Penalized Weighted Least-Squares Approach for Low-Dose X-Ray Computed Tomography

Jing Wang[*1,2], Tianfang Li[3], Hongbing Lu[4], and Zhengrong Liang[1,2]

Departments of Radiology[1] and Physics and Astronomy[2], State University of New York, Stony Brook, NY 11794, USA
Department of Radiation Oncology[3], Stanford University, Stanford, CA 94305, USA
Department of Biomedical Engineering/Computer Applications[4], Fourth Military Medical University, Xi'an, Shaanxi 710032, China

ABSTRACT

The noise of low-dose computed tomography (CT) sinogram follows approximately a Gaussian distribution with nonlinear dependence between the sample mean and variance. The noise is statistically uncorrelated among detector bins at any view angle. However the correlation coefficient matrix of data signal indicates a strong signal correlation among neighboring views. Based on above observations, Karhunen-Loève (KL) transform can be used to de-correlate the signal among the neighboring views. In each KL component, a penalized weighted least-squares (PWLS) objective function can be constructed and optimal sinogram can be estimated by minimizing the objective function, followed by filtered backprojection (FBP) for CT image reconstruction. In this work, we compared the KL-PWLS method with an iterative image reconstruction algorithm, which uses the Gauss-Seidel iterative calculation to minimize the PWLS objective function in image domain. We also compared the KL-PWLS with an iterative sinogram smoothing algorithm, which uses the iterated conditional mode calculation to minimize the PWLS objective function in sinogram space, followed by FBP for image reconstruction. Phantom experiments show a comparable performance of these three PWLS methods in suppressing the noise-induced artifacts and preserving resolution in reconstructed images. Computer simulation concurs with the phantom experiments in terms of noise-resolution tradeoff and detectability in low contrast environment. The KL-PWLS noise reduction may have the advantage in computation for low-dose CT imaging, especially for dynamic high-resolution studies.

Keywords: penalized weighted least-squares (PWLS), KL transform, low-dose CT, noise-resolution tradeoff, ROC curve.

1. INTRODUCTION

Low-dose X-ray computed tomography (CT) imaging is clinically desired and has been under investigation in the last decade [1]. Reconstructing the low-dose (either low mA or shorter acquisition period) CT images is essentially a noise problem. Up to now, many noise-reduction strategies have been proposed to address this problem. One of the major strategies models the data noise property by a cost function in image space and then minimizes the cost function for image reconstruction by iterative numerical algorithms [2, 3]. An alternative strategy models the noise property by a cost function in sinogram space, seeks an optimal solution for the Radon transform, and then inverts the Radon transform for image reconstruction by the well established filtered backprojection (FBP) algorithm [4-7]. These two strategies of modeling the data statistics are mathematically equivalent with different implementations. Another major strategy applies a sophisticated linear or non-linear filter directly on either sinogram image, ramp-filtered sinogram image [8-12], or reconstructed image [13, 14] which considers the image characteristics rather than the data properties.

Based on repeated phantom experiments, the low-mA (or low dose) CT calibrated projection data were found to follow approximately a Gaussian distribution with a nonlinear dependence between the sample mean and sample variance, *i.e.*, the noise is signal-dependent [15]. The sample mean and variance dependence can be described by an analytical

[*] Email: jingwang@mil.sunysb.edu; Phone: (631) 444 7921; Fax: (631) 444 6450.

Medical Imaging 2006: Physics of Medical Imaging, edited by Michael J. Flynn, Jiang Hsieh,
Proceedings of SPIE Vol. 6142, 614247, (2006) · 1605-7422/06/$15 · doi: 10.1117/12.653903

formula [5, 15]. Filtering this signal-dependent noise by spatially-invariant low-pass filters, such as the Hanning, Butterworth, *etc*, has shown ineffectiveness, as expected [5, 15]. Sophisticated edge-preserving noise filtering by the use of the local statistics can improve the results [10, 13], but was observed incapable to handle the noise-induced streak artifacts, which mimics the edges [16]. Modeling the signal-dependent noise properties with penalized weighted least-squares (PWLS) cost function in either sinogram space [4, 5] or image domain [17] has shown very promising results. Minimizing the PWLS cost function can be performed iteratively in the image domain for direct image reconstruction [17] or either analytically [4] or iteratively in the sinogram space [5], followed by FBP for image reconstruction. This work compares the performances of these three different PWLS minimization approaches [4, 5, 17] for low-dose CT imaging using both phantom experiments and computer simulations.

2. METHODS

2.1 Noise model

Analysis of repeated measurements from the same phantom indicates that the calibrated projection data of low-mA CT follow approximately a Gaussian distribution with an associated relationship between the data sample mean and variance which can be described by the following analytical formula [5, 15],

$$\sigma_i^2 = f_i \times \exp(\mu_i / \eta) \tag{1}$$

where μ_i is the mean and σ_i^2 is the variance of the calibrated repeated-measurements of projections at detector channel or bin i, η is a scaling parameter and f_i is a parameter adaptive to different detector bins. Further analysis on the repeated measurements reveals that there is no statistical correlation for the data noise among the detector bins at all the projection views. The statistical correlation coefficient of data noise (*i.e.*, the datum minus its mean) between the *i*th and the *j*th detector bins is defined by [18]:

$$\rho_{i,j} = \frac{\mathrm{cov}(y_i, y_j)}{\sigma_i \sigma_j} \tag{2}$$

where $\mathrm{cov}(y_i, y_j)$ is the covariance of repeated measurements of projection data between the *i*th and the *j*th detector bins, σ_i and σ_j are the standard deviations for the projection data of the *i*th and *j*th detector bins respectively. The statistical correlation coefficients matrix of the data noise among the detector bins is shown in Fig. 1. Only the diagonal elements have a value of 1, other elements have values close to zero, indicating that there is no correlation of noise among different detector bins. Similar results were obtained for the noise among projection views, indicating that there is no correlation of noise among different projection views.

