

# Binary Tomography with Deblurring

Stefan Weber<sup>1</sup>, Thomas Schüle<sup>1,3</sup>, Attila Kuba<sup>2</sup>, and Christoph Schnörr<sup>1</sup>

<sup>1</sup> University of Mannheim,  
Dept. of Mathematics and Computer Science, CVGPR-Group,  
D-68131 Mannheim, Germany

{wstefan, schuele, schnoerr}@uni-mannheim.de  
www.cvgpr.uni-mannheim.de

<sup>2</sup> University of Szeged,  
Dept. of Image Processing and Computer Graphics,  
H-6720 Szeged, Hungary

kuba@inf.u-szeged.hu,  
www.inf.u-szeged.hu/~kuba

<sup>3</sup> Siemens Medical Solutions,  
D-91301 Forchheim, Germany  
www.siemensmedical.com

**Abstract.** We study two scenarios of limited-angle binary tomography with data distorted with an unknown convolution: Either the projection data are taken from a blurred object, or the projection data themselves are blurred. These scenarios are relevant in case of scattering and due to a finite resolution of the detectors. Assuming that the unknown blurring process is adequately modeled by an isotropic Gaussian convolution kernel with unknown scale-parameter, we show that parameter estimation can be combined with the reconstruction process. To this end, a recently introduced Difference-of-Convex-Functions programming approach to limited-angle binary tomographic reconstruction is complemented with Expectation-Maximization iteration. Experimental results show that the resulting approach is able to cope with both ill-posed problems, limited-angle reconstruction and deblurring, simultaneously.

## 1 Introduction

It is a general characteristic of imaging systems that the acquired images are some distorted versions of the ideal images of real objects. The distortion is due to physical limitations, e.g., finite resolution in space and time, non-uniform sensitivity in the field of view, etc. In many cases the distorted image can be modeled as the convolution of the ideal image with some function describing the distortion [1].

The situation is the same in tomography when the cross-sections of some 3D object are reconstructed from its projections. The pixel values in the projection images are usually only some approximations of the line integrals to be measured by a perfect imaging system in an ideal physical situation. In different application areas of tomography there are several correction methods to improve the

quality of the reconstructed images. The correction strategies can be divided into two classes roughly. The first class contains the methods aiming to correct the projection data before reconstruction (let us call them preprocessing) and then the reconstruction is performed from the corrected projection data. The second class is the family of special methods when the correction is included into the reconstruction process. We believe that both strategies can be useful. If the correction can be done as a preprocessing step before reconstruction then one of the methods from the first class is preferable. However, there are situations when the correction is impossible or too complicated before reconstruction, e.g., scatter correction in CT or in SPECT, then the correction during the reconstruction can still give a good solution.

The situation is very similar in the case of binary tomography, when the range of the function to be reconstructed is just the set  $\{0, 1\}$  (as a summary of binary tomography see [2]). The known discrete range can be used in the reconstruction process as a kind of a priori information, and binary functions can be reconstructed effectively from very few projections (e.g., 2-5). As binary tomography is getting to be applied in several areas, the problem of distortion of such tomography images becomes an important problem to be studied. There are publications discussing different corrections in DT, e.g. in X-ray and neutron tomography [3, 4], and electron microscopy [5].

In this paper we deal with the general distortion model when the distortion can be described by the convolution with a Gaussian kernel  $G_\sigma(\cdot)$ . If the parameter  $\sigma$  is known in advance then the correction (deconvolution) can be done as a preprocessing step before the reconstruction. However, if the parameter is not known then we are going to show that there is still a way to binary tomography by including this parameter as an unknown value to be determined. To motivate our approach we present some reconstructions, see figure 1, performed without deblurring.

Section 2 shows the mathematical model of distorted DT and the reconstruction problem to be solved. Our reconstruction approach including the deconvolution adaptation is described in Section 3. The optimization algorithm is specified in section 4. Several experiments have been done to test our reconstruction procedure for both noiseless and noisy data. The corresponding results are presented in Section 5. We conclude and indicate further work in Section 6.

