Tomographic reconstruction with spatially varying parameter selection

Yiqiu Dong\textsuperscript{1} and Carola-Bibiane Schönlieb\textsuperscript{2}

\textsuperscript{1} Department of Applied Mathematics and Computer Science, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
\textsuperscript{2} Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom

E-mail: yido@dtu.dk and cbs31@cam.ac.uk

Received 19 November 2018, revised 16 January 2020
Accepted for publication 4 February 2020
Published 22 April 2020

Abstract

In this paper we propose a new approach for tomographic reconstruction with spatially varying regularization parameter. Our work is based on the SA-TV image restoration model proposed by Dong \textit{et al} (2011 \textit{J. Math. Imag. Vis.} \textbf{40} 82–104) where an automated parameter selection rule for spatially varying parameters has been proposed. Their parameter selection rule, however, only applies if measured imaging data are defined in the image domain, e.g. for image denoising and image deblurring problems. By introducing an auxiliary variable in their model, we show here that this idea can indeed be extended to general inverse imaging problems such as tomographic reconstruction where measurements are not in the image domain. With a spatially varying regularization parameter, the new method can suppress artifacts due to limited data and noise while preserving more details. Using numerical simulations on synthetic and real data, we demonstrate the validity of the proposed approach and its effectiveness for computed tomography reconstruction, delivering reconstruction results that are significantly improved compared to the state-of-the-art.

Keywords: computed tomography, inverse problems, variational methods, spatially varying parameter, total variation regularization

1. Introduction

The computed tomography (CT) technique has been involved in many clinical and industrial applications. By sending x-rays through the object of interest, we are able to measure the reduction in intensity on the opposite side due to the attenuation of the x-ray when traveling through
the object. Then using reconstruction methods we will obtain a 2D or 3D image of the x-ray attenuation coefficients in the object.

Since x-rays travel along straight lines, mathematically CT scans are modeled by line integrals, and the corresponding reconstruction problem is to find the function from the knowledge of its line integrals for which an explicit inversion formula was derived in [2]. The assumption for obtaining this analytical solution is to have complete and clean data, i.e. the data includes the line integrals in the continuous setting from all directions and without noise. But in real applications the data are limited, and noise is unavoidable. In order to reconstruct high-quality images from a limited number of noisy measurements, several reconstruction methods are proposed in the literature, see e.g. [3–6] as well as the monograph [2] and the many references therein.

In recent years, much attention has been given to variational methods for CT reconstruction [7–11]. One of the most commonly used variational models for CT reconstruction, which is based on the maximum a posterior (MAP) estimator and the assumption of white additive Gaussian noise, is

\[
\min_{u \in L^2(\Omega)} \frac{\alpha}{2} \|Au - f\|^2 + R(u),
\]

where \(u \in L^2(\Omega)\) is the reconstructed image supported in an open subset \(\Omega = (-b, b)^2 \subset \mathbb{R}^2\), \(f \in L^2([0, \pi] \times (-b, b))\) denotes the CT data (often called the sinogram), and \(A : L^2(\Omega) \rightarrow L^2([0, \pi] \times (-b, b))\) is the x-ray transform, which is a bounded linear operator in \(L^2(\Omega)\). In this paper, we focus on the continuous setting where the data \(f\) and the reconstruction \(u\) are elements in infinite dimensional function spaces. For information on discrete models as well as continuous-discrete models for CT reconstruction we refer to [3]. In (1), \(R(u)\) is the regularization term, and \(\alpha > 0\) is the regularization parameter. Typical examples for \(R\) are Tikhonov regularization [12], total variation (TV) [13, 14], and several extensions [15–20]. The regularization parameter \(\alpha\) in (1) controls the trade-off between a good fit to the data and the regularization induced on a minimiser of (1) by \(R\). Note that strictly speaking the noise in CT data is Poisson due to the fact that the number of transmitted photons can be modeled as a Poisson process. For large incident photon flux, however, the model (1) which was derived with a white Gaussian distribution assumption on the noise can serve as a good approximation to the Poisson noise model. Moreover, the corresponding least-squares problem associated with the Gaussian assumption can be solved easily by a lot of efficient algorithms, so many reconstruction methods for CT are based on the model (1), e.g. [7, 9–11, 14, 19].

The choice of \(\alpha\) is critical for receiving a desirable reconstruction. If \(\alpha\) is chosen too large the reconstruction will be under-regularised and might still contain noise and other artifacts due to imperfections in the data. On the other hand, an \(\alpha\) that is chosen too small will render an over-regularized solutions in which structural information is lost. How to choose \(\alpha\), therefore, is an important and moreover challenging question. Even more so, as is the case in (1), when the regularization term is defined in the image domain but the data-fitting term is defined in the measurement space. A classic approach for regularization parameter selection in this case is Morozov’s discrepancy principle [21], which has been used for CT reconstruction, e.g. in [22, 23].

