SUMMARY OF THE PH.D. THESIS

Texture, Nonogram, and Convexity Priors in Binary Tomography

Author:
Judit SZŰCS

Supervisor:
Péter BALÁZS, Ph.D.
associate professor

Doctoral School of Computer Science
Department of Image Processing and Computer Graphics
Faculty of Science and Informatics
University of Szeged, Hungary

Szeged, 2021

1 Introduction

The basic aim of tomography is to reconstruct the description (or the model) of the source object from its slices. Slices are 2D cross-sectional images that are sectioned from the 3D object by a 2D plane. Due to the non-destructive restrictions, slices cannot be accessed directly, only secondary information is available about them. This information is usually the sum of the material densities in a given 2D slice along certain directions.

In this study, we worked with transmission tomography, meaning the radiation comes from outside of the examined object. During an experiment, the object to be examined stands between the source of the X-ray and the detectors, which measure the exiting intensity of radiation. When an X-ray beam passes through an object, its intensity decreases due to physical mechanisms, i.e. a part of it is absorbed. Passing through different materials, the degree of absorption is different and is always specific to the given material.

The image to be reconstructed can be created with several continuous techniques (e.g. analytical methods, algebraic reconstruction methods). Unfortunately, these techniques need several hundreds of equiangularly positioned projections to achieve reasonable image quality, meaning they are almost always useless when only few projections are available. A projection is a collection of line integrals of the image taken along predefined straight lines.

In some cases, one can assume that the object to be reconstructed consists of just a few known materials, i.e. only a few different grayscale intensity values that are known in advance can appear in the image representing the object. Discrete Tomography (DT) is concerned with exactly such issues, thus, fewer projections could potentially be enough for a satisfactory reconstruction.

1.1 DART

Discrete Algebraic Reconstruction Technique (DART) is based on the ART method, taking advantage of the properties of the discrete images: the range of the image function is finite and contains just a small number of elements [3]. Its fundamental principle is that thresholding the result of an arbitrary continuous reconstruction method usually produces an approximately satisfactory result, even for discrete problems, except for some inaccuracies at the boundary of the object in the image. Therefore, after performing a continuous reconstruction and thresholding it, DART iteratively refines the boundary. Due to the fact that it uses prior information about the image to be reconstructed, DART is effective if the number of gray levels is small (five or fewer), and it can also treat noisy projections. We used this method as a reference.

2 Binary Tomography

Binary Tomography (BT) is a special case of DT, which can be utilized when the object to be reconstructed consists of a single material, yielding a corresponding image containing only 0s and 1s.

2.1 Ryser's algorithm

One of the earliest methods for solving a binary reconstruction problem is Ryser's algorithm [24], which creates a binary matrix based on the horizontal and the vertical projections of the source image. This algorithm runs in polynomial time, but usually does not result in a uniquely determined solution.

2.2 Binary Tomography with Prior Information

Owing to the insufficient amount of available data, the binary reconstruction can be extremely underdetermined from two projections. One way to reduce the search space of feasible solutions is to exploit some prior knowledge of the image to be reconstructed. There is a wide array of literature regarding image reconstruction algorithms in BT using different kinds of geometrical or shape priors, depending on the application.

Although both DART and Ryser's algorithm (and several other methods) are well-functioning algorithms, unfortunately, they are unable to use prior information. Tracing the binary reconstruction back to an optimization problem could lead to an approach where prior information can be incorporated into the reconstruction process.

2.3 Binary Reconstruction by Optimization

The binary reconstruction can be traced back to the equation known from algebraic reconstruction. To overcome the difficulties arising from underdeterminedness and inconsistency, the binary reconstruction problem is commonly reformulated as a minimization problem

$$C(\mathbf{X}) = ||\mathbf{A}\mathbf{x} - \mathbf{b}||_2 + \gamma \cdot \Phi(\mathbf{X}) \ \rightarrow \min\,, \tag{1}$$

where matrix $\mathbf{A} \in \mathbb{R}^{k \times mn}$ describes the relationship between the k beams and the pixels (a_{ij} is a weight that represents how the i^{th} beam affects the j^{th} pixel); while vector $\mathbf{b} \in \mathbb{R}^k$ contains the measured projection values; $\mathbf{x} \in \{0,1\}^{mn}$ represents the unknown image of size $m \times n$, in a row-by-row vector form ($||\mathbf{A}\mathbf{x} - \mathbf{b}||_2$ ensures that the projections of the reconstructed image are close to the measured ones); and $\Phi(\mathbf{X}) : \{0,1\}^{mn} \rightarrow \mathbb{R}$ is a regularization function expressing how image $\mathbf{x}$ fits with prior information (e.g. homogeneity, convexity, circularity, etc.). The smaller the value, the better the result. Finally, $\gamma \geq 0$ is a scaling constant. If it is established that the measured projection data is not trustworthy enough (e.g. due to noise), and that the prior information is trustworthy, then γ is assigned higher value. In case the projection data is more reliable than the prior, γ is assigned lower value.

Eq. 1 is a discrete global optimization problem, which is known to be NP-hard to solve. There are several methods to solve this optimization problem, but unfortunately, in this case, conventional gradient-based search algorithms can easily get stuck in a local optimum, as they only accept solutions of decreasing value during optimization. Hence, the use of metaheuristics (e.g. simulated annealing or a genetic algorithm) is needed.

2.4 Simulated Annealing

Simulated Annealing (SA) is a general stochastic heuristic for solving global optimization problems with suitable approximation in a large search space [18]. The search only runs on one thread. The pseudocode of SA is provided in Alg. 1.

In general, the algorithm usually starts with a random solution in its variables (initial state s), and with a given initial temperature T_0, which is decreased by the cooling factor α during the iterations. The method runs until one of the stopping criteria is met, which depends on the specification of the problem. This can be, for example, to reach a number of iterations or a certain magnitude of temperature. Moving to a neighbor means selecting a solution from the feasible ones in the current iteration. The concept of 'neighbor' depends on the specification of question. In the case of binary images, it can be defined, for instance, by inverting a randomly selected pixel. In each iteration, the algorithm randomly selects a neighbor and evaluates whether to accept or deny it and keep the previous one. The decision between acceptance or denial depends on which one can keep the system at a lower energy level. If the objective function value assigned to the new proposal is not clearly better than the previous one, an exponential expression of these values is compared to a randomly generated number between 0 and 1 (see line 12 of Alg. 1). If this value is greater than the generated number, the new solution will be accepted, otherwise it is denied. It is necessary to accept a solution with worse value than the previous one for the function not to get stuck in a local optimum. The lower the temperature is, the less chance to adopt poor solutions. This way, approximating the global optimum is guaranteed. In some iterations, the temperature is not decreased: β defines a maximum rate of the possibly denied proposals on the current temperature, while k denotes the number of iterations. With a suitable cooling strategy, the method is proven to reach the global optimum. However, the algorithm has countless parameters, and fine-tuning them can be really challenging.

Algorithm 1 Simulated Annealing

1: $s \leftarrow$ initial state
2: $T_0 \leftarrow$ initial temperature
3: $k \leftarrow 1$
4: $\alpha \in (0.5, 1)$
5: $\beta \in (0, 1)$
6: **while** $(stoppingCriteria == FALSE)$ **do**
7:     $tempStay := 0$
8:     **while** $(tempStay < \beta \cdot f(s))$ **do**
9:         $actual := neighbor(s)$
10:         **if** $(f(actual) < f(s))$ **then**
11:             $s := actual$
12:         **else if** $(e^{\frac{f(s)-f(actual)}{T_k}} > rand(0,1))$ **then**
13:             $s := actual$
14:         **end if**
15:         $tempStay := tempStay + 1$
16:     **end while**
17:     $T_{k+1} := T_k \cdot \alpha$
18:     $k := k + 1$
19: **end while**

2.5 Quality Measurement in Binary Tomography

If the original image is known, a commonly calculated figure-of-merit for determining the quality of a reconstruction is the Relative Mean Error (RME) [19], specified by

$$RME = \frac{\sum_i |x_i^o - x_i^r|}{\sum_i x_i^o} \cdot 100\% \,, \tag{2}$$

where x_i^o and x_i^r stand for the i^{th} pixel of the original and reconstructed image, respectively. RME value provides the ratio of pixel difference between the original and resulting image, and the number of object pixels of the original image. *'Pixel error'* can be defined as

$$E_p = \frac{\sum_i |x_i^o - x_i^r|}{m \cdot n} \cdot 100\% \,, \tag{3}$$

where $m \times n$ is the size of the image. Note that pixel error differs from RME in the normalizing factor only (in RME, it is the sum of the pixel values; whereas in pixel error, it is the size of the image). Pixel error ensures a more accurate comparison between the images of different sizes.

