

Simultaneous Reconstruction and Segmentation of CT Scans with Shadowed Data

François Lauze, Yvain Quéau, Esben Plenge

▶ **To cite this version:**

François Lauze, Yvain Quéau, Esben Plenge. Simultaneous Reconstruction and Segmentation of CT Scans with Shadowed Data. Sixth International Conference on Scale Space and Variational Methods in Computer Vision (SSVM), Jun 2017, Kolding, Denmark. pp.308-319, 10.1007/978-3-319-58771-4_25 . hal-02118564

HAL Id: hal-02118564

https://normandie-univ.hal.science/hal-02118564

Submitted on 3 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Simultaneous Reconstruction and Segmentation of CT Scans with Shadowed Data

F. Lauze[1] and Y. Quéau[2] and E. Plenge[1]

[1] Department of Computer Science, Universtity of Copenhagen
[2] Computer Vision Group, TU München

Abstract. We propose a variational approach for simultaneous reconstruction and multiclass segmentation of X-ray CT images, with limited field of view and missing data. We propose a simple energy minimisation approach, loosely based on a Bayesian rationale. The resulting non convex problem is solved by alternating reconstruction steps using an iterated relaxed proximal gradient, and a proximal approach for the segmentation. Preliminary results on synthetic data demonstrate the potential of the approach for synchrotron imaging applications.

1 Introduction

X-ray Computerised Tomography (CT) attempts at reconstructing a 3D structure from a set of planar projections that consist of attenuated measurements of X-rays across an object. In many domains, especially material sciences and geosciences, this reconstructed structure is often a building block for segmentation, which is further used to analyse some physical properties of the sample, such as porosity and tortuosity, or to perform some numerical simulations. Traditionally, and because of its low cost, Filtered Back Projection reconstruction (FBP) is performed, followed by segmentation. This has the disadvantage to propagate reconstruction errors in the segmentation. Other approaches have thus been proposed that do not decouple these steps. Discrete tomography [1] reconstructs an image with a small finite number of values, while Yoon et al. reconstruct a segmented image, using a level set approach [12]. In this work we follow a simultaneous reconstruction and segmentation (SRS) approach to produce both reconstructed image and its segmentation. This approach was first proposed by van Sompel and Brady [10], using a Hidden Markov Measure Field Model (HMMFM). Another was proposed by Ramlau and Ring in [8]. Romanov et al. [9] assumed that information about the segmentation is known in the form of prior knowledge on the parameters of a Gaussian Mixture Model for intensity distribution, and some partial convexification is performed. A Potts-Model based algorithm was proposed by Storah et al. in [11]. The present work is somewhat similar, though we place ourselve in a limited Field of View situation, where the detector extent is less than the object diameter. Algebraic Reconstruction Techniques (ART) are know to perform properly ([7], Chap VI), though an extra difficulty is that the signal may be *partially* blocked at certain angles, a situation

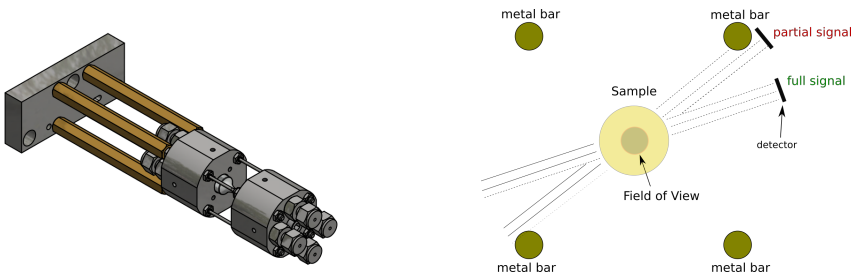


Fig. 1. Acquisition setup. Left: device holding the sample to be imaged. Right: limited Field of View and obstacle.

somewhat similar, but not identical to limited-angle tomography. FBP solvers can also reasonably deal with limited field of view by proper extension of the sinogram data (see also [7], Chap. VI), however, noise is still propagated and the presence of holes in the sinogram produces potentially severe streak artifacts in the reconstruction affecting thereafter the segmentation quality. One of the reasons for these streak artifacts is the fact that the absence of measurements is interpreted as a zero-measurement in FBP. Ongoing work on extending microlocal analysis [5] from the limited angle problem to this type of sinogram missing data is also providing clues on potential singularities arising from them.

This work is targeted towards Synchrotron X-Ray tomography, as often the case in material science. This means a simple, parallel beam geometry. We come back to the SRS problem where we attempt to estimate all parameters during the reconstruction and segmentation process. A clear modelling of what is available allows to properly take into account missing data, both for FoV and unrecorded data. Spatial regularisation, via Total Variation minimisation, and positivity constraints reduce noise and artifacts, providing an inpainting-like mechanism for the missing data. We couple it to a segmentation term which, in the reconstruction phase, also acts as a reaction towards specific values. Because out of the FoV, reconstructed values are not reliable, the segmentation is run only inside the FoV.

The paper is organised as follows. Section 2 formalises SRS with limited field of view and occluding geometry as a Bayesian inference problem. This yields an energy-minimisation formulation which is numerically tackled in Section 3. Starting from a rough initial reconstruction and segmentation obtained by standard methods, the proposed algorithm iteratively refines the reconstruction, given the current segmentation, and then the segmentation, given the current reconstruction. Both these steps are efficiently achieved by lagged proximal iterations. The preliminary results presented in Section 4 confirm that this joint approach is promising, in comparison with the standard sequential strategy.

