

Discrete Tomography by Continuous Multilabeling Subject to Projection Constraints

Matthias Zisler, Stefania Petra, Claudius Schnörr, and Christoph Schnörr

Heidelberg University

Abstract. We present a non-convex variational approach to non-binary discrete tomography which combines non-local projection constraints with a continuous convex relaxation of the multilabeling problem. Minimizing this non-convex energy is achieved by a fixed point iteration which amounts to solving a sequence of convex problems, with guaranteed convergence to a critical point. A competitive numerical evaluation using standard test-datasets demonstrates a significantly improved reconstruction quality for noisy measurements from a small number of projections.

1 Introduction

Computed tomography [14] deals after spatial discretization in an algebraic set-up with the reconstruction of 2D- or 3D-images $u \in \mathbb{R}^N$ from a small number of noisy measurements $b = Au + \nu \in \mathbb{R}^m$. The latter correspond to line integrals that sum up all absorptions over each ray transmitted through the object. A given projection matrix $A \in \mathbb{R}^{m \times N}$ encodes this imaging geometry. Applications range from medical imaging [3] to natural sciences and industrial applications, like non-destructive material testing [7]. Many situations require to keep the number of measurements as low as possible, which leads to a small number of projections and hence to a severely ill-posed reconstruction problem.

To cope with such problems, a common assumption in the field of *discrete tomography* [8] concerns knowledge of a *finite* range of $u \in \mathcal{L}^N$, $\mathcal{L} := \{c_1, \dots, c_K\} \subset [0, 1]$, that is, u represents a piecewise constant function. Our main concern in this paper is to effectively exploit the additional prior knowledge in terms of \mathcal{L} , besides the projection constraints, in order to solve the *discrete reconstruction problem*

$$Au = b \quad \text{s.t.} \quad u_i \in \mathcal{L}, \quad \forall i = 1, \dots, N, \quad (1)$$

which generally is a NP-hard problem.

Related work on discrete tomography considers either *binary* or *non-binary* (multivalued) problems. The latter ones are considerably more involved.

Regarding *binary discrete tomography*, Weber et al. [29, 21] proposed to combine a quadratic program with a non-convex penalty which gradually enforces binary constraints. More recently, Kappes et al. [9] showed how a binary discrete graphical model and a sequence of s-t graph-cuts can be used to take into account the affine projection constraints and to recover high-quality reconstructions.

Regarding *non-binary discrete tomography*, an extension of the latter approach is not straightforward due to the nonlocal projection constraints. Weber [27, Chapter 6] proposed a non-convex term for non-binary discrete tomography that we derive in a natural way in the present work. However, Weber’s approach differs with respect to the data term for the projection constraints, regularization and optimization, and additionally requires parameter tuning.

Because u is assumed to be piecewise constant, an obvious approach is to consider sparsity promoting priors. The authors of [23] proposed a dynamic programming approach for minimizing the ℓ_0 -norm of the gradient. However, the set \mathcal{L} of feasible intensities is not exploited. In the convex setting, the integrality constraints are dropped and priors like the ℓ_1 -norm or the total variation (TV) are used [22, 6, 5], with a postprocessing step to round the continuous solution to a piecewise constant one. This approach connects discrete tomography and the fast evolving field of *compressive sensing* with corresponding recovery guarantees [5]. Again, however, the prior information of the range of the image to be reconstructed is not involved in the optimization process. We focus next on methods that make use of the set \mathcal{L} during the reconstruction process.

Tuysuzoglu et al. [25] casted the non-binary discrete reconstruction problem into a series of submodular binary problems within an α -expansion approach by linearizing the ℓ_2 -fidelity term around an iteratively updated working point. This local approximation discards a lot of information, and a significantly larger number of projections is required to get reasonable reconstructions. Maeda et al. [12] suggested a probabilistic formulation which couples a continuous reconstruction with the Potts model. Alternating optimization is applied to maximize the *a posteriori* probability locally. However, there is no guarantee that these alternating continuous and discrete block coordinate steps converge.

Ramlau et al. [10] investigated the theoretical regularization properties of the piecewise constant Mumford-Shah functional [13] applied to linear ill-posed problems. In earlier work [19], they considered discrete tomography reconstruction using this framework. The difficult geometric optimization of the partition is carried out by a level-set approach and additionally the intensities \mathcal{L} were estimated in an alternating fashion. By contrast, our approach is based on a convex relaxation of the perimeter regularization and the set \mathcal{L} is assumed to be known beforehand.