Calculating the different terms of the cost function separately (and paying attention to how they decrease during the iterations) can also result in a satisfactory measurement.

3 Binary Tomography with Local Binary Pattern Priors

First, the inherently insufficient amount of projection data was augmented by statistical image priors describing the approximate texture of the image to be reconstructed. The priors were extracted from sample images, ahead of the reconstruction.

Local Binary Patterns (LBP) is a visual descriptor suitable for finding image patterns or repetitions, mostly used for classification in computer vision. The basic LBP algorithm [23] examines the relation between the center pixel and its 8-neighbors in a 3×3 window for each pixel, which thereafter describes it as a sequence of 0s and 1s. This results in a binary code, which is converted to a decimal number. The resulted codes for all image pixels are aggregated in a 256-dimensional histogram (*LBP histogram* or *LBP vector*) representing the frequency of each possible pattern.

A soft extension of LBP is Soft LBP [1] or Fuzzy LBP (FLBP) [17], which integrates fuzzy logic into the basic LBP process. FLBP is robust and continuous with respect to the output, meaning minor changes in the input image only result in minor changes in the output. To improve the robustness of the basic LBP operator, the original threshold function is expanded by two fuzzy membership functions. Fuzzification allows FLBP to assign several (or even all) bins of the histogram to each pixel to some degree, not just a single one. While LBP is based on thresholding that makes it sensitive to noise, FLBP enables a more robust representation of image texture by fuzzifying the calculated LBP codes using fuzzy rules and membership functions.

The Shift Local Binary Patterns (SLBP) was introduced in [20] as a quick approximation of the computationally heavy FLBP. There, the authors described SLBP as generating a fixed number of local binary codes for each pixel position. An intensity limit is added to specify the shift interval. When this interval is modified, a new binary code is calculated and added to the histogram.

Since the basic LBP is not invariant to rotations and viewpoint changes caused by the fixed arrangement of weights, Dominant Rotated Local Binary Patterns (DRLBP) was proposed in [21] to ensure rotation-invariance. The main idea of the method is that the weights are aligned in a circular manner, so the effects of image rotations can be traced by rotating the weights by the same (unknown) angle.

3.1 Preliminary Studies

In some preliminary studies, we used the basic LBP descriptor as follows. Assuming that the image to be reconstructed belongs to a given class, we first selected representatives of that class and calculated their LBP vectors. Then, in the reconstruction, we took two priors into account (the image is to be smooth, and must have a texture similar to the previously observed ones). Having two different priors, the formula of Eq. 1 became

$$C(\mathbf{X}) = ||\mathbf{Ax} - \mathbf{b}||_2 + \gamma_1 \cdot \Phi_1(\mathbf{X}) + \gamma_2 \cdot \Phi_2(\mathbf{X}) \ \rightarrow \min . \tag{4}$$

$\Phi_1(\mathbf{X})$ measures the smoothness of the current solution by taking the sum of the convolution of each image pixel by a Gaussian-like kernel $\mathbf{K}$ Eq. 6. Formally,

$$\Phi_1(\mathbf{X}) = \sum_{i=3}^{m-2} \sum_{j=3}^{n-2} (\mathbf{K} * \mathbf{X}_{i,j}) , \tag{5}$$

where

$$\mathbf{K} = \frac{1}{\sqrt{2\pi}} \begin{pmatrix} 0 & 0 & e^{-2} & 0 & 0 \\ 0 & e^{-\sqrt{2}} & e^{-1} & e^{-\sqrt{2}} & 0 \\ e^{-2} & e^{-1} & 0 & e^{-1} & e^{-2} \\ 0 & e^{-\sqrt{2}} & e^{-1} & e^{-\sqrt{2}} & 0 \\ 0 & 0 & e^{-2} & 0 & 0 \end{pmatrix} , \tag{6}$$

and $\mathbf{X}_{i,j}$ represents a 5×5 sized submatrix of $\mathbf{X}$ with center point (i, j). $\Phi_2(\mathbf{X})$ provides the minimal Euclidean distance of the LBP vector of the current image compared to all representative LBP vectors. Formally,

$$\Phi_2(\mathbf{X}) = \min_t \{ ||LBP(\mathbf{X}) - LBP(\mathbf{X}_t)||_2 \} , \tag{7}$$

where $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_t$ are the t sample images from the same class as $\mathbf{X}$, known a priori. Again, γ_1 and γ_2 are scaling weights.

To test the efficiency of the method, we conducted experiments on software phantom images from 6 classes of various textures with SA. The sizes of the images are diverse, the smallest one is 430×300, while the largest one is 1230×1120.

All parameters of the SA algorithm were set manually, in an empirical manner. Each time, the method started out with a random binary image. The stopping criteria of the algorithm were to reach 800 000 iterations or to perform 50 000 iterations without improving the optimal result. The initial temperature was set to $T_1 = 5\,800$, while the cooling schedule was controlled by $\alpha = 0.99$. The SA algorithm was allowed to remain the same temperature for some iterations. In our case, this was fine-tuned by $\beta = 0.035$. Selecting a neighbor meant randomly selecting and inverting some of the image pixels. We found that inverting a single pixel per iteration was unable to decrease the cost value significantly at the beginning. The number of pixels to be inverted (l) in iteration t depended on the cost value $f(s)$ in the previous iteration $(t - 1)$. More precisely, $l = \frac{\beta}{2} \cdot f(s)$. We found that the best parameter values were $\gamma_2 = 0.5$ and $\gamma_1 \in [0.03, 0.06]$ (depending on the investigated image class).

To determine the quality of the reconstruction, RME (see Eq. 2) was used. The reconstructions were performed on random image patches of size 64×64 from each image class using 2, 3, 4, and 8 projections, assuming parallel beam geometry, with one pixel distance between the projection lines. The observed and the test dataset contained 150 ($t = 150$) and 10 images, respectively, from each image class. Due to the stochastic nature of SA, each reconstruction task was repeated 5 times, and the mean values of the results were utilized. The DART algorithm was used as a comparative study.

We observed that DART usually provided better results than SA, but in some cases, the differences were not significant.

Performing experimental analysis on software phantom images, we found the concept promising, especially in case of few available projections. However, we also deduced that in some cases, the expressive power of the standard LBP descriptor was not satisfactory. Fortunately, an advantage of this approach is that it can be improved by using other versions of LBP. Consequently, we decided to conduct experiments with the aforementioned LBP versions. Based on our observations, we opted to use SLBP in the subsequent investigations.

3.2 Results with SLBP

Previously, in the reconstruction formula (see Eq. 4), two priors were taken into account: the image must be smooth, and – at the same time – it must have a texture similar to the previously observed ones. Since SLBP is a more informative and more powerful texture descriptor than the basic LBP, we decided to omit the smoothness prior. Therefore, we turned back to the original cost function defined in Eq. 1,

$$C(\mathbf{X}) = ||\mathbf{Ax} - \mathbf{b}||_2 + \gamma \cdot \Phi(\mathbf{X}) \ \rightarrow \min \, ,$$

where $\Phi(\mathbf{X})$ provides the minimal Euclidean distance of the SLBP vector of the current image compared to all representative SLBP vectors of the sample images observed in advance. Formally,

$$\Phi(\mathbf{X}) = \min_t \{ ||SLBP(\mathbf{X}) - SLBP(\mathbf{X}_t)||_2 \} \, , \tag{8}$$

where $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_t$ are the t sample images from the same class as $\mathbf{X}$, known a priori. In our settings, we observed 150 random patches of size 64×64 as the prior dataset, meaning $t = 150$.

