STATISTICAL X-RAY CT RECONSTRUCTION USING A SPLITTING-BASED ITERATIVE ALGORITHM WITH ORTHONORMAL WAVELETS

Sathish Ramani and Jeffrey A. Fessler

EECS Department, University of Michigan, Ann Arbor, MI 48109-2122, USA.

ABSTRACT

Regularized image reconstruction methods are of increasing interest for X-ray CT imaging, but are hampered by the long computation times of iterative algorithms. We recently developed a variable splitting-based *alternating direction method of multipliers* (ADMM) that provides superior convergence speeds for statistical X-ray CT reconstruction compared to conventional methods. ADMM however demands storing auxiliary *constraint variables* and can become memory-expensive when shift-invariant regularization operators (e.g., finite differences) are used, especially in 3-D CT. Since an orthonormal wavelet transform (OWT) is memory-efficient, in this work, we employ OWT with nonquadratic regularization in ADMM for CT reconstruction. We propose a practical strategy for performing iteration-dependent random shifting in ADMM to (partially) compensate for the shift-variance of OWT and reduce block-artifacts therefrom. Preliminary evaluations with a 2-D synthetic phantom and real 2-D *in-vivo* human head data indicate that the proposed strategy provides CT reconstructions that are comparable in quality to those obtained using nonquadratic regularization with finite differences.

Index Terms— X-ray CT imaging, statistical image reconstruction, method of multipliers, alternating minimization, orthonormal wavelet transform.

1. INTRODUCTION

Statistical X-ray CT reconstruction using penalized weighted least squares criteria can provide improved image-quality compared to filtered back-projection (FBP) [1]. The goal usually is to minimize a cost function composed of a (statistically) weighted quadratic data-fidelity term and a (possibly nonquadratic) regularization term Ψ to suppress noise,

$$\mathbf{P0} \;:\; \arg\min_{\mathbf{x}} \left\{ J(\mathbf{x}) \triangleq \frac{1}{2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|_{\mathbf{W}}^2 + \Psi(\mathbf{R}\mathbf{x}) \right\}, \quad (1)$$

where $\mathbf{y}$ is the $M \times 1$ data vector (log of transmission data), $\mathbf{A}$ is the $M \times N$ system matrix, $\mathbf{A}\mathbf{x}$ represents the forward projection operation (e.g., line integrals), $\mathbf{W} = \text{diag}\{w_i\}$ is a $M \times M$ diagonal matrix consisting of statistical weights (for simplicity we used $w_i = e^{-y_i}$ in our experiments), $\|\mathbf{u}\|_{\mathbf{W}}^2 \triangleq \mathbf{u}^\top \mathbf{W}\mathbf{u}$ and $\mathbf{R}$ denotes a regularization operator with shift-invariant blocks, e.g., finite differences, wavelet frames, etc.

Conventional gradient-descent methods (e.g., nonlinear conjugate gradient) for $\mathbf{P0}$ depend on the Hessian $\mathbf{A}^\top \mathbf{W}\mathbf{A}$ that is highly shift-variant in CT particularly due to the large dynamic range of $\{w_i\}$. As a result, it becomes difficult to precondition and accelerate such methods [2]. We recently introduced an iterative algorithm—alternating direction method of multipliers (ADMM) [3]—to circumvent this difficulty. The philosophy underlying ADMM is to convert $\mathbf{P0}$ in to an equivalent constrained problem using variable splitting [3]:

$$\min_{\mathbf{x},\mathbf{u},\mathbf{v}} \{ \frac{1}{2}\|\mathbf{y}-\mathbf{u}\|_{\mathbf{W}}^2 + \Psi(\mathbf{v}) \} \text{ s.t. } \mathbf{u} = \mathbf{A}\mathbf{x},\ \mathbf{v} = \mathbf{R}\mathbf{x}, \quad (2)$$

where $\mathbf{u}$, $\mathbf{v}$ are auxiliary constraint variables: $\mathbf{u}$ "isolates" the shift-variant component $\mathbf{W}$ from $\mathbf{A}\mathbf{x}$ and $\mathbf{v}$ splits the regularization term in $\mathbf{P0}$, respectively. The constrained problem (2) is then solved in an augmented Lagrangian (AL) framework with alternating minimization that decouples (2) in terms of $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{x}$. As a result, ADMM involves simple operations such as inverting a diagonal matrix, minimizing 1-D denoising cost functions, and the solving of a "nearly" shift-invariant linear system [3] that is amenable to FFT-based preconditioning using cone-type filters [4]. These key features enable accelerated convergence of ADMM as demonstrated in [3]. Since ADMM is based on the AL framework, it requires storing the constraint variables $(\mathbf{u}, \mathbf{v})$ and associated Lagrange multiplier vectors at each iteration [3, Sec. III, VI-A]. When $\mathbf{R}$ is a "tall" matrix of large proportions (e.g., finite differences along many directions), storing $\mathbf{v}$ (and its Lagrange multiplier) can represent a significant memory overhead, especially in 3-D CT.