Varga et al. [26] suggested a heuristic algorithm which is adaptively combining an energy formulation with a non-convex polynomial in order to steer the reconstruction towards the feasible values. Batenburg et al. [2] proposed the *Discrete Algebraic Reconstruction Technique (DART)* algorithm which starts with a continuous reconstruction by a basic algebraic reconstruction method, followed by thresholding to ensure a piecewise constant function. These steps interleaved with smoothing are iteratively repeated to refine the locations where u jumps. This heuristic approach yields good reconstructions in practice but cannot be characterized by an objective function that is optimized.

We regard [2, 26] as state-of-the-art approaches for the experimental comparison.

Contributions. We present a novel variational approach to the discrete tomography reconstruction problem in the general non-binary case. Contrary to existing work, we utilize both the non-local projection constraints and the feasible set of intensities \mathcal{L} in connection with an established convex relaxation of the multilabeling approach with a Potts prior. We show how the resulting non-convex overall energy can be optimized efficiently by a fixed-point iteration which requires to solve a convex problem at each step. In this way, the derivation of our non-convex data and its local updates arise naturally. We also propose a suitable rounding procedure as post-processing step, because the integrality constraints are relaxed. A comprehensive numerical evaluation demonstrates the superior reconstruction performance of our approach compared to related work.

2 Reconstruction by Constrained Multilabeling

In this section, we first reformulate the discrete reconstruction problem (1) as a constraint combinatorial multilabeling problem. Then we derive a tractable variational approximation and suggest a proper rounding procedure.

2.1 Constrained Multilabeling Problem

We assume that there are less measurements than pixels $m \ll N$ and hence that the discrete reconstruction problem (1) is ill-posed and requires regularization. A common choice is the Potts model [18], $R(u) = \|\nabla u\|_0 := |\{i \mid (\nabla u)_i \neq 0\}|$ for sparse gradient regularization which favours piecewise constant images. In presence of noisy measurements b , we use the more general constraints $\underline{b}(\epsilon) \leq Au \leq \bar{b}(\epsilon)$ instead of $Au = b$, where ϵ is an upper bound of the noise level. As a result, the discrete reconstruction problem can be rewritten as

$$E(u) = \lambda \cdot \|\nabla u\|_0 \quad \text{s.t.} \quad \underline{b}(\epsilon) \leq Au \leq \bar{b}(\epsilon) \quad \wedge \quad u_i \in \mathcal{L} \quad \forall i = 1 \dots N. \quad (2)$$

We refer to problem (2) as a *constrained multilabeling problem* with Potts regularization but point out that, from the viewpoint of graphical models, the system of affine inequalities induces (very) high-order potentials. This high-order interaction induced by the non-local constraints results in a non-standard labeling problem which becomes intractable for discrete approaches and larger problem sizes. We adopt, therefore, the strategy of solving a sequence of convex relaxations in order to minimize a non-convex energy, which properly approximates the original problem.

2.2 Approximate Variational Problem

Our starting point is the established convex relaxation of the multilabeling problem [30, 11, 17]. Minimizing the energy in (3a) below with respect to z over the set of relaxed indicator vectors (3b) assigns to each given image pixel from u^0 a label of the set $\mathcal{L} = \{c_1, \dots, c_K\}$. The discretized total variation, weighted

by λ , and the simplex constraints G constitute a basic convex relaxation of the integrality constraints with respect to z .

$$E(z, u^0) = \sum_{i=1}^N \sum_{k=1}^K z_{ik} (u_i^0 - c_k)^2 + \lambda \sum_{k=1}^K \|\nabla z_k\|_1 \quad (3a)$$

$$\text{s.t. } z \in G := \left\{ z \in [0, 1]^{N \times K} : \sum_{k=1}^K z_{ik} = 1, \forall i = 1, \dots, N \right\}. \quad (3b)$$

Regarding the notation, we denote by z_k , $k \in \{1, \dots, K\}$ the k -th *column vector* of z and by $z_{ik} = (z_k)_i$ the entries of the *matrix* z .