To optimize the cost function, we used Simulated Annealing again. All parameters of the SA algorithm were set manually in an empirical manner, similar to the previous case, except that the stopping criteria of the algorithm were to reach 800 000 iterations or to perform 25 000 iterations without improving the optimal result. We also found that the value of the scaling parameter γ in the cost function had to be determined individually for each image class.

As a first step, we generated four different synthetic image classes with simple textures of size 1024×1024. As prior information, we used the SLBP vectors of 150 randomly selected patches of size 64×64 from the same image classes. Then, we selected 10 random patches of size 64×64 from each class and attempted to reconstruct them from the horizontal and vertical projections complemented by the prior data using different γ weights. Finally, we attempted to reconstruct the images using the prior information only (i.e. omitting the projection term), to emphasize the importance of both terms. Based on the experiments, we can deduce the following. For image classes, which are uniquely determined by their horizontal and vertical projections, both Ryser's algorithm and DART provided exact reconstructions. Unfortunately, SA did not lead to perfect reconstruction in each case using only the projection term, but complementing it with the prior information could potentially improve the quality. In case of the other two classes, by increasing the γ weight of the SLBP term, the reconstructions could be forced to be more similar in texture to the observed patches. In contrast, DART and Ryser's algorithm were unable to reconstruct the images.

In the next experiment, we used the aforementioned software phantom images from 6 classes of different textures. We observed that increasing the γ weight to 2 could greatly improve the numerical quality of the reconstruction. We note that further incrementation of the γ weight did not result in additional enhancement. However, the visual results were not always satisfying, not even using DART or Ryser's algorithm. The reason for this is, the small patterns of the original image classes do not have local regular textures, even though the original images follow some regularities in the global sense. Furthermore, using solely SLBP term provided poorer results than using both terms. Therefore, we can deduce that both terms are needed for an acceptable reconstruction.

As a next step, we conducted tests on real images of 12 different classes from the Brodatz texture database [4]. The size of the original images varied between 256×256 and 1230×1120. First, we attempted to reconstruct 10-10 random image patches of size 64×64 from each class. We were faced again with the fact that these image classes, as observed previously, do not have regular textures in this size; consequently, we went on to test image patches of size 128×128. Unfortunately, there were image classes that were hard to reconstruct due to the above mentioned issue. In contrast, there were classes that provided satisfactory results. Increasing the γ parameter again provided better results.

Finally, tests were conducted on real images of various porous micro-structures [14]. The size of the original images (after applying image preprocessing methods) varied between 228×228 and 350×350. The SA algorithm was tested on 10-10 random image patches of size 64×64 from each class. Once again, we deduced that increasing the γ weight to 2 resulted in better final objective function value, i.e. both terms of the objective function were acceptable.

In contrast, RME was still high, but E_p was lower. The reconstruction did not become more accurate when γ was assigned a higher value than 2 or the projection term was omitted. We attempted to modify the size of the patches to 128×128. The results were similar to those of smaller in size.

Therefore, we found that SLBP as texture prior information can significantly improve the reconstruction quality, especially when only very few projections are available. Naturally, we cannot expect perfect reconstructions from merely two projections. Still, the improvement in quality when applying the method is indisputable.

The findings of this thesis point have been published in [28, 29].

4 Binary Tomography Based on Nonograms

Logic puzzles are very popular nowadays. One of them is the so-called Nonogram, where the aim is either to fill in or leave the cells blank of an image grid according to the numbers at the side of the grid to reveal a hidden picture. The numbers familiar from DT denote how many consecutive filled-in squares there are in the given rows or columns. These images are usually black and white, describing a binary image.

Inspired by Nonogram puzzles, in this thesis point, we introduced a novel prior to describe the expected structure of the reconstructed image: the number of strips (consecutive 1s of maximal length) in each row and column. We considered the problem of reconstructing a binary image from its row and column sums with the additional constraint that the number of strips is given for each row and column.

4.1 Problem Outline

Based on the BINARY TOMOGRAPHY and NONOGRAM problems, we defined the following intermediate problem.

Problem. STRIP CONSTRAINED BINARY TOMOGRAPHY (SCBT)
Input: Four non-negative integer vectors $H \in \mathbb{Z}^m$, $V \in \mathbb{Z}^n$, $SH \in \mathbb{Z}^m$, and $SV \in \mathbb{Z}^n$.
Output: A binary matrix of size $m \times n$, if it exists, with row sum vector H, column sum vector V, and in each row and column having the number of strips prescribed by SH and SV, respectively, where a 'strip' is defined as follows.

Definition 4.1. *Given a binary matrix* $\mathbf{X}$ *of size* $m \times n$*, a sequence of consecutive positions* $(i, j_s), (i, j_{s+1}), \ldots, (i, j_{s+l-1})$ *(where* l *is a positive integer, and* $1 \leq j_s \leq n$*) in the* i^{th} *row* $(1 \leq i \leq m)$ *form a* strip *if* $x_{i,j_s} = x_{i,j_{s+1}} =, \ldots, = x_{i,j_{s+l-1}} = 1$*, and* $x_{i,j_s-1} = x_{i,j_s+l} = 0$ *(if the latter two positions exist). The length of the strip is provided by* l*. Strips of columns can be defined analogously.*

Definition 4.2. *A switching component in a binary matrix* $\mathbf{X} \in \{0,1\}^{m \times n}$ *is a set of four positions* $(i, j), (i', j), (i, j'), (i', j')$ $(1 \leq i, i' \leq m,\ 1 \leq j, j' \leq n)$*, such that* $x_{ij} = x_{i'j'}$ *and* $x_{i'j} = x_{ij'} = 1 - x_{ij}$*.*

The problem is connected both to BT and Nonogram puzzles, and is in general NP-hard. We proposed the following.

Proposition 4.1. *The solution of the* SCBT *problem is not always uniquely determined.*

Proposition 4.2. *For certain inputs, the* SCBT *problem can have several solutions, such that one cannot be transformed into the other by a sequence of elementary switchings in a way that all switching produces a matrix that is a solution of the same* SCBT *problem as well.*

Proposition 4.3. *For an arbitrary solution of a fixed instance of the* SCBT *problem, it holds that the number of pairs of adjacent 1s in the* i^{th} *row* $(i = 1, \ldots, m)$ *is equal to* $h_i - sh_i$*, and the number of pairs of adjacent 1s in the* j^{th} *column* $(j = 1, \ldots, n)$ *is equal to* $v_j - sv_j$*.*

4.2 Proposed Methods

4.2.1 Constraint Satisfaction

In line with [13], SCBT is formulated as a Constraint Satisfaction Problem. Let $\mathbf{YH}^{m\times(n-1)}$ and $\mathbf{YV}^{(m-1)\times n}$ two unknown binary matrices, indicate the locations of adjacent 1s in the solution, i.e. the $\begin{bmatrix}1 & 1\end{bmatrix}$ and $\begin{bmatrix}1\\1\end{bmatrix}$ components of the solution image, respectively. The conditions of the model are as follows:

$$\textstyle\sum_{j=1}^{n} x_{ij} = h_i \quad (i = 1,\ldots,m)\,, \tag{9}$$

$$\textstyle\sum_{i=1}^{m} x_{ij} = v_j \quad (j = 1,\ldots,n)\,, \tag{10}$$

$$\textstyle\sum_{j=1}^{n-1} yh_{ij} = h_i - sh_i \quad (i = 1,\ldots,m)\,, \tag{11}$$

$$\textstyle\sum_{i=1}^{m-1} yv_{ij} = v_j - sv_j \quad (j = 1,\ldots,n)\,, \tag{12}$$

$$yh_{ij} \le x_{ij} \text{ and } yh_{ij} \le x_{i(j+1)} \quad (i = 1,\ldots,m; j = 1,\ldots,(n-1))\,, \tag{13}$$

$$yv_{ij} \le x_{ij} \text{ and } yv_{ij} \le x_{(i+1)j} \quad (i = 1,\ldots,(m-1); j = 1,\ldots,n)\,. \tag{14}$$

To implement the constraint satisfaction model, we used the **intlinprog** MATLAB built-in function, which is a mixed-integer linear programming solver.