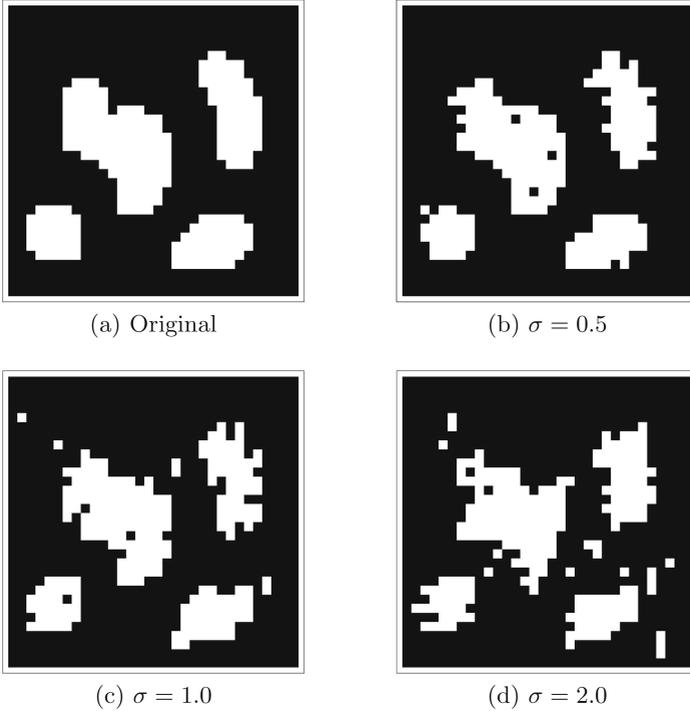
## 2 Problem Statement

### 2.1 Binary Tomography and Reconstruction by DC-Programming

We consider the reconstruction problem of transmission tomography for binary objects. As explained in figure 2, the imaging process is represented by the algebraic system of equations

$$Ax = b, \quad A \in \mathbb{R}^{m \times n}, \quad x \in \{0, 1\}^n, \quad b \in \mathbb{R}^m, \quad (1)$$

where  $A$  and  $b$  are given, and the binary indicator vector  $x$  representing the unknown object has to be reconstructed. To this end, we introduced in [6] the variational approach



**Fig. 1. Reconstruction without deblurring fails.** Panel (a) shows an object which was blurred with a Gaussian convolution kernel  $G_\sigma$  at three different scales  $\sigma \in \{0.5, 1.0, 2.0\}$ , and then projected along 5 directions  $0^\circ, 22.5^\circ, 45^\circ, 67.5^\circ, 90^\circ$ . Panels (b)-(d) show the reconstruction results *without* deblurring. The performance considerably deteriorates for increasing  $\sigma$ . Note that the original object (a) can be reconstructed without error from three projections only.

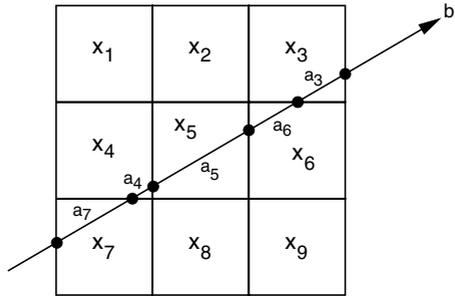
$$x_\mu^* = \operatorname{argmin}_{x \in [0,1]^n} J_\mu(x), \quad J_\mu(x) = D(x) + \alpha S(x) - \mu \frac{1}{2} \langle x, x - e \rangle \quad (2)$$

where

$$D(x) = \frac{1}{2} \|Ax - b\|^2 \quad (3)$$

and  $S(x)$  is a convex smoothness prior (see section 3.2) which favors spatially homogeneous objects as reconstructions and  $e := (1, 1, \dots, 1)^\top \in \mathbb{R}^n$ .

Problem (2) constitutes a numerically convenient relaxation of the combinatorial problem (1) because the set of feasible solutions  $[0, 1]^n$  is convex. Starting with the global optimum  $x_0$  of the *convex* functional  $J_0$ , the last term in (2) gradually enforces a locally optimal binary solution for increasing values of parameter  $\mu$ . Although global optimality cannot be guaranteed, experimental results showed an excellent reconstruction performance [6, 7]. For further details of this framework and an overview from the optimization point of view, we refer to [8].



**Fig. 2. Discretization model for transmission tomography.** The measured projection data are given in terms of a vector  $b \in \mathbb{R}^m$ . Each component  $b_i$  corresponds to a projection ray measuring the absorption along the ray through the volume which is discretized into cells. The absorption  $a_j$  in each cell is assumed to be proportional to the density of the unknown object.  $x_1, x_2, \dots$  are binary variables indicating whether the corresponding cells belong to the object ( $x_k = 1$ ) or not ( $x_k = 0$ ). Assembling all projection rays into a linear system gives  $Ax = b$ ,  $x \in \{0, 1\}^n$ , from which the unknown binary object, represented by  $x$ , has to be determined.

### 2.2 Binary Tomography with Blurred Data

Let  $G_\sigma$  denote the matrix that represents the linear mapping of some data by convolving it with an isotropic Gaussian kernel and scale-parameter  $\sigma$ . We generalize problem (1) along two directions:

#### Reconstruction from Projections of Blurred Objects

The corresponding generalization of the reconstruction problem (1) reads:

$$AG_\sigma x = b, \quad A \in \mathbb{R}^{m \times n}, \quad x \in \{0, 1\}^n, \quad b \in \mathbb{R}^m \tag{4}$$

#### Reconstruction from Blurred Projection Data

The corresponding generalization of the reconstruction problem (1) reads:

$$G_\sigma Ax = b, \quad A \in \mathbb{R}^{m \times n}, \quad x \in \{0, 1\}^n, \quad b \in \mathbb{R}^m \tag{5}$$

For notational simplicity, we used in both cases the same symbol  $G_\sigma$ , although  $G_\sigma$  denotes a block-circulant matrix in (4) corresponding to the convolution of multi-dimensional data  $x$ , whereas  $G_\sigma$  represents the one-dimensional convolution of the projection data in (5).