2 Statement of the problem

For a monochromatic X-ray travelling along a line L, Lambert-Beer's law asserts that

$$I_L = I_0 \exp\left(-\int_L f(x)\,dx\right) \iff y_L := \log\frac{I_0}{I_L} = \int_L f(x)\,dx \tag{2.1}$$

where I_0 is the initial intensity and f the attenuation coefficient function. Recovering f accounts thus to solving (2.1) for f when we have observed $(y_L)_L$. More precisely, the *Radon transform* $R : L^1(\mathcal{D}) \to L^1(\mathbb{S}^1 \times \mathbb{R})$ of a function f is defined as

$$Rf(\theta, s) = \int_{\theta^\perp} f(s\theta + y)\,dy, \quad \theta \in \mathbb{S}^1, s \in \mathbb{R}. \tag{2.2}$$

with $\mathcal{D}$ a bounded domain of $\mathbb{R}^2$, $\mathbb{S}^1$ the unit circle of $\mathbb{R}$, and $\theta^\perp$ the line orthogonal to θ. The well celebrated inversion formula proven by Radon in 1917 states that when f is smooth with compact support, it can be recovered from its Radon transform by the *inverse Radon transform*:

$$f(x) = \frac{1}{4\pi^2}\int_{\mathbb{S}^1}\int_{-\infty}^{\infty} \frac{\frac{\partial}{\partial s}Rf(\theta, t)}{x\cdot\theta - t}\,dtd\theta. \tag{2.3}$$

This inversion is however ill-posed, and in our setting, we observe only $Rf(\theta, s)$ for values (θ, s) in a potentially complex sub-domain of $\mathbb{S}^1 \times \mathbb{R}$, as illustrated in Fig. 2. Filtered backprojection (FBP) is used to stabilise the inverse Radon transform, and relatively simple tricks can deal with the limited Field of View problem in the (FBP) settings. The shadowing caused by the presence of metallic bars induces however serious artifacts in FBP. This has motivated us to follow a discrete, direct approach.

Notations. We start by introducing notations used in the sequel. In the 2-dimensional setting, the sought function f can be viewed as a discrete $N\times N \to \mathbb{R}$ image $\boldsymbol{x}$ on a regular square grid, called *tomogram* and represented as a vector of $\mathbb{R}^{N^2}$. It is assumed to be large enough to cover the object we want to image. Each $\boldsymbol{x}_n$, $n = 1 \dots N^2$, satisfies

$$x_{\min} \leq \boldsymbol{x}_n \leq x_{\max} \tag{2.4}$$

with $x_{\min} \geq 0$, $x_{\max} \leq +\infty$. This is a simple box constraint and we denote by $\mathcal{C}$ the box $[x_{\min}, x_{\max}]^{N^2}$. In the 3D setting, $\boldsymbol{x}$ would be a $N \times B \times P$ discrete image, with similar box constraints. Yet, in this preliminary work we restrict ourselves to the 2D case.

Because we are in the parallel beam geometry setting, each angular projection is inherently 2D to 1D. Assume ℓ directions and a detector array of length d with M elements, where d is larger that the diameter of the object under scrutiny. Our aim is to recover the image $\mathbf{x}$ from samples of the Radon transform of the underlying attenuation function f, which form an $\ell \times M$ discrete image called *sinogram*, and represented as a vector $\boldsymbol{y} \in \mathbb{R}^{\ell M}$. The corresponding projection

matrix / discrete Radon Transform $\mathbf{R}$ is thus an element of $\mathbb{R}^{\ell M \times N^2}$ such that the sinogram is obtained as

$$\boldsymbol{y} = \mathbf{R}\boldsymbol{x}. \tag{2.5}$$

This is a discrete analogue to the standard FBP setting. We choose here a simple model where entry $\mathbf{R}_{ij}$ represents the contribution of grid element j to the line-ray encoded by i, as for instance the length of the intersection of this line and the grid element. Other less local representations can be chosen, see [4] for instance, for discussions on the generation of the discrete Radon transform matrices.

When the detector array has a length d' smaller than the object diameter, some parts of the object are not traversed by X-rays, or these rays are not recorded. If assuming the same resolution as for building $\mathbf{R}$, the resulting projection matrix $\mathbf{Q}$ is made of a subset of lines of $\mathbf{R}$. With m rays recorded per view, $\mathbf{Q}$ is a $\mathbb{R}^{\ell m \times N^2}$ matrix obtained from $\mathbf{R}$ by a linear projector $\mathbf{P}_F \in \mathbb{R}^{\ell m \times \ell M}$. The reduced sinogram $\boldsymbol{y}_F$ is $\boldsymbol{y}_F = \mathbf{P}_F \boldsymbol{y}$.

In the situation described in Figure 1, the geometry of the missing data in the sinogram will vary according to the ratio between bar radius and detector extent. This is easily computed by projecting orthogonally the disks along the detector, at each angle. Figure 2 illustrates this in two extreme cases, a very large and a very narrow detector.