To study the data signal correlation among difference views, we calculated the correlation coefficient of the data signal among difference views. The correlation coefficient of data signal between the *k*th view and the *l*th view is defined by [18]:

$$r_{kl} = \frac{B\sum_{i=1}^{B}(y_{i,k} - \bar{y}_k)(y_{i,l} - \bar{y}_l)}{\sqrt{\left[B\sum_{i=1}^{B} y_{i,k}^2 - (\sum_{i=1}^{B} y_{i,k})^2\right]\left[B\sum_{i=1}^{B} y_{i,l}^2 - (\sum_{i=1}^{B} y_{i,l})^2\right]}} \tag{3}$$

where

$$\bar{y}_k = \frac{1}{B}\sum_{i=1}^{B} y_{i,k}\ , \qquad \bar{y}_l = \frac{1}{B}\sum_{i=1}^{B} y_{i,l} \tag{4}$$

and B is the number of bins for each view and indices k and l run over the nearby views. Figure 2 shows the correlation coefficients matrix of the data signal among the projection views. The values of the diagonal band or diagonal element groups are close to 1, which indicates a strong signal correlation among the neighboring views.

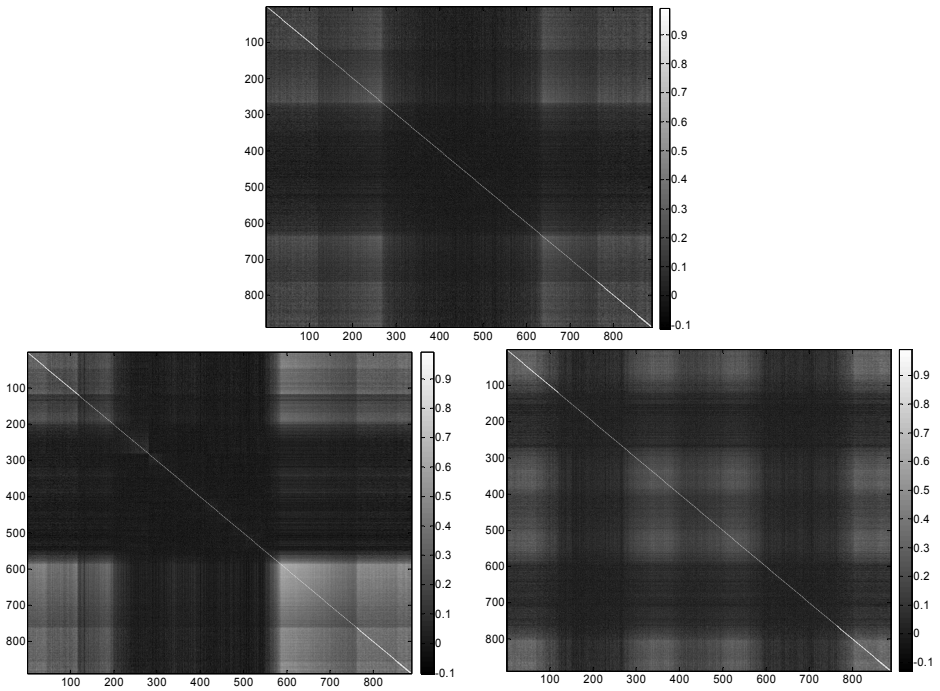

Fig. 1: Correlation coefficients matrix of data noise among detector bins from repeated measurements of the GE QA nearly-symmetric cylinder phantom (top), the Shoulder asymmetric phantom at 3 o'clock position (bottom left), and the Shoulder asymmetric phantom at 12 o'clock position (bottom right).

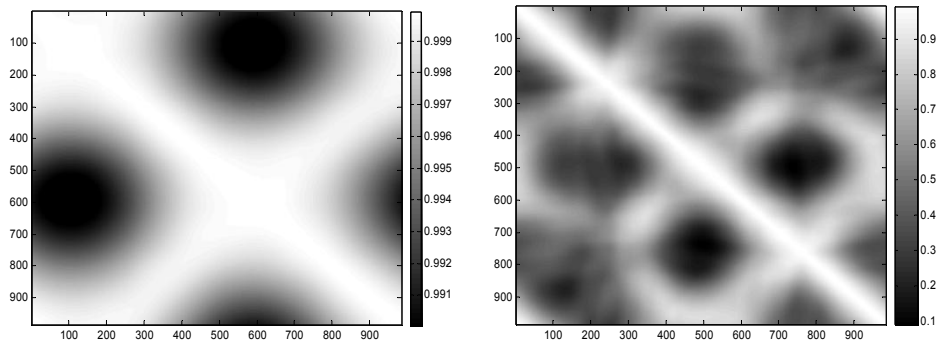

Fig. 2: Correlation coefficients matrix of data signal among projection views from the sinogram of the GE QA nearly-symmetric cylinder phantom (left) and the asymmetric Shoulder phantom (right).

These observations on the first- and second-order noise moments lead to the well-known weighted least-squares (WLS) cost function

$$\Phi(\mu) = (\hat{y} - \mu)'\Sigma^{-1}(\hat{y} - \mu) \tag{5}$$

in the sinogram space and

$$\Phi(\lambda) = (\hat{y} - P\lambda)'\Sigma^{-1}(\hat{y} - P\lambda) \tag{6}$$

in the image domain to address the noise problem of low-dose CT imaging, where $\mu = P\lambda$ is the vector of projection data $\{\mu_i\}$ to be estimated in the sinogram space and λ is the vector of attenuation coefficients $\{\lambda_j\}$ to be reconstructed in the image domain. P is the system or projection matrix and vector $\hat{y}$ is the system-calibrated projection measurements. Σ is a diagonal matrix with the ith element of σ_i^2, *i.e.*, an estimate of the variance of

repeatedly measured y_i at detector bin *i* which is determined from equation (1). Symbol $'$ denotes the transpose operator.