4.2.2 Variants of Simulated Annealing

Although, the mixed integer programming-based method always provides an exact solution, it is difficult to apply in practice, as the SCBT problem is NP-complete. Therefore, as an alternative approach, we reformulated it into an optimization problem, where the aim was to minimize

$$C(\mathbf{X}) = ||H - H'||_2 + ||V - V'||_2 + ||SH - SH'||_2 + ||SV - SV'||_2\,, \tag{15}$$

where vectors H, V, SH, SV are given as input; and H', V', SH', SV' are the corresponding vectors belonging to the current solution image $\mathbf{X}$. We suggested SA again to solve the problem. The effectiveness of the SA method was influenced by several factors, among which initialization strategy is one of the key issues. Therefore, different initializing strategies were proposed as follows. **BasicSA** starts out with a completely random binary image, **FixpixelSA** begins with a random binary image with a fixed number of object points, **RyserSA** uses Ryser's algorithm to identify the initial solution and elementary switchings to refine it, and **CurveballSA** is based on the idea of the Curveball method [9, 10, 26]. A local update of the number of strips in each iteration was also presented, which reduced the running time of the methods significantly.

4.3 Results

For the evaluation, three different datasets were used: random images of different densities (binary images of sizes 3×3, $4 \times 4, \ldots, 256 \times 256$, containing $0\%, 10\%, \ldots, 100\%$ randomly selected object pixels); real Nonogram puzzles [34]; and so-called TomoPhantom images [33]. We found that the exact method is seldom applicable in practice, since it requires high running time even for small matrices, and cannot be used for larger matrices. In contrast, the stochastic approach is much more suitable for practical purposes, even though on occasion, it merely provides approximate solutions.

We found that although **FixpixelSA** is faster than the other methods, it does not yield the best results. **RyserSA** often finds a perfect solution, especially in case of dense and/or small matrices, but sometimes gets stuck in a local minimum. **CurveballSA** performs better for real images than **RyserSA** and is able to avoid getting stuck in local minima. Unfortunately, **CurveballSA** requires much higher running time.

We also found that the pixel error is high even if the objective function value is relatively small. This is due to the presence of switching components, which can yield many different solutions for the same input data. The reason for this is, a structure close to the uniform random one provides more freedom for switching components to occur, and the solution may be highly ambiguous. If a large portion of the pixels is fixed to belong to the foreground/background, this happens less often. This theory was also confirmed by the calculation of the number of switching components for different matrix densities of size $n \times n$. It turned out that there are, in fact, a large number of switching components for every image class, even for small sizes.

The results obtained in this chapter provide insight into binary tomographic reconstruction with structural priors.

The findings of this thesis point have been published in [2, 27, 32].

5 Global and Local Quadrant-Convexity

Shape representation is a relevant topic in digital image analysis, due to, for instance, object recognition and classification issues. Suitable approaches for handling these problems include designing new shape descriptors as well as measures for descriptors sensitive to distinguishing shapes, while also robust to noise. Some methods provide a unified approach that can be applied to determine a variety of shape measures, but more often they are specific to a single aspect of a shape.

Remark 5.1. *Note that in image analysis as well as in this thesis, the technical term 'measure' is used as a synonym for figure-of-merit rather than referring to the measure concept of mathematical analysis. Therefore, shape measures often do not satisfy, for example, monotonicity.*

Over the years, a variety of measures for descriptors based on convexity has been developed, one of them is the so-called Quadrant-convexity (Q-convexity for short) [7, 8].

In this thesis point, we proposed a Quadrant-convexity-based measure to define a scalar shape descriptor based on spatial relative positions of points that permits a definition of enlacement and interlacement between two objects. The term *enlacement* between F and G indicates that object G is somehow inside object F (i.e. the reference), while the term *interlacement* is defined a mutual (symmetrical) enlacement between the two objects.

5.1 Global Quadrant-Convexity

Consider a two-dimensional object F represented by a non-empty *lattice set*, i.e. a finite subset of $\mathbb{Z}^2$, or equivalently, a function $f : \mathbb{Z}^2 \to \{0,1\}$. Let $\mathcal{R}$ be the smallest discrete rectangle containing F and suppose it is of size $m \times n$. Without loss of generality, we can assume that the bottom-left corner of $\mathcal{R}$ is in the origin $(0,0)$, i.e. $\mathcal{R} = \{0,\dots,m-1\} \times \{0,\dots,n-1\}$. Alternatively, F can be viewed as a binary image, i.e. as a union of white unit squares (foreground pixels) corresponding to points of F, and $\mathcal{R} \setminus F$ being the union of black unit squares (background pixels).

Each position (i,j) in rectangle $\mathcal{R}$, along with the horizontal and vertical directions, determines the following four quadrants:

$$
\begin{aligned}
Z_0(i,j) &= \{(l,k) \in \mathcal{R} : \ 0 \le l \le i, \ 0 \le k \le j\} \ , \\
Z_1(i,j) &= \{(l,k) \in \mathcal{R} : \ i \le l \le m-1, \ 0 \le k \le j\} \ , \\
Z_2(i,j) &= \{(l,k) \in \mathcal{R} : \ i \le l \le m-1, \ j \le k \le n-1\} \ , \\
Z_3(i,j) &= \{(l,k) \in \mathcal{R} : \ 0 \le l \le i, \ j \le k \le n-1\} \ .
\end{aligned}
$$

Let us denote the number of object points (foreground pixels) of F in $Z_p(i,j)$ by $n_p(i,j)$, for $p = 0,\ldots,3$, meaning

$$n_p(i,j) = |Z_p^w(i,j) \cap F| \ (p = 0,\ldots,3)\ . \tag{16}$$

Definition 5.1. *Lattice set F is Q-convex, if for each (i,j) $\quad (n_0(i,j) > 0 \wedge n_1(i,j) > 0 \wedge n_2(i,j) > 0 \wedge n_3(i,j) > 0)$ implies $(i,j) \in F$.*

If F is not Q-convex, then a position (i,j) exists, which violates the Q-convexity property, i.e. $n_p(i,j) > 0$ for all $p = 0,\ldots,3$ and $(i,j) \notin F$. A lattice set that is not Q-convex is called *Q-concave*. We defined the Q-concavity measure of F as the sum of the contributions of non-Q-convexity for each point in $\mathcal{R}$. Formally,

$$\varphi_F(i,j) = n_0(i,j)n_1(i,j)n_2(i,j)n_3(i,j)(1 - f(i,j))\ , \tag{17}$$

where (i,j) is an arbitrary point of $\mathcal{R}$, and $f(i,j) = 1$ if the point in position (i,j) belongs to the object, otherwise $f(i,j) = 0$. Moreover, let

$$\varphi_F = \varphi_F(F) = \sum_{(i,j)\in\mathcal{R}} \varphi_F(i,j)\ . \tag{18}$$

Remark 5.2. *If $f(i,j) = 1$, then $\varphi_F(i,j) = 0$. Moreover, if $f(i,j) = 0$ and $n_p(i,j) = 0$ exist, then $\varphi_F(i,j) = 0$. Consequently, F is Q-convex if and only if $\varphi_F = 0$.*

5.1.1 Obtaining Enlacement Descriptors by Normalization

In order to measure the degree of Q-concavity, or equivalently the degree of landscape enlacement for a given object F, we normalized φ so that it ranges in $[0,1]$. We proposed two possible normalizations gained by normalizing each contribution. A global one is obtained as follows [11]:

$$\mathcal{E}_F^{(1)}(i,j) = \frac{\varphi_F(i,j)}{\max_{(i',j')\in\mathcal{R}}\ \varphi_F(i',j')}\ . \tag{19}$$

Second (local) one is based on [5]. There, the authors proved

Proposition 5.1. *Let $f(i,j) = 0$, and h_i^F and v_j^F be the i^{th} column and j^{th} row sums, respectively. Then, $\varphi_F(i,j) \leq ((|F| + h_i^F + v_j^F)/4)^4$.*

As a consequence, we are able to normalize each contribution as

$$\mathcal{E}_F^{(2)}(i,j) = \frac{\varphi_F(i,j)}{((|F| + h_i^F + v_j^F)/4)^4}\ . \tag{20}$$

In order to obtain a normalization for the global enlacement for both Eqs. 19 and 20, we sum up each single contribution and then divide by the number of nonzero contributions.