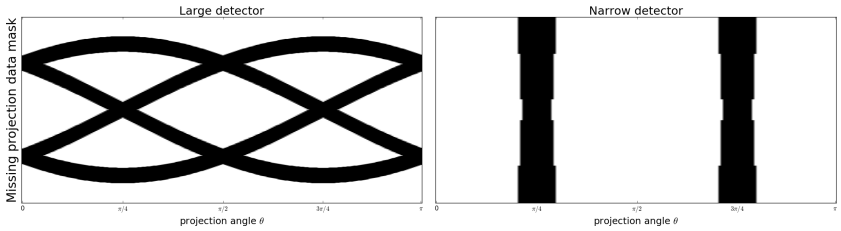


Fig. 2. Sinogram geometry induced by the occluding geometry of the bars of Figure 1.

Each point in the sinogram corresponds to a unique angle and detector position and thus to a line of the projection matrix. Lines corresponding to missing data points in the sinogram space should also be removed from the projection matrix. This can be modelled as another projector $\mathbf{P}_M \in \mathbb{R}^{S \times \ell m}$ with $S \leq \ell m$, and the predicted incomplete sinogram data $\boldsymbol{y}$ can be written as

$$\boldsymbol{y} = \mathbf{P}_M \boldsymbol{y}_F = \mathbf{P}_M \mathbf{P}_F \mathbf{R}\boldsymbol{x} = \mathbf{A}\boldsymbol{x}. \tag{2.6}$$

where $\mathbf{A}$ is of size $S \times N^2$. If the field of view is limited or if the geometry induces occlusions, then $S < N^2$ and the image $\mathbf{x}$ cannot be recovered directly from the sinogram $\mathbf{y}$ by inverting $\mathbf{A}$. It is necessary to introduce additional constraints: the Bayesian framework provides a natural framework for such a task. We thus use this framework to formulate the reconstruction problem, but also for segmenting the reconstructed image $\mathbf{x}$ i.e., grouping its pixels into different segments $\boldsymbol{\delta}$.

Bayesian Inference and MAP. Bayesian inference has been used in several SRS work to propose a posterior distribution of $\boldsymbol{x}$ and of a segmentation $\boldsymbol{\delta}$. In the limited field of view, and limited data problem we proceed the same way, incorporating the knowledge of the field of view F and occluding geometry M.

$$\begin{aligned} p(\boldsymbol{x}, \boldsymbol{\delta}|\boldsymbol{y}, F, M) &= \frac{p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\delta}, F, M)p(\boldsymbol{x}, \boldsymbol{\delta}|F, M)}{p(\boldsymbol{y}|F, M)} \\ &\propto p(\boldsymbol{y}|\boldsymbol{x}, F, M)p(\boldsymbol{x}, \boldsymbol{\delta}|F, M) \end{aligned} \tag{2.7}$$

The likelihood term $p(\boldsymbol{y}|\boldsymbol{x}, F, M)$ expresses how the observed sinogram data may deviate from the one predicted in (2.6). The conditional prior $p(\boldsymbol{x}, \boldsymbol{\delta}|F, M)$ links image values, segmentation and geometry of the field of view and occluding data. There is in general no hope to get a proper reconstruction out of the field of view. This is what Figure 3, which records the amount of recorded X-rays beams crossing each pixel, illustrates.

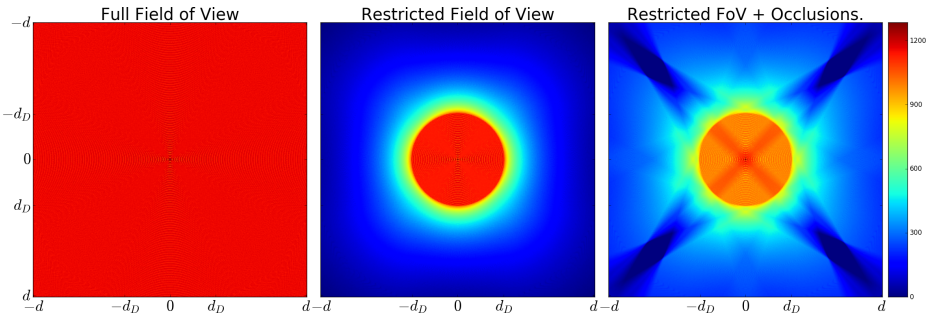


Fig. 3. X-ray visits of pixels during the acquisition.

It makes clear that between full field of view and restricted field of view (left and centre image), a large amount of information is lost, while this loss is much more moderate between centre and right image (restricted field of view and occluding geometry), and thus suggests that if a restricted field of view reconstruction is possible, then with proper priors, a reconstruction is also possible with (moderate) occluding geometry. So in order to reconstruct a tomogram and segment it, we choose a joint prior on $(\boldsymbol{x}, \boldsymbol{\delta})$ independent of M, i.e., we assume $p(\boldsymbol{x}, \boldsymbol{\delta}|F, M) = p(\boldsymbol{x}, \boldsymbol{\delta}|F)$. We may want to factorise it further by writing $p(\boldsymbol{x}, \boldsymbol{\delta}|F) = p(\boldsymbol{x}|\boldsymbol{\delta}, F)p(\boldsymbol{\delta}|F) = p(\boldsymbol{\delta}|\boldsymbol{x}, F)p(\boldsymbol{x}|F)$, but the unfactorised expression keeps the symmetry between $\boldsymbol{\delta}$ and $\mathbf{x}$.