In this paper, we propose a variational method with spatially varying regularization parameter for CT reconstruction. The idea behind the spatially varying regularization parameter is that objects imaged with CT usually contain structures of different scales. In order to preserve textures with different scales in the CT image reconstruction while still removing noise and other artifacts, different regularization parameters should be assigned to different structural scales in the reconstruction. Our proposed method is based on the work in [1] for image
denoising with spatially adaptive total variation regularisation. Therein, the structural scales are defined in the image domain via a spatially varying regularisation parameter that is built into the data fitting term. Since in CT the data fitting term in (1) is defined in measurement space, it is not immediately clear how this approach can be applied to this case, or more generally to inverse imaging problems in which reconstruction (image) space and measurement space do not coincide. The variational model that we propose here circumvents this problem by introducing a new variable in (1) to split the tomographic reconstruction step and the spatially varying parameter estimation and regularisation step. This new variational model is introduced in (5) in section 3. By using the spatially varying regularization parameter, we will demonstrate with numerical examples that the new method provides much better CT reconstruction results compared to total variation regularisation with scalar $\alpha$. Moreover, the proposed model can be used for reconstruction problems for other inverse imaging problems.

Our paper is organized as follows. In section 2 we review the work on spatially varying regularization parameter selection in [1]. Section 3 describes our new variational method for CT reconstruction in detail. In section 4 we present numerical results from simulated and real measured data. Finally, conclusions are drawn in section 5.

2. Review of the SA-TV method

Since objects are usually comprised of structures with different scales, a locally varying regularization is desirable. In [1] a fully automated adjustment strategy for a spatially varying regularization parameter for image denoising and deblurring was proposed that is based on local variance estimators. The resulting regularization method is called the spatially adaptive total variation (SA-TV) method, and the corresponding variational model is

$$\min_{u \in \text{BV}(\Omega)} \frac{1}{2} \int_{\Omega} \lambda(x) |Ku(x) - z(x)|^2 \, dx + \int_{\Omega} |Du|,$$

where $K \in L(L^2(\Omega))$ is a blurring operator, $z \in L^2(\Omega)$ is a blurred image corrupted by additive white Gaussian noise with mean 0 and variance $\sigma^2$, $\lambda(x)$ is in $L^\infty(\Omega)$ and bounded by $[\varepsilon, \bar{\lambda}]$ with $\varepsilon > 0$, and $\text{BV}(\Omega)$ is the space of functions of bounded variation. Here, $u \in \text{BV}(\Omega)$ if $u \in L^1(\Omega)$ and its total variation (TV)

$$\int_{\Omega} |Du| = \sup \left\{ \int_{\Omega} u \, \text{div} \vec{v} \, dx : \vec{v} \in (C^\infty_0(\Omega))^2, \|\vec{v}\|_{L^\infty} \leq 1 \right\}$$

is finite, where $(C^\infty_0(\Omega))^2$ is the space of vector-valued functions with compact support in $\Omega$. The space $\text{BV}(\Omega)$ endowed with the norm $\|u\|_{\text{BV}(\Omega)} = \|u\|_{L^1(\Omega)} + \int_{\Omega} |Du|$ is a Banach space; see, e.g. [24]. The capability of the SA-TV method strongly depends on the correct selection of the parameter function $\lambda$.

In order to obtain a spatially varying $\lambda$, the idea behind the SA-TV method is to find $\lambda(x_0)$ for all $x_0 \in \Omega$ such that the corresponding restored image $u_{\lambda}$ satisfies the local constraint

$$\int_{\Omega_{x_0}} |Ku_{\lambda}(x) - z(x)|^2 \, dx \leq \sigma^2 |\Omega_{x_0}|,$$

where $\Omega_{x_0}$ denotes a subset of $\Omega$ with size $[-\frac{x_0}{2}, \frac{x_0}{2}] \times [-\frac{x_0}{2}, \frac{x_0}{2}]$ and centered at $x_0$, and $|\Omega_{x_0}|$ gives its area. Roughly, the constraint (3) means that in each local region $\Omega_{x_0}$ we expect that
the local variance of the residual is less than the noise variance, which can be understood as
claiming that for a correctly chosen $\lambda$ nearly only noise is left in the residual and consequently
the selected $\lambda(x_0)$ would automatically depend on the noise level and the scale of textures in
this region. For example, if $z$ is rather homogeneous in $\Omega_{\omega_0}$, then we expect that the constraint
(3) would be satisfied for a small $\lambda$; on the other hand, if $z$ features a lot of small scale textures
inside $\Omega_{\omega_0}$, a larger $\lambda$ is needed to preserve these textures in the restored image and only leave
the noise in the residual.