2.2. Penalized Weighted Least-Squares Image Reconstruction in Image Domain

Minimizing the WLS cost function of equation (5) or (6) usually leads to unacceptable results, similar to the maximum likelihood (ML) approach [19]. A penalty is desired for a penalized ML (pML) or MAP (maximum *a posteriori* probability) solution [6]. A PWLS image reconstruction method was first proposed by Sauer and Bouman [17] for iterative reconstruction of X-ray CT images and then was extended by Fessler [20] to PET (positron emission tomography) reconstruction. Sukovic and Clinthorne [21] later applied the iterative PWLS method to dual-energy X-ray CT reconstruction. Mathematically, the PWLS cost function can be written, in the image domain, as [17, 20]:

$$\Phi(\lambda) = (\hat{y} - P\lambda)'\Sigma^{-1}(\hat{y} - P\lambda) + \beta R(\lambda) . \tag{7}$$

The first term in equation (7) is the WLS measure of equation (5). The second term is a penalty, where β is a smoothing parameter which controls the degree of agreement between the estimated and the measured data. In this work, the quadratic penalty [20] is adapted,

$$R(\lambda) = \lambda' R\lambda = \sum_j \sum_{m \in N_j} w_{jm}(\lambda_j - \lambda_m)^2 \tag{8}$$

where index *j* runs over all image elements in the image domain, N_j represents the set of eight neighbors of the *j*th image pixel in two dimensions (in three dimensions, N_j indicates the 26 neighbors of the *j*th image voxel). In the two-dimensional (2D) cases, the weight w_{jm} is equal to 1 for the vertical and horizontal first-order neighbors and $1/\sqrt{2}$ for the diagonal (second-order) neighbors. The task for image reconstruction is to estimate the attenuation coefficient distribution map λ from projections $\hat{y}$:

$$\lambda = \arg\min_{\lambda \geq 0} \Phi(\lambda) . \tag{9}$$

In this work, we adapt the iterative Gauss-Seidel (GS) update algorithm [17] or more generally the iterative successive over-relaxation (SOR) algorithm [20] to calculate the solution of equation (8). Implementation of the iterative PWLS+SOR procedure [20] for our low-dose CT imaging can be summarized as follows:

Initialization:

$$\hat{\lambda} = FBP\{\hat{y}\}$$
$$\mu = P\hat{\lambda}$$
$$\Sigma = diag\{\sigma_i^2(\mu_i)\} = diag\{f_i \exp(\mu_i / \eta)\}$$
$$\hat{r} = \hat{y} - \mu$$
$$s_j = p_j'\Sigma^{-1}p_j, \forall j$$
$$\alpha_j = s_j + \beta \sum_{m \in N_j} w_{jm}.$$

For each iteration: (10)

begin

For each pixel *j*:

begin

$$\hat{\lambda}_j^{old} := \hat{\lambda}_j$$
$$\hat{\lambda}_j^{new} := \frac{p_j'\Sigma^{-1}\hat{r} + s_j\hat{\lambda}_j^{old} + \beta\sum_{m \in N_j} w_{jm}\hat{\lambda}_m}{\alpha_j}$$
$$\hat{\lambda}_j := \max\{0, (1-\omega)\hat{\lambda}_j^{old} + \omega\hat{\lambda}_j^{new}\}$$

$$\hat{r} := \hat{r} + p_j(\hat{\lambda}_j^{old} - \hat{\lambda}_j)$$

end

$$\hat{\sigma}_i^2 := f_i \exp(\sum_j p_{ij}\lambda_j / \eta)$$

end

where $0 \le \omega \le 1$. If ω is set to 1, the above procedure is equivalent to the GS update algorithm as described in [17], *i.e.*, a GS-PWLS approach. In our implementation, the variance σ_i^2 is updated during each iteration according to equation (1) for a more accurate estimation of the variance of projection. The iterative process is terminated if certain convergence criteria are satisfied [17, 20].

2.3. Penalized Weighted Least-Squares Sinogram Noise Reduction

Another statistical image reconstruction strategy is to find an optimal estimation of the line integral or the Radon transform from the noisy sinogram, and then reconstruct the CT image by FBP method, which is theoretically derived for the inversion of the Radon transform. This statistical approach in sinogram space is more computationally efficient than the statistical reconstruction through the image domain. Li *et al*. [5] modeled the noise properties by a PWLS cost function in the sinogram space:

$$\Phi(\mu) = (\hat{y} - \mu)'\Sigma^{-1}(\hat{y} - \mu) + \ln(2\pi\,\Sigma) + \beta\, R(\mu) \tag{11}$$

where μ is the vector of ideal projection data, the first term is the WLS measure of equation (5), the second term $\ln(2\pi\,\Sigma)$ is the normalization for a Gaussian probability distribution, and $R(\mu)$ is the penalty term, similar to equation (8),

$$R(\mu) = \mu' R\mu = \sum_i \sum_{m \in N_i} w_{im}(\mu_i - \mu_m)^2 \tag{12}$$

where N_i indicates the set of four nearest (or first-order) neighbors of the ith pixel in the sinogram (or a 2D projection image from a 2D object. For a 3D object, its 3D sinogram is a stack of the 2D sinograms). In the following, we will limit our presentation in 2D cases for simplicity. The weight w_{im} is equal to 1 for the two horizontal neighbors (along the bin direction) and 0.25 for the two vertical neighbors (along the angular direction) [5]. By this design, the sinogram is smoothed less in the angular direction than in the radial direction. The task for sinogram noise reduction is to estimate μ from $\hat{y}$:

$$\mu = \arg\min_{\mu \ge 0} \Phi(\mu). \tag{13}$$

There are many numerical algorithms which can be employed to calculate the solution of equation (13). In this work, we adapt the iterated conditional mode (ICM) algorithm as used in [5]. The iterative ICM formula for the solution of equation (10) is given by:

$$[\mu_i^{(n)} - (y_i + \eta)]^2 - \eta^2 = [1 + 2\beta\,\eta(2.5\mu_i^{(n)} + \sum_{m \in N_i} w_{im}\mu_m^{(n-1)})](\sigma_i^2)^{(n-1)} \tag{14}$$

where index *n* represents the iterative number.

Implementation of the ICM-PWLS sinogram approach for the solution of nonlinear equation (14) can be summarized as follows:

(1) Initialize an estimate of the noise-free sinogram with the observed data $\hat{y}$, *e.g.*, by a local non-linear filter [16].
(2) Update the previous estimate in a pixel-by-pixel manner based on equation (11) and the variance σ_i^2 of equation (1).
(3) Repeat step 2 until it converges to certain criteria, which is usually determined empirically.

Usually 10 iterations are sufficient for a relatively stable solution, which does not change visually for further iterations as reported in [5]. The convergence after 10 iterations is not proved theoretically, however, many numerical tests, including ours in this work, have not shown any divergence after 10 iterations for all the experimental data.

