Sparsity Constrained Sinogram Inpainting for Metal Artifact Reduction in X-ray Computed Tomography

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

Abstract–In this paper, we proposed a new projection completion metal artifact reduction (MAR) algorithm in x-ray computed tomography (CT) using a sparsity based sinogram inpainting (interpolation) technique. We developed the MAR algorithm on a Bayesian framework in which a wavelet-based generalized Gaussian (ℓ^p) prior was applied and then the inpainting problem was formulated as a constrained optimization problem. For the optimization, we derived a projected gradient descent algorithm using a majorization-minimization technique. The gradient step was performed by a soft thresholding operator for an ℓ^1 prior, and a hard thresholding with a decaying threshold for an ℓ^0 prior. We utilized a tight frame of translation-invariant wavelets implemented by undecimated discrete wavelet transform. As in the clinical setting there is no ground truth CT image to objectively evaluate the performance of a proposed MAR algorithm, we also introduced a novel approach to simulate metal artifacts in a real CT dataset. The results showed that the proposed MAR algorithm using hard thresholding efficiently recovers and inpaints the sinogram projections corrupted by metallic implants.

I. INTRODUCTION

THE non-linear absorption of polychromatic x-ray beams used in Computed Tomography (CT) often results in beam hardening artifacts in the CT images reconstructed by the conventional filtered-back projection (FBP) algorithm [1]. In the presence of strongly attenuating objects such as metallic implants, the beam hardening is usually so strong that the detectors sensing the implants experience a serious photon starvation, and as a result, the relevant projection data become corrupted and inconsistent. As a result, in the CT image reconstruction process, this data inconsistency appears as streaking metal artifacts that can obscure diagnostically crucial information in the surrounding tissues of the implants and reduce the applicability of this valuable modality [2]. Hence, in the past three decades, there have been considerable efforts toward developing efficient metal artifact reduction (MAR)

algorithms and compensating for the corrupted or missing projection data.

Generally, the MAR algorithms fall into two classes: projection completion algorithms and model-based iterative reconstruction algorithms. The projection completion techniques, which are followed by reconstruction, aim at interpolating the missing projections based on their adjacent projections [3, 4]. These algorithms are fast and computationally appealing; however, if not efficiently developed, they may result in new artifacts. In fact, their efficiency depends entirely on how robustly they exploit the still available projections for the recovery of missing ones. On the other hand, the model-based iterative reconstruction techniques model the image and imaging process using prior knowledge and assumptions, and therefore hold considerable potential for artifact reduction [5, 6]. However, this class of algorithms is computationally demanding and requires further progress to accept reliable utility.

In this study, we proposed a new projection completion MAR algorithm based on inpainting techniques[1] that are well-established in the recovery of missing parts of image and video data [7]. We formulated the inpainting of missing projections as a constrained optimization within a Bayesian framework. A wavelet-based generalized Gaussian prior was incorporated in this framework, which allowed exploiting the sparsity of the sinograms in the wavelet domain for the recovery process. In [8], the authors proposed a wavelet-based MAR algorithm in which the missing projections are one-dimensionally interpolated (along projection profiles) by a weighted sum of the wavelet coefficients of the corrupted projections and those of linearly interpolated projections. Although theoretically well-founded, their algorithm did not significantly reduce metal artifacts. In comparison, as mentioned above, our algorithm is a wavelet-based Bayesian optimization problem through which the missing projections are two-dimensionally optimally estimated by iterative soft/hard thresholding schemes instead of pure interpolations.

The rest of this paper is organized as follows. In Section II, we first formulate the constrained optimization problem in a Bayesian framework and then describe a novel approach in simulating metal artifacts in real CT datasets through which the algorithm was evaluated. In Section III, we elaborate a projected gradient descent algorithm for solving the

---

[1] Inpainting is an artistic synonym for (image) interpolation.

formulated problem. In Section IV, we compare the performance of the MAR algorithms for three different tight frames of translation invariant wavelets. Finally, conclusions are drawn in Section V.