Because the decision on the acceptance or rejection of a local parameter value relies on the
scale of textures in the local region, the method potentially requires that the residual is defined
in the image domain and the textures in $z$ can be easily distinguished, i.e. the operator $K$ is
limited to small transformations in image space, e.g. a blurring operator that performs only
slight blurring. Due to this limitation, the SA-TV method cannot be directly applied to general
inverse problems where $z$ is not in the image domain.

3. Tomographic reconstruction method with spatially varying $\lambda$

In this section, we follow the same adjustment strategy for a spatially varying regularization
parameter as proposed in [1] and presented in section 2 but propose a novel extension of their
approach that can be applied to general inverse imaging problems. We exemplify our proposed
scheme for the problem of CT reconstruction. By using a spatially varying $\lambda$, which assigns
different values to different texture scales in the reconstruction, the new method can reconstruct
better small textures while still removing noise and other artifacts compared to L2-TV which
uses a fixed scalar $\lambda$, compare the numerical results in section 4.

Considering the TV regularization in the CT reconstruction model (1), we obtain

$$
\min_{u \in BV^+(\Omega)} \frac{\alpha}{2} \|Au - f\|^2 + \int_{\Omega} |Du|,
$$

(4)

where $BV^+(\Omega) = \{ u \in BV(\Omega) : u(x) \geq 0 \}$. In (4) the regularization is based on the smoothness
assumption on the reconstructed image $u$, and the data-fitting term arises from the CT forward
model

$$
f = Au + \epsilon
$$

with the assumption that the noise $\epsilon$ follows a Gaussian distribution with mean 0 and
variance $\sigma^2$. Since $A$ denotes the x-ray transform, i.e. a line integral operator, each point in
the sinogram $f$ is the value of a line integral. This value is made up of global information, such
that $f$ is smoother than $u$ and in particular, carries different singularities compared to $u$, see the
detailed study using microlocal analysis in [25]. Due to this disparity between the properties of
the data $f$ and the reconstruction $u$, we cannot select the regularization parameter in the same
way as in (3) according to the scale of textures in the residual.

To circumvent this problem, we introduce a new variable $w$ and split the data-fitting term and regularization term in (4). The new variational model that we propose is as follows:

$$
\min_{u, w \in BV^+(\Omega)} J(u, w) := \frac{\alpha}{2} \|Aw - f\|^2 + \frac{1}{2} \int_{\Omega} \lambda(x)(w(x) - u(x))^2 \, dx + \int_{\Omega} |Du|,
$$

(5)
where $\alpha > 0$ and $0 < \varepsilon \leq \lambda(x) \leq \bar{\lambda}$. Because the sinogram $f$ shows global information on the objects, it is difficult to define a local constraint for choosing a locally different $\alpha$. With the extra variable $w$, however, the new added quadratic term is defined in the image domain and $w$ can take over the former role of $z$ in (3). We therefore can introduce a spatially varying parameter based on local image textures in $w$. Note that in the new model the residual $w - u$ generally carries the artifacts from the CT reconstruction due to the noise in the sinogram and the image textures.

Since $\lambda(x) \geq \varepsilon > 0$, the model is strictly convex. Hence, existence and uniqueness of solutions to (5) is a straightforward exercise, whose result is summarised in the following theorem.

**Theorem 3.1.** Let $f$ be in $L^2([0, \pi) \times (-b, b))$, $A: L^2(\Omega) \to \in L^2([0, \pi) \times (-b, b))$ be the x-ray transform, which is a bounded linear operator, and $\lambda \in L^\infty(\Omega)$ be bounded in $[\varepsilon, \bar{\lambda}]$. Then, the model (5) has a unique solution.

To solve the minimization problem in (5) numerically, we consider its discretized version

\[
\min_{u, w \in \mathbb{R}^{m \times n}} J^D(u, w) := \frac{\alpha}{2} \|Aw - f\|^2_F + \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \lambda_{ij} (w_{ij} - u_{ij})^2 + TV(u),
\]  

(6)

where we define $\mathbb{R}^{m \times n} = \{z \in \mathbb{R}^{m \times n} : z_{ij} \geq 0 \text{ for all } i = 1, \ldots, m, j = 1, \ldots, n\}$ and the image size as $m$-by-$n$. For the sake of simplicity, we keep the same notations from the continuous context. TV under discrete setting is given by

\[
TV(u) = \sum_{i=1}^m \sum_{j=1}^n \sqrt{(\nabla_{x1} u_{ij})^2 + (\nabla_{x2} u_{ij})^2},
\]

where $\nabla_{x1}$ and $\nabla_{x2}$ denote the derivatives along horizontal and vertical directions, respectively. Then, we use an alternating optimisation algorithm, which starts from an initial guess $u^0 \in \mathbb{R}^{m \times n}$ and follows an iterative scheme:

\[
u^{k+1} = \arg \min_{u \in \mathbb{R}^{m \times n}} \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \lambda_{ij} (w_{ij}^k - u_{ij})^2 + TV(u),
\]

(7)

\[
u^{k+1} = \arg \min_{w \in \mathbb{R}^{m \times n}} \frac{\alpha}{2} \|Aw - f\|^2_F + \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \lambda_{ij} (w_{ij} - u^{k+1}_{ij})^2,
\]

(8)

which is solved until numerical convergence, i.e. the difference of iterates is smaller than a prescribed tolerance.