2.4. Penalized Weighted Least-Squares Sinogram Noise Reduction in KL Domain

In the GS-PWLS approach, the penalty is implemented in the image domain for direct image reconstruction. The ICM-PWLS approach implements the penalty in the sinogram space for ideal sinogram estimate, followed by FBP for image reconstruction. The penalty plays a very important role and implicitly assumes some kind of data signal correlation in either image domain or sinogram space. By analyzing the phantom sinograms, we observed a strong data signal correlation among the nearby views, see Fig. 2. This correlation can be effectively utilized by the Karhunen-Loève (KL) transform for two purposes. One is to de-composite the correlated signals for adaptive noise treatment and the other is to reduce the above 2D sinogram noise-filtering operation into a series of 1D procedures, each on a KL principal component, respectively.

In this study, the KL transform is first applied to account for the correlative information among the neighboring views of tomographic projections as reported in [4]. For each (or *v*th) view of the projection data, its neighboring views (*i.e.*, *n* views before the *v*th view and *n* views after the *v*th view, $n = 1, 2, 3, \ldots$) are selected to perform the KL transform. For the nearby $2n+1$ views, the element kl of the covariance matrix K_t of the projection data can be calculated by [4, 22]:

$$[K_t]_{kl} = \frac{1}{B-1}\sum_{i=1}^{B}(y_{i,k} - \bar{y}_k)(y_{i,l} - \bar{y}_l) \tag{15}$$

where $\bar{y}_k$ and $\bar{y}_l$ were defined by equation (4).

From the covariance matrix K_t, the KL transform matrix A is calculated based on

$$K_t A' = A' D . \tag{16}$$

In equation (16), $D=diag\{d_l\}_{l=1}^{2n+1}$, where d_l is the lth eigenvalue of K_t. The dimension of K_t is usually very small so that the eigenvector in equation (16) can be very efficiently computed. In this work, n is chosen to be 1 so that the dimension of matrix K_t is 3×3, (*i.e.*, $k, l = 1, 2, 3$). The KL transform is then defined as:

$$\tilde{y} = A\, y \tag{17}$$

where $\tilde{y}$ denotes the KL transformed y. After the KL transform, the covariance $\tilde{K}_t$ of $\tilde{y}$ will be:

$$\tilde{K}_t = AK_tA' = diag\{d_l\}_{l=1}^{2n+1} \tag{18}$$

where all of the symbols have been defined previously. Equation (18) implies that the covariance matrix of the KL transformed data signal is diagonal, *i.e.*, the covariance of the signal between different views after the KL transform will be zero. Therefore, the data signals of different KL components are no longer correlated so that PWLS objective function can be constructed for each KL principal component separately. In the KL domain, the PWLS cost function becomes [4, 22], from the WLS of equation (4),

$$\Phi_l(\tilde{\mu}_l) = (\tilde{y}_l - \tilde{\mu}_l)'\tilde{\Sigma}_l^{-1}(\tilde{y}_l - \tilde{\mu}_l) + (\beta / d_l)\tilde{R}(\tilde{\mu}_l) \tag{19}$$

where $\tilde{y}_l$ and $\tilde{\mu}_l$ are the lth KL components of $\hat{y}$ and μ respectively, $\tilde{R}(\tilde{\mu}_l)$ is the penalty term which will be defined later and $\tilde{\Sigma}_l$ is the diagonal variance matrix of $\tilde{y}_l$. The inverse variance matrix $\tilde{\Sigma}_l^{-1}$ can be estimated by [4, 22]:

$$\tilde{\Sigma}_l^{-1} = diag\{\varphi_l' Q_i^{-1}\varphi_l\}_{i=1}^{B} \tag{20}$$

where $Q_i = diag\{\sigma_{i,k}^2\}_{k=1}^{2n+1}$ is the variance matrix of the projection at bin i and φ_l is the lth KL basis vector. For a given pixel in the sinogram, a 3×3 window has been used to estimate the sample mean, then the estimated sample mean was employed by equation (1) to calculate the sample variance $\sigma_{i,k}^2$ for the weight $\tilde{\Sigma}_l^{-1}$ with excellent results [4]. Equation (19) reveals that the regularization parameter should be chosen as (β / d_l) to perform the PWLS minimization adaptively on each KL component. This choice is favorable because the regularization parameter varies adaptively according to the SNR (signal-to-noise ratio) of that component. A smaller KL eigenvalue is usually associated with a component having a lower SNR and, therefore, a larger regularization value should be used to penalize

this noisier data component [4, 22]. By the KL transform along the angular direction, the above 2D sinogram noise reduction now becomes a 1D process on each principal component, where the 1D process is performed along the bin direction [4]. For the 1D process along the bin direction, the quadratic penalty in equation (20) can be defined as:

$$\tilde{R}(\tilde{\mu}_l) = \tilde{\mu}_l^{'} \tilde{R} \tilde{\mu}_l = \sum_i \sum_{m \in N_i} w_{im} (\tilde{\mu}_{i,l} - \tilde{\mu}_{m,l})^2 \tag{21}$$

where N_i indicates the two nearest neighbors of the ith pixel in the KL domain along the 1D bin direction (*i.e.*, $m = i \pm 1$) and the weight w_{im} is equal to 1 for the two neighbors. This penalty considers the data signal correlation among nearby bins and plays the same role as the penalty terms in the above GS-PWLS and ICM-PWLS approaches.

Taking the derivative of equation (19) with respect to $\tilde{\mu}_{i,l}$ in the KL domain, to minimize equation (19) becomes to solve $\tilde{\mu}_l$ from the following tri-diagonal system of equations of

$$(\tilde{\Sigma}_l^{-1} + \frac{\beta}{d_l} \hat{R}) \tilde{\mu}_l = \tilde{\Sigma}_l^{-1} \tilde{y}_l \tag{22}$$

where $\hat{R}$ is a tri-diagonal matrix of

$$\hat{R} = \begin{pmatrix} 1 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & 0 & -1 & 2 & -1 \\ 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix} \tag{23}$$

with dimension of $B \times B$, where B was defined before as the number of detector.

For the tri-diagonal equation sets of equation (22), the solution of $\tilde{\mu}_l$ can be computed analytically and efficiently through the procedures of LU decomposition [23]. After the PWLS noise treatment for each component in the KL domain, the processed data are inverse KL transformed, followed by FBP reconstruction for the low-dose CT images.

