Available online at www.sciencedirect.com
SCIENCE DIRECT®
Discrete Applied Mathematics 151 (2005) 229–243
DISCRETE APPLIED MATHEMATICS
www.elsevier.com/locate/dam

Discrete tomography by convex–concave regularization and D.C. programming

T. Schüle[a, c], C. Schnörr[a], S. Weber[a], J. Hornegger[b]

[a] *Department of M&CS, University of Mannheim, CVGPR-Group, D-68131 Mannheim, Germany*
[b] *Department of M&CS, Friedrich-Alexander University Erlangen-Nürnberg, D-91058 Erlangen, Germany*
[c] *Siemens Medical Solutions, D-91301 Forchheim, Germany*

Received 6 October 2003; received in revised form 4 May 2004; accepted 10 February 2005

Abstract

We present a novel approach to the tomographic reconstruction of binary objects from few projection directions within a limited range of angles. A quadratic objective functional over binary variables comprising the squared projection error and a prior penalizing non-homogeneous regions, is supplemented with a concave functional enforcing binary solutions. Application of a primal-dual subgradient algorithm to a suitable decomposition of the objective functional into the difference of two convex functions leads to an algorithm which provably converges with parallel updates to binary solutions. Numerical results demonstrate robustness against local minima and excellent reconstruction performance using five projections within a range of 90°.

Our approach is applicable to quite general objective functions over binary variables with constraints and thus applicable to a wide range of problems within and beyond the field of discrete tomography.

Keywords: Discrete tomography; Combinatorial optimization; Concave minimization; D.C. programming

E-mail addresses: schuele@uni-mannheim.de (T. Schüle), schnoerr@uni-mannheim.de (C. Schnörr), wstefan@uni-mannheim.de (S. Weber), joachim@hornegger.de (J. Hornegger).
URLs: http://www.cvgpr.uni-mannheim.de (T. Schüle), http://www.siemensmedical.com (T. Schüle).

0166-218X/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2005.02.028

1. Introduction

1.1. Discrete tomography

Discrete tomography is concerned with the reconstruction of discrete-valued functions from projections. Historically, the field originated from several branches of mathematics like, for example, the combinatorial problem to determine binary matrices from its row and column sums (see the survey [14]). Meanwhile, however, progress is not only driven by challenging theoretical problems [7,10] but also by real-world applications where discrete tomography might play an essential role (cf. [12, Chapters 15–21]).

The work presented in this paper is motivated by the reconstruction of objects from *few* projection directions within a limited range of angles which, from the viewpoint of established mathematical models [16], is a severely ill-posed problem. For further information about the specific application area of our work, *Digital Substraction Angiography*, and its medical relevance we refer to [21]. Mathematically, it is reasonable to assume in this connection that the function f to be reconstructed is *binary-valued*. This poses one of the essential questions of discrete tomography: how can knowledge of the discrete range of f be exploited in order to regularize the reconstruction problem?

1.2. Related work

Related work in the field of discrete tomography include [15,3,6] where the reconstruction problem is formulated as an optimal estimation problem based on Markov-random-field models. Accordingly, stochastic sampling (Metropolis and Gibbs sampling, respectively) is used in [15,3] for the purpose of optimization which, when properly applied, is notoriously slow, whereas a multiscale implementation of a coordinate-wise sequential update technique (a special version of the well-known ICM-technique) is employed in [6]. In this connection, our parallel and deterministic mathematical programming approach presented below provides an alternative. Furthermore, our way to enforce a discrete-valued solution by concave minimization may be considered as a binary steering technique which is less heuristic than the approach in [2].

2. Problem statement

2.1. Projection equations

To each pixel x_i we assign a value a_i representing the absorption of a particular projection ray. For the example depicted in Fig. 1, the values a_3, a_4, a_5, a_6, and a_7 corresponding to the ray are simply measured by the lengths of the ray's sections through the pixel areas (indicated by black dots). This agrees with our assumption of a binary image to be reconstructed. Similarly, in the 3D case one would measure the intersections of each projection ray with each voxel x_i.

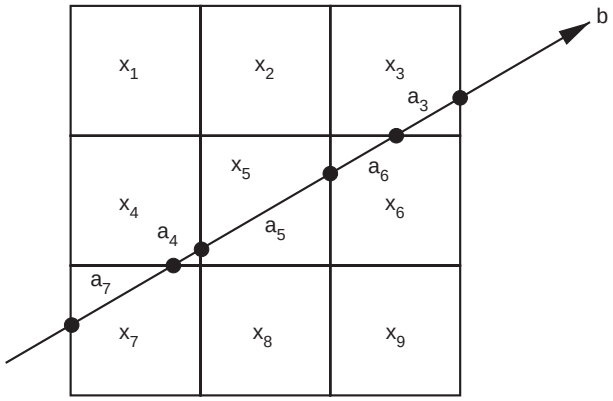


Fig. 1. Discretization model applied in order to obtain a mathematical formulation of the reconstruction problem.

Accordingly, the projection of a single ray is represented by the scalar product $\langle a, x\rangle = b,\ x_i \in \{0, 1\}, \forall i$, and the set of all projection rays yields the linear system

$$Ax = b, \quad x = (x_1, \ldots, x_n)^\top \in \{0, 1\}^n. \tag{1}$$

2.2. Linear objective functions

Due to noise in the measurement vector b when dealing with real data, (1) is likely to have no feasible solution. In order to take advantage of continuous problem formulations, Fishburn et al. [5] considered the relaxation $x_i \in [0, 1],\ i = 1, \ldots, n$, and investigated the following linear programming approach:

$$\min_{x\in[0,1]^n} \langle 0, x\rangle, \quad Ax = b. \tag{2}$$

In particular, the information provided by feasible solutions in terms of additivity and uniqueness of subsets $S \subset \mathbb{Z}^n$ is considered (see [5]).