Accordingly, the variational approach (2) generalizes to

$$x_\mu^* = \operatorname{argmin}_{x \in [0, 1]^n} J_\mu(x; \sigma), \quad J_\mu(x; \sigma) = D(x; \sigma) + \alpha S(x) - \mu \frac{1}{2} \langle x, x - e \rangle, \tag{6}$$

where the data term  $D(x; \sigma)$  indicates the dependency on the unknown convolution operator in (4) and (5), respectively.

### 3 Approach

#### 3.1 Data Term and Scale Estimation

Optimization of  $J_\mu$  in (6) is complicated through the unknown scale-parameter  $\sigma$  of the convolution operator  $G_\sigma$ . A common and natural approach to solve this problem is to apply the well-known Expectation-Maximization (EM) iteration (cf., e.g. [9]) to the probabilistic interpretation of the data term  $D(x; \sigma)$  as a likelihood term, provided this is computationally feasible. We elaborate this approach in this section.

We regard minimization of  $J_\mu$  in (6) as Maximum-A-Posteriori (MAP) estimation of  $x$ , given the data  $b$ :

$$p(x|b) \propto \exp(-J_\mu(x; \sigma)) \propto p(b|x)p(x) \quad (7a)$$

$$p(x) \propto \exp(-\alpha S(x) + \mu \frac{1}{2}(x, x - e)) \quad (7b)$$

**Remark.** The normalizing term missing in (7a) only depends on  $b$  and therefore it is unessential for estimating  $x$ .

The data likelihood  $p(b|x)$  is unknown due to the dependency of the data term  $D(x; \sigma)$  on the unknown parameter  $\sigma$ . Given some estimate  $\hat{x}$ , the standard EM-approach is then to maximize instead the following lower bound, commonly called  $Q$ -function, which does not depend on  $\sigma$ :

$$\log p(b|x) \geq \int_{\mathbb{R}_+} p(\sigma|b, \hat{x}) \log \frac{p(b, \sigma|x)}{p(\sigma|b, \hat{x})} d\sigma$$

Expanding the log-expression shows that only the first term, commonly called  $Q$ -function, depends on  $x$  and therefore is relevant:

$$Q(x|\hat{x}, b) := \int_{\mathbb{R}_+} p(\sigma|b, \hat{x}) \log p(b, \sigma|x) d\sigma \quad (8)$$

To compute (8), the first term under the integral is evaluated via Bayes' rule

$$p(\sigma|b, \hat{x}) = \frac{p(b|\sigma, \hat{x})p(\sigma|\hat{x})}{p(b|\hat{x})}.$$

The denominator does not depend on  $\sigma$  and therefore it is unessential for marginalizing  $\sigma$  on the right in (8). The first term of the numerator is given by the data term  $p(b|\sigma, \hat{x}) = Z^{-1} \exp(-D)$ , where  $Z$  is a normalizing constant. Furthermore, it is reasonable to assume independency  $p(\sigma|x) = p(\sigma)$ . Thus, we obtain

$$p(\sigma|b, \hat{x}) \propto \frac{1}{Z} \exp(-D(\hat{x}; \sigma)) p(\sigma). \quad (9)$$

For the second term under the integral in (8), we compute

$$\log p(b, \sigma|x) \propto \log p(b|\sigma, x) + \log p(\sigma) \propto -D(x; \sigma) + \log p(\sigma) \quad (10)$$

using again  $p(\sigma|x) = p(\sigma)$ , and dropping the normalizing constant of the first term on the right, as explained above in the remark after eqns. (7). Furthermore, we can drop the last term  $\log p(\sigma)$  in (10) because it neither depends on  $x$ , nor does it contribute to the averaging of  $D(x; \sigma)$  with respect to  $\sigma$ .

As a result, we insert the remaining term  $-D(x; \sigma)$ , together with (9), into (8) and denote the resulting expression again with  $Q$ :

$$Q(x|\hat{x}, b) := \int_{\mathbb{R}_+} \frac{1}{Z} \exp(-D(\hat{x}; \sigma)) p(\sigma) (-D(x; \sigma)) d\sigma \tag{11}$$

This expression shows clearly how the unknown dependency on  $\sigma$  of the objective criterion (6) is dealt with: Given a current estimate  $\hat{x}$  and a prior distribution  $p(\sigma)$ , the unknown data term  $D(x; \sigma)$  is replaced by maximizing the average (11). Consequently, we replace the functional  $J_\mu(x; \sigma)$  in (6) by the approximation

$$E_\mu(x; \hat{x}) := -Q(x|\hat{x}, b) + \alpha S(x) - \mu \frac{1}{2} \langle x, x - e \rangle. \tag{12}$$

In practice, we choose the prior  $p(\sigma)$  to be uniform within a reasonable interval  $[\sigma_{\min}, \sigma_{\max}]$ , and  $\hat{x}$  is the current estimate on  $x$ .  $Q(x|\hat{x}, b)$  is then evaluated by computing the one-dimensional integral (11) numerically using the trapezoidal rule.

