Metal Artifact Reduction in CT-Based Attenuation Correction of PET using Sobolev Sinogram Restoration

A. Mehranian, *Member, IEEE,* M. R. Ay, *Member, IEEE,* A. Rahmim, *Member, IEEE,* H. Zaidi, *Senior Member, IEEE*

***Abstract*–In patients bearing metallic implants, CT images reconstructed by the filtered back-projection algorithm usually suffer substantially from streaking metal artifacts. The CT-based attenuation correction of PET using such images can lead to pseudo-uptakes and thus equivocal findings. In this paper, we introduce a new metal artifact reduction (MAR) algorithm based on Bayesian iterative restoration techniques applied in the sinogram space. We proposed a Sobolev prior in the maximum *a posteriori* (MAP) estimation of the projections corrupted by metallic implants. The Sobolev prior was invoked in order to impose the a priori knowledge that a CT sinogram is a smooth dataset in which it is highly probable that neighboring projections have similar photon counts. We compared the proposed prior with a total variation (TV) one which imposes piece-wise smoothness on the sinograms being restored. We also compared it with a smoothed TV (TVs) prior which ranks between the Sobolev and TV ones. We formulated the MAP estimation as a convex constrained optimization problem and solved it for the Sobolev and TVs priors by a projected gradient descent algorithm and for the TV/s priors by a sophisticated primal-dual projected gradient algorithm. The results of artifact simulations in a real CT image showed that the Sobolev-based MAR algorithm outperforms its TV and TVs-based counterparts in terms of convergence rate and is comparable with the TVs in projection recovery. It was demonstrated that the proposed MAR algorithm has high applicability in fast and efficient CT-based attenuation correction of PET data.**

I. Introduction

ATTENUATION correction of positron emission tomography (PET) measurements plays a pivotal role in improving the quantitative and diagnostic accuracy of PET images. In dual modality PET/CT imaging systems, the measurements are corrected for attenuation using attenuation maps derived from co-registered CT images. In patients with metallic implants, this CT-based attenuation correction (CTAC) results in overestimation and/or underestimation of tracer uptake in regions corresponding to metal artifacts in the CT images reconstructed by filtered back-projection (FBP) algorithm [2]. Accurate and reliable CTAC calls for almost artifact-free CT images, which in turn have spurred extensive efforts on developing metal artifact reduction (MAR) algorithms to restore missing projections due to the strong absorption of X-ray photons by metallic objects.

Generally, MAR algorithms fall into two classes: projection completion methods and model-based iterative reconstruction algorithms. In the first class, the missing projection data are recovered by estimating from neighboring projections using different interpolating algorithms [3-5]. Although projection completion methods are computationally appealing, they may not completely eradicate metal artifacts, particularly when multiple metallic objects like dental fillings are present in the imaging field of view. In fact, the efficiency of the interpolating algorithms depends totally on how robustly they make use of the still available projections. Once the missing projections are completed, the artifact-reduced CT images are obtained by a FBP reconstruction. The model-based iterative reconstruction techniques utilize prior knowledge of the photon counting statistics and the geometry of the imaging scanner. They can easily be adapted to truncated and missing data situations, as the weights (contributions) associated with the regions of missing sinogram data can be down-weighted or simply set to zero [6, 7]. However, these algorithms are computationally intensive, inasmuch as they have not yet largely superseded the conventional FBP algorithms in commercial CT scanners.

In this study, we developed a new projection completion algorithm based on iterative image restoration techniques. Specifically, we formulated the problem of restoring missing projections, as an inverse problem, on the framework of Bayesian estimation and made use of a Sobolev smooth prior in order to incorporate our a priori knowledge. We proposed the Sobolev prior, as a smooth (quadratic) and convex prior, on the account that CT sinograms are rather smooth images in which it is highly probable that neighboring projections have similar photon counts. The proposed prior was compared with the well known total variation (TV) prior [8], which has been previously studied for recovering missing projections [9, 10]. The TV prior is a non-smooth, convex prior and favors piece-wise smooth images. Due to its non-smoothness, this prior is not continuously differentiable which results in intractable optimization problems. In order to handle this, a smoothed TV prior is usually employed instead. By contrast, Sobolev-based

This work was supported by the Research Center for Science and Technology in Medicine (RCSTIM), Tehran University of Medical Sciences, Tehran, Iran, under grant No. 12001.

A. Mehranian is with the RCSTIM, Tehran University of Medical Sciences, Tehran, Iran (e-mail: mehranian@razi.tums.ac.ir).

M. R. Ay is with the RCSTIM and the Department of Medical Physics and Biomedical Engineering, Tehran University of Medical Sciences, Tehran, Iran, (e-mail: mohammadreza_ay@sina.tums.ac.ir).