In this work, we propose to use regularizers involving orthonormal wavelet transform (OWT) for $\mathbf{P0}$ as OWT requires relatively less memory compared to products with (a "tall") $\mathbf{R}$. OWTs however are shift-variant and tend to yield block-artifacts in reconstructions [5]. Iteration-dependent random shifting (IDRS) [5] is a simple means of reducing block-artifacts due to OWTs and has been successfully used for image restoration [5] and MRI reconstruction [6] in conjunction with iterative shrinkage-thresholding (IST) type methods. Here, we present a strategy to incorporate IDRS in ADMM for statistical CT reconstruction using OWT-based regularizers. We show that the proposed strategy (with OWT and IDRS) retains the efficacy of the ADMM originally developed in [3] and demonstrate with experiments on a 2-D synthetic phantom and real 2-D *in-vivo* human head data that it yields CT reconstructions that are comparable in quality to those obtained using nonquadratic regularizers (e.g., total variation) with shift-invariant operators such as finite differences. The proposed method can also be readily applied to 3-D CT using a 3-D OWT.

2. ALTERNATING DIRECTION METHOD OF MULTIPLIERS (ADMM)

We summarize[1] below the ADMM developed in [3] for solving $\mathbf{P0}$. At iteration j, one performs:

This work was supported by the National Institutes of Health under Grant R01-HL-098686. The authors would like to thank J. Seamans, J. B. Thibault, and B. DeMan, GE Healthcare, for providing the *in-vivo* human head data-set used in the experiments and code for the DD-projector.

[1] Detailed derivation of ADMM, convergence theory and empirical rules for setting μ and ν are available in [3].

$$\mathbf{x}^{(j+1)} = \mathbf{G}_\nu^{-1}(\mathbf{A}^\top(\mathbf{u}^{(j)} - \boldsymbol{\eta}_{\mathbf{u}}^{(j)}) + \nu\mathbf{R}^\top(\mathbf{v}^{(j)} - \boldsymbol{\eta}_{\mathbf{v}}^{(j)})), \quad (3)$$

$$\mathbf{u}^{(j+1)} = \mathbf{D}_\mu^{-1}(\mathbf{W}\mathbf{y} + \mu\boldsymbol{\varrho}_{\mathbf{u}}^{(j)}), \quad (4)$$

$$\mathbf{v}^{(j+1)} = \mathbf{d}_\Psi\{\boldsymbol{\varrho}_{\mathbf{v}}^{(j)}\} \triangleq \arg\min_{\mathbf{v}} \left\{\Psi(\mathbf{v}) + \frac{\mu\nu}{2}\|\mathbf{v} - \boldsymbol{\varrho}_{\mathbf{v}}^{(j)}\|_2^2\right\}, \quad (5)$$

$$\boldsymbol{\eta}_{\mathbf{u}}^{(j+1)} = \boldsymbol{\varrho}_{\mathbf{u}}^{(j)} - \mathbf{u}^{(j+1)}, \quad (6)$$

$$\boldsymbol{\eta}_{\mathbf{v}}^{(j+1)} = \boldsymbol{\varrho}_{\mathbf{v}}^{(j)} - \mathbf{v}^{(j+1)}, \quad (7)$$

where $\boldsymbol{\eta}_{\mathbf{u}}^{(\cdot)}$, $\boldsymbol{\eta}_{\mathbf{v}}^{(\cdot)}$ are Lagrange-multiplier-like vectors for the constraints in (2), $\boldsymbol{\varrho}_{\mathbf{u}}^{(j)} \triangleq \mathbf{A}\mathbf{x}^{(j+1)} + \boldsymbol{\eta}_{\mathbf{u}}^{(j)}$, $\boldsymbol{\varrho}_{\mathbf{v}}^{(j)} \triangleq \mathbf{R}\mathbf{x}^{(j+1)} + \boldsymbol{\eta}_{\mathbf{v}}^{(j)}$, $\mu > 0$, $\nu > 0$ are parameters[1] that govern only the convergence speed of the algorithm, $\mathbf{G}_\nu \triangleq (\mathbf{A}^\top\mathbf{A} + \nu\mathbf{R}^\top\mathbf{R})$, $\mathbf{D}_\mu \triangleq (\mathbf{W} + \mu\mathbf{I}_M)$ and $\mathbf{I}_M$ is the identity matrix of size M. Since $\mathbf{W}$ is diagonal, $\mathbf{D}_\mu$ is also diagonal and can be inverted exactly in (4). The term $\mathbf{A}^\top\mathbf{A}$ in $\mathbf{G}_\nu$ is "nearly" shift-invariant [4] (unlike $\mathbf{A}^\top\mathbf{W}\mathbf{A}$), so for $\mathbf{R}$ such that $\mathbf{R}^\top\mathbf{R}$ is shift-invariant, an iterative solver such as the conjugate gradient (CG) method for (3) is amenable to FFT-based preconditioning using suitable *cone filters* [4].