II. Methods

A. Problem Formulation

We represent a two-dimensional sinogram by either an $r \times \phi$ matrix – where r and ϕ are the number of detector channels and projection angles – or by a $N = r \times \phi$-length vector formed by lexicographically ordering the sinogram matrix. Let $\tilde{x} \in \mathbb{R}^N$ denote an observed CT sinogram with metal traces that corrupt the projection data at the locations indexed by a set Ω, and y be identical to $\tilde{x}$ but with removed projections over Ω. In the recovery of corrupted or missing data, we 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 matrix that zeroes the projections at the locations indexed by Ω, i.e. $\Phi_{ii} = 1$ if $i \in \Omega$ and $\Phi_{ii} = 0$ if $i \notin \Omega$. The n is often a realization of a zero-mean Gaussian noise field with covariance matrix Λ.

In order to obtain a best estimate x of $\hat{x}$, we derive an optimization algorithm for the inversion of (1) within a Bayesian framework. In Bayesian modeling, the unknown sinogram x is treated as a stochastic quantity with a prior probability distribution $P(x)$, which models our *a priori* knowledge, and therefore can be exploited to enforce a sparsity constraint in the estimation process. The most widely used sparsity-promoting prior is a generalized Gaussian distribution of the form

$$P(x) \propto \exp\left\{ -\beta \|x\|_p^p \right\},$$

where $\|x\|_p^p = \sum_i |x_i|^p$ is the p-th power of the ℓ^p norm of x and β is the prior hyperparameter. The observation y is also a stochastic process with a *conditional* probability distribution that under white Gaussian noise ($\Lambda = \sigma^2 I$, σ^2 is the variance of noise and I is the identity matrix) is described by

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

In the context of Bayesian estimation under zero-one loss, one can find a maximum *a posteriori* (MAP) estimate of x by the following maximization:

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

which by taking the logarithm of the probability distributions can be equivalently obtained by

$$\hat{x} = \arg\min_x \frac{\alpha}{2} \|y - \Phi x\|_2^2 + \beta \|x\|_p^p \tag{2}$$

where α and β control the balance between the solution's data fidelity and sparsity.
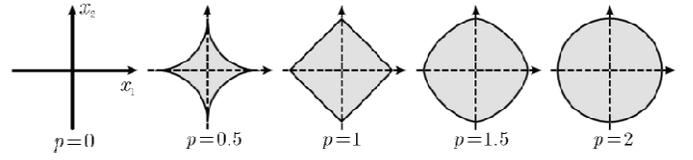

Fig. 1. The unit ball ($r = 1$) of an ℓ^p prior in $\mathbb{R}^2$.

As the variance of noise is decreased, α is increased and vise versa. In the limit, when there is no noise ($\alpha \rightarrow \infty$), as in the present context, the problem (2) reduces to the following equality constrained optimization problem

$$\hat{x} = \arg\min_{x \in X} \|x\|_p^p, \quad X = \left\{ x \in \mathbb{R}^N : \Phi x = y \right\}, \tag{3}$$

where X is a constraint set inside which the solution resides. Geometrically, this set appears as a set of hyperplanes whose intersection with an ℓ^p-ball of minimum radius r, $\{ x : \|x\|_p^p \leq r^p \}$, defines the solution of (3).

Generally, an ideal sparsity promoting prior is an ℓ^0 norm, $\|x\|_0$, which counts the number of non-zero elements in x. However, this prior is non-convex and results in an intractable minimization problem that may have multiple local minima. As shown in Fig. 1, according to the definition 1 in section III.A, there are ℓ^p convex priors for $p \geq 1$ among which the ℓ^1 prior (Laplace prior) favors the highest degree of sparsity and thus gained the most popularity in compressed sensing and sparse regularization problems [9].