Definition 5.2. *For a given binary image F, its global enlacement landscape $\mathcal{E}_F^{(\cdot)}$ is defined by*

$$\mathcal{E}_F^{(\cdot)} = \sum_{(i,j)\in\bar{F}} \frac{\mathcal{E}_F^{(\cdot)}(i,j)}{|\bar{F}|}\ ,$$

where $\bar{F}$ denotes the subset of (landscape) points in $\mathcal{R} \setminus F$ for which the contribution is not null.

5.1.2 Object Enlacement and Interlacement

Up to this point, we have defined shape measure φ_F based on the concept of Q-convexity, which provides a quantitative measure for examining relationships with a reference object. We modified this into a spatial relationship between two objects. Let F and G be two objects. The idea was to determine how many occurrences of points of G are somehow between points of F. This is achieved by restricting the points in $\mathcal{R} \setminus F$, while also taking the points in G into account. Therefore,

$$\varphi_{FG}(i,j) = \begin{cases} \varphi_F(i,j) & \text{if } (i,j) \in G \\ 0 & \text{otherwise} \end{cases} . \tag{21}$$

Note that if $G = \mathcal{R} \setminus F$, then trivially $\varphi_{FG}(i,j) = \varphi_F(i,j)$. The enlacement descriptors of G by F are thus

$$\mathcal{E}^{(1)}_{FG}(i,j) = \frac{\varphi_{FG}(i,j)}{\max_{(i,j)\in G}\ \varphi_{FG}(i,j)} , \tag{22}$$

and

$$\mathcal{E}^{(2)}_{FG}(i,j) = \frac{\varphi_{FG}(i,j)}{((|F| + h^F_i + v^F_j)/4)^4} . \tag{23}$$

To clarify, note that, by definition, $\varphi_{FG}(i,j) = 0$ if $(i,j) \notin \bar{F}$, thus the maximum of $\varphi_{FG}(i,j)$ in G is equal to the maximum in $G \cap \bar{F}$.

Definition 5.3. *Let F and G be two objects. The enlacement of G by F is*

$$\mathcal{E}^{(\cdot)}_{FG} = \sum_{(i,j)\in G} \frac{\mathcal{E}^{(\cdot)}_{FG}(i,j)}{|G \cap \bar{F}|} .$$

Undoubtedly, the enlacement of two objects is an asymmetrical relation, so the enlacement of G by F and the enlacement of F by G provide different 'views'. Combining by their harmonic mean results in a symmetrical relation, which is the description of mutual enlacement (in other words, interlacement).

Definition 5.4. *Let F and G be two objects. The interlacement of F and G is*

$$\mathcal{I}^{(\cdot)}_{FG} = \frac{2\mathcal{E}^{(\cdot)}_{FG}\mathcal{E}^{(\cdot)}_{GF}}{\mathcal{E}^{(\cdot)}_{FG} + \mathcal{E}^{(\cdot)}_{GF}} , \tag{24}$$

where $\mathcal{E}^{(\cdot)}_{FG}$ and $\mathcal{E}^{(\cdot)}_{GF}$ are the enlacement of G by F and the enlacement of F by G, respectively.

The measures can be implemented efficiently in linear time in the size of the image. By definition, $Z_0(l,k) \subseteq Z_0(i,j)$ if $l \leq i$ and $k \leq j$, and hence $n_0(l,k) \leq n_0(i,j)$ with $l \leq i$ and $k \leq j$. Analogous relations hold for $Z_1,\ Z_2,\ Z_3$ and for $n_1,\ n_2,\ n_3$ accordingly. Exploiting this property and proceeding line by line, we are able to count the number of points in F for $Z_p(i,j)$, for each (i,j) in linear time, and store them in a matrix for any $p = \{0,1,2,3\}$. Then, $\varphi(i,j)$ can be computed in constant time for any (i,j). Normalization is straightforward.

5.1.3 Experiments

In the following experiments, we used public datasets of fundus photographs of the retina for classification issues. We opted to compare our descriptor with its counterpart in [12] since they model a similar idea based on a quantitative concept of convexity. The main difference is that we provide a fully 2D approach, whereas the directional enlacement landscape in [12] is one-dimensional. The CHASEDB1 [16] dataset is composed of 20 binary images with centered optic disks, while the DRIVE [25] dataset contains 20 images, where the optic disk is shifted from the center. Following the strategy of [12], we gradually added different types of random noise (which can also be interpreted as increasingly stronger segmentation errors) to size 1000×1000 images. Gaussian and Speckle noises were added with 10 increasing variances $\sigma^2 \in [0,2]$, while Salt & Pepper noise was added with 10 increasing amounts in $[0, 0.1]$. Then, we attempted to divide the images into two classes (CHASEDB1 and DRIVE) based on their interlacement values by the 5-nearest neighbor classifier (5NN) with inverse Euclidean distance. For the implementation, we used the WEKA Toolbox [15] and leave-one-out cross-validation to evaluate accuracy.

For a more complex classification problem, we used the High-Resolution Fundus (HRF) dataset [22] composed of 45 images of fundus: 15 healthy, 15 with glaucoma symptoms, and 15 with diabetic retinopathy symptoms. Using the same classifier as before, we attempted to separate the 15 healthy images from the 30 diseased cases.

To conclude the results we obtained, our descriptor provided comparable (and in some cases even better) results than the descriptor from [12]. This is especially promising, since – being two-dimensional based – our descriptor only uses two directions (four quadrants), whereas that of [12] – being one-dimensional based – uses 180 directions, also taking more time to compute.

5.2 Local Quadrant-Concavity Histograms for Image Classification

Histograms built on local features are able to provide much richer information on the geometry and structural properties of the shape than single scalar descriptors do. Therefore, we extended the global Q-concavity measure to histograms collecting Q-concavity values calculated under all possible positions of an image window of predefined size.

The aim was refining Eqs. 17 and 18 to achieve a local and thus more informative measure of Q-concavity. To this end, consider a $(2w+1) \times (2w+1)$ window ($w \in \mathbb{Z}^+$). The quadrants around (i,j) restricted to this window size can then be defined as

$$\begin{aligned}
Z_0^w(i,j) &= \{(l,k) \in \mathcal{R} : \ (i-w) \le l \le i, \ (j-w) \le k \le j\} \, , \\
Z_1^w(i,j) &= \{(l,k) \in \mathcal{R} : \ i \le l \le (i+w), \ (j-w) \le k \le j\} \, , \\
Z_2^w(i,j) &= \{(l,k) \in \mathcal{R} : \ i \le l \le (i+w), \ j \le k \le (j+w)\} \, , \\
Z_3^w(i,j) &= \{(l,k) \in \mathcal{R} : \ (i-w) \le l \le i, \ j \le k \le (j+w)\} \, ,
\end{aligned}$$

where $\mathcal{R} = \{0, \ldots, m-1\} \times \{0, \ldots, n-1\}$ and $m \times n$ is the size of the image. The number of object points in Z_p^w is

$$n_p^w(i,j) = |Z_p^w(i,j) \cap F| \ \ (p = 0, \ldots, 3) \, , \tag{25}$$

and the local Q-concavity contribution at point (i,j) is

$$\varphi_F^w(i,j) = n_0^w(i,j) n_1^w(i,j) n_2^w(i,j) n_3^w(i,j)(1 - f(i,j)) \, . \tag{26}$$

Finally, the *local Q-concavity histogram* of F (LQH) is a mapping $hist_{F,w} : \mathbb{Z} \to \mathbb{Z}$ which can be defined in various ways. The first approach focuses on the background points in $\mathcal{R}$. In this case,

$$hist_{F,w,bg}(r) = |(i,j) \in \mathcal{R} \setminus F \ : \ \varphi_F^w(i,j) = r| \, , \tag{27}$$

i.e. we take each background point, calculate its local Q-concavity value, and increase the value of the corresponding bin by 1. Alternatively, we can take all points in $\mathcal{R}$ into account. Then,

$$hist_{F,w,all}(r) = |(i,j) \in \mathcal{R} \ : \ \varphi_F^w(i,j) = r| \, . \tag{28}$$

A third approach focuses exclusively on the points that truly violate local Q-convexity (approach 'nonzero'). Then we get

$$hist_{F,w,nz}(r) = |(i,j) \in \mathcal{R} \ : \ \varphi_F^w(i,j) = r > 0| \, , \tag{29}$$

i.e. in this case, the 0 bin is omitted.