Moreover, $\mathbf{v}$ (2) leads to the denoising problem (5) (represented by a denoising operator $\mathbf{d}_\Psi$) that can be solved exactly for several instances of Ψ [7, Sec. 4] including popular criteria such as total variation (TV):

$$\Psi_{\mathrm{TV}}(\mathbf{R}\mathbf{x}) = \lambda \sum_r \sqrt{\sum_{p=0}^{P-1} |[\mathbf{R}_p\,\mathbf{x}]_r|^2}, \quad (8)$$

ℓ_1-regularization:

$$\Psi_{\ell_1}(\mathbf{R}\mathbf{x}) = \lambda \sum_r |[\mathbf{R}\,\mathbf{x}]_r|\,, \quad (9)$$

and smoothed-Laplacian (SL) that has been successfully used in tomography [2]:

$$\Psi_{\mathrm{SL}}(\mathbf{R}\mathbf{x}) = \lambda \sum_r \Phi_{\mathrm{SL}}\left(|[\mathbf{R}\,\mathbf{x}]_r|\right), \quad (10)$$

where $\lambda > 0$ is the regularization parameter, $\Phi_{\mathrm{SL}}(x) = x/\delta - \log(1 + x/\delta)$, $\delta > 0$ [2] and $\mathbf{R} \triangleq [\mathbf{R}_0{}^\top \cdots \mathbf{R}_{P-1}{}^\top]^\top$ is a block matrix of $N \times N$ shift-invariant regularization operators $\mathbf{R}_p$ (e.g., finite differences). For the regularizers in (8)-(10), it can be shown that (5) decouples in terms of the components, $\{v_r\}$ and $\varrho_{\mathbf{v},r}^{(j)}$, of $\mathbf{v}$ and $\boldsymbol{\varrho}_{\mathbf{v}}^{(j)}$, respectively, i.e., $v_r^{(j+1)} = d_\Psi(\varrho_{\mathbf{v},r}^{(j)})$, $r = 0 \dots PN-1$ [8], and leads to the following respective mappings [8]:

$$d_{\Psi_{\mathrm{TV}}}(\varrho_{\mathbf{v},r}^{(j)}) = \varrho_{\mathbf{v},r}^{(j)} \max\left(1 - \frac{\lambda}{\mu\nu\beta_r^{(j)}}, 0\right), \quad (11)$$

$$d_{\Psi_{\ell_1}}(\varrho_{\mathbf{v},r}^{(j)}) = \mathrm{sign}\{\varrho_{\mathbf{v},r}^{(j)}\} \max\left(|\varrho_{\mathbf{v},r}^{(j)}| - \frac{\lambda}{\mu\nu}, 0\right), \quad (12)$$

$$d_{\Psi_{\mathrm{SL}}}(\varrho_{\mathbf{v},r}^{(j)}) = \mathrm{sign}\{\varrho_{\mathbf{v},r}^{(j)}\} \frac{\zeta_r^{(j)} + \sqrt{(\zeta_r^{(j)})^2 + 4\delta|\varrho_{\mathbf{v},r}^{(j)}|}}{2}, \quad (13)$$

where, for $r = 0 \dots PN-1$, $\zeta_r^{(j)} \triangleq |\varrho_{\mathbf{v},r}^{(j)}| - \delta - \lambda/(\delta\mu\nu)$ and $\beta_r^{(j)} \triangleq \sqrt{\sum_{p=0}^{P-1} |\varrho_{\mathbf{v},pN+(r \bmod N)}^{(i)}|^2}$.

In addition to the $N \times 1$ reconstruction $\mathbf{x}^{(\cdot)}$, ADMM also requires storing the $M \times 1$ vectors $\mathbf{u}^{(\cdot)}$, $\boldsymbol{\eta}_{\mathbf{u}}^{(\cdot)}$ and the $PN \times 1$ vectors $\mathbf{v}^{(\cdot)}$, $\boldsymbol{\eta}_{\mathbf{v}}^{(\cdot)}$. Depending on the type of scanner geometry and reconstruction setup, $PN > M \gtrapprox N$. For instance, in 3-D CT, when finite-differences are used for $\mathbf{R}$ with $P = 13$ (there are 13 nearest-neighbors on one side of any voxel), $\mathbf{v}^{(\cdot)}$, $\boldsymbol{\eta}_{\mathbf{v}}^{(\cdot)}$ together correspond to 26 image-volumes: storing $\mathbf{v}^{(\cdot)}$, $\boldsymbol{\eta}_{\mathbf{v}}^{(\cdot)}$ might thus set a practical limitation on the applicability of ADMM for $\mathbf{R}$ (with shift-invariant blocks), especially in 3-D CT.