When x is not immediately sparse on its own, one can still exploit an inherent sparsity by decomposing it in an appropriate sparse representation dictionary. By definition, if x is sparse (or compressible) in a given dictionary, $\Psi = [\psi_1, ..., \psi_M] \in \mathbb{C}^{N \times M}$, most of its energy is represented by and concentrated in only a few coefficients. In other words, when x is expressed as a linear combination of $M \leq N$ elementary functions or atoms,

$$x = \Psi\theta = \sum_{m=1}^M \theta_m \psi_m,$$

most of the decomposition coefficients, θ_m, are zero and peaked around it. The dictionary Ψ is usually (i) a unitary basis whose columns (the atoms ψ_m) are pair-wise orthogonal, (ii) a tight frame or (iii) a union of bases and/or frames. For example, Ψ can be a Fourier or an orthogonal wavelet basis, a wavelet or a curvelet frame; etc. In the case of a basis set, the dictionary is a square ($M = N$) full-rank matrix whose inverse is the same as its adjoint, i.e. $\Psi^{-1} = \Psi^*$. Whereas for a tight frame, it is redundant and a non-square ($M < N$) matrix with a Moore–Penrose pseudo inverse Ψ^+. From the dictionary Ψ, we define an inverse (e.g. wavelet) transform by $W = \Psi$ and a forward transform by $W^T = \Psi^*$

or Ψ^+. In the case of a basis set, $W^TW = WW^T = I$ holds, while for a tight frame we have $WW^T = cI$, where $1/c$ is the frame's redundancy factor [10]. As a sinogram is not so sparse, we sparsify it with the forward transform W^T. Subsequently, the following compressed sensing problem is obtained from (3):

$$\hat{x} = \arg\min_{x \in X} \left\| W^T x \right\|_1, \quad X = \left\{ x \in \mathbb{R}^N \mid \Phi x = y \right\} \quad (4)$$

In fact, compressed sensing attempts to exploit the sparsity or compressibility of a true signal within a given transform domain in order to recover it from its samples. In this study, we made use of a tight frame of translation invariant (TI) wavelets as the dictionary Ψ, which was implemented by the undecimated discrete wavelet transform (UDWT), with 4 resolution levels, yielding a redundancy factor of 1:13[1]. The wavelets that we considered included: Daubechies wavelets with a vanishing moment of 4 (ψ_{D4}) and 8 (ψ_{D8}), and Daubechies 7-9 biorthogonal wavelets ($\psi_{D7}, \tilde{\psi}_{D9}$)[2].

B. Simulated CT data

Because in the clinical settings there is no ground truth CT image to objectively evaluate the performance of a proposed MAR algorithm, we simulated metal artifacts in a real chest CT dataset. Let x_t be the sinogram of the true image obtained by its forward projection In order to obtain an artifact-degraded CT image (Fig. 3b), we first artificially implanted a shoulder prosthesis in the true image, shown in Fig.3a, and then forward projected the resulting image. In the next step, the metal trace and thus the inpainting domain, Ω, was identified by thresholding. Then, the count levels in the projections on Ω were saturated according to the following empirical scheme, which gave rise to a prominent data inconsistency in the projection space. Letting $i = (r, \phi)$, the observed sinogram $\tilde{x}$ with metal trace was synthesized by:

$$\tilde{x}(i) = \begin{cases} 0.4\, x_t(i) + 0.6 \max\left\{ x_t(i), \forall i \in \Omega \right\}, & i \in \Omega \\ x_t(i), & i \notin \Omega \end{cases}$$

Finally, after FBP reconstruction, the CT numbers that overflowed the dynamic range of a typical CT display system were clipped at 3071 HU.

III. THE OPTIMIZATION ALGORITHM

The constraint set defined in (4) as well as the ℓ^1 prior are both convex (proofs follows below). Hence, the problem (4) can be considered as a convex optimization (programming) problem. We first give some basic definitions on convexity and then elaborate our optimization algorithm.