### 3.2 Smoothness Term

As smoothness prior  $S(x)$  in (12), we use a discrete approximation of the total-variation (TV) measure

$$\int_{\Omega} |\nabla x| d\Omega$$

of  $x$  (here temporarily regarded as a function), whose edge-preserving properties are well-known in image processing [10]. Recently, it has also been successfully used in connection with discrete tomography [11].

## 4 Optimization

The problem to minimize the functional  $E_\mu(x; \hat{x})$  in (12) over the convex set of feasible solutions  $B := [0, 1]^n$  can be written with a corresponding indicator function  $I_B(x) = 0$  if  $x \in B$  and  $I_B(x) = +\infty$  if  $x \notin B$ , as

$$\inf_{x \in \mathbb{R}^n} E_\mu(x; \hat{x}), \quad E_\mu(x; \hat{x}) = F(x; \hat{x}) - H_\mu(x), \tag{13}$$

where

$$F(x; \hat{x}) := -Q(x|\hat{x}, b) + \alpha S(x) + I_B(x)$$

is a proper lower-semicontinuous convex functional, and where

$$H_\mu(x) := \mu \frac{1}{2} \langle x, x - e \rangle$$

is convex as well, thus concave when subtracted in (13). Therefore, a natural minimization approach is DC (Difference of Convex functions) programming [12, 13].

To specify the algorithm, recall the following definitions from convex analysis [14] for a function  $f$ :

$$\begin{aligned} \text{dom}(f) &:= \{x \in \mathbb{R}^n \mid f(x) < +\infty\} && \text{effective domain of } f \\ \partial f(\bar{x}) &:= \{v \mid f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle\} && \text{subdifferential of } f \text{ at } \bar{x} \\ f^*(y) &:= \sup_{x \in \mathbb{R}^n} \{\langle x, y \rangle - f(x)\} && \text{conjugate function} \end{aligned}$$

We apply to (13) the following algorithm adopted from [13]:

**Subgradient Algorithm**

Choose  $x^0 \in \text{dom}(F)$  arbitrary (this choice does not dependent on the second argument of  $F$ ).

For  $k = 0, 1, \dots$  compute until convergence:

$$y^k \in \partial H_\mu(x^k) \tag{14}$$

$$x^{k+1} \in \partial F^*(y^k; \hat{x}) \tag{15}$$

The investigation of this algorithm in [13] includes the following results:

**Proposition 1.** [13] Assume  $F(\cdot; \hat{x}), H_\mu : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  to be proper, lower-semicontinuous and convex, and  $\text{dom}(F) \subset \text{dom}(H_\mu), \text{dom}(H_\mu^*) \subset \text{dom}(F^*)$ .

Then

- (i) the sequences  $\{x^k\}, \{y^k\}$  according to the equations (15) and (14) are well-defined,
- (ii)  $\{F(x^k; \hat{x}) - H_\mu(x^k)\}$  is decreasing,
- (iii) every limit point  $x^*$  of  $\{x^k\}$  is a critical point of  $E_\mu(x; \hat{x}) = F(x; \hat{x}) - H_\mu(x)$ .

**Remarks**

Concerning the full reconstruction algorithm, as listed on the subsequent page, we point out:

- Estimation of the unknown scale-parameter  $\sigma$  through the EM-iteration (cf. section 3.1) is done as part of step (15) – see lines 9-14 of the reconstruction algorithm listed on the following page.
- The global optimum of the convex optimization problem in line 11 of the reconstruction algorithm (cf. subsequent page) can be computed using any method. In our implementation, we used a dedicated algorithm [15] in view of the simple structure of the box-constraints  $x \in [0, 1]^n$ .

**Reconstruction Algorithm**

```

1 Choose  $x^0$  arbitrary (for example  $x^0 := (\frac{1}{2}, \dots, \frac{1}{2})^\top$ )
2 Choose  $\delta_\mu \in \mathbb{R}_+$  (our choice:  $\delta_\mu \in (0, 0.5]$ )
3 Choose  $\epsilon > 0$  (our choice:  $10^{-4} \leq \epsilon \leq 10^{-2}$ )
4 Set  $i := 0, \mu^0 := 0$ 
5 Do (μ-loop)
6   Set  $k := 0$ 
7   Do (DC-loop)
8      $y^k := \nabla H_{\mu^i}(x^k)$ 
9     Set  $l := 0, \hat{x}^0 := x^k$ 
10    Do (EM-loop)
11       $\hat{x}^{l+1} := \operatorname{argmin}_{x \in [0,1]^n} \{F(x; \hat{x}^l) - \langle y^k, x \rangle\}$ 
12       $l := l + 1$ 
13      while  $\|\hat{x}^l - \hat{x}^{l-1}\|_2 > \epsilon$  (EM-loop)
14       $x^{k+1} := \hat{x}^l$ 
15       $k := k + 1$ 
15      while  $\|x^k - x^{k-1}\|_2 > \epsilon$  (DC-loop)
16       $\mu^{i+1} := \mu^i + \delta_\mu$ 
17 while  $\exists x_j^k \in [\epsilon, 1 - \epsilon], j = 1, \dots, n$  (μ-loop)