We compared the above mentioned three histogram variants on three different datasets: 1) the one – also used previously – with 6 software phantom image classes; 2) TomoPhantom images [33]; 3) and real images from the Brodatz texture database [4]. Based on the intraclass and interclass Euclidean distances, the *background* approach proved to be the best one. Therefore, we preferred to use the histogram based on the background points.

5.2.1 Experiment with the Retina Dataset

As a case study, in order to investigate the classification power of LQH, the previous experiment was repeated with exactly the same dataset, settings, and classifier (CHASEDB1 and DRIVE dataset, different types of added random noise, etc.). In this case, the question was whether the local Q-concavity histogram could outperform the global Q-concavity measure. We attempted to classify the images into two categories (CHASEDB1 and DRIVE) based on their local Q-concavity histograms using the background points (see Eq. 27), by the 5-nearest neighbor classifier with inverse Euclidean distance. To avoid overfitting, we used leave-one-out cross-validation.

As we attempted to develop a fully locally computable descriptor, we opted for the local normalization form of Eq. 23. To compare the performance of the local descriptor to the global Q-concavity measure, the formula of local Q-concavity contribution at point (i, j) (provided in Eq. 26) was modified to

$$\varphi^w_{F,norm}(i,j) = \frac{n^w_0(i,j)n^w_1(i,j)n^w_2(i,j)n^w_3(i,j)(1-f(i,j))}{(\alpha^w(i,j)+h^w(i,j)+v^w(i,j))^4}\,, \tag{30}$$

where

$$\alpha^w(i,j) = \sum_{l=i-w}^{i+w}\sum_{k=j-w}^{j+w} f(l,k) \tag{31}$$

denotes the number of points belonging to F in the current window,

$$v^w(i,j) = \sum_{k=j-w}^{j+w} f(i,k) \tag{32}$$

is the number of points of F in the i^{th} column, and

$$h^w(i,j) = \sum_{l=i-w}^{i+w} f(l,j) \tag{33}$$

is the number of points of F in the j^{th} row, restricted to the window, respectively. Finally, the *local Q-concavity histogram (LQH)* of F is a mapping $hist_{F,w} : \mathbb{R} \to \mathbb{Z}$ with

$$r \mapsto |(i,j) \in \mathcal{R} \setminus F \;:\; \varphi^w_{F,norm}(i,j) = r|\,. \tag{34}$$

Since $\varphi^w_{F,norm}(i,j)$ is not an integer anymore, the bins of the histogram are not straightforward to index. To overcome this issue, we carry out quantization. Let q be the number of quantization levels, i.e. the number of bins to occur in the histogram. Then we obtain the bin indices by the following formula:

$$bin\_index = \left\lfloor \varphi^w_{F,norm}(i,j) \Big/ \frac{\max \varphi^w_{F,norm}}{q-1} \right\rfloor\,. \tag{35}$$

Tests were made with 2 and 10 quantization levels ($q = 2$ and $q = 10$). We observed that in the presence of Gaussian noise, in nearly all instances, the local, histogram-based method performs significantly better than the one based on the global Q-concavity measure, even better than SLBP, both for $q = 2$ and $q = 10$. For Salt & Pepper noise, SLBP is best; however, especially for $q = 10$, the local method is almost as advantageous as SLBP. We stress that SLBP is a 256-dimensional descriptor, whereas the histogram-based method only uses 10-dimensional vectors (when $q = 10$). In the case of Speckle noise, the global Q-concavity measure seems to be the best choice, although in some cases, and especially for smaller window sizes, the local approach as well as SLBP ensures comparable accuracy.

Next, we turned to the question of which window size and quantization level could ensure the highest classification accuracy. We presented a strategy to determine the appropriate window size and quantization level to calculate LQH when these are used in classification, i.e. in separating images of different classes.

For the Brodatz dataset and the software phantom image classes, 7×7 proved to be a suitable window size, but 5×5 and 9×9 could be also viable options. Regarding the quantization levels, even 10-16 bins can provide relatively satisfactory classification accuracy for these window sizes. Therefore, we opted to use 10 bins.

5.3 Local Q-Concavity Histograms for Binary Image Reconstruction

Finally, we examined how LQH improves the quality of binary tomographic reconstruction using two projections.

This time, the regularization utilized in solving Eq. 1 was based on the concept of Q-convexity. Assuming that the image to be reconstructed belongs to a certain class, representatives of this class can be selected and their LQH vectors can be calculated. In the reconstruction process, the LQH vectors were taken into account as prior information. In this case, we also worked with two projections: the horizontal (H) and vertical (V) ones. Hence, Eq. 1 became

$$C(\mathbf{X}) = ||H - H'||_2 + ||V - V'||_2 + || + \gamma \cdot \Phi(\mathbf{X}) \ \rightarrow \min \, , \tag{36}$$

where H and V are input vectors (row and column sums of the image, respectively); and H' and V' are the corresponding vectors belonging to the current solution image $\mathbf{X}$. $\Phi(\mathbf{X})$ measures the Euclidean distance between the LQH vector of image $\mathbf{X}$ and all representative LQH vectors, and ultimately takes the minimum. Formally,

$$\Phi(\mathbf{X}) = \min_t \{ ||LQH(\mathbf{X}) - LQH(\mathbf{X}_t)||_2 \} \, , \tag{37}$$

where $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_t$ are the t sample images from the same class as $\mathbf{X}$, known a priori.

For calculating LQH (with normalization and quantization), we used the implementation based on Eqs. 30 - 35. To optimize the cost function (see Eq. 36) we used SA, again.

5.3.1 Fast LQH Calculation

Calculating the LQH vectors from scratch in each iteration of SA is quite time-consuming. Furthermore, fast LQH calculation would also be beneficial when an abundance of such descriptors must be processed, e.g. in classification issues. Three approaches were presented to speed up this process.

The first one is based on integral images and dynamic programming (DP). The idea is to decrease the computational complexity by calculating LQH for one pixel only; whereas for the other pixels, only the differences are taken into consideration. We deduced that for smaller window sizes, DP required higher running time than the original method due to the fact that DP calculates and stores LQHs for all pixels, while the original method considers pixels with 0 values only. On the other hand, for larger window sizes, running time was significantly reduced. Therefore, this approach is useful if larger windows are required for practical reasons.

The second idea was, instead of calculating LQH for all image pixels, only a predefined amount (e.g. 10%, 20%, etc.) is examined. Pixels are randomly selected. The resulting LQHs gained by random sampling were similar to the fully evaluated ones, even when only 20% of the pixels were examined. With this strategy, running time can be reduced drastically. Due to the fact that this approach is largely random in itself, unfortunately, it is not applicable with the SA algorithm. Still, the advantage of this approach is indisputable. When a multitude of images need to be processed from the viewpoint of Q-convexity, this method may save a considerable amount of time.

In the third approach, we took advantage of the observation that the LQH vector does not need to be entirely recalculated in each iteration of SA. Since only one pixel is modified per iteration, we are able to update the LQH vector by focusing only on those parts of the image where the modified pixel has a potential impact. Using this method, running time can be reduced significantly. Furthermore, this strategy is fully in accordance with the concept of SA, which is why this method is selected to speed up reconstruction.

5.3.2 Results

The tests were conducted again on two different datasets: software phantom image classes and real images from the Brodatz texture database [4].

