Science & Technology Facilities Council

Preprint
RAL-P-2015-001

Direct edge-preserving regularization in tomographic image reconstruction

D Kazantsev, E Ovtchinnikov, WRB Lionheart,
PJ Withers, PD Lee

January 2015

Submitted for publication in the Proceedings of the Royal Society A

RAL Library
STFC Rutherford Appleton Laboratory
R61
Harwell Oxford
Didcot
OX11 0QX

Tel: +44(0)1235 445384
Fax: +44(0)1235 446403
email: libraryral@stfc.ac.uk

ISSN 1361- 4762

Direct edge-preserving regularization in tomographic image reconstruction

D. Kazantsev[1,2], E. Ovtchinnikov[3], W. R. B. Lionheart[4], P. J. Withers[1,2] and P. D. Lee[1,2]

January 22, 2015

[1] The Manchester X-ray Imaging Facility, School of Materials, The University of Manchester, Manchester, M13 9PL

[2] Research Complex at Harwell, Didcot, Oxfordshire, OX11 0FA, UK

[3] Numerical Analysis Group, STFC, Rutherford Appleton Laboratory, Oxfordshire, OX11 0QX, UK

[4] School of Mathematics, Alan Turing Building, The University of Manchester, M13 9PL, UK

Abstract

In this paper ,we present a novel regularization technique for iterative tomographic reconstruction. The proposed penalty is built on the edge-preserving Laplacian and encourages piecewise-smooth solutions with sharp edges. Our penalty outperforms total variation regularization, resulting in higher signal-to-noise ratio and visually more appealing reconstructed images. Simple and effective implementation of the proposed regularization technique within the conjugate gradient algorithm makes it a competitive alternative to conventional regularization methods.

1 Introduction

Frequently in X-ray computed tomography (CT) the amount of collected projection data is lower than is required by the Nyquist sampling theorem [1]. In medical imaging, the data restrictions are applied to minimize ionizing radiation that can harm living tissue cells. In material science, the aim is to better resolve temporal resolution via higher frame rate acquisition. In such cases of limited data, iterative techniques can provide better reconstructions than analytical methods [2].

Dealing with ill-posed and ill-conditioned inverse problems, iterative methods require regularization to constrain the desirable space of solutions [3]. Due to edge-preserving properties, total variation (TV) regularization [4] has been extensively used in tomographic iterative reconstruction for the last three decades. However, TV penalty produces piecewise-constant (so-called "cartoon" effect) images even if the original object is smooth [5, 6].

In this paper, we propose a novel penalty that reduces "cartoon" effect and increases the signal-to-noise ratio (SNR) of reconstructed images. We use a preconditioned and regularized conjugate gradient method to minimize the least squares problem [7]. Both TV and the proposed penalty are compared quantitatively and visually.

2 Image reconstruction problem

The tomographic image reconstruction problem consists of determining the shape of an object based on its X-ray observations from several different angular positions. Incoming photons with different energies are registered by detectors and information about the path length a photon has travelled along the line can be decoded. Therefore, by solving an inverse problem where projection data is given, the level of absorption or attenuation coefficient of the object can be recovered. In mathematical terms, this problem can be formulated as the least squares problem:

$$\hat{u} = \arg\min_{u} \|Au - b\|_2^2, \tag{1}$$

where u is a function of spacial variables describing the observed object (e.g. density of the object's material), b is a function of the number of the detector bins and the observation angle describing the projection data (sinogram), and A is a linear operator mapping the "space of objects" to the "space of observations". The operator A is an integral operator, and zero is a condensation point of its singular values, which makes the problem (1) ill-posed.

The spatial and angular disretization replaces continuous functions with discrete (or grid) functions that can be represented by vectors and the operator A with a matrix. Depending on the numbers of spacial grid points, detectors and angles, the matrix A can be "fat" (the number of rows is less than the number of columns) or "tall" (the number of rows is greater than the number of columns). Irrespective of the shape, the singular values of A form a very tight cluster near zero owing to the same property of the integral operator from which A derives [8].