Gritzmann et al. [11] introduced the following linear integer programming problem for binary tomography:

$$\max_{x\in\{0,1\}^n} \langle e, x\rangle, \quad e := (1, \ldots, 1)^\top, \quad Ax \leqslant b, \tag{3}$$

and suggested a range of greedy approaches within a general framework for local search. Compared to (2), the objective function (3) which is called *best-inner-fit* (*BIF*) in [11], looks for the maximal set compatible with the measurements. Also, the formulation of the constraints is better suited to cope with measurement errors and noise.

In [20,21], we considered the relaxation of (3) $x_i \in [0, 1], \forall i$, supplemented with a standard smoothness prior enforcing spatial coherency of solutions

$$\sum_{i=1}^{n} \sum_{j\in\mathcal{N}(i)} (x_i - x_j)^2, \tag{4}$$

where $\mathcal{N}(i)$ denotes grid positions in a 4-neighborhood (6-neighborhood in the 3D-case) around position i. In order to incorporate this prior into the linear programming approach (3), we had to use a corresponding approximation by means of auxiliary variables along with a randomized post-processing step for computing binary solutions.

2.3. Nonlinear objective functions

The main objective of this paper is to focus directly on a larger class of objective functions over binary variables. As a prototype problem, we consider the quadratic integer programming problem

$$\min_{x\in\{0,1\}^n} \frac{1}{2}\left\{ \|Ax-b\|^2 + \alpha \sum_{i=1}^{n} \sum_{j\in\mathcal{N}(i)} (x_i - x_j)^2 \right\}, \quad 0<\alpha\in\mathbb{R}, \tag{5}$$

where $\mathcal{N}(i)$ denotes grid positions in a neighborhood around i.

Objective function (5) combines the squared projection error with a standard smoothness prior enforcing spatial coherency of solutions. While similar functionals are common in connection with image restoration [4], the integer constraint $x_i \in \{0,1\}, \forall i$, renders the optimization problem a combinatorially difficult one.

In Section 3, we develop a relaxation of (5) based on the constraints $x_i \in [0,1], \forall i$, and an additional concave functional enforcing integer solutions. A corresponding mathematical programming approach is developed in Section 4 which allows to treat large-scale problems and parallel implementations. Numerical results demonstrating the feasibility of the approach are discussed in Section 5.

3. Relaxation and convex–concave regularization

Dropping constant terms, the objective function in (5) reads

$$E(x) := \tfrac{1}{2}\langle x, Qx\rangle + \langle q, x\rangle, \tag{6}$$

$$Q := A^\top A + \alpha L, \tag{7}$$

$$\langle x, Lx\rangle := \sum_{i=1}^{n} \sum_{j\in\mathcal{N}(i)} (x_i - x_j)^2, \tag{8}$$

$$q := -A^\top b. \tag{9}$$

The function E in (6) is convex because the matrix Q in (7) is positive semidefinite. As a consequence, we cannot expect to obtain an integer-valued minimizer $x \in \{0,1\}^n$ based on the relaxation $\min_x E(x), x \in [0,1]^n \subset \mathbb{R}^n$. Therefore, we supplement our objective function with a term enforcing integer solutions and associate with the integer programming problem

$$\min_{x\in\{0,1\}^n} E(x), \tag{10}$$

the nonlinear problem

$$\min_{x\in[0,1]^n} F(x;\mu), \quad F(x;\mu) := E(x) + \tfrac{1}{2}\mu\langle x, e-x\rangle, \quad 0<\mu\in\mathbb{R}, \tag{11}$$

with $e := (1,\ldots,1)$. Note that this family of functionals, parametrized by μ, contains two terms regularizing the underconstrained problem of minimizing the squared projection error $\|Ax-b\|^2$: the convex smoothness term $\frac{1}{2}\langle x, Lx\rangle$ with weight α that penalizes non-homogeneous reconstructions, and the concave term $\frac{1}{2}\mu\langle x, e-x\rangle$. The rationale for the latter term is given in the following proposition (see [8,13] for a proof).

Proposition 1. *Suppose that E is Lipschitzian on an open set $A \supset [0,1]^n$ and twice continuously differentiable on $[0,1]^n$. Then there exists a $\mu_Q \in \mathbb{R}$ such that for all $\mu > \mu_Q$:*

(i) (10) *and* (11) *are equivalent,*
(ii) *$F(x)$ is concave on $[0,1]^n$.*

This connection between integer programming and concave minimization is well-known. Its applicability depends on how severely the concave minimization problem suffers from local minima. Numerical results indicate (Section 5) that (11) is fairly well-behaved in this respect.

The reason, we believe, is the convexity of the functional E. Starting with a sufficiently small parameter value μ, we can easily compute a *global* minimum of F in (11). Next, we gradually enforce the integer constraint $x_i \in \{0,1\}, \forall i$, by locally minimizing $F(x;\mu)$ for each value of an increasing sequence $\{\mu^k\} \to \mu_Q$, where μ_Q equals an upper bound [9] of the largest eigenvalue of Q in (7). A corresponding algorithm will be derived in the following section. The concave term which becomes "active" through $\{\mu^k\}$ does not introduce any bias, i.e. prefer one local minimum over another one. Finally, due to Proposition 1, we obtain a local integer-valued minimum $x \in \{0,1\}^n$ of $F(x;\mu)$ which is also a local minimum of $E(x)$, since $F(x;\mu) = E(x)$ for $x \in \{0,1\}^n$.

