



University of Szeged
Faculty of Science and Informatics
PhD School in Computer Science

THESIS

Geometric and Visual Reconstruction of Binary Shapes

Author:

József Németh

Supervisors:

János Csirik, DSc

Zoltán Kató, DSc

A DISSERTATION SUBMITTED FOR THE DEGREE OF DOCTOR OF
PHILOSOPHY OF THE UNIVERSITY OF SZEGED

Szeged, 2015

Acknowledgements

First of all, I would like to thank János Csirik for his guidance and support and for that he made it possible for me to carry out my doctoral studies, and Zoltán Kató who introduced me to the wonderful world of scientific research. I am also very thankful to Csaba Domokos for his invaluable help and for sharing his knowledge and experiences.

I am grateful to Péter Balázs, László Varga, Márton Balaskó, Antal Nagy, László Nyúl, Attila Tanács and Kálmán Palágyi who gave me some good advice and also often provided me program codes and data for my experiments. I also appreciate the friendship and support of Erika Griechisch, Norbert Hantos, Gábor Németh, András London and Tamás Németh.

Last but not least, I am thankful to my parents who always supported and encouraged my education, to my brother Ádám and to my girlfriend Eszter for their love and support.

My work was supported by the Doctoral School of the University of Szeged, Foundation for Education and Research in Informatics.

Contents

1	Introduction	1
1.1	Binary Image Processing	3
1.1.1	Discrete Tomographic Reconstruction	3
1.1.2	Binary Image Deconvolution	4
1.1.3	Binary Shape Registration	5
1.2	Summary by Chapters	7
1.3	Summary of the Author's Contributions	8
2	Discrete Tomography with Unknown Intensity Levels	11
2.1	Basic Concepts and Notations	12
2.2	The Reconstruction Method	13
2.2.1	Optimization	15
2.2.2	Estimation of the Discretization Parameter	19
2.3	Gray-Level Independence and Implementational Details	21
2.4	Numerical Experiments	24
2.4.1	Comparison to the PDM-DART Method	26
2.4.2	Limited Angular Range Experiments	28
2.4.3	Comparison to a Convex Programming Approach	30
2.4.4	Experiments with Valid Directon Sets of Uniqueness	32
2.4.5	Real Data Experiments	33
2.5	Discussion	35
2.5.1	Complexity	35
2.6	Summary	36
2.7	Appendix: The Lipschitz-Continuity of the Discretization Term	36
3	Binary Shape Deconvolution Using Discrete Tomography	41
3.1	The Restoration Method	41
3.1.1	Deconvolution of the Projections	43
3.1.2	Binary Tomographic Reconstruction	44
3.2	Experiments and Comparison	47
3.2.1	Experiments on Real Out of Focus Images	50

3.3	Summary	50
4	Nonlinear Registration of Binary Shapes	53
4.1	A Parametric Reconstruction Framework	53
4.1.1	Construction of the System of Equations	56
4.2	The Studied Deformation Models	57
4.2.1	Planar Homography	57
4.2.2	The Taylor Series Expansion of Planar Homography	59
4.2.3	Polynomial Transformations	60
4.2.4	Thin Plate Spline	61
4.3	Implementational Details	62
4.3.1	Numerical Implementation	63
4.3.2	Solution and Complexity	64
4.4	Experiments on Synthetic Images	65
4.4.1	Comparison of Different ω Functions	66
4.4.2	Comparative Tests	68
4.5	Experiments on Real Images	73
4.5.1	Planar Homography	73
4.5.2	Thin Plate Spline	77
4.5.3	An Industrial Application	78
4.6	Summary	79
5	Affine Invariants Based Projective Registration of Binary Shapes	81
5.1	The Registration Method	81
5.1.1	Decomposition of the Transformation	82
5.1.2	Estimation of the Perspective Distortion	83
5.1.3	Estimation of the Affine Component	84
5.2	Implementational Issues	85
5.3	Experiments	86
5.4	Discussion	87
5.5	Summary	88
	Appendices	89
A	Summary in English	91
B	Summary in Hungarian	95

List of Figures

2.1	The discretization model and typical geometry types	12
2.2	Demonstration of how the minimization of the discretization term en- forces binary sets	15
2.3	Demonstration of the convergence of the proposed method	16
2.4	Estimation of discretization parameter $\hat{\alpha}$	20
2.5	Four of the phantom images used to evaluate the performance of the proposed method.	23
2.6	Example noiseless and noisy projections for different incident beam in- tensities.	24
2.7	Some reconstruction results of the PDM-DART method and the pro- posed method	27
2.8	Limited angular range experiments	29
2.9	Comparison to a convex programming approach	31
2.10	A further comparison to a convex programming approach	32
2.11	Reconstruction of a gas pressure regulator	33
2.12	Gas pressure regulator: 3D visualization of the reconstruction results. .	34
3.1	The degradation model and the basic idea of the proposed method . .	42
3.2	A typical L-curve and its curvature	43
3.3	The evolution of the estimation of the projection vector during the iteration.	44
3.4	Associated network for a 3×3 case	46
3.5	Reconstruction results for different $s \in \mathcal{S}$ scale values.	47
3.6	Example reconstruction results on synthetic images	49
3.7	Results on letters extracted from out-of-focus document images.	50
4.1	The registration problem	54
4.2	The effect of various ω functions	56
4.3	Example deformation fields	57
4.4	Example perspective distortion	58
4.5	Projective transformations	59
4.6	Example polynomial transformation.	60

4.7	Example thin plate spline transformation.	61
4.8	Coverage of transformed shapes during the minimization process	62
4.9	Plots of tested $\{\omega_i\}$ function sets.	66
4.10	<i>Planar homographies</i> : Example images from the synthetic data set and registration results	69
4.11	<i>Polynomial transformations</i> and <i>Thin plate spline transformations</i> : Ex- ample images from the synthetic data set and registration result	70
4.12	Sample observations with various degradations for testing robustness. .	72
4.13	Registration results on traffic signs	74
4.14	Registration results on hip prosthesis X-ray images	75
4.15	Registration results on traffic signs	76
4.16	Sample images from the MNIST dataset and registration results	77
4.17	Registration results of printed signs	78
5.1	The registration process	85
5.2	Example images from the synthetic data set and registration results . .	87
5.3	Registration results on traffic signs	88

List of Tables

1.1	Correspondence between the thesis points and the publications.	8
2.1	The rNMP measures (%) and runtime performances (sec.) provided by the PDM-DART method and the proposed method	25
2.2	The rNMP measures (%) and runtime performances (sec.) provided by the PDM-DART method and the proposed method	26
3.1	Test results and comparison on the synthetic dataset	48
4.1	Comparison of the tested $\{\omega_i\}$ sets	67
4.2	<i>Planar homography</i> : Comparative tests of the proposed method on the synthetic dataset	69
4.3	<i>Polynomial transformations</i> and <i>Thin plate spline transformations</i> : Registration results	71
4.4	Test of the robustness against segmentation errors	73
4.5	Comparative results on the MNIST database	77
5.1	Test results on the synthetic dataset	86
A.1	Correspondence between the thesis points and the publications.	94
B.1	A tézispontok és a szerző publikációinak kapcsolata.	98

List of Algorithms

1	The tomographic reconstruction method	17
2	Minimization with backtracking line search	18
3	Deconvolution of a projection vector	45
4	Shape restoration using discrete tomography	47
5	Pseudo code of the nonlinear registration framework	65
6	Pseudo code of the affine invariants based projective registration method	84

Chapter 1

Introduction

Binary shapes play important role in the field of image processing, due to that

1. the images of many real world objects are basically binary shapes, e.g., text characters, traffic signs, bones or implants on X-ray images, etc.,
2. in most of the image processing applications at some point of the processing pipeline the images are binarized (*i.e.*, segmented).

In many cases, the additional information that we are working with a binary valued image limits the number of possible solutions (the search space) thus it can help to obtain more accurate results. For example, the linear inverse problems like tomography and deconvolution are sometimes under-determined, and have many possible solutions. In many other cases, these problems are over-determined and inconsistent due the the noisy input data. Knowing that we are seeking a binary image restricts the search space, and as a result, helps to find the (most accurate) solution. For example in the case of tomography, binary images can be usually reconstructed accurately only from a few projections.

In many other cases, however, the lack of rich intensity information makes it difficult to deal with these binary images. For example image registration techniques commonly work with previously established point pairs and these correspondences are usually obtained based on the intensity patterns around the points. Thus in the case of binary images these methods can not provide appropriate point correspondences. On the other hand, in such cases we do not need to deal with the intensity change between the images. Therefore many techniques have been presented previously to register binary images using statistics of the point coordinates of the shapes.

This work is a summary of the author's research results in the visual reconstruction and the geometric registration of binary shapes. Most of the presented results and techniques are related to statistical moments. In the analysis of binary shapes, these

statistics have many advantages. They give an easy-to-compute, compact, and robust representation of an image or a comparable measure of one property of the shape. Invariant moments, for example, keep their value when the shape undergoes in some kind of deformation, thus they can be considered as similarity measures and can be used for shape description or matching. Moreover, they also provide an efficient tool for binary image registration.

For example, in Chapter 2, we use a kurtosis-like fourth-order statistics to enforce binary solutions in a discrete tomography reconstruction method. This discretization term plays two roles in the algorithm. First, it gives a measure of the binarity of the intermediate solutions. At the same time, it is used to drive the reconstruction towards a binary one, by gradually minimizing it as a functional of the solution. This can be done using simple gradient based optimization technique, since this functional is continuously differentiable. The advantage of the method is that it does not require the estimation of the two intensities of the images. Instead, it estimates the mid-level of the intensities using the discretization term.

A binary image registration method is proposed in Chapter 4 which involves covariant function sets. These functions are integrated over the foreground (object) regions, that results in generalized shape moments of the shapes. More specifically, using power functions we obtain the geometric shape moments. These functions are covariant with respect to the transformation, *i.e.*, their values vary with the transformation, thus each function constraints the transformation parameters. These constraints are formularized as a system of nonlinear equation, which is solved by standard nonlinear optimizer and the solution gives the parameters of the aligning transformation.

In Chapter 5, different kind of statistics are used for projective registration of binary images. In the first step, to estimate the perspective component of the transformation, affine moment invariants (AMIs) are used. These shape statistics are invariant to the affine transformations (translation, rotation, scaling and shearing), while generally covariant to perspective distortions, thus they can be used to set up a system of equation, and its solution gives the perspective parameters. In the second step, using third-order geometric moments we obtain a system of third-order equations, and its solution gives the parameters of the affine component of the transformation.

The problems addressed in this work involve different mathematical tools. In each chapter we used notation systems that we found to be appropriate to discuss the given topic in a simple and comprehensible way. Nevertheless, it resulted in somewhat inconsistent notations between the parts. Therefore we build up the notation system in each chapter as we introduce the basic concepts of the given topic.

1.1 Binary Image Processing

While intensity levels or colors allow better representation of a real-world scene, in many cases the silhouettes provide enough information *e.g.* to recognize or align the objects. Moreover binary images are usually easier to store, process and analyse than grayscale images, and therefore many algorithms have been developed to deal specifically with binary images. In this section, we briefly introduce the topics that we address in this work and overview the related literatures.

1.1.1 Discrete Tomographic Reconstruction

Tomography focuses on the reconstruction of images from a set of their projections taken from different angles. The projections are usually taken by exposing the object to some radiation and measuring the loss of the particles on its other side [28, 49, 48]. This way, the interior structure of the object can be examined. Tomography is of high importance in different areas, *e.g.*, in medicine, biology, while it also can be applied in industry. The theory of the tomographic reconstruction has been heavily studied, and there are many algorithms that can provide accurate results when a sufficiently high amount of projection data is available. Filtered Back Projection [49], Simultaneous Algebraic Reconstruction Technique (SART [3]) and Simultaneous Iterative Reconstruction Technique (SIRT [102, 39]) are some of the most widely used methods.

In the case of discrete tomography [46, 47], it is assumed that the image contains only a limited number of different intensity levels, *i.e.*, the object consists of only a few different materials. This additional information limits the number of possible solutions, thus helps to obtain high quality reconstructions even from few projections. Discrete tomography can also be considered as a combinatorial problem. The combinatorial properties of binary sets were addressed in [88]. In his work, Ryser introduced an efficient method for the reconstruction of binary images from vertical and horizontal projections.

The reconstruction of binary images from vertical and horizontal projections can be traced back to the network flow problem, which can be solved in polynomial time. This efficient solvability can also be generalized to more projections [15, 9, 68]. These methods iteratively update the solution in each iteration step using only two projections and the result of the previous iteration. Another network flow based approach described in [37], Gesù *et al.* first generates a set of initial reconstructions from different projection pairs, then an evolutionary algorithm is used to find an optimal solution.

Discrete tomographic reconstruction problems are often under-determined due to the low number of projections and inconsistent due to measurement errors [103]. However, there can be found exact solutions in the case of special classes of binary im-

ages [19, 26, 69]. Furthermore, the projection angle selection also affects the quality of the reconstructions [107, 98]. In [106], different projection angle selection strategies have been examined. Since the reconstruction of binary images from more than two projections is proved to be NP-hard [36], the researchers examined approximation approaches to optimize the objective functionals, e.g., evolutionary algorithms [8, 6, 37] and simulated annealing [71].

In [25], a heuristic for binary tomography, called *Binary Steering* has been proposed. The basic idea is to gradually enforce binary solutions between the consecutive steps of a continuous reconstruction algorithm. In [92], Schüle *et al.* proposed a less heuristic binary steering approach based on D.C. (difference of convex functions) programming. The D.C. method can provide accurate reconstructions even from few projections. The objective functional prescribes that the image should satisfy the projection data, but also enforces binary solutions at the same time. The reconstruction is performed by a primal-dual sub-gradient algorithm, applied to the convex-concave decomposition of the functional. The authors also extended their work to multi-valued objects [93]. Another extension of this approach for multi-valued discrete tomography has been proposed in [108] with a modified energy function and an appropriate optimization strategy. Some further approaches can be found in [111] and [63]. In [12], Batenburg *et al.* introduced the Discrete Algebraic Reconstruction Technique (DART). The main advantage of this method is that it is not restricted to binary images. The algorithm first segments the initial continuous reconstruction, then iteratively improves the result by updating only the boundary pixels. Experiments showed that the algorithm provides accurate solutions. The method has also been successfully applied on three-dimensional data [13].

Most of the discrete tomography algorithms assume that the intensity values are accurately known [12, 92, 63, 26, 19, 71, 87]. However, in practice, these gray-levels are usually unknown and can not be estimated in a straightforward way. Even when the attenuation coefficients (*i.e.*, the densities of the materials) are known, the imaging device still has to be calibrated, and these calibration parameters usually change over time due to the usage of the equipment. The non-uniqueness of the binary images in the case of absorbed projections was examined in [53].

1.1.2 Binary Image Deconvolution

Enhancement of degraded images is a crucial step in many applications. Such distortions usually come in many forms such as motion blur, camera misfocus and noise. Misfocus and motion blur can be described as the convolution of the image with a so called blurring function (filters while noise usually follows a specific distribution as a

good approximation.

Many techniques can be found in the literature for digital image enhancement including simple methods and more sophisticated algorithms. Unsharp masking [91] is one of the widely used algorithms to enhance image contrast. The basic idea of this method can from photography in which it is used since the first half of the 20th century to increase the sharpness of images. Its digital version is applied in many image processing softwares. Inverse filtering (or deconvolution) of images is another basic idea for recovering an image that is blurred by a known low-pass filter. It tries approximately inverting the process that caused the image to be blurred. Researchers have been studying deconvolution methods for several decades, and have approached the problem from different directions.

Deconvolution is linear inverse problem and thus in the case of noisy images it requires regularization. This can be handled using the Wiener filter [112] which can be used when the point-spread-function (PSF, the filter which was used to blur the image) and the signal-to-noise ratio (SNR) are known. Blind deconvolution [56] is a technique for those cases when the PSF is unknown. In this method the PSF and the images is updated in each iteration step, and thus its convergence depends on how accurate is the initial estimation of the PSF.

The deconvolution of binary images like document images has been addressed by several researchers (see [51], [55], [59] and [95]). In most of these algorithms, the problem is traced back to the solution of a system of equations which then solved by some iterative optimization technique. Although these methods are intensely studied, they are still time-consuming and require good parameter settings which makes the applications of them difficult in practice. The size of the point spread function is usually also limited.

1.1.3 Binary Shape Registration

In most of the image processing applications a key step is the registration of images, *i.e.*, the estimation of the transformation which aligns one image to the other (see [119] for a good survey). The overlapped images can be then combined or compared. Most of the techniques assume linear transformation (*i.e.*, rigid-body, similarity or affine) between the shapes, however in many applications nonlinear deformations [118] (*e.g.*, projective, polynomial, elastic) need to be considered. For example, the estimation of the parameters of a projective transformation (also known as planar homography) between two views of the same planar object has a fundamental importance in computer vision. Other typical application areas include visual inspection [109], object matching [16] and medical image analysis [27]. Good surveys can be found in [119, 64].

Classical landmark based (or correspondence based) methods usually trace back the problem into the solution of a system of linear equations set up using the coordinates of point pairs [45]. These point pairs are usually established by matching the intensity value patterns around the points [62]. On the other hand featureless methods estimate the transformation parameters directly from image intensity values over corresponding regions [65].

In many cases, however, the images do not contain sufficient variety of graylevel values (e.g., images of traffic signs or letterings), or suffered from intensity value distortions (e.g., X-ray images). Although there are some time consuming methods to cope with brightness change across image pairs [50], these conditions make the classical brightness-based methods unreliable. In [35], Francos *et al.* propose a method for the estimation of a homeomorphism between graylevel images. They showed how to transform the problem into the solution of a linear system of equations, however they assumed that the intensity values differ only by a zero mean Gaussian noise.

When the segmentations are available it is reasonable to solve the registration problem using the binary versions of the images [96, 40]. Most of the current approaches are restricted to affine transformations. For example Domokos *et al.* showed that it is possible to trace back the affine matching problem to an exactly solvable polynomial system of equations [30]. Affine moments and invariants can also be used to recover linear transformation [100]. In [116] Yezzi *et al.* proposed a variational framework that uses active contours to simultaneously segment and register features from multiple images and applied it to medical image registration, where 2D and 3D rigid body transformations are considered. A statistics-based technique is described in [97] for the registration of edge-detected images, which includes a well-defined measure of the statistical confidence associated with the solution. Edge pixel matching is used to determine the "best" translations. The statistical method called the McNemar test is used to determine which candidate solutions are not significantly worse than the best ones. This way confidence regions of the registration parameters can be determined. This method, however, is limited to solving for 2D translations only [97].

Most of the nonlinear registration techniques use point correspondences [40, 114, 16]. Although there are robust keypoint detectors like SIFT [62] or SURF [41], these can not be used for binary registration due to the lack of rich intensity patterns. Belongie *et al.* proposed a nonlinear shape matching algorithm in [16]. The method first establish point correspondences between the binary shapes using a novel similarity metric, called *shape context*, which consists in constructing a log-polar histogram of surrounding edge pixels. Then it uses the generic thin plate spline model to align the shapes. The main advantage of this approach is that it does not require intensity information. In [40], Guo *et al.* introduced a method to estimate diffeomorphic

distortions between shapes, where point correspondences between the boundary points of the shapes are estimated using simulated annealing.

In [1], Bronsetin *et al.* proposed a similarity metric for non-rigid deformable shapes and they extended it to partial similarity. It has been showed that their similarity metric can be used to solve the correspondence problem. In [114], Wörz *et al.* proposed a novel approximation approach for landmark-based elastic registration using Gaussian elastic body splines. Other methods use variational techniques [67]. In [42], Tagare *et al.* proposed a non-rigid registration algorithm based on L_2 norm and information-theory.

1.2 Summary by Chapters

Chapter 2 addresses the problem of discrete tomography. In this field, it is commonly assumed that the intensity values of the images (*i.e.*, the attenuation coefficients of the materials) are accurately known a priori. In practice, however, this information is usually not available and hard to determine. The author proposes novel reconstruction method for those cases, when the intensity values are unknown. Using higher order statistics based discretization term, the solution automatically converges to binary sets. Comparative tests showed, that the proposed method is more robust to the projection noise as a state-of-the-art algorithm. The method has also been successfully applied on real data.

In Chapter 3, a binary tomography based image deblurring method is introduced. Tikhonov regularization based 1-dimensional deconvolution is proposed to restore the horizontal and vertical projections. In this approach, the L-curve method is used to determine the regularization parameter which provides the best trade-off between the norm of the residual (data-fit term) and the norm of the solution (smoothness). To reconstruct the shapes from the restored projections, a maximum-flow based binary tomography method is applied. Comparative synthetic tests showed that the method can provide reliable result even in the case of low signal-to-noise ratio images, thanks to that the L-curve method can provide a good regularization parameter for any noise level. Results on real out-of-focus images are also presented.

In Chapter 4, the nonlinear registration of binary images is addressed. A general framework is proposed to estimate the parameters of a diffeomorphism that aligns two binary shapes. The framework is applied to different widely-used transformation classes. Comparisons to other state-of-the-art methods are conducted and its robustness to different type of segmentation errors is examined. The algorithm has also been successfully applied in different real-world applications.

Chapter 5 introduce a technique to estimate the parameters of a planar homog-

	[73]	[74]	[76]	[75]	[32]	[72]
I.	•					
II.		•				
III.			•	•	•	
IV.						•

Table 1.1: Correspondence between the thesis points and the publications.

raphy transformation between binary shapes. In this method the planar homography is decomposed into the perspective distortion and the affine transformation and the parameters of these two parts are estimated separately in two consecutive steps. Comparative synthetic tests and results on real images are also presented.

1.3 Summary of the Author's Contributions

In the followings the list of the key points of the dissertation is given. Table 1.3 shows the connection between the thesis points and the publications of the author. The planar homography related results of the thesis point III. was also previously presented in [77].

- I.) The author proposes a novel higher order statistics based binary tomography reconstruction method for those cases, when the intensities on the images are unknown. He proposes an objective functional in which a discretization term is applied to enforce binary solutions. He propose to minimize the objective functional by a graduated non-convexity optimization approach, in which the weight of the discretization term is increased during the optimization process to gradually enforce binary solutions. The mid-level of the intensities is estimated directly to the intermediate solutions in each iteration step. He proposes to estimate the mid-level as the minimum of the discretization term. The author examines the convergence properties of the method and shows that the behaviour of the method is independent of the value of the intensities. The robustness of the algorithm against different strengths of projection noise is also demonstrated. The author compares his algorithm to state-of-the-art methods and shows that his approach is a good alternative. He also successfully applies his algorithm to real projection data.
- II.) A binary tomography base method for the deconvolution of binary images is introduced. The author proposes to deconvolve the projections of the blurred

images using a Tikhonov regularization based algorithm. The optimal regularization parameter is found by the L-curve method. A maximum-flow based binary tomography method is proposed to reconstruct the shapes from their deblurred projections, . He shows in comparative tests that his method provides more reliable results than another widely-used method. He also presents results on real out-of-focus images.

- III.) The problem of nonlinear registration of binary images is addressed. A general registration framework is applied to trace back the registration problem to a system of nonlinear equations. The author applies the framework to different nonlinear transformation classes such as planar homography, polynomial transformations, and thin plate splines. He goes into the implementational details and shows that the equations can be written in three alternative forms which improves the registration results. In the case of planar homography, the author shows that the performance of the algorithm can be improved if the framework is applied to the Taylor series expansion of the transformation. The author proposes different ω function sets to construct the system of equations and compares them. He proposes to normalize the equations to guarantee equal contribution to the objective functional. The author compares the method to other state-of-the-art methods on synthetic datasets. He also examines the robustness of the algorithm against different types of segmentation errors. The author shows that the method can be applied in different real world applications.
- IV.) The author proposes an affine moment invariants based algorithm to estimate the parameters of a planar homography transformation between binary shapes. He shows how to decompose the transformation into a perspective and an affine component and estimate their parameters separately in two consecutive steps. He shows how to estimate the projective parameters of the transformation using affine moment invariants. In the second step, an affine registration method is used to determine the remaining parameters of the transformation. The author compares the method to the algorithm presented in the previous thesis point and shows results obtained by the algorithm on real images.

Chapter 2

Discrete Tomography with Unknown Intensity Levels

Batenburg *et al.* proposed a semi-automatic solution called Discrete Gray Level Selection (DGLS) method for the estimation of the gray-levels [10]. This algorithm can be used as a preprocessing step before the discrete tomography reconstruction. An initial continuous reconstruction is obtained, and an expert user is asked to select regions on the image that can be expected to contain constant gray-levels. This information is then used to estimate the gray-levels corresponding to the selected regions. In those cases, however, when the topology of the object is complex and does not contain sufficiently large homogeneous regions, this method can not be used.

An automatic method to estimate the segmentation parameters (gray-levels and threshold values) for the DART method has been proposed in [105]. This approach is based on the projection difference as a cost function. The main drawback of this approach is that it requires full DART reconstruction in each iteration step, therefore it is computationally inefficient. In [104], the DART method has been extended with a more advanced segmentation technique called Projection Distance Minimization (PDM) [11]. While in the original DART method the segmentation parameters are fixed, the PDM-DART method re-estimates these parameters in each iteration step by minimizing the distance between the projections of the intermediate segmented image and the projection data. The results are comparable to those, that can be obtained with the original DART, when the segmentation parameters are accurately known a priori.

In signal processing and image processing, entropy and high order statistics play important role to regularize the solutions [113, 22]. For binary deconvolution such measures are used to prescribe binary values for the solutions in [115] and [58]. In [51], Kim *et al.* proposed a blind deconvolution method for binary images with unknown

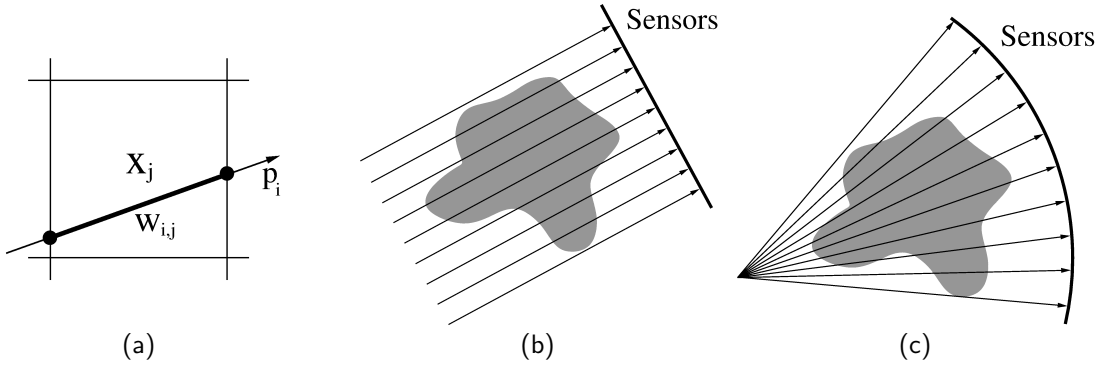


Figure 2.1: The discretization model (a) and the illustration of the parallel beam geometry (b) and the fan beam geometry (c).

intensity levels. The objective functional included a discretization term (a normalized kurtosis statistics) which enforces binary solutions. The only parameter of this term is the mid-level of the intensity levels which is estimated simultaneously with the solution using a gradient based trust-region-reflective method.

In this chapter, we introduce a binary tomography reconstruction method which applies the same higher order statistics based discretization term described in [51]. In our method, the weight of this term is increased during the optimization process to gradually enforce binary solutions which results in a *graduated non-convexity* optimization scheme [17]. The discretization parameter (mid-level) is estimated directly to each intermediate solution, as an alternating minimization step.

2.1 Basic Concepts and Notations

We introduce the basic concepts of the binary tomography reconstruction problem. Let the intensity values of the background pixels and the foreground (object) pixels denoted by c_1 and c_2 respectively. The digital image of size $h \times w$ is represented by the column vector $\mathbf{x} \in \{c_1, c_2\}^n$, where $n = hw$ is the total number of the pixels. Suppose that projections have been taken from k different angles and let l_i the number of measurements in the i th projection vector. All projection data are contained by the column vector $\mathbf{p} \in \mathbb{R}^m$, where $m = \sum_{i=1}^k l_i$ is the total number of measurements. The projection acquisition process, as usually is modeled by the system of linear equations

$$\mathbf{W}\mathbf{x} = \mathbf{p}, \quad (2.1)$$

where the matrix $\mathbf{W} \in \mathbb{R}^{m \times n}$ describes the projection geometry, i.e., w_{ij} represents the contribution of the image pixel x_j to the projection value p_i (see Fig. 2.1). The most

frequently used beam geometry types in tomography as well as in binary tomography include parallel beam, fan beam, and cone beam.

Basically a reconstruction of \mathbf{x} can be obtained by solving the system of linear equations Eq. (2.1). However, due to the low number of projections and the measurement errors (noisy projection data), the system in Eq. (2.1) is usually under-determined and inconsistent [103]. While the fact that the image that we are seeking is binary valued theoretically limits the number of possible solutions, in a combinatorial point of view, the accurate restoration of binary matrices is generally NP-hard, even when the intensity levels are accurately known. Moreover, herein we assume that the intensity levels c_1 and c_2 are unknown, which makes the reconstruction problem even more difficult.

2.2 The Reconstruction Method

The reconstruction is performed by the minimization of the following objective functional:

$$E(\mathbf{x}, \alpha, \mu) = F(\mathbf{x}) + \lambda S(\mathbf{x}) + \mu D(\mathbf{x}, \alpha), \quad (2.2)$$

in which we formulated three expected properties of the solution. First, the *data fidelity term* represents that \mathbf{x} should satisfy the projections:

$$F(\mathbf{x}) = \frac{1}{2} \|\mathbf{W}\mathbf{x} - \mathbf{p}\|_2^2. \quad (2.3)$$

It expresses that the solution should have projections close to the input projection data \mathbf{p} in the least squares sense and it is commonly used in the field of discrete tomography (e.g., in [92, 108]). As usually in the case of linear inverse problems, minimizing Eq. (2.3) alone would lead to non-feasible solutions, due to noisy projection data [103]. To regularize the solution, the objective functional contains a *smoothness prior term*

$$S(\mathbf{x}) = \frac{1}{2} \|L\mathbf{x}\|_2^2, \quad (2.4)$$

where L is the discrete Laplacian regularization matrix, i.e., $L\mathbf{x}$ gives the same result as the 2-dimensional convolution with the Laplacian filter. This term penalizes solutions with high norm of second derivatives but allows the formation of edges. Setting λ high enforces smooth regions even if the projections are noisy. Using only the data and the smoothness terms, the functional $F(\mathbf{x}) + \lambda S(\mathbf{x})$ is convex, and its minimum provides a continuous solution.

Binary reconstruction is imposed by the *discretization term*

$$D(\mathbf{x}, \alpha) = n \frac{\|\mathbf{x} - \alpha \mathbf{1}_n\|_4^4}{(\|\mathbf{x} - \alpha \mathbf{1}_n\|_2^2)^2} - 1, \quad (2.5)$$

where α is the *discretization parameter*, $\mathbf{1}_n$ denotes the length- n column vector of ones, and $\|\mathbf{x}\|_p = (\sum_{i=1}^n x_i^p)^{1/p}$ is the general vector norm. This functional is minimized by binary images, if α is the mid-level of the intensity values, *i.e.*, there exists d for which $|x_i - \alpha| = d$, for every $i = 1, \dots, n$ [22]. In such cases, it has a value of 0, while it reaches its maximum value $n - 1$ if all the pixel intensities are equal to α except one pixel. Thus it is bounded, and moreover, two times continuously differentiable except at $\mathbf{x} = \alpha \mathbf{1}_n$. Therefore standard gradient based optimization techniques can minimize it efficiently. The discretization term $D(\mathbf{x}, \alpha)$ is closely related to the fourth order standardized moment (or kurtosis), except that the fourth and second moments are normalized using the mid-level α instead of the mean of the values.

Setting the weight μ of the discretization term large allows only binary solutions as it suppresses the other components of the energy functional. The following theorem states that for any $\xi > 0$ a sufficiently high value of μ can be chosen such that the discretization level (*i.e.*, the value of the discretization term) will be at most ξ at the minimum point of the energy functional.

Theorem 1. *For any $\xi > 0$, there exists $\mu(\xi) \in \mathbb{R}$, such that if $\mu \geq \mu(\xi)$ and $(\hat{\mathbf{x}}, \hat{\alpha}) = \arg \min_{\mathbf{x}, \alpha} E(\mathbf{x}, \alpha, \mu)$, then $D(\hat{\mathbf{x}}, \hat{\alpha}) \leq \xi$.*

Proof. Let $\xi > 0$ be fixed and let $\mathbf{y} \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ be arbitrary such that $D(\mathbf{y}, \beta) = 0$. Choose $\mu(\xi) = (F(\mathbf{y}) + \lambda S(\mathbf{y}))/\xi$. If $\mu \geq \mu(\xi)$ and $(\hat{\mathbf{x}}, \hat{\alpha}) = \arg \min_{\mathbf{x}, \alpha} E(\mathbf{x}, \alpha, \mu)$, then

$$\mu D(\hat{\mathbf{x}}, \hat{\alpha}) \leq E(\hat{\mathbf{x}}, \hat{\alpha}, \mu) \leq E(\mathbf{y}, \beta, \mu) = F(\mathbf{y}) + \lambda S(\mathbf{y}) = \mu(\xi)\xi \leq \mu\xi \quad (2.6)$$

and the statement follows. \square

The reconstruction problem for given $\xi > 0$ and $\hat{\mu} \geq \mu(\xi)$ can be defined as the following minimization problem:

$$(\hat{\mathbf{x}}, \hat{\alpha}) = \arg \min_{\mathbf{x}, \alpha} E(\mathbf{x}, \alpha, \hat{\mu}). \quad (2.7)$$

In Section 2.2.1, we propose a graduated optimization scheme to minimize Eq. (2.7) by iteratively increasing the discretization weight μ and re-estimating $\hat{\mathbf{x}}$ and the mid-level $\hat{\alpha}$ in each iteration step. While convergence to the global minimum is not guaranteed, our experiments showed the efficiency of the method. Since Theorem 1 does not provide a sharp (minimum) threshold $\mu(\xi)$, thus instead of increasing μ until it reaches a specific $\hat{\mu} \geq \mu(\xi)$, it is more reasonable to terminate the minimization process once the value of the discretization term gets below ξ . In this case, the result will be a local minimum of the energy functional corresponding to the value of μ at the moment of the termination.

