 15 projection directions of images containing 3 gray level are presented.

Reconstruction results of phantoms with 6 different gray levels (Table 3) show that, compared to TRDT and DART, GCDT method prevails in 10 out of 12 cases, whilst DART performes the best in 2 cases. In Figs. 3 and 5 reconstructions from 6 and 15 projection directions respectively are presented.

In addition to multi level discrete tomography, we have tested our algorithm on binary images. It can be noticed on the Fig. 6 that GCDT method gives poor results in the cases of the reconstruction from two projections. On the other hand, already from higher number of projections, GCDT shows competitive performance. We have tried to avoid this drawback by adding orientation as

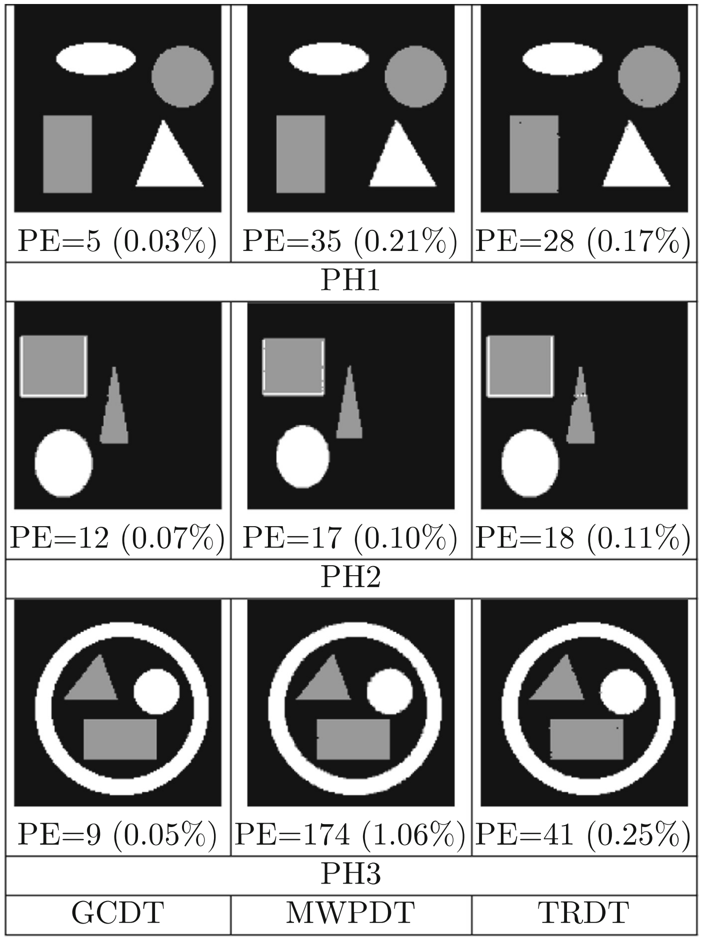


Fig. 4. Reconstructions of the test images using data from 15 projection directions.

a priori information to GCDT method, thus building the GCORIENTBT algorithm. Later, we have compared the GCORIENTBT algorithm with three other reconstruction methods (BTO, DART, GCDT) (Tables 4 and 5). In 12 out of 15 cases, GCORIENTBT gives the best reconstruction (smallest PE/m.r.). It can be noticed that, as expected, by adding the orientation prior to GCDT method, significantly better results for BT are obtained. The advantage of GCORIENTBT is in running time as well. Execution time of GCORIENTBT is in most of the cases even more than 50% shorter compared to its best competitor BTO.

Summarizing the results obtained by the total of 24 multi- gray level analyzed reconstruction tasks, see Tables 1 and 3, the quality of the reconstruction, indicated by m.r.for the proposed GCDT method was the best in 20 cases, i.e., in 83% of the analyzed cases. We emphasize that the results of the GCDT method, in cases when they are the best, are significantly better (in the most of the cases more than 50% better). Farther, GCORIENTBT performed better in 80% of the cases.

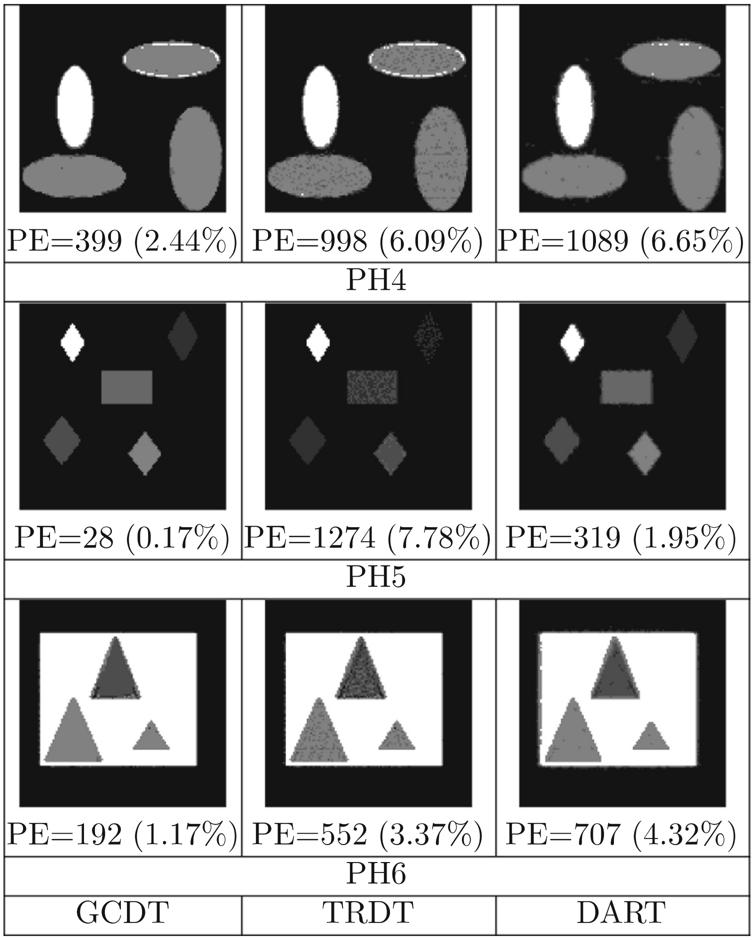


Fig. 5. Reconstructions of the test images using data from 15 projection directions.

Table 1. Experimental results for PH1, PH2 and PH3 images, using three different reconstruction methods. The abbreviation d indicates the number of projections. The best performance is bold fonted.

d		PH1				PH2				PH3			
		6	9	12	15	6	9	12	15	6	9	12	15
	(PE)	**255**	159	59	35	**143**	138	20	18	655	456	275	174
MWP	(m.r. %)	**1.55**	0.97	0.36	0.21	**0.87**	0.84	0.12	0.11	3.99	2.78	1.67	1.06
	(PE)	412	175	48	28	209	141	17	17	412	301	101	41
TRDT	(m.r. %)	2.51	1.06	0.29	0.17	1.28	0.86	0.10	0.10	2.51	1.83	0.61	0.25
	(PE)	272	**69**	**8**	**5**	225	**124**	**12**	**12**	**272**	**116**	**20**	**9**
GCDT	(m.r. %)	1.66	**0.42**	**0.04**	**0.03**	1.37	**0.76**	**0.07**	**0.07**	**1.66**	**0.70**	**0.12**	**0.05**