A Maximum a Posteriori approach leads to minimisation of the neg-log-posterior

$$(\boldsymbol{x}^*, \boldsymbol{\delta}^*) = \operatorname{argmin}_{\boldsymbol{x}, \boldsymbol{\delta}} - \log p(\boldsymbol{y}|\boldsymbol{x}, F, M) - \log p(\boldsymbol{x}, \boldsymbol{\delta}|F). \tag{2.8}$$

Sinogram noise in X-ray tomography is usually modelled as Poisson, and in the high number of X-rays quanta, the limiting distribution can be considered

Gaussian [4], hence a least-squares term can be chosen for the first component in (2.8). The joint prior encodes both priors on $\boldsymbol{x}$ and $\boldsymbol{\delta}$ on F as well as their mutual dependence. We propose a discrete TV prior for $\boldsymbol{x}$, with box-positivity constraints. For the interplay between $\boldsymbol{x}$ and $\boldsymbol{\delta}$, a simple Pott's model term is used. And for the prior on $\boldsymbol{\delta}$, we ask smoothness of the segments within a HMMFM framework. For that purpose write $\boldsymbol{\delta} = (\boldsymbol{v}, \boldsymbol{c})$ where $\boldsymbol{v} : F \to \Delta_K$ is the label field, Δ_K is the standard simplex of $\mathbb{R}^K$ where K is the number of labels (supposed to be known in advance), $\boldsymbol{v}_n = (v_{n1}, \dots, v_{nK}) \in \Delta_k \iff v_{nk} \geq 0$, $k = 1 \dots K$ and $\sum_{k=1}^{K} v_{nk} = 1$. A discrete squared-gradient magnitude is used for regularisation of $\boldsymbol{v}$. Putting all pieces together, we obtain the following cost function:

$$\mathcal{E}(\boldsymbol{x}, \mathbf{c}, \mathbf{v}; \boldsymbol{y}) = \frac{1}{2}\|\mathbf{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \alpha J_F(\boldsymbol{x}) + \frac{\beta}{2}\left(\lambda \sum_{n \in F} \sum_{k=1}^{K} \mathbf{v}_{nk} \left(\boldsymbol{x}_n - \mathbf{c}_k\right)^2 + \frac{1}{2}\|\mathbf{D}\mathbf{v}\|_2^2\right), \\ 0 \leq \boldsymbol{x}_{\min} \leq \boldsymbol{x}_n \leq \boldsymbol{x}_{\max} \leq +\infty, \quad \mathbf{v}_n \in \Delta_K. \quad (2.9)$$

In this expression, $\mathbf{c} \in \mathbb{R}^K$ is the vector of mean segment values, $J_F(\boldsymbol{x})$ is the discrete TV-semi norm of $\boldsymbol{x}$ over F, $\mathbf{D}$ is a matrix representing the finite differences stencil used to approximate the gradient, and α, β and λ are user-defined hyper-parameters that need to be tuned appropriately.

3 Optimisation

The cost function is convex in $\boldsymbol{x}$, in $\mathbf{v}$ and in $\mathbf{c}$, but not jointly convex in $\boldsymbol{x}, \mathbf{v}, \mathbf{c}$. We propose to optimise it iteratively by alternating between updates of $\boldsymbol{x}$, $\mathbf{v}$ and $\mathbf{c}$. We first describe updates for the different arguments, from the simplest to the most complex one. Then we describe the full algorithm. As segmentation is difficult to obtain without reconstruction, a few iterations of a TV-reconstruction algorithm, followed by a K-means segmentation, are performed prior to start the joint optimisation. These iterations are actually simplifications of the joint optimisation when one or the other arguments are known.

3.1 The problem in c

The problem in $\mathbf{c}$ is classical and trivial: each $\mathbf{c}_k$ is given by

$$\mathbf{c}_k = \frac{\sum_{n=1}^{N^2} \mathbf{v}_{nk} \mathbf{x}_n}{\sum_{n=1}^{N^2} \mathbf{v}_{nk}}. \quad (3.1)$$

3.2 The problem in v

Call $U_K = \Delta_K^F$ the set of functions from F to Δ_K. After simplification of (2.9), the part of the cost function of interest in $\mathbf{v}$ is $\mathcal{F}(\mathbf{v}) = \frac{1}{2}\|\mathbf{D}\mathbf{v}\|_2^2 + \lambda \langle \mathbf{g}, \mathbf{v} \rangle + \iota_{U_K}(\mathbf{v})$ with $\mathbf{g} = \left((\boldsymbol{x} - c_1)^2, \dots, (\boldsymbol{x} - c_K)^2\right)^T$, $\langle -, - \rangle$ is the Euclidean inner product on

Algorithm 1 Sketch of the full algorithm.

Input: Sinogram $\boldsymbol{y}$, field of view F, sinogram mask M, number of classes K, weight parameters α, β, λ, ART parameter $\rho \in (0, 2)$, maximum number of iterations L_{RS}.
Output: Reconstruction $\boldsymbol{x}$ and segmentation $(\mathbf{v}, \mathbf{c})$ of $\boldsymbol{x}$ on F .
Initialisation: Run approximate reconstruction to produce $\boldsymbol{x}^0$. Run a K-means or Otsu clustering to produce $(\boldsymbol{c}^0, \mathbf{v}^0)$ from $\boldsymbol{x}^0$ and F.
for $i = 0$ to L_{RS} **do**
▷ **Solution in $\boldsymbol{x}$**
Solve for $\boldsymbol{x}^{i+1}$ from $\boldsymbol{x}^i$, $\mathbf{c}^i$ and $\mathbf{v}^i$.
▷ **Solution in c**
Solve for $\mathbf{c}^{i+1}$ from $\boldsymbol{x}^{i+1}$ and $\mathbf{v}^i$.
▷ **Solution in v**
Solve for $\mathbf{v}^{i+1}$ from $\boldsymbol{x}^{i+1}$ and $\mathbf{c}^{i+1}$.
end for