4. D.C. programming

The objective of this section is to derive an algorithm for minimizing $F(x;\mu)$ in (11). The presence of both convex and concave terms suggests a mathematical programming approach to the minimization of the difference of two convex functions (*D.C. programming*). We sketch a corresponding approach [17,18] in Section 4.1 and work out its application to the optimization problem (11) in Section 4.2.

We point out that we deliberately choose a mathematical formulation general enough to encompass not only problem (11) but also other combinations of terms and constraints discussed in Section 2.

The following notation and basic concepts of convex analysis [19] will be used for a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and a set $C \subset \mathbb{R}^n$:

$$\operatorname{dom} f = \{x \in \mathbb{R}^n \mid f(x) < +\infty\} \quad \text{effective domain of } f,$$

$$f^*(y) = \sup_{x\in\mathbb{R}^n}\{\langle x, y\rangle - f(x)\} \quad \text{(conjugate function)},$$

$$\hat{\partial} f(\overline{x}) = \{v \mid f(x) \geqslant f(\overline{x}) + \langle v, x - \overline{x}\rangle, \forall x\} \quad \text{subdifferential of } f \text{ at } \overline{x},$$

$$\delta_C(x) = \begin{cases} 0, & x \in C \\ +\infty, & x \notin C \end{cases} \quad \text{(indicator function of } C).$$

Recall that when f is proper, lower-semicontinuous (LSC), and convex, then $f = f^{**} = (f^*)^*$ and:

$$\hat{\partial} f(\overline{x}) = \{y \mid f(x) \geqslant f(\overline{x}) + \langle y, x - \overline{x}\rangle, \ \forall x\} \tag{12}$$

$$= \operatorname{argmax}_y\{\langle y, \overline{x}\rangle - f^*(y)\} \tag{13}$$

$$\hat{\partial} f^*(\overline{y}) = \{x \mid f^*(y) \geqslant f^*(\overline{y}) + \langle x, y - \overline{y}\rangle, \ \forall y\} \tag{14}$$

$$= \operatorname{argmax}_x\{\langle \overline{y}, x\rangle - f(x)\} \tag{15}$$

4.1. DC-algorithm

Let $g, h : \mathbb{R}^n \to \overline{R}$ be proper, LSC and convex,

$$\operatorname{dom} g \subset \operatorname{dom} h, \quad \operatorname{dom} h^* \subset \operatorname{dom} g^*, \tag{16}$$

and

$$f(x) = g(x) - h(x). \tag{17}$$

Consider the optimization problem

$$\inf_x\{g(x) - h(x)\}. \tag{18}$$

Using conjugate functions, we write

$$\begin{aligned} \inf_x\{g(x) - h(x)\} &= \inf_x\left\{g(x) - \sup_y\{\langle x, y\rangle - h^*(y)\}\right\} \\ &= \inf_x \inf_y\{g(x) - [\langle x, y\rangle - h^*(y)]\} \end{aligned}$$

and obtain the dual problem of (18):

$$\inf_y\{h^*(y) - g^*(y)\}. \tag{19}$$

We use the following primal-dual subgradient algorithm [18]:

DC-algorithm (*DCA*):
Choose $x^0 \in \operatorname{dom} g$ arbitrary.
For $k = 0, 1, \ldots$ compute:

$$y^k \in \hat{\partial} h(x^k), \tag{20}$$

$$x^{k+1} \in \hat{\partial} g^*(y^k). \tag{21}$$

The investigation of DCA in [18] includes the following results:

Proposition 2 (*Pham Dinh and Hoai An [18]*). *Assume* $g, h : \mathbb{R}^n \to \overline{\mathbb{R}}$ *be proper, LSC and convex, and* (16). *Then*

(i) *the sequences* $\{x^k\}, \{y^k\}$ *according to* (20), (21) *are well-defined,*
(ii) $\{g(x^k) - h(x^k)\}$ *is decreasing,*
(iii) *every limit point* x^* *of* $\{x^k\}$ *is a critical point of* $g - h$.

Step (21) corresponds to the affine majorization defined by $y^k \in \partial h(x^k)$ of the concave—hence "difficult"—part in (18). From (12), we have

$$\inf_x\{g(x) - h(x)\} \leqslant \inf_x\{g(x) - [h(x^k) + \langle y^k, x - x^k\rangle]\},$$

thus, by (15)

$$\begin{aligned} x^{k+1} &\in \operatorname{argmin}_x\{g(x) - [h(x^k) + \langle y^k, x - x^k\rangle]\} \\ &= \partial g^*(y^k) \end{aligned}$$

which is a convex optimization problem. Similarly, from $x^k \in \partial g^*(y^{k-1})$ and (14) follows

$$\inf\{h^*(y) - g^*(y)\} \leqslant \inf\{h^*(y) - [g^*(y^{k-1}) + \langle x^k, y - y^{k-1}\rangle]\},$$

thus, by (13)

$$\begin{aligned} y^k &\in \operatorname{argmin}_y\{h^*(y) - [g^*(y^{k-1}) + \langle x^k, y - y^{k-1}\rangle]\} \\ &= \partial h(x^k) \end{aligned}$$

Again we point out that this optimization problem is convex.

4.2. Application to the relaxed problem

Both the application and the efficiency of the algorithm (20), (21) depend on the choice of the decomposition (17), which is not unique. In this paper, we decompose the functional $F(x; \mu) = g(x) - h(x; \mu)$ in (11) as follows:

$$g(x) = \tfrac{1}{2}\langle x, \lambda I x\rangle + \delta_C(x), \quad C = [0, 1]^n \tag{22}$$

$$h(x; \mu) = \tfrac{1}{2}\langle x, (\lambda I - Q)x\rangle - \langle q, x\rangle - \tfrac{1}{2}\mu\langle x, (e - x)\rangle \tag{23}$$

$$= \tfrac{1}{2}\langle x, [(\lambda + \mu)I - Q]x\rangle - \langle q + \tfrac{1}{2}\mu e, x\rangle \tag{24}$$

with λ equal to an upper bound of the largest eigenvalue of Q.