In (7), the solution satisfies the minimum–maximum principle, i.e. $\min w^k \leq u^{k+1} \leq \max w^k$, and we know that $u^k \in \mathbb{R}^{m \times n}$ according to the constraint in (8), then the $u$-subproblem is equivalent to

\[
u^{k+1} = \arg \min_{u \in \mathbb{R}^{m \times n}} \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n \lambda_{ij} (w_{ij}^k - u_{ij})^2 + TV(u),
\]

(9)

which is identical to the model (2) in the SA-TV method under the discrete setting, with $K$ being the identity operator. Hence, we can apply the SA-TV method proposed in [1] to solve the $u$-subproblem with automatically adjusted regularization parameter $\lambda$. The $w$-subproblem in (8) is a least-squares problem, and we can solve it efficiently by using, for example, the CGLS.
Algorithm 1. SA-TV for image reconstruction from indirect measurements.

1: Input $f$ and initialize $w^0$.
2: For $k = 1, 2, \ldots, k_0$, calculate iteratively $u^{k+1}$ and $w^{k+1}$. Output the final $\lambda$ from the SA-TV method for obtaining $u^{k+1}$ in each iteration, and set as $\lambda^{k+1}$.

\[ u^{k+1} = \arg \min_{u \in [0, \infty)^m} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_i (u^{k+1}_{ij} - u_{ij})^2 + TV(u), \]

\[ w^{k+1} = \arg \min_{w \in \mathbb{R}^{m \times n}} \frac{\alpha}{2} \|Aw - f\|_F^2 + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_i (w_{ij} - u_{ij})^2, \]

3: For $k = k_0 + 1, \ldots$, calculate iteratively $u^{k+1}$ and $w^{k+1}$.

\[ u^{k+1} = \arg \min_{u \in [0, \infty)^m} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_i (u^{k+1}_{ij} - u_{ij})^2 + TV(u), \]

\[ w^{k+1} = \arg \min_{w \in \mathbb{R}^{m \times n}} \frac{\alpha}{2} \|Aw - f\|_F^2 + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_i (w_{ij} - u_{ij})^2. \]

4: If $u^{k+1}$ satisfies the stopping criteria, it returns $u^{k+1}$ and stops.

method [26] followed by a non-negativity projection. The selection of the parameter $\alpha$ is done by using the discrepancy principle. In order to reduce the computational complexity, after a few iterations we stop the automatic selection of $\lambda$ in the SA-TV method during solving the subproblem (9). That is, in our method $\lambda$ is only automatically selected in the first $k_0$ iterations, after that we keep $\lambda$ fixed. According to our numerical tests, in most cases $\lambda$ stops changing significantly after 5 iterations, hence we set $k_0 = 5$. The overall algorithm is given in algorithm 1.

Another advantage of our approach of keeping $\lambda$ fixed after $k_0$ iterations is that the convergence of the sequence $\{u^k, w^k\}$ is guaranteed according to

\[ J^D(u^{k+1}, w^{k+1}) \leq J^D(u^{k+1}, w^{k}) \leq J^D(u^k, w^k) \]

as well as the fact that $J(u, w)$ is bounded below by zero. Then, taking the strict convexity of the objective function $J$ into account, we obtain the convergence of the algorithm based on the convergence results from [27].

**Theorem 3.2.** Assume that $u^0 \in \mathbb{R}^{m \times n}$ and $\lambda \in \mathbb{R}^{m \times n}$ is fixed and bounded in $[\varepsilon, \bar{\lambda}]$. Then the sequence $\{u^k, w^k\}$ generated by algorithm 1 converges sublinearly to the unique minimizer of the problem (6).

**Proof.** Set

\[ F(u, w) = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_i (w_{ij} - u_{ij})^2, \]

\[ G_1(u) = TV(u), \]

\[ G_2(w) = \frac{\alpha}{2} \|Aw - f\|_F^2. \]

Obviously, the functions $G_1$ and $G_2$ are closed and proper convex subdifferentiable functions. In addition, the function $F$ is a continuously differentiable convex function. Since $\lambda$ is bounded
in \([\varepsilon, \bar{\lambda}]\), the function \(\mathcal{F}\) is Lipschitz continuous with respect to both \(u\) and \(w\) with the same Lipschitz constant \(\bar{\lambda}\).