$\mathbb{R}^{N^2K}$ and ι_{U_K} the indicator function of U_K incorporating the constraints on $\mathbf{v}$. We write $h(\mathbf{v}) = \frac{1}{2}\|\mathbf{D}\mathbf{v}\|_2^2 + \lambda \langle \mathbf{g}, \mathbf{v} \rangle$. A classical solution is the proximal method which consists in computing $\mathbf{v}^{i+1}$ iteratively, by using the following lagged j-iterations:

$$\begin{aligned} \mathbf{v}^{j+1} &= \mathrm{prox}_{t_i \mathcal{F}}(\mathbf{v}^j) = \mathrm{argmin}_{\mathbf{v} \in U_K} \left\{ h(\mathbf{v}) + \frac{1}{2t_i}\|\mathbf{v} - \mathbf{v}^j\|_2^2 \right\} \\ &= \mathcal{P}_{U_k}\left(\mathbf{v}^j - t_i \tilde{\nabla} h(\mathbf{v}^{j+1}) \right) \end{aligned} \tag{3.2}$$

with t_i a diminishing gradient step, constant over a sweep i and set to $t_i = 1/(1+i)$, and $\tilde{\nabla} h(\mathbf{v}^{j+1})$ the subgradient of h at $\mathbf{v}^{j+1}$ [3]. As h is in fact smooth, the subgradient is just its usual gradient $\nabla h(\mathbf{v}^{j+1})$: $\nabla h(\mathbf{v}) = \mathbf{D}^T\mathbf{D}\mathbf{v} + \lambda \mathbf{g}$ with $-\mathbf{D}^T\mathbf{D}$ a discrete "vector Laplacian". In practice replace it by the two-steps method

$$\tilde{\mathbf{v}}^{j+1} = \left(\mathbf{D}^T\mathbf{D} + t_j^{-1}\, id_F \right)^{-1} \left(t_j^{-1}\mathbf{v}^j - \lambda \mathbf{g} \right) \tag{3.3}$$

$$\mathbf{v}^{j+1} = \mathcal{P}_{U_K}\left(\tilde{\mathbf{v}}^{j+1} \right). \tag{3.4}$$

Reflective boundary conditions are used to solve (3.3) while the projection onto U_K can for instance be implemented using classical simplex projection algorithms.

3.3 The problem in $\boldsymbol{x}$

With $\mathbf{c}$ and $\mathbf{v}$ fixed, the problem in $\boldsymbol{x}$ still presents two difficulties that prevent a direct approach: 1) the TV-seminorm term $J_F(\mathbf{x})$ and the size of the matrix $\mathbf{A}$. Row / block of rows action methods and iterated proximal algorithms provide a way to deal with it, following [6]. Before describing it, we rewrite the segmentation term in a more compact way.

Let $\Pi_F : \boldsymbol{x} \mapsto \boldsymbol{x}_{|F}$ the restriction to F. Let also N_F denote the amount of pixels in the field of view F, so that Π_F can be identified with a projection

$\mathbb{R}^{N^2} \to \mathbb{R}^{N_F}$. Then

$$\frac{1}{2}\sum_{n\in F}\sum_{k=1}^{K}\mathbf{v}_{nk}\left(x_n-\mathbf{c}_k\right)^2=\frac{1}{2}\|\mathbf{V}\Pi_F\boldsymbol{x}-L(\mathbf{V},\mathbf{c})\|_2^2 \tag{3.5}$$

where we have set $\mathbf{V}=(\mathbf{V}_1,\ldots,\mathbf{V}_K)^T$ with $\mathbf{V}_k=\mathrm{diag}(\mathbf{v}_{1k},\ldots,\mathbf{v}_{nk})^{\frac{1}{2}}$ and $L(\mathbf{V},c)=\left((\mathbf{c}_1\mathbf{V}_1 1_{N_F})^T,\ldots,(\mathbf{c}_K\mathbf{V}_K 1_{N_F})^T\right)^T$, $1_{N_F}=(1,\ldots,1)^T\in\mathbb{R}^{N_F}$. Note that $\mathbf{V}^T\mathbf{V}=\mathrm{id}_{\mathbb{R}^{N_F}}$. Dividing $\mathbf{A}$ in p blocks (one per viewing angle, if the view is non-empty), we have

$$\mathcal{E}(\boldsymbol{x})=\sum_{q=1}^{p}\frac{1}{2}\|\mathbf{A}_q\boldsymbol{x}-\boldsymbol{y}_q\|_2^2+\alpha J_F(\boldsymbol{x})+\frac{\beta\lambda}{2}\|\mathbf{V}\boldsymbol{x}-L(\mathbf{V},\mathbf{c})\|_2^2 \tag{3.6}$$

where we abusively denote $\mathbf{V}:=\mathbf{V}\Pi_F$ for sake of compactness.