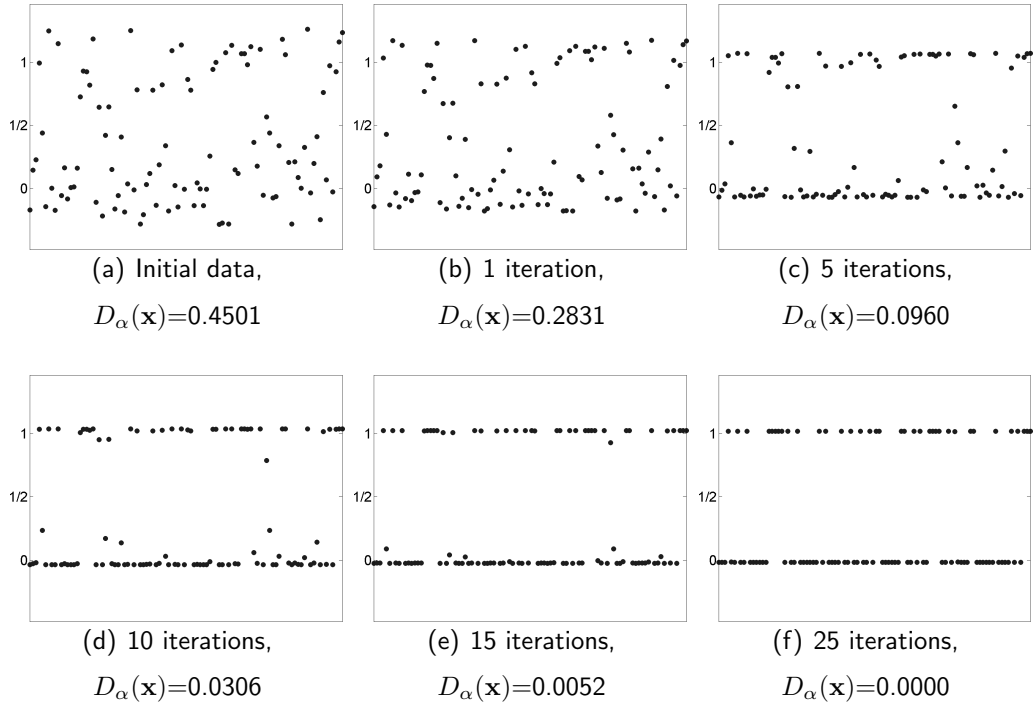


Figure 2.2: Demonstration of how the minimization of $D_\alpha(\mathbf{x})$ enforces binary sets. **(a)**: The initial data $\mathbf{x} \in \mathbb{R}^n, n = 100$ is created as a mixture of two normal distributions ($\mu_1 = 0, \sigma_1 = 0.2, \mu_2 = 1, \sigma_2 = 0.3$). **(b-f)**: The resulting vectors after different number of gradient descent iterations with $\alpha = 0.5$.

For the sake of simple notation we will refer to the discretization term and the objective functional by $D_\alpha(\mathbf{x})$ and $E_{\alpha,\mu}(\mathbf{x})$, in those cases, when the discretization weight μ and the discretization parameter α are fixed. Furthermore, the notation $E_{\alpha,\mu}^{\mathbf{p}}(\mathbf{x})$ will be used whenever we want to emphasize that the objective functional is constructed using the projection vector \mathbf{p} .

2.2.1 Optimization

The proposed optimization approach is based on the observation that while the discretization term is minimized by binary images, it does not have other local minimums (w.r.t. \mathbf{x}). Fig. 2.2 demonstrates how the minimization of the discretization term by simple gradient descent steps

$$\mathbf{x}' = \mathbf{x} - \nabla D_\alpha(\mathbf{x}), \quad (2.8)$$

initiating with a non-binary vector converges to a binary set. The original binary image, that we are seeking, minimizes the discretization term, if the discretization parameter α is equal to $(c_1 + c_2)/2$. Obviously, the value of this mid-level is initially unknown.

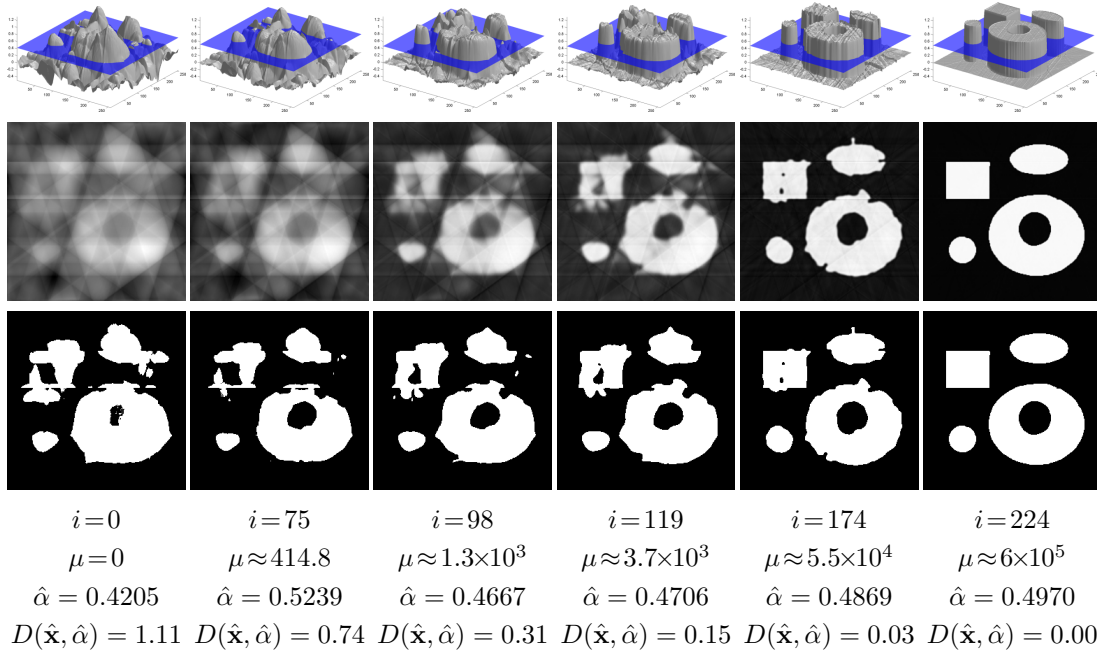


Figure 2.3: Demonstration of the convergence of the proposed method. The columns show intermediate reconstructions of Fig. 2.5(b) from 5 projections after different number of iterations. In the first row, 3-dimensional plots of the images are shown, in which the estimated mid-levels $\hat{\alpha}$ are indicated by horizontal planes. The second and third rows show the images and their $\hat{\alpha}$ -thresholded versions. Below the images the corresponding discretization weight μ , the estimated mid-level $\hat{\alpha}$ and the value of the discretization term $D(\hat{\mathbf{x}}, \hat{\alpha})$ can be found.

In the case of a binary image deconvolution method it was proposed [51], to optimize the value of the mid-level along with the solution and the point-spread-function in the same gradient descent based optimization process. However, we found that in the case of our method it is more efficient to estimate the mid-level directly to each intermediate solution $\hat{\mathbf{x}}$ based on the analysis of the discretization function. Therefore we propose the following graduated optimization approach for solving the optimization problem in Eq. (2.7). Starting from an initial reconstruction $\hat{\mathbf{x}}$ which is obtained by the minimization of the objective functional without the discretization term (*i.e.*, with $\mu = 0$), in each iteration step

1. Estimate the mid-level $\hat{\alpha}$ for the current solution $\hat{\mathbf{x}}$.
2. Increase the weight of the discretization term (μ).
3. Refine the reconstruction $\hat{\mathbf{x}}$ by locally minimizing the objective functional using the new parameters $\hat{\alpha}$ and μ .

Algorithm 1: The tomographic reconstruction method

Input : The projection data \mathbf{p} and the projection geometry matrix \mathbf{W} .**Output:** The reconstructed image $\hat{\mathbf{x}}$.Minimize $E_{0,0}(\mathbf{x})$ with the gradient descent method in Algorithm 2 to find the initial reconstruction $\hat{\mathbf{x}}$.Estimate $\hat{\alpha}$ (see Section 2.2.2) and choose a sufficiently small initial discretization weight μ .**while** $D_{\hat{\alpha}}(\hat{\mathbf{x}}) > \xi$ **do**Re-estimate $\hat{\mathbf{x}}$ by locally minimizing the objective functional $E_{\hat{\alpha},\mu}(\mathbf{x})$ using the gradient descent method in Algorithm 2 initiated with the previous minimum $\hat{\mathbf{x}}$.Estimate $\hat{\alpha}$ for the current solution $\hat{\mathbf{x}}$.Increase the discretization weight (we set $\mu \leftarrow 1.05\mu$).**end**Return with $\hat{\mathbf{x}}$.

This process increasingly enforces binary solutions while iteratively re-estimates the mid-level. Note that the method does not require the estimation of the intensity values during this process neither the thresholding of the intermediate solutions. The main steps of the method can be found in Algorithm 1, while the implementational details are discussed in Section 2.3.

The proposed minimization strategy follows a graduated optimization scheme. In the first step, without the discretization term (*i.e.*, with $\mu = 0$), the gradient based minimization can easily find a global minimum of the objective functional. Then as we slightly increase the weight of the discretization term μ and re-estimate the mid-level $\hat{\alpha}$, it is expected, that the local minimization finds the new global optimum, and as experimental results show, the proposed method converges to accurate solutions. Fig. 2.3 demonstrates the evolution of the intermediate solutions.

In each iteration step, the objective functional is minimized using the current discretization parameter $\hat{\alpha}$ and discretization weight μ . For minimization we used gradient descent method with backtracking line search[78], which guarantees convergence by finding a gradient descent step size t that not increases the objective functional. Since $E_{\hat{\alpha},\mu}$ is continuous (except at $\mathbf{x} = \alpha \mathbf{1}_n$), such a decreasing step size always exists in the negative gradient direction. The pseudo code of the minimization process can be

Algorithm 2: Minimization with backtracking line search

```

repeat
  Let the search direction be  $\mathbf{g} = -\nabla E_{\hat{\alpha},\mu}(\hat{\mathbf{x}})$ 
  Set  $t = h/K^*(\nabla C)$  and choose  $\beta, c \in (0, 1)$  (we set
   $h = 8, \beta = 0.5, c = 0.5$ ).
  while  $E_{\hat{\alpha},\mu}(\hat{\mathbf{x}} + t\mathbf{g}) > E_{\hat{\alpha},\mu}(\hat{\mathbf{x}}) - ct\|\mathbf{g}\|_2^2$  do
     $t \leftarrow \beta t$ 
  end
   $\hat{\mathbf{x}} \leftarrow \hat{\mathbf{x}} + t\mathbf{g}$ 
until  $\|t\mathbf{g}\|_2^2 \leq \delta\sigma^2(\hat{\mathbf{x}})$ 

```

found in Algorithm 2, in which the gradient of the objective functional is computed as

$$\begin{aligned}
\nabla E_{\alpha,\mu}(\mathbf{x}) &= \nabla F(\mathbf{x}) + \lambda \nabla S(\mathbf{x}) + \mu \nabla D_\alpha(\mathbf{x}), \\
\nabla F(\mathbf{x}) &= \mathbf{W}^T(\mathbf{W}\mathbf{x} - \mathbf{p}), \\
\nabla S(\mathbf{x}) &= \mathbf{L}^T \mathbf{L} \mathbf{x}, \\
\nabla_i D_\alpha(\mathbf{x}) &= 4n \left(\frac{(x_i - \alpha)^3}{(\|\mathbf{x} - \alpha \mathbf{1}_n\|_2^2)^2} - \frac{(x_i - \alpha)\|\mathbf{x} - \alpha \mathbf{1}_n\|_4^4}{(\|\mathbf{x} - \alpha \mathbf{1}_n\|_2^2)^3} \right),
\end{aligned} \tag{2.9}$$

where $\nabla_i D_\alpha(\mathbf{x})$ denotes the i th element of $\nabla D_\alpha(\mathbf{x})$.

An appropriate initial step-length for the line search process can be found by considering the Lipschitz-continuity [110] of $\nabla E_{\hat{\alpha},\mu}$. Let the Lipschitz-constant of an arbitrary Lipschitz-continuous function $f : \Omega \rightarrow \mathbb{R}^n$, $\Omega \subseteq \mathbb{R}^n$ denoted by $K(f)$. If f is continuous on Ω , then

$$K(f) = \sup_{\mathbf{x} \in \Omega} \|J(f)(\mathbf{x})\|, \tag{2.10}$$

where $J(f)$ denotes the Jacobian matrix of f .

Let $C(\mathbf{x}) = F(\mathbf{x}) + \lambda S(\mathbf{x})$ the convex component of $E_{\hat{\alpha},\mu}(\mathbf{x})$. Since

$$J(\nabla C) = \mathbf{W}^T \mathbf{W} + \lambda \mathbf{L}^T \mathbf{L} \tag{2.11}$$

is a constant matrix, thus an upper bound estimation for $K(\nabla C) = \|J(\nabla C)\|$ can be given by

$$K^*(\nabla C) = \|J(\nabla C)\|_1, \tag{2.12}$$

where $\|J(\nabla C)\|_1$ is equal to the maximum column sum of the positive semidefinite $J(\nabla C)$.

It can be shown [110], that choosing step-length $t \leq 1/K(\nabla E_{\hat{\alpha},\mu})$, the gradient descent step will not increase the function value. Since $K(\nabla E_{\hat{\alpha},\mu}) \leq K(\nabla C) + \mu K(\nabla D_{\hat{\alpha}})$, thus $K(\nabla C) \leq K(\nabla E_{\hat{\alpha},\mu})$. Therefore we propose to choose

$t = h/K^*(\nabla C)$ as an initial step-length guess in Algorithm 2, with $h \geq 1$ for the sake of a faster convergence (we set $h = 8$).

The discretization term is also Lipschitz-continuous on an appropriately defined subset $\Lambda \subseteq R^n$, and in Section 2.7 we estimate an upper bound $K^*(\nabla D_{\hat{\alpha}})$ for its Lipschitz-constant. Therefore, $K^*(\nabla E_{\hat{\alpha},\mu}) = K^*(\nabla C) + \mu K^*(\nabla D_{\hat{\alpha}})$ is an upper bound for $K(\nabla E_{\hat{\alpha},\mu})$ and a step size $t^* = 1/K^*(\nabla E_{\hat{\alpha},\mu})$ can be chosen directly (*i.e.*, without the line search process) with which the gradient descent step will not increase the objective functional. However, for fast convergence Λ should be restricted to the set of the next possible solutions in the direction of $-\nabla E_{\hat{\alpha},\mu}$, and $K^*(\nabla D_{\hat{\alpha}})$ should be estimated locally on that Λ before each descent step. In our experiments, this estimation was more time consuming, thus we present our numerical results using the simple line search approach. However in a parallel implementation this direct step-length estimation could be a good alternative. Moreover the Lipschitz-continuity of the gradient of the whole energy functional is an important property which indicates that it can be efficiently minimized by standard optimization techniques.

2.2.2 Estimation of the Discretization Parameter

In this section we show how to assign a suitable value of $\hat{\alpha}$ to the given intermediate reconstruction. It should be emphasized that it is not an attempt to estimate the mid-level of the unknown original image, we only propose a heuristic way to assign a value to a given intermediate solution in each iteration step. This $\hat{\alpha}$ then can be used in the discretization term in the next iteration step, and thus binary solutions can be enforced relatively to this value. For that purpose, for a given intermediate solution $\hat{\mathbf{x}}$ we consider the discretization term $D_{\hat{\mathbf{x}}}(\alpha) = D(\hat{\mathbf{x}}, \alpha)$, as a function of α . For the case of a binary vector $\hat{\mathbf{x}} \in \{c_1, c_2\}^n$, this function is plotted on Fig. 2.4(a). It is minimized by $\alpha = (c_1 + c_2)/2$, while it converges to (but never reaches) 0, if $\alpha \rightarrow \pm\infty$. Thus in the case of an intermediate solution $\hat{\mathbf{x}} \in \mathbb{R}^n$, it is reasonable to select $\hat{\alpha}$ corresponding to the local minimum of $D_{\hat{\mathbf{x}}}(\alpha)$ (see Fig. 2.4(b)). Another reason why this approach is efficient is that determining the three extremas of $D_{\hat{\mathbf{x}}}(\alpha)$ can be done in a straightforward way by simply solving the 3th order equation

$$D'_{\hat{\mathbf{x}}}(\alpha) = 0, \quad (2.13)$$

for α in algebraic way. If it has three real roots, then the second one in order gives $\hat{\alpha}$.

In some cases, typically when there is no suitable amount of projection data available, the initial continuous solution and some further intermediate solutions can be inaccurate, and thus object pixels and the background pixels are not separated sufficiently. Therefore $D_{\hat{\mathbf{x}}}(\alpha)$ has no local minimum indicating the mid-level of the intensities, *i.e.*, Eq. (2.13) has only one real root (see Fig. 2.4(c)). In such cases, however,

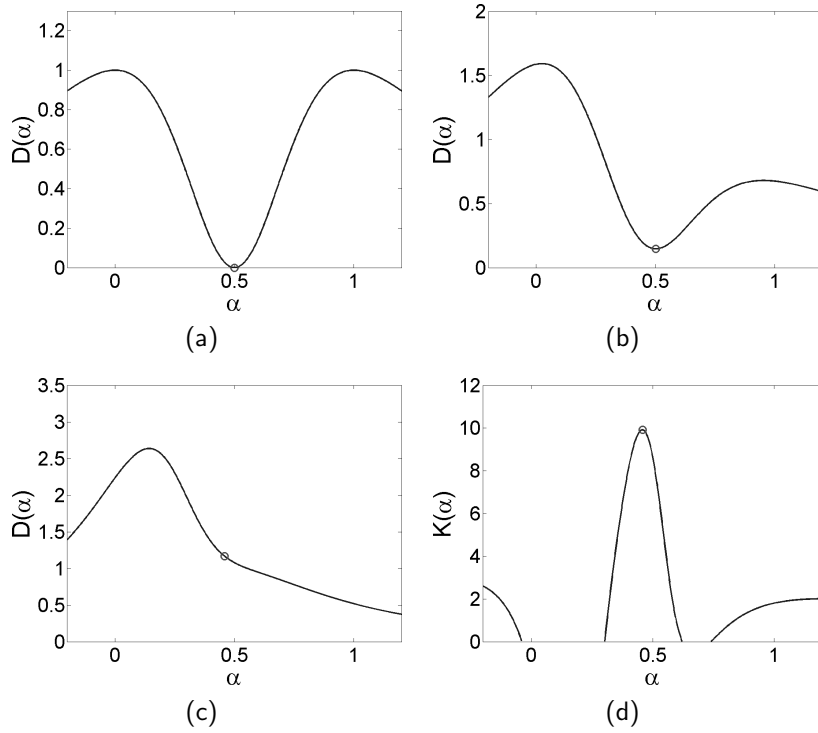


Figure 2.4: Estimation of discretization parameter $\hat{\alpha}$ by analyzing the discretization term $D_{\hat{x}}(\alpha)$ as a function of α . **(a)-(b)**: For a binary set and for a non-binary intermediate reconstruction, its minimum indicates $\hat{\alpha}$. **(c)-(d)**: If the intermediate reconstruction is inaccurate, *i.e.*, $D_{\hat{x}}(\alpha)$ has no minimum, we propose to choose $\hat{\alpha}$ that maximizes $K_{\hat{x}}(\alpha)$ (here we plotted only its positive values).

there is still a point with large curvature, which indicates an appropriate discretization parameter, *i.e.*, we choose $\hat{\alpha}$ that maximizes the following curvature-like function:

$$K_{\hat{x}}(\alpha) = \frac{D_{\hat{x}}''(\alpha)}{(\varepsilon + D_{\hat{x}}'(\alpha))}, \quad \varepsilon \ll 1. \quad (2.14)$$

Fig. 2.4(d) shows the plot of $K_{\hat{x}}(\alpha)$ corresponding to $D_{\hat{x}}(\alpha)$ in Fig. 2.4(c). While the maximum of $K_{\hat{x}}(\alpha)$ can not be determined in an algebraic way, the simple numerical method that evaluates Eq. (2.14) for α values sampled uniformly in the domain $[\min\{\hat{x}_i\}, \max\{\hat{x}_i\}]$ can find a good approximation.

Other approaches to provide the discretization parameter $\hat{\alpha}$ are also feasible. In our experiments, when sufficient amount of projection data was available, and thus the initial continuous solution was accurate, the proposed method was able provide accurate reconstructions by simply setting $\hat{\alpha} = (\min\{\hat{x}_i\} + \max\{\hat{x}_i\})/2$ in every iteration step, while other pixel classification and threshold selection approaches [85] could also be applied. On the other hand, as experimental results show, our method provided

accurate reconstructions even from low number of projections using the discretization parameter estimation approach described in this section.

2.3 Gray-Level Independence and Implementational Details

In this section, we discuss the implementational details and a practically important property of the proposed method. In practice, the gray-levels can take values of different orders of magnitude. It was a main purpose to develop our algorithm to provide the same results, when it is applied to reconstruct images with different gray-levels containing the same object. Without this property the algorithm would require to set its certain parameters (e.g., the termination criterion thresholds and the initial discretization weight), to the specific application.

Definition 1. *The images $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are equivalent, if there exist $u \in \mathbb{R}$ and $v \in \mathbb{R}$, such that $\mathbf{y} = u\mathbf{x} + \mathbf{v}$, where $\mathbf{v} = v\mathbf{1}_n$.*

For better understanding, it is worth pointing out that if $\mathbf{x} \in \{c_1, c_2\}^n$ and $\mathbf{y} = u\mathbf{x} + \mathbf{v}$ for some $u, v \in \mathbb{R}$ then $\mathbf{y} \in \{uc_1 + v, uc_2 + v\}^n$.

Definition 2. *The projection data $\mathbf{p} \in \mathbb{R}^m$ and $\mathbf{q} \in \mathbb{R}^m$ are equivalent, if there exist $u \in \mathbb{R}$ and $v \in \mathbb{R}$, such that $\mathbf{q} = u\mathbf{p} + \mathbf{W}\mathbf{v}$, where $\mathbf{v} = v\mathbf{1}_n$.*

It can be easily seen that for a given projection matrix W , the projections of equivalent images are also equivalent with the same u and v values. Moreover, the binary images of the same object are equivalent.

By the *gray-level independence* of the proposed method we mean, that if it is applied on two equivalent projection vectors, then the resulting reconstructed images will also be equivalent. This property can be very useful in practice, because the attenuation coefficients and thus the gray-levels can be highly variable. For example, thanks to this independence, in our experiments on synthetic data and on real data, the proposed method could be applied without any modification or fine-tuning the termination criteria. Nevertheless, providing this property of the method is far from trivial. The following statement ensures that applying gradient descent process on equivalent intermediate solutions keeps their equivalence.

Theorem 2. *Let the projection matrix \mathbf{W} be fixed and let \mathbf{p} and \mathbf{q} equivalent projection vectors (in Def. 2.) for some u and v . Suppose that we apply the proposed method on \mathbf{p} and \mathbf{q} simultaneously using the same regularization parameter λ , and at certain points of the iteration processes*

1. the intermediate solutions $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are equivalent (in Def. 1.), and $\hat{\mathbf{y}} = u\hat{\mathbf{x}} + \mathbf{v}$, where $\mathbf{v} = v\mathbf{1}_n$.
2. μ and τ are the discretization weights, and $\tau = u^2\mu$.
3. $\hat{\alpha}$ and $\hat{\beta}$ are the mid-level estimations, and $\hat{\beta} = u\hat{\alpha} + v$.

After one gradient descent step in Algorithm 2, the resulting solutions $\hat{\mathbf{x}}'$ and $\hat{\mathbf{y}}'$ will also be equivalent, and $\hat{\mathbf{y}}' = u\hat{\mathbf{x}}' + \mathbf{v}$.

Proof. First consider the correlation of the two objective functionals:

$$\begin{aligned}
 E_{\hat{\beta},\tau}^{\mathbf{q}}(\hat{\mathbf{y}}) &= \frac{1}{2}\|W\hat{\mathbf{y}} - \mathbf{q}\|_2^2 + \frac{\lambda}{2}\|\mathbf{L}\hat{\mathbf{y}}\|_2^2 + \tau D_{\hat{\beta}}(\hat{\mathbf{y}}) \\
 &= \frac{1}{2}\|W(u\hat{\mathbf{x}} + \mathbf{v}) - (u\mathbf{p} + \mathbf{W}\mathbf{v})\|_2^2 + \frac{\lambda}{2}\|\mathbf{L}(u\hat{\mathbf{x}} + \mathbf{v})\|_2^2 + u^2\mu D_{\hat{\alpha}}(\hat{\mathbf{x}}) \\
 &= \frac{1}{2}\|u(W\hat{\mathbf{x}} - \mathbf{p})\|_2^2 + \frac{\lambda}{2}\|u(\mathbf{L}\hat{\mathbf{x}})\|_2^2 + u^2\mu D_{\hat{\alpha}}(\hat{\mathbf{x}}) = u^2 E_{\hat{\alpha},\mu}^{\mathbf{p}}(\hat{\mathbf{x}}), \quad (2.15)
 \end{aligned}$$

since $D_{\hat{\beta}}(\hat{\mathbf{y}}) = D_{\hat{\alpha}}(\hat{\mathbf{x}})$ and $\mathbf{L}\mathbf{v} = \mathbf{0}_n$. In a similar way, it can be shown that $\nabla E_{\hat{\beta},\tau}^{\mathbf{q}}(\hat{\mathbf{y}}) = u\nabla E_{\hat{\alpha},\mu}^{\mathbf{p}}(\hat{\mathbf{x}})$.

Let $\mathbf{g}_1 = -\nabla E_{\hat{\alpha},\mu}^{\mathbf{p}}(\hat{\mathbf{x}})$ and $\mathbf{g}_2 = -\nabla E_{\hat{\beta},\tau}^{\mathbf{q}}(\hat{\mathbf{y}})$. Considering the termination criterion of the line search method in Algorithm 2, it can be easily checked that:

$$\begin{aligned}
 E_{\hat{\beta},\tau}^{\mathbf{q}}(\hat{\mathbf{y}} + t\mathbf{g}_2) &> E_{\hat{\beta},\tau}^{\mathbf{q}}(\hat{\mathbf{y}}) - ct\|\mathbf{g}_2\|_2^2 \iff \\
 &\iff E_{\hat{\alpha},\mu}^{\mathbf{p}}(\hat{\mathbf{x}} + t\mathbf{g}_1) > E_{\hat{\alpha},\mu}^{\mathbf{p}}(\hat{\mathbf{x}}) - ct\|\mathbf{g}_1\|_2^2, \quad (2.16)
 \end{aligned}$$

thus the line search method will terminate with the same step-length t , for both cases. Performing one gradient descent step:

$$\hat{\mathbf{y}}' = \hat{\mathbf{y}} + t\mathbf{g}_2 = u\hat{\mathbf{x}} + \mathbf{v} + t u\mathbf{g}_1 = u(\hat{\mathbf{x}} + t\mathbf{g}_1) + \mathbf{v} = u\hat{\mathbf{x}}' + \mathbf{v}. \quad (2.17)$$

□

In the followings, we discuss the implementation details of the algorithm. Using these settings, one can easily check, that performing the algorithm on equivalent projection vectors, the conditions of Lemma 2 will always be fulfilled, independently of what are the actual values of u and v . Therefore the method will perform gray-level independently, *i.e.*, applying it on two equivalent projection vectors, the resulting images will also be equivalent.

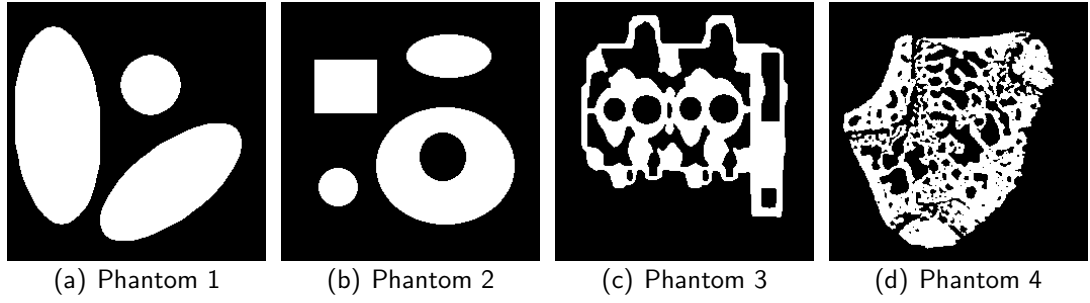


Figure 2.5: Four of the phantom images used to evaluate the performance of the proposed method.

Initialization: The first gradient descent minimization should be started from some initial $\hat{\mathbf{x}}$. Since the energy functional without the discretization term is convex, the initial point can be arbitrary. However, to provide gray-intensity independence we choose the constant vector containing the average intensity value in each pixel:

$$\hat{x}_i = \frac{\sum_{j=1}^m p_j}{\sum_{r=1}^m \sum_{i=1}^n w_{ri}}. \quad (2.18)$$

It can be easily checked, that the initial vectors calculated from equivalent projections will be also equivalent.

Discretization weight: The graduated optimization scheme requires a small initial discretization weight μ with which the discretization term does not suppress the other components of the energy functional in the early iteration steps. Since the discretization term is bounded, such μ can be selected based on the value of the energy functional after the first minimization process. In our experiments $\mu = 10^{-2} E_{0,0}(\hat{\mathbf{x}})$, where $\hat{\mathbf{x}}$ is the initial continuous reconstruction, provided a sufficiently small initial weight. From Eq. (2.15) it follows that this choice fulfills the 2nd condition of Theorem 2.

Discretization parameter: If $\hat{\mathbf{y}} = u\hat{\mathbf{x}} + \mathbf{v}$, then the method described in Section 2.2.2 will provide $\hat{\alpha}$ and $\hat{\beta}$ as the discretization parameters for $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, such that $\hat{\beta} = u\hat{\alpha} + v$. This satisfies the 3rd condition of Lemma 2.

Gradient descent processes: To detect the convergence of the gradient descent process in Algorithm 2, the step-length is compared to the normalized threshold $\delta\sigma^2(\hat{\mathbf{x}})$. We used $\delta = 2^{-25}$ in all of our experiments.

Termination criterion: As a termination criterion for the main loop of the algorithm, the value of the discretization term is compared to $\xi = 10^{-4}$, to detect if the

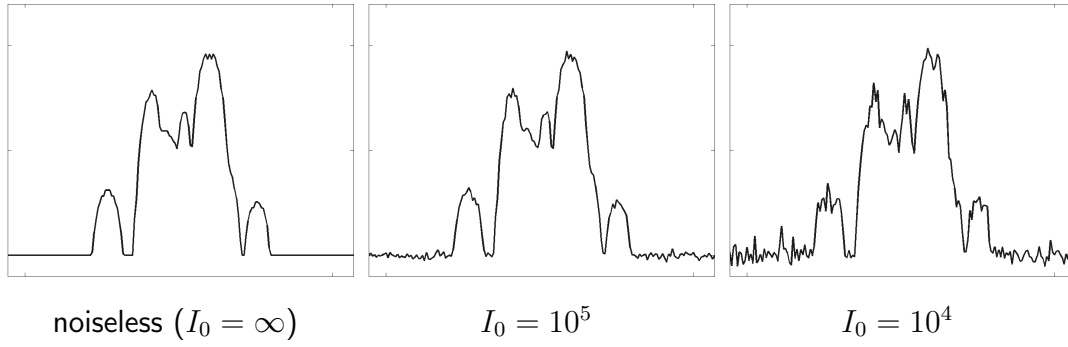


Figure 2.6: Example noiseless and noisy projections for different incident beam intensities.

solution reached a sufficient discretization level. The value of the discretization term is gray-level independent, *i.e.*, if $\hat{\mathbf{y}} = u\hat{\mathbf{x}} + \mathbf{v}$ and $\hat{\beta} = u\hat{\alpha} + v$, then $D_{\hat{\alpha}}(\hat{\mathbf{x}}) = D_{\hat{\beta}}(\hat{\mathbf{y}})$.

With these settings, performing the reconstruction algorithm on two equivalent projections, the corresponding intermediate solutions and the final reconstructions will also be equivalent. Implementing the algorithm this way helps to apply it on images containing even highly different gray-levels, without any modification or parameter refinement.