```

**5 Evaluation**

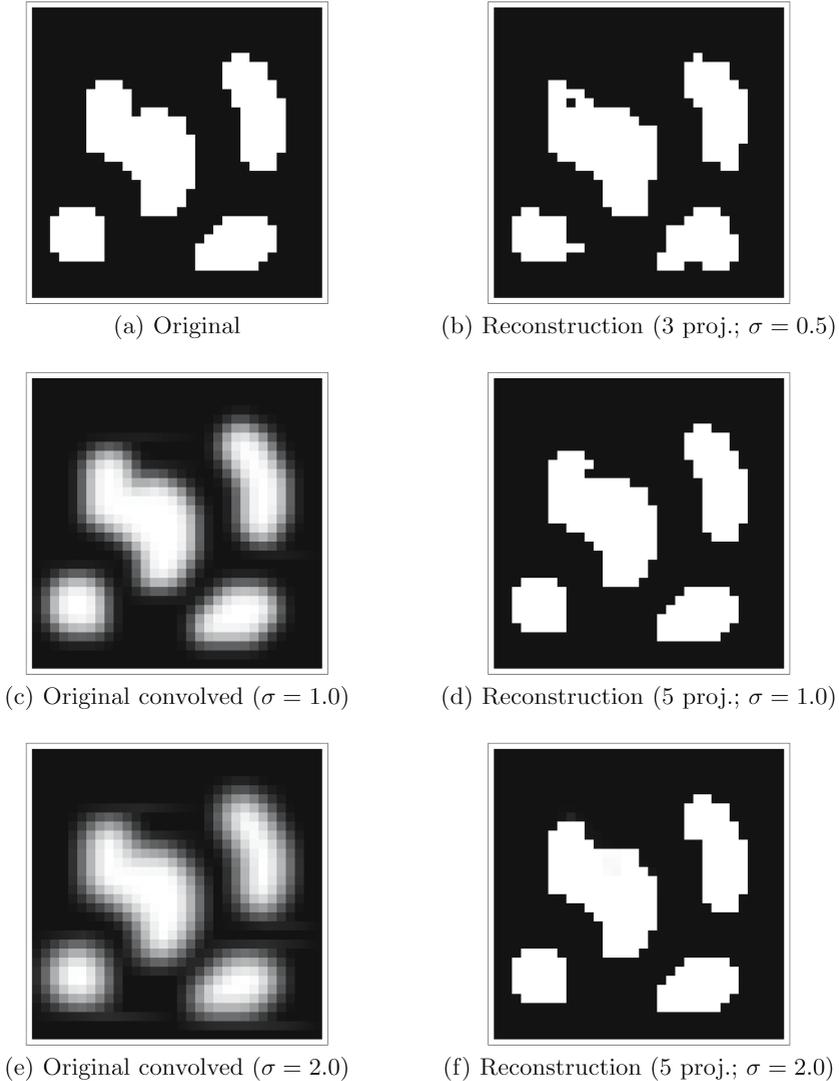
**5.1 Reconstruction from Projections of Blurred Objects**

In figure 1, we showed that binary reconstruction fails in case of blurred objects. We repeated the experiment, however, this time taking deblurring into account. The results shown in figure 3 reveal that our novel reconstruction algorithm copes with both problems, deblurring by scale-parameter estimation and binary reconstruction, at the same time.

Further experiments showed, that the original object can be reconstructed even with four projections only ( $0^\circ, 45^\circ, 90^\circ,$  and  $135^\circ$ , for  $\sigma = 1.0$ ).

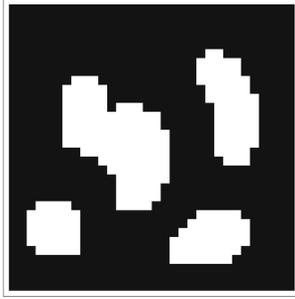
**5.2 Reconstruction from Blurred Projections**

The upper-left image shown in figure 4 was projected along for directions  $0^\circ, 45^\circ, 90^\circ, 135^\circ$ . Panel (b) shows these projections for illustration, and panel (d) the blurred version ( $\sigma = 1.5$ ). The latter data was used to compute the reconstruction shown in panel (c). Panels (e) and (f) show the reconstructions for  $\sigma = 1.0$  with and without deblurring, respectively. While the latter result clearly shows the ill-posedness of the combined deblurring-reconstruction problem, the results (c) and (e) demonstrate the stability of our new reconstruction algorithm even under such severe conditions.

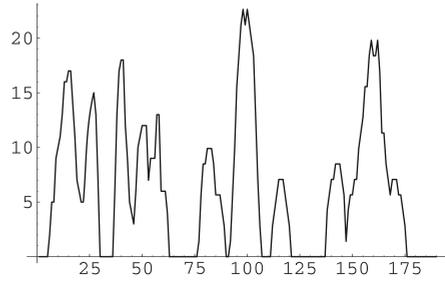


**Fig. 3. Reconstruction from blurred objects.** (a) Original image,  $32 \times 32$ . (c) and (e): original image convolved with different Gaussian kernels,  $\sigma \in \{1.0, 2.0\}$ . 5 projections were taken for both images  $0^\circ, 22.5^\circ, 45^\circ, 67.5^\circ, 90^\circ$ . Figures (d), and (f) show the corresponding results of our reconstruction algorithm. Since we obtained for  $\sigma = 0.5$  the original image we present in this case the reconstruction from only three projections,  $0^\circ, 45^\circ$ , and  $90^\circ$ . Throughout the experiments the smoothing parameter  $\alpha$  was set to 0.01.