3 Regularization

The quadratic functional $\|Au - b\|_2^2$ can be minimized by a suitable iterative minimization algorithm, e.g. conjugate gradient (CG) algorithm [7], however, the convergence of iterations can be very slow because of the poor conditioning of

the Hessian A^*A. The slow convergence of iterations is closely related to another difficulty in dealing with this kind of problem, which is the ill-posedness of the problem. Generally, the solution $\hat{u}$ of the discretized problem (1) is not unique (if A is "fat"), and even if it is, in practical computation $\hat{u}$ is indistinguishable from any $\hat{u}+h$ if $\|Ah\|_2$ is below the round-off error level, which, for A that we are dealing with, may happen even if $\|h\|_2$ is substantial.

Both difficulties can be tackled by a regularization technique, whereby (1) is replaced with

$$\hat{u}_\alpha = \arg\min \psi_\alpha(u), \quad \psi_\alpha(u) = \|Au-b\|_2^2 + \alpha R(u), \tag{2}$$

where $R(u)$ is a suitable regularization functional, and α is a positive scalar parameter. A prime example of the regularization is that known as Tikhonov's, where $R(u) = \|u\|_2^2$, which makes the minimization problem (2) well-posed, the eigenvalues of the Hessian of $\psi_\alpha(u)$ being not less than α. Tikhonov regularization is quadratic, therefore high frequencies that are related to the object boundaries are penalized less, resulting in smooth recovery of $\hat{u}$. To preserve boundaries one needs to consider non-quadratic penalties, e.g. the use of total variation (TV) [4] can significantly improve oscillatory solutions. The differentiable (due to small ϵ constant) TV penalty is given as:

$$R_{TV}(u) = \||\nabla u|\|_1 = \sqrt{u_x^2 + u_y^2 + \epsilon^2}. \tag{3}$$

We will use the TV penalty (3) for comparison with the proposed regularization term.

It is intuitively obvious that the regularization parameter α must not be large. In order to get some further insight into the issue, let us estimate the difference between $\hat{u}_\alpha$ and $\hat{u}$ assuming for simplicity that all singular values of A are positive.

Let us assume $R(u) = \|Ru\|_2^2$, where R is a square non-degenerate matrix, and denote $M = A^*A$, and $N = R^*R$. Then $\hat{u}_\alpha$ (where α may be zero) satisfies

$$(M+\alpha N)\hat{u}_\alpha = A^*b, \tag{4}$$

which implies the following equation for the regularization error, $h_\alpha = \hat{u}_\alpha - \hat{u}$:

$$(M+\alpha N)h_\alpha = -\alpha N\hat{u}. \tag{5}$$

Multiplication by $M^{-1}Nh_\alpha$ yields

$$((N+\alpha NM^{-1}N)h_\alpha, h_\alpha) = -\alpha(N\hat{u}, M^{-1}Nh_\alpha). \tag{6}$$

Now, in the left-hand side of (6) we have

$$((N+\alpha NM^{-1}N)h_\alpha, h_\alpha) \geq (Nh_\alpha, h_\alpha) = (R^*Rh_\alpha, h_\alpha) = \|Rh_\alpha\|_2^2,$$

and in the right-hand side

$$(N\hat{u}, M^{-1}Nh_\alpha) = (NM^{-1}N\hat{u}, h_\alpha) = (RM^{-1}N\hat{u}, Rh_\alpha) \leq \|RM^{-1}N\hat{u}\|_2\|Rh_\alpha\|_2.$$

Thus, (6) implies

$$\|Rh_\alpha\|_2 \leq \alpha\|RM^{-1}N\hat{u}\|_2. \tag{7}$$

4 Edge-preserving piecewise-smooth regularization

At first glance, the regularization appears to be merely a compromise move that distorts the problem so that it becomes solvable. While this is certainly so in the case of Tikhonov's regularization, an alternative viewpoint can be offered, which is helpful in designing a proper regularization. Observe that the problem (1) can only be solved approximately, if only because of the inexactness of computer arithmetic. Observe further that it may have infinitely many approximate solutions that are indistinguishable in practical computation, as pointed out in the previous section, if one is only guided by the smallness of the data fidelity functional $\|Au-b\|_2^2$. The regularization can be viewed as some kind of additional criterion that helps to verify whether a particular computed solution is acceptable. This viewpoint is supported by the fact that in the case where $R(u) = \|Ru\|_2^2$, the regularized problem (2) is equivalent to the original problem (1) for these extended A and b:

$$A_\alpha = \begin{bmatrix} A \\ \alpha R \end{bmatrix}, \quad b_\alpha = \begin{bmatrix} b \\ 0 \end{bmatrix}. \tag{8}$$