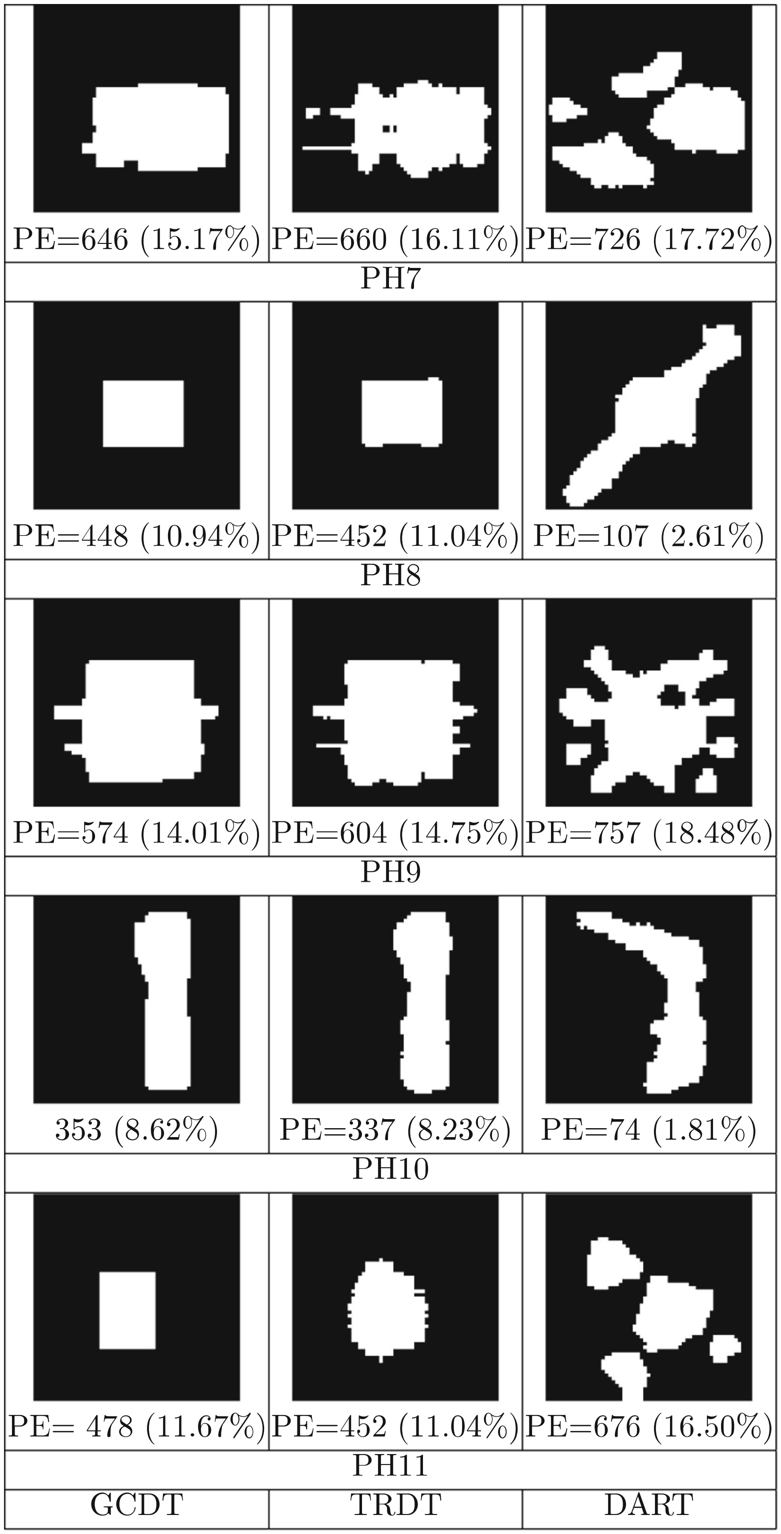


Fig. 6. Reconstructions of the binary test images using data from 2 projection directions.

			
PE=466 (11.38%) e.t.= 227	PE=813 (19.85%) e.t.= 0.87	PE=929 (22.68%) e.t.= 0.46	PE=464 (11.32%) e.t.= 82.61
PH7			
			
PE=669(16.33%) e.t.= 32	PE=817 (19.95%) e.t.=0.84	PE=817 (19.95%) e.t.= 0.47	PE=652(15.92%) e.t.= 31.95
PH8			
			
PE=970 (23.68%) e.t.= 177	PE=564 (13.77%) e.t.= 0.5	PE=990 (24.17%) e.t.= 0.69	PE=976(23.83%) e.t.= 60.17
PH9			
			
PE=672 (16.41%) e.t.= 328	PE=722 (17.63%) e.t.= 0.54	PE=722 (17.63%) e.t.= 0.41	PE=657 (16.04%) e.t.= 60.44
PH10			
			
PE=209 (5.10%) e.t.= 280	720 (17.58%) e.t.= 0.44	PE=720 (17.58%) e.t.= 0.42	PE=180 (4.39%) e.t.= 66.89
PH11			
BTO	DART	GCDT	GCORIENTBT

Fig. 7. Reconstructions of the binary test images using data from 1 projection direction, $\alpha = 0^{\circ}$. The abbreviation e.t. means elapsed time in seconds.

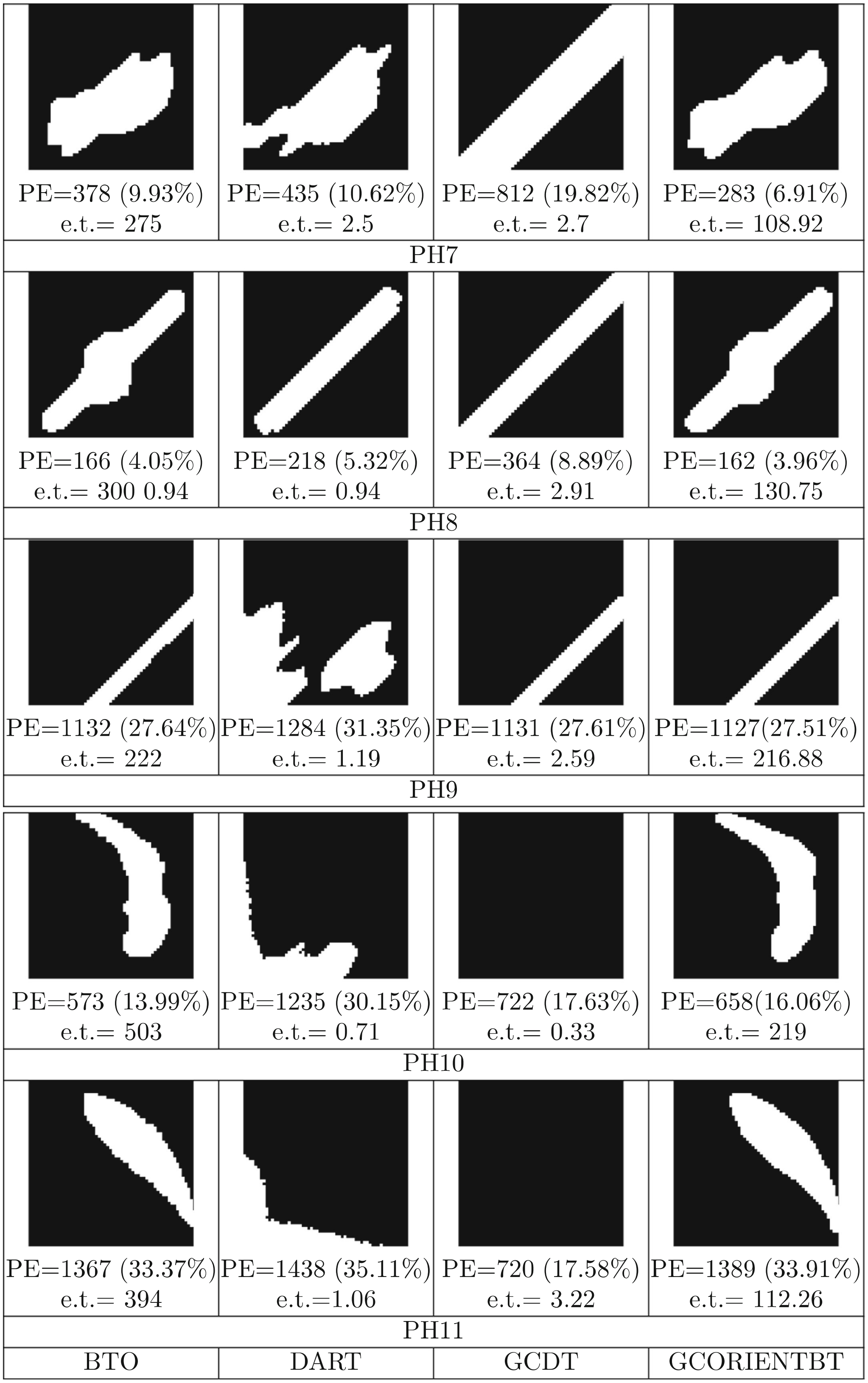


Fig. 8. Reconstructions of the binary test images using data from 1 projection direction, $\alpha = 45°$. The abbreviation e.t. means elapsed time in seconds.