Based on theorem 3.1, \(\mathcal{J}^D\) has a unique minimizer. We define its minimizer as \((u^*, w^*)\) and its minimum as \(\mathcal{J}^{D*}\). Further, we define the ‘diameter’ of the sublevel set of \(\mathcal{J}^D\) at \(\mathcal{J}^D(u^0, w^0)\) as

\[
R = \max_{u, w \in \mathbb{R}^{m \times n}} \left\{ \|u - u^*\|_F + \|w - w^*\|_F : \mathcal{J}^D(u, w) \leq \mathcal{J}^D(u^0, w^0) \right\}.
\]

According to theorem 3.7 in [27], the sequence \(\{(u^k, w^k)\}\) generated by algorithm 1 satisfies the following inequality for any \(k \geq 1\):

\[
\mathcal{J}^D(u^k, w^k) - \mathcal{J}^{D*} \leq \frac{3 \max\{\mathcal{J}^D(u^0, w^0) - \mathcal{J}^{D*}, \bar{\lambda}R^2\}}{k}.
\]

That is, the sequence \(\{(u^k, w^k)\}\) converges to the unique minimizer of \(\mathcal{J}^D\) sublinearly. □

**Remark 1.** Note that due to the \(\lambda\)-term in (8), the proposed method is different to a two-stage method, which reconstructs image \(w\) first by solving a least-squares problem \(\min_w \|Aw - f\|_F^2\) and then uses the SA-TV method to post-process the image. In our approach the variables \(u\) and \(w\) are correlated through the \(\lambda\)-term in the model.

**Remark 2.** In our method, the spatially varying parameter is determined according to the ‘noise’ variance in the residual and the scale of textures in local regions, which is the same as in the SA-TV method. Hence, similar as in [28] the new method also can be extended to apply other regularization terms.

### 4. Numerical results

In this section we provide numerical results for simulated as well as real data to study the behavior of the proposed reconstruction method with a spatially varying regularization parameter. The main focus of our experiments is to show that our proposed approach which uses a spatially varying regularisation parameter \(\lambda\) achieves a good trade-off between reconstructing textures of small scale while at the same time suppressing artifacts also in homogeneous areas in the image due to limited data and noise in the measurements.

In our method, the parameter \(\alpha\) is chosen by applying the discrepancy principle with given or estimated noise variance \(\sigma^2\). We apply the SA-TV method to solve the \(u\)-subproblem (7), using the default setting suggested in [1] where the only input is the ‘noise’ level in the reconstruction, which for the real data is estimated by calculating the variance of the dark background. To solve the \(w\)-subproblem (8), we apply the CGLS method [26] followed by a non-negativity projection. The stopping criterion for CGLS is

\[
\alpha^2 \|A^*(Aw_l - f)\|_F^2 + \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda^2_{ij} (w_{ij} - u_{ij}^{k+1})^2 \leq 10^{-6},
\]

where \(A^*\) denotes the adjoint operator of \(A\) and \(l\) is the iteration index for CGLS. In addition, the stopping criterion for the overall method is

\[
\frac{\|u_{k+1} - u_k\|_F}{\|u_k\|_F} \leq 10^{-4}.
\]
Figure 1. Phantoms for tests. (a) Head phantom used for simulation [29], (b) gelatin phantom used for real x-ray scan.

For the experiments, we test our method on a simulated head phantom from [29] and a real gelatin phantom shown in figure 1. Here, we consider a 2D CT parallel beam. The detector has full coverage of the object at any projection angle, and a constant angular spacing of the rays is set in the interval of [0, π].

Example 1. Our first test example is on the simulated head phantom, which is generated in a square domain of 256 × 256 pixels, i.e. there are 256² = 65,536 unknowns. With 362 rays and 45 projection angles that are equally spaced in the range [0, π], the corresponding CT reconstruction problem has an under-determined rate of 25%. The measurements are given by \( f = \bar{A}u + \epsilon \), where \( \bar{u} \) is the ground truth (the true attenuation coefficients in the object) and \( \epsilon \) denotes the additive white Gaussian noise with the relative noise level \( \| \epsilon \|_F / \| \bar{A}u \|_F \).