2.4 Numerical Experiments

To evaluate the performance of the proposed method, we used four phantom images of size 256×256 (see Fig. 2.5). For these experiments we considered the parallel beam projection geometry with equidistant line spacing. Different number of projections distributed equiangularly in the full angular range have been used. The distance between the projection beams was 1 pixels in every direction. We performed our experiments for both noiseless and noisy cases, where projection noise of two different levels has been applied. In tomography, the projection data measured as the amount of particles reaching the detectors, and these particle counts are typically assumed to follow Poisson distribution. The strength of the noise can be characterized by the *incident beam intensity* (I_0), which is the maximal number of particle count in the beams. This noise model was recently used in [104]. Here, we tested the methods for $I_0 = 10^5$ and $I_0 = 10^4$, and the noiseless cases are referred by $I_0 = \infty$. As Fig. 2.6 shows, the first amount provides somewhat weaker noise, while in the second case we get stronger distortions. The weight of the smoothness prior was set empirically to $\lambda = 4.0$, and this

Phantom 1, $I_0 = \infty$					Phantom 2, $I_0 = \infty$				
#p	PDM-DART		Proposed		#p	PDM-DART		Proposed	
	rNMP	time	rNMP	time		rNMP	time	rNMP	time
4	0.79	9	1.07	48	5	0.69	16	0.68	91
5	0.43	13	0.66	51	6	0.32	11	0.25	50
8	0.09	9	0.07	37	8	0.14	11	0.10	43
12	0.01	20	0.02	33	16	0.01	10	0.02	39

Phantom 1, $I_0 = 10^5$					Phantom 2, $I_0 = 10^5$				
#p	PDM-DART		Proposed		#p	PDM-DART		Proposed	
	rNMP	time	rNMP	time		rNMP	time	rNMP	time
4	1.25	11	1.15	48	5	1.43	22	1.04	98
5	0.81	15	1.02	51	6	0.93	13	0.69	56
8	0.50	11	0.37	39	8	0.74	10	0.38	46
12	0.40	9	0.33	34	16	0.33	11	0.17	41

Phantom 1, $I_0 = 10^4$					Phantom 2, $I_0 = 10^4$				
#p	PDM-DART		Proposed		#p	PDM-DART		Proposed	
	rNMP	time	rNMP	time		rNMP	time	rNMP	time
4	2.87	11	3.34	46	5	3.62	22	5.75	85
5	2.35	13	2.87	54	6	3.27	15	2.53	54
8	1.71	8	1.73	40	8	2.57	20	2.34	46
12	1.21	20	1.45	36	16	1.45	8	1.46	45

Table 2.1: The rNMP measures (%) and runtime performances (sec.) provided by the PDM-DART method and the proposed method on Phantom 1 and Phantom 2 (Fig. 2.5) using different number of projections. The three rows show the results for the three different strengths of noise applied on the projections. Bold values indicate the better rNMP results for each test case.

value was used in all of these experiments. To evaluate the reconstruction results we used the error measure called *relative Number of Misclassified Pixels* (rNMP), which gives the ratio of the number of erroneous pixels and the number of object pixels in the original phantom image. This error measure was recently used in [104] to evaluate the results of the PDM-DART method. The rNMP measures and runtime performances provided by the proposed method can be found in Table 2.1 and Table 2.2, while Fig. 2.7 shows some reconstruction results.

Phantom 3, $I_0 = \infty$					Phantom 4, $I_0 = \infty$				
#p	PDM-DART		Proposed		#p	PDM-DART		Proposed	
	rNMP	time	rNMP	time		rNMP	time	rNMP	time
20	1.41	32	0.72	69	28	10.46	41	9.68	88
28	0.74	23	0.19	66	36	7.03	66	6.23	96
40	0.22	47	0.08	73	40	5.89	86	4.99	98
80	0.00	58	0.04	97	80	2.16	64	1.78	127

Phantom 3, $I_0 = 10^5$					Phantom 4, $I_0 = 10^5$				
#p	PDM-DART		Proposed		#p	PDM-DART		Proposed	
	rNMP	time	rNMP	time		rNMP	time	rNMP	time
20	2.09	25	1.39	71	28	11.82	63	10.10	86
28	1.21	23	0.57	65	36	8.51	55	7.05	102
40	0.90	48	0.25	73	40	7.16	50	5.60	98
80	0.08	53	0.05	99	80	2.74	113	1.98	126

Phantom 3, $I_0 = 10^4$					Phantom 4, $I_0 = 10^4$				
#p	PDM-DART		Proposed		#p	PDM-DART		Proposed	
	rNMP	time	rNMP	time		rNMP	time	rNMP	time
20	5.61	21	4.67	75	28	16.30	60	13.53	86
28	4.41	30	3.83	71	36	12.70	91	11.00	97
40	3.52	24	2.94	82	40	12.48	103	10.00	101
80	1.91	76	2.01	123	80	8.58	60	5.37	146

Table 2.2: The rNMP measures (%) and runtime performances (sec.) provided by the PDM-DART method and the proposed method on Phantom 3 and Phantom 4 (Fig. 2.5) using different number of projections. The three rows show the results for the three different strengths of noise applied on the projections. Bold values indicate the better rNMP results for each test case.

2.4.1 Comparison to the PDM-DART Method

Van Aarle *et al.* introduced an elegant solution for the automatic estimation of the intensity levels [104]. The main advantage of their method is that it is not restricted to binary images, and its efficiency for reconstructing images containing three unknown gray-levels was also demonstrated. Their method is based on the DART algorithm [12]. In each iteration step, the DART method segments the actual intermediate solution, and fixes the non-boundary pixels. In the next reconstruction step, it updates only the non-fixed pixels. For the reconstruction step the SIRT algorithm [39] is proposed, but theoretically other methods could also be applied. To obtain more accurate results, certain amount of randomly chosen non-boundary pixels are also selected to be

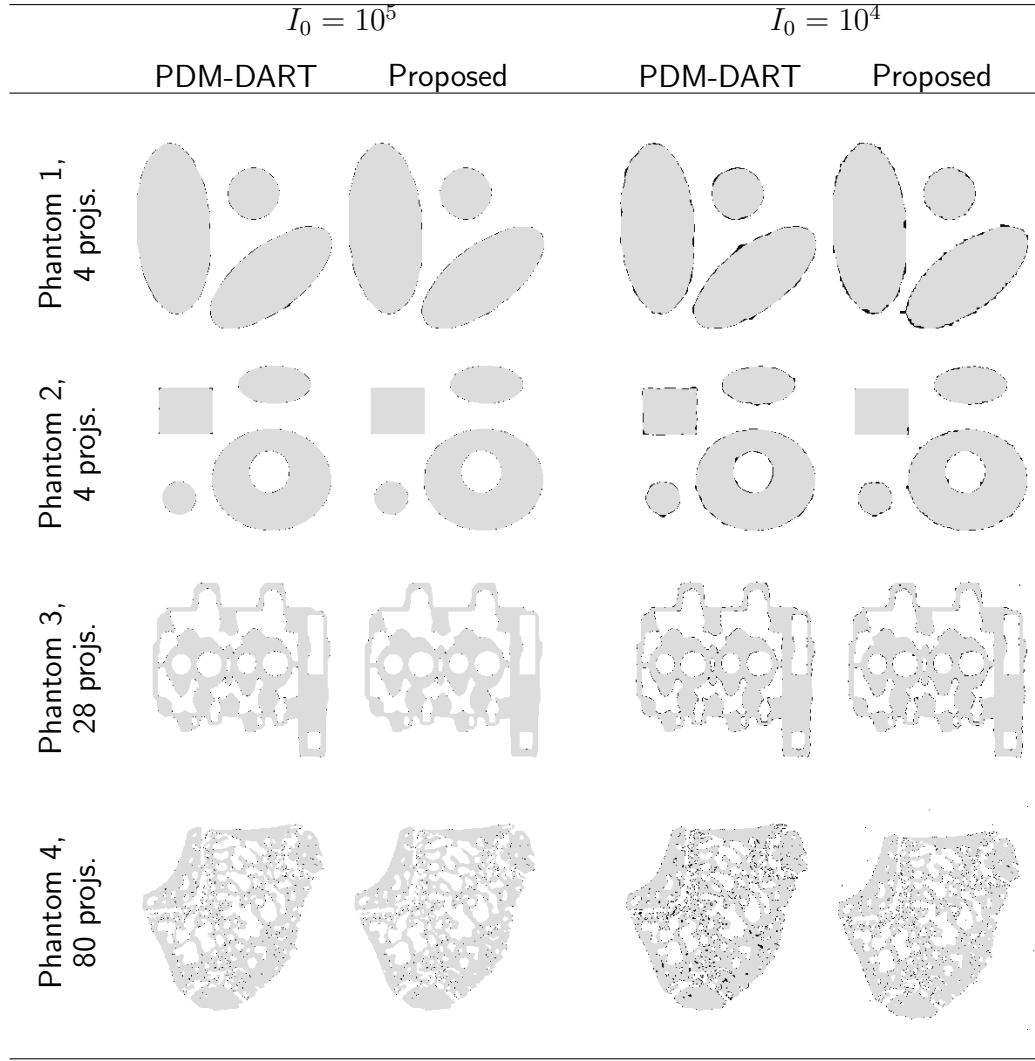


Figure 2.7: Some reconstruction results of the PDM-DART method and the proposed method. The misclassified pixels are depicted with dark.

non-fixed. This allows the formation of new boundaries that are not connected to the boundary of the current solution. The noise of the projections and the errors caused by the incorrectly fixed pixels can result in fluctuations in the updated values of the non-fixed pixels. To overcome this issue, a Gaussian filter is applied to the intermediate solutions. While in the original description of the DART method this smoothing operation was performed only on the non-fixed pixels, in the more recent work of the authors [104] it was proposed to apply it on the whole image. We also found in our experiments that this latter approach provides more accurate results. For the estimation of the segmentation parameters, the *projection distance minimization* (PDM) method has been proposed. The gray-levels and the threshold values are selected to obtain a segmented image that satisfies the projection data the most.

For our tests, we basically set the parameters of the PDM-DART method according to the author's recommendations. As it was proposed, for termination criterion the *total projection error* (TPE) was considered and the DART iterations was stopped if TPE had not been decreased in the last 30 iterations. We selected randomly 1% of the non-boundary pixels to remain non-fixed in each step. As it was recommended, to reduce the additional computational cost introduced by the PDM method, the segmentation parameter estimation was performed only once in every 10 iterations. Using these settings in our experiments, the PDM-DART method was able to find the correct segmentation parameters and to converge to the right boundaries of the objects. Nevertheless, we found that the choice of the radius of the Gaussian smoothing filter strongly affects the results as follows. Applying weaker smoothing allows the PDM-DART to converge to the right boundaries of the objects, and thus it provides accurate reconstructions. At the same time, in the presence of stronger projection noise, the large pixel-value fluctuations can only be handled by applying stronger smoothing, which however blurs the important details in the case of complex shapes. The influence of the choice of the smoothing kernel on the reconstruction quality has been recently investigated in [70]. In our experiments, the PDM-DART method was not able to provide reliable results with the same smoothing kernel in each test case. Since the adaptive selection of the kernel might be feasible, thus in order to make a fair comparison, we performed the PDM-DART method using smoothing kernels with different deviations ($\sigma = 0.5, 1.0, \dots, 3.0$), and in each test case we chose the best result. As it can be seen in Table 2.1 and Table 2.2, the proposed method is still more robust to the projection noise, as it has been performed with the same weight of the regularization term.

Both algorithms were implemented in Matlab R2008b and the experiments were performed on an Intel Core2 Quad 2.66GHz CPU. As the results show in Table 2.1 and Table 2.2, PDM-DART required less computational time especially on the less complex Phantom 1 and Phantom 2. This is due to that it updates only a relatively small subset of the pixels in each step, while the proposed method updates all pixels of the image. However, an efficient GPU based implementation, in which the update steps could be performed on every pixel in a parallel way, could equalize the runtime differences.

2.4.2 Limited Angular Range Experiments

In many applications, it is not possible to collect projection data from the full angular range due to the size or the shape of the investigated object or the design of the imaging device. To examine the performance of the methods in such scenarios,

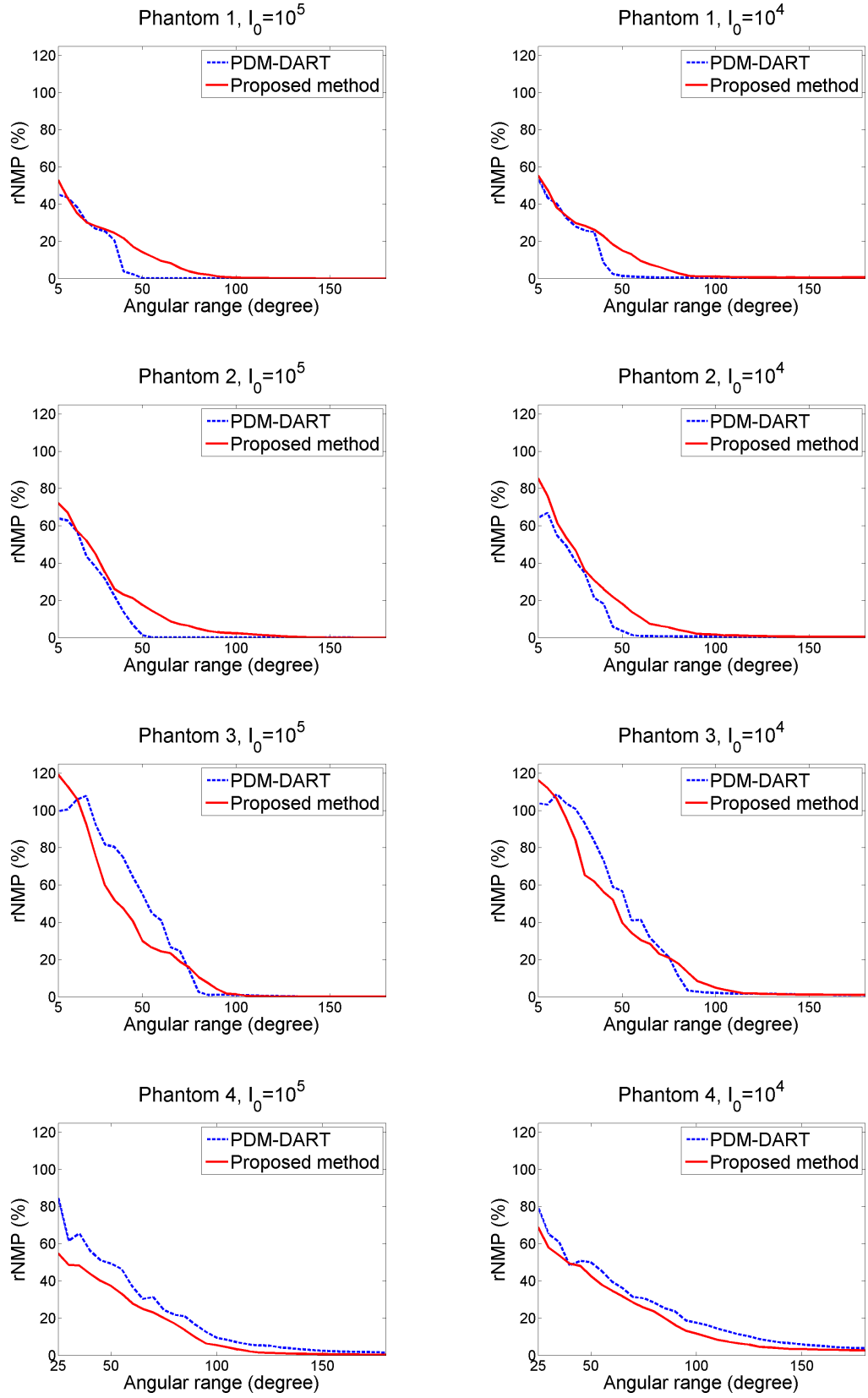


Figure 2.8: Limited angular range experiments. The rNMP measures as functions of the angular ranges are plotted.

we generated 180 projections in the full angular range, corresponding to the angles $0^\circ, 1^\circ, 2^\circ, \dots, 179^\circ$. We performed reconstructions using the projections in the intervals of $[0^\circ, \phi]$ for $\phi = 4^\circ, 9^\circ, 14^\circ, \dots, 179^\circ$, *i.e.*, the number of projections increased with the angular range. In Fig. 2.8, the rNMP errors are plotted against the angular ranges. Since we obtained similar results for $I_0 = \infty$ and $I_0 = 10^5$, thus only the results for $I_0 = 10^5$ and $I_0 = 10^4$ are presented. The PDM-DART method was able to provide almost perfect reconstructions of Phantom 1 and Phantom 2, even from as small range as $[0^\circ, 49^\circ]$. This is due to the efficiency of the DART method to converge to the right boundaries of the objects, even in those cases, where the methods based on the minimization of objective functionals (like the proposed algorithm) stuck in local minima. However, on the more complex Phantom 3 and Phantom 4, the proposed method provided more accurate results.

2.4.3 Comparison to a Convex Programming Approach

In [23], Capricelli and Combettes introduced a convex programming based approach for noisy discrete tomography. The method traces back the problem to the minimization of a convex function on the intersection of some appropriately constructed convex sets. The advantage of this approach is that these sets can represent a wide range of prior information (projection data, noise level estimation, binarity constraints, support image, etc.). The minimized objective function was the square of the ℓ^2 -norm of the image in their experiments. The binarity promoting constraint prescribed upper bound for the ℓ_1 -norm (total variation) of the image. In their experiments the projections was multiplied by 256 and Poisson noise was applied directly on the projection values. The implementation of the method is available online [24] as a part of a convex programming toolkit with an easy to use GUI. For discrete tomography the program allows only 2, 3 or 4 equiangular projections, the measurements are given as the discrete pixels sums along the four directions (horizontal, vertical and two diagonal pixel sums, that consist of $256 - 256$ and $511 - 511$ measurements in the case of a 256×256 image). Therefore, we tested their algorithm on Fig. 2.5(a) using 4 projections. The application provides an option to set the a priori values according to [23]. We tested different combinations of a priori informations and settings, and we obtained the most accurate result by using the a priori informations ℓ_1 -norm, ℓ_2 -norm, and the *pixel range* together and the default algorithm settings. The method basically provides a "close to binary" solution but the segmentation of the results is also implemented. Fig. 2.9(a) and Fig. 2.9(b) shows the output of the application and its binarization, the rNMP error was 3.94%. We applied the proposed method on the same projection data using the same projection geometry, and we obtained a 1.10% rNMP error (see Fig. 2.9(c)).

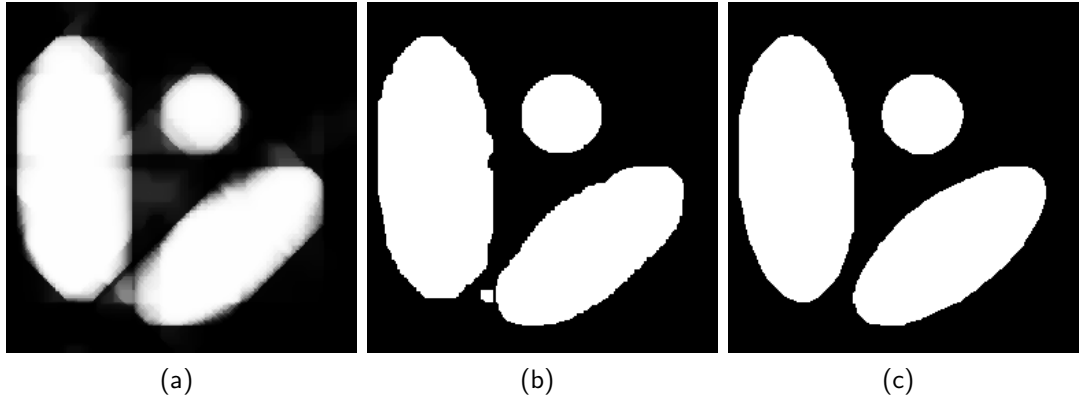


Figure 2.9: **(a)**: The reconstruction result obtained by the convex programming method in [23] using four projections and **(b)**: its discretization (rNMP=3.94%). **(c)**: The result obtained by the proposed method (rNMP=1.10%) using the same projection data and projection geometry.

For a further comparison, we considered one of the numerical simulations described in [23]. In this experiment, 4 projections were used to reconstruct the 32×32 image in Fig. 2.10(a). The histogram of the number of erroneous pixels on 2000 different realization of the projection noise showed that their method misclassified about an average of 8 pixels. Unfortunately the applied projection geometry is not clarified in [23] for this experiment. Thus we conducted two tests using our method assuming two different projection geometries:

1. In the first case, we considered parallel beam geometry with equidistant beam spacing. The horizontal and vertical projections consisted of $32 - 32$ measurements, while the diagonal projections consisted of $45 - 45$ values.
2. In the second case, we applied the discrete pixel sums that is used by their online available implementation. Here the horizontal and vertical projections consisted of $32 - 32$, while the diagonal projections consisted of $63 - 63$ measurements.

For both cases we conducted a similar test on Fig. 2.10(a) using the proposed method with $\lambda = 0.1$. 2000 realization of the projection noise was generated in the same way as described in [23]. The histograms of the results can be found in Fig. 2.10(b) and Fig. 2.10(c). It can be seen that for the equidistant beam spacing geometry the proposed method produced less than 4 wrong pixels in more than 80% of the test cases, while using discrete pixel sums it obtained perfect reconstruction in almost all cases. Although this comparison was not performed on the same data as in [23], the results are comparable due to the large number of randomly generated test cases.

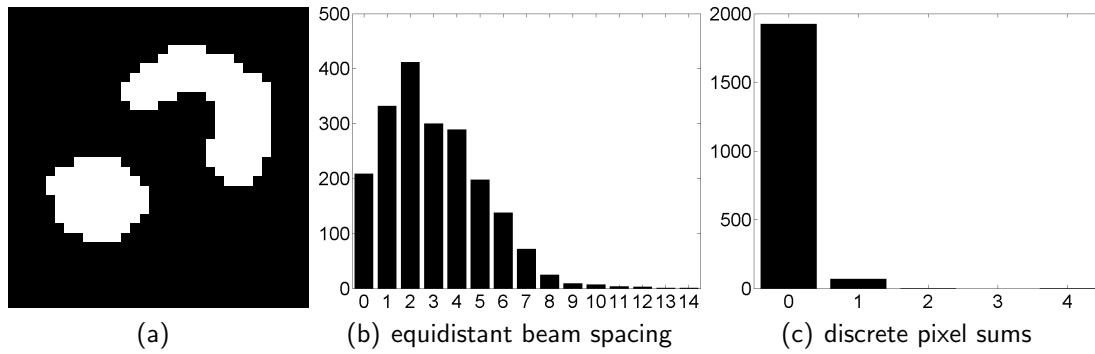


Figure 2.10: (a): The test image used in [23]. (b)-(c): The results obtained by our method on 2000 realization of the Poisson noise using 4 projections of the image on the left with the two types of projection geometries. The histogram of the number of missclassified pixels is plotted.

2.4.4 Experiments with Valid Direction Sets of Uniqueness

In [43], Hajdu addressed the problem of uniqueness in discrete tomography. It was shown that for any image size $h \times w$ there exists a valid (*i.e.*, non-trivial) set of four lattice directions, such that any two objects can be distinguished by means of their projections taken in those directions. In [20], it was shown that in fact whole families of such suitable sets of four directions can be found.

Let (a, b) be a lattice direction with $a, b \in \mathbb{Z}$ and a, b coprime. The projection value along a line parallel to (a, b) is computed as the sum of the intensities of those pixels (grid points) that fall on the given line. This kind of projection model can play an important role in the reconstruction of crystalline structures [14, 52].

In the first experiment, we considered 32×32 images and the direction set $\{(1, 3), (4, 5), (9, 7), (14, 15)\}$ which is valid for this size. This direction set provides a relatively large amount of projection information. It results in a total number of 1560 projection lines and an average of only ≈ 2.63 pixels contribute to each projection value. In the second case, we considered 255×255 images and the direction set $\{(1, 0), (0, 1), (126, 127), (127, 126)\}$ with 97536 projection lines and an average of only ≈ 2.66 pixels per projection value. Although theoretically these direction sets ensure the uniqueness of the reconstructions, it is not guaranteed that a heuristic algebraic reconstruction technique like the proposed method is able to find the exact solutions in all cases. Indeed, the problem of finding the exact reconstruction from four projections is generally NP-hard [36].

The aim of this experiment was to examine how often the proposed method is able to find the exact solution. Due to the high amount of perfect (*i.e.*, noiseless) projection information regularization was not required, thus the algorithm was applied

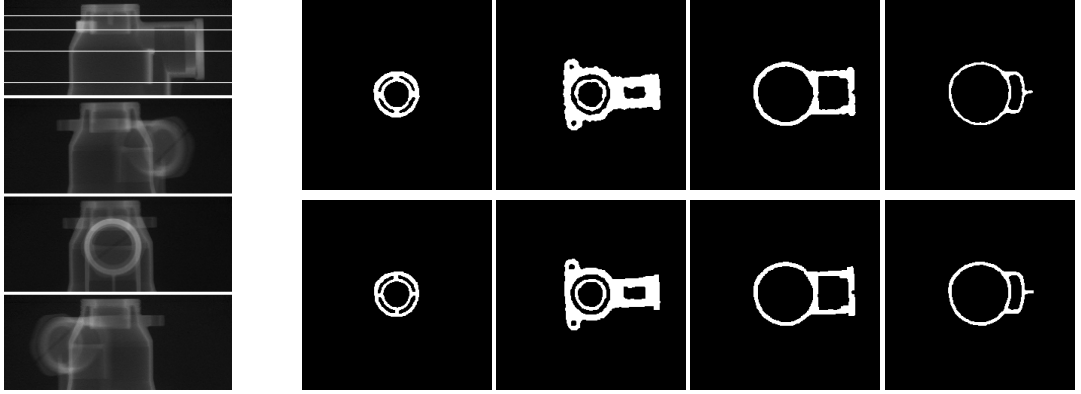


Figure 2.11: **First column:** Four of the 18 projections of the homogeneous part of a gas pressure regulator. **Other columns:** Reconstruction results of the PDM-DART method (first row) and the proposed method (second row). The slices are indicated on the first projection image.

with $\lambda = 0$. The proposed method was able to find the exact solutions on all of the 32×32 and the 255×255 versions of our phantom images (Fig. 2.5). For a further experiment, we generated 100 random binary images of size 255×255 . Each pixel of the images were set to 0 or 1 with probability 50%. The proposed method reconstructed 88 images perfectly, while misclassified 4-4 pixels on 12 images. We tested the PDM-DART method as well, which was also able to find the exact solutions on the phantom images, while perfectly reconstructed 87 random images and misclassified a total of 34 pixels. These results show that these methods are able to converge to the exact reconstructions in most cases when the images are uniquely determined by their projections.

2.4.5 Real Data Experiments

In [5], Balasko *et al.* compared the radiography and tomography possibilities of FRM-II and Budapest Research Reactor. Classical Filtered Back Projection reconstruction algorithm as well as discrete tomography methods have been tested. One of the experiments has been presented on a gas pressure regulator of which upper part contained only one material, thus binary tomography was possible. The reconstructions of the horizontal slices of the gas pressure regulator were performed using 18 projections distributed on the half circle. These neutron radiography pictures were acquired in the Radiography Station of the Budapest Research Reactor. The properties of the imaging system are presented in [5] and [4]. The slices of the gas pressure regulator were reconstructed separately, and fan beam geometry was assumed on each slice. Due to the distorting properties of the image acquiring systems, some preprocessing steps (e.g., intensity and uniformity corrections) were performed to prepare the projection

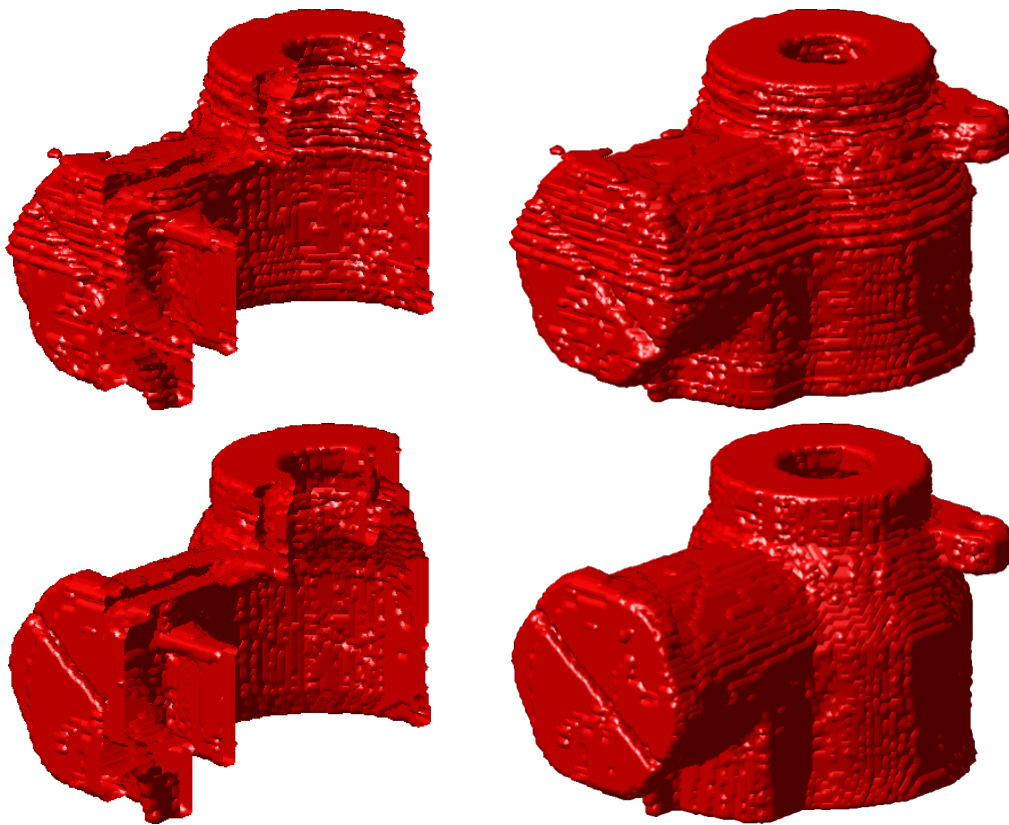


Figure 2.12: Gas pressure regulator. 3D visualization of the reconstruction results of the PDM-DART method (1st row) and the proposed method (2nd row). In the first column, half of the object is removed to show its inner structure.

data for reconstruction. Rigid-body registration algorithm was applied to correct image plate misplacements. To reduce complexity, the projections were cropped and the relevant parts were re-sampled. For discrete tomography, simulated annealing was chosen to minimize the energy function. One of the main difficulties was that the attenuation coefficients were unknown. These parameters were determined experimentally, by performing reconstructions assuming different intensity levels.

We tested both the PDM-DART algorithm and the proposed method using 18 preprocessed projections of the gas pressure regulator provided by the authors of [5]. Four of the projections and some reconstruction results can be seen in Fig. 2.11. Each line of the projections contained 309 measurements and the slices were reconstructed on an image grid of size 309×309 . A total of 88 horizontal slices of the homogeneous part of the gas pressure regulator has been reconstructed. To visualize the separately reconstructed slices altogether, 3D models has been created using marching cubes algorithm, see Fig. 2.12. Similarly to the synthetic phantom experiments, the main difficulty with the PDM-DART was to select the strength of the smoothing filter, that provides smooth results but preserves the small details of the images at the same time,

while the proposed method could be applied using the same parameter settings as in the synthetic tests.

2.5 Discussion

In the previous sections, we examined the performance of the PDM-DART method and the proposed method, and while both of the methods provided reliable results, the proposed method is found to be more robust to the projection noise. The two approaches are very different, and both methods have their advantages and drawbacks. The main advantage of the PDM-DART method is that it can be applied to multi-level images, while the proposed method has been developed to binary cases. In PDM-DART there are some parameters (e.g., the width of the boundaries on which the update steps are performed, the amount of randomly chosen non-fixed pixels, the strength of the smoothing operations applied in each iteration steps) that affects the accuracy of the reconstructions. These parameters have to be fine-tuned, and in our experiments we had to perform the PDM-DART with different smoothing kernels in each test case to find the settings that provides the best reconstruction. In contrary, the only parameter of the proposed method that could affect its results is the weight of the smoothing term, however, our method was able to provide competitive results using its same setting in all test cases, which shows, that the proposed method less sensitive to the strength of the projection noise.

In our experiments the proposed method provided reliable results using the smoothness prior $S(\mathbf{x})$ in both the weak and the strong noise cases. However, we did not focus on the choice of $S(\mathbf{x})$, which could be the subject of a future work. Theoretically other regularization terms can also be applied directly in the proposed energy functional. Also, the weight of the regularization can be chosen based on a priori information on the noise, while there are automatic regularization parameter estimation techniques [7].

2.5.1 Complexity

In the case of two-dimensional images, the algorithm stores the $\mathcal{O}(m\sqrt{n})$ non-zero entries of the sparse matrix W , since each projection measurement is computed using $\mathcal{O}(\sqrt{n})$ pixels. For the sparse matrix L and the current solution $\hat{\mathbf{x}}$ the method requires $\mathcal{O}(n)$ space, while the gradient of the energy function needs $\mathcal{O}(m)$ space. Computing the gradient and evaluating the energy functional in the gradient descent process requires $\mathcal{O}(m\sqrt{n} + n)$ operations (the matrix-vector multiplications and the evaluation of the polynomials in the discretization term and its gradient). The convergence of the

gradient based minimization processes is guaranteed, moreover, in our experiments the solutions always reached the desired discretization level. However this property as well as the number of required iterations and gradient descent steps should be examined theoretically in a future work.

2.6 Summary

In this chapter we presented a novel discrete tomography method which can be used in those cases when the gray-levels of the images are unknown. The reconstruction problem was traced back to the minimization of an appropriately constructed objective functional, in which, a higher-order statistics based discretization term enforces binary solutions. We introduced a graduated optimization scheme that increasingly enforces discrete solutions, while re-estimates the mid-level of the intensities in each iteration step. Experimental results on synthetic phantom images showed that the proposed optimization scheme can efficiently minimize the objective functional and the method provides high quality reconstructions. Limited angle experiments have been conducted, and the method has also been successfully applied on real data. Compared to state-of-the-art algorithms, the proposed method was found to be less sensitive to the strength of the projection noise since it was able to provide competitive results in each test case with the same parameter settings. The extension to multiple gray-levels might be feasible and can be an important direction of the future research.