A. Rahmim is with the Department of Radiology, School of Medicine, Johns Hopkins University, Baltimore, USA (e-mail: arahmim1@jhmi.edu).

H. Zaidi is with the Division of Nuclear Medicine, Geneva University Hospital, CH-1211 Geneva, Switzerland (e-mail: habib.zaidi@hcuge.ch).

optimization problems are amenable to minimization and therefore are computationally effective.

This paper is organized as follows. In Section II, we first formulate a constrained optimization problem for estimating missing projections in a Bayesian framework and then describe discrete Sobolev and TV priors and simulated CT data. In Section III, we elaborate a projected gradient descent algorithm for solving the formulated problems of Sobolev and smoothed TV priors. We also describe a primal-dual gradient algorithm for smoothed, as well as exact TV minimizations. In Section IV, we compare the performance of the MAR algorithms, and then present the PET images and their attenuation maps obtained from the proposed algorithm. Finally, conclusions are drawn in Section V.

II. Methods

A. Problem Formulation

Let $\tilde{x} \in \mathbb{R}^N$, $N = t \times \theta$ be an observed sinogram with some metal traces, where t and θ denote the number of detector bins and angular samplings, respectively. Also let $\Omega \in \mathbb{R}^N$ be a binary mask indicating the region of metal traces on $\tilde{x}$, where projection data are missing or corrupted and y be the same as $\tilde{x}$ but with zeroed (removed) metal traces. In order to recover missing projections, we first formulate the following forward model

$$y = \Phi x + n \tag{1}$$

where x is an ideal sinogram in which there are no missing projection data and $\Phi \in \mathbb{R}^{N\times N}$ is a diagonal lossy operator, $\Phi = \mathrm{diag}\{1 - \Omega\}$, which zeroes the projections on the regions corresponding to metal traces in Ω, and n denotes the noise, often additive zero-mean Gaussian with covariance matrix Λ, i.e. $\mathcal{N}(0,\Lambda)$. In order to find the best estimate, $\hat{x}$, of x, one can follow a Bayesian estimation scheme. According to Bayes' law, the a posteriori probability of x given y is defined as $P(x \mid y) = P(y \mid x)P(x) \,/\, P(y)$. In Bayesian models, all unknown parameters are treated as stochastic quantities with probability distributions describing them. Given the model (1) under white Gaussian noise ($\Lambda = \sigma^2 \mathrm{I}$, σ^2 is the variance of noise and I identity matrix), the conditional distribution of the y given x is obtained by

$$P(y \mid x) \propto \exp\left\{-\frac{\alpha}{2}\left\|y - \Phi x\right\|_2^2\right\}, \quad \alpha = 1 \,/\, \sigma^2 \tag{2}$$

Our prior knowledge of the ideal image x is expressed as a prior probability distribution $P(x)$. The prior distribution is properly described within Gibbs random fields, which are characterized by neighboring-pixel interactions, and is given by $P(x) \propto \exp(-\beta G(x))$, where the *prior energy* $G(x)$ is a sum of neighboring-interaction terms, called clique potentials, and β is the prior hyperparameter. The prior energy is generally of the form $G(x) = \sum_{c\in C} V_c(x)$ where C is a set of cliques (neighboring pixels). In connection with the prior energy the residual norm in (2) is sometimes referred to as *likelihood energy*. $P(y)$ is simply a probability normalization constant that can be dropped. Having an estimate of the hyperparameters α and β, one can obtain a maximum *a posteriori* (MAP) estimate of x by maximizing the *a posterior* probability i.e.

$$\hat{x} = \arg\max_x P(y \mid x) \times P(x)$$

which can equivalently be obtained by the following energy minimization

$$\hat{x} = \arg\min_{x\in\mathbb{R}^N} \frac{\alpha}{2}\left\|y - \Phi x\right\|_2^2 + \beta G(x) \tag{3}$$

where the hyperparameters weight the importance of the likelihood and prior energies. The residual norm measures the proximity of x to y if observed through Φ, while the prior enforces our *a priori* knowledge about it (e.g. smoothness). Generally, the prior penalizes noisy solutions which are of higher energies due to their stronger neighboring-pixel interactions. It should be noted that the hyperparameters in (3) are mostly encapsulated in a single regularization parameter $\lambda = \beta \,/\, \alpha$ between the two energies. It therefore controls the denoising effect of the prior and should be adapted to the variance of noise. As the variance of noise is decreased, the proximity of Φx to y is increased. In the limit, where there is no noise ($\alpha \to \infty$), as in the present context, the optimization problem (3) reduces to the following constrained problem

$$\hat{x} = \arg\min_{x\in\Theta} G(x) \tag{4}$$

where the constraint set inside which the solution must reside is defined as

$$\Theta = \{x \in \mathbb{R}^N : \Phi x = y\} \tag{5}$$

In fact, the problem (4) seeks a *minimum prior energy* estimate subjected to $\Phi x = y$.