1. Select $\mathbf{x}^{(0)}$, $\boldsymbol{\mathcal{S}}^{(0)}$, μ, $\nu > 0$ and set $j = 0$
2. Set $\mathbf{u}^{(0)} = \mathbf{A}\mathbf{x}^{(0)}$, $\widehat{\mathbf{v}}^{(0)} = \boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{S}}^{(0)}\mathbf{x}^{(0)}$, and $\boldsymbol{\eta}_{\mathbf{u}}^{(0)} = \boldsymbol{\eta}_{\widehat{\mathbf{v}}}^{(0)} = \mathbf{0}$

Repeat:

3. Obtain $\mathbf{x}^{(j+1)}$ by applying preconditioned CG to (20)
4. Compute $\mathbf{u}^{(j+1)}$ using (4)
5. Compute $\widehat{\mathbf{v}}^{(j+1)}$ using (21) with $\widetilde{\boldsymbol{\mathcal{W}}}^{(j)} = \boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{S}}^{(j)}$
6. Update $\boldsymbol{\eta}_{\mathbf{u}}^{(j+1)}$ using (5)
7. Update $\boldsymbol{\eta}_{\widehat{\mathbf{v}}}^{(j+1)}$ using (22)
8. Set $j = j + 1$

Until stop criterion is met

Fig. 1. ADMM with IDRS for statistical X-ray CT reconstruction using OWT. The algorithm uses a pregenerated sequence of random translations represented by $\{\boldsymbol{\mathcal{S}}^{(j)}\}_j$.

3. PROPOSED STRATEGY

3.1. Orthonormal Wavelets for CT Reconstruction Using ADMM

We propose to use orthonormal wavelet transform (OWT) as a memory-efficient alternative to $\mathbf{R}$ for CT reconstruction using ADMM. OWT has several attractive properties (e.g., regularity, sparsity) including its ability to represent an image at multiple scales. For brevity, we consider the following nonquadratic regularizers with an OWT denoted by the $N \times N$ matrix $\boldsymbol{\mathcal{W}}$:

$$\Psi_{\ell_1,\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{W}}\mathbf{x}) = \sum_{r=1}^{N} \lambda_r\,|[\boldsymbol{\mathcal{W}}\,\mathbf{x}]_r|\,, \quad (14)$$

$$\Psi_{\mathrm{SL},\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{W}}\mathbf{x}) = \sum_{r=1}^{N} \lambda_r \Phi_{\mathrm{SL}}\left(|[\boldsymbol{\mathcal{W}}\,\mathbf{x}]_r|\right), \quad (15)$$

where $\{\lambda_r\}_{r=1}^N$ are scale-dependent regularization parameters that offer the flexibility of excluding the approximation coefficients (this can be achieved by setting $\lambda_r = 0$ corresponding to the approximation coefficients) from $\Psi_{\ell_1,\boldsymbol{\mathcal{W}}}$ and $\Psi_{\mathrm{SL},\boldsymbol{\mathcal{W}}}$ as these coefficients are not sparse.

We employ $\bar{\mathbf{v}} = \boldsymbol{\mathcal{W}}\mathbf{x}$ in (2), so that $\bar{\mathbf{v}}^{(\cdot)}$ and $\boldsymbol{\eta}_{\bar{\mathbf{v}}}^{(\cdot)}$ in the corresponding ADMM are of the same size as the reconstruction $\mathbf{x}^{(\cdot)}$ and require relatively less storage compared to $\mathbf{v}^{(\cdot)}$ and $\boldsymbol{\eta}_{\mathbf{v}}^{(\cdot)}$ that are associated with $\mathbf{R}$. The denoising rules (12)-(13) are also applicable for the regularizers (14)-(15), respectively, at Step (5). Moreover, orthonormality of $\boldsymbol{\mathcal{W}}$ (i.e., $\boldsymbol{\mathcal{W}}^\top\boldsymbol{\mathcal{W}} = \mathbf{I}_N$) facilitates Step (3) since $\mathbf{G}_\nu$ in (3) becomes $\mathbf{G}_{\nu,\boldsymbol{\mathcal{W}}} \triangleq (\mathbf{A}^\top\mathbf{A} + \nu\boldsymbol{\mathcal{W}}^\top\boldsymbol{\mathcal{W}}) = (\mathbf{A}^\top\mathbf{A} + \nu\mathbf{I}_N)$ that is still "nearly" shift-invariant and can be effectively preconditioned using cone-type filters [4]. Thus, (3)-(7) can be directly employed for regularized statistical CT reconstruction using (14)-(15).