All parameters of the SA algorithm were set manually, in an empirical manner. Each time, the method started out with a random binary image. The stopping criteria of the algorithm were to reach 1 000 000 iterations or to perform a given number of iterations without improving the optimal result (parameter ξ). The initial temperature was set to $T_1 = 350$, while the cooling schedule was controlled by $\alpha = 0.99$. The SA algorithm was allowed to remain the same temperature for some iterations; in our case, this was fine-tuned by $\beta = 0.035$. Selecting a neighbor meant randomly selecting and inverting an image pixel. Based on empirical investigations, we found that a satisfactory parameter setting is $\gamma \in [0.4, 0.7]$, $\xi \in [3000, 6000]$ and $u \times u \in \{3 \times 3, 5 \times 5, \ldots, 13 \times 13\}$ for the window size, depending on image class. The quantization level was $q = 10$. We observed 150 random patches of size 128×128 as the prior dataset, meaning $t = 150$ (see Eq. 37).

We found that using LQH as prior information significantly improves the reconstruction quality from the visible point of view. Even though the reconstructed image did not fully resemble the structure of the original image, the low final values of the terms in the cost function explain that this phenomenon is caused by the inherent loss of information in the projections, not by the optimization process. Indeed, the reconstruction from the horizontal and vertical projections is an extremely ill-posed problem.

The findings of this thesis point have been published in [6, 31] and are currently under review to be published as a journal article [30].

6 Summary of the Author's Contributions

The findings of this dissertation can be divided into three thesis groups. Table 1 provides an overview of the results in relation to publications of the author.

In the first thesis group, I studied how texture information improves the quality of binary reconstruction using very few projections. These results were published in conference proceedings [29], and in a journal article [28].

I/1. I examined and integrated the basic LBP prior into the reconstruction process and solved the corresponding optimization problem by SA. I tested the algorithm on different datasets and evaluated the results.

I/2. I examined and compared the different LBP variants, and found that SLBP is the most useful in this case. I defined the proper objective function and integrated the SLBP prior into the reconstruction process and solved the corresponding optimization problem by SA. I tested the algorithm on different datasets and evaluated the results. I also defined alternative quality measurements for the results to be comparable.

In the second thesis group, I introduced a novel prior, based on Nonogram puzzles, to describe the expected structure of the reconstructed image: the number of strips (consecutive 1s of maximal length) in each row and column. These results were published in three conference proceedings [2, 27, 32].

II/1. I first solved the reconstruction problem with a deterministic integer programming approach; then proposed the stochastic method based on SA to resolve the issue; and finally, compared the results.

II/2. I defined different strategies for the SA algorithm, implemented and compared them; tested the algorithm on different datasets, and evaluated the results. I also examined how switching components affect the reconstruction process.

The third thesis group is based on global and local Quadrant-convexity. These results were published in two conference proceedings [6, 31] and are currently under review to be published as a journal article [30].

III/1. I tested and analyzed the classification accuracy of the global Q-convexity-based method.

III/2. I suggested three different approaches for calculating local Q-concavity histogram (LQH), then formalized the novel proposals and their corresponding equations. I tested the classification accuracy of the suggested methods and evaluated the results.

III/3. I adapted the appropriate LQH approach into the reconstruction process, implemented it, and tested on diverse datasets. I defined various approaches for speedy calculation of LQH. Finally, I evaluated the results.

	[28]	[29]	[2]	[32]	[27]	[6]	[31]	[30]
I/1.	•							
I/2.		•						
II/1.			•					
II/2.				•	•			
III/1.						•		
III/2.							•	
III/3.								•

Table 1: Overview of thesis points in relation to the author's publications.

References

[1] T. Ahonen and M. Pietikäinen. “Soft histograms for local binary patterns”. In: *Proceedings of the Finnish signal processing symposium, FINSIG*. Vol. 5. 9. 2007, pp. 1–4.

[2] P. Balázs and J. Szűcs. “Reconstruction of Binary Images with Fixed Number of Strips”. In: *International Conference Image Analysis and Recognition*. Springer. 2018, pp. 11–19.

[3] K. J. Batenburg and J. Sijbers. “DART: a practical reconstruction algorithm for discrete tomography”. In: *IEEE Transactions on Image Processing* 20.9 (2011), pp. 2542–2553.

[4] P. Brodatz. *Textures: a photographic album for artists and designers*. Dover Pubns, 1966.

[5] S. Brunetti, P. Balázs, and P. Bodnár. “Extension of a one-dimensional convexity measure to two dimensions”. In: *International Conference on Discrete Geometry for Computer Imagery*. Springer. 2017, pp. 219–230.

[6] S. Brunetti, P. Balázs, P. Bodnár, and J. Szűcs. “A Spatial Convexity Descriptor for Object Enlacement”. In: *International Conference on Discrete Geometry for Computer Imagery*. Springer. 2019, pp. 330–342.

[7] S. Brunetti and A. Daurat. “An algorithm reconstructing convex lattice sets”. In: *Theoretical Computer Science* 304.1-3 (2003), pp. 35–57.

[8] S. Brunetti and A. Daurat. “Reconstruction of convex lattice sets from tomographic projections in quartic time”. In: *Theoretical Computer Science* 406.1-2 (2008), pp. 55–62.

[9] C. J. Carstens. “Proof of uniform sampling of binary matrices with fixed row sums and column sums for the fast curveball algorithm”. In: *Physical Review E* 91.4: 042812 (2015).

[10] C. J. Carstens, A. Berger, and G. Strona. “A unifying framework for fast randomization of ecological networks with fixed (node) degrees”. In: *MethodsX* 5 (2018), pp. 773–780.

[11] M. Clément, C. Kurtz, and L. Wendling. “Fuzzy directional enlacement landscapes”. In: *International Conference on Discrete Geometry for Computer Imagery*. Springer. 2017, pp. 171–182.

[12] M. Clément, A. Poulenard, C. Kurtz, and L. Wendling. “Directional Enlacement Histograms for the Description of Complex Spatial Configurations between Objects”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.12 (2017), pp. 2366–2380.

[13] G. Dahl and T. Flatberg. “Optimization and reconstruction of hv-convex (0, 1)-matrices”. In: *Discrete applied mathematics* 151.1-3 (2005), pp. 93–105.

[14] *Digital Rocks Portal*. `https://www.digitalrocksportal.org`.

[15] F. Eibe, M. A. Hall, and I. H. Witten. “The WEKA workbench. Online appendix for data mining: practical machine learning tools and techniques”. In: *Morgan Kaufmann*. 2016.

[16] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman. “An ensemble classification-based approach applied to retinal blood vessel segmentation”. In: *IEEE Transactions on Biomedical Engineering* 59.9 (2012), pp. 2538–2548.

[17] D. Iakovidis, E. Keramidas, and D. Maroulis. “Fuzzy local binary patterns for ultrasound texture characterization”. In: *Image Analysis and Recognition* (2008), pp. 750–759.

[18] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. “Optimization by simulated annealing”. In: *Science* 220.4598 (1983), pp. 671–680.

[19] A. Kuba, G. T. Herman, S. Matej, and A. Todd-Pokropek. “Medical applications of discrete tomography”. In: *Discrete Mathematical Problems with Medical Applications, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, AMS* 55 (2000), pp. 195–208.

[20] G. Kylberg and I.-M. Sintorn. “Evaluation of noise robustness for local binary pattern descriptors in texture classification”. In: *EURASIP Journal on Image and Video Processing* 2013.1: 17 (2013).

[21] R. Mehta and K. Egiazarian. “Dominant rotated local binary patterns (DRLBP) for texture classification”. In: *Pattern Recognition Letters* 71 (2016), pp. 16–22.

[22] J. Odstrcilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Cernosek, O. Svoboda, and E. Angelopoulou. “Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database”. In: *IET Image Processing* 7.4 (2013), pp. 373–383.

[23] T. Ojala, M. Pietikäinen, and D. Harwood. “A comparative study of texture measures with classification based on featured distributions”. In: *Pattern Recognition* 29.1 (1996), pp. 51–59.