2.7 Appendix: The Lipschitz-Continuity of the Discretization Term

Here we provide computational details on how to locally estimate an upper bound of the Lipschitz-constant $K^*(\nabla D)$ of the gradient of the discretization term. Using this estimation, a gradient descent step-length is given by $t^* = 1/K^*(\nabla E) = 1/(K^*(\nabla C) + \mu K^*(\nabla D))$, which does not increase the value of the objective functional.

The computations are independent of the value of the mid-level α , thus for the sake of simplicity we assume that $\alpha = 0$. All the equations remain valid if we perform substitutions $x_i = z_i - \alpha$ and $\mathbf{x} = \mathbf{z} - \alpha \mathbf{1}_n$. The discretization term is now given by:

$$D(\mathbf{x}) = n \frac{\|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^2} - 1, \quad (2.19)$$

where we use the notation of the general vector norm $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$. The first order derivatives can be computed as

$$\nabla_i D(\mathbf{x}) = \frac{\partial D(\mathbf{x})}{\partial x_i} = 4n \left(\frac{x_i^3}{(\|\mathbf{x}\|_2^2)^2} - \frac{x_i \|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^3} \right) \quad (2.20)$$

The discretization term D and its first and second order derivatives are continuous everywhere except at $\mathbf{0}_n$. Thus on a subset $\Omega \subseteq \mathbb{R}^n \setminus \mathbf{0}_n$, the Lipschitz constant of ∇D is given by $K(\nabla D) = \sup_{\mathbf{x} \in \Omega} \|J(\nabla D)(\mathbf{x})\|$, where $J(\nabla D)$ is the Jacobian matrix of ∇D (we note that $J(\nabla D)$ is equal to the Hessian of D). The Jacobian $J(\nabla D)$ is a $n \times n$ symmetric matrix and its diagonal and non-diagonal entries are given by:

$$J_{i,i}(\nabla D)(\mathbf{x}) = \frac{\partial^2 D(\mathbf{x})}{\partial x_i^2} = 4n \left(\frac{6x_i^2 \|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^4} - \frac{8x_i^4}{(\|\mathbf{x}\|_2^2)^3} + \frac{3x_i^2}{(\|\mathbf{x}\|_2^2)^2} - \frac{\|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^3} \right) \quad (2.21)$$

and

$$J_{i,j}(\nabla D)(\mathbf{x}) = \frac{\partial^2 D(\mathbf{x})}{\partial x_i \partial x_j} = 4n \left(\frac{6x_i x_j \|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^4} - \frac{4x_i^3 x_j + 4x_i x_j^3}{(\|\mathbf{x}\|_2^2)^3} \right), \text{ if } i \neq j. \quad (2.22)$$

We will estimate an upper bound for the $\|\cdot\|_1$ -norm of the Jacobian which is equal to its maximum absolute column sum (we note that $\|J(\nabla D)(\mathbf{x})\|_1 = \|J(\nabla D)(\mathbf{x})\|_\infty$ since $J(\nabla D)$ is symmetric). For that purpose consider the i th absolute column sum:

$$\begin{aligned} K_i &= |J_{i,i}(\nabla D)(\mathbf{x})| + \sum_{\substack{j=1 \\ j \neq i}}^n |J_{j,i}(\nabla D)(\mathbf{x})| \\ &\leq 4n \left(\frac{6x_i^2 \|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^4} + \frac{8x_i^4}{(\|\mathbf{x}\|_2^2)^3} + \frac{3x_i^2}{(\|\mathbf{x}\|_2^2)^2} + \frac{\|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^3} \right) \\ &\quad + 4n \sum_{\substack{j=1 \\ j \neq i}}^n \left(\frac{6|x_i x_j| \|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^4} + \frac{4|x_i^3 x_j| + 4|x_i x_j^3|}{(\|\mathbf{x}\|_2^2)^3} \right) \\ &= 4n \left(\frac{6x_i^2 \|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^4} + \frac{8x_i^4}{(\|\mathbf{x}\|_2^2)^3} + \frac{3x_i^2}{(\|\mathbf{x}\|_2^2)^2} + \frac{\|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^3} \right) \\ &\quad + 4n \frac{6|x_i| \|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^4} \left(\left(\sum_{j=1}^n |x_j| \right) - |x_i| \right) + 4n \frac{4|x_i|^3}{(\|\mathbf{x}\|_2^2)^3} \left(\left(\sum_{j=1}^n |x_j| \right) - |x_i| \right) \\ &\quad + 4n \frac{4|x_i|}{(\|\mathbf{x}\|_2^2)^3} \left(\left(\sum_{j=1}^n |x_j|^3 \right) - |x_i|^3 \right) \\ &= 4n \left(\frac{3x_i^2}{(\|\mathbf{x}\|_2^2)^2} + \frac{\|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^3} + \frac{6|x_i| \|\mathbf{x}\|_4^4}{(\|\mathbf{x}\|_2^2)^4} \left(\sum_{j=1}^n |x_j| \right) \right. \\ &\quad \left. + \frac{4|x_i|^3}{(\|\mathbf{x}\|_2^2)^3} \left(\sum_{j=1}^n |x_j| \right) + \frac{4|x_i|}{(\|\mathbf{x}\|_2^2)^3} \left(\sum_{j=1}^n |x_j|^3 \right) \right). \end{aligned} \quad (2.23)$$

Let $X = \max_i \{|x_i|\}$. Using the inequalities $|x_i| \leq X$, $\|\mathbf{x}\|_4^4 \leq nX^4$ and $\sum_{j=1}^n |x_j| \leq nX$, we obtain a simple upper bound for all of the absolute column sums, thus an upper bound for the $\|\cdot\|_1$ -norm of the Jacobian:

$$\|J(\nabla D)(\mathbf{x})\|_1 = \max_i \{K_i\} \leq 4n \left[\frac{3X^2}{(\|\mathbf{x}\|_2^2)^2} + \frac{9nX^4}{(\|\mathbf{x}\|_2^2)^3} + \frac{6n^2X^6}{(\|\mathbf{x}\|_2^2)^4} \right] \quad (2.24)$$

Hence, knowing only the maximal absolute value X and the norm $\|\mathbf{x}\|_2^2$ we can compute an upper bound for $\|J(\nabla D)(\mathbf{x})\|$ using Eq. (2.24). This also follows that if on a set $\Omega_{S,M} \subseteq \mathbb{R}^n$, the maximum absolute elements $\max_i \{|x_i|\}$ are upper bounded while $\|\mathbf{x}\|_2^2$ has a lower bound, then ∇D is Lipschitz continuous on that set, i.e., if

$$\Omega_{S,M} = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \max_i \{|x_i|\} \leq M, S \leq \|\mathbf{x}\|_2^2 \right\}, \quad (2.25)$$

then

$$K(\nabla D) = \sup_{\mathbf{x} \in \Omega_{S,M}} \|J(\nabla D)(\mathbf{x})\| \leq 4n \left[\frac{3M^2}{S^2} + \frac{9nM^4}{S^3} + \frac{6n^2M^6}{S^4} \right]. \quad (2.26)$$

Now we construct the set of the next possible solutions. Suppose that the current solution is $\hat{\mathbf{x}}$ and let $\hat{\mathbf{y}} = \hat{\mathbf{x}} - (1/K^*(\nabla C))\nabla E_{\hat{\alpha},\mu}(\hat{\mathbf{x}})$. Consider the line segment that connects $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$:

$$\Lambda = \{\ell\hat{\mathbf{x}} + (1-\ell)\hat{\mathbf{y}} \mid 0 \leq \ell \leq 1\}. \quad (2.27)$$

Since we are looking for a step-length $t = 1/(K^*(\nabla E)) = 1/(K^*(\nabla C) + \mu K^*(\nabla D))$, thus $t \leq 1/K^*(\nabla C)$ and after the gradient descent step the next solution $\hat{\mathbf{x}}'$ will fall on the line segment Λ . Thus we will compute an upper bound of $K(\nabla D)$ on Λ . For that purpose, consider the largest absolute value of the components of the vectors in Λ , i.e., let $M = \max_{\mathbf{x} \in \Lambda} \{\max_i |x_i|\}$ and the square of the smallest vector norm in Λ , i.e., let $S = \min_{\mathbf{x} \in \Lambda} \{\|\mathbf{x}\|_2^2\}$. Since Λ is a line segment, thus M can be computed considering only its end points $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$:

$$M = \max \left\{ \max_i \{|\hat{x}_i|\}, \max_i \{|\hat{y}_i|\} \right\}, \quad (2.28)$$

while S can be computed as the square of the minimal norm of the points on the line segment Λ :

$$S = \begin{cases} \|\hat{\mathbf{x}}\|_2^2 & \text{if } k \leq 0 \\ \|\hat{\mathbf{y}}\|_2^2 & \text{if } k \geq 1 \\ \|\hat{\mathbf{x}} + k(\hat{\mathbf{y}} - \hat{\mathbf{x}})\|_2^2 & \text{otherwise,} \end{cases} \quad (2.29)$$

where $k = \frac{\langle \hat{\mathbf{x}}, \hat{\mathbf{x}} - \hat{\mathbf{y}} \rangle}{\|\hat{\mathbf{x}} - \hat{\mathbf{y}}\|_2^2}$ and $\langle \cdot | \cdot \rangle$ denotes the dot-product operation. It can be easily seen, that $\Lambda \subseteq \Omega_{S,M}$, thus an upper bound of $K^*(\nabla D)$ on Λ is given by Eq. (2.26).

In the first step of the proposed method where the initial continuous reconstruction is obtained by the minimization of $E_{0,0}(\mathbf{x}) \equiv C(\mathbf{x})$, setting $t = 1/K^*(\nabla C)$ as the gradient descent step size will provide convergence to the global optimum. In the discretization phase where the objective function $E_{\hat{\alpha},\mu}$ is minimized with $\mu > 0$, before each gradient descent step, M and S can be computed for the current solution $\hat{\mathbf{x}}$ using Eq. (2.28) and Eq. (2.29), and then $K^*(\nabla D)$ can be estimated using Eq. (2.26). Then using step size $t^* = 1/K^*(\nabla E) = 1/(K^*(\nabla C) + \mu K^*(\nabla D))$ ensures that the gradient descent step will not increase the function value, thus the iterations will converge. In our experiments, the proposed method provided similar results using this step-length estimation as with the line search method, we found that the estimation of S and M before each descent step was computationally less efficient in our implementation. On the other hand in a parallel implementation, the computation of S and M could be accelerated, and the step-length t^* can be considered as an alternative to the line search.

Chapter 3

Binary Shape Deconvolution Using Discrete Tomography

In [94], Sharif *et al.* described a discrete tomography based binary deconvolution algorithm. The size of the point-spread-function (PSF) was restricted to 3×3 and the restoration of the projections was handled as a simple linear inverse problem. Moreover, the method applied the Ryser's algorithm [88] to reconstruct the shapes, followed by a further time-consuming step to resolve the switching component ambiguities.

In this chapter a Tikhonov-regularization [101] based method is proposed to deconvolve the projections of the binary shape. The main advantage of this approach is that the best value of the regularization parameter can be automatically set using the L-curve method [44]. In the second step, the binary shape is reconstructed from the deconvolved projections using efficient network-flow discrete tomography algorithm [15]. The examine the efficiency of the method on synthetic as well as on real images and compare it to a well-known, widely-used method.

3.1 The Restoration Method

The goal of image restoration is to reconstruct an original image $f \in \mathbb{R}^{u \times v}$ from its degraded observation $g \in \mathbb{R}^{u \times v}$ which is blurred and noisy. The degradation process (as usually) is modeled as the convolution of the original image with the points-spread-function (PSF) and the addition of some noise:

$$g = h * f + n, \quad (3.1)$$

where $h \in \mathbb{R}^{p \times q}$ is the point spread function, $*$ denotes the 2-dimensional convolution operator and $n \in \mathbb{R}^{u \times v}$ is the additive noise. Here we assume that the PSF is known. Our goal is to restore binary images, such as $b \in \{0, 1\}^{u \times v}$. However, in the case of

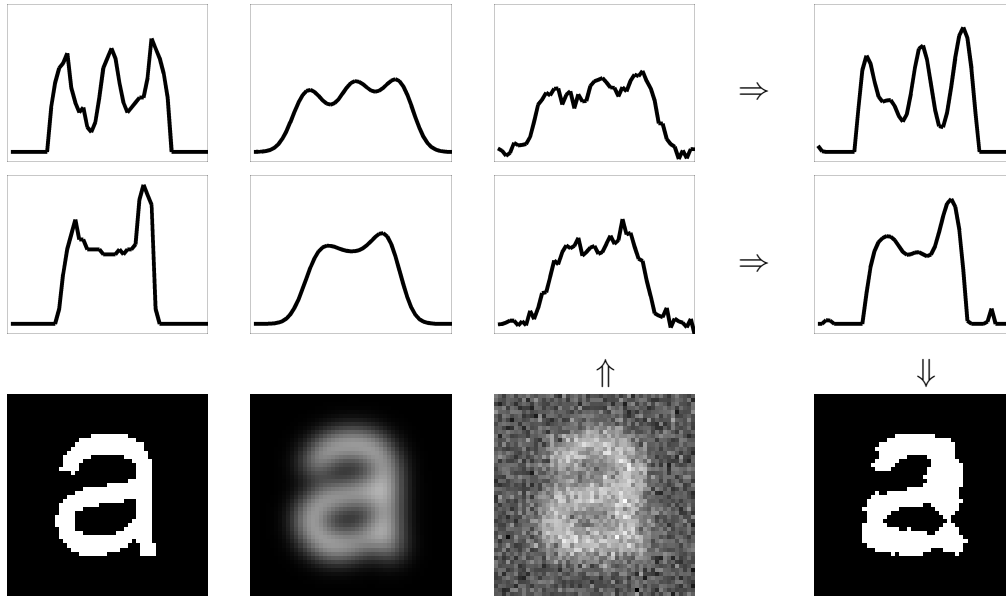


Figure 3.1: The degradation model and the basic idea of the proposed method. **First three columns:** the original binary image, the blurred image and the blurry and noisy image and their horizontal (top) and vertical (middle) projections. **Fourth column:** the method restores the projections of the binary image from the projections of the degraded image then reconstructs the binary shape.

real images the scaling of the intensity values can be arbitrary. Therefore we introduce the unknown scale factor $s \in \mathbb{R}$ and

$$f = sb \quad (3.2)$$

in the degradation model Eq. (3.1).

Let denote by \check{f}^θ the projection vector of an image f along an arbitrary directional angle θ , i.e., \check{f}^θ is the discrete approximation of the well-known Radon transform [28] in direction θ (we simply write \check{f} when the actual direction is not important). Here we take advantage of the following property of the Radon transform: the result of the Radon transform of a two dimensional convolution is the same as the one dimensional convolution of the Radon transformed functions, i.e., if $g = h * f$ then $\check{g} = \check{h} * \check{f}$, where $*$ denotes the 1-dimensional convolution operator. A similar relation is true for the addition. Thus the relationship between the projections of the original and the degraded images can be expressed as

$$\check{g} = \check{h} * \check{f} + \check{n}. \quad (3.3)$$

For a direction set $\Omega = \{\theta_i | i = 1, \dots, k\}$ the proposed method first estimates each projection \check{f}^{θ_i} of the unknown original image from the projections \check{g}^{θ_i} of the degraded

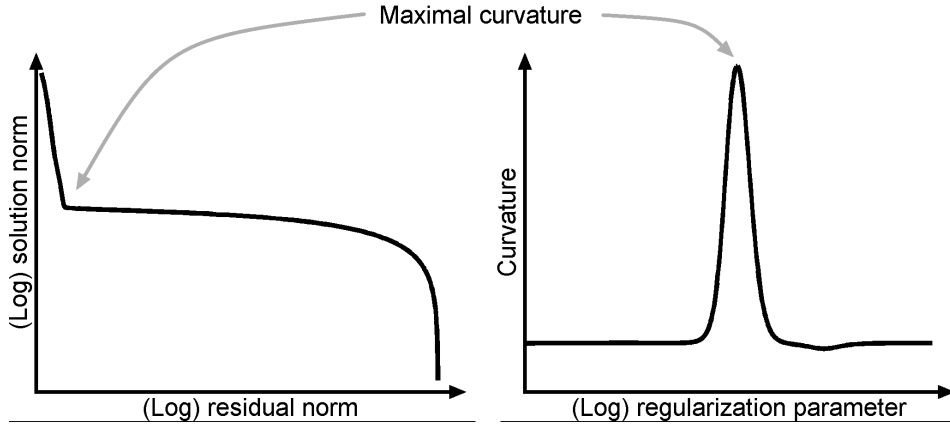


Figure 3.2: A typical L-curve and its curvature. The best trade-off between the norm of the residual and the norm of the solution is assigned by the maximal curvature value.

image and then as a second step reconstructs f from its estimated projections (see Fig. 3.1).

3.1.1 Deconvolution of the Projections

In this section we introduce a Tikhonov-regularization [101] based method for the deconvolution of the projections. Convolution is a linear operation, thus Eq. (3.3) can be written as

$$\check{g} = H\check{f} + \check{n}, \quad (3.4)$$

where the H matrix represents the convolution operation with the filter \check{h} . Since \check{n} is unknown, this system of linear equations is ill-posed and it requires regularization which penalizes solutions of large norm. The standard version of the Tikhonov regularization takes the form

$$\check{f}_\lambda = \arg \min_{\check{f}} \|H\check{f} - \check{g}\|_2^2 + \lambda^2 \|\check{f}\|_2^2, \quad (3.5)$$

where λ is a positive constant, the regularization parameter, that controls the smoothness of the solution. As λ increases the norm of the residual $\|H\check{f}_\lambda - \check{g}\|_2^2$ also increases (i.e., \check{f}_λ becomes less accurate), while the norm of the solution $\|\check{f}_\lambda\|_2^2$ decreases (i.e., \check{f}_λ becomes smoother). An explicit solution for a given value of λ is given by

$$\check{f}_\lambda = (H^T H + \lambda^2 I)^{-1} H^T \check{g}, \quad (3.6)$$

where I denotes the identity matrix. To determine a suitable value of the regularization parameter λ we used the so-called L-curve method [44]. The L-curve is the log-log plot of the norm of the residual and the norm of the solution for different regularization parameters:

$$\mathcal{L} = \{(\log_2 \|H\check{f}_\lambda - \check{g}\|_2^2, \log_2 \|\check{f}_\lambda\|_2^2), \lambda \geq 0\}. \quad (3.7)$$

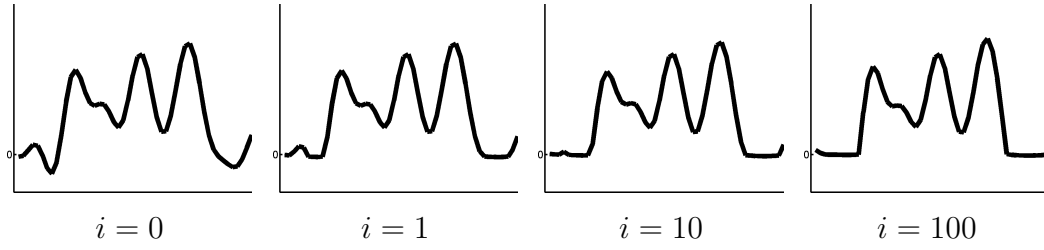


Figure 3.3: The evolution of the estimation of the projection vector during the iteration.

To obtain an optimal trade-off between the two values, the L-curve method proposes to choose λ^* which maximizes the curvature κ_λ of the curve \mathcal{L} (see Fig. 3.2). Although the L-curve method requires for a large set of different λ values to evaluate Eq. (3.6), it is still computationally efficient approach.

Since the elements of the projection vectors are non-negative we are interested in a solution of Eq. (3.4) such that $\check{f} > 0$. Unfortunately there is no simple explicit solution for this problem, but a typical iterative approach can be applied. For that purpose, we denoted by \check{f}^+ the version of the vector \check{f} in which all negative values were set to 0. Let $\check{f}_0 = \check{f}_{\lambda^*}$ the initial estimation. In each iteration step the method estimates \check{f}_i that approximates \check{f}_{i-1}^+ (see Fig. 3.3). This can be formulated by the following minimization problem:

$$\check{f}_i = \arg \min_{\check{f}} \|H\check{f} - \check{g}\|_2^2 + \lambda^{*2} \|\check{f}\|_2^2 + \|\check{f} - \check{f}_{i-1}^+\|_2^2 \quad (3.8)$$

which introduces an additional term to Eq. (3.5) that prefers similarity to the non-negative version of the previous estimation. This form of the *generalized* Tikhonov regularization has an explicit solution:

$$\check{f}_i = (H^T H + (1 + \lambda^{*2})I)^{-1} (H^T \check{g} + \check{f}_{i-1}^+). \quad (3.9)$$

The iteration ends if $\|\check{f}_i - \check{f}_{i-1}\|_2^2 < \epsilon \|\check{f}_{i-1}\|_2^2$ for an appropriately small $\epsilon > 0$. The resulting vector $\check{f} = \check{f}_i^+$ will be the estimation of the projection vector. The pseudo code of this algorithm can be found in Algorithm 3. Furthermore the residual $\|H\check{f} - \check{g}\|_2^2$ approximates \check{n} and thus the variance σ_n^2 of the additive noise n can be easily estimated.

3.1.2 Binary Tomographic Reconstruction

In this section we propose a method to reconstruct the binary image from the vectors \check{f}^{θ_i} using a standard discrete tomography technique. The reconstruction requires the $\check{b}^{\theta_i} = \check{f}^{\theta_i}/s$ binary projections. Unfortunately the scale factor s can not be determined explicitly but the upper and lower bounds of its possible values can be easily estimated. For that purpose, consider the maximal scaled projection value $M = \max_{\theta_i} \{\max\{\check{f}^{\theta_i}\}\}$

Algorithm 3: Deconvolution of a projection vector**Input** : The \check{g} projection of the degraded image and the PSF**Output:** The reconstructed vector \check{f}

- 1 Find λ^* and \check{f}_{λ^*} using the L-curve method and standard Tikhonov method (Eq. (3.6))
- 2 Let $\check{f}_0 = \check{f}_{\lambda^*}$
- 3 **repeat**
- 4 Update the estimation using generalized Tikhonov method (Eq. (3.9))
- 5 **until** $\|\check{f}_i^+ - \check{f}_{i-1}^+\|_2^2 < \epsilon \|\check{f}_{i-1}^+\|_2^2$
- 6 Return with $\check{f} = \check{f}_i^+$.

and let D be the length of the segment of the image that the projection ray corresponding to M intersects. It can be assumed that $1 \leq M/s \leq D$ so we can define a set of the possible values of the scale factor s as follows:

$$S = \{k/M, k = 1, \dots, \lfloor D \rfloor\}. \quad (3.10)$$

For each $s \in S$ we obtain the vectors $\check{b}^{\theta_i} = \lfloor \check{f}^{\theta_i}/s \rfloor$ as estimations of the projections of the unknown binary image b .

Here we consider the reconstruction from the vertical and horizontal projection vectors \check{b}^0 and $\check{b}^{\pi/2}$. The discrete tomography technique described in this section requires that these two vectors have equal sum. However, in practice this is not guaranteed since these vectors are obtained by rounding. To equalize the two vectors we uniformly decreased the elements of the vector of the largest sum. We decreased by 1 the highest elements of the same number as the difference between the two vectors.

It is well known, that usually two projections are not enough to reconstruct binary images, *i.e.*, there can be many binary images that have the same projections while in some other cases there is no binary image that satisfies the projections. We define the tomographic equivalence class:

$$U = U(\check{b}^0, \check{b}^{\pi/2}) = \{z \in \{0, 1\}^{u \times v} : \check{z}^0 = \check{b}^0, \check{z}^{\pi/2} = \check{b}^{\pi/2}\}. \quad (3.11)$$

Although in general case $|U| > 1$, we are interested in a solution which is the most similar to the input image g . Fortunately finding such a solution is quite straightforward. For that purpose a model image m has been created from the input image g by removing noise of variance σ_n^2 (which has been estimated after the estimation of the projections, see Section 3.1.1). The expectation that the solution z_s should be similar to the model image means that on those (x, y) positions where $z(x, y) = 1$ it is most likely that the model image has high grayscale values. This can be formulated by the

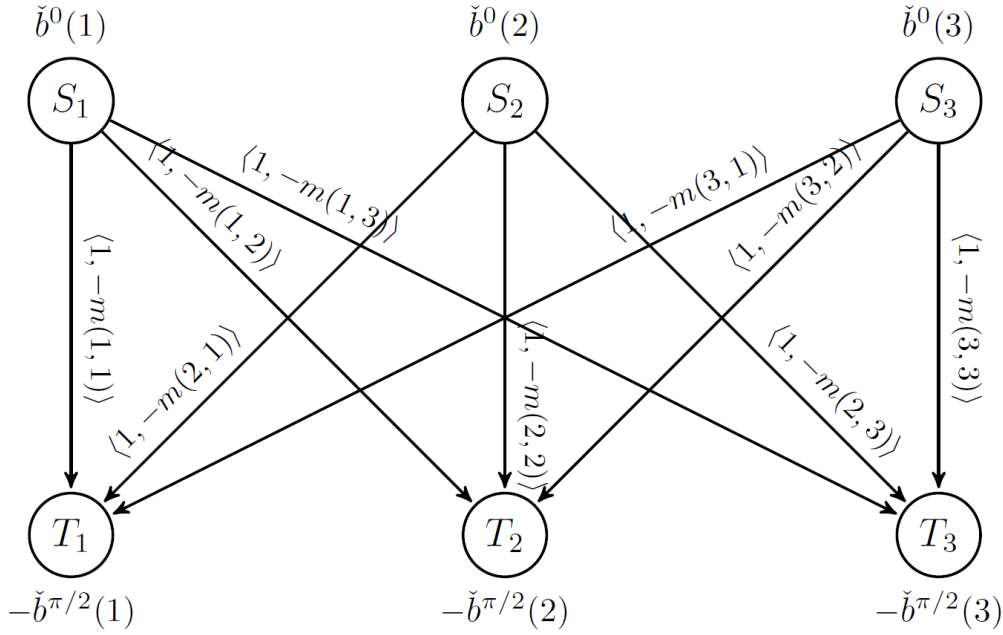


Figure 3.4: Associated network for a 3×3 case. The supply and demand values for the source nodes S_i and the sink nodes T_j are given by the vectors \check{b}^0 and $\check{b}^{\pi/2}$ respectively. Each of the edges has a flow capacity equal to 1 while the flow costs are given by the model image $m \in \mathbb{R}^{3 \times 3}$.

following minimization problem:

$$z_s = \arg \min_{z \in U} \left(- \sum_{x,y} z(x,y) m(x,y) \right). \quad (3.12)$$

This way the reconstruction problem can be traced back to the *minimum cost maximal flow* (MCMF) problem [15]. In this network the *supply* and *demand* nodes are representing the projection vectors while the edges are representing the image pixels (see Fig. 3.4). The edge $S_x \rightarrow T_y$ corresponding to the (x,y) pixel has a flow capacity equal to 1 and a flow cost equal to $-m(x,y)$. The minimum cost maximal flow can be found in polynomial time [80] and it determines the solution z_s of the discrete tomography problem. The pixel $z_s(x,y)$ gets a value of 1, if and only if the flow passes through the edge $S_x \rightarrow T_y$. As it has been noted previously in some cases the projections can not be satisfied, *i.e.*, U is empty. In such cases we simply discard the current scale factor and the corresponding projections. We also note that the original input image g could also be used as model image, however the results would be somewhat more noisy.

For different scale factors $s \in S$ the method finds different binary solutions z_s (see Fig. 3.5). To choose an optimal solution z , the method compares each z_s to the input

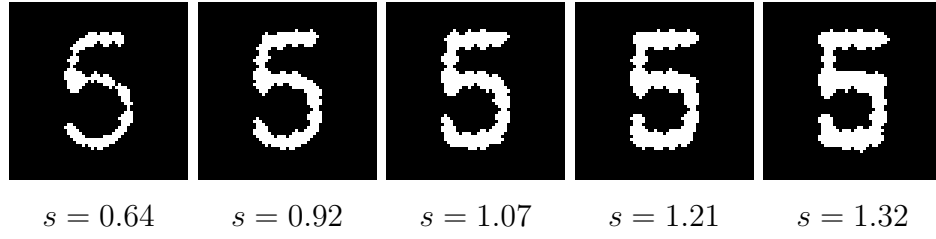


Figure 3.5: Reconstruction results for different $s \in S$ scale values.

image g in least-squares sense

$$z = \arg \min_{z_s} \|s(h * z_s) - g\|^2. \quad (3.13)$$

It should be noted that theoretically it is not impossible that $|U| = 0$ for each scale factor $s \in S$. In this case the proposed method could not provide any solution. However in practice we found that U contains at least one solution for most of $s \in S$.

The pseudo code of the algorithm can be found in Algorithm 4. As for the time complexity of the overall algorithm, the reconstruction of the projections require a matrix inversion in each iteration step which has a polynomial time consumption while the reconstruction of the binary shapes for each scale factor $s \in S$ using the MCMF approach has also polynomial time complexity. Hence the overall method is computationally efficient.

3.2 Experiments and Comparison

We examined the performance of the proposed method on synthetic dataset of images of 62 alphanumeric characters of size 59×59 pixels and their 1550 degraded versions. Each image was blurred by Gaussian filters with standard deviations $\sigma = 0, 1, \dots, 5$.

Algorithm 4: Shape restoration using discrete tomography

Input : The degraded grayscale image g and the PSF h

Output: The reconstructed binary image z

- 1 Restore the projections \check{f}^0 and $\check{f}^{\pi/2}$ using the method described in Section 3.1.1
 - 2 **foreach** $s \in S$ **do**
 - 3 Reconstruct the binary image z_s from the vectors $\check{b}^0 = \lceil \check{f}^0 / s \rceil$ and $\check{b}^{\pi/2} = \lceil \check{f}^{\pi/2} / s \rceil$ using the method described in Section 3.1.2
 - 4 **end**
 - 5 Return with $z = z_s$ according to Eq. (3.13)
-

Richardson-Lucy method					
σ	5dB	2.5dB	0dB	-2.5dB	-5dB
1.0	9.93	12.66	15.94	21.78	30.58
2.0	12.01	14.57	17.80	23.99	31.21
3.0	10.30	13.06	16.85	21.96	29.98
4.0	14.47	16.98	20.30	26.44	35.26
5.0	21.28	24.04	27.07	32.87	41.44

Proposed method					
σ	5dB	2.5dB	0dB	-2.5dB	-5dB
1.0	10.83	12.31	14.44	18.88	24.15
2.0	14.67	15.69	17.19	20.80	22.61
3.0	17.41	18.34	19.94	21.66	24.58
4.0	21.53	23.38	25.17	27.53	29.44
5.0	26.47	28.43	30.90	32.50	35.75

Table 3.1: Test results on the synthetic dataset of the Richardson-Lucy method and the proposed method. The average δ measures on 62 alphanumeric characters for different levels of blur and SNR are shown. The cases when the proposed method provided better results are depicted in bold.

White noises of different levels was added to each blurred image to implement signal-to-noise ratios of 5, 2.5, 0, -2.5, -5 dB. The reconstruction results were evaluated using the following normalized symmetric difference

$$\delta = \frac{|B \Delta Z|}{|B| + |Z|} \cdot 100\%, \quad (3.14)$$

where B and Z denote the set of the foreground pixels of the original image b and the reconstructed image z respectively.

For comparison, we have chosen the Richardson–Lucy algorithm [84] because this method also assumes that the PSF is known and does not require an estimation of the strength of the noise. On the other hand this method was developed for grayscale images thus thresholding is required. In our experiments Otsu [81] thresholding provided the best solutions. To improve the results of both methods, small isolated pixel groups has been removed from the reconstructed images by the following way. First, a morphological closing has been applied to connect regions and isolated pixels that are close to each other. Then 8-connected pixel groups containing at most a certain

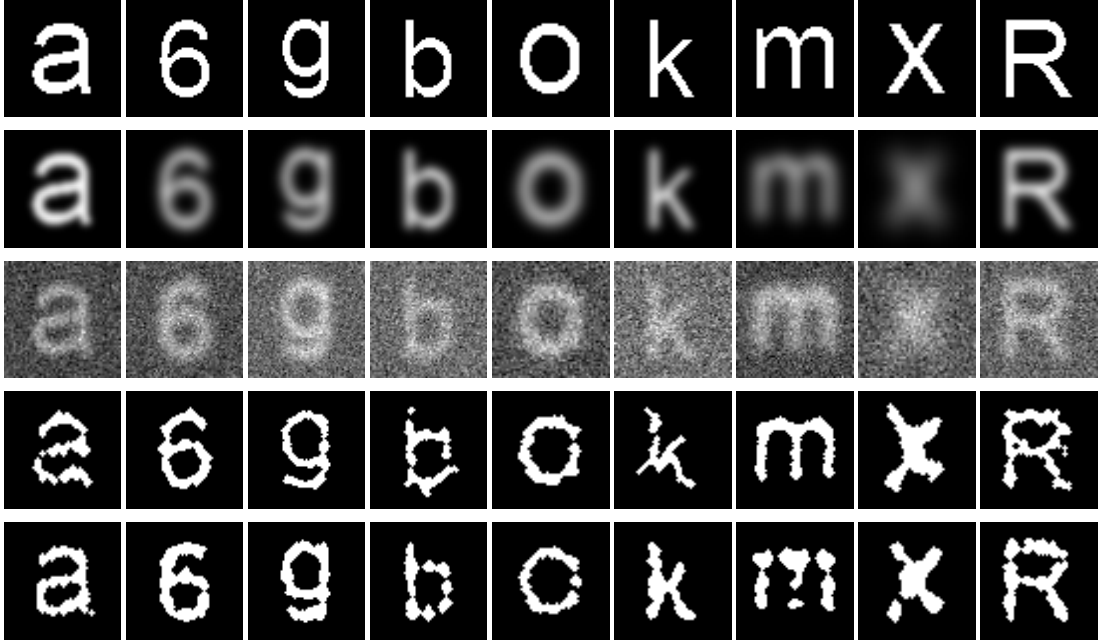


Figure 3.6: Example reconstruction results on synthetic images. **First three rows:** The original binary shapes, the blurred images, and the blurred and noisy version (from top to middle). **Fourth and fifth rows:** The results of the Richardson-Lucy method and the proposed method respectively.

number of pixels has been removed. Finally, a *logical AND* operation with the original shape has been applied to undo the morphological closing. It has to be noted that the Richardson-Lucy method produced much more of such artifacts. The summary of the test results is shown in Table 3.1. It can be seen that the proposed method provided better results in low SNR cases. This is due to the L-curve based estimation of the regularization parameter which implicitly gives a robust estimation of the level of the noise. On the other hand, in higher SNR cases the 2-dimensional deconvolution methods (like the Richardson-Lucy algorithm) can provide more accurate restorations.