Note that both g and h are convex. Since h is smooth, $\partial h(x) = \{\nabla h(x)\}$ and step (20) amounts to evaluate the gradient

$$y^k = \nabla h(x^k; \mu) \tag{25}$$

$$= [(\lambda + \mu)I - Q]x^k - (q + \tfrac{1}{2}\mu e). \tag{26}$$

Function g, on the other hand, is non-smooth due to the constraint $x \in C$, and we have to solve problem (15)

$$\begin{aligned} x^{k+1} &\in \partial g^*(y^k) && (27)\\ &= \operatorname{argmin}_x\{g(x) - \langle y^k, x\rangle\} && (28)\\ &= \operatorname{argmin}_x \left\{\frac{\lambda}{2}\|x\|^2 - \langle y^k, x\rangle + \delta_C(x)\right\} && (29)\\ &= \operatorname{argmin}_x\{\tilde{g}(x) + \delta_C(x)\}. && (30) \end{aligned}$$

As this amounts to minimize a *strictly* convex function over the convex set C, the global minimum x^{k+1} is unique and satisfies [19]

$$\begin{aligned} 0 &\in \nabla\tilde{g}(x) + \delta_C(x)\\ &= \lambda x^{k+1} - y^k + \partial\delta_C(x^{k+1}). \end{aligned} \tag{31}$$

Let $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^\top$, $i = 1, \ldots, n$, denote the ith unit vector, and $I_0(x)$, $I_1(x)$ the sets of indices where x meets the boundary of C, i.e. $i \in I_0(x)$ (resp. $I_1(x)$) if $x_i = 0$ (resp. $x_i = 1$), $\forall i$. Then $\partial\delta_C(x^{k+1})$ is the polyhedral convex cone generated by the vectors $\{-e_i\}_{i\in I_0(x^{k+1})}$ and $\{e_i\}_{i\in I_1(x^{k+1})}$. Thus, from (31), we compute the global minimum

$$(x^{k+1})_i = \begin{cases} 0, & y_i^k \leqslant 0\\ 1, & y_i^k \geqslant \lambda\\ \frac{1}{\lambda}y_i^k, & \text{otherwise} \end{cases}, \quad i = 1, \ldots, n. \tag{32}$$

The iteration rules (26) and (32) show why the decomposition (22), (24) is useful. For fixed value of the parameter μ, a local minimum of $F(x;\mu)$ in (11) can be computed by iterating the steps (26) and (32), both of which are computationally cheap and suitable for parallel implementations. In particular, we only need a sparse matrix vector product in (26) but no inversion of the matrix Q.

To summarize, we obtain the following overall algorithm for solving the binary tomography reconstruction problem by means of (11):

Algorithm.

(1) Choose $l = 0$, $\mu^0 = 0$, $x^0 = \frac{1}{2}e$, μ_Δ
(2) *Inner loop*:
    for $k = 0, 1, \ldots$ compute (26), (32) until convergence: $\{x^k\} \to x_{\mu^l}$
(3) *Outer loop*:
    if $x_{\mu^l} \in \{0, 1\}^n$ then stop; else increment $\mu^{l+1} = \mu^l + \mu_\Delta$, put $x^0 = x_{\mu^l}$, and repeat (2).

Condition (16) is satisfied because obviously dom $g \subset$ dom h and dom h^*=dom g^*=$\mathbb{R}^n$. Consequently, the results stated in Proposition 2 hold.

Remark. For the specific decomposition (22), (24), the algorithm above turns out to be a special instance of the Goldstein–Levitin–Polyak projection method [1]

$$x^{k+1} = P_C[x^k - a^k\nabla_x F(x^k, \mu)], \quad k = 0, 1, \ldots, \tag{33}$$

where $P_C[\cdot]$ denotes the unique projection onto the convex set C. In general, the damping parameter a^k has to be chosen with care in order to achieve convergence. In this connection, our approach proves convergence for the choice $a^k = 1/\lambda$.

5. Numerical results

We demonstrate the feasibility and the typical behavior of our approach with two examples of size 64 × 64 and 256 × 256. The algorithm was implemented in C++ and run on a Pentium 4 with 3 GHz. The smaller image was reconstructed from only three projections with horizontal, diagonal and vertical rays, whereas for the larger one five projections at 0°, 22.5°, 45°, 67.5°, and 90° were used. Both problems led to highly underdetermined linear systems (1), with 4096 unknowns and 224 equations for the 64 × 64 image and 65 536 unknowns and 1453 equations for the 256 × 256 example.

Parameter α was set to 0.1. For a fixed value of μ, the stopping criterion of the inner loop was $\|x^{k+1} - x^k\|_2 \leqslant \varepsilon_{\text{in}}$ with $\varepsilon_{\in} = 0.0001$ for the 64×64 image and $\varepsilon_{\in} = 0.0025$ for the larger one. After convergence of the inner loop μ was incremented with $\mu_{\Delta} = 0.00005 \cdot \mu_Q$ and the program terminated (convergence of the outer loop) when each component of x was within a neighborhood of $\varepsilon_{\text{out}} = 0.001$ around either 0 or 1: $\max_{i=1,\ldots,n}\{\min\{x_i, 1 - x_i\}\} < \varepsilon_{\text{out}}$.