Implementation of the KL-PWLS sinogram noise-reduction approach for low-dose CT can be summarized as follows:

(1) For a chosen v-th view from the projection data, the (v-1)-th and the (v+1)-th views of the projections are selected for the KL transform.
(2) Compute the spatial covariance matrix K_t from the selected neighboring views of the projections according to equation (15).
(3) Calculate the KL transform matrix A according to equation (16).
(4) Apply the KL transform on the selected neighboring views.
(5) Perform PWLS minimization on each KL component via equation (22).
(6) Apply inverse KL transform on the processed KL components for the estimate of the ideal sinogram.

The estimated ideal sinogram by either iterative ICM-PWLS or analytical KL-PWLS minimization is then reconstructed by FBP algorithm, which is well established for the inversion of the Radon transform by both speed and consistence.

3. RESULTS

3.1. Experimental Results

To compare the above three different strategies for noise reduction for low-dose CT imaging, two projection datasets from two different phantoms were acquired by a GE Hi-Speed multi-slice CT scanner with 10 mA and 20 mA

respectively. The number of bins per view is 888 with 984 views evenly spanned on a circular orbit of 360°. The detector arrays are on an arc concentric to the X-ray source with a distance of 949.075 mm. The distance from the rotation center to the X-ray source is 541 mm. The detector cell spacing is 1.0239 mm. The reconstructed image is of 512×512 array size.

For comparison purpose, reconstructed images by a conventional FBP method (*i.e.*, using a spatially-invariant low-pass Hanning filter with cutoff at 80% Nyquist frequency -- selected by visual inspection for a good result) are shown in Fig. 3(a) and Fig. 4(a). From the 10 mA Shoulder phantom study of Fig. 3, severe noise-induced streak artifacts can be observed in the conventional FBP reconstructed image. Images reconstructed by the GS-PWLS approach after 40 iterations are shown in Fig. 3(b) and Fig. 4(b). (The choice of 40 iterations has also been shown to produce a relatively stable solution, which did not change visually for further iterations [17, 20]). Standard FBP (with Ramp at the Nyquist frequency cutoff) reconstructed images from the ICM-PWLS smoothed sinograms (after 10 iterations) are shown in Fig. 3(c) and Fig. 4(c). (The choice of 10 iterations has also been shown to produce a relatively stable solution, which did not change visually for further iterations [5]). Standard FBP reconstructed images from the analytical KL-PWLS filtered sinograms are shown in Fig. 3(d) and Fig. 4(d). All the three PWLS minimization reconstructions demonstrated excellent noise reduction with satisfactory resolution preservation (see Fig. 4). The noise-induced streak artifacts were removed by the PWLS minimization as shown in Fig. 3.

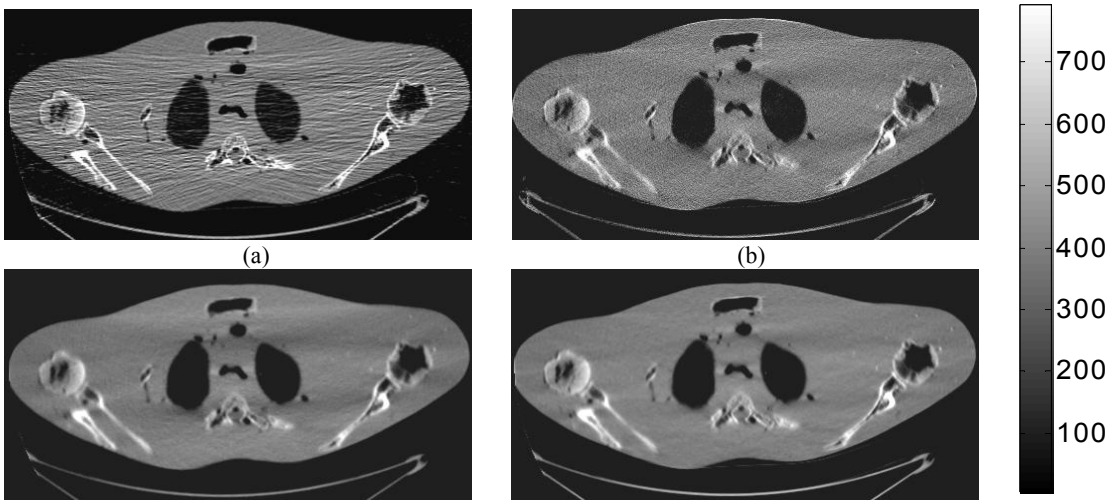

Fig. 3: Shoulder phantom study with 10 mA protocol. (a): FBP reconstructed image using the Hanning filter with cutoff at 80% Nyquist frequency. (b): Iterative GS-PWLS reconstructed image with $\beta=1\times10^{-4}$. (c): FBP reconstructed image from the iterative ICM-PWLS smoothed sinogram with $\beta=1\times10^{-5}$. (d): FBP reconstructed image from the analytical KL-PWLS filtered sinogram with $\beta=500$.

To quantitatively compare these three noise suppression approaches, we performed computer simulation studies in terms of noise-resolution trade-off, contrast to noise ratio (CNR) and receiver operating characteristic (ROC) studies.

3.2. Resolution-Noise Tradeoffs

The noise-resolution tradeoffs of the presented three PWLS approaches were computed by computer simulation using an ellipse digital phantom of Fig. 5(left) similar to that one used in [6, 9]. The detector array and X-ray source configuration is exactly the same as the GE scanner as described before. The ellipse phantom was sampled by the detector system on a circular orbit. Noise-free sinogram was computed based on the known densities and intersection lengths of the projection rays with the geometric shapes of the objects in the phantom. Noisy sinograms were generated by adding signal-dependent Gaussian noise according to equation (1). Image reconstruction was performed on a 512×512 array size in the image domain.