---

[1] There is one approximation subband for all resolution levels and three detail subbands for each one.
[2] This wavelet is used in the JPEG-2000 image compression standard.
[3] Also known as bound optimization or optimization transfer.

A. Definitions

Definition 1 A set $\Theta \subseteq \mathbb{R}^N$ is convex if $\forall x_1, x_2 \in \Theta$ and $\forall a \in [0,1]$, the point $x = ax_1 + (1-a)x_2$ is also in Θ.

Preposition 1 The set $X = \{x \mid \Phi x = y\}$ is a convex set.

Proof. Suppose $x_1, x_2 \in X$ i.e. $\Phi x_1 = y$, $\Phi x_2 = y$. Then $\forall a \in [0,1]$, we have

$$\Phi(ax_1 + (1-a)x_2) = a\Phi x_1 + (1-a)\Phi x_2$$
$$= ay + (1-a)y = y$$

which shows the point $ax_1 + (1-a)x_2$ is also in X.

Definition 2 A function $F(x) : \Theta \to \mathbb{R}$ is convex if $\forall x_1, x_2 \in \Theta$

$$F(ax_1 + (1-a)x_2) \le aFx_1 + (1-a)Fx_2$$

or equivalently if F is continuously differentiable,

$$F(x_2) \ge F(x_1) + \nabla F(x_1)^T (x_2 - x_1)$$

where $\nabla : \Theta \to \mathbb{R}^{N \times 2}$ is the gradient operator.

Definition 3 A continuously differentiable convex function F has a Lipschitz continuous gradient with Lipschitz constant L, if $\forall x_1, x_2 \in \Theta$

$$F(x_2) \le F(x_1) + \nabla F(x_1)^T (x_2 - x_1) + (L/2)\left\| x_2 - x_1 \right\|_2^2$$

Remark 1 The above condition is equivalent to the more standard definition of the Lipschitz continuity of the gradient [11], namely, $\forall x_1, x_2 \in \Theta$

$$\left\| \nabla F(x_2) - \nabla F(x_1) \right\|_2 \le L \left\| x_2 - x_1 \right\|_2$$

Remark 2 If $\nabla F(x)$ is L-Lipschitz, F is of class C^2 (twice continuously differentiable) and its Hessian at point x is bounded by the constant L i.e. $\nabla^2 F(x) \le LI$ [12].

Definition 4 The *indicator function* of a nonempty convex set Θ is a convex function that is defined as $D(x) = 0$ if $x \in \Theta$ and $D(x) = \infty$ if $x \notin \Theta$.

B. The Algorithm

To solve the problem (4), we follow a majorization-minimization technique[3] which replaces the original objective function with a surrogate function at each iteration [13]. Let us recast (4) into the following more general problem:

$$\hat{x} = \arg\min_x \{F(x) + D(x)\} \quad (5)$$

where $F(x) = \left\| W^T x \right\|_1$ and $D(x)$ is the indicator of the set X. Consider $\nabla F(x)$ to be L-Lipschitz, and therefore, according to the definition 3, we replace the objective function (5) by the following paraboloidal surrogate

$$Q_L(x, x^k) = F(x^k) + \nabla F(x^k)^T (x - x^k) + \frac{L}{2} \left\| x - x^k \right\|_2^2 + D(x)$$

which majorizes it near the point x^k, i.e. $Q_L(x, x^k) \ge F(x) + D(x)$. The minimization of this majorizer guarantees that of the problem (5). By completing

the squares and dropping terms independent of x, one arrives at:

$$x^{k+1} = \arg\min_x Q_L(x, x^k)$$
$$= \arg\min_x \frac{L}{2} \left\| x - \left(x^k - (1/L)\nabla F(x^k) \right) \right\|_2^2 + D(x) \quad (6)$$

The optimization (6) involves two steps: (i) gradient descent and (ii) projection onto a convex set (POCS). Hence, it is often referred to as gradient projection or projected gradient descent (PGD) [14, 15].