In both cases the algorithm was able to reconstruct the original image (Figs. 4 and 5). Fig. 6 shows the error $\|Ax - b\|_2$ after each iteration step. The overall computational time for the results (Figs. 4(f) and 5(f)) was about 99 s for the small image and 379 s for the large image.

The typical distribution of the computational costs and convergence as a function of μ is discussed in Figs. 2 and 3. (Figs. 4–6)

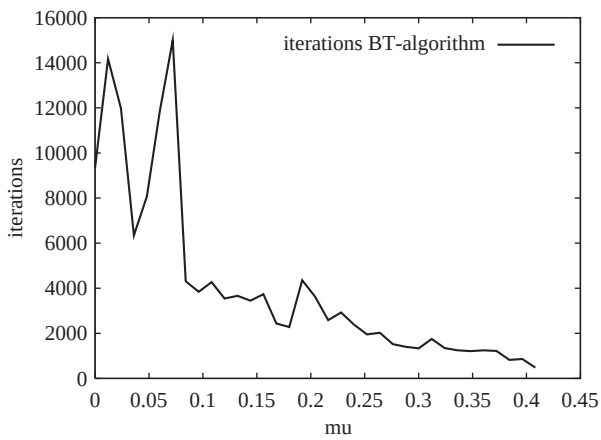


Fig. 2. Number of iterations of the inner loop as a function of μ (recorded for the reconstruction shown in Fig. 4). Most computational work has to be done within a small interval of "critical" values of μ. Outside this interval, that is for further increasing values of μ, most local decisions have already been made, and each x_i quickly converges to either 0 or 1 (cf. Fig. 3).

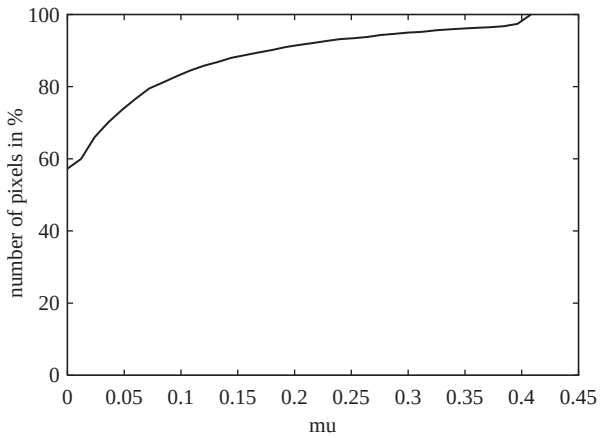


Fig. 3. Number of pixels $\{x_i\}$ as a function of μ which do satisfy the outer loop convergence criterion $\min\{x_i, 1 - x_i\} < \varepsilon_{\text{in}}$ (recorded for the reconstruction shown in Fig. 4). After a critical μ-phase shown in Fig. 2, most local decision have already been made, and each x_i quickly converges to either 0 or 1.

The binary constraint $x_i \in \{0, 1\}$ is satisfied for each pixel with accuracy ε_{out}. Consequently, no rounding scheme as a post-processing step is necessary.

To substantiate the robustness of our approach against local minima, we focused on a 4×4 toy problem for which exhaustive enumeration of all binary vectors $x \in \{0, 1\}^{16}$ is feasible. A reconstruction problem was created using horizontal and vertical projection rays (which corresponds to row and column sums), and α was again set to 0.1. Fig. 7 shows as possible solutions the subset of all 34 binary images consistent with the given projection constraints $Ax = b$.

The images (1)–(34) in Fig. 7 are ordered according to the respective values of $E(x)$ in (6). Note that lower function values corresponds to smoother images, i.e. the amount of horizontally and vertically connected structures decreases as $E(x)$ increases. The reconstruction computed by our algorithm is depicted as image (35) in Fig. 7. It coincides with image (2) and minimizes $E(x)$ among all 34 consistent solutions.

6. Conclusion and further work

We presented a novel framework for solving combinatorial binary reconstruction problems of discrete tomography. Based on a given objective functional over binary variables for measuring the quality of reconstructions, the main components of the approach are (i) a smoothness prior to achieve spatial coherent reconstructions, (ii) concave regularization for enforcing binary solutions, and (iii) a sound mathematical programming approach based on primal-dual subgradient minimization.

Numerical results show an excellent performance for reconstruction problems using three or five projection directions within 90°. We also illustrated the robustness of the approach with respect to local minima. At present, however, we have no proof regarding a performance