The Richardson-Lucy algorithm requires a couple of Fourier transforms in each iteration step, thus for fixed number of iterations, its time complexity is $\mathcal{O}(uv \log(uv))$. As we discussed in the previous section, the proposed method has a polynomial time complexity. We implemented our method in Matlab, but the network-flow algorithm was written in C. This inefficient mixed solution resulted longer running times (an average of ≈ 5 sec.), while the Richardson-Lucy algorithm was very fast (below 0.1 sec.). However a more efficient implementation of our method is probably feasible.

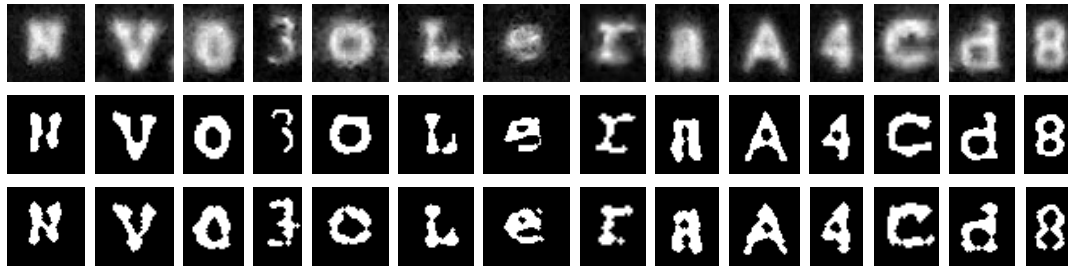


Figure 3.7: Example results on letters extracted from out-of-focus document images taken with EDoF camera. The input images are shown in the first row. The reconstruction results of the Richardson-Lucy method can be found in the second row and the results of the proposed method are in the third row.

3.2.1 Experiments on Real Out of Focus Images

Nowadays optical character recognition (OCR) is one of the most important functionalities of mobile devices. Nevertheless many modern devices are equipped with fixed-focus or full-focus (Extended Depth of Field - EDoF) lens. While these cameras in many cases provide better image quality than many auto-focus cameras, they usually can not be used for OCR because the images of A4 pages or business cards taken from a distance about 20 cm are out-of-focus. Therefore the enhancement of such images is an interesting task. We extracted a set of letters from images of A4 papers taken with an EDoF camera. Gaussian filters has been used as PSFs and their radius has been determined empirically. The results of the Richardson-Lucy method and the proposed method are shown in Fig. 3.7. It can be observed that the proposed method provided better results in same cases, thus it can be regarded as a promising alternative. We note that it is not clear that Gaussian filter approximates the best the out-of-focus effect. Disk shaped low-pass filters was also tested but while we obtained similar results with the proposed method, the Richardson-Lucy method provided much weaker restoration.

3.3 Summary

In this chapter we described a binary tomography based image restoration method. The first step of the method restores the projections of the shape using 1-dimensional deconvolution while the second step reconstructs the binary image using a discrete tomography technique. The main advantage of the method is that it does not require parameter setting or prior knowledge like the estimation of the signal-to-noise ratio. The optimal trade-off between the residual and the accuracy is determined automati-

cally using L-curve method. Numerical experiments on synthetic images showed that the method is robust to the level of the noise and can outperform a widely-used method in low SNR cases. The efficiency of the method has also been demonstrated on real alphanumeric images taken with fixed-focus camera.

In the described version of the algorithm, the intensity level of the shapes is found by performing tomography reconstructions assuming different possible values. For each value, we perform a tomographic reconstruction, and the best intensity level is found by comparing the result to the input image. This procedure is required because the applied binary tomographic technique assumes that the intensity levels are known (*i.e.*, 0 for the background and 1 for the shape pixels). Furthermore, the described reconstruction method is restricted to two projections.

Therefore in the future research, the algorithm could be improved by applying other tomography reconstruction algorithms. For example, the method described in Chapter 2 can deal with the problem of unknown intensity levels, and moreover, it allows to use more than two projections, which would probably improve the accuracy of the restoration.

Chapter 4

Nonlinear Registration of Binary Shapes

In this chapter, a general registration framework is proposed which can be used to estimate the parameters of nonlinear transformations between binary shapes without established point correspondences. The basic idea of the method is to set up a system of nonlinear equations by integrating a set of linearly independent nonlinear functions over the image domains. The system is then solved using the classical *Levenberg-Marquardt algorithm* [66]. The framework is applied to different nonlinear transformation models (e.g., planar homography, polynomial or thin plate spline transformations), and the performance and robustness of the method is demonstrated using tests on large set of synthetic images. The method has also been successfully applied in the case of different real-world applications, e.g., registration of X-ray images, traffic signs and handwritten characters, and the visual inspection of printed signs on tubular structures.

4.1 A Parametric Reconstruction Framework

In this section we introduce a novel general registration method. Suppose that $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a differentiable and invertible transformation (see Fig. 4.1), and its inverse is also differentiable (i.e., φ is a *diffeomorphism*). Our goal is to estimate its parameters with which the transformation aligns the shapes. Let the corresponding point coordinates on the template and observation shapes denoted by $\mathbf{x} = [x_1, x_2]^T \in \mathbb{R}^2$ and $\mathbf{y} = [y_1, y_2]^T \in \mathbb{R}^2$. Then the identity relation between these points can be written as followings:

$$\mathbf{y} = \varphi(\mathbf{x}) \quad \Leftrightarrow \quad \mathbf{x} = \varphi^{-1}(\mathbf{y}), \quad (4.1)$$

where $\varphi^{-1} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is the corresponding inverse transformation.

Consider the shapes as their characteristic function $\mathbb{1} : \mathbb{R}^2 \rightarrow \{0, 1\}$, where 0 and

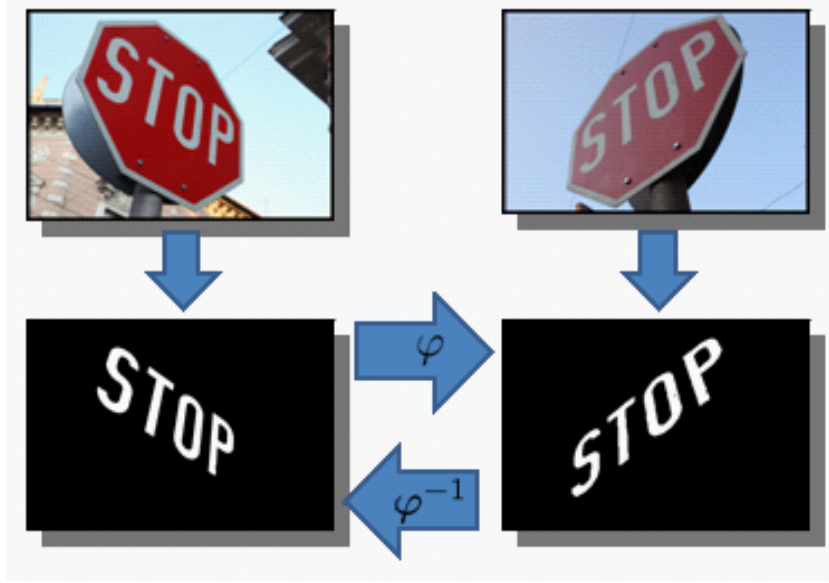


Figure 4.1: Problem statement: Estimation of the parameters of the aligning transformation (in this case a planar homography transformation) using the segmented versions of the images.

$\mathbb{1}$ denotes the background and the foreground points respectively. If the *template* and the *observation* shapes are denoted by $\mathbb{1}_t$ and $\mathbb{1}_o$, then

$$\mathbb{1}_o(\mathbf{y}) = \mathbb{1}_o(\varphi(\mathbf{x})) = \mathbb{1}_t(\mathbf{x}), \quad (4.2)$$

since \mathbf{x} and \mathbf{y} are corresponding point coordinates.

Classical correspondence based registration approaches would set up a system of equations from Eq. (4.1), of which solution gives the aligning parameters. Herein, however, we are looking for a correspondenceless solution. We can obtain a relation between the area of the shapes by integrating out individual point correspondences over the foreground regions and applying the integral transformation $\mathbf{y} = \varphi(\mathbf{x})$, $d\mathbf{y} = |J_\varphi(\mathbf{x})| d\mathbf{x}$:

$$\int_{\mathbb{R}^2} \mathbb{1}_o(\mathbf{y}) d\mathbf{y} = \int_{\mathbb{R}^2} \mathbb{1}_t(\mathbf{x}) |J_\varphi(\mathbf{x})| d\mathbf{x}, \quad (4.3)$$

where the Jacobian $|J_\varphi| : \mathbb{R}^2 \rightarrow \mathbb{R}$ of the transformation φ is given by:

$$|J_\varphi(\mathbf{x})| = \begin{vmatrix} \frac{\partial \varphi_1}{\partial x_1} & \frac{\partial \varphi_1}{\partial x_2} \\ \frac{\partial \varphi_2}{\partial x_1} & \frac{\partial \varphi_2}{\partial x_2} \end{vmatrix}. \quad (4.4)$$

The value of the Jacobian $|J_\varphi(\mathbf{x})|$ gives the area changing caused by the transformation φ at the point \mathbf{x} (*i.e.*, the measure of the transformation). We note that in the case of affine (linear) transformations this measure is constant, and can be estimated simply

as the ratio of the areas of the shapes [31]. However the Jacobian of a nonlinear transformation is a non-constant function of the point coordinates.

Note that the effect of the multiplication with the characteristic functions basically only restricts the integral domains to the foreground regions. Let the foreground regions of the *template* and *observation* shapes denoted by $\mathcal{F}_t = \{\mathbf{x} \in \mathbb{R}^2 | \mathbb{I}_t(\mathbf{x}) = 1\}$ and $\mathcal{F}_o = \{\mathbf{y} \in \mathbb{R}^2 | \mathbb{I}_o(\mathbf{y}) = 1\}$. Thus

$$\int_{\mathbb{R}^2} \mathbb{I}_o(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{F}_o} d\mathbf{y}, \quad (4.5)$$

and Eq. (4.3) can be written as

$$\int_{\mathcal{F}_o} d\mathbf{y} = \int_{\mathcal{F}_t} |J_\varphi(\mathbf{x})| d\mathbf{x}. \quad (4.6)$$

Thus we obtained an equation with unknowns corresponding to the parameters of the transformation (e.g., in the case of a planar homography, the H_{11}, \dots, H_{32} parameters, see Section 4.2.1). Obviously, the left hand side of the equation gives the area of the *observation*, while the right hand side gives the area of the transformed *template*. Thus Eq. (4.6) expresses the identity of these areas, i.e., that $|\varphi(\mathcal{F}_t)| = |\mathcal{F}_o|$. Since generally the transformation φ has more than 1 parameters (or *degree of freedom*), only one equation is not enough to estimate these unknowns.

Additional constraints can be obtained by multiplying Eq. (4.1) and Eq. (4.2):

$$\mathbf{y} \mathbb{I}_o(\mathbf{y}) = \varphi(\mathbf{x}) \mathbb{I}_t(\mathbf{x}). \quad (4.7)$$

Integrating both sides gives an equation similar to Eq. (4.6):

$$\int_{\mathcal{F}_o} \mathbf{y} d\mathbf{y} = \int_{\mathcal{F}_t} \varphi(\mathbf{x}) |J_\varphi(\mathbf{x})| d\mathbf{x}. \quad (4.8)$$

The transformation φ can be decomposed as

$$\varphi(\mathbf{x}) = [\varphi_1(\mathbf{x}), \varphi_2(\mathbf{x})]^T, \quad (4.9)$$

where $\varphi_1, \varphi_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ are the coordinate functions. Thus the integrands in Eq. (4.8) are (two dimensional) vectors and the values of the integrals are also vectors. Thus Eq. (4.8) can be decomposed into a system of two equations using these coordinate functions:

$$\int_{\mathcal{F}_o} y_i d\mathbf{y} = \int_{\mathcal{F}_t} \varphi_i(\mathbf{x}) |J_\varphi(\mathbf{x})| d\mathbf{x}, \quad i = 1, 2. \quad (4.10)$$

Although these equations give additional constraints, this system of equations is still not enough to estimate all of the unknown parameters of φ .

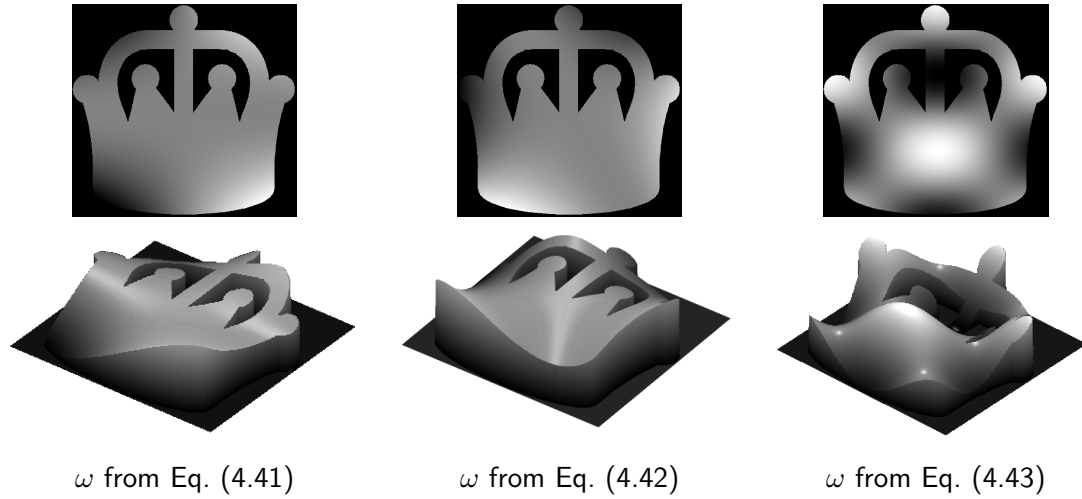


Figure 4.2: The effect of various ω functions. **Top**: the generated coloring of a binary shape. **Bottom**: the corresponding volumes.

4.1.1 Construction of the System of Equations

To obtain more equations, note, that the identity relation in Eq. (4.1) remains valid if we apply a function $\omega : \mathbb{R}^2 \rightarrow \mathbb{R}^n$, ($n \geq 1$) on both sides of the equation [30, 76, 75]:

$$\omega(\mathbf{y}) = \omega(\varphi(\mathbf{x})) \Leftrightarrow \omega(\mathbf{x}) = \omega(\varphi^{-1}(\mathbf{y})). \quad (4.11)$$

In a similar way as in the case of Eq. (4.8), we can integrate on both sides and obtain the following equation:

$$\int_{\mathcal{F}_o} \omega(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{F}_t} \omega(\varphi(\mathbf{x})) |J_\varphi(\mathbf{x})| d\mathbf{x}. \quad (4.12)$$

The basic idea of the proposed method is to generate a sufficient number of equations using a set of linearly independent ω functions. Since an equation set up using $\omega : \mathbb{R}^2 \rightarrow \mathbb{R}^n$ ($n > 1$) can be decomposed into n equations (similarly to Eq. (4.10)), we can assume that $n = 1$ without any restriction.

Suppose that φ has k parameters and let $\omega_i : \mathbb{R}^2 \rightarrow \mathbb{R}$, ($i = 1, \dots, \ell$) the set of applied functions. To solve for all unknowns, we need at least k equations, hence $\ell \geq k$. Thus we can obtain the following system of equations

$$\int_{\mathcal{F}_o} \omega_i(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{F}_t} \omega_i(\varphi(\mathbf{x})) |J_\varphi(\mathbf{x})| d\mathbf{x}, \quad i = 1, \dots, \ell, \quad (4.13)$$

where each ω_i function provides one new equation. Each applied ω_i function can be considered as a coloring of the shape (see Fig. 4.2) and the integrals in Eq. (4.13) give the volumes of the ω_i functions over the area of the shape. The equations thus match these volumes.

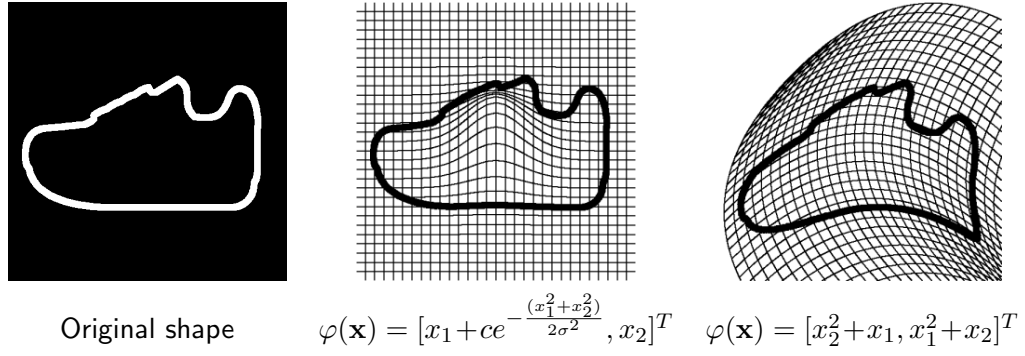


Figure 4.3: Example deformation fields

Each equation provides additional constraints. The solution of this system gives the estimation of the parameters of the aligning transformation. Unfortunately the algebraic independence of the equations can not be guaranteed. It can be easily seen, however, that the linear independency of the ω_i functions is an important criterion, which can be verified much easier. Therefore, in practice, we propose to apply a set of linearly independent nonlinear ω_i functions.

4.2 The Studied Deformation Models

In this section we introduce the transformation models that has been studied in our experiments. We note that, while on grayscale (or color) images most of the transformation models cause visible deformation, in the case of binary images, the effect of some transformations is not always observable (*i.e.*, on binary images the transformation is not always bijective, see Fig. 4.3). Herein we focus on some such a widely used transformation classes, but the framework can be applied to any diffeomorphic transformation. All we need to adopt a transformation model into our framework are the formular description of the transformation φ and its Jacobian $|J_\varphi(\mathbf{x})|$.

4.2.1 Planar Homography

Planar homography is the projective transformation between the images of the same planar object. It plays an important role in computer vision. The most simple method to estimate its parameters traces back the problem to the solution of a system of linear equations. Its coefficients are determined by the coordinates of point correspondences. The solution requires at least 4 point correspondences, but an overdetermined system solved in the least-squares sense gives more accurate result in spite of the inaccurate coordinates. A survey of planar homography estimation techniques can be found in [2].



Figure 4.4: Example perspective distortion

In many cases, when the perspective distortion is weak, affine transformation model is used to approximate the planar homography or to provide an initial estimation [54].

Planar homography is a projective transformation (see Fig. 4.5 and Fig. 4.4) which can be given as a linear transformation on the projective plane \mathbb{P}^2 using the

$$\mathbf{H} = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & 1 \end{pmatrix} \quad (4.14)$$

3×3 non-singular ($|\mathbf{H}| \neq 0$) matrix. Since such a projective transformation has 8 degree of freedom, we can assume without any restriction that $H_{33} = 1$. Then the transformation between the corresponding points $\mathbf{x}' = [x'_1, x'_2, x'_3]^T \in \mathbb{P}^2$ and $\mathbf{y}' = [y'_1, y'_2, y'_3]^T \in \mathbb{P}^2$ of the *template* and the *observation* shapes is given by

$$\mathbf{y}' = \mathbf{H}\mathbf{x}' \Leftrightarrow \mathbf{x}' = \mathbf{H}^{-1}\mathbf{y}'. \quad (4.15)$$

The Euclidean coordinates $\mathbf{y} = [y_1, y_2]^T \in \mathbb{R}^2$ of the homogeneous point $\mathbf{y}' = [y'_1, y'_2, y'_3]^T \in \mathbb{P}^2$ are obtained by applying the perspective division:

$$\begin{aligned} y_1 &= \frac{y'_1}{y'_3} = \frac{H_{11}x_1 + H_{12}x_2 + H_{13}}{H_{31}x_1 + H_{32}x_2 + 1} \equiv h_1(\mathbf{x}) \\ y_2 &= \frac{y'_2}{y'_3} = \frac{H_{21}x_1 + H_{22}x_2 + H_{23}}{H_{31}x_1 + H_{32}x_2 + 1} \equiv h_2(\mathbf{x}), \end{aligned} \quad (4.16)$$

where $h_i : \mathbb{R}^2 \rightarrow \mathbb{R}$. Thus, while planar homography is a linear transformation in the projective plane \mathbb{P}^2 , it becomes nonlinear in the Euclidian plane \mathbb{R}^2 . Let the nonlinear transformation corresponding to \mathbf{H} denoted by $h : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $h(\mathbf{x}) = (h_1(\mathbf{x}), h_2(\mathbf{x}))^T$. Thus Eq. (4.1) can be written as

$$\mathbf{y} = \varphi(\mathbf{x}) \equiv h(\mathbf{x}) \Leftrightarrow \mathbf{x} = h^{-1}(\mathbf{y}), \quad (4.17)$$

where $h^{-1} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ a \mathbf{H}^{-1} is the inverse transformation corresponding to the inverse matrix. The Jacobian $|J_h| : \mathbb{R}^2 \rightarrow \mathbb{R}$ is given by

$$|J_h(\mathbf{x})| = \begin{vmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} \end{vmatrix} = \frac{|\mathbf{H}|}{(H_{31}x_1 + H_{32}x_2 + 1)^3}. \quad (4.18)$$

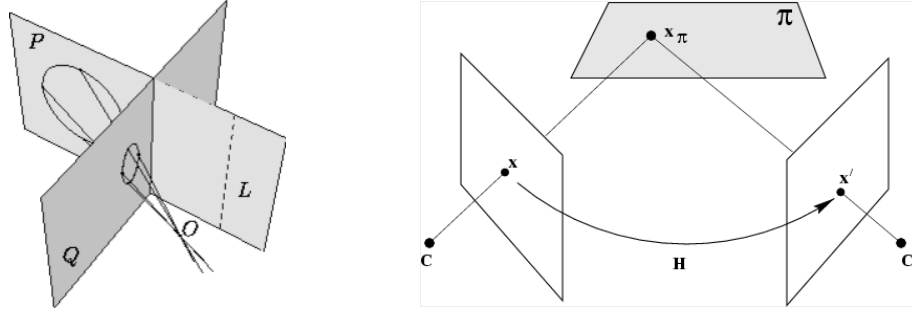


Figure 4.5: Projective transformations. **Left:** the images of the circular shape on the planes P and Q are related by a center-point projection. **Right:** The planar homography as the relation between two images of the same plane π taken with different cameras. (Figure from: [57], [83])

4.2.2 The Taylor Series Expansion of Planar Homography

In the case of planar homography, the integrands in Eq. (4.13) can be approximated by a Taylor series expansion and this way we can obtain polynomial equations. For example, consider the term $\omega_i(h(\mathbf{x})) |J_h(\mathbf{x})|$. If ω_i is polynomial, then both the numerator and the denominator of $\omega_i(h(\mathbf{x}))$ will be also polynomial and remains a polynomial when multiplied by $|J_h(\mathbf{x})|$. Thus

$$\omega_i(h(\mathbf{x})) |J_h(\mathbf{x})| = p(H_1, \dots, H_6) \chi(H_7, H_8), \quad (4.19)$$

where p is a polynomial, while χ is the remaining non-polynomial term. For example assuming the ω_i set from Eq. (4.41), we get

$$\chi(H_7, H_8) = \frac{1}{(H_7 x_1 + H_8 x_2 + 1)^{3+n+m}}. \quad (4.20)$$

where n and m are the exponents in ω_i . Since p is polynomial, in order to obtain a polynomial approximation of the whole term $\omega_i(h(\mathbf{x})) |J_h(\mathbf{x})|$, we only need to rewrite $\chi(H_7, H_8)$ in its n^{th} order multivariate Taylor series form. Let us denote the n^{th} order Taylor series expansion of the function $f : \mathbb{R}^k \rightarrow \mathbb{R}$ about the point (a_1, \dots, a_k) by $T_{f(a_1, \dots, a_k)}^n$. Then $\chi(H_7, H_8)$ can be approximated as follows:

$$\chi(H_7, H_8) \approx T_{\chi(0,0)}^n, \quad (4.21)$$

and thus the whole integrand can be rewritten as:

$$\omega_i(h(\mathbf{x})) |J_h(\mathbf{x})| \approx p(H_1, \dots, H_6) T_{\chi(0,0)}^n. \quad (4.22)$$

Substituting each integrand in Eq. (4.37) with the Taylor series expansion approximation, we obtain a system of polynomial equations. This increases the computational efficiency, as discussed in Section 4.2.3. In our experiments we obtained good results with $n = 5$.



Figure 4.6: Example polynomial transformation.

4.2.3 Polynomial Transformations

The reason of the importance of the polynomial transformations (see Fig. 4.3 and Fig. 4.6) is that they can be used to approximate other transformation models and general deformations. Let $p : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ a polynomial transformation and suppose that it is invertible over the foreground region \mathcal{F}_t of the template. Thus Eq. (4.1) can be written as:

$$\mathbf{y} = \varphi(\mathbf{x}) \equiv p(\mathbf{x}) \quad \Leftrightarrow \quad \mathbf{x} = p^{-1}(\mathbf{y}), \quad (4.23)$$

where $p(\mathbf{x}) = [p_1(\mathbf{x}), p_2(\mathbf{x})]^T$. Let $d = \max \{\deg(p_1), \deg(p_2)\}$, where $\deg(p)$ is the degree of p . Then

$$p_1(\mathbf{x}) = \sum_{i=0}^d \sum_{j=0}^{d-i} a_{ij} x_1^i x_2^j, \quad \text{és} \quad p_2(\mathbf{x}) = \sum_{i=0}^d \sum_{j=0}^{d-i} b_{ij} x_1^i x_2^j,$$

where a_{ij} and b_{ij} are the unknown parameters of the transformation and the number of these parameters is $(d+2)(d+1)$. The Jacobian $|J_p| : \mathbb{R}^2 \rightarrow \mathbb{R}$ is also polynomial

$$\begin{aligned} |J_p(\mathbf{x})| = & \left(\sum_{i=1}^d \sum_{j=0}^{d-i} i a_{ij} x_1^{i-1} x_2^j \right) \left(\sum_{j=1}^d \sum_{i=0}^{d-j} j b_{ij} x_1^i x_2^{j-1} \right) \\ & - \left(\sum_{j=1}^d \sum_{i=0}^{d-j} j a_{ij} x_1^i x_2^{j-1} \right) \left(\sum_{i=1}^d \sum_{j=0}^{d-i} i b_{ij} x_1^{i-1} x_2^j \right). \end{aligned}$$

In the followings we show, that in the case of the polynomial transformations, if the applied ω_i functions are polynomial, then the integrals in the system of equations can be decomposed. For that purpose, consider the system of equations in Eq. (4.13). Using the following set of polynomial ω_i functions:

$$\omega_i(\mathbf{x}) = x_1^{n_i} x_2^{m_i}, \quad n_i, m_i \in \mathbb{N}_0, \quad i = 1, \dots, \ell, \quad (4.24)$$

the system of equations in Eq. (4.13) becomes:

$$\int_{\mathcal{F}_o} y_1^{n_i} y_2^{m_i} d\mathbf{y} = \int_{\mathcal{F}_t} p_1(\mathbf{x})^{n_i} p_2(\mathbf{x})^{m_i} |J_p(\mathbf{x})| d\mathbf{x}, \quad i = 1, \dots, \ell. \quad (4.25)$$



Figure 4.7: Example thin plate spline transformation.

The left hand side of the equations are independent of the parameters of the transformations and can be computed using only the pixel coordinates of the *observation* shape. Furthermore since $p_1(\mathbf{x})$, $p_2(\mathbf{x})$ and $|J_p(\mathbf{x})|$ are polynomials, thus the right hand sides are also polynomial:

$$\int_{\mathcal{F}_t} p_1(\mathbf{x})^{n_i} p_2(\mathbf{x})^{m_i} |J_p(\mathbf{x})| d\mathbf{x} = \int_{\mathcal{F}_t} \sum_{q=0}^{c_i} \sum_{r=0}^{c_i-q} t_{igr} x_1^q x_2^r d\mathbf{x} = \sum_{q=0}^{c_i} \sum_{r=0}^{c_i-q} t_{igr} \int_{\mathcal{F}_t} x_1^q x_2^r d\mathbf{x}, \quad (4.26)$$

where $c_i = d(n_i + m_i) + d(d-1)$ is the order of the polynomial in the i th equation, while t_{igr} is a polynomial expression that consists of the parameters of p . Therefore Eq. (4.26) is a polynomial system of equations, in which the integrals $\int_{\mathcal{F}_t} x_1^q x_2^r d\mathbf{y}$ are independent of the transformation parameters, they have to be computed only once, in the beginning of the registration process, which highly increases the computational efficiency of the method.

4.2.4 Thin Plate Spline

The *Thin Plate Spline* (TPS) model is widely used to approximate general non-rigid deformations [38]. Let a set of control points $\mathbf{c}_k \in \mathbb{R}^2$ and corresponding coefficients (weights) $a_{ij}, w_{ki} \in \mathbb{R}$, $i = 1, 2$, $j = 1, \dots, 3$, $k = 1, \dots, K$. Then

$$s_i(\mathbf{x}) = a_{i1}x_1 + a_{i2}x_2 + a_{i3} + \sum_{k=1}^K w_{ki} Q(\|\mathbf{c}_k - \mathbf{x}\|), \quad i = 1, 2, \quad (4.27)$$

where $Q : \mathbb{R} \rightarrow \mathbb{R}$ is the so called *radial basis function*:

$$Q(r) = r^2 \log r^2. \quad (4.28)$$

The physical analogy of this model is a bended thin metal plate. Additional constraints are proposed to avoid unwanted global effects, such as the movement and the rotation of the plate under the imposition of the loads [118]:

$$\sum_{k=1}^K w_{ik} = 0, \quad \sum_{k=1}^K c_{ki} w_{ki} = 0, \quad i = 1, 2. \quad (4.29)$$

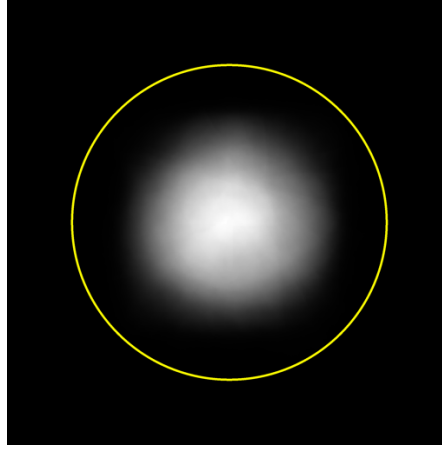


Figure 4.8: Coverage of transformed shapes of ≈ 1500 synthetic observations during the minimization process. Pixel values represent the number of intermediate shapes that included a particular pixel. For reference, we also show the circle with radius $\frac{\sqrt{2}}{2}$ used for normalization.

Now $s(\mathbf{x}) = [s_1(\mathbf{x}), s_2(\mathbf{x})]^T$,

$$\mathbf{y} = \varphi(\mathbf{x}) \equiv s(\mathbf{x}) \quad \Leftrightarrow \quad \mathbf{x} = s^{-1}(\mathbf{y}) \quad (4.30)$$

defines a transformation with $6 + 2K$ parameters (6 affine parameter a_{ij} and $2K$ coefficients w_{ki} for the control points).

The Jacobian $|J_s| : \mathbb{R}^2 \rightarrow \mathbb{R}$ of the transformation is given by:

$$|J_s(\mathbf{x})| = \begin{vmatrix} \frac{\partial s_1}{\partial x_1} & \frac{\partial s_1}{\partial x_2} \\ \frac{\partial s_2}{\partial x_1} & \frac{\partial s_2}{\partial x_2} \end{vmatrix}, \quad (4.31)$$

where

$$\frac{\partial s_i}{\partial x_j} = a_{ij} + \sum_{k=1}^K w_{ik} \left[-2(c_{kj} - x_j)(1 + \log(\|\mathbf{c}_k - \mathbf{x}\|)) \right]. \quad (4.32)$$

4.3 Implementational Details

It is required to avoid extreme (large or small) numerical values of the integrals. It can be done by applying some normalization on both the pixel coordinates and the ω_i functions. For that purpose, we normalized the coordinates of both shapes into the square $[-0.5, 0.5] \times [-0.5, 0.5]$ and chose ω_i with a range limited to the interval (e.g., $[-1, 1]$).