B. Discrete Sobolev and TV priors

We define the ℓ^p norm of a vector field $u = (u_1, u_2)$ where $u_1, u_2 \in \mathbb{R}^N$ as:

$$\left\|u\right\|_p = \left(\sum\nolimits_{ij} \left|u(i,j)\right|^p\right)^{1/p}, \quad \left|u(i,j)\right| = \sqrt{u_1(i,j)^2 + u_2(i,j)^2}.$$

The TV and Sobolev priors are then respectively defined as the ℓ^1 and one-half of the squared ℓ^2 norm of the discrete gradient, $\nabla : \mathbb{R}^N \to \mathbb{R}^{N\times 2}$, of the image x, i.e. $G_{TV}(x) = \left\|\nabla x\right\|_1$ and $G_{Sob}(x) = (1 \,/\, 2)\left\|\nabla x\right\|_2^2$. Via a forward finite difference approximation to the first-order partial derivatives of x with Neumann boundary conditions, we define the gradient of x as

$$\nabla x(i,j) = \begin{pmatrix} x(i+1,j) - x(i,j) \\ x(i,j+1) - x(i,j) \end{pmatrix}$$

which is thus a vector field. The adjoint of the gradient is $\nabla^* = -\mathrm{div}$, where $\mathrm{div} : \mathbb{R}^{N\times 2} \to \mathbb{R}^N$ is the divergence operator such that $\left\langle \nabla x, u \right\rangle_{\mathbb{R}^{N\times 2}} = -\left\langle x, \mathrm{div}(u) \right\rangle_{\mathbb{R}^N}$, where

$\langle \cdot,\cdot \rangle_X$ is the inner product in a vector space X. The discrete divergence operator is also defined as

$$\mathrm{div}(u) = u_1(i,j) - u_1(i-1,j) + u_2(i,j) - u_2(i,j-1)$$

The divergence of the gradient is the Laplacian operator that is denoted by Δx. The Laplacian of an image x is also the trace of its Hessian matrix $\nabla^2 x$ which contains the second-order partial derivatives. For a gradient-based energy minimization, one can derive the gradient of the Sobolev and TV priors from the diffusion equation as follows [11]:

$$\nabla G_{Sob}(x) = -\mathrm{div}(\nabla x) = -\Delta x \tag{6}$$

$$\nabla G_{TV}(x) = -\mathrm{div}\left(\nabla x \,/\, \|\nabla x\|_1\right) \tag{7}$$

In fact, Sobolev and TV energy minimizations correspond respectively to isotropic and anisotropic diffusion filtering with diffusion coefficients $D = 1$ and $D = 1 / \|\nabla x\|_1$, respectively. As (7) is not defined where $\nabla x = 0$, the TV prior is usually conditioned to $\sqrt{\delta^2 + \|\nabla x\|_2^2}$, where $\delta > 0$ is a small parameter that controls the edge preserving power of the resulting *smoothed* TV prior. Therefore, we would have a continuously differentiable TV prior with the following gradient

$$\nabla G_{TV}^{\delta}(x) = -\mathrm{div}\left(\nabla x \,/\, \sqrt{\delta^2 + \|\nabla x\|_2^2}\right) \tag{8}$$

A continuously differentiable convex function $G(x)$ is defined to have a Lipschitz continuous gradient with Lipschitz constant L, if

$$\|\nabla G(x) - \nabla G(x')\|_2 \le L\|x - x'\|_2, \text{ for } x, x' \in \mathbb{R}^N \tag{9}$$

The Lipschitz constant is of importance in the convergence rate of the gradient-based optimization algorithms and equals the square of the induced norm of the gradient operator, i.e. $\|\nabla\| = \max\{\|\nabla x\| : x \in \mathbb{R}^N, \|x\| \le 1\}$. For Sobolev and smoothed TV priors, we have $L \le 8$ and $L \le 8 / \delta$, respectively. A proof can be found in [12].