The Gradient descent step. The gradient of the prior $F = \left\| W^T x \right\|_1$ is given by $\nabla F(x) = W \left(W^T x / \left\| W^T x \right\|_1 \right)$. Hence, the gradient descent step with a step size $\tau = 1/L$ is given by

$$z^k = x^k - \tau W \left(W^T x^k / \left\| W^T x^k \right\|_1 \right)$$
$$= cW \left[W^T x^k \left(1 - \tau / c \left\| W^T x^k \right\|_1 \right) \right]$$

where $1/c$ is the tight frame's redundancy factor, here $c = 13$. The term inside the bracket now corresponds to a shrinkage estimator, called the soft-thresholding estimator, as applied to the wavelet coefficients $\theta^k = W^T x^k$ by a threshold $t = \tau/c$ [16]. It is defined as

$$\mathrm{T}^S(\theta^k, t) = \theta^k \left[1 - t / \left| \theta^k \right| \right]_+ \quad (7)$$

which non-linearly shrinks the coefficients toward zero. $[x]_+ = \max(0, x)$ denotes a non-negativity operator. Furthermore, for an ℓ^p wavelet-based prior, when p goes to zero, i.e. the prior approaches $\left\| W^T x^k \right\|_0$, the shrinkage estimator asymptotically becomes a hard thresholding estimator:

$$\mathrm{T}^H(\theta^k, t) = \theta^k \left(\left| \theta^k \right| > t \right) \quad (8)$$

which zeroes the coefficients having an absolute value lower than t [17]. However, this estimator leads to a non-convex optimization problem whose convergence is only guaranteed by a global search technique such as simulated annealing. For convergence, the gradient step size should satisfy $0 < \tau < 2(y_{\max})/L$, where $y_{\max}$ is the maximum photon count in y and $L = \lambda_{\max}(WW^T) \leq c$, $\lambda_{\max}(\cdot)$ is maximum eigenvalue. In this study, we also made use of a hard thresholding operator with a decaying step size (threshold) and compared it with the soft thresholding operator. In this approach which is known as graduated non-convexity [18], the decaying threshold ($t \to 0$) can play a role similar to the cooling parameter of the simulated annealing techniques, therefore allowing the solution to escape from local minima.

POCS step. The second step of the PGD algorithm in (6) is a proximity operator associated with the function $D(x)$ i.e.

$$\mathrm{prox}_{L,D}(z^k) = \arg\min_x (L/2)\left\| x - z^k \right\|_2^2 + D(x) \quad (9)$$

As $D(x)$ is an indicator function, (9) reduces to a projection onto the (convex) set X, that is, the closest point in X to z^k.

$$\mathrm{proj}_X(z^k) = \arg\min_{x \in X} \left\| x - z^k \right\|_2^2$$

For the constraint sets that are hyperplanes, this POCS operator becomes a Kaczmarz's algorithm [19] which gives

$$\mathrm{proj}_X(z^k) = z^k + \Phi^T(y - \Phi z^k) \quad (10)$$

In fact, given the definition of Φ in the present context, it can be seen that what (10) does is to merely insert the known projection data of y which are in the complement of Ω into the iterate z^k. In other words, operation (10) does the following:

$$\mathrm{proj}_X(z^k(i)) = \begin{cases} y(i), & i \in \Omega \\ z^k(i), & i \notin \Omega \end{cases} \quad (11)$$

We summarize the overall PGD algorithm in the following table.

PGD ALGORITHM

Choose $\tau^0 \in (0, 2\, y_{\max}/L)$, initialize x^0, $k > 0$,

While *stop criterion is not met* **do**

1. $\theta^k = W^T x^k$.
2. $\theta^k \leftarrow \mathrm{T}(\theta^k, \tau^k/c)$. // Soft/hard thresholding by (7) or (8).
3. Decay τ^k monotonically to zero, if hard thresholding.
4. $z^k = cW\theta^k$.
5. $x^k = \mathrm{proj}_X(z^k)$. // POCS by (10).
6. $k = k + 1$.