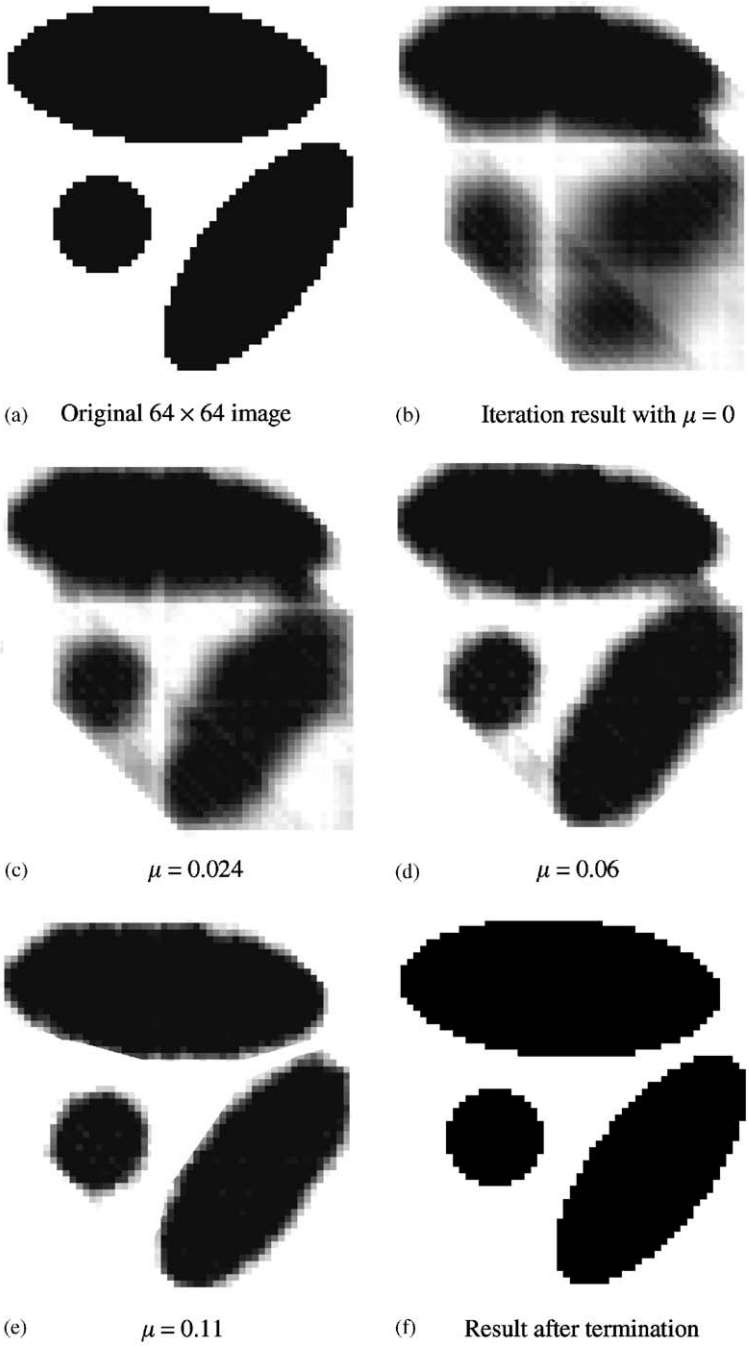


Fig. 4. (a) Original image (64 × 64), (b)–(f) Reconstructions from three projections (0°, 45°, 90°). Each image shows the reconstruction result after termination of the inner loop for a fixed value of μ.

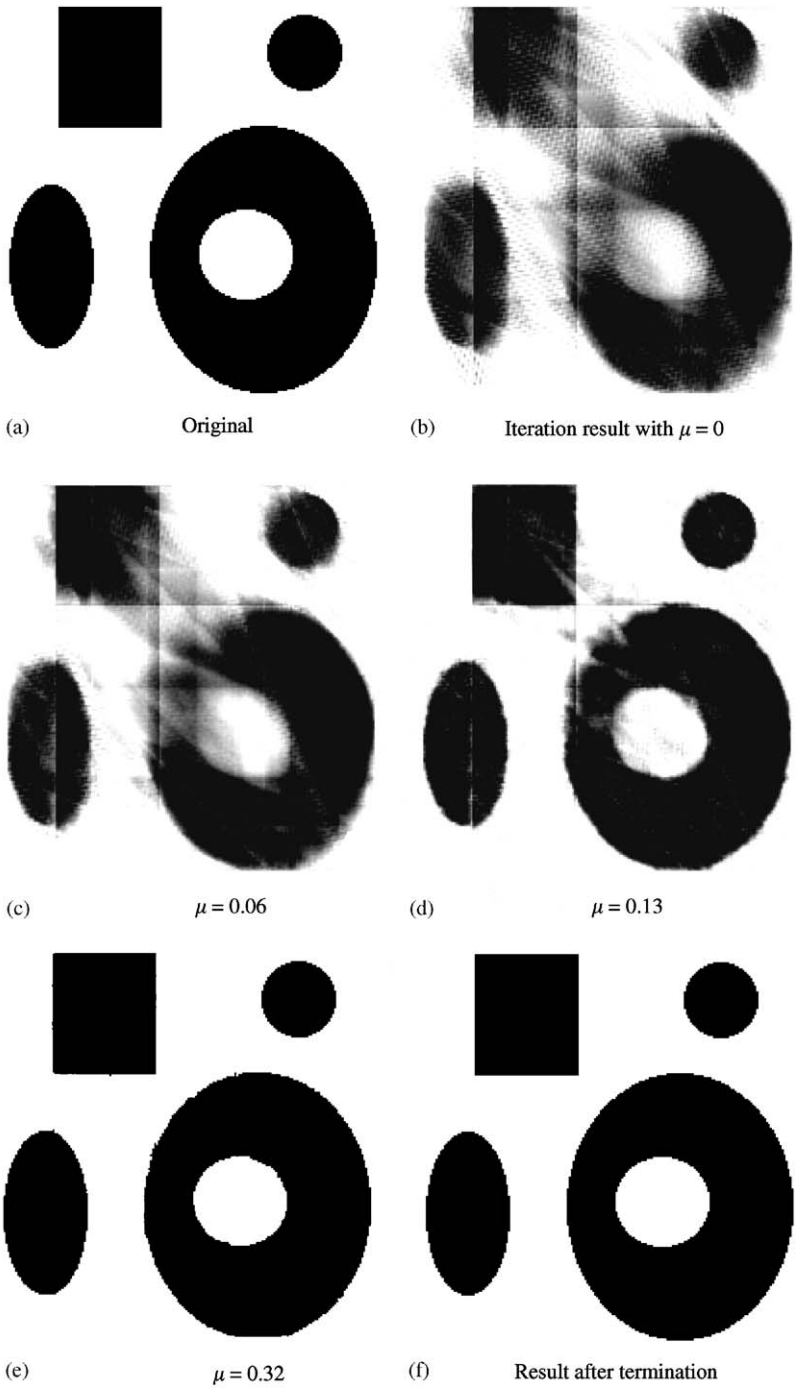
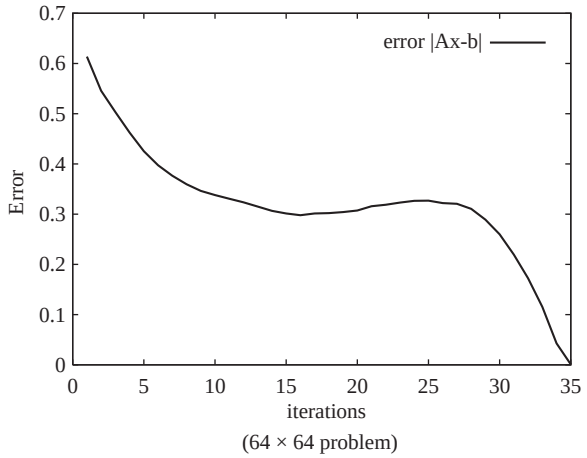
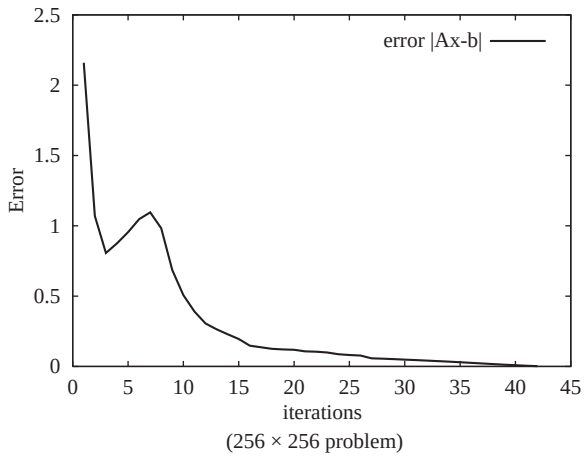