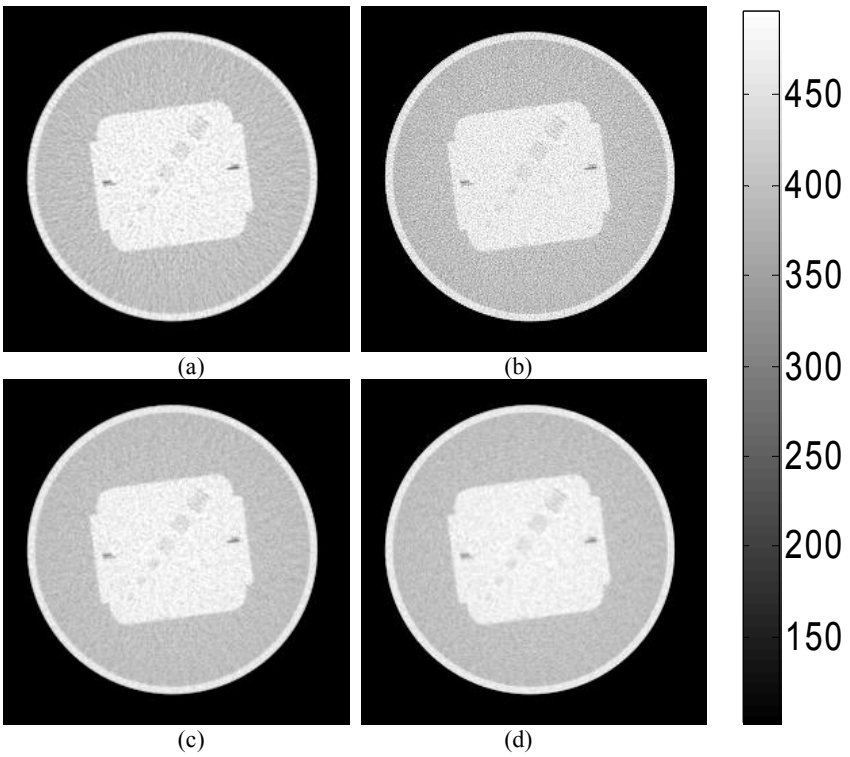

Fig. 4: Cylinder phantom study with 20 mA protocol. (a): FBP reconstructed image using the Hanning filter with cutoff at 80% Nyquist frequency. (b): Iterative GS-PWLS reconstructed image with β=1×10^{-4}. (c): FBP reconstructed image from the iterative ICM-PWLS smoothed sinogram with β=1×10^{-5}. (d): FBP reconstructed image from the analytical KL-PWLS filtered sinogram with β=500.

The reconstructed image resolution was analyzed by the edge spread function (ESF) along the central vertical profile on the left disk in the reconstructed ellipse-phantom image, which is similar to that used in [6]. The ESF is a measure of the broadening of a step edge. Assume that the broadening kernel is a Gaussian function with standard deviation σ_b, the ESF can be described by an error function (erf) parameterized by σ_b. By fitting the vertical profiles through the center of the left disk to an error function, we can obtain the parameter σ_b. The obtained parameter σ_b reflects the full-width at half-maximum (FWHM) of the broadening Gaussian function and, therefore, the reconstructed image resolution.

The reconstructed image noise was characterized by the standard deviation of a uniform region around the left disk in the reconstructed ellipse-phantom image. By varying the penalty parameter β for the three different PWLS approaches, we obtained the noise-resolution tradeoff curve for each approach. The noise-resolution tradeoff curves are shown in Fig. 5(right). The noise-resolution tradeoff curves of the ICM-PWLS and KL-PWLS sinogram smoothing approaches have the same trend. The KL-PWLS strategy shows a slightly better performance than the ICM-PWLS approach in all the resolution range. This may be due to the consideration of the signal correlation by the KL transform. The noise-resolution tradeoff curve of the GS-PWLS image reconstruction is slightly different from the other two approaches. At higher resolution (<1.5), the GS-PWLS image reconstruction is better than the PWLS sinogram smoothing approaches. At lower resolution (>1.5), the PWLS sinogram restoration is better than the GS-PWLS image reconstruction. This difference may be due to their implementation of penalty in different spaces. The penalty in GS-PWLS is implemented in image domain among the nearby image pixels. It is a local effect on the reconstructed image. However, the penalty of the PWLS minimization in the sinogram space may have a global effect on the FBP reconstructed image. Further investigation for different penalties in the sinogram space is needed.

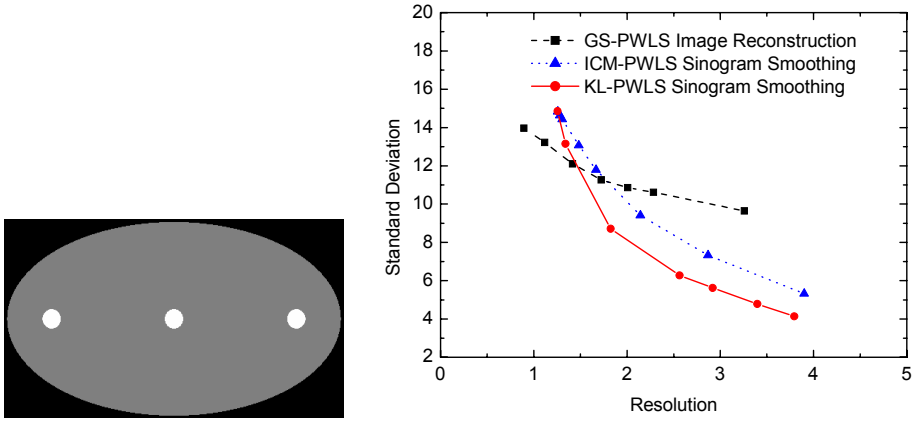

Fig. 5: Left is the phantom used for the noise-resolution tradeoff studies. Right shows the noise-resolution tradeoff curves for the presented three PWLS approaches. The resolution is measured by FWHM in pixel units.

3.4. Receiver Operating Characteristic (ROC) Studies by Channelized Hotelling Observer

One generally accepted method for evaluation of the performance of a medical imaging system or procedure is to evaluate the ability of an observer to detect an abnormality. By the observer study, a variety of pairs of true positive fraction (TPF) and false positive fraction (FPF) are generated as an observer changes the confidence threshold, resulting in a ROC curves [24]. Each ROC curve describes the inherent discrimination capacity of an imaging system or procedure. A common merit for comparing the ROC curves is the area under the curve (AUC) or A_Z. The reconstruction scheme, which generates a larger AUC, usually reflects a better detectability on abnormality. In this study, the channelized Hotelling observer (CHO) [25] was employed. This may eliminate the intra human observer variation.