Despite of these normalizations, the integrals in our equations could take values of different range of magnitude yielding a different contribution from each equation to

$$\begin{aligned}
h_1^{-1}(\mathbf{y}) &= x_1 = \frac{(H_{22} - H_{32}H_{23})y_1 - (H_{12} - H_{32}H_{13})y_2 + H_{23}H_{12} - H_{22}H_{13}}{(H_{32}H_{21} - H_{31}H_{22})y_1 - (H_{32}H_{11} - H_{31}H_{12})y_2 + H_{22}H_{11} - H_{21}H_{12}} \\
h_2^{-1}(\mathbf{y}) &= x_2 = \frac{-(H_{21} - H_{31}H_{23})y_1 + (H_{11} - H_{31}H_{13})y_2 - (H_{23}H_{11} - H_{21}H_{13})}{(H_{32}H_{21} - H_{31}H_{22})y_1 - (H_{32}H_{11} - H_{31}H_{12})y_2 + H_{22}H_{11} - H_{21}H_{12}} \\
|J_{h^{-1}}(\mathbf{y})| &= \frac{|\mathbf{H}|^2}{((H_{32}H_{21} - H_{31}H_{22})y_1 - (H_{32}H_{11} - H_{31}H_{12})y_2 + H_{22}H_{11} - H_{21}H_{12})^3}
\end{aligned} \tag{4.36}$$

the objective function value in the optimization process. To overcome this issue, we also normalized the equations by dividing the integrals with an appropriate constant. We found experimentally (see Fig. 4.8), that the transformations occurring during the least-squares minimization process do not transform the shapes out of a circle with center in the origin and a radius $\frac{\sqrt{2}}{2}$ (i.e., the circumscribed circle of the unit square). We thus adopt the following constant

$$N_i = \int_{\|\mathbf{x}\| \leq \frac{\sqrt{2}}{2}} |\omega_i(\mathbf{x})| d\mathbf{x}, \tag{4.33}$$

and the normalized version of Eq. (4.13) becomes

$$\frac{\int_{\mathcal{F}_o} \omega_i(\mathbf{y}) d\mathbf{y}}{N_i} = \frac{\int_{\mathcal{F}_t} \omega_i(\varphi(\mathbf{x})) |J_\varphi(\mathbf{x})| d\mathbf{x}}{N_i}, \quad i = 1, \dots, \ell. \tag{4.34}$$

4.3.1 Numerical Implementation

The system of equations in Eq. (4.13) has been constructed in the continuous space, in practice, however, we have limited resolution digital images. Therefore the integrals can only be approximated by finite sums over the foreground pixels. This approximation obviously introduces a negligible error into the equations. Let us denote by F_t and F_o the finite pixel sets corresponding to the continuous pixel regions \mathcal{F}_t and \mathcal{F}_o . Then Eq. (4.34) can be approximated by the following system of equations:

$$\frac{1}{N_i} \sum_{\mathbf{y} \in F_o} \omega_i(\mathbf{y}) = \frac{1}{N_i} \sum_{\mathbf{x} \in F_t} \omega_i(\varphi(\mathbf{x})) |J_\varphi(\mathbf{x})|, \quad i = 1, \dots, \ell. \tag{4.35}$$

This system of equations can be supplemented with further equations depending on the transformation model. These additional equations are either to improve numerical stability (like in the case of planar homography) or to enforce additional constraints required by the model (e.g., thin plate spline).

Planar homography: Since the inverse of a planar homography is also a planar homography, each equation of Eq. (4.35) can also be written in three alternative forms

by making use of the corresponding inverse transformation h^{-1} and the reverse integral transformation $\mathbf{x} = h^{-1}(\mathbf{y})$, $d\mathbf{x} = |J_{h^{-1}}(\mathbf{y})| d\mathbf{y}$:

$$\begin{aligned} \frac{1}{N_i} \sum_{\mathbf{x} \in F_i} \omega_i(\mathbf{x}) &= \frac{1}{N_i} \sum_{\mathbf{y} \in F_o} \omega_i(h^{-1}(\mathbf{y})) |J_{h^{-1}}(\mathbf{y})| \\ \frac{1}{N_i} \sum_{\mathbf{x} \in F_i} \omega_i(\mathbf{x}) |J_h(\mathbf{x})| &= \frac{1}{N_i} \sum_{\mathbf{y} \in F_o} \omega_i(h^{-1}(\mathbf{y})) \\ \frac{1}{N_i} \sum_{\mathbf{x} \in F_i} \omega_i(h(\mathbf{x})) &= \frac{1}{N_i} \sum_{\mathbf{y} \in F_o} \omega_i(\mathbf{y}) |J_{h^{-1}}(\mathbf{y})|. \end{aligned} \quad (4.37)$$

where h^{-1} and $|J_{h^{-1}}(\mathbf{y})|$ are defined in Eq. (4.36). While these equations are equivalent to the original one, including them provides additional constraints and thus helps the optimizer to find the optimal solution.

Thin plate spline: In the case of this transformation model, the number of required equations depends on the number of the control points. Furthermore, the constraints in Eq. (4.29) has to be included in the system of equations.

4.3.2 Solution and Complexity

The parameters of the aligning transformations are obtained as the solution of the system of equations. While the equations are nonlinear, we found that it can be solved efficiently by *Levenberg-Marquardt algorithm* [66] in least-squares sense. This method is commonly applied in function fitting, thus in the case of a function $f : \mathbb{R}^k \rightarrow \mathbb{R}^\ell$, it is suitable to solve:

$$\min_{\mathbf{x}} \sum_{i=1}^{\ell} (f_i(\mathbf{x}))^2, \quad \mathbf{x} \in \mathbb{R}^k. \quad (4.38)$$

The system of equations in Eq. (4.13) can be simply transformed into such a minimization problem, by substituting its two sides. Thus we obtain an objective function of class $\mathbb{R}^k \rightarrow \mathbb{R}^\ell$, where k is the number of the transformation parameters and ℓ is the number of equations (*i.e.*, the number of the applied ω_i functions). As an iterative optimizing method, the Levenberg-Marquardt method requires an initial solution. We found, that the parameter vector corresponding to the identity transformation is appropriate for that purpose. For example, in the case of planar homography it is the values of the $\mathbf{H}_0 = \mathbf{I}_{3 \times 3}$ identity matrix. On the other hand, in many cases one can find a better initial solution. For example the scaling between the shapes can be initially approximated as the ratio of the shape areas. For example, for planar homography

$$\mathbf{H}_0 = \begin{pmatrix} s_0 & 0 & 0 \\ 0 & s_0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{where } s_0 = \sqrt{\frac{M}{N}}, \quad (4.39)$$

Algorithm 5: Pseudo code of the nonlinear registration framework**Input** : Binary images of the *template* and *observation***Output**: k parameters of the estimated transformation $\hat{\phi}$

- 1 Normalize each input shape into the unit square $[-0.5, 0.5] \times [-0.5, 0.5]$ such that the shape's center of mass becomes the origin.
- 2 Choose a set of $\omega_i : \mathbb{R}^2 \rightarrow \mathbb{R}$ ($i = 1, \dots, \ell \geq k$) functions as described in Section 4.4.1.
- 3 Construct the system of equations of Eq. (4.35) and include any additional equations if needed (e.g., Eq. (4.37)).
- 4 Find a least-squares solution of the system using a *Levenberg-Marquardt* algorithm. The solver is initialized with the parameters of the identity transformation.
- 5 Unnormalizing the solution gives the parameters of the aligning transformation.

where N and M are the numbers of the foreground pixels on the *observation* and the *template* respectively. The proposed method has been implemented in Matlab and we set up the Levenberg-Marquard algorithm as the followings:

- maximum number of function evaluations: 1400
- maximum number of iterations: 1000
- minimal change in the function value: 10^{-12}
- minimum change in variables: 10^{-12}
- search type: quadratic polynomial

The pseudo code of the proposed method can be found in Algorithm 5. We use a fixed number of optimization steps, and the time consumption of each equation evaluation is depends only the number of the shape pixels, thus the time complexity of the algorithm $\mathcal{O}(N + M)$.

4.4 Experiments on Synthetic Images

The perfomance of the method has been examined on synthetic datasets as well as on real images. We also compared our algorithm to other nonlinear registration solutions. The first one was the *Shape Context* [16] method, which is developed for general non-linear alignment. It is based on a log-polar histogram based similarity measure (called shape context) of contour points of the shapes. The point correspondence problem (as

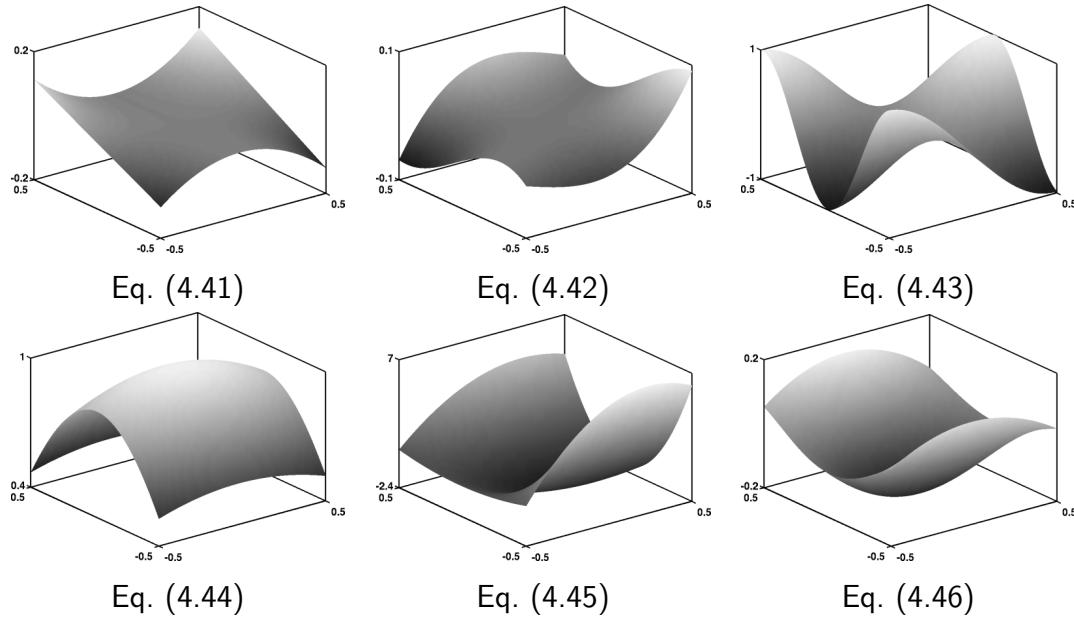


Figure 4.9: Plots of tested $\{\omega_i\}$ function sets.

an assignment problem) is then solved by the Hungarian method. The main advantage of this approach is that it does not require intensity information. Its demo implementation is available on the internet [89], we set the parameters empirically to their optimal value (`beta_init=30`, `n_iter=30`, `r=1`).

The registration results were evaluated using two different error measures. The δ measure gives the normalized symmetric difference of shapes, while ε measures the distance between the true φ and the estimated $\hat{\varphi}$ transformation:

$$\delta = \frac{|F_r \Delta F_o|}{|F_r| + |F_o|} \cdot 100\%, \quad \varepsilon = \frac{1}{|F_t|} \sum_{\mathbf{x} \in F_t} \|\varphi(\mathbf{x}) - \hat{\varphi}(\mathbf{x})\|, \quad (4.40)$$

where F_r , F_t and F_o are the sets of the object pixels of the registered, the template, and the observation shapes, while Δ denotes the symmetric difference. Obviously, ε can be used only in the case of synthetic tests, where the original parameters of the transformation used to generate the distorted shape are known. To examine the test results, the *median*, *mean*, and *standard deviation* of the error measures have been calculated.

4.4.1 Comparison of Different ω Functions

Theoretically any set of linearly independent ω_i functions could be used to set up the system of equations in Eq. (4.13). Obviously, ω_i must be an integrable function over the finite domains \mathcal{F}_o and \mathcal{F}_t , and the equations should be properly normalized. In

$\{\omega_i\}$ set	$\delta(\%)$			$\varepsilon(\text{pixel})$		
	m	μ	σ	m	μ	σ
1)	0.09	0.53	3.38	0.08	3.03	22.36
2)	0.11	1.01	5.01	0.10	4.40	24.14
3)	0.21	12.28	19.61	0.19	20.14	41.73
4)	0.12	1.52	6.25	0.11	6.02	25.79
5)	0.10	0.80	4.75	0.08	3.27	18.60
6)	0.10	0.99	4.84	0.08	4.17	20.78

Table 4.1: Comparison of the tested $\{\omega_i\}$ sets. The median (m), mean (μ), and standard deviation (σ) values are shown.

this section we examine the registration results using different $\{\omega_i\}$ function sets. We consider power, polynomial, and trigonometric function sets (see Fig. 4.9):

1. Power functions

$$\omega_i(\mathbf{x}) = x_1^{n_i} x_2^{m_i} \quad (4.41)$$

with $(n_i, m_i) \in \{(0, 0), (1, 0), (0, 1), (1, 1), (2, 0), (0, 2), (2, 1), (1, 2), (2, 2), (3, 0), (0, 3), (3, 1), (1, 3)\}$

2. Rotated power functions

$$\omega_i(\mathbf{x}) = (x_1 \cos \alpha_i - x_2 \sin \alpha_i)^{n_i} (x_1 \sin \alpha_i + x_2 \cos \alpha_i)^{m_i} \quad (4.42)$$

with $\alpha_i \in \{0, \frac{\pi}{6}, \frac{\pi}{3}\}$ and $(n_i, m_i) \in \{(1, 2), (2, 1), (1, 3), (3, 1)\}$

3. Mixed trigonometric functions

$$\omega_i(\mathbf{x}) = \sin(n_i x_1 \pi) \cos(m_i x_2 \pi) \quad (4.43)$$

with $(n_i, m_i) \in \{(1, 2), (2, 1), (2, 2), (1, 3), (3, 1), (2, 3), (3, 2), (3, 3), (1, 4), (4, 1), (2, 4), (4, 2)\}$

4. Trigonometric functions

$$\omega_i(\mathbf{x}) = Q_i(n_i x_1) R_i(m_i x_2) \quad (4.44)$$

with $Q_i(x), R_i(x) \in \{\sin(x), \cos(x)\}$ and $(n_i, m_i) \in \{(1, 1), (1, 2), (2, 1)\}$

5. Polynomials

$$\omega_i(\mathbf{x}) = P_{n_i}(x_1) P_{m_i}(x_2) \quad (4.45)$$

with $(n_i, m_i) \in \{(1, 2), (2, 1), (1, 3), (3, 1), (2, 3), (3, 2), (1, 4), (4, 1), (2, 4), (4, 2), (3, 4), (4, 3)\}$ composed of the following random polynomials:

$$\begin{aligned} P_1(x) &= 2x^2 - x - 1 \\ P_2(x) &= 2x^3 - x^2 \\ P_3(x) &= x^3 - 30x^2 + 3x + 2 \\ P_4(x) &= 3x^5 - x^2 + 5x - 1 \end{aligned}$$

6. Polynomials

$$\omega_i(\mathbf{x}) = L_{n_i}(x_1)L_{m_i}(x_2) \quad (4.46)$$

with $(n_i, m_i) \in \{(2, 3), (3, 2), (2, 4), (4, 2), (3, 4), (4, 3), (2, 5), (5, 2), (3, 5), (5, 3), (4, 5), (5, 4)\}$ composed of the following Legendre polynomials:

$$\begin{aligned} L_2(x) &= \frac{1}{2}(3x^2 - 1) \\ L_3(x) &= \frac{1}{2}(5x^3 - 3x) \\ L_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3) \\ L_5(x) &= \frac{1}{8}(63x^5 - 70x^3 + 15x) \end{aligned}$$

The results of these function sets in the case of planar homography can be found in Table 4.1. Basically, all median δ error measures are between 0.1 – 0.2. Although the mean values have a slightly bigger variance, this is mainly caused by a few outliers rather than a systematic error. It is thus fair to say that the considered ω_i functions perform equally well, which confirms that the results are independent of the choice of these functions. The computational efficiency of the method, however, can be improved by choosing polynomial ω_i functions (see Section 4.2.3 and Section 4.2.2). Therefore we recommend to use low order polynomials for computational efficiency. In our experiments, we have used the set 1), unless otherwise noted.

4.4.2 Comparative Tests

Planar homography: In the case of planar homography the synthetic dataset contained 37 different template shapes of size 256×256 and their 1480 transformed versions used as observations. To generate these observations, the applied plane projective transformations were randomly composed of 0.5, ..., 1.5 scalings; $-\frac{\pi}{4}, \dots, \frac{\pi}{4}$ rotations along the three axes; $-1, \dots, 1$ translations along both x and y axis and 0.5, ..., 2.5 along the z axis; and a random focal length chosen from $[0.5, 1.5]$. Note that these are projective transformations mapping a *template* shape from a plane placed

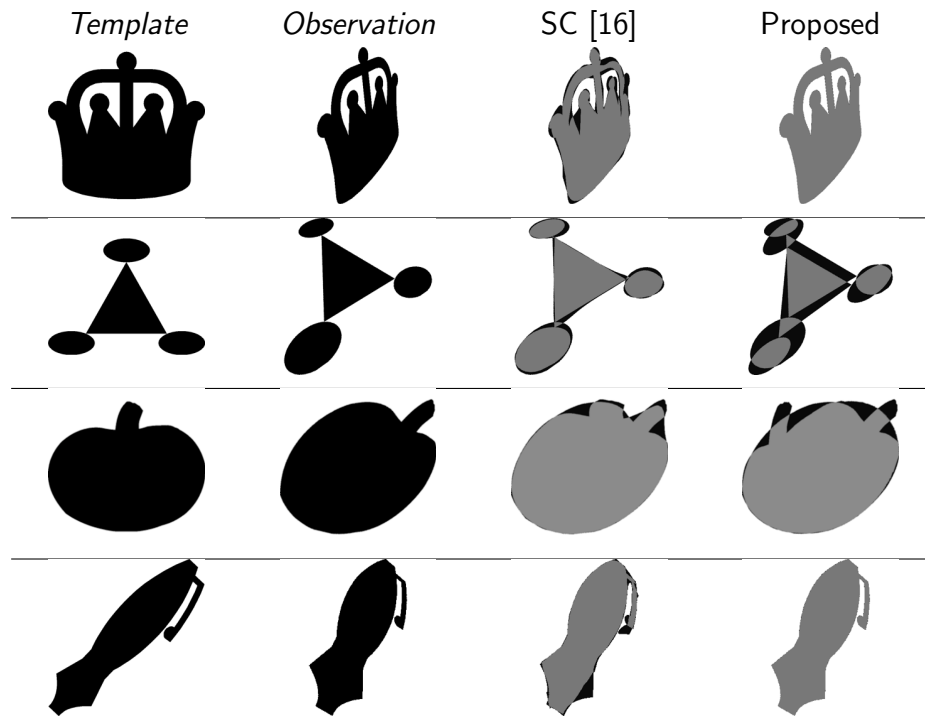


Figure 4.10: *Planar homographies*: Example images from the synthetic data set and registration results obtained by *Shape Context* [16] and the proposed method. The *observation* and the registered *template* were overlaid, overlapping pixels are depicted in gray whereas non-overlapping ones are shown in black.

	Runtime (sec.)			δ (%)			ε (pixel)	
	SC	P	T	SC	P	T	P	T
m	98.72	16.04	5.67	2.69	0.09	0.16	0.08	0.14
μ	102.78	27.04	6.52	4.41	0.54	0.88	2.97	3.79
σ	28.26	45.34	3.62	4.79	3.42	3.34	22.04	20.26

Table 4.2: Planar homography: Comparative tests of the proposed method on the synthetic dataset. SC – *Shape Context* [16]; P – proposed method using Eq. (4.35)–(4.36); and T – using its *Taylor expanded form*. m , μ , and σ denote the median, mean, and deviation.

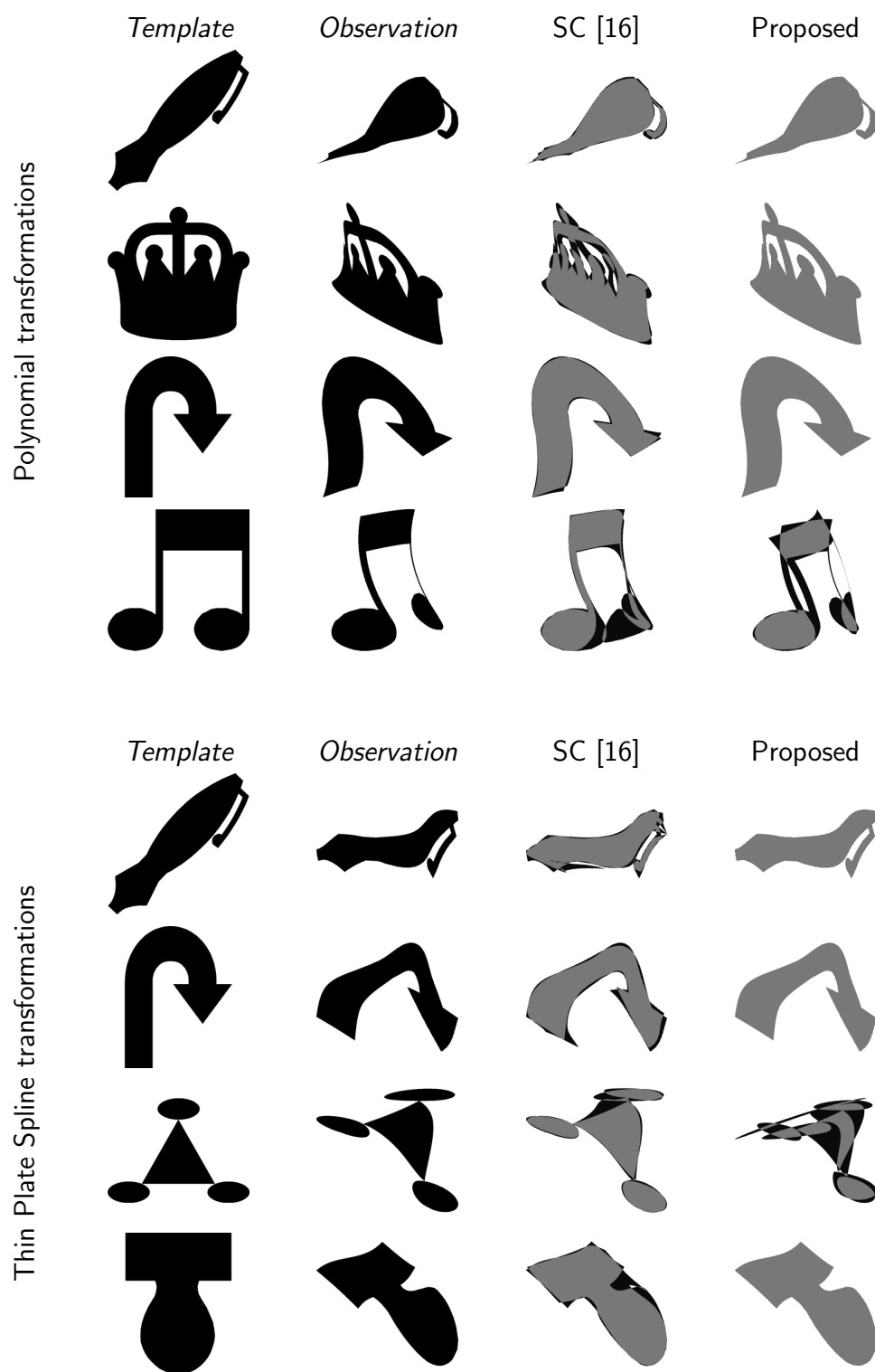


Figure 4.11: *Polynomial transformations* and *Thin plate spline transformations*: Example images from the synthetic data set and registration results obtained by *Shape Context* [16] and the proposed method. The *observation* and the registered *template* were overlaid, overlapping pixels are depicted in gray whereas non-overlapping ones are shown in black.

Polynomial transformations						
	Runtime (s.)			δ (%)		
	m	μ	σ	m	μ	σ
Shape Context [90]	100.91	105.54	30.14	2.50	2.99	1.81
Proposed method	14.72	15.42	6.95	0.33	2.04	6.12

Thin Plate Spline transformations						
	Runtime (s.)			δ (%)		
	m	μ	σ	m	μ	σ
Shape Context [90]	98.87	106.80	74.31	4.86	6.72	5.77
Proposed method	58.95	75.59	75.65	0.35	7.35	16.58

Table 4.3: *Polynomial transformations* and *Thin plate spline transformations*: The results of the proposed method and comparison to the *Shape Context* [90] method. (m: median, μ : mean, σ standard deviation)

in the 3D Euclidean space to the xy plane. Some typical examples of these images can be seen in Fig. 4.10, while a summary of registration results is presented in Table 4.2.

We compared the results of our method to the Shape Context method [16]. It can be seen, that the proposed method provided better δ measures (the only output of the demo implementation of Shape Context is the registered shape, hence ε can not be computed). The results of the proposed method using the Taylor expanded form of the planar homography transformation is also shown in Table 4.2. As it can be seen, the running times are considerably reduced at the price of a negligible loss in quality. In these tests, we have used the 5th order approximation of the integrands.

Polynomial transformations: The proposed method has been tested in the case of second order polynomial transformations. The synthetic dataset consisted of 37 template shapes and ≈ 1900 observations. The parameters of the transformations were uniformly randomly chosen: the $a_{20}, a_{02}, a_{11}, b_{20}, b_{02}, b_{11}$ parameters from $[-0.5; 0.5]$, the a_{10}, b_{01} parameters from $[0.5; 1.5]$, while the a_{01}, b_{10} parameters from the $[-0.25; 0.25]$ interval. Example images and registration results are shown in Fig. 4.11, while the summary of the results can be found in Table 4.3. It can be seen, that the proposed method provided more accurate results while required less runtime.

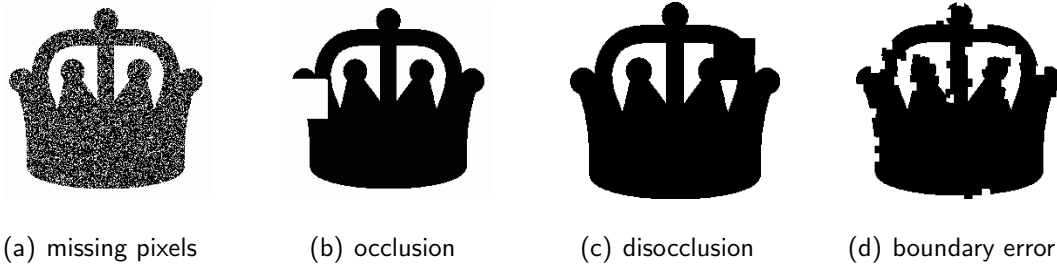


Figure 4.12: Sample observations with various degradations for testing robustness.

Thin plate splines: For the synthetic test we used the thin plate spline transformation model with 5 control points placed in the following points:

$$(0, 0), (-0.25, -0.25), (-0.25, 0.25), (0.25, -0.25), (0.25, 0.25)$$

. The synthetic dataset was consisted of the 37 template shapes and their ≈ 1000 observations. The affine parts (the a_{ij} parameters) of the transformations were randomly composed of $0.5, \dots, 1.5$ scalings, $-\frac{\pi}{4}, \dots, \frac{\pi}{4}$ rotations, and $-0.3, \dots, 0.3$ shearing, while the weights of the control points were chosen from $-0.58, \dots, 0.58$. Example images and registration results are shown in Fig. 4.11, while the summary of the results can be found in Table 4.3. It should be noted, that while the Shape Context placed the control points on the boundaries of the objects, the proposed method used the same 5 control point model with which the observation shapes were generated, which is a non-negligible advantage.

Robustness: We examined the robustness of the proposed method against segmentation errors on the synthetic dataset of planar homography. In the first testcase, $5\%, \dots, 20\%$ of the foreground pixels has been removed from the *observations* before registration. In the second case, we occluded continuous square-shaped regions of size equal to $1\%, \dots, 10\%$ of the shape, while in the third case we disoccluded a similar region. Finally, we randomly added or removed squares uniformly around the boundary of a total size $1\%, \dots, 10\%$ of the shape. Sample images can be found in Fig. 4.12.

The summary of the results can be found in Table 4.4. Results show that the proposed method is quite robust in those cases where the segmentation errors are uniformly distributed on the shapes (first and fourth testcases), and outperformed the Shape Context method [16] in these cases. In those cases, however, when the segmentation errors are consisted of larger continuous regions, like occlusion and disocclusion, it provides less accurate results compared to the Shape Context. The reason is that these large regions drastically change the value of the integrals in the system of equations.

(a) missing pixels		5%	10%	15%	20%
Shape Context [16]	m	21.85	24.91	26.38	27.2
	σ	5.97	6.14	6.37	6.56
Proposed method	m	2.98	5.69	8.51	11.57
	σ	4.13	5.23	6.09	6.74
(b) size of occlusion		1%	2.5%	5%	10%
Shape Context [16]	m	3.03	3.55	4.55	6.79
	σ	4.79	4.79	5.09	7.03
Proposed method	m	1.41	3.40	6.19	11.27
	σ	3.49	4.18	5.09	6.6
(c) size of disocclusion		1%	2.5%	5%	10%
Shape Context [16]	m	3.63	4.52	6.25	9.28
	σ	5.19	5.61	6.84	7.78
Proposed method	m	1.93	4.54	8.28	13.62
	σ	4.31	5.13	6.16	7.09
(d) size of boundary error		1%	5%	10%	20%
Shape Context [16]	m	2.86	3.78	4.68	6.92
	σ	4.72	4.83	5.04	5.92
Proposed method	m	0.54	1.67	2.67	4.03
	σ	3.28	3.5	3.9	4.47

Table 4.4: Median (m) and standard deviation (σ) of δ error (%) vs. various type of segmentation errors as shown in Fig. 4.12.

4.5 Experiments on Real Images

In this section we present the results of our algorithm in the case of various real problems.

4.5.1 Planar Homography

Traffic signs: We tested the proposed method to register traffic signs. Nowadays, many safety systems like radars and cameras are included in modern cars, thus one possible application of the proposed method could be the pose-estimation of the signs. In the case of a calibrated camera, the recovery of the exact spacial position of the sign could be also possible, thus the system could alert the driver. While in the literature there are automatic solutions for traffic sign detection and segmentation [82], we used classical thresholding and some morphological operations to extract the shapes

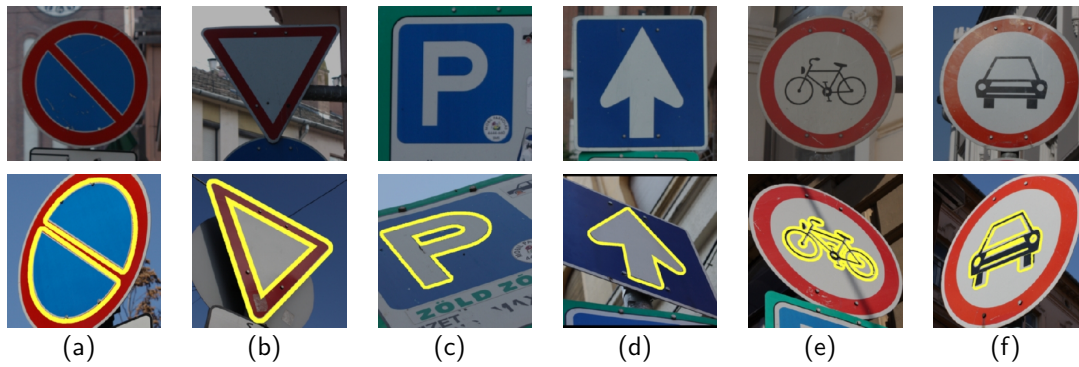


Figure 4.13: Registration results on traffic signs. The first row shows the *templates* while below them the corresponding *observations* with the overlaid contour of the registration results.

from the images. Each template and observation image pair was taken from different signboards. The main challenges were strong deformations, segmentation errors. In addition, the style of the same object is also often vary between different signboards. For example, the *observations* in Fig. 4.13(f), Fig. 4.15(b) and Fig. 4.15(c) do not contain exactly the same shape as the object on the *template*. In particular, the *STOP* sign in Fig. 4.15(c) uses different fonts. In spite of these difficulties, our method was able to recover a quite accurate transformation. The average δ error was 12,66% on these images.

Medical X-ray images: Hip replacement is an important especially for elderly patients whose hip joint has been broken. In this surgical procedure the hip joint is replaced by a prosthetic implant. After the operation inflammation often occurs, causing further loosening of fracture. In current practice, investigation of these loosening requires the inspection and comparison of X-ray images taken at different times after the operation. Automatic registration can help this visual inspection as clinically significant prosthesis movement can be very small [33].

The registration is challenged by highly non-linear radiometric distortion which makes the application of intensity-based methods difficult. On the other hand, the segmentation of the implants is possible [79], we used active contours [18]. Another problem is that the implant is positioned differently in 3D between the recording of the X-ray images, leading to nonlinear transformation between the planar images. Fortunately, these X-ray images are always taken in a well defined position of the patient's leg, the planar homography transformation can be assumed to be a good approximation here. Some registration results can be found in Fig. 4.14.