Next, we add the projection constraints $\underline{b} \leq Au \leq \bar{b}$ to the relaxed energy (3a) by transforming the indicator variables z back to their corresponding intensities with the linear operator $W: G \rightarrow \mathbb{R}^N$, $z \mapsto \sum_{l=1}^K c_l z_l$ which preserves convexity of the resulting energy

$$E(z, u^0) = \sum_{i=1}^N \sum_{k=1}^K z_{ik} (u_i^0 - c_k)^2 + \lambda \sum_{k=1}^K \|\nabla z_k\|_1 \quad (4)$$

$$+ \delta_{\mathbb{R}_+^m}(AWz - \underline{b}) + \delta_{\mathbb{R}_-^m}(AWz - \bar{b}) + \delta_G(z).$$

Note that the constraints $\underline{b} \leq AWz \leq \bar{b}$ and $z \in G$ are implemented by indicator functions $\delta_{\mathbb{R}_+^m}$ and $\delta_{\mathbb{R}_-^m}$.

In tomography, no image u^0 is given, however. Therefore, we cannot drop the unary data term in Eq. (4), since the constraints are feasible for all convex combinations of prototypes c_k . In other words, the constraints only constrain the value of a pixel but do not indicate how the indicator variables should realize this value (similar to estimating a vector given only its magnitude).

A straightforward approach would be to start with some initial guess u^0 , e.g. computed using some another reconstruction method, followed by iteratively applying this approach above. This gives the fixed point iteration

$$z^{n+1} = \arg \min_z E(z, Wz^n). \quad (5)$$

At every iteration a convex problem has to be solved whose solution updates the unary data term. This raises the question whether the iteration converges and which overall energy is actually optimized?

To address these questions, we first eliminate u^0 in a principled way. Note that $E(z, \cdot)$ in (4) is differentiable with respect to the second argument. We invoke Fermat's (first order) optimality condition $u^* = Wz$ which says that the optimal u^* must be equal to the weighted average of the labels c_k . Substituting this optimality condition back into the energy (4) results in the final version of the proposed energy which only depends on z ,

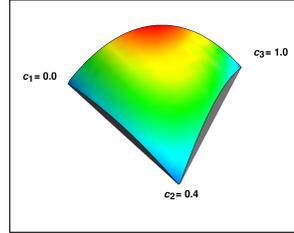
$$E(z) = \sum_{i=1}^N \sum_{k=1}^K z_{ik} ((Wz)_i - c_k)^2 + \lambda \sum_{k=1}^K \|\nabla z_k\|_1 \quad (6)$$

$$+ \delta_{\mathbb{R}_+^m}(AWz - \underline{b}) + \delta_{\mathbb{R}_-^m}(AWz - \bar{b}) + \delta_G(z).$$

This new energy function, Eq. (6), is non-convex because of the products in the first term which measures the discreteness of z . We call this term *phase data term* and denote it by

$$D(z) := \sum_{i=1}^N \sum_{k=1}^K z_{ik} ((Wz)_i - c_k)^2. \quad (7)$$

Fig. 1. Visualization of the phase data term $D(z)$ for $N = 1$ over the probability simplex G , the vertices correspond to the values $\mathcal{L} = \{0.0, 0.4, 1.0\}$. Note that the minimum is attained at the vertices of the simplex which correspond to unit vectors.



Using the notation introduced after Eq. (3), the i -th summand of (7) reads

$$\sum_{k=1}^K z_{ik} \left(\sum_{l=1}^K z_{il} c_l - c_k \right)^2 = \sum_{k=1}^K z_{ik} c_k^2 - \left(\sum_{k=1}^K z_{ik} c_k \right)^2 \quad (8)$$

which is concave with respect to the vector z_i . Consequently, $D(z)$ given by (7) is concave as well. Figure 1 shows a plot of $D(z)$. Weber [27, Chapter 6] proposed this term for discrete tomography which arises here in a natural way, whereas his overall approach differs with respect to data term for the projection constraints, regularization and optimization.

3 Optimization

In this section, we reformulate the objective function (6) as a DC program [15] and work out a corresponding optimization algorithm.

DC Programming. A large subclass of non-convex optimization problems are DC functions (difference of convex functions) which can be solved by DC Programming [15]. This generalizes subgradient optimization of convex functions to local optimization of DC functions. Accordingly, basic concepts of convex optimization like duality and KKT conditions were extended to DC functions [24]. The basic form of a DC program is given by

$$z^* = \arg \min_z g(z) - h(z), \quad (9)$$

where $g(z)$ and $h(x)$ are proper, lower semicontinuous, convex functions. There exists a simplified version of the DC algorithm [16] for minimizing (9) which guarantees convergence to a critical point by starting with $z^0 \in \text{dom}(g)$ and then alternatingly applying the updates

$$v^n \in \partial h(z^n) \quad \text{and} \quad z^{n+1} \in \partial g^*(v^n) \quad (10)$$

until a termination criterion is reached, where g^* denotes the Legendre-Fenchel conjugate [20] of g . To apply the DC algorithm to our non-convex energy $E(z)$ in (6), we rewrite $E(z) = g(z) - h(z)$ as a DC function. We set $h(z) = -D(z)$ since the phase data term (7) is concave by (8), and we denote by $g(z)$ the remaining convex terms from Eq. (6).