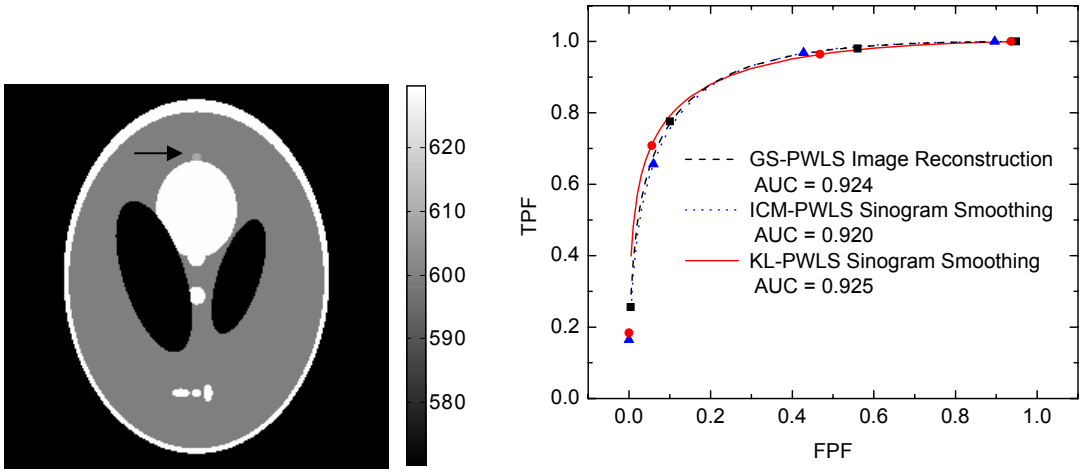

Fig. 6: Left picture is the modified Shepp-Logan head phantom used for ROC study. The right picture shows the results of the ROC evaluation and the fitted ROC curves by the binormal model.

Figure 6(left) shows a modified Shep-Logan head phantom (256×256 mm^2) used for the ROC study, where a low-contrast small lesion grows from the big ellipse as indicated by the arrow in the picture. The density of the lesion is about 1.5% of the background density and the radius of the lesion is 3 mm. The noise-free sinograms of this modified phantom with the lesion and the original phantom without the lesion were simulated, respectively. A total of 250 noisy sinograms were generated from each noise-free sinogram by adding Gaussian noise according to equation (1). Then these noisy sinograms were reconstructed by the three different PWLS reconstruction schemes. The CHO [25] gives a rating value for each image. Here we used four octave-wide rotationally symmetric frequency channels proposed by

Myers and Barrett [25]. This channel model was found to give good predictions of a human observer study in myocardial defect detection and in comparing fan- and parallel-beam collimators [26, 27]. In this CHO study, each reconstructed image generated a four-element feature vector, and the CHO was trained separately for each PWLS reconstruction approach, using subsets of these feature vectors given their class category (*i.e.*, with or without lesion). Then the CHO was applied to a different, independent ensemble of feature vectors to produce an ensemble of scalar rating values, which were subsequently analyzed using the CLABROC code [24], resulting in four pairs of FPF and TPF for each reconstruction method. These four pairs were then fitted by a ROC curve using the bi-normal model. By varying the penalty parameter β, we generated a series of ROC curves for each reconstruction method. The ROC curves of optimized penalty parameters are plotted in Fig. 6(right). The area under the three curves is 0.924 for the GS-PWLS image reconstruction, 0.920 for the ICM-PWLS sinogram restoration approach, and 0.925 for the KL-PWLS sinogram restoration strategy. The one-tailed p-values between each pair of two approaches are all larger than 0.05, which indicates that the difference between these three PWLS approaches is statistically not significant.

4. DISCUSSION AND CONCLUSION

Based on the noise properties of calibrated CT sinogram, where the calibration is expected to be well-established for current high-quality CT images, we compared three different noise reduction strategies for low-dose CT application. The GS-PWLS strategy is based on the PWLS cost function in image domain and the CT image is reconstructed by minimizing the cost function iteratively with penalty on the neighboring image pixels. The ICM-PWLS and KL-PWLS sinogram noise-reduction approaches aim to estimate the ideal sinogram (or the Radon transform) by minimizing the PWLS cost function with local penalty in sinogram space and reconstruct the CT images by inverting the Radon transform via the standard FBP algorithm (which has been widely used because of its speed and consistence). The ICM-PWLS strategy utilizes the penalty term to consider the difference of neighborhood between angular direction and radial direction (*i.e.*, inside the sinogram) and performs the estimation of the ideal sinogram by the iterative ICM algorithm. The KL-PWLS strategy employs the KL transform to further consider the signal correlation among the nearby projections and performs the estimation of the ideal sinogram analytically. Both phantom experiments and computer simulations demonstrate a similar performance for all these three PWLS-based approaches. The KL-PWLS strategy shows some improvement over the ICM-PWLS, and this may be due to the consideration of additional signal correlation via the KL transform. The GS-PWLS strategy shows some difference on resolution-noise tradeoff from the ICM-PWLS and KL-PWLS strategies, and this may be due to their different implementations of the penalty terms in difference spaces.

These three different PWLS strategies were coded by C++ language and executed in a PC Pentium IV with 2.4 GHz CPU speed and 2GB RAM memory. The GS-PWLS image reconstruction took approximately 10 minutes per iteration. The ICM-PWLS noise reduction in sinogram space took approximately 3 seconds per iteration. The analytical KL-PWLS noise treatment in the sinogram space took approximately 5 seconds. The gain in computing speed by the KL-PWLS strategy can be seen.

Alternative to the presented PWLS approaches above, the pML approaches by modeling the raw data directly (before system calibration) with maximization in the image domain [3] and in the sinogram space (followed by FBP) [6] have also shown potential for low-dose CT applications. Comparing these pML approaches to the presented PWLS strategies would be an interesting research topic and will be investigated.

ACKNOWLEDGEMENT

The authors wish to express their gratitude to Dr. J. Hsieh for providing the experimental phantom data. This work was supported in part by the NIH National Cancer Institute under Grant # CA82402 and Grant # CA110186. Dr. Lu was supported by the National Nature Science Foundation of China under Grant 30470490.

REFERENCES

1. Linton O.W., Fred A., and Mettler F.A., "National Conference on Dose Reduction in CT, with an Emphasis on Pediatric Patients", *American Journal of Roentgenology* Vol. 181, pp. 321-329, 2003.