End do

IV. RESULTS

Figs. 3a-b show the ground truth chest image (True) and its corresponding artificially degraded 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 in a real CT acquisition. Figs. 1c shows the performance of the proposed MAR algorithm with the 7/9 biorthogonal wavelets with a hard thresholding. For quantitative assessment of performance of the considered wavelets, we defined and measured the following signal-to- noise ratio (SNR) metric in the sinogram domain:

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

Furthermore, as an indicator of the amount of artifacts in the reconstructed images, we additionally defined the following normalized total variation error metric in the image domain:

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

where g and g_t denote the image under evaluation and the true image shown in Fig. 3a, respectively. where x and x_t are respectively the sinograms of g and g_t. Table I summarizes the %TV and SNR performance of the proposed MAR algorithm for the considered wavelets and thresholding operators. As could be expected, the hard thresholding corresponding to an ℓ^0 prior far surpasses the soft

thresholding, which corresponds to an ℓ^1 prior. The performance of the wavelets in the recovery of missing projections, assessed by SNR, and reducing metal artifact, assessed by %TV, is nearly the same; however, the biorthogonal wavelets achieved better performance.
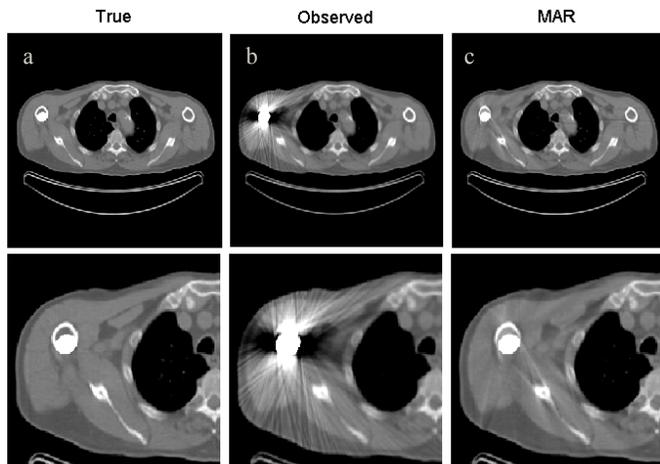

Fig. 3. a) The ground truth CT image (True) in which a shoulder prosthesis has been numerically implanted. b) The Observed CT image following filtered backprojection reconstruction algorithm. c) The metal-artifact reduced CT image reconstructed by the proposed MAR algorithm using Daubechies 7-9 biorthogonal wavelets and a hard-thresholding operator. The display window level/width is the same for all images ($WL = 40HU$, $WW = 800HU$).

TABLE I. THE %TV AND SNR PERFORMANCE OF THE PROPOSED MAR ALGORITHM.

Wavelet type	Hard Thresholding		Soft Thresholding	
	%TV	SNR (dB)	%TV	SNR (dB)
Daubechies 8	32.71	39.29	41.09	29.22
Daubechies 4	32.04	43.14	37.49	33.23
7-9 Biorthogonal	31.60	43.20	34.41	37.39

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

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

Fig. 4 illustrates the convergence rate of the studied wavelet tight frames in the recovery of missing projections as a function of the number of iterations. As seen, the 7-9 biorthogonal wavelets provide for better convergence rate. It can be explained by the fact that a CT sinogram is a smooth image dataset which can be better sparsified by those dictionaries that favor smooth textures. In fact, if a CT sinogram was piece-wise smooth, it would be better sparsified by curvelets or cubic B-spline wavelets. This also explains the better performance of the Daubechies wavelets with 4 vanishing moments over those of 8. Fig. 5 shows the computation (CPU) time of the MAR algorithm for the employed wavelets versus iteration number. For faster computation, the transforms were executed using MATLAB executable (MEX) files. However, the biorthogonal wavelets require more CPU time per iteration in comparison with the