In order to make the step $z^{n+1} \in \partial g^*(v^n)$ explicit, we apply the subgradient inversion rule of convex analysis to obtain

$$z^{n+1} \in \partial g^*(v^n) \Leftrightarrow v^n \in \partial g(z^{n+1}) \Leftrightarrow 0 \in \partial g(z^{n+1}) - v^n \quad (11)$$

which is equivalent to the convex optimization problem

$$z^{n+1} = \arg \min_z g(z) - \langle v^n, z \rangle. \quad (12)$$

Because h is differentiable, the first step of (10) reads

$$v^n \in \partial h(z^n) \Leftrightarrow v^n = -\nabla D(z^n), \quad (13)$$

where the gradient of D at z for pixel i and label c_k is given by (see Lemma 1 from the supplementary material)

$$(\nabla D(z))_{ik} = \frac{\partial D(z)}{\partial z_{ik}} = ((Wz)_i - c_k)^2, \quad i = 1, \dots, N, \quad k = 1, \dots, K. \quad (14)$$

Combining equations (13) and (14) and inserting into equation (12) yields

$$z^{n+1} = \arg \min_z E(z, Wz^n) = \arg \min_z g(z) + \sum_{i=1}^N \sum_{k=1}^K z_{ik} ((Wz^n)_i - c_k)^2. \quad (15)$$

We notice that the DC algorithm, summarized as Algorithm 1 below, agrees with the iteration (5), and hence *proves its convergence*. We apply the primal

Algorithm 1: DC Fixed Point Algorithm

1. Initialization: choose any $z^0 \in \mathbb{R}^{n \times k}$
2. Generate a sequence $(z^n)_{n \in \mathbb{N}}$ by solving the convex problems

$$z^{n+1} = \arg \min_z E(z, Wz^n) \quad (16)$$

until a termination criterion is met.

dual (PD) algorithm proposed by [4] to solve each convex subproblem (16).

Rounding Step. Recall that the data term $D(z)$ of (6) only steers the solution to the finite set of feasible values \mathcal{L} . As a consequence, for vanishing regularization parameter λ , the minimizer z will correspond to indicator vectors

z_i that assign a unique label to each pixel i . For larger values of λ which are more common in practice, however, the minimizing vectors z_i will not be integral in general. Therefore, a post-processing step for rounding the solution is required.

Given the minimizer z^* of (6), we propose to select a label for each pixel i as a post-processing step by solving the local problems

$$\hat{u}_i^* = \arg \min_{c \in \mathcal{L}} |(Wz^*)_i - c|, \quad i = 1, \dots, N. \quad (17)$$

Note that this method differs from the common rounding procedure of multilabeling approaches which select the label c_k if $z_{ik} = \max\{z_{i1}, \dots, z_{iK}\}$.

4 Numerical Experiments

Set-up. In this section, we compare our approach to state-of-the-art approaches for non-binary discrete tomography in limited angles scenarios. Specifically, we considered the Discrete Algebraic Reconstruction Technique (*DART*) [2] and the energy minimization method from Varga et al. [26] (*Varga*). As multivalued

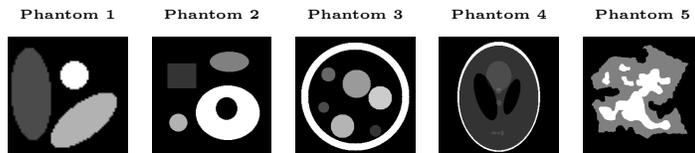


Fig. 2. The 5 different phantoms used for the numerical evaluation.

test-datasets we adapted the binary phantoms from Weber et al. [28] to more labels, shown as phantom 1,2 and 3 by Figure 2. Phantom 5 in Figure 2 was taken from [2] and phantom 4 is the well-known Shepp-Logan phantom. We created noisy scenarios by applying Poisson noise to the measurements b with a signal-to-noise ratio of $SNR = 20$ db. The geometrical setup was created by the ASTRA-toolbox [1], where we used parallel projections along equidistant angles between 0 and 180 degrees. Each entry a_{ij} of the matrix A corresponds to the length of the line segment of the i -th projection ray passing through the j -th pixel in the image domain. The width of the sensor-array was set 1.5 times the image size, so that every pixel intersects with a least a single projection ray.