Table 2. Experimental results for PH1, PH2 and PH3 images, using three different reconstruction methods. The abbreviation e.t. means elapsed time in minutes and d indicates the number of projections. The best performance is bold fonted.

d		PH1				PH2				PH3			
		6	9	12	15	6	9	12	15	6	9	12	15
MWPDT	(PRE)	**14.70**	12.19	9.96	9.08	**14.11**	**18.94**	6.08	7.71	**19.83**	18.77	18.80	16.43
	(e.t.)	1.76	2.63	3.17	4.06	5.34	8.17	6.36	11.62	2.19	2.87	4.30	4.66
TRDT	(PRE)	18.66	14.72	10.61	8.87	17.98	17.30	7.09	7.09	23.64	17.87	13.66	10.61
	(e.t.)	7.73	12.58	14.55	17.77	6.24	10.82	16.01	17.74	7.28	11.07	13.39	16.00
GCDT	(PRE)	23.24	**11.12**	**6.52**	**4.39**	26.77	21.04	**6.01**	**6.00**	25.87	**14.96**	**7.59**	**5.60**
	(e.t.)	7.73	12.58	14.55	17.77	6.25	10.82	16.01	17.74	7.29	11.07	13.40	16.01

Table 3. Experimental results for PH4, PH5 and PH6 images, using three different reconstruction methods. The abbreviation d indicates the number of projections. The best performance is bold fonted.

d		PH4				PH5				PH6			
		6	9	12	15	6	9	12	15	6	9	12	15
GCDT	(PE)	1976	**804**	**551**	**399**	**219**	**134**	**42**	**28**	727	**473**	**251**	**192**
	(m.r. %)	12.06	**4.91**	**3.36**	**2.44**	**1.34**	**0.82**	**0.26**	**0.17**	4.44	**2.89**	**1.53**	**1.17**
TRDT	(PE)	2435	1415	1188	998	1364	1330	1286	1274	889	807	587	552
	(m.r. %)	14.86	8.64	7.25	6.09	8.32	8.12	7.85	7.78	5.43	4.92	3.58	3.37
DART	(PE)	**1695**	1242	1177	1089	488	379	288	319	**649**	836	596	707
	(m.r. %)	**10.34**	7.58	7.18	6.65	2.98	2.31	1.76	1.95	**3.96**	5.10	3.64	4.32

Table 4. Experimental results for PH7, PH8 and PH9 images, using three different reconstruction methods. α indicates the direction of projections. The best performance is bold fonted.

α		PH7			PH8			PH9		
		0°	45°	90°	0°	45°	90°	0°	45°	90°
BTO	(PE)	466	378	383	669	166	362	970	1132	369
	(m.r. %)	11.38	9.23	9.35	16.33	4.06	8.84	23.68	27.64	9.01
DART	(PE)	813	435	935	817	218	817	**564**	1284	670
	(m.r. %)	19.85	10.62	22.83	19.95	5.32	19.95	**13.77**	31.35	16.36
GCDT	(PE)	929	812	935	817	364	817	990	1131	1180
	(m.r. %)	22.68	19.82	22.83	19.95	8.89	19.95	24.17	27.61	28.81
GCORIENTBT	(PE)	**464**	**283**	**378**	**652**	**162**	**324**	976	**1127**	**373**
	(m.r. %)	**11.32**	**6.91**	**9.23**	**15.93**	**3.96**	**7.91**	23.83	**27.51**	**9.11**

According to the above presented results, we conclude that experiments confirm the capability of the proposed methods to provide high quality reconstructions both for gray-level (GCDT) and binary (GCORIENTBT) images.

Table 5. Experimental results for PH10 and PH11 using three different reconstruction methods. α indicates the direction of projections. The best performance is bold fonted.

α		PH10			PH11		
		0°	45°	90°	0°	45°	90°
BTO	(PE)	672	**573**	597	209	1367	379
	(m.r. %)	16.41	**13.99**	14.58	5.10	33.37	9.25
DART	(PE)	722	1235	414	720	1438	720
	(m.r. %)	17.63	30.15	17.63	17.58	35.11	17.58
GCDT	(PE)	722	722	534	720	**720**	720
	(m.r. %)	17.63	17.63	13.04	17.58	**17.58**	17.58
GCORIENTBT	(PE)	**657**	658	**377**	**180**	1389	**340**
	(m.r. %)	**16.04**	16.06	**9.20**	**4.39**	33.91	**8.30**

5 Conclusions

In this paper we presented an approach for solving problem posed by discrete tomography. The approach uses a gradient based method to obtain a smooth reconstruction of an image and then uses graph cuts optimization method for discretization. In cases of lowered projection directions, the method uses orientation as an a priori information. Conducted experiments gave priority in reconstruction quality to the proposed methods compared to the formerly published reconstruction methods. Based on the obtained experimental results and analysis presented in this paper, we have concluded that the combination of a gradient based method with the graph cuts optimization method is suitable for providing high quality reconstructions in discrete tomography.

Acknowledgement. Authors acknowledge the Ministry of Education and Sciences of the R. of Serbia for support via projects OI-174008 and III-44006. T. Lukić acknowledges support received from the Hungarian Academy of Sciences via DOMUS project.

References

1. Batenburg, K.J., Sijbers, J.: DART: a fast heuristic algebraic reconstruction algorithm for discrete tomography. In: Proceedings of International Conference on Image Processing (ICIP), pp. 133–136 (2007)
2. Birchfield, S., Tomasi, C.: Multiway cut for stereo and motion with slanted surfaces. In: Proceedings of the International Conference on Computer Vision, pp. 489–495 (1999)
3. Birgin, E.G., Martínez, J.M., Raydan, M.: Algorithm: 813: SPG - software for convex-constrained optimization. ACM Trans. Math. Softw. **27**, 340–349 (2001)
4. Birgin, E., Martínez, J.: A box-constrained optimization algorithm with negative curvature directions and spectral projected gradients. Computing **15**, 49–60 (2001)
5. Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In: Proceedings of the International Conference on Computer Vision, pp. 105–112 (2001)