C. Simulated CT data

As in clinical settings, there is no a ground truth CT image in order to objectively evaluate the performance of a proposed MAR algorithm, we simulated metal artifacts induced by dental fillings in a real head CT image. Let x_t be the sinogram of the true image, obtained by its forward projection. In order to obtain an artifact-degraded CT image, shown in Fig.1b, we first artificially implanted four dental fillings in the dental arch of the true image (Fig. 1a) and then forward projected the resulting image. In the next step, metal traces and thus the binary mask Ω were identified by thresholding. Then, the count number of missing projections on Ω was saturated according to the following empirical scheme, which gave rise to a prominent data inconsistency in the projection space. Letting $i = (t,\theta)$, the observed sinogram $\tilde{x}$ with metal traces was synthesized as

$$\tilde{x}(i) = \begin{cases} 0.4\, x_t(i) + 0.6 \max\{ x_t(i), \forall \Omega(i) = 1\}, & \Omega(i) = 1 \\ x_t(i), & \Omega(i) = 0 \end{cases}$$

The observed image (Fig. 1b) was reconstructed by the FBP of $\tilde{x}$ and then, as in the GE Discovery LS PET/CT scanners, its maximum CT number was truncated to 3071 HU.

III. Optimization Algorithms

The Sobolev and TV priors, as well as the constraint set defined in (5) are convex (proof follows from below definitions). Hence, the problem (4) can be considered as a convex optimization (programming) problem for which local minimizers are global. We first give some basic definitions and then elaborate the optimization algorithms.

A. Definitions

A set $S \subset \mathbb{R}^N$ is a convex set if for every point $x, x' \in S$ and any $\theta \in (0,1)$, the point $\theta x + (1-\theta)x'$ also lies inside S. We define the domain of a function F as $\mathrm{dom}F = \{x : F(x) < \infty\}$. F is proper, if $\mathrm{dom}F \ne \emptyset$. The function F is convex, if $\mathrm{dom}F$ is convex and if for all $x, x' \in \mathrm{dom}F$, and any $\theta \in (0,1)$, we have $F(\theta x + (1-\theta)x') \le \theta F(x) + (1-\theta)F(x')$. Let S be a nonempty convex set, the (convex) *indicator function* $D(x)$ of S is defined as, $D(x) = 0$ if $x \in S$ and $D(x) = \infty$ if $x \notin S$. We also define the *conjugate* of a proper, closed and convex function as

$$F^*(x') = \max_{x \in \mathrm{dom}F} \langle x', x \rangle - F(x) \tag{10}$$

The conjugate of $F(x) = \|x\|_1$ is an indicator function defined as $F^*(x') = 0$ if $\|x'\|_* \le 1$, $F^*(x') = \infty$ otherwise, where the dual norm is given by $\|x'\|_* = \max\{\langle x', x \rangle : \|x\|_1 \le 1\}$.

B. Projected Gradient Descent (PGD)

To solve (4), we recast it into the following more general optimization problem

$$\min_{x \in \mathbb{R}^N} G(x) + D(x) \tag{11}$$

where $D(x)$ is an indicator function of the set Θ in (5). Let $\nabla G(x)$ be L-Lipschitz, thereby $G(x)$ is twice continuously differentiable over $\mathbb{R}^N$ and its Hessian matrix is thus symmetric, positive definite; i.e. $0 < \nabla^2 G(x) \le LI$ [13]. The constant L can be interpreted as an upper bound on $\nabla^2 G(x)$, and thus measures how well $G(x)$ can be (quadratically) approximated from above. We replace $G(x)$ by a quadratic function (paraboloidal surrogate) near point x^k as

$$G(x^k) + \left\langle \nabla G(x^k), x - x^k \right\rangle + \frac{1}{2}\left\langle x - x^k, \nabla^2 G(x^k)(x - x^k) \right\rangle$$

and then approximate it from above

$$G(x) \leq G(x^k) + \left\langle \nabla G(x^k), x - x^k \right\rangle + \frac{L}{2}\left\| x - x^k \right\|_2^2 \quad (12)$$

It is worth to note that the above condition is equivalent to the definition of the Lipschitz continuity of the gradient in (9) [14].By some algebra, dropping constant terms and adding the function $D(x)$, one can derive the parabola $Q_\tau(x, x^k)$ whose minimization acts as a proxy for the problem (11). Therefore,

$$\begin{aligned} x^{k+1} &= \arg\min_{x \in \mathbb{R}^N} Q_\tau(x, x^k) \\ &= \arg\min_{x \in \mathbb{R}^N} \left\{ \frac{1}{2\tau}\left\| x - (x^k - \tau \nabla G(x^k)) \right\|_2^2 + D(x) \right\} \end{aligned} \quad (13)$$

where $\tau = 1 / L$. Equation (13) is composed of an operator called the *proximal map* of $D(x)$ defined as

$$prox_{\tau D}(z^k) = \arg\min_{z \in \mathbb{R}^N} \left\{ \frac{1}{2\tau}\left\| z - z^k \right\|_2^2 + D(z) \right\} \quad (14)$$

and a gradient descent step with step size τ. It is therefore called a *proximal gradient* algorithm [15] which reads

$$x^{k+1} = prox_{\tau D}(x^k - \tau \nabla G(x^k)) \quad (15)$$

As $D(x)$ is an indicator function here, (14) turns into a projector $P_\Theta(z^k)$, that projects iterates back into the set Θ. This reduces the proximal gradient to a *projected gradient descent* (PGD) algorithm, in which the projector is defined as $P_\Theta(z^k) = z^k + \Phi^T(y - \Phi z^k)$. This projection in effect re-insert the already known projection data of y into the iterate z^k. In other words, this operation (10) acts as follows:

$$P_\Theta(z^k(i)) = \begin{cases} y(i), & \text{if } \Omega(i) = 1 \\ z^k(i), & \text{otherwise} \end{cases}$$