Daubechies. It should be emphasized that orthogonal wavelets execute much faster than translation-invariant wavelets; however, their performance in the recovery of missing projections are reciprocally poorer resulting in the generation of artifacts and subsequently incorrect estimations.
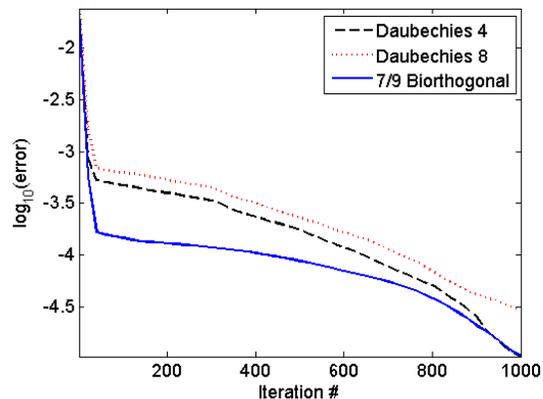

Fig. 4. The logarithm of the normalized ℓ^2 error as a function of the number of iterations indicating the convergence rate.
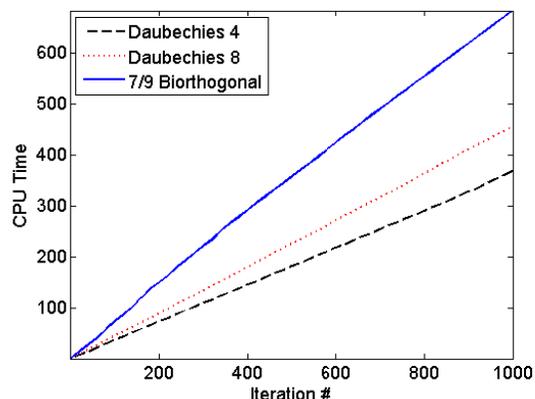

Fig. 5. Computation (CPU) time as a function of the number of iterations.

V. DISCUSSION

We proposed a new wavelet-based projection completion metal artifact reduction algorithm in x-ray CT of patients with metallic implants. In comparison with other existing MAR algorithms in this category (sinogram-based), our approach favorably exploits the following properties: first, by decomposing a sinogram into several resolution levels, this approach attempts to estimate the information hidden in missing projections, hence increasing the accuracy of the recovered data. Second, we formulated the recovery of the missing or corrupted projections as iterative hard and soft thresholding within a Bayesian paradigm, which leads to an optimal projection recovery. In fact, as noted in [17], the soft- and hard-thresholded estimators appear as MAP estimators with generalized Gaussian densities serving as a prior distribution for wavelet coefficients. For our approach, we exploited translation invariant wavelets implemented by the undecimated (stationary) wavelet transform in which the decimation (down sampling) is eliminated in favor of

invariance to shifts of the input image [20]. However due to their redundancy, these wavelets are of higher computational complexity. This renders the computational cost of our iterative MAR algorithm relatively expensive. In order to reduce the total number of the required arithmetic operations, we pursued an approach of block-wise sinogram inpainting with overlapping square blocks, schematically depicted in Fig. 5.
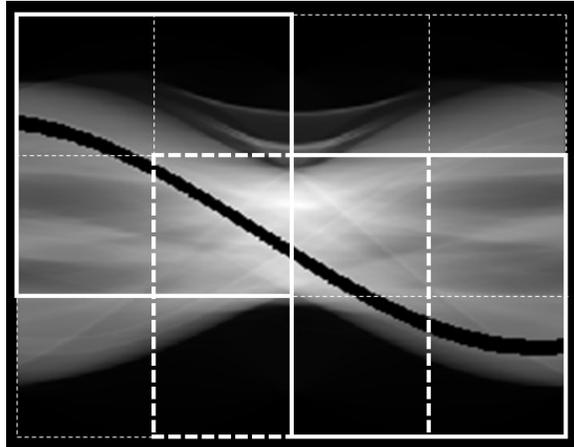

Fig. 5. A schematic of block-wise sinogram inpainting. The missing projections (the black band) are block-wise and sequentially recovered.