6. Boykov, Y., Kolmogorov, V.: Computing geodesics and minimal surfaces via graph cuts. In: Proceedings of the International Conference on Computer Vision, pp. 26–33 (2003)
7. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Trans. PAMI **26**(9), 1124–1137 (2004)
8. Boykov, Y., Veksler, O., Zabih, R.: Markov random fields with efficient approximations. In: Proceedings of the International Conference on Computer Vision and Pattern Recognition, pp. 648–655 (1998)
9. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Trans. PAMI **23**(11), 1222–1239 (2001)
10. Delong, A., Osokin, A., Isack, H.N., Boykov, Y.: Fast approximate energy minimization with label costs. In: Proceedings of the International Conference on Computer Vision and Pattern Recognition, vol. 96(1), pp. 1–27 (2010)
11. Herman, G.T., Kuba, A.: Discrete Tomography: Foundations, Algorithms and Applications. Birkhäuser, Boston (1999)
12. Herman, G.T., Kuba, A.: Advances in Discrete Tomography and Its Applications. Birkhäuser, Boston (2006)
13. Kim, J., Zabih, R.: Automatic Segmentation of Contrast- Enhanced Image Sequences. In: Proceedings of the International Conference on Computer Vision, pp. 502–509 (2003)
14. Kolmogorov, V., Zabih, R.: Visual correspondence with occlusions using graph cuts. In: Proceedings of the International Conference on Computer Vision, pp. 508–515 (2001)
15. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts? IEEE Trans. PAMI **26**(2), 147–159 (2004)
16. Kwatra, V., Schoedl, A., Essa, I., Turk, G., Bobick, A.: Graphcut Textures: Image and Video Synthesis Using Graph Cuts. In: Proc. SIGGRAPH 2003. pp. 277–286. ACM Trans. Graphics (2003)
17. Lukić, T.: Discrete tomography reconstruction based on the multi-well potential. In: Aggarwal, J.K., Barneva, R.P., Brimkov, V.E., Koroutchev, K.N., Korutcheva, E.R. (eds.) IWCIA 2011. LNCS, vol. 6636, pp. 335–345. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21073-0_30
18. Lukić, T., Balázs, P.: Binary tomography reconstruction based on shape orientation. Pattern Recogn. Lett. **79**, 18–24 (2016)
19. Lukić, T., Lukity, A.: A spectral projected gradient optimization for binary tomography. In: Rudas, I.J., Fodor, J., Kacprzyk, J. (eds.) Computational Intelligence in Engineering. SCI, vol. 313, pp. 263–272. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15220-7_21
20. Lukić, T., Marčeta, M.: Gradient and graph cuts based method for multi-level discrete tomography. In: Brimkov, V.E., Barneva, R.P. (eds.) IWCIA 2017. LNCS, vol. 10256, pp. 322–333. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59108-7_25
21. Lukić, T., Nagy, B.: Deterministic discrete tomography reconstruction method for images on triangular grid. Pattern Recogn. Lett. **49**, 11–16 (2014)
22. Nagy, B., Lukić, T.: Dense projection tomography on the triangular tiling. Fundam. Inform. **145**, 125–141 (2016)
23. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis, and Machine Vision. Thomson-Engineering, Toranto (2007)

24. Varga, L., Balázs, P., Nagy, A.: An energy minimization reconstruction algorithm for multivalued discrete tomography. In: Proceedings of 3rd International Symposium on Computational Modeling of Objects Represented in Images, pp. 179–185. Taylor & Francis, Rome (2012)
25. Zisler, M., Petra, S., Schnörr, C., Schnörr, C.: Discrete tomography by continuous multilabeling subject to projection constraints. In: Rosenhahn, B., Andres, B. (eds.) GCPR 2016. LNCS, vol. 9796, pp. 261–272. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45886-1_21
26. Žunić, J., Rosin, P.L., Kopanja, L.: On the orientability of shapes. IEEE Trans. Image Process. **15**, 3478–3487 (2006)

APPENDIX C

Publication III

Lukić, Tibor, and **Marina Marčeta**. "Gradient and graph cuts based method for multi-level discrete tomography." International Workshop on Combinatorial Image Analysis. Springer, Cham, 2017.

Gradient and Graph Cuts Based Method for Multi-level Discrete Tomography

Tibor Lukić[1] and Marina Marčeta[1]

Faculty of Technical Sciences, University of Novi Sad, Serbia
tibor@uns.ac.rs, marina.marceta@uns.ac.rs

Abstract. In this paper, we are proposing a new energy-minimization reconstruction method for the multi gray level discrete tomography. The proposed reconstruction approach combines a gradient based algorithm with the graph cuts optimization. This new technique is able to reconstruct images that consist of an arbitrary number of gray levels. We present the experimental evaluation of the new method, where we compare its performance with performance of the already suggested methods for multi-level discrete tomography. The obtained experimental results give an advantage to the proposed approach, especially regarding the quality of the reconstructed test images.

1 Introduction

Tomography [14] reconstructs images of non accessible or non visible objects. It deals with recovering images from a number of projections. Tomography will be our focus in this paper. From a mathematical point of view, the object corresponds to a function. The problem posed, is to reconstruct this function from its integrals, or its sums over subsets of its domain. In general, the tomographic reconstruction problem may be continuous or discrete. In *Discrete Tomography* (DT) [15, 16] the range of the function is a finite set. In practice, DT often deals with reconstructions of digital images that consist of a number of gray levels. DT has a wide range of applications in areas where the materials of the object under investigation are known before, such as industrial non-destructive testing or electron tomography [15, 16].

To the best of our knowledge, there are only a few reconstruction algorithms suggested for this DT problem, that deal with multi gray level tomography image reconstruction. These are the Discrete Algebraic Reconstruction Technique (DART) [2], the Multi-Well Potential based method (MWPDT) [22], method which combines non-local projection constraints, continuous convex relaxation of the Multilabeling problem and DC programming (MDC) [25], and the Non-Linear Discretization function based reconstruction algorithm (NLD) [30]. The DART method uses a fixed threshold function for the discretization process (without any regularization), which can lead to radical solutions and less accurate reconstructions, especially in the case of reduced projection data. The MDC is a powerful method, but less flexible related to adding new regularization terms, because the energy function has to be expressed as a difference of convex functions. The MWPDT and NLD methods applies a non-convex energy function in the reconstruction process, which can stuck in local minimum, i.e., in a semi-continuous solution. The proposed method in this paper is developed in such a way to avoids the above listed disadvantages.

One of the approaches used for solving problems in image processing and computer vision has been developed based on graph cuts. The core of this approach is to construct a specialized graph for the energy function to be minimized such that the minimum cut on the graph also minimizes the energy (either globally or locally). The minimum cut, in turn, can be computed very efficiently by max-flow algorithms. The output of these algorithms is generally a solution with some interesting theoretical quality guarantees. In [20] is given, which conditions the energy function needs to satisfy in order to be minimized via graph cuts.

In this paper, we propose a new deterministic reconstruction method for the DT problem, which combines a gradient based method, with a graph cuts type optimization method. The proposed method uses a smooth regularization prior and allows reconstruction of images that contain an arbitrary number of different gray levels.

The structure of the paper is the following. In Section 2, the basic reconstruction problem is described. In Section 3, we present the new reconstruction method based on the graph cuts approach. Our experimental results are provided in Section 4 and finally, Section 5 is the conclusion.

2 Reconstruction Problem

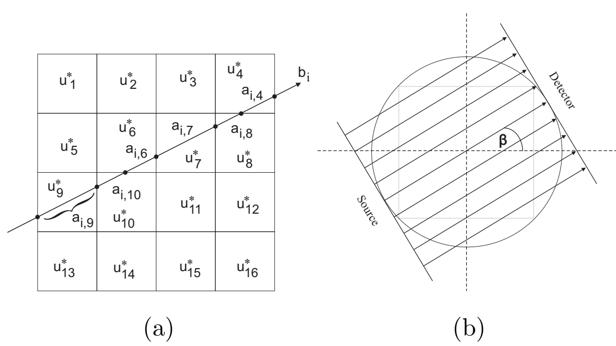


Fig. 1: (a) Example of a projection value calculation on an image u^* of size $N = 4 \times 4 = 16$. A projection ray penetrates through the image pixels. The projection value b_i is calculated by $b_i = a_{i,4}u_4^* + a_{i,6}u_6^* + a_{i,7}u_7^* + a_{i,8}u_8^* + a_{i,9}u_9^* + a_{i,10}u_{10}^*$. (b) Parallel beam projection. The source-detector system can rotate around a center point. The projection direction is determined by the angle β.