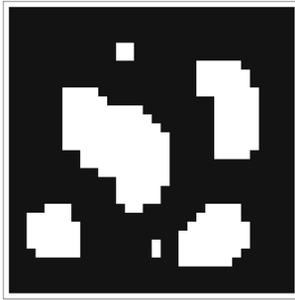
To illustrate the deblurring process further, figure 5 depicts the expressions  $\exp(-D(\hat{x}; \sigma))/Z$  and  $D(\hat{x}; \sigma)$ , respectively, as a function of  $\sigma$  during the reconstruction process. It can be clearly seen that the former expression peaks most



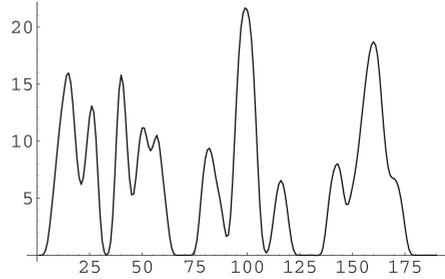
(a) Original



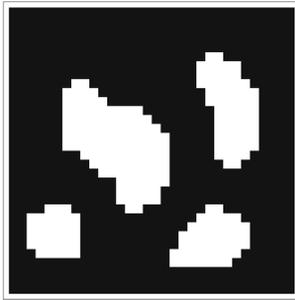
(b) Correct projections



(c) Reconstruction ( $\sigma = 1.5 \wedge \alpha = 0.05$ )



(d) Blurred projections ( $\sigma = 1.5$ )



(e) Reconstruction ( $\sigma = 1.0 \wedge \alpha = 0.01$ )