Fig. 5. (a) Ground truth 256 × 256 image, (b)–(f) Reconstructions from five projections $(0°, 22.5°, 45°, 67.5°, 90°)$. Each image shows the reconstruction result after termination of the inner loop for a fixed value of μ.



(64 × 64 problem)



(256 × 256 problem)

Fig. 6. Projection error $\|Ax-b\|$ after each outer loop iteration for both examples. Note that error does not decrease monotonously, due to enforcing the integer constraint $x_i \in \{0, 1\}, \forall i$, during the iteration.

bound of combinatorial solutions. Finally, we showed that suitable decompositions of the objective function into the difference of convex functions lead to a two-step iteration, both steps of which are computationally cheap and allow for parallel implementations.

So far, we applied the approach to the particular objective function (5). However, our formulation is sufficiently general to include many variations of (5) by combining terms discussed in Section 2 and further constraints. Note, for example, that Proposition 1 also holds for objective functions defined on any binary subset $\{0, 1\}^n \cap C$, and that the formulation (20), (21) in terms of subgradients admits using general non-smooth objective functions.

Finally, we point out that, typically, the decomposition (17) is not unique but allows to derive various instances of the subgradient minimization scheme with different properties of algorithms and corresponding implementations. Explorations of these possibilities will be the objective of our future work.