In this paper we consider the DT reconstruction problem, represented by a linear system of equations

$$\begin{gathered} A\,u = b, \text{where} \\ A \in \mathbb{R}^{M\times N},\ \ u \in \Lambda^N,\ b \in \mathbb{R}^M,\ \ \Lambda = \{\lambda_1, \lambda_2, ..., \lambda_k\},\ \ \lambda_i \in [0,1],\ \ k \geq 2. \end{gathered} \tag{1}$$

The value of k represents the number of different gray level values. The set Λ is given by the user. The matrix A is a so-called *projection matrix*, whose each row corresponds to one projection ray. The corresponding components of the vector b contain the detected projection values, while the vector u represents the unknown binary image to be reconstructed. The i-th row entries $a_{i,\cdot}$ of A represent the length of the intersection of the pixels and the i-th projection ray passing through them, see Figure 1(a). The projection value measured by a projection ray is calculated as a sum of products of the pixel's intensity and the corresponding length of the projection ray through that pixel. The side length of each pixel is one. Therefore, vertical and horizontal projection rays represent the sum of the gray intensity values of pixels in corresponding columns and rows, respectively. Projections are taken from different directions. For each projection direction, a number of parallel projection rays are taken (parallel beam projection), as shown in Figure 1(b). The distance between two adjacent parallel projection rays can vary depending on the reconstruction problem. We set this distance to be equal to the side length of pixels.

The reconstruction problem means finding the solution image u of the linear system of equations (1), where the projection matrix A and the projection vector b are given. This system is often undetermined ($N > M$), and therefore additional regularization (based on a priori information) is needed for the determination of quality and acceptable solutions.

3 Reconstruction Method Based on the Graph Cuts Method

A directed, weighted graph $G = (X, \rho)$, consists of a set of nodes X and a set of directed edges ρ that connect them. The nodes, in image processing interpretations, mostly correspond to pixels or voxels in 3D. All edges of graph are assigned some weight or cost.

Let $G = (X, \rho)$ be a directed graph with non-negative edge weights that has two special nodes or terminals, the source A and the sink B. An $a-b$-cut (which is referred informally as a cut) $C = A, B$ is a partition of the terminals in X into two disjoint sets A and B so that $a \in A$ and $b \in B$. The cost of the cut is the sum of the costs of all edges that go from A to B:

$$c(A, B) = \sum_{x\in A, y\in B, (x,y)\in\rho} c(x, y).$$

The minimum $a - b$-cut problem is to find a cut C, with the minimum cost among all cuts. Algorithms to solve this problem can be found in [8].

The approach that uses graph cuts for energy minimization has, as a basic technique, the construction of a specialized graph for the energy function to be minimized, so that the minimum cut on the graph also minimizes the energy. The form of the graph depends on the exact form of X and on the number of labels. The minimum cut, in turn, can be computed very efficiently by max flow algorithms.

These methods have been successfully used in the last 20 years for a wide variety of problems, naming image restoration [9, 10], stereo and motion [3, 19], image synthesis [21], image segmentation [7] and medical imaging [6, 18].

3.1 Potts model

The Potts model in graph cuts theory is based on the minimization of the following energy

$$E(d) = \sum_{p \in \mathcal{P}} D(p, d_p) + \sum_{(p,q) \in \mathcal{N}} K_{(p,q)} \cdot T(d_p \neq d_q), \tag{2}$$

where $d = \{d_p | p \in \mathcal{P}\}$ represents the labelling of the image pixels $p \in \mathcal{P}$. By $D(p, d_p)$ we denote the data cost term, where $D(p, d_p)$ is a penalty or cost for assigning a label d_p to a pixel p. $K_{(p,q)}$ is an interaction potential between neighboring pairs p and q, $\mathcal{N}$ is a set of neighboring pairs. Function $T(\cdot)$ is 1 if the condition inside parenthesis is true and 0 otherwise.

3.2 Proposed reconstruction method

Our tomography reconstruction approach is a combination of the graph cuts method and the quadratic iterative minimization method. In the first step, we determine the data cost values for each image pixels. The data cost values are determined as intensity values of the continuous/smooth approximation of the final reconstruction image, obtained as a solution of the following energy-minimization problem

$$\min_{u \in [0,1]^N} E_Q(u) := \|Au - b\|^2. \tag{3}$$

Function E_Q is quadratic type and $\Omega = [0, 1]^N$ is a feasible set. Therefore, the problem (3) is a constrained and quadratic type energy-minimization problem. This minimization problem can be solved by several optimization methods. According to our earlier experiences in similar problems [23, 24, 26] we chose the Spectral Projected Gradient (SPG) optimization algorithm [4] for this task.

For THE application of this algorithm two conditions must be satisfied [4]: i) The objective function has continuous partial derivatives on an open set that contains Ω; ii) The projection function P_Ω of an arbitrary vector onto the set Ω is provided. The objective function in (3) is a multiple differentiable function in $\mathbb{R}^N$, therefore requirement i) is satisfied. The projection P_Ω of an arbitrary vector $u \in \mathbb{R}^n$ onto the set Ω we define as

$$[P_\Omega(u)]_i = \begin{cases} 0, & u_i \leq 0 \\ 1, & u_i \geq 1 \\ u_i, \text{ elsewhere} & \end{cases}, \qquad \text{where } \ i = 1, \ldots, N\,.$$

P_Ω is a projection with respect to the Euclidean distance, i.e.
$P_\Omega(x) = \arg\{\min_{y \in \Omega} d_2(x, y)\}$. Hence, requirement ii) is also satisfied.

The pseudo-code of the SPG is presented in Alg. 1. The reconstruction process, starts with the initial solution u^0, where each pixel intensity is set as 0.5, as the middle of the interval $[0, 1]$. The SPG algorithm combines the non-monotone line search algorithm [13] and the spectral gradient step-length selection [1, 5, 27].

Algorithm 1: SPG optimization algorithm.

$u^0 = [0.5, 0.5, ..., 0.5]^T$; $d^0 = P_\Omega(u^0 - \nabla E_Q(u^0)) - u^0$; $k = 0$;
repeat
  Determine the step-length $\lambda^k > 0$ by a line search approach, see [4];
  $u^{k+1} = u^k + \lambda^k d^k$;
  Calculate the gradient spectral step-length $\theta_{k+1} > 0$, see [4];
  $d^{k+1} = P_\Omega(u^{k+1} - \theta_{k+1} \nabla E_Q(u^{k+1})) - u^{k+1}$; $k = k + 1$;
until $\|u^k - u^{k-1}\|_\infty < 10^{-2}$;
$u^{new} = u^k$;

PH1 PH2 PH3 Shepp-Logan

Fig. 2: Original test images (128x128). Phantoms PH1, PH2 and PH3 contain 3 different gray levels (0,0.5,1), while Shepp-Logan contains 6 different gray levels (0,0.1,0.2,0.3,0.4,1).