In the problem (1) that we are dealing with here, A is a discretization of an integral operator, owing to which $\|Ah\|_2$ is small on oscillating grid functions h with wavelengths that are close to the grid step. Hence, if one directly applies e.g. CG algorithm to the minimization of the quadratic functional $\psi_0 = \|Au-b\|_2^2$, then after sufficiently many iterations one is likely to end up with an oscillating short-wavelength approximate solution $\hat{u}$. But most images that one encounters in practice do not feature such oscillations and can be actually represented by piecewise-smooth functions $\hat{u}$. This suggests that the value of $R(u)$ should be large on short-wavelength functions u. At the same time, $R(u)$ should remain small on the boundaries (walls) between objects constituting the image, where u is discontinuous or has large gradients.

The following regularizer is therefore suggested:

$$R_{EL}(u) = \left\| w_x \frac{\partial^2 u}{\partial x^2} \right\|_2^2 + \left\| w_y \frac{\partial^2 u}{\partial y^2} \right\|_2^2, \tag{9}$$

where the weights w_x and w_y are given by

$$w_x = \left(1 + \beta \left(\frac{1}{a_x} \frac{\partial u}{\partial x} \right)^2 \right)^{-1}, \quad w_y = \left(1 + \beta \left(\frac{1}{a_y} \frac{\partial u}{\partial y} \right)^2 \right)^{-1}, \tag{10}$$

β is a positive scalar parameter, and a_x and a_y are the average x- and y-derivatives of u. These averages are introduced purely for the sake of scale-invariance, and can be computed e.g. as $a_x = 2u_{max}/d_x$ and $a_y = 2u_{max}/d_y$, where u_{max} is the maximum of u and d_x and d_y are the sizes of the square containing the image.

By design, $R_{EL}(u)$ (where EL stands for edge-preserving Laplacian) is large on short-wavelength functions u with small oscillation amplitude (equivalent to Laplacian smoothing), and small at the edges between objects where the x- and y-derivatives of u are large.

5 Preconditioning

In order to simplify the computation of the gradient of $\psi_\alpha(u)$ and to deal with the minimization of a quadratic functional, we resort to an inner-outer iterative scheme consisting of restarted CG iterations, with w_x and w_y only updated on restarts.

Let us denote by L_x and L_y the matrices representing the discretized second partial x- and y-derivatives, and by W_x and W_y the diagonal matrices representing w_x and w_y. The Hessian H_{EL} of the functional $\psi_\alpha(u)$ is (ignoring the dependence of w_x and w_y on u – cf. inner-outer iterative scheme)

$$H_{EL} = A^*A + \alpha(L_x^* W_x^* W_x L_x + L_y^* W_y^* W_y L_y) = A^*A + \alpha R_{EL}^* R_{EL}, \qquad (11)$$

where

$$R_{EL} = \begin{bmatrix} W_x L_x \\ W_y L_y \end{bmatrix}. \qquad (12)$$

We compare the proposed EL penalty with TV regularization (3) with the following Hessian H_{TV}:

$$H_{TV} = A^*A + \alpha(D_x^* \Phi(u) D_x + D_y^* \Phi(u) D_y), \qquad (13)$$

where D_x and D_y are matrices representing the discretized first partial x- and y-derivatives, $\Phi(u) = diag(\phi'(u))$ is a diagonal matrix whose diagonal elements are $\phi'(u)$, $\phi(t) = 2\sqrt{t + T^2}$.