For convergence, the gradient step size should satisfy $0 < \tau < 2 / \lambda_{\max}(\nabla^2 G(x))$ (see [1] for a proof), where $\lambda_{\max}$ is the largest eigenvalue of the Hessian matrix which is bounded by L (recall L is ≤ 8 and $\leq 8 / \delta$ for our Sobolev and smoothed TV priors). We summarize this algorithm as follows.

PGD ALGORITHM

Choose $\tau \in (0, 2 / L)$, compute $\nabla G(x^k)$ according to (6) or (8), update x^k as $x^{k+1} = P_\Theta(x^k - \tau \nabla G(x^k))$

Note that in smoothed TV energy minimization, the parameter δ should be a multiplier of the maximum count of y, denoted by $y_{\max}$, thereby it can make the TV prior smooth.

C. Primal-Dual Projected Gradient (PDPG)

The smoothed TV energy minimization with a PGD algorithm suffers from slow convergence (see Results). In order to compare the Sobolev-PGD algorithm with a more robust algorithm, we implemented the energy minimization of the smoothed and the exact TV priors using primal-dual algorithms which are particularly targeted for non-smooth convex optimization problems. Consider the problem (11) as

$$\min_{x \in \mathbb{R}^N} F(\nabla x) + D(x) \quad (16)$$

where $F : \mathbb{R}^{N \times 2} \rightarrow \mathbb{R}_+$ is $F(u) = \sqrt{\delta^2 + \left\| u \right\|_2^2}$. Using (10), the primal-dual formulation of the above problem reads

$$\begin{aligned} &\min_x \left\{ \max_u \left\langle u, \nabla x \right\rangle - F^*(u) \right\} + D(x) = \\ &\quad \max_u \left\{ \min_x \left\langle \nabla^* u, x \right\rangle + D(x) \right\} - F^*(u) \end{aligned} \quad (17)$$

where $u \in \mathbb{R}^{N \times 2}, x \in \mathbb{R}^N$. To solve (17), we follow the algorithm proposed in [16] and summarize it as follows

PDPG ALGORITHM

Choose τ, σ such that $\tau\left(\frac{\sigma}{y_{\max}}\right) L < 1$, update $u^k, x^k, \underline{x}^k$

$$\begin{aligned} u^{k+1} &= prox_{\sigma F^*}(u^k + \sigma \nabla \underline{x}^k) \\ x^{k+1} &= prox_{\tau D}(x^k - \tau \nabla^* u^{k+1}) \\ \underline{x}^{k+1} &= 2x^{k+1} - x^k \end{aligned} \quad (18)$$

The idea of (18) is to alternate between a gradient ascent in the dual variable u and a gradient descent in the primal variable x. The $prox_{\tau D}$ is the same as P_Θ defined before. The dual function $F^*(u)$ is an indicator function (see section III.A), hence, its proximal map (recall (14)) is also a projector. One can derive $prox_{\sigma F^*}$ according to Moreau's decomposition [17] as follows

$$z^k = prox_{\sigma F}(z^k) + \sigma\, prox_{1/\sigma F^*}(z^k / \sigma) \quad (19)$$

The $prox_{\sigma F}(z^k)$ has a closed-from solution, which is $z^k\left[1 - \sigma / \sqrt{\delta^2 + \left\| z^k \right\|_2^2}\right]_+$, where $[x]_+ = \max(0, x)$ denotes a non-negativity operator. For an exact TV prior, where $\delta = 0$, the solution corresponds to a soft-thresholding operator commonly used in wavelet denoising (see [17] for a proof). It is worth noting that a function $F(u) = (1 / 2)\left\| u \right\|_2^2$ would have a $z^k / (1 + \sigma)$ solution. From (19), the proximal map of F^* is thus given by

$$prox_{\sigma F^*}(z^k) = z^k - z^k\left[1 - 1 / \sqrt{(\sigma\delta)^2 + \left\| z^k \right\|_2^2}\right]_+$$