In the next step we have to discretize the smooth solution of the problem 3 u, obtained by the SPG algorithm. For this task we apply the graph cuts method based on the Potts model, described in Section 3.1. The energy model in (2) is successfully used in many energy minimization problems with similar energy structure: sum of a data and a regularization/neighboring interaction terms. We mention discrete tomography reconstruction algorithms proposed by Schüle et al. [28, 31] and Lukić et al. [23, 24]. The Potts interaction model (second term in (2)) showed good ability to enhance compactness of the solution (see [8, 12, 29]), in a similar way as the compactness saving regularization terms do in already suggested reconstruction methods [24, 28], which also motivate our choice for application of this model. We note that other interaction models, for example the linear model [8], can also be taken into consideration, but this issue is out of focus of this paper. The data cost term D in (2) is determined using information provided by the smooth solution. More precisely, we define it in the following way

$$\begin{aligned}
&D(p,0) = |u(p) - \lambda_1|,\\
&D(p,1) = |u(p) - \lambda_2|,\\
&D(p,2) = |u(p) - \lambda_3|,\\
&\vdots\\
&D(p,k-1) = |u_p - \lambda_k|,
\end{aligned}$$

where $u(p)$ represents the intensity of a pixel p. The idea is to make data cost small/cheap in the vicinity of the given gray values. The neighbor pairs are defined based on 1-neighboring system, i.e., $(p,q) \in \mathcal{N}$ if the image coordinates of p and q differs for one value only. The interaction potential $K_{(p,q)}$ (see (2)) in our experiments is set as a constant and its value is 1. Now, the energy function in (2) is determined and ready to be minimized. For this task we use the GCO graph cuts based optimization algorithm, introduced in [10] and further analyzed in [8, 11, 20]. The GCO algorithm determines the label values d_p for each pixel p. Each label value is assigned to one predefined gray level in the following way: $d_p = 0 \to \lambda_1$, $d_p = 1 \to \lambda_2$, ..., $d_p = (k-1) \to \lambda_k$. Therefore, the obtained label values also determine intensities of pixels (from the given set of gray levels) in the final (discrete) solution, therefore the reconstruction process is terminated. We denote this method by Graph Cuts Discrete Tomography (GCDT) reconstruction method.

Naturally arises the simplest, but less powerful, way for discretization of the smooth solution u provided as a result of the minimization problem (3). This approach is based on the application of the thresholding function, defined by

$$t(v) = \begin{cases} \lambda_1 & v < \tau_1 \\ \lambda_2 & \tau_1 \leq v < \tau_2 \\ ... & \\ \lambda_k & \tau_{l-1} \leq v \end{cases},$$

where $v \in \mathbb{R}$ and $\tau_l = \frac{\lambda_i + \lambda_{i+1}}{2}$, $\; l = 1, 2, ..., k-1$. The final solution u^r is obtained by application of the thresholding function to the smooth solution u, i.e., $u^r = [t(u_1), t(u_2), t(u_3), ..., t(u_N)]$. We denote this method by TRDT, and use it in experimental work as a control method.

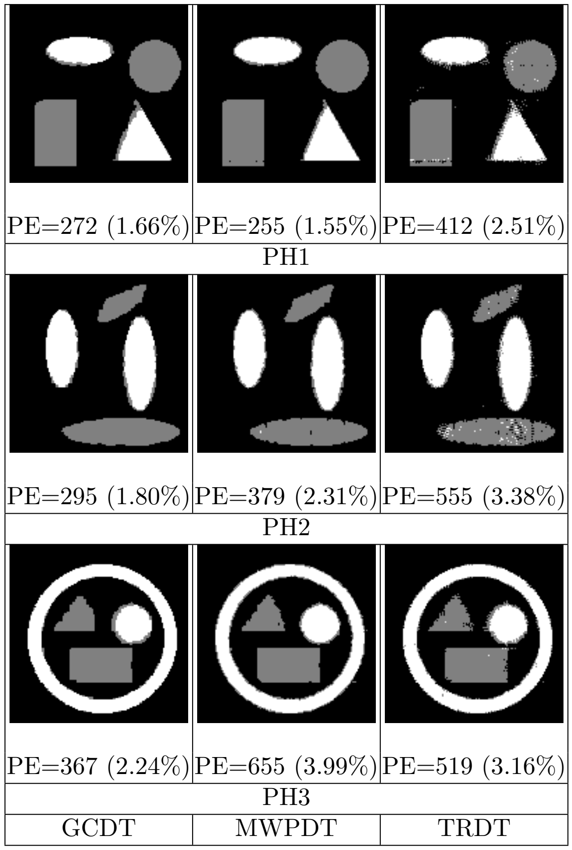


Fig. 3: Reconstructions of the test images using data from 6 projection directions.

4 Experimental Results

In this section we experimentally evaluate the proposed graph cuts based reconstruction method, denoted by GCDT. In the experiments we use 4 test images (phantoms), as originals in reconstructions, presented in Figure 2. Phantoms PH1, PH2 and PH3 contain 3 gray levels, while the well-known Shepp-Logan phantom [17] contains 6 gray levels. We consider reconstructions of these images obtained from different projection directions. A total of 128 parallel rays are taken for each projection direction. In all cases, the projection directions are uniformly selected between 0 and 180 degrees. The obtained results are compared with the results provided by the Multi Well Potential based method (MWPDT) [22], already suggested for multi-level discrete tomography reconstruction, and with the simple method based on the classical thresholding, denoted by TRDT.

Related to the Shepp-Logan test image, the DART method, proposed in [2], is also included into the evaluation process.

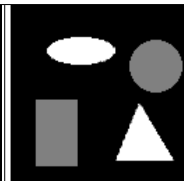
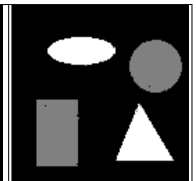
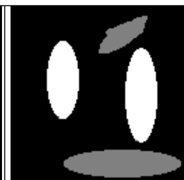
		
PE=5 (0.03%)	PE=35 (0.21%)	PE=28 (0.17%)
PH1		
		
PE=6 (0.03%)	PE=5 (0.03%)	PE=11 (0.06%)
PH2		
		
PE=9 (0.05%)	PE=174 (1.06%)	PE=41 (0.25%)
PH3		
GCDT	MWPDT	TRDT

Fig. 4: Reconstructions of the test images using data from 15 projection directions.

In the evaluation process, we analyze the quality of the reconstructions and required running times. The quality of the reconstructions are expressed by the pixel error (PE), i.e., the absolute number of the misclassified pixels, and by the misclassification rate ($m.r.$), i.e., the pixel error measure relative to the total number of image pixels. Also, as a qualitative error measure, we consider the projection error, defined by $PRE = \|Au^r - b\|$, where u^r represents the reconstructed image. This error expresses the accordance of the reconstruction with the given projection data.

In Table 2 we present pixel errors for reconstructions of three phantom images (PH1, PH2 and PH3) obtained from a different number of projections by three different methods (MWPDT, GCDT and TRDT). In Figures 3 and 4 reconstructions from 6 and 15 projection directions are presented. From a total of 12 different reconstruction problems, GCT method provided the best results in 10 cases, while in 2 cases the dominant was the MWPDT method. We emphasize that the results of the GCT method, in cases when they are the best, are significantly better, at least by 50%, compared with other results. Table 3 presents the obtained projection errors (PRE) and the needed running times in these experiments. Regrading the PRE values, the proposed GCT method dominated in 8 cases, while MWPDT in 4 cases.

Fig. 5: Reconstructions of the Shepp-Logan test images by the proposed GCDT method.

All reconstruction methods (MWPDT, GCDT and TRDT) are implemented completely in Matlab. The best running times in all of the experiments was achieved by the MWPDT method (see Table 3). GCDT and TRDT methods uses the smooth solution/reconstruction as a first step, before the "binarization process" starts by GCO graph cuts optimization [10]. This smooth solution is achieved as a final termination, with high precision. This process, because of the high precision, requires significantly higher number of iterations than is needed for MWPDT method in total, resulting in a greater consumption of time.