We write $\mathcal{E}(\boldsymbol{x})=\sum_{q=1}^{p+2}f_q(\boldsymbol{x})$ with

$$f_q(\boldsymbol{x})=\frac{1}{2}\|\mathbf{A}_q\boldsymbol{x}-\boldsymbol{y}_q\|_2^2,\quad q=1\ldots p \tag{3.7}$$

$$f_{p+1}(\boldsymbol{x})=\alpha J_F(\boldsymbol{x}), \tag{3.8}$$

$$f_{p+2}(\boldsymbol{x})=\frac{\beta\lambda}{2}\|\mathbf{V}\boldsymbol{x}-L(\mathbf{V},\mathbf{c})\|_2^2. \tag{3.9}$$

Using one of the schemes (RIPG-I or RIPG-II) from [6], we obtain the following iterative scheme. With $\rho\in]0,2[$ and an initial estimate set for the i-th sweep, a full update on $\boldsymbol{x}^i$ is given as follows.
Set $\boldsymbol{x}_i^0=\boldsymbol{x}^i$. Then for $q=1$ up to $p+2$:

$$\begin{aligned}\mathbf{z}^q&=\mathrm{prox}_{t_i f_q}(\boldsymbol{x}_i^q)\\ \boldsymbol{x}_i^{q+1}&=\mathcal{P}_{\mathcal{C}}(\rho\mathbf{z}^q+(1-\rho)\boldsymbol{x}_i^q)\end{aligned}$$

and set $\boldsymbol{x}^{i+1}=\boldsymbol{x}_i^{p+2}$. The first p steps are updates from the Radon transform, step $p+1$ is a TV regularisation, while the last step is a reaction towards the current segmentation. $\mathcal{P}_{\mathcal{C}}$ is the box constraint projection from (2.4).

The proximality operators for the first p equations and the last one have the form $\mathrm{prox}_{t\varphi}(\boldsymbol{x})$ where $\varphi(\boldsymbol{x})$ is quadratic, $\varphi(\boldsymbol{x})=\frac{\delta}{2}\|\mathbf{H}\boldsymbol{x}-\mathbf{z}\|_2^2$ with $\mathbf{H}\in\mathbb{R}^{L\times N^2}$ and $\mathbf{z}\in\mathbb{R}^L$. By definition of the prox operator, one has

$$\mathrm{prox}_{t\varphi}(\boldsymbol{x})=\underset{\mathbf{u}\in\mathbb{R}^{N^2}}{\mathrm{argmin}}\left\{\frac{t\delta}{2}\|\mathbf{H}\mathbf{u}-\mathbf{z}\|_2^2+\frac{1}{2}\|\mathbf{u}-\boldsymbol{x}\|_2^2\right\}. \tag{3.10}$$

Writing the normal equations, one get

$$\mathrm{prox}_{t\varphi}(\boldsymbol{x})=\left(\mathbf{H}^T\mathbf{H}+\frac{1}{t\delta}\,\mathrm{id}_{\mathbb{R}^{N^2}}\right)^{-1}\left(\mathbf{H}^T\mathbf{z}+\frac{1}{t\delta}\boldsymbol{x}\right). \tag{3.11}$$

Using the classical relation valid for all $\tau \notin \mathrm{spec}(\mathbf{M}\mathbf{M}^T) \cup \{0\}$ and all $\mathbf{M} \in \mathbb{R}^{n\times m}$,

$$\left(\mathbf{M}^T\mathbf{M} + \tau\,\mathrm{id}_{\mathbb{R}^m}\right)^{-1}\mathbf{M}^T = \mathbf{M}^T\left(\mathbf{M}\mathbf{M}^T + \tau\,\mathrm{id}_{\mathbb{R}^n}\right)^{-1} \tag{3.12}$$

Eq. (3.11) can be rewritten as

$$\mathrm{prox}_{t\varphi}(\boldsymbol{x}) = \boldsymbol{x} - \mathbf{H}^T\left(\mathbf{H}\mathbf{H}^T + \frac{1}{t\delta}\,\mathrm{id}_{\mathbb{R}^L}\right)^{-1}(\mathbf{H}\boldsymbol{x} - \mathbf{z}) \tag{3.13}$$

and $\mathbf{H}^T\left(\mathbf{H}\mathbf{H}^T + \frac{1}{t\delta}\,\mathrm{id}_{\mathbb{R}^n}\right)^{-1}$ is a regularised or damped pseudoinverse of $\mathbf{H}$. We use it for the first p equations with $\mathbf{H} = \mathbf{A}_q$, $\delta = 1$, $t = t_i$. This correspond to a block-damped ART approach. With a standard discretization of the Radon transform, the image resolution being given by detector resolution, $\mathbf{A}_q\mathbf{A}_q^T$ is at most (and most often) tridiagonal, its values can be cached for reuse in subsequent updates.

A FISTA scheme [2] is used for the TV-regularization proximal step originbating from decomposition (3.8) (see also discussion in [6]).