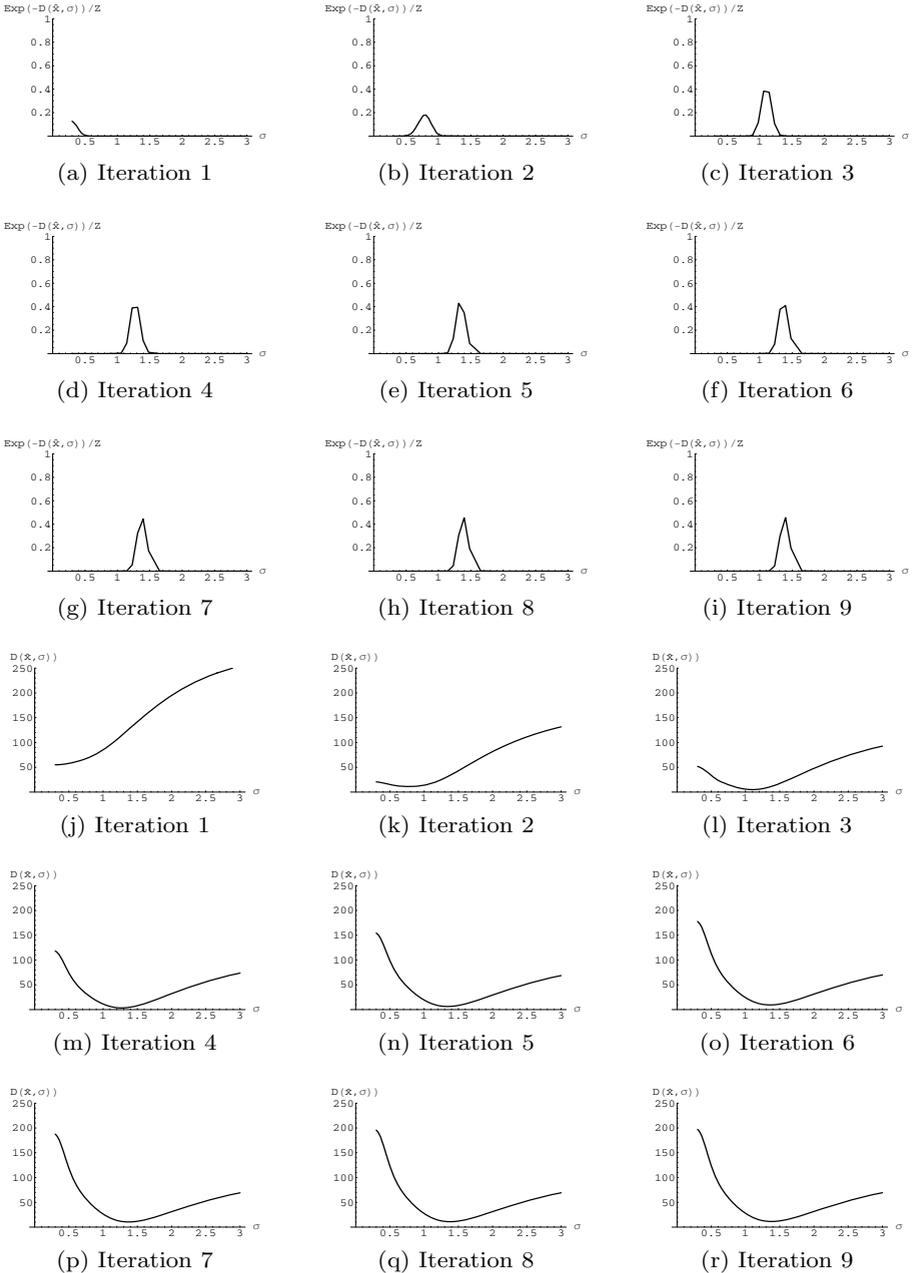


(f) Without deblurring ( $\sigma = 1.0 \wedge \alpha = 0.01$ )

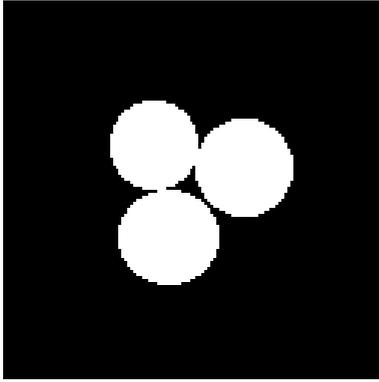
**Fig. 4. Reconstruction from blurred projections.** Projections at  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$  were taken from the image shown in panel (a) and convolved with a Gaussian kernel,  $\sigma = 1.5$ . Panels (b) and (d) show the correct projections and the blurred projections, respectively. Panel (c) shows the reconstruction result ( $\alpha = 0.05$ ). Panel (e) shows the reconstruction from projection data that were blurred with  $\sigma = 1.0$ . Panel (f) shows the erroneous reconstruction without taking deblurring into account.

around the correct value  $\sigma = 1.5$ , whereas the latter term attains its global minimum there.

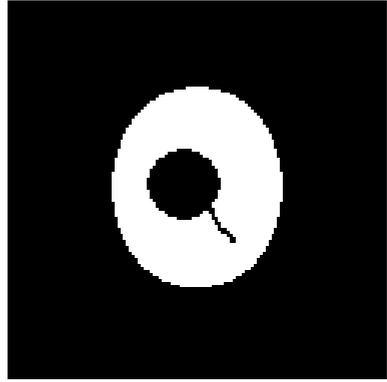
The experiments also revealed that reconstruction from blurred projections is more difficult than reconstruction from projections of blurred objects, as in the



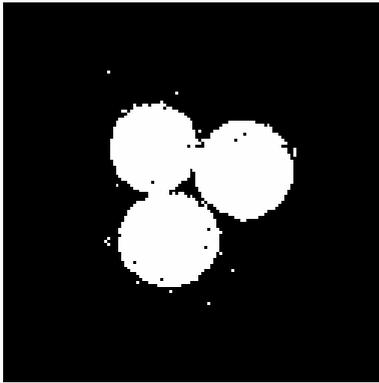
**Fig. 5. Upper half:** The term  $\exp(-D(\hat{x}; \sigma))/Z$  as a function of  $\sigma$  during the iteration. **Lower half:** The term  $D(\hat{x}; \sigma)$  as a function of  $\sigma$  during the iteration. While the former term peaks most near the correct value  $\sigma = 1.5$ , the latter attains its global minimum there. This illustrates that the inner EM-loop of the overall reconstruction algorithm is well-defined and robust.



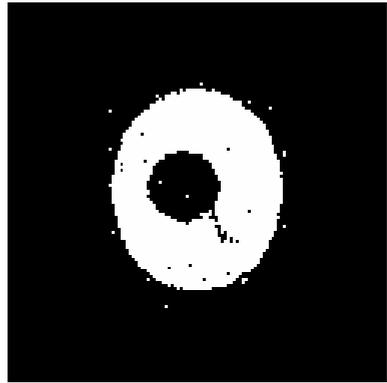
(a) Original



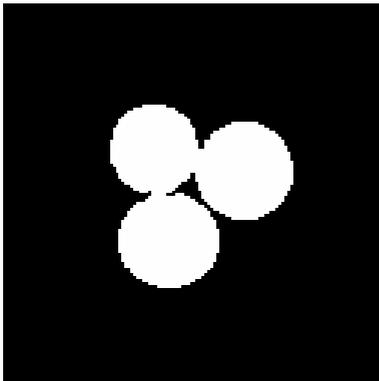
(b) Original



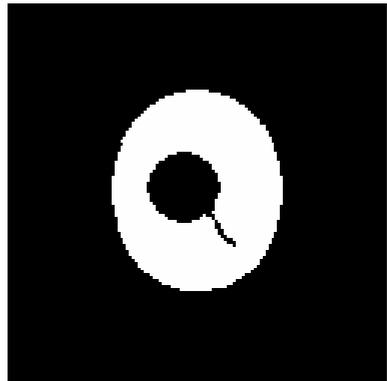
(c) Without deblurring ( $\sigma = 1.0$ )