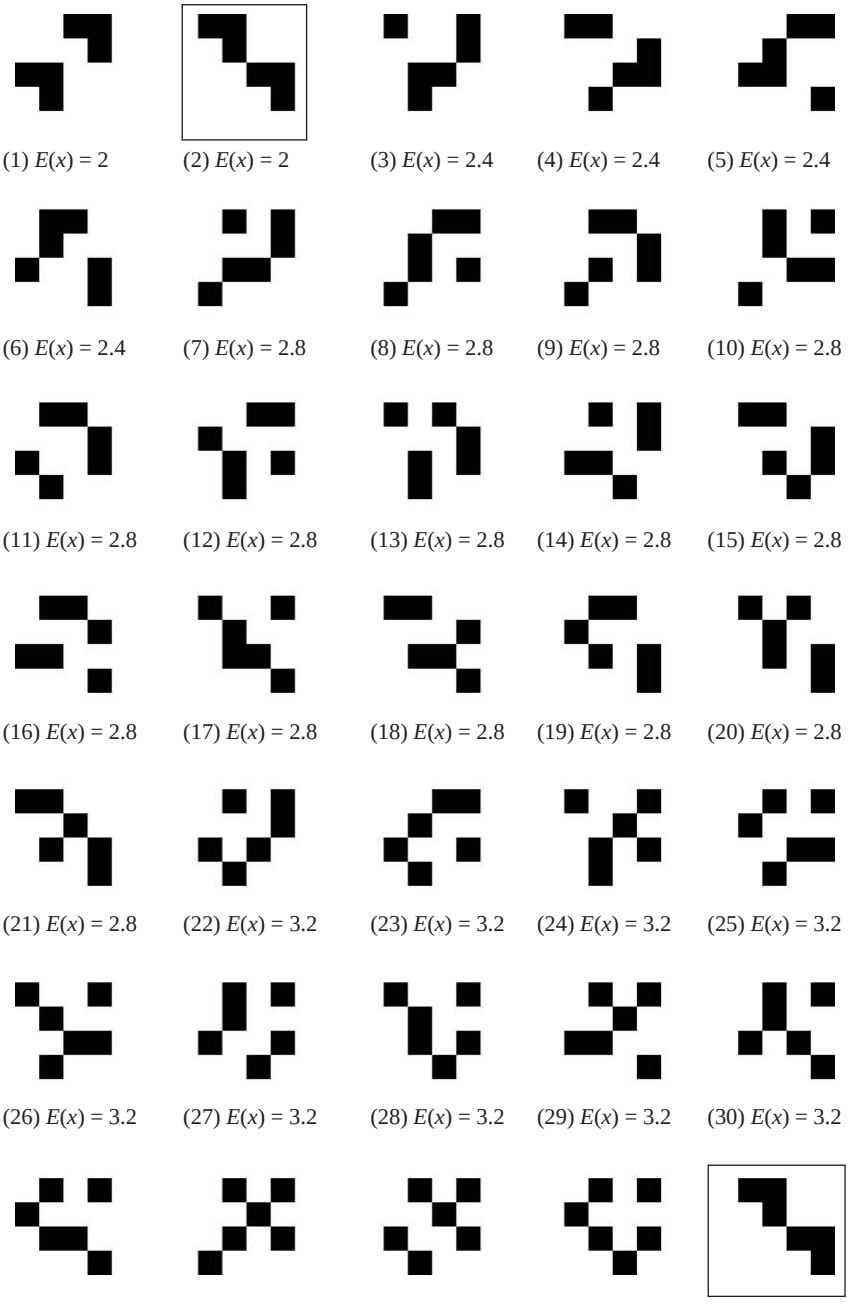


Fig. 7. (1)–(34) show image x with the same horizontal and vertical projections. The images are ordered according to their objective value $E(x)$. Image (35) is the reconstruction result of our algorithm which, indeed, minimizes E. This result also illustrates the robustness of the approach against poor local minima.

References

[1] D.P. Bertsekas, On the Goldstein–Levitin–Polyak gradient projection method, IEEE Trans. Automat. Control 21 (2) (1976) 174–184.
[2] Y. Censor, Binary steering in discrete tomography reconstruction with sequential and simultaneous iterative algorithms, Linear Algebra Appl. 339 (2001) 111–124.
[3] M.T. Chan, G.T. Herman, E. Levitan, Bayesian image reconstruction using image-modeling gibbs priors, Internat. J. Imag. System Technol. 9 (1998) 85–98.
[4] G. Demoment, Image reconstruction and restoration: overview of common estimation structures and problems, IEEE Trans. on Acoust. Speech Signal Process. 37 (12) (1989) 2024–2036.
[5] P. Fishburn, P. Schwander, L. Shepp, R. Vanderbei, The discrete radon transform and its approximate inversion via linear programming, Discrete Appl. Math. 75 (1997) 39–61.
[6] T. Frese, C.A. Bouman, K. Sauer, Multiscale Bayesian methods for discrete tomography, in: G.T. Herman, A. Kuba (Eds.), Discrete Tomography, Birkhäuser, Basel, 1999, pp. 237–267.
[7] R.J. Gardner, P. Gritzmann, Discrete tomography: determination of finite sets by X-rays, Trans. Amer. Math. Soc. 349 (6) (1997) 2271–2295.
[8] F. Giannessi, F. Niccolucci, Connections between nonlinear and integer programming problems, in: Symposia Mathematica, Istituto Nazionale di Alta Mathematica, 1976, pp. 161–176.
[9] G.H. Golub, C.F. VanLoan, Matrix Computations, third ed., Johns Hopkins University Press, Baltimore, MD, 1997.
[10] P. Gritzmann, D. Prangenberg, S. de Vries, M. Wiegelmann, Success and failure of certain reconstruction and uniqueness algorithms in discrete tomography, Internat. J. Imag. System Technol. 9 (1998) 101–109.
[11] P. Gritzmann, S. de Vries, M. Wiegelmann, Approximating binary images from discrete X-rays, SIAM J. Optim. 11 (2) (2000) 522–546.
[12] G. Herman, A. Kuba (Eds.), Discrete Tomography: Foundations, Algorithms, and Applications, Birkhäuser, Boston, 1999.
[13] R. Horst, H. Tuy, Global Optimization: Deterministic Approaches, third ed., Springer, Berlin, 1996.
[14] A. Kuba, G.T. Herman, Discrete tomography: a historical overview, in: G.T. Herman, A. Kuba (Eds.), Discrete Tomography, Birkhäuser, Boston, 1999, pp. 3–34.
[15] S. Matej, G.T. Herman, A. Vardi, Binary tomography on the hexagonal grid using gibbs priors, Internat. J. Imag. System Technol. 9 (1998) 126–131.
[16] F. Natterer, F. Wübbeling, Mathematical Methods in Image Reconstruction, SIAM, Philadelphia, PA, 2001.
[17] T. Pham Dinh, S. Elbernoussi, Duality in d.c. (difference of convex functions) optimization, subgradient methods, in: Trends in Mathematical Optimization, vol. 84, International Series of Numerical Mathematics, Birkhäuser, Basel, 1988, pp. 277–293.
[18] T. Pham Dinh, L.T. Hoai An, A d.c. optimization algorithm for solving the trust-region subproblem, SIAM J. Optim. 8 (2) (1998) 476–505.
[19] R.T. Rockafellar, Convex Analysis, second ed., Princeton University Press, Princeton, NJ, 1972.
[20] S. Weber, C. Schnörr, J. Hornegger, A linear programming relaxation for binary tomography with smoothness priors, in: Proceedings of International Workshop on Combinatorial Image Analysis (IWCIA'03), Palermo, Italy, May 14–16, 2003.
[21] S. Weber, T. Schüle, C. Schnörr, J. Hornegger, A linear programming approach to limited angle 3D reconstruction from DSA projections, Methods of Information in Medicine, 43 (2004) 320–326.