For the segmentation reaction equation, $\mathbf{H} = \mathbf{V}\Pi_F$ and $\mathbf{H}^T\mathbf{H} = \Pi_F^T\mathbf{V}^T\mathbf{V}\pi_F = \Pi_F^T\Pi_F$ and $\boldsymbol{y} = \Pi_F^T\Pi_F\boldsymbol{x}$ is given by $\boldsymbol{y}_s = \boldsymbol{x}_s$ if $s \in F$ and $\boldsymbol{y}_s = 0$ if not. The proximal calculation is actually straightforward. Pixels out of F are not modified (in accordance of course with our prior hypothesis and the form (2.9) of the cost function), while in F, they are updated via the simple formula: the s-component of $\boldsymbol{z}^q$ is given by

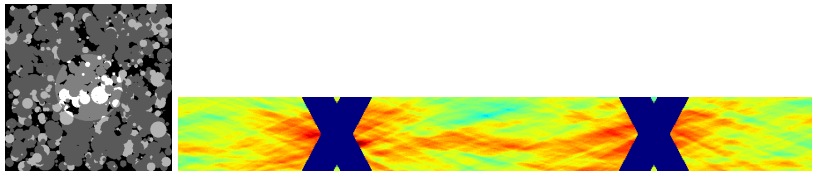
$$\boldsymbol{z}_s^q = \frac{\sum_{k=1}^K \mathbf{v}_{sk}\mathbf{c}_k + (t_i\beta\lambda)^{-1}\left(\boldsymbol{x}_i^q\right)_s}{1 + (t_i\beta\lambda)^{-1}} \tag{3.14}$$

4 Experiments

To evaluate our variational SRS strategy, we consider the synthetic dataset from Figure 4. We created a 2D tomogram of size $N \times N$ with $N = 300$, containing randomly generated circular patterns. To simulate various materials, each circular pattern was attributed one out of $K = 3$ scalar values within the set $\{0, 128, 255\}$. From this tomogram, we generated the corresponding sinogram using $l = 720$ directions (corresponding to a equally reparted 1D slices at every angle of 0.25 radians within $[0, \pi]$). To simulate limited field of view, only the central part of the tomogram is considered for reconstruction and segmentation, leading to a sinogram of size 720×282. Gaussian noise was then added to the sinogram. Eventually, some measurements in the sinogram were set to the null value, in order to simulate occluding geometry.

Figure 5 shows that standard FBP is not only sensitive to noise, but also efficient only when there is no missing data in the sinogram. This justifies the need for a joint reconstruction and segmentation strategy, instead of a sequential treatment.

Figure 6 presents our initial guess for the reconstruction (TV-reconstruction, without segmentation) and for the segmentation (K-means applied to the TV-reconstruction result). The results are already much better than the standard



(a) (b)

Fig. 4. Dataset used for evaluation. (a) Ground truth image, of size 1000×1000. The field of view is indicated by the central circle. (b) Sinogram computed from (a), of size 720×282, with missing data and additive, zero-mean, Gaussian noise (with standard deviation equal to 1% of the signal amplitude.

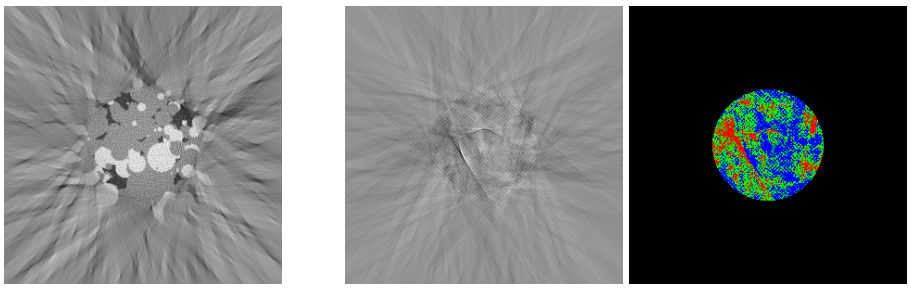
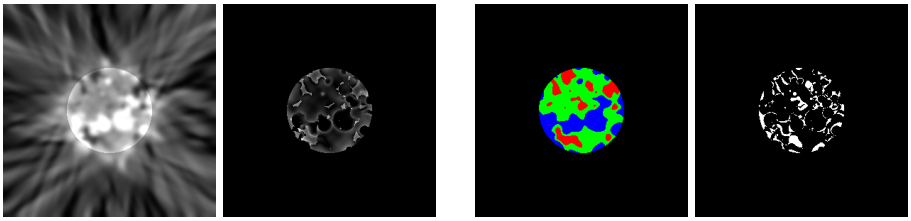


Fig. 5. Left: FBP reconstruction result, in the case where the sinogram does not contain missing data. Although it is quite sensitive to noise, the reconstruction is satisfactory and can be used for segmentation. Middle: same, when the sinogram contains missing data. Right: K-means segmentation applied to the middle reconstruction, which is not informative.

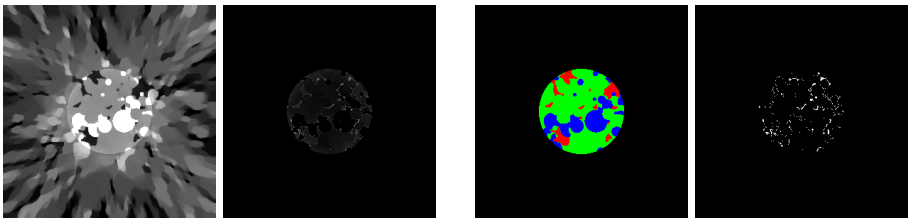
FBP approach of Figure 5, but thin structures are missed by the TV-reconstruction, which biases the subsequent segmentation.