Table 1: Experimental results for Shepp-Logan image, using three different reconstruction methods. The abbreviation m.r. indicates misclassification rate and d indicates the number of projections.

	d	TRDT (m.r. %)	DART (m.r. %)	GCDT (m.r. %)
Shepp-Logan	12	12.74	14.21	5.72
	15	10.44	8.44	3.17
	18	10.03	2.56	2.14

Reconstruction results of the well-known Shepp-Logan [17] phantom image is presented in Table 1. This phantom is considered to be one of the most complex, containing 6 different gray levels. We compare the results obtained by the three different reconstruction methods: TRDT, DART and GCDT. The results for DART are taken from [2]. The projection data is acquired from 12, 15 and 18 projection directions. The GCDT method provides the best results in all cases (smallest m.r. values).

Table 2: Experimental results for PH1, PH2 and PH3 images, using three different reconstruction methods. The abbreviation d indicates the number of projections.

	PH1				PH2				PH3			
d	6	9	12	15	6	9	12	15	6	9	12	15
(PE)	255	159	59	35	379	242	56	5	655	456	275	174
MWP (m.r. %)	1.55	0.97	0.36	0.21	2.31	1.47	0.34	0.03	3.99	2.78	1.67	1.06
(PE)	412	175	48	28	555	290	37	11	412	301	101	41
TRDT (m.r. %)	2.51	1.06	0.29	0.17	3.38	1.77	0.22	0.06	2.51	1.83	0.61	0.25
(PE)	272	69	8	5	295	121	15	6	272	116	20	9
GCDT (m.r. %)	1.66	0.42	0.04	0.03	1.80	0.73	0.09	0.03	1.66	0.70	0.12	0.05

Summarizing the results obtained by the total of the 15 analyzed reconstruction tasks, see Tables 1 and 2, the quality of the reconstruction, indicated by m.r., for the proposed GCDT method was the best in 13 cases, i.e., in 87% of the cases. According to these results, we conclude that the experiments confirm the capability of the proposed method to provide high quality reconstructions.

Table 3: Experimental results for PH1, PH2 and PH3 images, using three different reconstruction methods. The abbreviation e.t. means elapsed time in minutes and d indicates the number of projections.

	PH1				PH2				PH3			
d	6	9	12	15	6	9	12	15	6	9	12	15
(PRE)	14.70	12.19	9.96	9.08	15.32	15.50	9.68	3.09	19.83	18.77	18.80	16.43
MWPDT (e.t.)	1.76	2.63	3.17	4.06	1.78	2.82	3.21	3.06	2.19	2.87	4.30	4.66
(PRE)	18.66	14.72	10.61	8.87	19.01	19.09	8.89	6.23	23.64	17.87	13.66	10.61
TRDT (e.t.)	7.73	12.58	14.55	17.77	5.44	10.90	12.67	15.74	7.28	11.07	13.39	16.00
(PRE)	23.24	11.12	6.52	4.39	18.31	13.94	7.10	4.57	25.87	14.96	7.59	5.60
GCDT (e.t.)	7.73	12.58	14.55	17.77	5.45	10.91	12.67	15.74	7.29	11.07	13.40	16.01

5 Conclusions

In this paper, a new energy-minimization based reconstruction method for multi-level tomography is proposed. It combines a gradient based method, with the graph cuts optimization method. Experiments show advantages of the proposed method in comparison with three formerly published reconstruction methods. Based on the obtained experimental results and analysis presented in this paper, we conclude that the combination of a gradient based method with graph cuts optimization method is suitable for providing high quality reconstructions.

Acknowledgement

Tibor Lukić acknowledges the Ministry of Education and Sciences of the R. of Serbia for support via projects OI-174008 and III-44006. Marina Marčeta acknowledges the Ministry of Education and Sciences of the R. of Serbia for support via project OI-174008.

References

1. Barzilai, J., Borwein, J.M.: Two point step size gradient methods. IMA Journal of Numerical Analysis 8, 141–148 (1988)
2. Batenburg, K.J., Sijbers, J.: DART: A Fast Heuristic Algebraic Reconstruction Algorithm for Discrete Tomography. In: Proceedings of International Conference on Image Processing (ICIP). pp. 133–136 (2007)
3. Birchfield, S., Tomasi, C.: Multiway cut for stereo and motion with slanted surfaces. Proc. International Conf. Computer Vision pp. 489–495 (1999)
4. Birgin, E.G., Martínez, J.M., Raydan, M.: Algorithm: 813: SPG - software for convex-constrained optimization. ACM Transactions on Mathematical Software 27, 340–349 (2001)
5. Birgin, E., Martínez, J.: Spectral Conjugate Gradient Method for Unconstrained Optimization. Appl. Math. Optimization 43, 117–128 (2001)
6. Boykov, Y., Jolly, M.P.: Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images. In: Proc. International Conf. Computer Vision. pp. 105–112 (2001)
7. Boykov, Y., Kolmogorov, V.: Computing Geodesics and Minimal Surfaces via Graph Cuts. In: Proc. International Conf. Computer Vision. pp. 26–33 (2003)
8. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Transactions on PAMI 26(9), 1124–1137 (2004)
9. Boykov, Y., Veksler, O., Zabih, R.: Markov random fields with efficient approximations. In: Proceedings of IEEE Conf. Computer Vision and Pattern Recognition. pp. 648–655 (1998)
10. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Transactions on PAMI 23(11), 1222–1239 (2001)
11. Delong, A., Osokin, A., Isack, H.N., Boykov, Y.: Fast approximate energy minimization with label costs. Proc. IEEE Conf. Computer Vision and Pattern Recognition 96(1), 1–27 (2010)

12. Greig, D., Porteous, B., Seheult, A.: Exact maximum a posteriori estimation for binary images. Journal of the Royal Statistical Society 51(2), 271–279 (1989)
13. Grippo, L., Lampariello, F., Lucidi, S.: A Nonmonotone Line Search Technique for Newton's Method. SIAM J. Numer. Anal. 23, 707–716 (1986)
14. Herman, G.T.: Fundamentals of computerized tomography: Image reconstruction from projection (2nd ed.). Springer-Verlag (2009)
15. Herman, G.T., Kuba, A.: Discrete Tomography: Foundations, Algorithms and Applications. Birkhäuser (1999)
16. Herman, G.T., Kuba, A.: Advances in Discrete Tomography and Its Applications. Birkhäuser (2006)
17. Kak, A.C., Slaney, M.: Principles of Computerized Tomographic Imaging. SIAM (2001)
18. Kim, J., Zabih, R.: Automatic Segmentation of Contrast- Enhanced Image Sequences. In: Proc. International Conf. Computer Vision. pp. 502–509 (2003)
19. Kolmogorov, V., Zabih, R.: Visual Correspondence with Occlusions Using Graph Cuts. In: Proc. International Conf. Computer Vision. pp. 508–515 (2001)
20. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts? IEEE Transactions on PAMI 26(2), 147–159 (2004)
21. Kwatra, V., Schoedl, A., Essa, I., Turk, G., Bobick, A.: Graphcut Textures: Image and Video Synthesis Using Graph Cuts. In: Proc. SIGGRAPH 2003. pp. 277–286. ACM Trans. Graphics (2003)
22. Lukić, T.: Discrete Tomography Reconstruction Based on the Multi-well Potential. In: Proceedings of Combinatorial Image Analysis - 14th International Workshop (IWCIA). LNCS, vol. 6636, pp. 335–345. Springer-Verlag, Madrid, Spain (2011)
23. Lukić, T., Balázs, P.: Binary tomography reconstruction based on shape orientation. Pattern Recognition Letters 79, 18–24 (2016)
24. Lukić, T., Nagy, B.: Deterministic discrete tomography reconstruction method for images on triangular grid. Pattern Recognition Letters 49, 11–16 (2014)
25. M. Zisler, S. Petra, C.S.C.S.: Discrete Tomography by Continuous Multilabeling Subject to Projection Constraints. In: Proc. of 38th German Conference on Pattern Recognition. pp. 261–272. LNCS, Springer-Verlag (2016)
26. Nagy, B., Lukić, T.: Dense projection tomography on the triangular tiling. Fundamenta Informaticae 145, 125–141 (2016)
27. Raydan, M.: The Barzilai and Browein Gradient Method for the Large Scale Unconstrained Minimization Problem. SIAM J. of Optimization 7, 26–33 (1997)
28. Schüle, T., Schnörr, C., Weber, S., Hornegger, J.: Discrete Tomography by Convex-concave Regularization and D.C. Programming. Discrete Appl. Math. 151, 229–243 (2005)
29. Snow, D., Viola, P., Zabih, R.: Exact voxel occupacy with graph cuts. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. vol. 1, pp. 345–352 (2000)
30. Varga, L., Balázs, P., Nagy, A.: An energy minimization reconstruction algorithm for multivalued discrete tomography. In: Proceedings of 3rd International Symposium on Computational Modeling of Objects Represented in Images. pp. 179–185. Taylor & Francis, Rome, Italy (2012)
31. Weber, S., Nagy, A., Schüle, T., Schnörr, C., Kuba, A.: A benchmark evaluation of large-scale optimization approaches to binary tomography. In: Proc. of 13th International Conference on Discrete Geometry for Computer Imagery (DGCI). LNCS, vol. 4245, pp. 146–156. Springer-Verlag, Szeged, Hungary (2006)