For studying the performance of our method, we compare it with four commonly used reconstruction methods: the filtered back-projection (FBP) algorithm [2], the simultaneous algebraic reconstruction technique (SART) [30] as a representation of simultaneous algebraic methods, Kaczmarz’s method (also known as ART) [3] as a representation of sequential algebraic methods, and the L2-TV reconstruction method which solves the variational model (1) with TV regularization as proposed in [13]. All methods are solved under a non-negativity constraint. We use the AIR Tools II toolbox [31] for obtaining the reconstruction results from FBP, SART and Kaczmarz’s method as well as the forward and back projections in the L2-TV and our methods. Note that the performance of SART and Kaczmarz’s method strongly depends on the choice of relaxation parameters. The selection of these parameters is an important and difficult research topic and is outside the scope of this paper. For the purpose of comparison with the TV-regularized methods we took the well-tested default parameter choices for SART and Kaczmarz’s method, i.e. 1.9 and 1, respectively. We also tried some other parameters but did not feel that they significantly improved the reconstruction and for simplicity therefore decided to neglect them. The stopping criterion in both SART and Kaczmarz is when the relative error (RErr) \( \frac{\| \bar{u} - \bar{u} \|_F}{\| \bar{u} \|_F} \) is less than 10⁻³. We show two reconstruction results from the L2-TV method. One is with the scalar regularization parameter \( \alpha \) that gives the smallest RErr values, and the other is with a slightly smaller scalar \( \alpha \) and is visually as smooth as the results from our method.
Figure 2. Results of different methods for reconstructing the head phantom with underdetermined rate 25% and relative noise level 0.2. (a) FBP (RErr = 0.3253), (b) SART (RErr = 0.1292), (c) Kaczmarz (RErr = 0.1406), (d) L2-TV with large scalar $\alpha$ (RErr = 0.0644), (e) L2-TV with small scalar $\alpha$ (RErr = 0.0737), (f) our method (RErr = 0.0525), (g) $\lambda$ in our method with its colorbar.

on the large homogeneous regions. The minimization problem in the L2-TV method is solved by the algorithm proposed in [32].

In figures 2 and 3 we give the reconstruction results, which are shown in the same intensity range as the original phantom in figure 1(a), from the simulated measurements with the noise level 0.2 and 0.8, respectively. Since the FBP algorithm is based on the analytical formulation of the inverse Radon transform, it implicitly requires continuously measured clean data from the whole 0 to $\pi$ angular range. Although we can use different filters to reduce the effect of noise, it is usually difficult to avoid the artifacts due to limited data [4]. We can clearly see many stripe artifacts due to the noise and sparse projection angles in the FBP results. Both the SART and
Kaczmarz’s methods perform better than FBP, but there are still some visible stripe artifacts in the reconstruction. Note that both SART and Kaczmarz belong to algebraic reconstruction methods, i.e. reconstruction results are obtained by solving the linear system $Au = f$ without any extra prior information on the solutions besides a non-negativity constraint. By using the TV regularization in the L2-TV and our method, we can evidently reduce the influence of the noise and avoids stripe artifacts. In addition, comparing the results from the L2-TV and our methods, we find that our method suppresses artifacts much better while reconstructing most details. For instance, the gray region in the head and the black dotted region on the right side. With respect to RErr, it is also clear that our method gives the best reconstruction results. In
Figure 4. Plots of the relative error through the iterations of our method for reconstructing the head phantom with underdetermined rate 25%. (a) With relative noise level 0.2, (b) with relative noise level 0.8.

Table 1. CPU-time in seconds of different reconstruction methods.

| Noise level | FBP  | SART | Kaczmarz | L2-TV | Ours  |
|-------------|------|------|----------|-------|-------|
| 0.2         | $0.47 \times 10^{-2}$ | 25.99 | 13.40    | 8.97  | 58.23 |
| 0.8         | $0.33 \times 10^{-2}$ | 12.77 | 6.74     | 10.23 | 61.51 |

figures 2(f) and 3(f), we also plot the function $\lambda$ obtained by our method with a colorbar. One can see that in the more textured regions $\lambda$ is large in order to preserve the details, and in the more homogeneous regions it is small to reduce artifacts. In addition, in figure 4 we show the behavior of the relative error through the iterations. It is clear that the relative error starts to level off before meeting the stopping criterion. Note that the relative error does not necessarily converge to zero, since the original phantom is not guaranteed to be the minimizer of our model (5).

For comparing computational times, in table 1 we list the CPU-times consumed by the different reconstruction methods in our comparison. All tests are run in MATLAB R2016a on a PC equipped with 2.2 GHz quad-core Intel Core i7 processor and 16 GB of 1600 MHz DDR3L onboard memory. It is clear that FBP is the most efficient method. Due to fast convergence, Kaczmarz’s method costs less time than SART. For the L2-TV method, we apply the semi-smoothing Newton method proposed in [32] which has been proved to have super-linear convergence rate, so it is also efficient compared to algebraic methods. Since in our method the regularization parameter $\lambda$ is spatially varying and adjusted iteratively, and in each iteration we need to solve a least-squares problem (8) and a TV-regularized problem (7), our method is the most time consuming.

Example 2. In order to further study the performance of our method, in this example we test our method on 50 randomly generated 256-by-256 phantoms, which are created based on the Shepp–Logan phantom with a random number of randomly-sized ellipses at random locations and random attenuation. The CT measurements are simulated by using the AIR Tools II toolbox [31] with 362 rays and 45 projection angles that are equally spaced in the range $[0, \pi]$. The relative noise level in the measurements is 0.8.
In figure 5, we show the RErr histogram of the reconstruction results by our method for these 50 phantoms. We can see that the RErr values are mainly in the range [0.105, 0.141]. In order to see the reconstruction results visually, in figure 6 we show three examples of the random phantoms that correspond to the smallest, the median and the largest RErr values and their reconstructions by applying our method. Here, the original phantoms and reconstruction results are all shown in the same intensity range [0, 1]. In all cases, we can see that our method can effectively suppress artifacts due to limited data and reconstruct most details. The plots of $\lambda$ show that in regions where small scale textures are present $\lambda$ values are large in order to preserve them. In the worst case reconstruction, the phantom is much more complicated for reconstruction due to many small scale textures, which are difficult to distinguish from noise. In addition, the ellipse that appears in the background also results in errors in the ‘noise’ level estimation of the SA-TV method when we are solving the $u$-subproblem.