3.2. Iteration-Dependent Random Shifting (IDRS)

While OWT is appealing from the standpoint of memory, its shift-variant nature can lead to block-artifacts in the reconstruction [5, 6]. Iteration dependent random shifting (IDRS) [5] is a simple technique to partially compensate for the shift-variance of OWT. The idea underlying IDRS is to first apply a (random) translation (at each iteration) to the reconstruction ($\mathbf{x}^{(\cdot)}$), then perform nonlinear processing associated with the reconstruction algorithm and finally undo the translation before proceeding to the next iteration. IDRS is computationally efficient as it only requires trivial translation operations and is easily incorporated in iterative shrinkage-thresholding

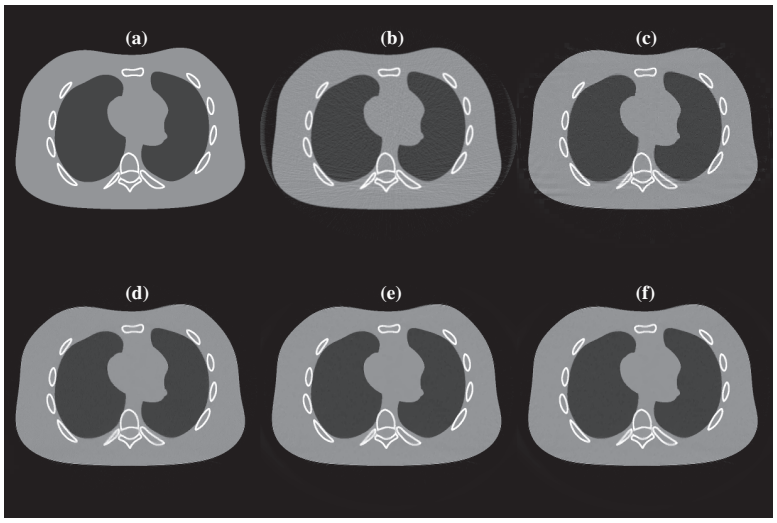


Fig. 2. Simulation (color scale is in cm^{-1}): (a) True NCAT phantom, (b) FBP reconstruction with Hanning filter ($\mathrm{RMSE} = 0.0322\,\mathrm{cm}^{-1}$), and reconstructions obtained using ADMM with (c) $\Psi_{\ell_1,\boldsymbol{\mathcal{W}}}$ (14) and no IDRS ($0.0294\,\mathrm{cm}^{-1}$), (d) $\Psi_{\ell_1,\boldsymbol{\mathcal{W}}}$ (15) with IDRS (proposed strategy, $0.0214\,\mathrm{cm}^{-1}$), (e) Ψ_{ℓ_1} (10) with finite differences ($0.0194\,\mathrm{cm}^{-1}$) and (f) total variation (TV, $0.0198\,\mathrm{cm}^{-1}$). The proposed method with OWT and IDRS produces a reconstruction (d) comparable to (e) and (f) obtained, respectively, using ℓ_1- and TV-regularizers with finite differences.

(IST) type algorithms [5, 6]. However, using IDRS with ADMM for OWT-based regularizers (14)-(15) is slightly more involved as ADMM (3)-(7) employs additional constraint variables. Below, we present a practical strategy to integrate IDRS in ADMM.

Let $\boldsymbol{\mathcal{S}}$ represent a generic $N \times N$ (block) permutation matrix such that $\boldsymbol{\mathcal{S}}\mathbf{x}$ is a (randomly) translated version of $\mathbf{x}$. By construction $\boldsymbol{\mathcal{S}}$ is an orthonormal matrix, i.e., $\boldsymbol{\mathcal{S}}^\top\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{S}}^\top = \mathbf{I}_N$, since $\boldsymbol{\mathcal{S}}^\top$ undoes the effect of $\boldsymbol{\mathcal{S}}$ and vice-versa. We then consider (14)-(15) with $\boldsymbol{\mathcal{S}}\mathbf{x}$ (as we want to "induce" shift-invariance in these OWT-based regularizers), which is equivalent to replacing $\boldsymbol{\mathcal{W}}$ with $\widetilde{\boldsymbol{\mathcal{W}}} = \boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{S}}$ in (14)-(15). Correspondingly, we use $\widetilde{\mathbf{v}} = \widetilde{\boldsymbol{\mathcal{W}}}\mathbf{x}$ in (2) that leads to a scheme similar to (3)-(7) but with the following respective update rules for Steps (3), (5), and (7):

$$\mathbf{x}^{(j+1)} = \mathbf{G}^{-1}_{\nu,\widetilde{\boldsymbol{\mathcal{W}}}}(\mathbf{A}^\top(\mathbf{u}^{(j)} - \boldsymbol{\eta}^{(j)}_{\mathbf{u}}) + \nu\widetilde{\boldsymbol{\mathcal{W}}}^\top(\widetilde{\mathbf{v}}^{(j)} - \boldsymbol{\eta}^{(j)}_{\widetilde{\mathbf{v}}})), \quad (16)$$

$$\widetilde{\mathbf{v}}^{(j+1)} = \arg\min_{\mathbf{v}} \left\{ \Psi_{\cdot,\widetilde{\boldsymbol{\mathcal{W}}}}(\mathbf{v}) + \frac{\mu\nu}{2}\|\mathbf{v} - \boldsymbol{\varrho}^{(j)}_{\widetilde{\mathbf{v}}}\|_2^2 \right\}, \quad (17)$$