Implementation details. Each subproblem of Algorithm 1 was approximately solved using the primal dual (PD) algorithm [4] limited to 1000 iterations or until the primal dual gap drops below 0.1. The outer iteration was terminated if the change of the energy between two subsequent iterations, normalized by the number of pixels, was smaller than 10^{-5} in the noiseless case and 10^{-4} in the noisy case. Additionally, we limited the number of outer iterations to 20. For DART we used the publicly available implementation included in the ASTRA-toolbox [1] and for the method of Varga [26] we used our own implementation

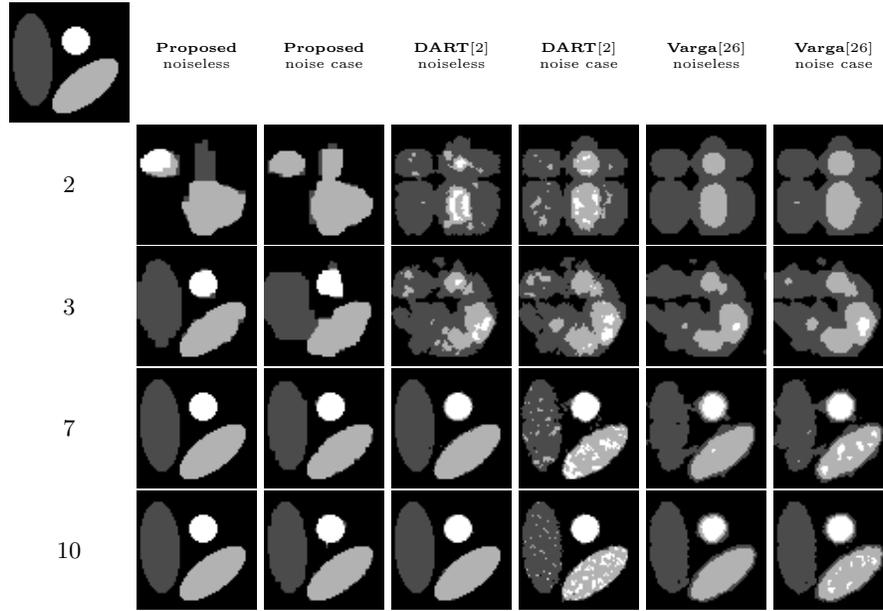
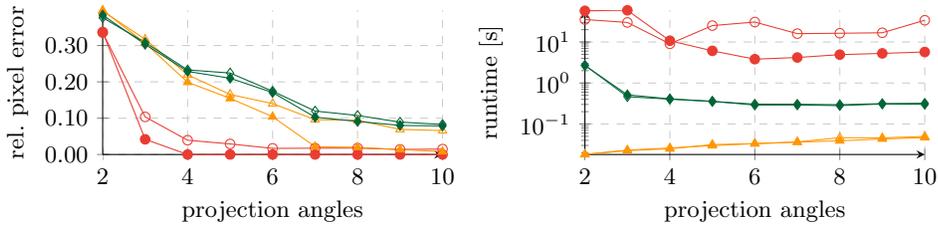


Fig. 3. Visual results of experiment phantom 1.

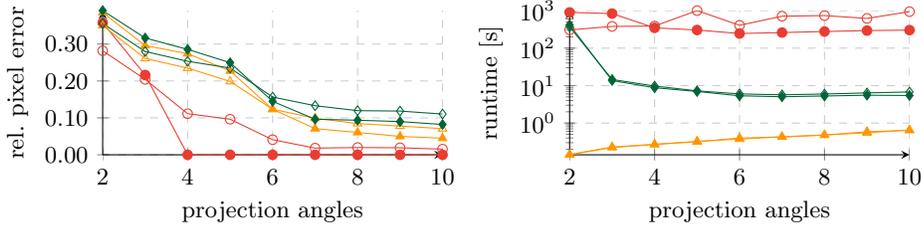
in MATLAB since no public code was available. We tried to use the default parameters of the competing approaches as proposed by their authors. However, since the test-datasets differ in size, we slightly adjusted the parameters in order to get best results for every algorithm and problem instance.