IV. RESULTS AND DISCUSSION

Figs. 1a-b show the ground truth head CT image (True) and its corresponding artificially degraded CT image (Observed). It is noticeable that the dark and bright streaking artifacts in the Observed CT image have impaired the anatomically relevant information much the same way as they do in a real CT acquisition. In Figs. 1c-d, the performance of the smoothed TV- and Sobolev-based MAR algorithms has been shown. As can be visually appreciated, both algorithms have

substantially reduced the artifacts, however, the Sobolev one have some more suppressed streaking artifacts. To objectively compare the algorithms, we defined the following normalized

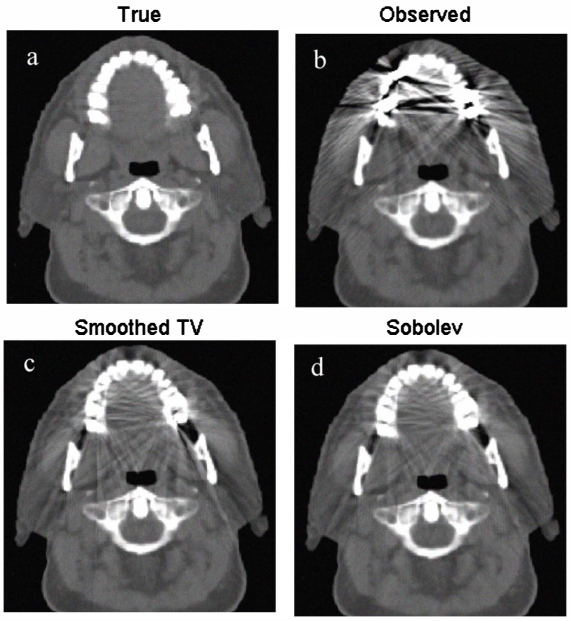

Fig. 1. a) The real head CT image used as a True image and b) its Observed CT image reconstructed by filtered backprojection algorithm after numerically implanting four amalgam dental fillings. The metal-artifact reduced CT images reconstructed using c) smoothed TV-based MAR and d) Sobolev-based MAR algorithms. The display window level/width is the same for all images ($\mathrm{WL} = 30\mathrm{HU}$, $\mathrm{WW} = 800\mathrm{HU}$).

total variation error in image space, adopted from [18],

$$\%\mathrm{TV}=100\times\left(\left\|\nabla(g-g_t)\right\|_1 / \left\|\nabla g_t\right\|_1\right)$$

where g and g_t are the image under evaluation and the true image shown in Fig. 1a, respectively. We also quantified the efficiency of the algorithms in the recovery of missing projections in projection space by the following metric

$$\mathrm{SNR}_{(\mathrm{dB})} = -20\log_{10}\left(\left\|x-x_t\right\|_2 / \left\|x_t\right\|_2\right)$$

where x and x_t are respectively the sinograms of g and g_t .

Table I summarizes the %TV and SNR performance of the algorithms. As could be expected, the smoothed TV priors and at the top of them the Sobolev prior demonstrate enhanced quantitative performance in comparison with the non-smooth TV prior.

TABLE I. THE %TV AND SNR PERFORMANCE OF THE MAR ALGORITHMS.

Algorithm	%TV	SNR (dB)
TV–PDPG	61.47	31.91
TVs–PDPG	53.95	33.15
TVs–PGD	53.83	33.20
Sob–PGD	53.66	33.26

TVs = smoothed TV, Sob = Sobolev

We further evaluated the MAR algorithms in terms of convergence rate and computation (CPU) time. To keep track of convergence, the logarithm of the following normalized ℓ^2 error was invoked

$$error = \left\|x^k - x_t\right\|_2 / \left\|x_t\right\|_2$$

where x^k is the sinogram being restored at iteration k and x_t is the sinogram of the true image.

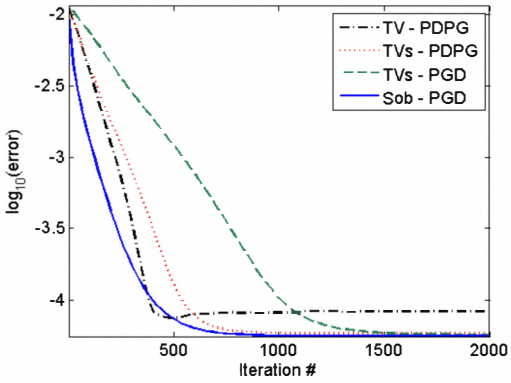

Fig. 2. The logarithm of the normalized ℓ^2 error versus iteration number.