(d) Without deblurring ( $\sigma = 1.0$ )



(e) With deblurring ( $\sigma = 1.0$ )



(f) With deblurring ( $\sigma = 1.0$ )

**Fig. 6. Reconstruction from blurred projections.** (a),(b) Original image,  $128 \times 128$ . For both images, reconstruction problems were set up using 5 projections,  $0^\circ$ ,  $36^\circ$ ,  $72^\circ$ ,  $108^\circ$ , and  $144^\circ$ , and blurring these projections with a Gaussian kernel,  $\sigma = 1.0$ . (c),(d) Reconstruction without deblurring. (e),(f) Reconstruction with deblurring.

former section: Perfect reconstruction of the original object was possible for a smaller blurring scale only ( $\sigma = 0.8$ ).

## 6 Conclusion and Further Work

We extended our reconstruction algorithm for binary tomography with an Expectation-Maximization (EM) step to improve its behavior in the presence of degradations during data acquisition. For evaluation purposes we defined two different degradation models. The same reconstruction algorithm can be applied to either of them which accurately estimates an unknown scale-parameter  $\sigma$ , during the reconstruction. Our results show that our approach stabilizes the reconstruction process in the presence of degradations.

Regarding the  $Q$  function in the EM-step, further work includes an adaptive sampling strategy of the supporting points. This is important for two reasons: First, it is expected to produce a more accurate approximation of the integral especially in areas where the true  $\sigma$  is suspected. Second, it should also reduce the number of supporting points since we can skip areas which are of low interest. The latter should further speed up our algorithm.

We suppose that our approach is sufficiently general to be applied to other combined reconstruction – missing parameter estimation scenarios as well. This will also be subject to our future work.

## Acknowledgments

The third author was supported during his stay in Mannheim by the Alexander von Humboldt-Foundation. This work was partially supported by the NSF Grant DMS 0306215 and the OTKA Grant T048476.

The CVGPR-group gratefully acknowledges support from Siemens Medical Solutions, Forchheim, Germany.

## References

1. Gonzales, R.C., Woods, R.E.: Digital Image Processing. 2 edn. Addison Wesley, Reading, MA (1992)
2. Herman, G.T., Kuba, A., eds.: Discrete Tomography: Foundations, Algorithms, and Applications. Birkhäuser Boston (1999)
3. Balaskó, M., Kuba, A., Nagy, A., Kiss, Z., Rodek, L., Ruskó, L.: Neutron-, gamma-, and X-ray three-dimensional computed tomography at the Budapest research reactor site. Nuclear Instruments and Methods in Physics Research A, **542** (2005) 22–27
4. Kuba, A., Ruskó, L., Rodek, L., Kiss, Z.: Preliminary studies of discrete tomography in neutron imaging. IEEE Trans. Nucl. Sci. **NS-52** (2005) 380–385
5. Carazo, J.M., Sorzano, C.O., Rietzel, E., Schröder, R., Marabini, R.: Discrete tomography in electron microscopy. In: Discrete Tomography. Foundations, Algorithms, and Applications (Eds.: Herman, G. T., Kuba, A.). Birkhäuser, Boston, MA (1999) 405–416

6. Schüle, T., Schnörr, C., Weber, S., Hornegger, J.: Discrete tomography by convex-concave regularization and d.c. programming. *Discrete Applied Mathematics* **151** (2005) 229–243
7. Weber, S., Schüle, T., Schnörr, C., Hornegger, J.: A linear programming approach to limited angle 3d reconstruction from dsa projections. *Special Issue of Methods of Information in Medicine* **4** (2004) 320–326
8. Schnörr, C., Schüle, T., Weber, S.: Variational Reconstruction with DC-Programming. In Herman, G.T., Kuba, A., eds.: *Advances in Discrete Tomography and Its Applications*. Birkhäuser, Boston (2006) To appear.
9. MacLachlan, G., Krishnan, T.: *The EM Algorithm and Extensions*. Wiley (1996)
10. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Physica D* **60** (1992) 259–268
11. Capricelli, T., Combettes, P.: Parallel block-iterative reconstruction algorithms for binary tomography. *Electr. Notes in Discr. Math.* **20** (2005) 263–280
12. Pham Dinh, T., Elberoussi, S.: Duality in d.c. (difference of convex functions) optimization subgradient methods. In: *Trends in Mathematical Optimization*, Int. Series of Numer. Math. Volume 84. Birkhäuser Verlag, Basel (1988) 277–293
13. Pham Dinh, T., Hoai An, L.T.: A d.c. optimization algorithm for solving the trust-region subproblem. *SIAM J. Optim.* **8**(2) (1998) 476–505
14. Rockafellar, R.T.: *Convex analysis*. 2 edn. Princeton Univ. Press, Princeton, NJ (1972)
15. Birgin, E.G., Martínez, J.M., Raydan, M.: Algorithm 813: SPG - software for convex-constrained optimization. *ACM Transactions on Mathematical Software* **27** (2001) 340–349