Performance measure. For the evaluation we measured the relative pixel error, that is the relative number of erroneously reconstructed pixels as compared to the groundtruth.

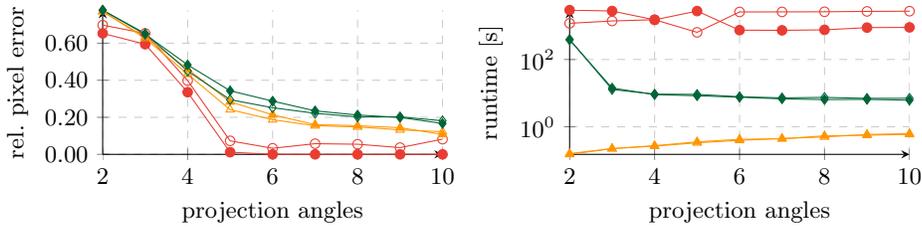
Results. Figure 4 shows all results of the numerical evaluation. For each test-datasets (phantoms 1 - 5), the left plot displays the relative pixel error for increasing numbers of projection angles. On the right, the corresponding runtime is shown as log-scaled plot. For each algorithm two curves are drawn: filled markers correspond to the noiseless case and non-filled markers correspond to the noisy case. The results show that the proposed approach returns a perfect reconstruction with the least number of projection angles in the noiseless case among all approaches. In the scenarios, the proposed algorithm is performing better, too. In the noiseless case phantom 1 can be almost perfectly reconstructed from only 3 projection angles and fully from 4 by the proposed approach whereas DART needs 7 projection to get an almost perfect reconstruction and the method of Varga needs at least 7 projections to get a reasonable reconstruction. These visual differences can be seen in Figure 3. This ranking of the performance of the approaches is similar for phantoms 2,3 and 5, except for the phantom 5 in the noiseless case where the approach of Varga performs better than DART. Figure 5 shows the results for phantom 4, where our approach is able to fully



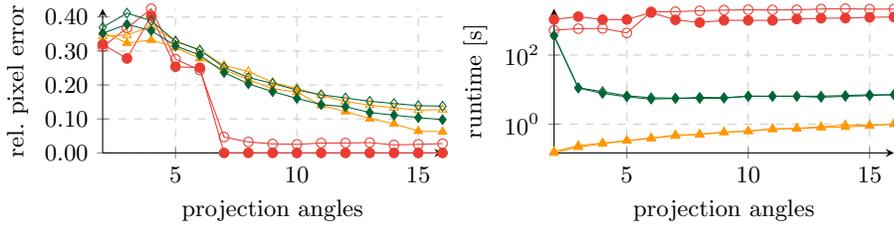
(a) Phantom 1



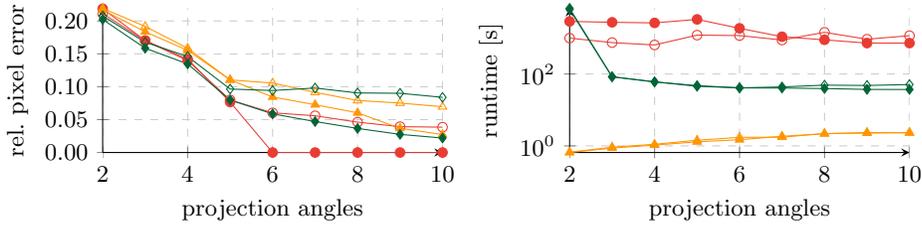
(b) Phantom 2



(c) Phantom 3



(d) Phantom 4



(e) Phantom 5



Fig. 4. Numerical evaluation of the approaches for the different test-datasets and increasing (but small) numbers of projections, in the noiseless case (filled markers) and in the noisy case (non-filled markers), with noise level $SNR = 20$ db. The relative pixel error is shown. The proposed approach gives perfect reconstructions with the least number of projection angles in the noiseless case and also returns high-quality reconstructions in the presence of noise, compared to the other approaches.