2. Lange K. and Carson R., "EM Reconstruction Algorithms for Emission and Transmission Tomography", *J. Computer Assisted Tomography*, Vol. 8, pp. 306-316, 1984.
3. Elbakri I.A. and Fessler J.A., "Efficient and Accurate Likelihood for Iterative Image Reconstruction in X-ray Computed Tomography", *Proc. SPIE Med. Imaging*, Vol. 5032, pp. 1839-1850, 2003.
4. Lu H., Li X., Hsiao I.T., and Liang Z., "Analytical Noise Treatment for Low-Dose CT Projection Data by Penalized Weighted Least-Square Smoothing in the K-L Domain", *Proc. SPIE Med. Imaging*, Vol. 4682, pp. 146-152, 2002.
5. Li T., Li X., Wang J., Wen J., Lu H., Hsieh J., and Liang Z., "Nonlinear Sinogram Smoothing for Low-Dose X-ray CT", *IEEE Trans. Nucl. Science*, Vol. 51, pp. 2505-2513, 2004.
6. La Rivière P.J. and Billmire D.M., "Reduction of Noise-Induced Streak Artifacts in X-ray Computed Tomography Through Spline-Based Penalized-Likelihood Sinogram Smoothing", *IEEE. Trans. Med. Imaging*, Vol. 24, pp. 105-111, 2005.
7. La Rivière P.J.: Penalized-Likelihood Sinogram Smoothing for Low-Dose CT, *Medical Physics*, Vol. 32, pp. 1676-1683, 2005.
8. Sauer K. and Liu B., "Non-Stationary Filtering of Transmission Tomograms in High Photon Counting Noise", *IEEE Trans. Med. Imaging*, Vol. 10, pp. 445-452, 1991.
9. Hsieh J., "Adaptive Streak Artifact Reduction in Computed Tomography Resulting from Excessive X-ray Photon Noise", *Medical Physics*, Vol. 25, pp. 2139-2147, 1998.
10. Demirkaya K., "Reduction of Noise and Image Artifacts in Computed Tomography by Nonlinear Filtration of the Projection Images", *Proc. SPIE Med. Imaging*, Vol. 4322, pp. 917-923, 2001.
11. Zhong J., Ning R., and Conover D., "Image Denoising Based on Multiscale Singularity Detection for Cone Beam CT Breast Imaging", *IEEE Trans. Med. Imaging*, Vol. 23, pp. 696-703, 2004.
12. Kachelrieβ M., Watzke O., and Kalender W.A., "Generalized Multi-Dimensional Adaptive Filtering for Conventional and Spiral Single-Slice, Multi-Slice, and Cone-Beam CT", *Medical Physics*, Vol. 28, pp. 475-490, 2001.
13. Rust G.F., Aurich V., and Reiser M., "Noise/Dose Reduction and Image Improvements in Screening Virtual Colonoscopy with Tube Currents of 20 mAs with Nonlinear Gaussian Filter Chains", *Proc. SPIE Med. Imaging*, Vol. 4683, pp. 186-197, 2002.
14. Lu H., Li X., Li L., Chen D., Xing Y., Wax M., Hsieh J., and Liang, Z., "Adaptive Noise Reduction Toward Low-Dose Computed Tomography", *Proc. SPIE Med. Imaging,* Vol.5030, pp. 759-766, 2003.
15. Lu H., Hsiao I., Li X., and Liang, Z., "Noise Properties of Low-Dose CT Projections and Noise Treatment by Scale Transformations", *Conf Record IEEE NSS-MIC*, in CD-ROM, 2001.
16. Wang J., Lu H., Li T., and Liang, Z., "Sinogram Noise Reduction for Low-Dose CT by Statistics-Based Nonlinear Filters", *Proc. SPIE Med. Imaging*, Vol. 5747, pp. 2058-2066, 2005.
17. Sauer K. and Bouman C., "A Local Update Strategy for Iterative Reconstruction form Projections", *IEEE Trans. Signal Processing*, Vol. 41, pp. 534-548, 1993.
18. Edwards, A.L.: *An Introduction to Linear Regression and Correlation.* San Francisco, CA: W. H. Freeman, pp. 33-46, 1976. (www.mathworld.wolfram.com/correlationCoefficient.html).
19. Snyder D., Miller M., Thomas L., and Politte D., "Noise and Edge Artifacts in Maximum-Likelihood Reconstruction for Emission Tomography", *IEEE Trans. Med. Imaging*, Vol. 6, pp. 228-238, 1987.
20. Fessler J.A., "Penalized Weighted Least-Squares Image Reconstruction for Positron Emission Tomography", *IEEE Trans. Med. Imaging*, Vol. 13, pp. 290-300, 1994.
21. Sukovic P. and Clinthorne N.H., "Penalized Weighted Least-Squares Image Reconstruction for Dual energy X-ray Computed Tomography", *IEEE Trans. Med. Imaging.*, Vol. 19, pp. 1075–1081, 2000.
22. Wernick M.N., Infusino E.J., and Milošević M., "Fast Spatio-Temporal Image Reconstruction for Dynamic PET", IEEE *Trans. Med. Imaging*, Vol. 18, pp. 185-195, 1999.
23. Press W.H., Teukolsky S.A., Vetterling W.T., and Flannnery B.P. *Numerical Recipes in C: The art of scientific computin*g Second Edition, Cambridge University Press, pp. 50-51, 1992
24. Metz C.E., "ROC Methodology in Radiological Imaging", *Invest. Radiology*, Vol. 21, pp. 720-733, 1986.
25. Myers K.J. and Barrett H.H., "Addition of a Channel Mechanism to the Ideal-Observer Model", *J. Opt. Soc. America*, Vol. 4, pp. 447-57, 1987.
26. Wollenweber S.D., Tsui B.M.W., Frey E.C., Lalush D.S., and LaCroix K.J., "Comparison of Human and Channelized Hotelling Observers in Myocardial Defect Detection in SPECT", *J. Nucl. Med.*, Vol. 39, p. 771A, 1998.
27. Wollenweber S.D., Lalush D.S., Tsui B.M.W., and Gullberg G.T. "Evaluation of Myocardial Defect Detection between Parallel-Hole and Fan Beam SPECT Using Hotelling Trace", *IEEE Trans. Nucl. Science*, Vol. 45, pp. 2205-2210, 1997.