The largest eigenvalue of the matrix R_{EL} is of the order $\mathcal{O}(h^{-2})$, $h = \min(h_x, h_y)$. If α is not very small (considerably larger than h^4), then the large condition number of $R_{EL}^* R_{EL}$ can slow down the convergence of CG iterations for the minimization of $\psi_\alpha(u)$. To alleviate this problem, we introduce preconditioning that consists of the multiplication of the gradient of $\psi_\alpha(u)$ by the inverse of $H_{EL,\sigma} = \sigma^2 I + \alpha R_{EL}^* R_{EL}$ for EL and $H_{TV,\sigma} = \sigma^2 I + \alpha(D_x^* \Phi(u) D_x + D_y^* \Phi(u) D_y)$ for TV in the course of CG iterations, where σ is a scalar value of the order of the largest singular value of A. Since the matrix H_σ is very sparse, the application of its inverse to a vector can be efficiently performed (via the factorization of H_σ) by modern state-of-the-art sparse direct solvers [9].

6 Numerical Results

Here we present numerical experiments with reconstruction of 2D Shepp-Logan (SL) phantom of $N, M = 256$ pixels size. Reconstructions performed from noisy

projection data (3% normally distributed random noise is added to sinogram). Projections were sampled significantly lower than it is required by the Nyquist theorem, therefore we consider only $90, 45$ and 30 projections. According to the Nyquist theorem, the number of projection angles θ should be at least $\theta \approx 2.83N$ [1].

Four reconstruction algorithms were used to reconstruct sparse and noisy projection data, namely: Filtered back-projection (FBP), conjugate gradients for least squares (CGLS), CGLS with TV penalty (CGLS-TV) and CGLS with the proposed edge-preserving Laplacian penalty (CGLS-EL). All reconstruction parameters (the number of inner-outer iterations, α, β, T) for the compared algorithms were thoroughly optimized with respect to the minimum value of the root mean square error (RMSE) given as:

$$RMSE(u, \hat{u}) = \sqrt{\frac{1}{N}\sum_{i=0}^{N}(u_i - \hat{u}_i)^2}. \tag{14}$$

In Fig. 1, we present RMSE curves for each of the compared methods and reconstructions for the best RMSE values are presented in Fig. 2. RMSE curves are given for experiments with $90, 45$ and 30 projections. Notably, the proposed CGLS-EL method outperforms CGLS-TV in all experiments. Moreover, CGLS-EL is more stable than CGLS-TV in cases of very limited noisy data (see Fig. 1 for 45 and 30 projections). In Table 1, the best achieved values of RMSE for each method are presented.

Table 1: RMSE values for FBP, CGLS, CGLS-TV and CGLS-EL methods

	FBP	CGLS	CGLS-TV	CGLS-EL
90 projections	0.148	0.057	0.037	0.029
45 projections	0.183	0.0765	0.055	0.041
30 projections	0.220	0.092	0.075	0.063

From the obtained reconstructions (see Fig. 2) some visual quality assessment can be done. FBP reconstruction in all cases is very noisy and CGLS reconstruction looks significantly better. However, noise is still present in CGLS reconstructed images and the most importantly ring artifacts are visible (see reconstructions from 45 projections). Those artifacts are related to prematurely stopped CGLS iterations to avoid increase of the RMSE. On the early iterations of CGLS, higher frequencies are not fully recovered, and the blurring effect is more visible (see reconstructions from 30 projections). Using regularization it is possible to iterate CGLS longer which leads to significant improvement of resolution. Notably, CGLS-EL is better at removing ring effects and regularizing solution (less noise, sharper boundaries) than CGLS-TV (see reconstructions from 45 projections).

Additionally, visual characteristics of reconstructed images can be better recognized on the magnified parts of the reconstructed SL phantom (see Fig.

3). Because the second derivatives are used in the EL model instead of the first order derivatives as with the TV regularization, the uniform areas look much smoother but prominent edges remain unsmoothed.

In the next experiment, we aim to show how the proposed regularization deals with recovering ramp like surfaces in images. In Fig. 4, the reconstruction experiment with the ramp phantom is presented (the magnified part of the phantom is shown). The full phantom is 300×300 pixels and 360 projections were collected. In reconstructed images one can see that with iterative methods noise is substantially reduced (see Fig. 4 (right)). However, the CGLS-TV method brings patchy appearance to the reconstruction, and image reconstructed with CGLS-EL is smoother. The RMSE is also lower for CGLS-EL than for CGLS-TV.