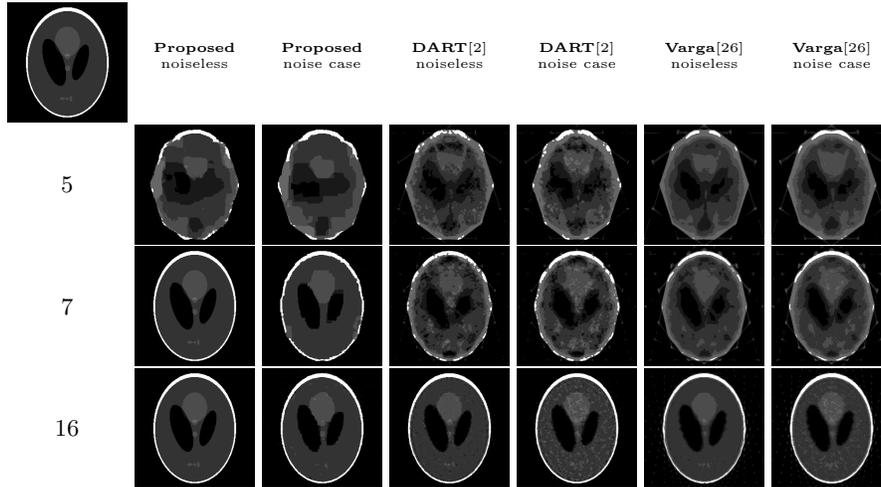


Fig. 5. Visual results of experiment phantom 4.

reconstruct from merely 7 projections in the noiseless case and returns a good piecewise-constant result in the noisy case. Due to a lack of space, we refer to Appendix B of the supplementary material for the visualization of the results for all test-datasets.

Regarding the runtime (right plots from figure 4), DART is the fastest approach, Varga is in between DART and the proposed approach, which is clearly consuming more runtime to return more accurate solutions. However, if computational performance is important, the proposed approach could be easily parallelized and implemented e.g. in CUDA to run on modern graphics cards.

5 Conclusion and Future Work

We presented a novel non-convex variational approach for solving the discrete tomography reconstruction in the general non-binary case. The approach combines a convex relaxation of the multilabeling problem with Potts prior and the non-local tomographic projection constraints. The feasible set of labels is taken into account by a non-convex data term which naturally emerges when the function to be reconstructed is represented as a convex combination of these values. A DC algorithm reliably minimizes the overall objective function and provably converges. The reconstruction performance turned out to be superior to the state of the art.

In future work, we plan to improve the running time and focus on the theoretical aspects of this approach. The proposed data term (7), in particular, fits nicely into spatially continuous variational formulations and thus may indicate ways for further improvement.

References

1. Aarle, W., Palenstijn, W., Beenhouwer, J., Altantzis, T., Bals, S., Batenburg, K., Sijbers, J.: The {ASTRA} Toolbox: A Platform for Advanced Algorithm Development in Electron Tomography. *Ultramicroscopy* 157, 35 – 47 (2015)
2. Batenburg, K., Sijbers, J.: DART: A Practical Reconstruction Algorithm for Discrete Tomography. *Image Processing, IEEE Transactions on* 20(9), 2542–2553 (Sept 2011)
3. Bushberg, J., Seibert, J., Leidholdt, E., Boone, J.: *The Essential Physics of Medical Imaging*. Wolters Kluwer, 3rd edn. (2011)
4. Chambolle, A., Pock, T.: A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging. *J. Math. Imaging Vis.* 40(1), 120–145 (May 2011)
5. Denitju, A., Petra, S., Schnörr, C., Schnörr, C.: Phase Transitions and Cospase Tomographic Recovery of Compound Solid Bodies from Few Projections. *Fundamenta Informaticae* 135, 73–102 (2014)
6. Goris, B., Broek, W., Batenburg, K., Mezerji, H., Bals, S.: Electron Tomography Based on a Total Variation Minimization Reconstruction Technique. *Ultramicroscopy* 113, 120 – 130 (2012)
7. Hanke, R., Fuchs, T., Uhlmann, N.: X-ray Based Methods for Non-Destructive Testing and Material Characterization. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 591(1), 14 – 18 (2008)
8. Herman, G., Kuba, A.: *Discrete Tomography: Foundations, Algorithms and Applications*. Birkhäuser (1999)
9. Kappes, J.H., Petra, S., Schnörr, C., Zisler, M.: TomoGC: Binary Tomography by Constrained GraphCuts. *Proc. GCPR* 30
10. Klann, E., Ramlau, R.: Regularization Properties of Mumford–Shah-Type Functionals with Perimeter and Norm Constraints for Linear Ill-Posed Problems. *SIAM Journal on Imaging Sciences* 6(1), 413–436 (2013)
11. Lellmann, J., Becker, F., Schnörr, C.: Convex Optimization for Multi-Class Image Labeling with a Novel Family of Total Variation Based Regularizers. In: *Proceedings of the IEEE Conference on Computer Vision (ICCV 09) Kyoto, Japan*. pp. 646–653 (2009), 1
12. Maeda, S., Fukuda, W., Kanemura, A., Ishii, S.: Maximum a posteriori X-ray computed tomography using graph cuts. *Journal of Physics: Conference Series* 233 (2010)
13. Mumford, D., Shah, J.: Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems. *Communications on Pure and Applied Mathematics* 42(5), 577–685 (1989)
14. Natterer, F., Wübbeling, F.: *Mathematical Methods in Image Reconstruction*. SIAM (2001)
15. Pham Dinh, T., El Bernoussi, S.: Algorithms for Solving a Class of Nonconvex Optimization Problems. *Methods of Subgradients*. In: Hiriart-Urruty, J.B. (ed.) *Fermat Days 85: Mathematics for Optimization*, North-Holland Mathematics Studies, vol. 129, pp. 249 – 271. North-Holland (1986)
16. Pham Dinh, T., Hoai An, L.: *Convex Analysis Approach to D.C. Programming: Theory, Algorithms and Applications*. *Acta Math. Vietnamica* 22(1), 289–355 (1997)