Example 3. Additionally to the simulated data, we also test our method on real CT measurements provided by the Industrial Mathematics Computed Tomography Laboratory at the University of Helsinki. In this experiment a gelatin phantom, which contains gelatin in a petri dish with two embedded aluminum heat sinks and three blackboard chalks, shown in figure 1(b), is measured with a custom-built $\mu$CT device. A set of cone-beam projections with resolution $2240 \times 2368$ and angular step of $\frac{\pi}{180}$ was measured. The exposure time was 1000 ms, the x-ray tube acceleration voltage was 50 kV and tube current 1 mA. From the 2D projection images the middle rows corresponding to the central horizontal cross-section of the gelatin phantom were taken to form a fan-beam sinogram of resolution $2240 \times 360$. This sinogram was further
down-sampled by binning and taking logarithms to obtain the sinograms with 560 beams and 360 or 180 equally spaced projection angles in the range \([0, 2\pi]\) for our tests. The reconstructions are in a square domain of \(512 \times 512\) pixels, which result in an under-determined rate of 77% and of 38%, respectively.

In figures 7 and 8 we compare our method with the FBP algorithm and the L2-TV method, and all results are shown in the same intensity range, with the same colorbar shown on the most left side in the figures. We use the ASTRA toolbox [33] for GPUs for computing the FBP results and for computing the forward as well as back projections in the L2-TV approach and our reconstruction method. Due to insufficient measurements and noise, FBP cannot provide satisfactory results. Comparing the results obtained by our method with the ones from the L2-TV method, we see that our method reduces more artifacts while keeping similar quality on reconstructed object textures. Furthermore, from the final values of \(\lambda\) obtained in our method we find that our method can correctly distinguish textured regions from homogeneous...
regions. Then, by setting different regularization parameter values, we vary the strength of the smoothness in the different regions.

5. Conclusion

In this paper, we introduce a new tomographic reconstruction method with a spatially varying regularization parameter. By introduce an auxiliary variable, the new approach extends the idea proposed in the SA-TV method to general inverse problems, where the data-fitting term and the regularization term fall in different domains. Numerically we have shown that the new approach can reduce the influence of the noise better as well as preserving more details comparing with the state-of-the-art. One limitation of the new approach is that an input of the ‘noise’ level in the reconstruction is required due to the use of the SA-TV method. A better artifact estimation method will be further researched in future work. In addition, an extension of our method to accommodate for Poisson distributed noise can be done by changing the data-fitting term. This is a more accurate model for representing CT measurements especially under a low incident photon flux case, and constitutes a straightforward extension to be left for future work. Further validating our method in real applications is outside the scope of this paper and will be a very important direction for future research.
Acknowledgments

We thank Samuli Siltanen and Alexander Meaney from the Industrial Mathematics Computed Tomography Laboratory at the University of Helsinki for creating the gelatin phantom and providing its CT data. We also thank Per Christian Hansen from Technical University of Denmark for the random phantom generator. YD acknowledges the support of the National Natural Science Foundation of China via Grant No. 11701388. CBS acknowledges support from the Leverhulme Trust project on Breaking the non-convexity barrier, EPSRC Grant No. EP/M00483X/1, the EPSRC Center No. EP/N014588/1, the RISE projects CHiPS and NoMADS, the Cantab Capital Institute for the Mathematics of Information and the Alan Turing Institute. Both authors would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the program ‘Variational methods and effective algorithms for imaging and vision’ where work on this paper was undertaken. This program was supported by EPSRC Grant number EP/K032208/1. We gratefully acknowledge the support of NVIDIA Corporation with the donation of a Quadro P6000 GPU used for this research.