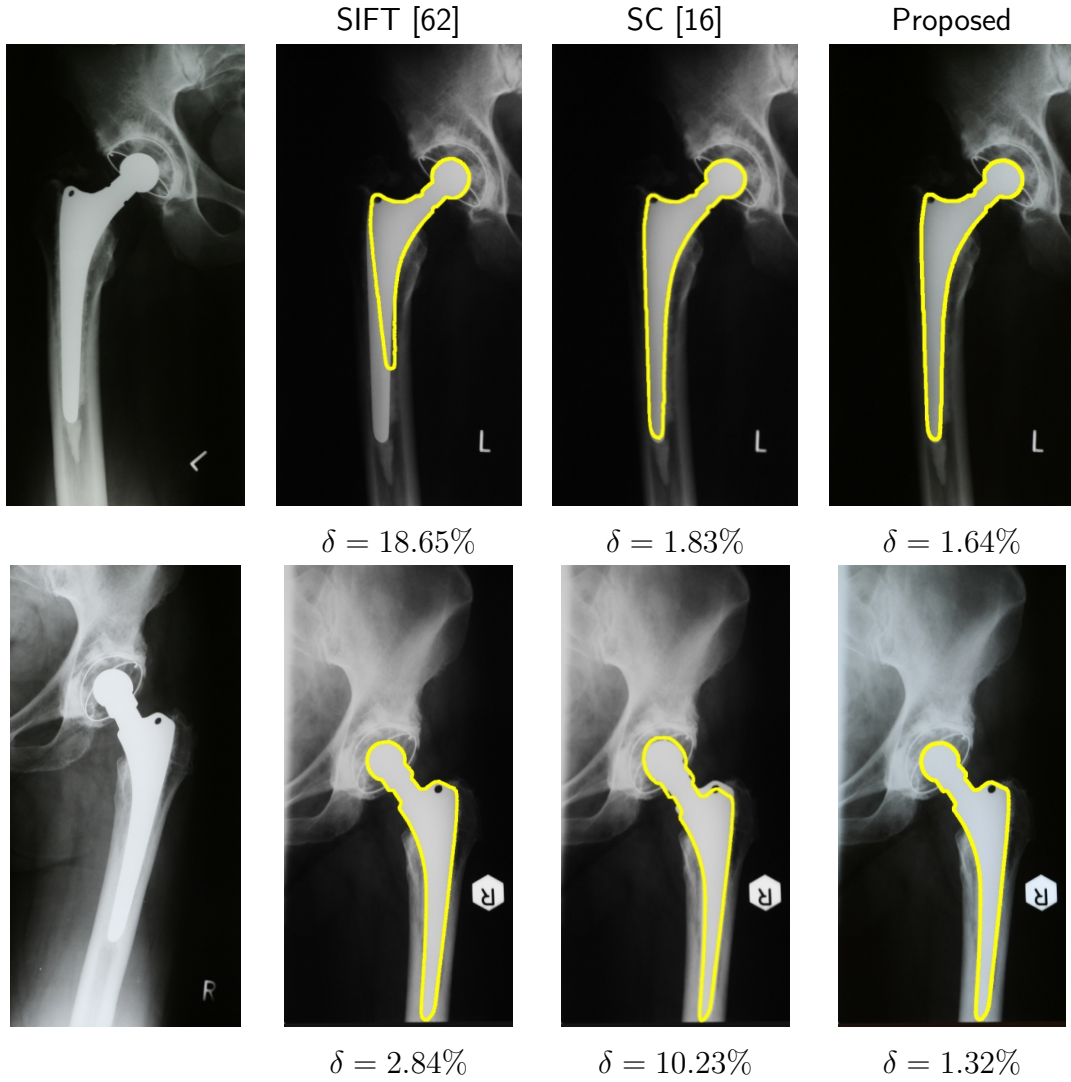


Figure 4.14: Registration results on hip prosthesis X-ray images. The overlaid contours show the aligned contours of the corresponding images on the left. Images in the second column show the registration results obtained by *SIFT* [62]+*homest* [60], in the third column the results of *Shape Context* [16]+*homest* [60], while the last column contains the results of the proposed method.

Comparison: We compared the results of our algorithm to the *homest* [60] software which is also available online. Using different methods [117, 45], it estimates a homography which minimizes a cost function obtained by point correspondences, and filters out outliers using regression [86]. For point matching, the authors propose to use the SIFT [61] method, which implements a scale and rotation invariant similarity measure [62]. This method requires the grayscale version of the images as input, thus we provided the masked signboard region for traffic sign matching and the prosthesis region for medical registration. Furthermore, we also tested *homest* using the point



Figure 4.15: Registration results on traffic signs. The *templates* are in the first row, then the results obtained by *SIFT* [62]+*homest* [60] (second row), where the images show point correspondences between the images found by *SIFT* [62] in the third row. The results obtained by *Shape Context* [16]+*homest* [60] (fourth row) and the proposed method in the last row. The contours of the registered images are overlaid.

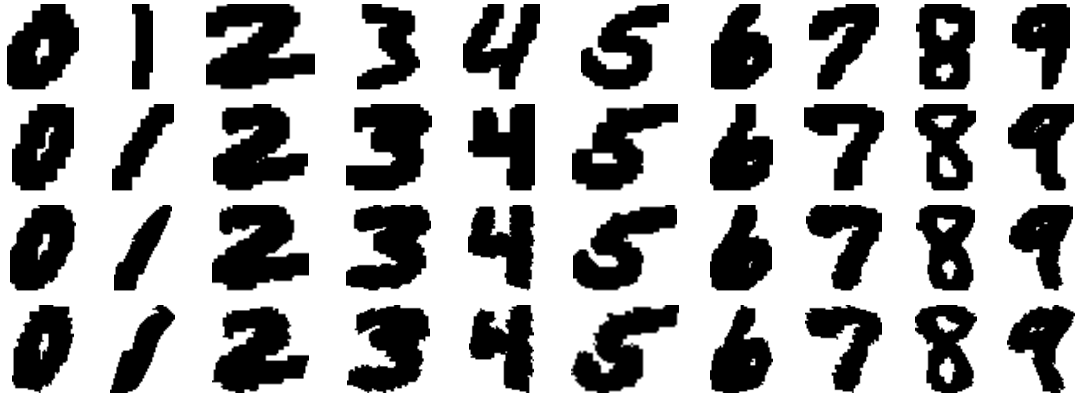


Figure 4.16: Sample images from the MNIST dataset and registration results using a thin plate spline model. First and second rows show the images used as *templates* and *observations* while the 3rd and 4th rows show the registration results obtained by Shape Context [16] and the proposed method, respectively.

	Runtime (sec.)			δ (%)		
	m	μ	σ	m	μ	σ
Shape Context [16]	35.02	34.43	7.58	7.86	9.40	4.71
Proposed method	10.00	9.81	1.47	7.66	8.93	4.22

Table 4.5: Comparative results on 2000 image pairs from the MNIST database. m , μ , and σ stand for the median, mean, and standard deviation.

correspondences obtained by *Shape Context*, which required the binary shapes as input.

Although the *SIFT* parameter called `distRatio`, controlling the number of the extracted correspondences, has been manually fine-tuned, we could not get reliable results due to the lack of rich radiometric features. Fig. 4.14 shows two results on X-ray images while on traffic signs (see Fig. 4.15), *SIFT* could not find enough correspondences in about half of the cases. As for *Shape Context*-based correspondences, we got somewhat better alignments (an average δ of 33.47% for the traffic sign images).

4.5.2 Thin Plate Spline

Character matching: The proposed method has been applied for aligning handwritten digits from the MNIST dataset. Aligning an observation to different templates can be used to classify the shape; the lowest deformation gives the right class of the questioned object [16]. Here we used 25 control points for the TPS model, these control points were placed on the 5×5 regular grid over the input images, which involves $2 \cdot 25 + 6 = 56$ model parameters. The equations were generated using the

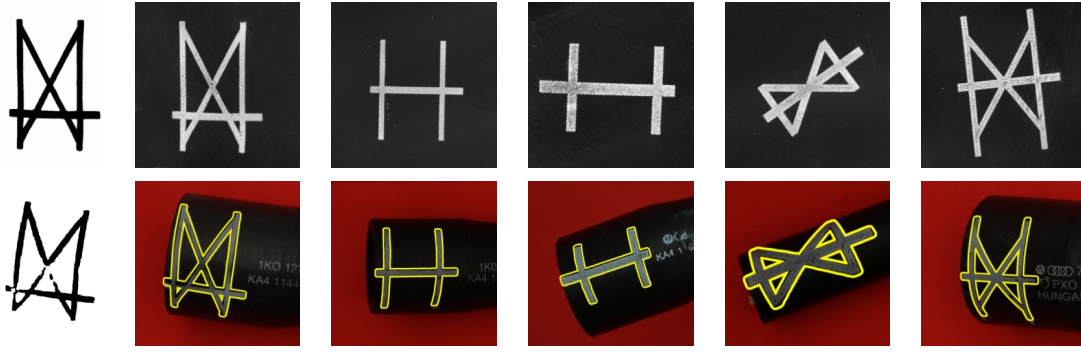


Figure 4.17: Registration results of printed signs. **Top:** planar *templates*. **Bottom:** the corresponding *observations* with the overlaid contour of the registration results. The first image pair shows the segmented regions used for registration. Note the typical segmentation errors. (Images provided by *ContiTech Fluid Automotive Hungária Ltd.*)

function set Eq. (4.41) with parameters $0 \leq m_i, n_i \leq 8$, $m_i + n_i \leq 8$ resulting in an overdetermined system of 81 equations. During this test, image pairs of the same digit has been aligned, it meant ≈ 2000 test cases. The images were of size 28×28 , some example images and registration results can be found in Fig. 4.16. The results were compared to the Shape Context [16] method, which also uses a thin plate spline model but control points are placed on the contours of the shapes. Comparative results in Table 4.5 show that our method provided slightly better results within 1/3 CPU time.

4.5.3 An Industrial Application

For easier installation, signs are printed on the surface of hoses used in car industry. Complete signs are expected for admissible quality, thus these signs are visually inspected during manufacturing. We were developing an automatic inspection system, in which the signs were compared to corresponding template shapes. In this process, the proposed method was used for the shape alignment. The main challenges were the segmentation errors and the complex distortions. The transformation model of the printing procedure is as follows:

1. The stamp (basically a planar template of the sign) is positioned on the hose surface. This can be described by a 2D rotation and scaling $S : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ of the *template*.
2. Then the stamp is pressed onto the surface, modeled as a transformation $\gamma :$

$\mathbb{R}^2 \rightarrow \mathbb{R}^3$ that maps the *template's* plane to a cylinder with radius r :

$$\gamma(\mathbf{x}) = \left[r \sin \frac{x_1}{r}, x_2, -r \cos \frac{x_1}{r} \right]^T.$$

3. Finally, a picture is taken with a camera, which is described by a classical projective transformation $\mathbf{P} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ with an unknown camera matrix.

Thus the transformation

$$\varphi(\mathbf{x}) = (\mathbf{P} \circ \gamma \circ \mathbf{S})(\mathbf{x}) \quad (4.47)$$

acting between a planar *template* and its distorted *observation* has 11 parameters: \mathbf{S} has 3 parameters, γ has one (r), and \mathbf{P} has 7 (six extrinsic parameters and the focal length). Different ω sets has been tested in order to obtain the best alignments. The function set in Eq. (4.42) provided the best results with parameters using all combinations for $\alpha_i \in \{0, \frac{\pi}{6}, \frac{\pi}{3}\}$ and $(n_i, m_i) \in \{(1, 2), (2, 1), (1, 3), (3, 1)\}$ yielding a system of 12 equations. The method has been tested on more than 150 image pairs and it proved to be efficient in spite of segmentation errors and severe distortions.

4.6 Summary

In this chapter, a binary image registration method has been described. This general framework can be applied to any diffeomorphic transformation model. Using only the coordinates of the foreground pixels of the shapes, it sets up a system of equations and the solution of this system provides the parameters of the aligning transformation. Thus the approach does not require any feature detection and established correspondences. The algorithm is easy to implement and experiments showed that it provides accurate alignments. The robustness of the method against segmentation errors has also been demonstrated, and it has been applied in the case of different real-world applications where it clearly outperformed state-of-the-art correspondence based methods.

Chapter 5

Affine Invariants Based Projective Registration of Binary Shapes

In the previous chapter, we described a robust method to estimate the parameters of nonlinear transformations between binary shapes. The framework was sufficiently applied to general nonlinear transformations, e.g., polynomial and thin plate spline, and it been proved to be efficient in case of many real applications, e.g., matching handwritten characters or aligning multimodal prostate images. Although this approach proved to be very robust against the strength of the deformation, in some cases (e.g., when the shapes are rotated more than 90 degrees), it can not find the right solution due to the iterative minimization involved.

In this chapter another method is described to estimate the parameters of projective transformations between binary shapes. The transformation is decomposed and the perspective and the affine parts are recovered in two sequential steps. Similarly as the method described in the previous chapter, it does not need any feature extraction or established correspondences, it works only with the coordinates of the foreground pixels.

5.1 The Registration Method

We are looking for the parameters of a two dimensional projective transformation (also called *planar homography*) $h : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $h(\mathbf{x}) = [h_1(\mathbf{x}), h_2(\mathbf{x})]^T$ that aligns the input binary images, as it was described in Section 4.2.1. The parameters of such a transformation are the H_{ij} elements of the 3×3 matrix \mathbf{H} , and $H_{33} = 1$ is fixed. The template and the observation shapes are represented by their foreground regions $\mathcal{F}_t \subset \mathbb{R}^2$ and $\mathcal{F}_o \subset \mathbb{R}^2$. Thus the correspondence between the two shapes can be

written as

$$\mathcal{F}_t = h(\mathcal{F}_o). \quad (5.1)$$

5.1.1 Decomposition of the Transformation

The H_{31} and H_{32} parameters of the transformation are responsible for the *perspective* distortion, while the others generate an *affine* transformation (translation, rotation, scaling and shearing). These two parts of the transformation can be performed one after the other, so thus h can be decomposed as follows:

$$h = h^a \circ h^p \quad (5.2)$$

where $h^p : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $h^p(\mathbf{x}) = [h_1^p(\mathbf{x}), h_2^p(\mathbf{x})]^T$ is a nonlinear transformation:

$$\begin{aligned} h_1^p(\mathbf{x}) &= \frac{x_1}{p_1 x_1 + p_2 x_2 + 1} \\ h_2^p(\mathbf{x}) &= \frac{x_2}{p_1 x_1 + p_2 x_2 + 1}, \end{aligned} \quad (5.3)$$

resulting only a perspective distortion. It has only two parameters p_1 and p_2 . The affine component $h^a : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $h^a(\mathbf{x}) = [h_1^a(\mathbf{x}), h_2^a(\mathbf{x})]^T$ is a linear transformation with 6 parameters:

$$\begin{aligned} h_1^a(\mathbf{x}) &= a_{11}x_1 + a_{12}x_2 + a_{13} \\ h_2^a(\mathbf{x}) &= a_{21}x_1 + a_{22}x_2 + a_{23}. \end{aligned} \quad (5.4)$$

The a_{13} and a_{23} parameters are responsible for translation, while scaling, shearing and rotation are induced by the remaining four parameters.

Now we can write the relationship between the shapes as follows:

$$\mathcal{F}_t = (h^a \circ h^p)(\mathcal{F}_o) = h^a(h^p(\mathcal{F}_o)). \quad (5.5)$$

The proposed method estimates the p_1, p_2 parameters of the perspective component h^p and the a_{ij} parameters of the affine component h^a in two distinct steps, then according to Eq. (5.2) we can obtain the H_{ij} parameters of h as follows:

$$\begin{aligned} H_{11} &= a_{11} + p_1 a_{13}, & H_{12} &= a_{12} + p_2 a_{13}, \\ H_{21} &= a_{21} + p_1 a_{23}, & H_{22} &= a_{22} + p_2 a_{23}, \\ H_{13} &= a_{13}, & H_{23} &= a_{23}, & H_{31} &= p_1, & H_{32} &= p_2 \end{aligned} \quad (5.6)$$

5.1.2 Estimation of the Perspective Distortion

If Eq. (5.5) stands then there is only an affine transformation between \mathcal{F}_t and $h^p(\mathcal{F}_o)$, thus for any affine-invariant function $I : \mathbb{R}^2 \rightarrow \mathbb{R}$:

$$I(\mathcal{F}_t) = I(h^p(\mathcal{F}_o)). \quad (5.7)$$

Note that the unknowns of this equation are the p_i parameters of h^p . Moreover as we will show for given values of p_1 and p_2 using traditional moment-based affine invariants the right hand side of the equation can be efficiently estimated using only the Jacobian of h^p , so it is not necessary to actually generate the image $h^p(\mathcal{F}_o)$ which would be very time-consuming.

The basic idea of the proposed method is that given a set of independent affine invariant functions $I_i : \mathbb{R}^2 \rightarrow \mathbb{R}$, $i = 1 \dots n$ we obtain a system of equations:

$$I_i(\mathcal{F}_t) = I_i(h^p(\mathcal{F}_o)). \quad (5.8)$$

The parameters of h^p are obtained as the solution of this system of equations. It is clearly a highly nonlinear system and thus do not have exact solution. However as experimental results show it can be efficiently solved by a general nonlinear solver.

Although any set of affine invariant functions could be appropriate, herein we use affine moment invariants [34], because it allows a very efficient numerical estimation of the system of equations in Eq. (5.8). The left hand sides of the system of equations Eq. (5.8) do not depend on the parameters of h^p so can be estimated directly using the point coordinates of the template image. The geometric moment m_{rs} of order $(r + s)$ of a shape \mathcal{F} is defined as

$$m_{rs}(\mathcal{F}) = \int_{\mathcal{F}} x_1^r x_2^s d\mathbf{x}. \quad (5.9)$$

The affine moment invariants I_i of a shape are rely on the so called central moments that are defined as follows:

$$\mu_{rs}(\mathcal{F}) = \int_{\mathcal{F}} (x_1 - c_1)^r (x_2 - c_2)^s d\mathbf{x} \quad (5.10)$$

where the coordinates of the center of mass of the shape are given by using the geometric moments:

$$c_1 = \frac{m_{10}(\mathcal{F})}{m_{00}(\mathcal{F})} \quad \text{and} \quad c_2 = \frac{m_{01}(\mathcal{F})}{m_{00}(\mathcal{F})}. \quad (5.11)$$

The affine moment invariants $I_i(F)$ then are obtained using these central moments. For example the first two affine moment invariants are given as follows:

$$\begin{aligned} I_1 &= (\mu_{20}\mu_{02} - \mu_{11}^2)/\mu_{00}^4 \\ I_2 &= (-\mu_{30}^2\mu_{03}^2 + 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} - 4\mu_{30}\mu_{12}^3 \\ &\quad - 4\mu_{21}^3\mu_{03} + 3\mu_{21}^2\mu_{12}^2)/\mu_{00}^{10}. \end{aligned} \quad (5.12)$$

Algorithm 6: Pseudo code of the affine invariants based projective registration method

Input : Binary images: *template* and *observation*

Output: The H_{ij} parameters of the aligning planar homography

- 1 Choose a set of affine invariants $I_i : \mathbb{R}^2 \rightarrow \mathbb{R}$ ($i = 1, \dots, n$)
 - 2 The solution of the system of equations Eq. (5.8) gives the perspective parameters p_1 and p_2
 - 3 Solve the system of equations Eq. (5.17) to determine the affine parameters a_{ij}
 - 4 Use Eq. (5.6) to obtain the H_{ij} parameters of the aligning planar homography
-

For more on affine moment invariants see [34].

Given fixed parameters of h^p we show how to compute the right hand side of the equations Eq. (5.8) avoiding the generation of the image $h^p(\mathcal{F}_o)$ by making use of the Jacobian determinant $|J_{h^p}|$ of the transformation. For a shape \mathcal{F} that is distorted by h^p the geometric moment can be estimated as follows:

$$m_{rs}(h^p(\mathcal{F})) = \int_{\mathcal{F}} [h_1^p(\mathbf{x})]^r [h_2^p(\mathbf{x})]^s |J_{h^p}(\mathbf{x})| d\mathbf{x} \quad (5.13)$$

where the Jacobian of the perspective distortion is given by

$$|J_{h^p}(\mathbf{x})| = \frac{1}{(p_1 x_1 + p_2 x_2 + 1)^3}, \quad (5.14)$$

(for more details on the usage of the Jacobian and mathematical derivation of such equations see Chapter 4 or [35]). On the perspectively distorted shape $h^p(\mathcal{F})$ the central moments are given by

$$\mu_{rs}(h^p(\mathcal{F})) = \int_{\mathcal{F}} [h_1^p(\mathbf{x}) - c_1]^r [h_2^p(\mathbf{x}) - c_2]^s |J_{h^p}(\mathbf{x})| d\mathbf{x}, \quad (5.15)$$

where

$$c_1 = \frac{m_{10}(h^p(\mathcal{F}))}{m_{00}(h^p(\mathcal{F}))} \quad \text{and} \quad c_2 = \frac{m_{01}(h^p(\mathcal{F}))}{m_{00}(h^p(\mathcal{F}))}. \quad (5.16)$$

For fixed values of the parameters p_1 and p_2 the affine moment invariants $I(h^p(F))$ in the right hand side of the system Eq. (5.8) can be obtained using the central moments in Eq. (5.15) that can be estimated using only the foreground points of the shape \mathcal{F} . Thus we avoid to generate the $h^p(\mathcal{F})$ images which would be very time consuming.

5.1.3 Estimation of the Affine Component

After the p_1 and p_2 parameters of the perspective distortion are estimated, the affine transformation h^a should be recovered between the shapes \mathcal{F}_t and $h^p(\mathcal{F}_o)$. For that

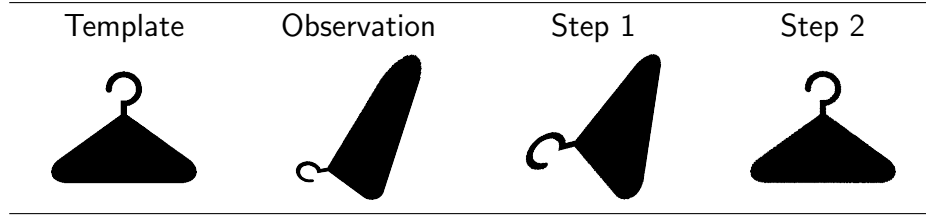


Figure 5.1: The registration process: The first step removes only the perspective distortion from the observation image while the second step restores the affine transformation and thus align it to the original template image.

purpose we used the approach proposed in [30]. In order to avoid the generation of $h^p(\mathcal{F}_o)$ we modified the method by making use of the Jacobian $|J_{h^p}|$ of the perspective component. The following system of equations

$$\int_{F_t} y_k^n dy = |J_{h^a}| \sum_{i=1}^n \binom{n}{i} \sum_{j=0}^i \binom{i}{j} a_{k1}^{n-i} a_{k2}^{i-j} a_{k3}^j \int_{F_o} h_1^p(\mathbf{x})^{n-i} h_2^p(\mathbf{x})^{i-j} |J_{h^p}(\mathbf{x})| d\mathbf{x}, \quad (5.17)$$

for $n = 1, 2, 3$ and $k = 1, 2$ can be written for the a_{ij} parameters of the affine transformation. This system contains six polynomial equations up to order three which is enough to solve for all the 6 unknowns.

The Jacobian of an affine transformation is constant over the whole plane, and it equals to the ratio of the areas of the shapes:

$$|J_{h^a}| = \frac{\int_{F_t} dy}{\int_{F_o} |J_{h^p}(\mathbf{x})| d\mathbf{x}} \quad (5.18)$$

Although the system of equations in Eq. (5.17) may have many solutions, we can select that real root which corresponds to the determinant what we computed in Eq. (5.18). Note that the solution is not unique if the shape is affine symmetric.

Putting together the projective transformation h^p and the affine transformation h^a using Eq. (5.6) we get the H_{ij} parameters of the aligning planar homography transformation h .

5.2 Implementational Issues

Although the two unknowns of the perspective part h^p would require only a system of two equations, *i.e.*, two invariants in system Eq. (5.8), the independence of the chosen set of affine moment invariants is not guaranteed. Accordingly we found that using overdetermined system of equations the method provides more accurate alignments.

	$\delta(\%)$			ε (pixel)		
	m	μ	σ	m	μ	σ
Shape Context	10.10	13.36	11.02	-	-	-
Parametric	10.54	14.31	13.99	151.60	122.72	87.21
AMI Based	2.38	6.23	12.34	1.88	23.46	53.00

Table 5.1: Test results on the synthetic dataset of Shape Context, the parametric framework (Chapter 4) and the AMI based method. m , μ , and σ denote the median, mean, and deviation.

The results shown in Section 5.3 was obtained using the $\{I_3, I_4, I_5, I_6\}$ set of invariants. While the system is clearly nonlinear we found that it could be solved efficiently using nonlinear optimization methods, and in our experiments, we used the standard differential evolution [99] method. The equations are constructed in the continuum but in practice we only have a limited precision digital image. Consequently, the integrals over the continuous domains \mathcal{F}_t and \mathcal{F}_o can only be approximated by discrete sums over the set of foreground pixel. The pseudo code of the proposed algorithm can be found in Algorithm 6 while Fig. 5.1 shows example registration results after the two subsequent steps of the method.

5.3 Experiments

We examined the performance of the proposed method on a large synthetic dataset consist of 35 different template shapes and their projectively distorted versions as observations a total of ≈ 1100 images of size 300×300 . The applied projective transformations were randomly composed of $0.5, \dots, 1.5$ scalings; $-\frac{\pi}{4}, \dots, \frac{\pi}{4}$ rotations along the x and y axes and $-\pi, \dots, \pi$ along the z axe; $-1, \dots, 1$ translations along both x and y axis and $0.5, \dots, 2.5$ along the z axis; and a random focal length chosen from the $[0.5, 1.5]$ interval.

Registration results were quantitatively evaluated using the two kind of error measures defined in Eq. (4.40). The summary of the results is shown in Table 5.1.

For comparison we examined the results of the parametric framework described in Chapter 4 on the same dataset. It provided weaker results since could not cope with cases when the images were rotated more than 90 degrees. We have also compared the performance of our method to that of *Shape Context* [16], a more general shape matching algorithm developed for nonlinear registration of binary images. For our experiments, we used the program provided by the authors, its parameters were set empirically to their optimal value ($\beta_{init} = 30$, $n_{iter} = 30$, annealing rate

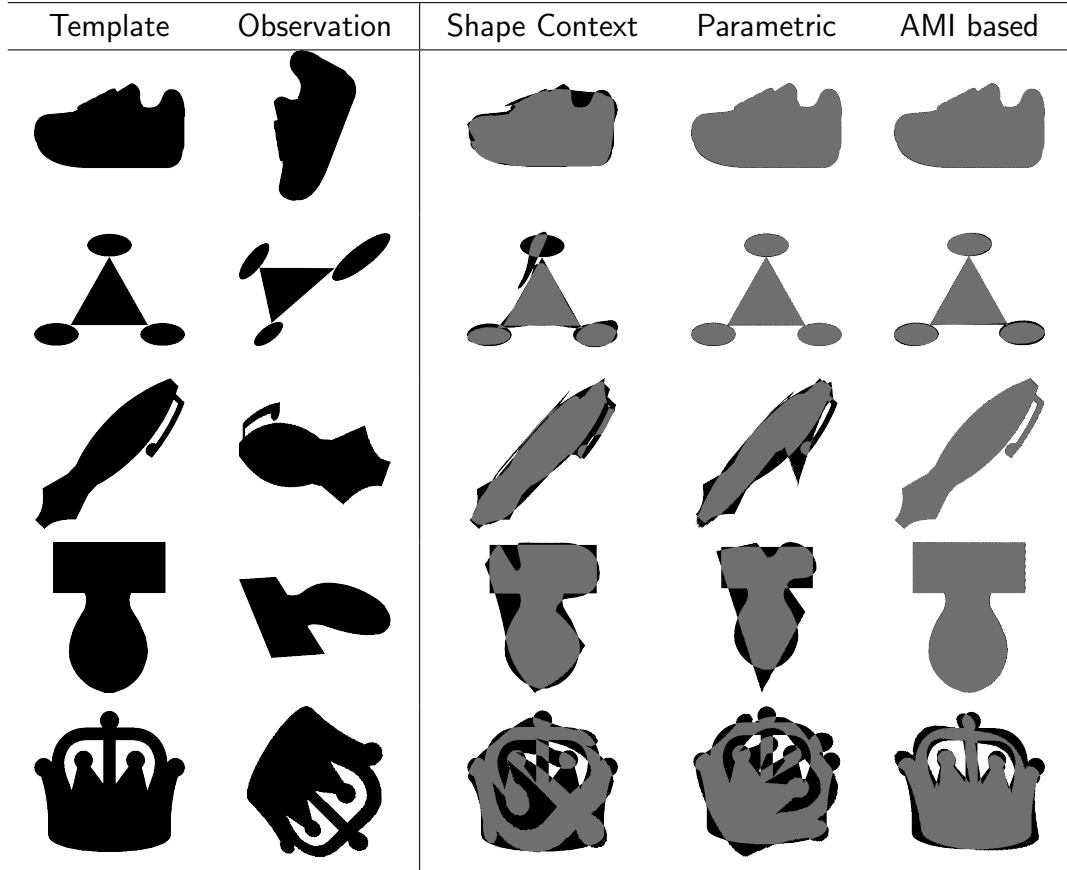


Figure 5.2: Example images from the synthetic data set and registration results obtained by Shape Context [16], the parametric framework (Chapter 4) and the AMI based method. The template and the registered observation were overlaid, overlapping pixels are depicted in gray whereas nonoverlapping ones are shown in black.

$r = 1$). The average running time of the parametric and the AMI based method (C implementations) were 4.75 sec. and 12.86 sec. respectively. The runtime of the Matlab implementation of Shape Context (68.87 sec.) is not authoritative. We also tested the proposed method on traffic signs. Some registration results can be found in Fig. 5.3

5.4 Discussion

In our experiments, we found that in its current form, the algorithm is not robust enough against segmentation errors. On the one hand, this is due to that the applied affine moment invariants are based on high-order geometrical moments, and these statistics are sensitive to small shape changes. On the other hand, when the first step could not estimate accurately the p_1 and p_2 parameters, and thus the perspective



Figure 5.3: Registration results on traffic signs. The images used as *observations* are shown in the first row, and below them the corresponding *templates* with the overlaid contours of the registration results.

distortion was not eliminated perfectly, the affine registration method in the second step failed to align the shapes. For example, in some cases, when the template and observation images contained slightly different variant of the same traffic sign, the proposed method failed to provide good results.

Both steps of the method could be improved to make the algorithm more robust against segmentation errors. First of all the choice of the invariants could be examined theoretically or even by an exhaustive series of experimental tests, to find the most robust set. On the other hand, different kind of affine invariant features could also be considered (e.g., [21]). At the same time, other affine registration methods (e.g., [29]) could be tested as second step of the method.

5.5 Summary

In this chapter, we proposed a method to estimate projective transformations between binary shapes. The two-step approach first estimates the perspective distortion independently of the affine part of the transformation which is recovered in the second step. While classical image registration methods usually use feature correspondences, the proposed method works with invariant moments estimated using the point coordinates of the objects. The efficiency of the algorithm has been demonstrated on synthetic dataset as well as on real traffic sign images. Comparative test showed that the algorithm provides better results as the method of Chapter 4, especially in the case

of strong deformations, *e.g.*, when the shapes are rotated more than 90 degrees. The method in its described form is sensitive to segmentation errors, however this can be improved in the future work.

Appendix A

Summary in English

Binary shapes play important role in the field of image processing, due to that

1. the images of many real world objects are basically binary shapes, e.g., text characters, traffic signs, bones or implants on X-ray images, etc.
2. in most of the image processing applications at some point of the processing pipeline the images are binarized (*i.e.*, segmented).

In many cases, the additional information that we are working with a binary valued image limits the number of possible solutions (the search space) thus it can help to obtain more accurate results. For example, the linear inverse problems like tomography and deconvolution are usually under-determined, and have many possible solutions. Knowing that we are seeking a binary image restricts the search space, and as a result, for example in the case of tomography, binary images can be usually reconstructed accurately only from a few projections.

In many other cases, however, the lack of rich intensity information makes it difficult to deal with these binary images. For example image registration techniques commonly work with previously established point pairs and these correspondences are usually obtained using the intensity patterns around points. Thus in the case of binary images these methods can not obtain appropriate point correspondences. On the other hand, in the case of binary images we do not need to deal with the intensity change between the images. Therefore many techniques has been presented previously to register binary images using statistics computed using the point coordinates of the shapes.

Chapter 2 addresses the problem of discrete tomography. In this field, it is commonly assumed that the intensity values of the images (*i.e.*, the attenuation coefficients of the materials) are a priori accurately known. In practice, however, this information is usually not available and hard to determine. The author proposes novel reconstruction method for those cases, when the intensity values are unknown. Using higher order

statistics based discretization term, the solution automatically converges to binary sets. Comparative tests showed, that the proposed method is more robust to the projection noise as a state-of-the-art algorithm. The method has also been successfully applied on real data.

In Chapter 3, a binary tomography based image deblurring method is introduced. Tikhonov regularization based 1-dimensional deconvolution is proposed to restore the horizontal and vertical projections. In this approach, the L-curve method is used to determine the regularization parameter which provides the best trade-off between the norm of the residual (data-fit term) and the norm of the solution (smoothness). To reconstruct the shapes from the restored projections, a maximum-flow based binary tomography method is applied. Comparative synthetic tests show that the method can provide reliable result even in the case of low signal-to-noise ratio images, thanks to that the L-curve method can provide a good regularization parameter for any noise level. Results on real out-of-focus images are also presented.

In Chapter 4, the nonlinear registration of binary images is addressed. A general framework is proposed to estimate the parameters of a diffeomorphism that aligns two binary shapes. The framework is applied to different widely-used transformation classes. Comparisons to other state-of-the-art methods are conducted and the robustness to different type of segmentation errors is examined. The algorithm has also been successfully applied in different real-world applications.

Chapter 5 introduce a technique to estimate the parameters of a planar homography transformation between binary shapes. In this method the planar homography is decomposed into the perspective distortion and the affine transformation and the parameters of these two parts are estimated separately in two consecutive steps. Comparative synthetic tests and results on real images are presented.

Summary of the Thesis Points

In the followings the list of the key points of the dissertation is given. Table A shows the connection between the thesis points and the publications of the author. The planar homography related results of the thesis point III. was also previously presented in [77].