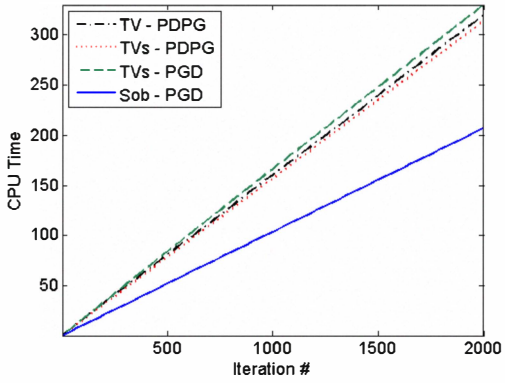

Fig. 3. The computation (CPU) time versus iteration number.

Fig. 2 shows the convergence rate of the studied MAR algorithms in the recovery of missing projection data as a function of iteration number. As seen, the Sobolev–PGD and smoothed TV (TVs)–PGD algorithms possess the fastest and the slowest convergence. The initial convergence rate of the non-smoothed TV–PDPG algorithm is comparable with the Sobolev–PGD algorithm, which demonstrates the ability of primal-dual schemes in non-smooth convex optimization problems. However, its convergence reaches a plateau in the earlier iterations; such that it can not guarantee even the level of convergence reached by the smoothed TV prior. In fact, our prior knowledge of the to-be-restored sinogram accounts for this convergence behavior. The exact TV prior allows for piece-wise smooth images and take into account discontinuities. However, a CT sinogram is rather a smooth image in which neighboring pixels are highly correlated and thus is well described by a Gibbs or Markov random field. The priors that favor this smoothness property should have a better performance in the recovery and estimation of missing projections. This explains the comparable %TV and SNR performance of the smoothed TV priors with the Sobolev one, and their enhanced performance compared to the standard non-smooth TV prior.

In comparison with the TVs–PGD, the TVs–PDPG algorithm has a faster convergence, however, still slower than the Sobolev–PGD. It should be noted that the convergence rate of the smoothed TV algorithms is highly dependent on the

choice of the parameter δ. We chose $\delta = 2\text{E-}02\ y_{\max}$ and $\delta = 12\text{E-}02\ y_{\max}$ for the TVs–PGD and TVs–PDPG algorithms, respectively. As mentioned earlier, the PDPG algorithm is not targeted for smooth priors. Our experiments showed that as the smoothness in the TVs-PDPG gets higher, the convergence of the resulting algorithm gets slower. Hence, this algorithm is not suited for such basically smooth priors as Sobolev. The computation time of the studied algorithms has been illustrated in Fig. 3. As seen, the Sobolev–PGD algorithm far surpasses the other algorithms. This property is of importance particularly in fast CTAC in PET/CT imaging systems.

We finally applied the Observed and Sobolev CT images shown in Figs. 1b-1d for attenuation correction of a simulated PET image corresponding to the True CT image in Fig. 1a. For this purpose, a bilinear energy mapping was employed and the attenuation corrections (μ-maps) of PET at 511 Kev were derived from the CT numbers, which are in essence the attenuation coefficients of tissues at CT energy, typically 80 Kev. Fig. 4 shows the CTAC PET images, together with their corresponding μ-maps obtained from the above-mentioned CT images. The PET images were statistically reconstructed by a penalized weighted least squares algorithm available in [19]. As demonstrated, the Sobolev-based MAR algorithm has considerably compensated for the pseudo-uptakes induced by streaking artifacts. The results showed that the algorithm reduced the %TV of the Observed μ-maps shown in Fig. 4, from 27.85 to 14.24.

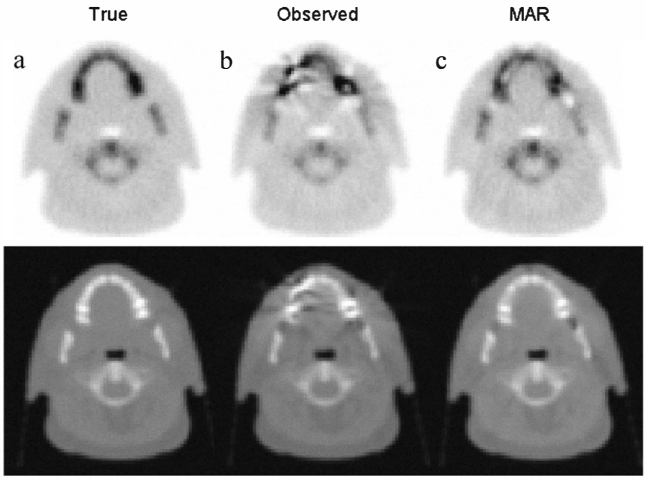

FIG. 4. Attenuation corrected PET images (Top), together with their attenuation maps (Bottom) derived from a) the True, b) Observed c) and Sobolev (MAR) CT images shown in Fig. 1.