$$\boldsymbol{\eta}^{(j+1)}_{\widetilde{\mathbf{v}}} = \boldsymbol{\varrho}^{(j)}_{\widetilde{\mathbf{v}}} - \widetilde{\mathbf{v}}^{(j+1)}, \quad (18)$$

where $\boldsymbol{\eta}^{(j)}_{\widetilde{\mathbf{v}}}$ is the Lagrange-multiplier-like vector associated with $\widetilde{\mathbf{v}}^{(j)}$, $\boldsymbol{\varrho}^{(j)}_{\widetilde{\mathbf{v}}} \triangleq \widetilde{\boldsymbol{\mathcal{W}}}\mathbf{x}^{(j+1)} + \boldsymbol{\eta}^{(j)}_{\widetilde{\mathbf{v}}}$, and $\mathbf{G}_{\nu,\widetilde{\boldsymbol{\mathcal{W}}}} \triangleq (\mathbf{A}^\top\mathbf{A} + \nu\widetilde{\boldsymbol{\mathcal{W}}}^\top\widetilde{\boldsymbol{\mathcal{W}}})$. Since $\boldsymbol{\mathcal{W}}$ and $\boldsymbol{\mathcal{S}}$ are both orthonormal, $\widetilde{\boldsymbol{\mathcal{W}}}^\top\widetilde{\boldsymbol{\mathcal{W}}} = \boldsymbol{\mathcal{S}}^\top\boldsymbol{\mathcal{W}}^\top\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{S}} = \mathbf{I}_N$, so we have that

$$\mathbf{G}_{\nu,\widetilde{\boldsymbol{\mathcal{W}}}} = (\mathbf{A}^\top\mathbf{A} + \nu\mathbf{I}_N) = \mathbf{G}_{\nu,\boldsymbol{\mathcal{W}}}. \quad (19)$$

Using the fact that $\widetilde{\boldsymbol{\mathcal{W}}}^{-1} = \widetilde{\boldsymbol{\mathcal{W}}}^\top$ (orthonormality of $\widetilde{\boldsymbol{\mathcal{W}}}$), and $\widehat{\mathbf{v}}^{(\cdot)} \triangleq \widetilde{\boldsymbol{\mathcal{W}}}^\top\widetilde{\mathbf{v}}^{(\cdot)}$, $\boldsymbol{\eta}^{(\cdot)}_{\widehat{\mathbf{v}}} \triangleq \widetilde{\boldsymbol{\mathcal{W}}}^\top\boldsymbol{\eta}^{(\cdot)}_{\widetilde{\mathbf{v}}}$, (16)-(18) can be written, respectively, as

$$\mathbf{x}^{(j+1)} = \mathbf{G}^{-1}_{\nu,\boldsymbol{\mathcal{W}}}(\mathbf{A}^\top(\mathbf{u}^{(j)} - \boldsymbol{\eta}^{(j)}_{\mathbf{u}}) + \nu(\widehat{\mathbf{v}}^{(j)} - \boldsymbol{\eta}^{(j)}_{\widehat{\mathbf{v}}})), \quad (20)$$

$$\widehat{\mathbf{v}}^{(j+1)} = \widetilde{\boldsymbol{\mathcal{W}}}^\top \times \left(\arg\min_{\mathbf{v}} \left\{ \begin{matrix} \Psi_{\cdot,\widetilde{\boldsymbol{\mathcal{W}}}}(\mathbf{v}) \\ + \frac{\mu\nu}{2}\|\mathbf{v} - \widetilde{\boldsymbol{\mathcal{W}}}(\mathbf{x}^{(j+1)} + \boldsymbol{\eta}^{(j)}_{\widehat{\mathbf{v}}})\|_2^2 \end{matrix} \right\} \right), \quad (21)$$

$$\boldsymbol{\eta}^{(j+1)}_{\widehat{\mathbf{v}}} = \boldsymbol{\eta}^{(j)}_{\widehat{\mathbf{v}}} - (\widehat{\mathbf{v}}^{(j+1)} - \mathbf{x}^{(j+1)}), \quad (22)$$

Due to the "nearly" shift-invariant structure of $\mathbf{G}_{\nu,\boldsymbol{\mathcal{W}}}$ (19), cone-filter-type preconditioners can again be used to accelerate CG-based solvers for (20). For the OWT-based regularizers (14)-(15), denoising rules (12)-(13) can be respectively used for solving (21). We see that the (random) translation (included in $\widetilde{\boldsymbol{\mathcal{W}}}$) is introduced in the input (i.e., $\widetilde{\boldsymbol{\mathcal{W}}}(\mathbf{x}^{(j+1)} + \boldsymbol{\eta}^{(j)}_{\widehat{\mathbf{v}}})$) to the denoising step (21) and is undone in the same step (using $\widetilde{\boldsymbol{\mathcal{W}}}^\top$) after performing the denoising operation, this is somewhat similar to how IDRS is typically applied to IST-type algorithms [5, 6]. Finally, since the development (16)-(22) applies to a generic $\boldsymbol{\mathcal{S}}$, we can employ a different $\boldsymbol{\mathcal{S}}^{(j)}$ (and correspondingly, $\widetilde{\boldsymbol{\mathcal{W}}}^{(j)} = \boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{S}}^{(j)}$) at each iteration as suggested in [5], thus leading to the ADMM algorithm with IDRS in Fig. 1 for CT reconstruction using OWT-based regularizers.