[24] H. J. Ryser. “Combinatorial properties of matrices of zeros and ones”. In: *Canadian Journal of Mathematics* 9 (1957), pp. 371–377.

[25] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken. “Ridge-based vessel segmentation in color images of the retina”. In: *IEEE transactions on medical imaging* 23.4 (2004), pp. 501–509.

[26] G. Strona, D. Nappo, F. Boccacci, S. Fattorini, and J. San-Miguel-Ayanz. “A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals”. In: *Nature communications* 5.1 (2014), pp. 1–9.

[27] J. Szűcs and P. Balázs. “An Improved Simulated Annealing Approach for Reconstructing Binary Images with Fixed Number of Strips”. In: *International Conference on Image Analysis and Recognition*. Springer. 2019, pp. 174–185.

[28] J. Szűcs and P. Balázs. “Binary Image Reconstruction Using Local Binary Pattern Priors”. In: *International Journal of Circuits, Systems and Signal Processing* 11 (2017), pp. 296–299.

[29] J. Szűcs and P. Balázs. “Binary Tomography Using Variants of Local Binary Patterns as Texture Priors”. In: *International Conference on Computer Analysis of Images and Patterns*. Springer. 2019, pp. 141–154.

[30] J. Szűcs and P. Balázs. “Local Q-Concavity Histograms for Binary Image Classification and Reconstruction”. In: ***Submitted** to The Visual Computer* (2021).

[31] J. Szűcs and P. Balázs. “Local Q-Convexity Histograms for Shape Analysis”. In: *International Workshop on Combinatorial Image Analysis*. Springer. 2020, pp. 245–257.

[32] J. Szűcs and P. Balázs. “Variants of Simulated Annealing for Strip Constrained Binary Tomography”. In: *Lecture Notes in Computer Science 10986*. 2019, pp. 1–11.

[33] L. G. Varga, L. G. Nyúl, A. Nagy, and P. Balázs. “Local and global uncertainty in binary tomographic reconstruction”. In: *Computer Vision and Image Understanding* 129 (2014). Special section: Advances in Discrete Geometry for Computer Imagery, pp. 52–62.

[34] J. Wolter. *Web Paint-by-Number*. `https://webpbn.com/`.

Publications of the Author

Journal publications

- Judit Szűcs, Péter Balázs: Binary Image Reconstruction Using Local Binary Pattern Priors; In: International Journal of Circuits, Systems and Signal Processing (NAUN) 11, pp. 296-299. (2017). ISSN: 1998-4464. SJR: Q4

- Judit Szűcs, Péter Balázs: Local Q-Concavity Histograms for Binary Image Classification and Reconstruction; *In: **Submitted** to The Visual Computer* SJR: Q2

Full papers in conference proceedings

- Péter Balázs, Judit Szűcs: Reconstruction of Binary Images with Fixed Number of Strips; In: Image Analysis and Recognition. Ed. by Aurélio Campilho, Fakhri Karray, and Bart ter Haar Romeny. Vol. 10882. Cham: Springer International Publishing, 2018, pp. 11-19.
(ISBN: 978-3-319-93000-8)
DOI: 10.1007/978-3-319-93000-8_2 CORE: C

- Judit Szűcs, Péter Balázs: Variants of Simulated Annealing for Strip Constrained Binary Tomography; In: Computational Modeling of Objects Presented in Images. Fundamentals, Methods, and Applications. Ed. by Reneta P. Barneva et al. Vol. 10986. Cham: Springer International Publishing, 2019, pp. 82-92.
(ISBN: 978-3-030-20805-9)
DOI: 10.1007/978-3-030-20805-9_8

- Sara Brunetti, Péter Balázs, Péter Bodnár, Judit Szűcs: A Spatial Convexity Descriptor for Object Enlacement; In: Proceedings of International Conference on Discrete Geometry for Computer Imagery. Ed. by Michel Couprie et al. Vol. 11414. Cham: Springer International Publishing, 2019, pp. 330-342.
(ISBN: 978-3-030-20805-9)
DOI: 10.1007/978-3-030-14085-4_26 CORE: C

- Judit Szűcs, Péter Balázs: An Improved Simulated Annealing Approach for Reconstructing Binary Images with Fixed Number of Strips; In: Image Analysis and Recognition. Ed. by Fakhri Karray, Aurélio Campilho, and Alfred Yu. Vol. 11662. Cham: Springer International Publishing, 2019, pp. 174-185.
(ISBN: 978-3-030-27202-9)
DOI: 10.1007/978-3-030-27202-9_15 CORE: C

- Judit Szűcs, Péter Balázs: Binary Tomography Using Variants of Local Binary Patterns as Texture Priors; In: Computer Analysis of Images and Patterns. Ed. by Mario Vento and Gennaro Percannella. Vol. 11678. Cham: Springer International Publishing, 2019, pp. 141-154.
(ISBN: 978-3-030-29891-3)
DOI: 10.1007/978-3-030-29888-3_12 CORE: B

- Judit Szűcs, Péter Balázs: Local Q-Convexity Histograms for Shape Analysis; In: Combinatorial Image Analysis. Ed. by Tibor Lukić et al. Vol. 12148. Cham: Springer International Publishing, 2020, pp. 245–257.
(ISBN: 978-3-030-51002-2)
DOI: 10.1007/978-3-030-51002-2_18 CORE: C

Other publications

- Balázs Péter, Lékó Gábor, Ozsvár Zoltán, Petrovszki Gábor, Szűcs Judit, Varga László: 3D ipari CT szkenner informatikai rendszerének tervezése és fejlesztése; In: Képfeldolgozók és Alakfelismerők Társaságának 11. konferenciája, Szováta, Romania, 24-27 January, 2017, Paper Nr. 58.
- Judit Szűcs, Péter Balázs: Binary Image Reconstruction based on Structural Information; In: 4th Winter School of PhD Students in Informatics and Mathematics, Pécs, Hungary, 24-26 February, 2017, pp. 33-34.
(ISBN: 978-615-5586-12-5)
- Szűcs Judit: Képrekonstrukció tanult strukturális információval; In: XXXIII. Országos Tudományos Diákköri Konferencia, Informatika Tudományi Szekció, Számítógépes jelfeldolgozás, alakfelismerés tagozat, Veszprém, Hungary, 22 April, 2017.
- Judit Szűcs, Péter Balázs: Binary Image Reconstruction with Strip Constraints; In: 5th Winter School of PhD Students in Informatics and Mathematics, Debrecen, Hungary, 2-4 March, 2018, pp. 42-43.
(ISBN: 978-615-5586-23-1)
- Judit Szűcs: Image Reconstruction from Picross-based Vectors; In: Tavaszi Szél Konferencia 2018 absztrakt kötet, pp. 324.
(ISBN: 978-615-5586-26-2)
- Judit Szűcs and Péter Balázs: Strip Constrained Binary Tomography; In: Volume of Short Papers of the 11th Conference of PhD Students in Computer Science, Szeged, Hungary, 25-27 June, 2018, pp. 96.
- Szűcs Judit, Balázs Péter: Bináris rekonstrukció nonogramszerű priorral; In: Képfeldolgozók és Alakfelismerők Társaságának 12. Konferenciája, Debrecen, Hungary, 28-31 January, 2019, Paper Nr. 13.
- Judit Szűcs, Péter Balázs: Binary Image Reconstruction Based on Texture Priors; In: 6th Winter School of PhD Students in Informatics and Mathematics; Szeged, Hungary, 22-24 February, 2019, pp. 30.
(ISBN: 978-615-5586-37-8)
- Judit Szűcs, Péter Balázs: A Q-Convexity Based Local Shape Descriptor; In: 7th Winter School of PhD Students in Informatics and Mathematics; Veszprém, Hungary, 24-26 January, 2020, pp. 34.
(ISBN: 978-615-5586-54-5)
- Szűcs Judit: Q-konvexitáson alapuló alakleíró; In: XXIII. Tavaszi Szél Konferencia 2020 absztrakt kötet I., pp. 365.
(ISBN: 978-615-5586-70-5)