Starting from this inital guess, we let our SRS algorithm iteratively refine the reconstruction and segmentation for 500 iterations, which required around 10 mn of computation on a recent I7 processor with 32 GB of RAM, using non-parallelised Python codes. Figure 7 shows that the final reconstruction and segmentation are much more accurate than the initial result from Figure 6. By plotting the evolution of the reconstruction and segmentation error as functions of the iteration number (Figure 8-a), we observe that most of the gain is achieved during the first iterations, a $\tau = 90\%$ segmentation accuracy being reached after only 50 iterations. This indicates that accurate results can be expected within a reasonable time, which could be further reduced using parallelisation. Next, we question in Figure 8-b and Figure 8-c the robustness of our method to the choice of parameters and to increasing noise level. The method is quite sensitive to the choice of parameters, but the same set of parameters can be used for a reasonable range of noise (here, it is only with $\sigma = 10\%$ that the reference set of parameters is inappropriate).



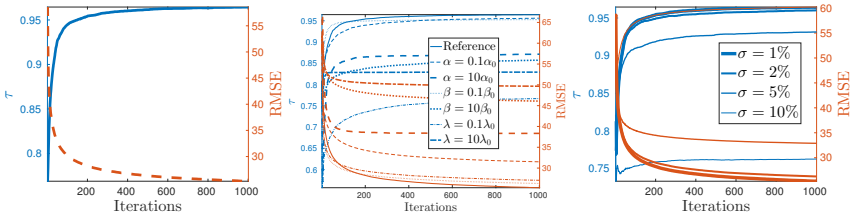
RMSE = 58.37 $\tau = 0.81$

Fig. 6. Initial 3D-reconstruction and segmentation. From left to right: TV-reconstruction of the image; absolute difference between (a) and ground truth (Fig. 4-a): black is 0, white is 255, and RMSE is the root mean square error; K-means Otsu clustering obtained from (a); classification map, where black indicates good classification, white indicates bad classification, and τ is the rate of good classification. Sharp structures are not very well recovered, and small areas are badly segmented.



RMSE = 25.28 $\tau = 0.96$

Fig. 7. Top: 3D-reconstruction and segmentation after 500 iterations, with the same convention as in Fig. 6. At convergence, fine structures are recovered and small areas are finely segmented.



(a) (b) (c)

Fig. 8. (a) Evolution of the RMSE on the reconstructed image, and of the good classification rate on the segmentation, as functions of the iteration. As the iterations go, both the 3D-reconstruction and the segmentation are improved. (b) Evaluation of the reconstruction and the segmentation as functions of iterations, for different sets of model parameters (noise level: $\sigma = 1\%$). (c) Ditto, with increasing noise level (the same reference parameters are used in all four experiments).

5 Conclusion

In this paper we proposed a joint image reconstruction and segmentation for Limited Field of View shadowed tomographic data. Data shadowing / occlusion makes it difficult, if not impossible, to recover a tomogram using the classical filtered backprojection approach. However, we showed that from an inverse problem viewpoint, when the amount of missing data is reasonable due to the shadowing effect, recovery is possible and by coupling it with segmentation, we avoid resolution loss in the latter stage.

Other types of segmentations can be used, the main memory and time bottleneck is the reconstruction part, though some method modifications may allow for a high degree of parallelism with reasonable block reconstruction approaches which open for fast GPU implementations. Investigating such acceleration techniques would be particularly worthwile regarding real-world applications to synchrotron imaging, which involve tremendous amount of data.

Aknowledgement

F. Lauze acknowledges funding from the Innovation Fund Denmark and Mærsk Oil and Gas A/S, for the P^3 Project. E. Plenge acknowledges funding from EUs 7FP under the Marie Skodowska-Curie grant (agreement no 600207), and from the Danish Council for Independent Research (grant ID DFF5054-00218).

References

1. Batenburg, K.J., Sijbers, J.: DART: a practical reconstruction algorithm for discrete tomography. IEEE Transactions on Image Processing 20(9), 2542–2553 (2011)
2. Beck, A., Teboulle, M.: Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. IEEE TIP 18(11), 2419–2434 (2009)
3. Bertsekas, D.: Incremental proximal methods for large scale convex optimization. Mathematical Programming, Ser. B. 129, 163–195 (2011)
4. Buzug, T.: Computed Tomography: from Photon Statistics to Modern Cone Beam CT. Sprin (2008)
5. Frickel, J., Quinto, E.T.: Characterization and reduction of artifacts in limited angle tomography. Inverse Problems 29(12) (December 2013)
6. andersen hansen:2014: Generalized row action methods for tomographic imagin. Numer. Algor 67, 121–144 (2014)
7. Natterer, F.: The Mathematics of Computerized Tomography, vol. 32. SIAM (1986)
8. Ramlau, R., Ring, W.: A mumford-shah approach for contour tomography. Journal of Computational Physics 221(2), 539–557 (February 2007)
9. Romanov, M., Dahl, A.B., Dong, Y., Hansen, P.C.: Simultaneous tomographic reconstruction and segmentation with class priors. Inverse Problems in Science and Engineering 24(8) (2016)
10. van de Sompel, D., Brady, M.: Simultaneous reconstruciton and segmentation algorithm for positron emission tomography and transmission tomography. In: Proceedings of the 2008 International Symposium on Biomedical Imaging (2008)
11. Storah, M., Weinmann, A., Frickel, J., Unser, M.: Joint image reconstruction and segmentation using the potts model. Inverse Problems 2(32) (February 2015)
12. Yoon, S., Pineda, A.R., Fahrig, R.: Simultaneous segmentation and reconstruction: A level set method approach for limited view computed tomography. Medical Physics 37(5), 2329–2340 (2010)