17. Pock, T., Chambolle, A., Cremers, D., Bischof, H.: A Convex Relaxation Approach for Computing Minimal Partitions. In: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. pp. 810–817 (June 2009)
18. Potts, R.B.: Some Generalized Order-Disorder Transformations. *Mathematical Proceedings of the Cambridge Philosophical Society* 48, 106–109 (1 1952)
19. Ramlau, R., Ring, W.: A Mumford–Shah Level-Set Approach for the Inversion and Segmentation of X-ray Tomography Data. *Journal of Computational Physics* 221(2), 539–557 (2007)
20. Rockafellar, R.T., Wets, R.J.B.: *Variational Analysis*, vol. 317. Springer Science & Business Media (2009)
21. Schüle, T., Schnörr, C., Weber, S., Hornegger, J.: Discrete Tomography by Convex-Concave Regularization and D.C. Programming. *Discrete Applied Mathematics* 151(13), 229 – 243 (2005)
22. Sidky, E.Y., Pan, X.: Image Reconstruction in Circular Cone-Beam Computed Tomography by Constrained, Total-Variation Minimization. *Physics in Medicine and Biology* 53(17), 4777 (2008)
23. Storath, M., Weinmann, A., Friel, J., Unser, M.: Joint Image Reconstruction and Segmentation Using the Potts Model. *Inverse Problems* 31(2), 025003 (2015)
24. Toland, J.F.: Duality in Nonconvex Optimization. *Journal of Mathematical Analysis and Applications* 66(2), 399–415 (1978)
25. Tuysuzoglu, A., Karl, W., Stojanovic, I., Castanon, D., Unlu, M.: Graph-Cut Based Discrete-Valued Image Reconstruction. *Image Processing, IEEE Transactions on* 24(5), 1614–1627 (May 2015)
26. Varga, L., Balázs, P., Nagy, A.: An Energy Minimization Reconstruction Algorithm for Multivalued Discrete Tomography. In: 3rd International Symposium on Computational Modeling of Objects Represented in Images, Rome, Italy, Proceedings (Taylor & Francis). pp. 179–185 (2012)
27. Weber, S.: Discrete Tomography by Convex-Concave Regularization Using Linear and Quadratic Optimization. PhD thesis, Ruprecht-Karls-Universität, Heidelberg, Germany (2009)
28. Weber, S., Nagy, A., Schüle, T., Schnörr, C., Kuba, A.: A Benchmark Evaluation of Large-Scale Optimization Approaches to Binary Tomography. In: *Discrete Geometry for Computer Imagery (DGCI 2006)*. LNCS, vol. 4245, pp. 146–156. Springer (2006)
29. Weber, S., Schnörr, C., Hornegger, J.: A Linear Programming Relaxation for Binary Tomography with Smoothness Priors. *Electronic Notes in Discrete Mathematics* 12, 243–254 (2003)
30. Zach, C., Gallup, D., Frahm, J., Niethammer, M.: Fast Global Labeling for Real-Time Stereo Using Multiple Plane Sweeps. In: *Proceedings of the Vision, Modeling, and Visualization Conference 2008, VMV 2008, Konstanz, Germany, October 8-10, 2008*. pp. 243–252 (2008)