This approach also favors the fact that discrete wavelet transforms are typically implemented for image sizes of integer powers of two. The overlapping blocks ensure continuity and correlation between the recovered projections and hence are preferred to those without overlaps. In [21], a similar block-wise approach was also used for sinogram denoising using learned dictionaries in the context of morphological diversity analysis.

VI. CONCLUSION

In this study, we introduced a sparsity-based sinogram inpainting algorithm for reducing metal artifacts in x-ray computed tomography. The algorithm was founded within a Bayesian framework in which we made use of a wavelet-based prior to impose a sparsity constraint in the recovery of the missing projections. To objectively assess the efficacy of the proposed algorithm with respect to a ground truth, we proposed a novel approach to simulate the metal artifacts arising from shoulder prosthesis in a real patient CT dataset. It was demonstrated that proposed MAR algorithm is promising and efficient. Future work will be focus on applying the proposed algorithm in the recovery of truncated and angular under-sampled projection data.

REFERENCES

[1] R. A. Brooks and G. D. Chiro, "Beam hardening in X-ray reconstruction tomography," *Phys. Med. Biol.,* vol. 21, pp. 390-398, 1976.

[2] E. M. Kamel, C. Burger, A. Buck, G. K. von Schulthess, and G. W. Goerres, "Impact of metallic dental implants on CT-based attenuation correction in a combined PET/CT scanner," *Eur Radiol,* vol. 13, pp. 724–728, 2003.

[3] 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.

[4] 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.

[5] B. De Man, "Iterative Reconstruction for Reduction of Metal Artifacts in Computed Tomography," in *Elektrotechniek*. vol. PhD, Belgium: Katholieke Universiteit Leuven, 2001.

[6] 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.

[7] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *SIGGRAPH 2000*, New Orleans, USA, 2000, pp. 417-424.

[8] S. Zhao, D. Robertson, G. Wang, and B. Whiting, "X-ray CT metal artifact reduction using wavelets: An application for imaging total hip prostheses," *IEEE Trans. Med. Imag.,* vol. 19, pp. 1238-1247, 2000.

[9] D. L. Donoho, "Compressed sensing," *IEEE Trans. on Info. Theory,* vol. 52, pp. 1289-1306, 2006.

[10] M. Elad, *Sparse and Redundant Representations, From Theory to Applications in Signal and Image Processing*. New York: Springer, 2010.

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

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

[13] M. A. T. Figueiredo, J. M. Bioucas-Dias, and R. D. Nowak, "Majorization–Minimization Algorithms for Wavelet-Based Image Restoration," *IEEE Signal Processing Society,* vol. 16, pp. 2980 - 2991, 2007.

[14] I. Daubechies, M. Fornasier, and I. Loris, "Accelerated projected gradient method for linear inverse problems with sparsity constraints," *Journal of Fourier Analysis and Applications,* vol. 14, pp. 764-792, 2008.

[15] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.,* vol. 18, pp. 2419-2434, 2009.

[16] A. Beck and M. Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems," *SIAM J. Imaging Sci.,* vol. 2, pp. 183-202, 2009.

[17] P. Moulin and J. Liu, "Analysis of multiresolution image denoising schemes using generalized-Gaussian and complexity priors," *IEEE Transactions on Information Theory,* vol. 45, pp. 909-919, 1999.

[18] A. Blake and A. Zisserman, *Visual Reconstruction*. Cambridge, MA: MIT Press, 1987.

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

[20] P. Moulin, "Multiscale Image Decompositions and Wavelets," in *The Essential Guide to Image Processing*, A. Bovik, Ed.: Academic Press, 2009.

[21] J. Shtok, M. Elad, and M. Zibulevsky, "Sparsity-Based Sinogram Denoising for Low-Dose Computed Tomography," in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* Prague, Czech Republic, 2011.