V. CONCLUSION

We proposed a Sobolev smooth prior for the recovery of missing projections in a Bayesian estimation framework and compared it with both a smoothed and an exact TV priors. The problem of restoring missing projections was casted as a constrained optimization problem and solved with a projection gradient descent algorithm for the Sobolev and a robust primal-dual hybrid projected gradient for the TV priors. As in the clinical settings, there is no a ground truth CT image to objectively evaluate the performance of a metal artifact reduction (MAR) algorithm, we simulated metal artifacts in a real CT data with an empirical formula to make the projection data corrupted and thus inconsistent. It was found that the more smoothness the priors impose on the sinograms, the more accurate the missing projections are being restored. As a result, the Sobolev prior outperforms the smoothed and the exact TV priors. It was demonstrated that the proposed prior results in a fast and computationally efficient MAR algorithm which is of interest in fast CT-based attenuation correction of PET data.

REFERENCES

[1] E. P. Chong and S. H. Zak, *An Introduction to Optimization*, 2 ed.: John Wiley & Sons, 2001.

[2] M. R. Ay, A. Mehranian, M. Abdoli, P. Ghafarian, and H. Zaidi, "Qualitative and Quantitative Assessment of Metal Artifacts Arising from Implantable Cardiac Pacing Devices in Oncological PET/CT Studies: A Phantom Study," *Mol. Imaging Biol.,* 2011.

[3] S. Zhao, K. T. Bae, B. Whiting, and G. Wang, "A wavelet method for metal artifact reduction with multiple metallic objects in the field of view," *J. X-Ray Sci. Technol.,* vol. 10, pp. 67-76, 2002.

[4] W. J. H. Veldkamp, R. M. S. Joemai, A. J. van der Molen, and J. Geleijns, "Development and validation of segmentation and interpolation techniques in sinograms for metal artifact suppression in CT," *Med. Phys.,* vol. 37, 2010.

[5] C. Xu, F. Verhaegen, D. Laurendeau, S. A. Enger, and L. Beaulieu, "An algorithm for efficient metal artifact reductions in permanent seed implants," *Med. Phys.,* vol. 38, pp. 47- 56, 2011.

[6] L. Ritschl, F. Bergner, C. Fleischmann, and M. Kachelrieß, "Improved total variation-based CT image reconstruction applied to clinical data," *Phys. Med. Biol.,* vol. 56, pp. 1545-1561, 2011.

[7] X. Zhang, J. Wang, and L. Xing, "Metal artifact reduction in x-ray computed tomography (CT) by constrained optimization," *Med. Phys.* , vol. 38, pp. 701-711, 2011.

[8] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D,* vol. 60, pp. 259-268, 1992.

[9] X. Duan, L. Zhang, Y. Xiao, J. Cheng, Z. Chen, and Y. Xing, "Metal Artifact Reduction in CT images by Sinogram TV inpainting," in *IEEE Nuclear Science Symposium Conference Record*, 2008, pp. 4175-4177.

[10] H. Xue, L. Zhang, Y. Xiao, Z. Chen, and Y. Xing, "Metal Artifact Reduction in Dual Energy CT by Sinogram Segmentation Based on Active Contour Model and TV Inpainting," in *IEEE Nuclear Science Symposium Conference Record*, 2009, pp. 904- 908.

[11] Y. You, W. Xu, A. Tannenbaum, and M. Kaveh, "Behavioral analysis of anisotropic diffusion in image processing," *IEEE Trans, on Image Process,* vol. 5, pp. 1539-1552, 1996.

[12] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imaging Vis.,* vol. 20, pp. 89-97, 2004.

[13] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, U.K.: Cambridge University Press, 2004.

[14] Y. Nesterov, *Introductory Lectures on Convex Optimization*: Kluwer Academic Publishers, Dordrecht, 2004.

[15] A. Beck and M. Teboulle, "Gradient-Based Algorithms with Applications in Signal Recovery Problems," in *Convex Optimization in Signal Processing and Communications*, D. Palomar and Y. Eldar, Eds.: Cambribge University Press, 2010, pp. 33--88.

[16] A. Chambolle and T. Pock, "A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging," *J Math Imaging Vis,* vol. 40, pp. 120-145, 2010.

[17] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *SIAM Multiscale Model. Simul.,* vol. 4, pp. 1168-1200, 2005.

[18] J. Tang, B. E. Nett, and G. H. Chen, "Performance comparison between total variation (TV)-based compressed sensing and

statistical iterative reconstruction algorithms," *Phys. Med. Biol.*, vol. 54, pp. 5781-5804, 2009.

[19] J. Fessler, "Image reconstruction toolbox," http://www.eecs.umich.edu/~fessler/code/index.html.