- I.) The author introduces a novel higher order statistics based binary tomography reconstruction technique, which can be used in those cases, when the intensity values of the images are unknown. He proposes an objective functional in which a discretization term is applied to prescribe binary solutions. He propose to minimize the objective functional by a graduated non-convexity optimization

approach, in which the weight of the discretization term is increased during the optimization process to gradually enforce binary solutions. He proposes to estimate the mid-level of the intensities directly to the intermediate solutions in each iteration step. This mid-level is estimated as the minimum of the discretization term. The author examines the convergence properties of the method and shows that the behaviour of the method is independent of the value of the intensities. He demonstrates the robustness of the algorithm against different strengths of projection noise. The author compares his algorithm to state-of-the-art methods and shows that his approach is a good alternative. He also successfully applies his algorithm to real projection data.

- II.) The author proposes a binary tomography based approach for the deconvolution of binary images. He proposes to deconvolve the projections of the blurred images using a Tikhonov regularization based approach, in which the optimal value of the regularization parameter is found by the L-curve method. To reconstruct the shapes from their deblurred projections, he applies a maximum-flow based binary tomography method. He shows in comparative tests that his method provides more reliable results than another widely-used method. He also presents results on real out-of-focus images.
- III.) The author addresses the problem of nonlinear registration of binary shapes. A general registration framework is used which traces back the registration problem to a system of nonlinear equations. He applies the framework to different nonlinear transformation classes such as planar homography, polynomial transformations, and thin plate splines. He goes into the implementational details and shows that the equations can be written in three alternative forms which improves the registration results. In the case of planar homography, the author shows that the performance of the algorithm can be improved if we apply it to the Taylor series expansion of the transformation. The author proposes different ω function sets to construct the system of equations and compares them. He proposes to normalize the equations to guarantee equal contribution to the objective functional. The author compares the method to other state-of-the-art methods on synthetic datasets, and examines the robustness of the algorithm against different types of segmentation errors. The author shows that the method can be applied in different real world applications.
- IV.) The author proposes an affine moment invariants based approach to estimate the parameters of a planar homography transformation between binary shapes. He shows how to decompose the transformation into a perspective and an affine part and recover their parameters separately in two consecutive steps. He shows

how to estimate the projective parameters of the transformation using affine moment invariants. As a second step, he applies an affine registration method to determine the remaining parameters of the transformation. The author compares the method to the algorithm presented in the previous thesis point and shows results obtained by the algorithm on real images.

	[73]	[74]	[76]	[75]	[32]	[72]
I.	•					
II.		•				
III.			•	•	•	
IV.						•

Table A.1: Correspondence between the thesis points and the publications.

Appendix B

Summary in Hungarian

A bináris képek fontos szerepet játszanak a képfeldolgozás területén, mivel

1. sok objektum képe alapján véve bináris, például karakterek, KRESZ táblák, csontok vagy implantátumok a röntgenképeken, stb,
2. a képfeldolgozó rendszerek többségében a feldolgozás bizonyos pontján a képeket binarizálják, azaz szegmentálják.

Sok esetben az a többletinformáció, hogy a kép, amivel dolgozunk csak két intenzitásértéket tartalmaz, leszűkíti a keresési teret, azaz a lehetséges megoldások halmazát, ezért segít, hogy pontosabb eredményekhez jussunk. Például a lineáris inverz problémák (mint például a tomográfia vagy a dekonvolúció) sokszor aluldefiniáltak és több lehetséges megoldásuk van. Tudva, hogy egy bináris képet keresünk, leszűkíthetjük a keresési teret. Ennek köszönhetően például a tomográfia esetében a bináris képek sokszor néhány vetületből is pontosan rekonstruálhatóak.

Más esetekben azonban az intenzitásértékek változatosságának hiánya megnehezíti a bináris képekkel való munkát. Például a képregisztrációs módszerek gyakran korábban kinyert pontmegfeleltetésekből dolgoznak és ezek a megfeleltetések általában a pontok körüli intenzitás mintázatok alapján történnek. Ezért bináris képeken ezek a módszerek nem tudnak megfelelő pontmegfeleltetéseket kinyerni. Másrészt bináris képek esetén nem kell foglalkoznunk a képek közötti intenzitás-változással. Ezért sok olyan módszert fejlesztettek ki, melyek a bináris alakzatok regisztrációjához a pontkoordinátákból számított statisztikákat használják.

A 2. fejezetben diszkrét tomográfiával foglalkozunk. Ezen a területen gyakran feltételezik, hogy a rekonstruálandó képeken lévő intenzitásértékek (s így az anyagok elnyelődési együtthatói) apriori ismertek. A gyakorlatban azonban ez az információ általában nem áll rendelkezésre és nehéz meghatározni. Egy új rekonstrukciós módszert mutatunk be azokra az esetekre, amikor az intenzitás szintek ismeretlenek. Magasabb

fokú statisztika alapú diszkretizációs függvény segítségével a megoldás automatikusan binárizálódik. Összehasonlító tesztek megmutatták, hogy a bemutatott módszer robosztusabb a vetületi zajjal szemben mint egy másik, korszerű módszer. A módszer valós adatokon elért eredményei is bemutatásra kerültek.

A 3. fejezetben bináris tomográfia alapú képélesítő módszert mutatunk be. Egy Tikhonov-regularizáció alapú 1-dimenziós dekonvolúciós eljárással helyreállítjuk a kép függőleges és vízszintes vetületeket. L-curve módszert használunk a megfelelő regularizációs paraméter meghatározásához. Az alakzatoknak a helyreállított vetületeikből való rekonstrukciójához, egy maximális folyam alapú bináris tomográfiai módszert használunk. Összehasonlító tesztek mutatják, hogy a módszer megbízható eredményeket nyújt még alacsony jel-zaj arányú képeken is, köszönhetően annak, hogy az L-curve módszer megfelelő regularizációs paramétert ad magas zajszintek esetén is. Valós életlen képeken elért eredményeket is bemutatunk.

A 4. fejezet bináris képek nemlineáris regisztrációjával foglalkozik. Egy általános keretrendszert mutatunk be, mely bináris képek közötti diffeomorfizmusok paramétereinek meghatározására használható. A keretrendszert különböző széleskörben használt transzformációs modell esetében is alkalmazzuk. Szintetikus tesztek megmutatták, hogy a módszer robosztus különböző szegmentálási hibákkal szemben. Az algoritmust sikeresen alkalmaztuk különböző valós problémák esetén.

A 5. fejezet bináris képek között síkhomográfia paramétereinek meghatározására mutat be módszert. A síkhomográfia transzformációt felbontjuk egy perspektív és egy affin transzformációra és ezek paramétereit külön-külön határozzuk meg két egymásutáni lépésben. Összehasonlító tesztek eredményeit és valós képeken elért eredményeket is bemutatunk

Az eredmények tézisszerű összefoglalása

A következőkben felsoroljuk a szerző eredményeit tézispontokba foglalva. A Table B. táblázat mutatja, hogy a szerző publikációi mely tézispontokhoz kapcsolódnak. A III. tézispont síkhomográfiára vonatkozó eredményei a [77] munkában is be lettek mutatva.

- I.) A szerző kidolgozott egy új, magasab fokú statisztikát használó bináris tomográfiai technikát, amely azokban az esetekben is alkalmazható, amikor a képek intenzitásértékei ismeretlenek. Egy olyan célfüggvényt javasol, amelyben egy diszkretizációs tag írja elő a bináris megoldásokat. A célfüggvény minimalizálásához egy fokozatos optimalizálási stratégiát mutat be, amelyben a diszkretizációs tag súlya növekszik, így fokozatosan kényszeríti ki a bináris megoldásokat. Az intenzitásértékek középszintjének becslése a diszkretizációs tag minimalizálásával

történik. A szerző megvizsgálta a módszer konvergencia-tulajdonságait és megmutatja hogy az algoritmus viselkedése független az intenzitásértékektől. Bemutatja a módszer robusztusságát a különböző mértékű vetületi zajokkal szemben. A módszert összehasonlította más korszerű módszerekkel és megmutatta, hogy az algoritmus jó alternatívát jelent. Sikeresen használta a módszert valós adatokon.

- II.) A szerző egy bináris tomográfia alapú bináris képhelyreállítási módszert mutat be. Az elmosódott képek vetületeit Tikhonov regularizáció segítségével állítja helyre. A megfelelő regularizációs paraméter kiválasztásához az L-curve módszert javasolja. A helyreállított vetületekből egy maximum-folyam-alapú bináris tomográfiai módszerrel rekonstruálja az alakzatot. Összehasonlító teszt mutatja, hogy a módszer sok esetben jobb eredményeket ér el mint egy másik, széleskörben használt módszer. A szerző bemutat valós kamerával készített, elmosódott képeken elért eredményeket is.
- III.) A szerző bináris alakzatok nemlineáris regisztrációjával foglalkozik. Egy általános regisztrációs keretrendszert használ, amely a regisztrációs problémát egy egyenletrendszer megoldására vezeti vissza. A keretrendszert különböző nemlineáris transzformációs osztályok esetében is alkalmazza, mint például síkhomográfia, polinomiális transzformációk és thin-plate-spline transzformáció. Bemutatja az implementációs részleteket és azt, hogy az egyenletek oda-vissza történő felírásával hogyan lehet javítani a módszer eredményein. A síkhomográfia esetében megmutatja, hogy az algoritmus hatékonysága javítható, ha a transzformáció Taylor sorba fejtett változatán alkalmazzuk a keretrendszert. Különböző ω függvény halmazokat vizsgál az egyenletrendszer felírásához, és összehasonlítja őket. A szerző az egyenletek normalizálásával éri el, hogy azok azonos módon járuljanak hozzá a célfüggvényhez. A szerző összehasonlítja a módszert más korszerű technikákkal szintetikus adatokon és megvizsgálja a javasolt algoritmus robusztusságát a szegmentálási hibákkal szemben. A szerző bemutatja a módszer eredményeit több valós probléma esetén, mint például röntgenképek vagy KRESZ táblák illesztése.
- IV.) A szerző egy affin momentum invariánsokon alapuló módszert mutat be bináris alakzatok közötti síkhomográfia transzformáció paramétereinek becslésére. Megmutatja, hogyan lehet felbontani a transzformációt egy perspektív és egy affin összetevőre. A paraméterek meghatározása két, egymást követő lépésben történik. Az első lépés csak a perspektív torzulást állítja helyre a megfigyelt képen, míg a második lépés meghatározza az affin transzformációt így megkapjuk

az eredeti sablon képet. A szerző összehasonlítja módszert az előző pontban bemutatott technikával és bemutat valós képeken elért eredményeket is.

	[73]	[74]	[76]	[75]	[32]	[72]
I.	•					
II.		•				
III.			•	•	•	
IV.						•

Table B.1: A tézispontok és a szerző publikációinak kapcsolata.

Bibliography

- [1] A. M. Bruckstein A. M. Bronstein, M. M. Bronstein and R. Kimmel. Analysis of two-dimensional non-rigid shapes. *International Journal of Computer Vision*, 78:67–88, June 2008.
- [2] Anubhav Agarwal, C. V. Jawahar, and P. J. Narayanan. A survey of planar homography estimation techniques. Technical report, International Institute of Informatics Technology, Hyderabad 500019 India, June 2005.
- [3] A.H. Andersen and A.C. Kak. Simultaneous algebraic reconstruction technique (SART): A superior implementation of the ART algorithm. *Ultrasonic Imaging*, 6(1):81–94, January 1984.
- [4] Márton Balaskó, Attila Kuba, Antal Nagy, Zoltán Kiss, Lajos Rodek, and László Ruskó. Neutron-, gamma- and x-ray three-dimensional computed tomography at the budapest research reactor site. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 542(1–3):22–27, 2005. Proceedings of the Fifth International Topical Meeting on Neutron Radiography ITMNR-5 Fifth International Topical Meeting on Neutron Radiography.
- [5] Marton Balasko, Attila Kuba, Attila Tanacs, Zoltan Kiss, Antal Nagy, and Burkhard Schillinger. Comparison radiography and tomography possibilities of FRM-II (20 mw) and Budapest (10 MW) Research Reactor. In Muhammed Arif, editor, *Proceedings of the Eight World Conference WCNR-8*, pages 18–27, October 2008.
- [6] Péter Balázs and Mihály Gara. An evolutionary approach for object-based image reconstruction using learnt priors. In *Proceedings of the 16th Scandinavian Conference on Image Analysis*, pages 520–529. Springer-Verlag Berlin, Heidelberg, 2009.

- [7] JohnathanM. Bardsley and John Goldes. Regularization parameter selection and an efficient algorithm for total variation-regularized positron emission tomography. *Numerical Algorithms*, 57(2):255–271, 2011.
- [8] K. J. Batenburg. An evolutionary algorithm for discrete tomography. *Discrete Applied Mathematics*, 151:36–54, October 2005.
- [9] K. J. Batenburg. A network flow algorithm for reconstructing binary images from continuous x-rays. *Journal of Mathematical Imaging and Vision*, 30(3):231–248, 2008.
- [10] K. J. Batenburg, W. Van Aarle, and J. Sijbers. A semi-automatic algorithm for grey level estimation in tomography. *Pattern Recognition Letters*, 32:1395–1405, July 2011.
- [11] K. J. Batenburg and J. Sijbers. Optimal threshold selection for tomogram segmentation by projection distance minimization. *IEEE Transactions on Medical Imaging*, 28:676–686, 2009.
- [12] K. J. Batenburg and J. Sijbers. DART: A practical reconstruction algorithm for discrete tomography. *IEEE Transactions on Image Processing*, 20(9):2542–2553, 2011.
- [13] K. J. Batenburg, J. Sijbers, H.F. Poulsen, and E. Knudsen. DART: A robust algorithm for fast reconstruction of 3D grain maps. *Journal of Applied Crystallography*, 43:1464–1473, 2010.
- [14] Kees J. Batenburg, Marta D. Rossell, Rolf Erni, Gustaaf Van Tendeloo, and Sandra Van Aert. Three-dimensional atomic imaging of crystalline nanoparticles. *Nature*, 470:374–377, 2011.
- [15] Kees Joost Batenburg. A network flow algorithm for reconstructing binary images from discrete x-rays. *Journal of Mathematical Imaging and Vision*, 27(2):175–191, 2007.
- [16] Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape matching and object recognition using shape context. *Transaction on Pattern Analysis and Machine Intelligence*, 24(4):509–522, April 2002.
- [17] A. Blake and A. Zisserman. *Visual Reconstruction*. Artificial Intelligence. MIT Press, 2003.

- [18] Thomas Boudier. The snake plugin for ImageJ. software <http://www.snv.jussieu.fr/~wboudier/softs/snake.html>.
- [19] S. Brunetti, A. Del Lungo, F. Del Ristoro, A. Kuba, and M Nivat. Reconstruction of 4- and 8-connected convex discrete sets from row and column projections. *Linear Algebra and its Applications*, 339(1–3):37–57, December 2001.
- [20] S. Brunetti, P. Dulio, and C. Peri. Discrete tomography determination of bounded lattice sets from four x-rays. *Discrete Applied Mathematics*, 161(15):2281 – 2292, 2013.
- [21] Darshan Bryner, Anuj Srivastava, and Eric Klassen. Affine-invariant, elastic shape analysis of planar contours. In *IEEE Proceedings of Conference on Computer Vision and Pattern Recognition*, pages 390–397, June 2012.
- [22] James A. Cadzow. Blind deconvolution via cumulant extrema. *IEEE Signal Processing Magazine*, 13:24–42, May 1996.
- [23] T.D. Capricelli and P.L. Combettes. A convex programming algorithm for noisy discrete tomography. In Gabor T. Herman and Attila Kuba, editors, *Advances in Discrete Tomography and Its Applications*, Applied and Numerical Harmonic Analysis, pages 207–226. Birkhäuser Boston, 2007.
- [24] Thomas Capricelli. Convex processing. software <http://labs.freehackers.org/projects/cvx-processing>.
- [25] Yair Censor. Binary steering in discrete tomography reconstruction with sequential and simultaneous iterative algorithms. *Linear Algebra and its Applications*, 339:111–124, December 2001.
- [26] Marek Chorbak and Christoph Dürr. Reconstructing hv-convex polyominoes from orthogonal projections. *Information Processing Letters*, 69:283–289, 1999.
- [27] M. Holden D. L. G. Hill, P. G. Batchelor and D. J. Hawkes. Medical image registration. *Physics in Medicine and Biology*, 46:R1–R45, March 2001.
- [28] S. R. Deans. *The Radon transform and some of its applications*. A Wiley-Interscience publication. Wiley, 1983.
- [29] Csaba Domokos and Zoltan Kato. Binary image registration using covariant gaussian densities. In Aurélio Campilho and Mohamed Kamel, editors, *Proceedings of International Conference on Image Analysis and Recognition*, volume 5112 of *Lecture Notes in Computer Science*, pages 455–464, Póvoa de Varzim, Portugal, June 2008. Springer.

- [30] Csaba Domokos and Zoltan Kato. Parametric estimation of affine deformations of planar shapes. *Pattern Recognition*, 43:569–578, March 2010.
- [31] Csaba Domokos, Zoltan Kato, and Joseph M. Francos. Parametric estimation of affine deformations of binary images. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, pages 889–892, Las Vegas, Nevada, USA, April 2008.
- [32] Csaba Domokos, Jozsef Nemeth, and Zoltan Kato. Nonlinear shape registration without correspondences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(5):943–958, May 2012.
- [33] M.R. Downing, P.E. Undrill, P. Ashcroft, D.W.L. Hukins, and J.D. Hutchison. Automated femoral measurement in total hip replacement radiographs. In *Proceedings of International Conference on Image Processing and Its Applications*, volume 2, pages 843–847, Dublin, Ireland, July 1997. IEEE.
- [34] Jan Flusser, Tomás Suk, and Barbara Zitová. *Moments and Moment Invariants in Pattern Recognition*. Wiley & Sons, October 2009.
- [35] J.M. Francos, R. Hagege, and B. Friedlander. Estimation of multidimensional homeomorphisms for object recognition in noisy environments. In *Proceedings of Conference on Signals, Systems and Computers*, volume 2, pages 1615–1619, Pacific Grove, California, USA, November 2003.
- [36] R.J. Gardner, P. Gritzmman, and D. Prangenberg. On the computational complexity of reconstructing lattice sets from their x-rays. *Discrete Mathematics*, 202(1–3):45–71, 1999.
- [37] Vito Di Gesù, Giosuè Lo Bosco, Filippo Millonz, and Cesare Valenti. A memetic algorithm for binary image reconstruction. In ValentinE. Brimkov, RenetaP. Barneva, and HerbertA. Hauptman, editors, *Combinatorial Image Analysis*, volume 4958 of *Lecture Notes in Computer Science*, pages 384–395. Springer Berlin Heidelberg, 2008.
- [38] Ardeshir Goshtasby. Registration of images with geometric distortions. *IEEE Tran. on Geoscience and Remote Sensing*, 26(1):60–64, January 1988.
- [39] J. Gregor and T. Benson. Computational analysis and improvement of SIRT. *IEEE Transactions on Medical Imaging*, 27(7):918–924, 2008.
- [40] H. Guo, A. Rangarajan, S. Joshi, and L. Younes. Non-rigid registration of shapes via diffeomorphic point matching. In *Proceedings of International Symposium on*

- Biomedical Imaging: From Nano to Macro*, volume 1, pages 924–927, Arlington, VA, USA, April 2004. IEEE.
- [41] T. Tuytelaars H. Bay, A. Ess and L. Van Gool. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110:346–359, August 2008.
- [42] D. Groisser H. D. Tagare and O. Skrinjar. Symmetric non-rigid registration: A geometric theory and some numerical techniques. *Journal of Mathematical Imaging and Vision*, 34:61–88, May 2009.
- [43] Lajos Hajdu. Unique reconstruction of bounded sets in discrete tomography. *Electronic Notes in Discrete Mathematics*, 20:15 – 25, 2005.
- [44] Per Christian Hansen and Dianne P. O’Leary. The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM Journal on Scientific Computing*, 14:1487–1503, 1993.
- [45] Richard Hartley. In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(6):580–593, June 1997.
- [46] Gabor T. Herman and Attila Kuba, editors. *Discrete Tomography: Foundations, Algorithms, and Applications*. Birkhäuser Boston, 1999.
- [47] Gabor T. Herman and Attila Kuba, editors. *Advances in Discrete Tomography and its Applications*. Birkhäuser Boston, 2007.
- [48] G.T. Herman. *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*. Advances in Pattern Recognition. Springer London, 2009.
- [49] A. C. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging*. IEEE Press, 1988.
- [50] S. Kaneko, Y. Satohb, and S. Igarashi. Using selective correlation coefficient for robust image registration. *Pattern Recognition*, 36:1165–1173, May 2003.
- [51] Jeongtae Kim and Soohyun Jang. High order statistics based blind deconvolution of bi-level images with unknown intensity values. *Optics Express*, 18(12):12872–12889, June 2010.
- [52] C. Kisielowski, P. Schwander, F.H. Baumann, M. Seibt, Y. Kim, and A. Ourmazd. An approach to quantitative high-resolution transmission electron microscopy of crystalline materials. *Ultramicroscopy*, 58(2):131–155, 1995.

- [53] Attila Kuba and Murice Nivat. A sufficient condition for non-uniqueness in binary tomography with absorption. *Theoretical Computer Science*, 346(2–3):335–357, 2005. In memoriam: Alberto Del Lungo (1965-2003).
- [54] M. Pawan Kumar, Sujit Kuthirummal, C. V. Jawahar, and P. J. Narayanan. Planar homography from fourier domain representation. In *Proceedings of International Conference on Signal Processing and Communications*, 2004.
- [55] Edmund Y. Lam. Blind bi-level image restoration with iterated quadratic programming. *IEEE Transactions on Circuits and Systems Part 2: Express Briefs*, 54(1):52–56, 2007.
- [56] Edmund Y. Lam and Joseph W. Goodman. Iterative statistical approach to blind image deconvolution. *Journal of the Optical Society of America A*, 17(7):1177–1184, 2000.
- [57] Silvio Levy. Projective transformations. web page, <http://www.geom.uiuc.edu/docs/reference/CRC-formulas/node16.html>.
- [58] Ta-Hsin Li. A joint estimation approach for two-tone image deblurring by blind deconvolution. *IEEE Transactions on Image Processing*, 11:847–858, August 2002.
- [59] Ta-Hsin Li and Ke-Shin Lii. Deblurring two-tone images by a joint estimation approach using higher-order statistics. In *Proceedings of IEEE SP Workshop on Higher-Order Statistics*, pages 108–111, 1997.
- [60] Manolis Lourakis. homest: A C/C++ library for robust, non-linear homography estimation. software, <http://www.ics.forth.gr/~lourakis/homest/>, 2008.
- [61] David Lowe. Demo software: SIFT keypoint detector. software, <http://www.cs.ubc.ca/~lowe/keypoints/>, July 2005.
- [62] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, November 2004.
- [63] Tibor Lukić. Discrete tomography reconstruction based on the multi-well potential. *Proceedings of the 14th International Conference on Combinatorial Image Analysis*, 6636:335–345, 2011.
- [64] J. B. Antoine Maintz and Max A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, March 1998.

- [65] Steve Mann and Rosalind W. Picard. Video orbits of the projective group a simple approach to featureless estimation of parameters. *IEEE Transactions on Image Processing*, 6(9):1281–1295, September 1997.
- [66] Donald W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 11(2):431–441, 1963.
- [67] S. Marsland and C. J. Twining. Constructing diffeomorphic representations for the groupwise analysis of nonrigid registrations of medical images. *IEEE Transactions on Medical Imaging*, 23:1006–1020, August 2004.
- [68] E.V. Moisi, B. Nagy, and V.I. Cretu. Maximum flow minimum cost algorithm for reconstruction of images represented on the triangular grid. In *IEEE 8th International Symposium on Applied Computational Intelligence and Informatics*, pages 35–40, May 2013.
- [69] Ábris Nagy and Csaba Vincze. Reconstruction of hv-convex sets by their coordinate x-ray functions. *Journal of Mathematical Imaging and Vision*, 49(3):569–582, 2014.
- [70] Antal Nagy. Smoothing filters in the dart algorithm. In RenetaP. Barneva, ValentinE. Brimkov, and Josef Llapal, editors, *Combinatorial Image Analysis*, volume 8466 of *Lecture Notes in Computer Science*, pages 224–237. Springer International Publishing, 2014.
- [71] Antal Nagy and Attila Kuba. Reconstruction of binary matrices from fan-beam projections. *Acta Cybernetica*, 17(2):359–385, January 2005.
- [72] Jozsef Nemeth. Recovering projective transformations between binary shapes. In *Proceedings of Advanced Concepts for Intelligent Vision Systems*, volume 7517, pages 374–383, Brno, Czech Republic, 2012.
- [73] Jozsef Nemeth. Discrete tomography with unknown intensity levels using higher-order statistics. *Journal of Mathematical Imaging and Vision*, 2015. Online first, DOI: 10.1007/s10851-015-0581-0.
- [74] Jozsef Nemeth and Peter Balazs. Restoration of blurred binary images using discrete tomography. In *Proceedings of Advanced Concepts for Intelligent Vision Systems*, volume 8192, pages 80–90, Poznan, Poland, 2013.
- [75] Jozsef Nemeth, Csaba Domokos, and Zoltan Kato. Nonlinear registration of binary shapes. In *Proceedings of International Conference on Image Processing*, pages 1101–1104, Cairo, Egypt, November 2009. IEEE.

- [76] Jozsef Nemeth, Csaba Domokos, and Zoltan Kato. Recovering planar homographies between 2D shapes. In *Proceedings of International Conference on Computer Vision*, pages 2170–2176, Kyoto, Japan, September 2009. IEEE.
- [77] József Németh. Síkhomográfia paramétereinek becslése bináris képeken. In *Proceedings of National Scientific Students' Associations Conference*, April 2009. Supervisors: Zoltán Kató and Csaba Domokos. Note: in Hungarian.
- [78] Jorge Nocedal and Stephen J. Wright. *Numerical optimization*. Springer series in operations research and financial engineering. Springer, New York, NY, 2. ed. edition, 2006.
- [79] Alina Oprea and Constantin Vertan. A quantitative evaluation of the hip prosthesis segmentation quality in x-ray images. In *Proceedings of International Symposium on Signals, Circuits and Systems*, volume 1, pages 1–4, Iasi, Romania, July 2007. IEEE.
- [80] James B. Orlin. A polynomial time primal network simplex algorithm for minimum cost flows. In *Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '96, pages 474–481, Philadelphia, PA, USA, 1996. SIAM.
- [81] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1):62–66, January 1979.
- [82] Carlos Filipe Paulo and Paulo Lobato Correia. Automatic detection and classification of traffic signs. In *Proceedings of Workshop on Image Analysis for Multimedia Interactive Services*, pages 11–14, Santorini, Greece, June 2007.
- [83] Marc Pollefeys. Scene planes and homographies. Presentation, www.cs.unc.edu/~marc/mvg/course16.ppt.
- [84] William Hadley Richardson. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62(1):55–59, Jan 1972.
- [85] A. Rosenfeld and A.C. Kak. *Digital Picture Processing*. 2.ed. Number v. 1-2. Academic Press, 1982.
- [86] Peter J. Rousseeuw. Least median of squares regression. *Journal of the American Statistical Association*, 79(388):871–880, December 1984.
- [87] Stephane Roux, Hugo Leclerc, and Francois Hild. Efficient binary tomographic reconstruction. *Journal of Mathematical Imaging and Vision*, 49(2):335–351, 2014.

- [88] H.J. Ryser. Combinatorial properties of matrices of zeros and ones. In *Classic Papers in Combinatorics*, Modern Birkhäuser Classics, pages 269–275. Birkhäuser Boston, 1987.
- [89] J. Malik S. Belongie and J. Puzicha. Shape context demo software. http://www.eecs.berkeley.edu/Research/Projects/CS/vision/shape/sc_digits.html, 2002.
- [90] J. Malik S. Belongie and J. Puzicha. Shape matching and object recognition using shape context. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):509–522, April 2002.
- [91] Robert J. Schalkoff. *Digital Image Processing and Computer Vision*. John Wiley & Sons, New York, NY, USA, 1989.
- [92] T. Schüle, C. Schnörr, S. Weber, and J. Hornegger. Discrete tomography by convex–concave regularization and D.C. programming. *Discrete Applied Mathematics*, 151:229–243, October 2005.
- [93] Thomas Schüle, Stefan Weber, and Christoph Schnörr. Adaptive reconstruction of discrete-valued objects from few projections. *Electronic Notes in Discrete Mathematics*, 20:365–384, July 2005.
- [94] Behzad Sharif and Behnam Sharif. Discrete tomography in discrete deconvolution: Deconvolution of binary images using ryser’s algorithm. *Electronic Notes in Discrete Mathematics*, 20:555–571, July 2005.
- [95] Yijiang Shen, Edmund Y. Lam, and Ngai Wong. Binary image restoration by positive semidefinite programming. *Optics Letters*, 32(2):121–123, 2007.
- [96] K. M. Simonson, S. M. Drescher, and F. R. Tanner. A statistics-based approach to binary image registration with uncertainty analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29:112–125, January 2007.
- [97] Katherine M. Simonson, Steven M. Drescher, and Franklin R. Tanner. A statistics-based approach to binary image registration with uncertainty analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29:112–125, January 2007.
- [98] A. Singer and H. Wu. Two-dimensional tomography from noisy projections taken at unknown random directions. *SIAM Journal on Imaging Sciences*, 6(1):136–175, 2013.

- [99] Rainer Storn and Kenneth Price. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359, December 1997.
- [100] Tomáš Suk and Jan Flusser. Affine normalization of symmetric objects. In Jacques Blanc-Talon, Wilfried Philips, Dan Popescu, and Paul Scheunders, editors, *Proceedings of International Conference on Advanced Concepts for Intelligent Vision Systems*, volume 3708 of *Lecture Notes in Computer Science*, pages 100–107, Antwerp, Belgium, September 2005. Springer.
- [101] A.N. Tikhonov and V.I.A. Arsenin. *Solutions of ill-posed problems*. Scripta series in mathematics. Winston, 1977.
- [102] Jeannot Trampert and Jean-Jacques Leveque. Simultaneous iterative reconstruction technique: Physical interpretation based on the generalized least squares solution. *Journal of Geophysical Research*, 95(B8):12553–12559, August 1990.
- [103] Cesare Valenti. A genetic algorithm for discrete tomography reconstruction. *Gen. Progr. Evolv. Mach.*, 9(1):85–96, March 2008.
- [104] Wim van Aarle, Kees Joost Batenburg, and Jan Sijbers. Automatic parameter estimation for the discrete algebraic reconstruction technique (DART). *IEEE Transactions on Image Processing*, 21(11):4608–4621, 2012.
- [105] Wim van Aarle, Karel Crombecq, Ivo Couckuyt, K. Joost Batenburg, and Jan Sijbers. Efficient parameter estimation for discrete tomography using adaptive modelling. In *Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, pages 229–232, July 2011.
- [106] Laszlo Varga, Peter Balazs, and Antal Nagy. Projection selection algorithms for discrete tomography. In J. Blanc-Talon, D. Bone, W. Philips, D. Popescu, and P. Scheunders, editors, *Proceedings of the Advanced Concepts for Intelligent Vision Systems*, volume 6474 of *Lecture Notes in Computer Science*, pages 390–401, Sydney, Australia, December 2010.
- [107] Laszlo Varga, Peter Balazs, and Antal Nagy. Direction-dependency of binary tomographic reconstruction algorithms. *Graphical Models*, 73:365–375, November 2011.
- [108] Laszlo Varga, Peter Balazs, and Antal Nagy. An energy minimization reconstruction algorithm for multivalued discrete tomography. In *Proceedings of the 3rd International Symposium on Computational Modeling of Objects Represented in Images*, pages 179–185, Rome, Italy, September 2012. Taylor & Francis.

- [109] J. A. Ventura and W. Wan. Accurate matching of two-dimensional shapes using the minimal tolerance zone error. *Image and Vision Computing*, 15:889–899, December 1997.
- [110] M.N. Vrahatis, G.S. Androulakis, J.N. Lambrinos, and G.D. Magoulas. A class of gradient unconstrained minimization algorithms with adaptive stepsize. *Journal of Computational and Applied Mathematics*, 114(2):367 – 386, 2000.
- [111] Stefan Weber. *Discrete tomography by convex–concave regularization using linear and quadratic optimization*. PhD thesis, Heidelberg University, 2009.
- [112] Norbert Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*. John Wiley & Sons, New York, NY, USA, 1949.
- [113] Ralph A Wiggins. Minimum entropy deconvolution. *Geoprospection*, 16(1–2):21–35, 1978.
- [114] S. Wörz and K. Rohr. Physics-based elastic registration using non-radial basis functions and including landmark localization uncertainties. *Computer Vision Image Understanding*, 111(3):263–274, September 2008.
- [115] H.S. Wu. Minimum entropy deconvolution for restoration of blurred two-tone images. *Electronic Letters*, 26:1183–1184, July 1990.
- [116] A. Yezzi, L. Zöllei, and T. Kapur. A variational framework for joint segmentation and registration. In *Proceedings of IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 44–51, Kauai, HI, USA, December 2001. IEEE.
- [117] O. Faugeras Z. Zhang, R. Deriche and Q. T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78(1-2):87–119, October 1995.
- [118] L. Zagorchev and A. Goshtasby. A comparative study of transformation functions for nonrigid image registration. *IEEE Transactions on Image Processing*, 15:529–538, March 2006.
- [119] Barbara Zitová and Jan Flusser. Image registration methods: A survey. *Image and Vision Computing*, 21(11):977–1000, October 2003.