ORCID iDs

Yiqiu Dong https://orcid.org/0000-0001-8363-9448

References

[1] Dong Y, Hintermüller M and Monserrat Rincon-Camacho M 2011 Automated regularization parameter selection in multi-scale total variation models for image restoration J. Math. Imag. Vis. 40 82–104
[2] Frank N and Frank W 2001 Mathematical Methods in Image Reconstruction (Philadelphia, PA: Society for Industrial and Applied Mathematics)
[3] Gabor T H 2009 Fundamentals of Computerized Tomography: Image Reconstruction from Projections 2nd edn (Berlin: Springer)
[4] Quinto E T 1988 Tomographic reconstructions from incomplete data-numerical inversion of the exterior Radon transform Inverse Problems 4 867
[5] Kachment P, Lancaster K and Mogilevskaya L 1995 On local tomography Inverse Problems 11 571
[6] Frikel J and Quinto E T 2013 Characterization and reduction of artifacts in limited angle tomography Inverse Problems 29 125007
[7] Sidky E Y and Pan X 2008 Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization Phys. Med. Biol. 53 4777–907
[8] Frikel J 2013 Sparse regularization in limited angle tomography Appl. Comput. Harmon. Anal. 34 117–41
[9] Romanov M, Dahl A B, Dong Y and Hansen P C 2016 Simultaneous tomographic reconstruction and segmentation with class priors Inverse Probl. Sci. Eng. 24 1432–53
[10] Jia L, Miao C, Shen Z, Wang G and Yu H 2016 Robust frame based x-ray CT reconstruction J. Comput. Math. 34 683
[11] Riis N A B, Frøsigh J, Dong Y and Hansen P C 2018 Limited-data x-ray CT for underwater pipeline inspection Inverse Problems 34 034002
[12] Tikhonov A N and Arsenin V Y 1977 Solutions of Ill-Posed Problems (Washington, DC: Winston and Sons)
[13] Rudin L, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms Physica D 60 259–68
[14] Yang X, Hofmann R, Dapp R, van de Kamp T, dos Santos Rolo T, Xiao X, Moosmann J, Kashef J and Stotzka R 2015 TV-based conjugate gradient method and discrete L-curve for few-view CT reconstruction of x-ray in vivo data Opt. Express 23 5368–87
[15] Weickert J, Romeny B M T H and Viergever M A 1998 Efficient and reliable schemes for nonlinear diffusion filtering IEEE Trans. Image Process. 7 398–410
[16] Scherzer O 1998 Denoising with higher order derivatives of bounded variation and an application to parameter estimation Computing 60 1–27
[17] Bredies K, Kunisch K and Pock T 2010 Total generalized variation SIAM J. Imaging Sci. 3 492–526
[18] Lefkimmiatis S, Roussos A, Maragos P and Unser M 2015 Structure tensor total variation SIAM J. Imaging Sci. 8 1090–122
[19] Yu W and Zeng L. 2015 ℓ0 gradient minimization based image reconstruction for limited-angle computed tomography PLoS One 10 e0130793
[20] Humphries T, Winn J L and Faridani A 2017 Superiorized algorithm for reconstruction of CT images from sparse-view and limited-angle polyenergetic data Phys. Med. Biol. 62 6762–83
[21] Morozov V A 1966 On the solution of functional equations by the method of regularization Soviet Math. Dokldy 7 414–17
[22] Hansen P C 2010 Discrete Inverse Problems: Insight and Algorithms (Philadelphia, PA: Society for Industrial and Applied Mathematics)
[23] Zhang C, Zhang T, Zheng J, Li M, Lu Y, You J and Guan Y 2015 A model of regularization parameter determination in low-dose x-ray CT reconstruction based on dictionary learning Comput. Math. Methods Med. 2015
[24] Giusti E 1984 Minimal Surfaces and Functions of Bounded Variation (Boston, MA: Birkhäuser)
[25] Krishnan V P and Quinto E T 2015 Microlocal analysis in tomography Handbook of Mathematical Methods in Imaging 2nd edn, ed O Scherzer (New York: Springer) pp 847–902
[26] Ake Björck Å 1996 Numerical Methods for Least Squares Problems (Philadelphia, PA: SIAM)
[27] Beck A 2015 On the convergence of alternating minimization for convex programming with applications to iteratively reweighted least squares and decomposition schemes SIAM J. Optim. 25 185–209
[28] Bredies K, Dong Y and Hintermüller M 2013 Spatially dependent regularization parameter selection in total generalized variation models for image restoration Int. J. Comput. Math. 90 109–123
[29] Yu Z, Noo F, Frank D, Adam W, Lauritsch G and Hornegger J 2012 Simulation tools for two-dimensional experiments in x-ray computed tomography using the forbild head phantom Phys. Med. Biol. 57 N237–52
[30] Andersen A H and Kak A C 1984 Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm Ultrasonic Imaging 6 81–94
[31] Hansen P C and Jørgensen J S 2018 AIR tools II: algebraic iterative reconstruction methods, improved implementation Numerical Algorithms 79 107–37
[32] Hintermüller M and Stadler G 2006 An infeasible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration SIAM J. Sci. Comput. 28 1–23
[33] van Aarle W, Jan Palenstijn W, Cant J, Janssens E, Blechrodt F, Dabravolski A, De Beenhouwer J, Batenburg K. J and Jan S 2016 Fast and flexible x-ray tomography using the ASTRA toolbox Opt. Express 24 25129–47