4. EXPERIMENTAL RESULTS

As a preliminary evaluation, we performed simulations with a 1024×1024 2-D slice of the NCAT phantom [9] and experiments with a 2-D *in-vivo* human head data-set (888×984-view sinogram) acquired with a GE scanner using 120 kVp source potential, 585 mA tube current and 0.6 s rotation. We used the DD-projector [10] (with 8 threads) for implementing matrix-vector products such as $\mathbf{Ax}$, $\mathbf{A}^\top\mathbf{u}$ and initialized ADMM with the image reconstructed using FBP with the ramp filter. Computation time (per iteration) of ADMM (scheme (3)-(7) and the proposed method in Fig. 1) is dominated by products with $\mathbf{A}$ and $\mathbf{A}^\top$ in (3) and (20), and is similar for the considered regularizers (8)-(10) and (14)-(15). So we concentrated on the image quality instead. We ran 100 iterations of ADMM (3)-(7) for regularizers (8)-(10) with finite-differences for $\mathbf{R}$ and the proposed scheme in Fig. (1) for (14)-(15) with 3 levels of the orthonormal Haar wavelet transform for $\boldsymbol{\mathcal{W}}$. We applied 2 CG iterations with appropriate cone-filter-type preconditioners [3] for "solving" (3) and (20), respectively. We set $\lambda_r = \lambda$ and excluded the approximation coefficients in (14)-(15).

In our simulations, we numerically generated a 888×984-view noisy sinogram with GE LightSpeed fan-beam geometry [3] corresponding to a monoenergetic source with 2.5×10^4 incident photons per ray and no background events. We reconstructed 512×512 images over a FOV of 65 cm. For the regularized methods, we chose λ by minimizing the RMSE between the true noisefree phantom (Fig. 2a) and the corresponding regularized reconstructions (Figs. 2c-2f). For completeness, we include FBP reconstruction (Fig. 2b, obtained using Hanning filter) that is blurred and streaked with artifacts. The ℓ_1-regularized (using Ψ_{ℓ_1} (9)) and TV-regularized reconstructions (Figs. 2e, 2f, respectively) have reduced noise, RMSE, and artifacts.

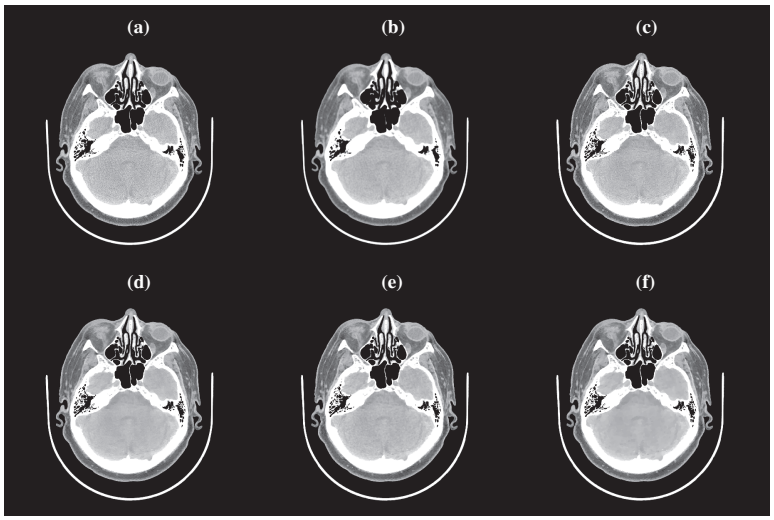


Fig. 3. Experiment with *in-vivo* human head data (color scale is in Hounsfield units): FBP reconstructions with (a) ramp filter, (b) Hanning filter, and regularized reconstructions obtained using ADMM with (c) $\Psi_{\mathrm{SL},\mathcal{W}}$ (15) and no IDRS, (d) $\Psi_{\mathrm{SL},\mathcal{W}}$ (15) with IDRS (proposed strategy), (e) Ψ_{SL} (10) with finite differences and (f) total variation (TV). Reconstruction (f) is "patchy" indicating that (strongly edge-preserving) TV is less appealing for this data. The proposed method with OWT (and smooth edge-preserving $\Psi_{\mathrm{SL},\mathcal{W}}$) produces a reconstruction (d) comparable to (e) obtained with Ψ_{SL} (10) and finite differences.