План третмана података

Назив пројекта/истраживања
Regularized Models for Tomographic Image Reconstruction, (наслов на српском језику: Модели са регуларизацијом за реконструкцију слика у томографији)
Назив институције/институција у оквиру којих се спроводи истраживање
а) Факултет техничких наука, Универзитет у Новом Саду б) в)
Назив програма у оквиру ког се реализује истраживање
Истраживање се врши у оквиру израде докторске дисертације на студијском програму Математика у техници.
1. Опис података
1.1 Врста студије *Укратко описати тип студије у оквиру које се подаци прикупљају* **У овој студији нису прикупљани подаци.** 1.2 Врсте података а) квантитативни б) квалитативни 1.3. Начин прикупљања података а) анкете, упитници, тестови б) клиничке процене, медицински записи, електронски здравствени записи в) генотипови: навести врсту __________________________________ г) административни подаци: навести врсту ________________________ д) узорци ткива: навести врсту___________________________________ ђ) снимци, фотографије: навести врсту______________________________ е) текст, навести врсту ____ ж) мапа, навести врсту _______________________________________ з) остало: описати ____

1.3 Формат података, употребљене скале, количина података

1.3.1 Употребљени софтвер и формат датотеке:

а) Excel фајл, датотека ____________________

b) SPSS фајл, датотека ____________________

c) PDF фајл, датотека _____________________

d) Текст фајл, датотека ____________________

e) JPG фајл, датотека _____________________

f) Остало, датотека ______________________

1.3.2. Број записа (код квантитативних података)

а) број варијабли ____

б) број мерења (испитаника, процена, снимака и сл.) ________________________

1.3.3. Поновљена мерења

а) да

б) не

Уколико је одговор да, одговорити на следећа питања:

а) временски размак између поновљених мера је _________________________________

б) варијабле које се више пута мере односе се на ___________________________________

в) нове верзије фајлова који садрже поновљена мерења су именоване као _____________

Напомене: ____

Да ли формати и софтвер омогућавају дељење и дугорочну валидност података?

а) Да

б) Не

Ако је одговор не, образложити ____

____

2. Прикупљање података

2.1 Методологија за прикупљање/генерисање података

2.1.1. У оквиру ког истраживачког нацрта су подаци прикупљени?

а) експеримент, навести тип ____

б) корелационо истраживање, навести тип ____________________________________

ц) анализа текста, навести тип ____

д) остало, навести шта ____

2.1.2 Навести врсте мерних инструмената или стандарде података специфичних за одређену научну дисциплину (ако постоје).

____

____

2.2 Квалитет података и стандарди

2.2.1. Третман недостајућих података

а) Да ли матрица садржи недостајуће податке? Да Не

Ако је одговор да, одговорити на следећа питања:

а) Колики је број недостајућих података? ____________________________

б) Да ли се кориснику матрице препоручује замена недостајућих података? Да Не

в) Ако је одговор да, навести сугестије за третман замене недостајућих података

____

2.2.2. На који начин је контролисан квалитет података? Описати

____

____

2.2.3. На који начин је извршена контрола уноса података у матрицу?

____

____

3. Третман података и пратећа документација

3.1. Третман и чување података

3.1.1. Подаци ће бити депоновани у ___________________________________ репозиторијум.

3.1.2. URL адреса ____

3.1.3. DOI ____

3.1.4. Да ли ће подаци бити у отвореном приступу?

а) *Да*

б) *Да, али после ембарга који ће трајати до* ___________________________________

в) *Не*

Ако је одговор не, навести разлог ____

3.1.5. Подаци неће бити депоновани у репозиторијум, али ће бити чувани.

Образложење

____

____

3.2 Метаподаци и документација података

3.2.1. Који стандард за метаподатке ће бити примењен? ___________________________________

3.2.1. Навести метаподатке на основу којих су подаци депоновани у репозиторијум.

____

____

Ако је потребно, навести методе које се користе за преузимање података, аналитичке и *процедуралне информације, њихово кодирање, детаљне описе варијабли, записа итд.*

____

____

____

____

3.3 Стратегија и стандарди за чување података

3.3.1. До ког периода ће подаци бити чувани у репозиторијуму? ________________________

3.3.2. Да ли ће подаци бити депоновани под шифром? Да Не

3.3.3. Да ли ће шифра бити доступна одређеном кругу истраживача? Да Не

3.3.4. Да ли се подаци морају уклонити из отвореног приступа после извесног времена?

Да Не

Образложити

____

____

4. Безбедност података и заштита поверљивих информација

Овај одељак МОРА бити попуњен ако ваши подаци укључују личне податке који се односе на учеснике у истраживању. За друга истраживања треба такође размотрити заштиту и сигурност података.

4.1 Формални стандарди за сигурност информација/података

Истраживачи који спроводе испитивања с људима морају да се придржавају Закона о заштити података о личности *(https://www.paragraf.rs/propisi/zakon_o_zastiti_podataka_o_licnosti.html)* и одговарајућег институционалног кодекса о академском интегритету.

4.1.2. Да ли је истраживање одобрено од стране етичке комисије? Да Не

Ако је одговор Да, навести датум и назив етичке комисије која је одобрила истраживање

____

4.1.2. Да ли подаци укључују личне податке учесника у истраживању? Да Не

Ако је одговор да, наведите на који начин сте осигурали поверљивост и сигурност информација везаних за испитанике:

а) Подаци нису у отвореном приступу

б) Подаци су анонимизирани

ц) Остало, навести шта

____

____

5. Доступност података

5.1. Подаци ће бити

а) јавно доступни

б) доступни само уском кругу истраживача у одређеној научној области

ц) затворени

Ако су подаци доступни само уском кругу истраживача, навести под којим условима могу да их користе:

____

____

Ако су подаци доступни само уском кругу истраживача, навести на који начин могу приступити подацима:

____

____

5.4. Навести лиценцу под којом ће прикупљени подаци бити архивирани.

____

6. Улоге и одговорност

6.1. Навести име и презиме и мејл адресу власника (аутора) података

____

6.2. Навести име и презиме и мејл адресу особе која одржава матрицу с подацима

___

6.3. Навести име и презиме и мејл адресу особе која омогућује приступ подацима другим истраживачима

___