Using $\Psi_{\ell_1,\mathcal{W}}$ (14) without IDRS leads to a reconstruction (Fig. 2c) with shading artifacts and higher RMSE, while the proposed method (using $\Psi_{\ell_1,\mathcal{W}}$ (14) with IDRS) yields a reconstruction (Fig. 2d) that is comparable in quality and RMSE to those in Figs. 2e, 2f.

In case of the real *in-vivo* human head data, we reconstructed 1024×1024 images with 50 cm FOV. We set $\delta = 10$ HU in Ψ_{SL} (10) and $\Psi_{\mathrm{SL},\mathcal{W}}$ (15), and adjusted λ manually for the regularized methods to obtain the "best" visual quality (in terms of suppressing noise and streak-artifacts) in each case individually. The FBP reconstructions in Figs. 3a, 3b corresponding to the ramp and Hanning filters, respectively, are either noisy or blurred and streaked with artifacts. The TV-regularized reconstruction (Fig. 3f) exhibits piecewise constant "patches" that are due to the strongly edge-preserving nature of the TV-regularizer. In contrast, the smooth edge-preserving regularizer Ψ_{SL} (10) with finite differences provides a visually more appealing reconstruction (Fig. 3e). The OWT-based reconstruction (Fig. 3c) obtained using $\Psi_{\mathrm{SL},\mathcal{W}}$ (15) without IDRS exhibits noise-like artifacts, but the proposed method (using $\Psi_{\mathrm{SL},\mathcal{W}}$ (15)) with IDRS suppresses such artifacts and yields a reconstruction (Fig. 3d) comparable to that in Fig. 3e obtained with Ψ_{SL} (10) and finite differences.

5. CONCLUSIONS & DISCUSSION

Iteration-dependent random shifting (IDRS) [5] is a simple procedure that can reduce block-artifacts in reconstruction problems involving an orthonormal wavelet transform (OWT) [5]. In this work, we proposed a practical scheme to incorporate IDRS in alternating direction method of multipliers (ADMM) [3] for statistical X-ray CT reconstruction using OWT-based regularizers. We demonstrated using experiments with a synthetic 2-D NCAT phantom and real 2-D *in-vivo* human head data that the proposed method provides reconstructions comparable in quality to those obtained using regularizers with shift-invariant operators such as finite differences. The proposed method can also be directly applied to 3-D CT using a 3-D OWT. Reconstruction quality may further be improved in statistical methods for CT reconstruction by using spatially-varying regularization parameters (SVRP) [11]. We are currently implementing SVRP in the proposed scheme with application to 3-D CT reconstruction.

6. REFERENCES

[1] J-B. Thibault, K. Sauer, C. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multi-slice helical CT," *Med. Phys.*, vol. 34, no. 11, pp. 4526–44, Nov. 2007.

[2] J. A. Fessler and S. D. Booth, "Conjugate-gradient preconditioning methods for shift-variant PET image reconstruction," *IEEE Trans. Im. Proc.*, vol. 8, no. 5, pp. 688–99, May 1999.

[3] S. Ramani and J. A. Fessler, "A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction," *IEEE Trans. Med. Imag.*, 2012, to appear, DOI: 10.1109/TMI.2011.2175233, `http://tinyurl.com/admmct2d`.

[4] N. H. Clinthorne, T. S. Pan, P. C. Chiao, W. L. Rogers, and J. A. Stamos, "Preconditioning methods for improved convergence rates in iterative reconstructions," *IEEE Trans. Med. Imag.*, vol. 12, no. 1, pp. 78–83, Mar. 1993.

[5] M. A. T. Figueiredo and R. D. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Trans. Im. Proc.*, vol. 12, no. 8, pp. 906–16, Aug. 2003.

[6] M. Guerquin-Kern, M. Haberlin, K.P. Pruessmann, and M. Unser, "A fast wavelet-based reconstruction method for magnetic resonance imaging," *IEEE Trans. Med. Imag.*, vol. 30, no. 9, pp. 1649–1660, 2011.

[7] C. Chaux, P. L. Combettes, J-C. Pesquet, and Valérie R Wajs, "A variational formulation for frame-based inverse problems," *Inverse Prob.*, vol. 23, no. 4, pp. 1495–518, Aug. 2007.

[8] S. Ramani and J. A. Fessler, "Parallel MR image reconstruction using augmented Lagrangian methods," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 694–706, Mar. 2011.

[9] W. P. Segars and B. M. W. Tsui, "Study of the efficacy of respiratory gating in myocardial SPECT using the new 4-D NCAT phantom," *IEEE Trans. Nuc. Sci.*, vol. 49, no. 3, pp. 675–9, June 2002.

[10] B. De Man and S. Basu, "Distance-driven projection and back-projection," in *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, 2002, vol. 3, pp. 1477–80.

[11] J. A. Fessler and W. L. Rogers, "Spatial resolution properties of penalized-likelihood image reconstruction methods: Space-invariant tomographs," *IEEE Trans. Im. Proc.*, vol. 5, no. 9, pp. 1346–58, Sept. 1996.