7 Discussion

In this paper, we claim that the piecewise-smooth images are more favourable than the piecewise-constant, however this might not be the case for all reconstructed objects. Therefore, some prior knowledge about the investigated object is needed to choose an appropriate regularization term. In terms of the choice of reconstruction parameters, complexity of computer implementation and the speed of computation, our method is very similar to the TV penalty. Our future work will be to explore further the parameter space of our method and give some recommendations for automated choice of parameters.

8 Conclusion

In this paper, we presented a novel regularization technique that outperforms in terms of signal-to-noise ratio the conventional total variation regularization. From the preliminary experiments the proposed method is more suitable for the limited data problems in tomography than total variation regularization. Additionally, due to properties of the proposed penalty term we reduce the "cartoon" effect substantially and achieve more realistic images.

Acknowledgment

This work has been supported by the Engineering and Physical Sciences Research Council under grants EP/J010553/1, EP/J010456/1 and EP/I02249X/1.

References

[1] Kak AC, Slaney M, 2009 *Principles of computerized tomographic imaging.* IEEE Press, New York.

[2] Buzug TM. 2008 *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT*. Springer.

[3] Qi J, Leahy RM. 2006, Iterative reconstruction techniques in emission computed tomography. *Phys Med. Biol.* **51**, pp. 541–578.

[4] Rudin L, Osher S, Fatemi E. 1992, Nonlinear total variation based noise removal algorithms, *Physica D*, **60(1-4)**, pp. 259–268.

[5] Wang Y, Yang J, Yin W, Zhang Y. 2008, A new alternating minimization algorithm for total variation image reconstruction, *SIAM Journal on Imaging Sciences*, **1(3)**, pp. 248–272.

[6] Papafitsoros K, Schönlieb C-B. 2014, A combined first and second order variational approach for image reconstruction, *Journal of Mathematical Imaging and Vision*, **48 (2)**, pp. 308–338.

[7] Nocedal J, Wright S. 2006 *Numerical Optimization*. Springer.

[8] Hansen PC. 2010. *Discrete inverse problems: insight and algorithms* **7**. SIAM.

[9] Davis T. A. 2006 *Direct Methods for Sparse Linear Systems*. SIAM.

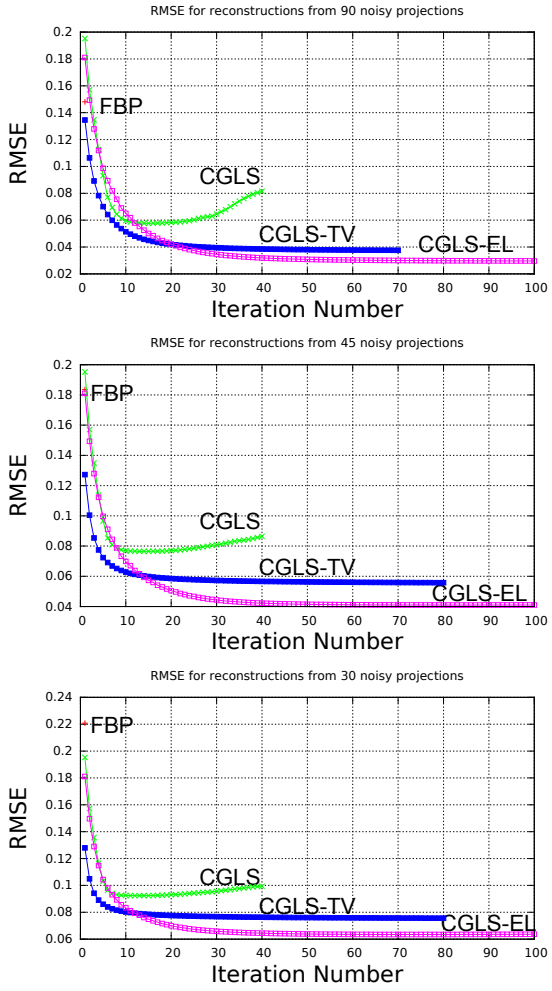

Figure 1: RMSE for FBP, CGLS, CGLS-TV and CGLS-EL algorithms for $90, 45$ and 30 noisy projections.

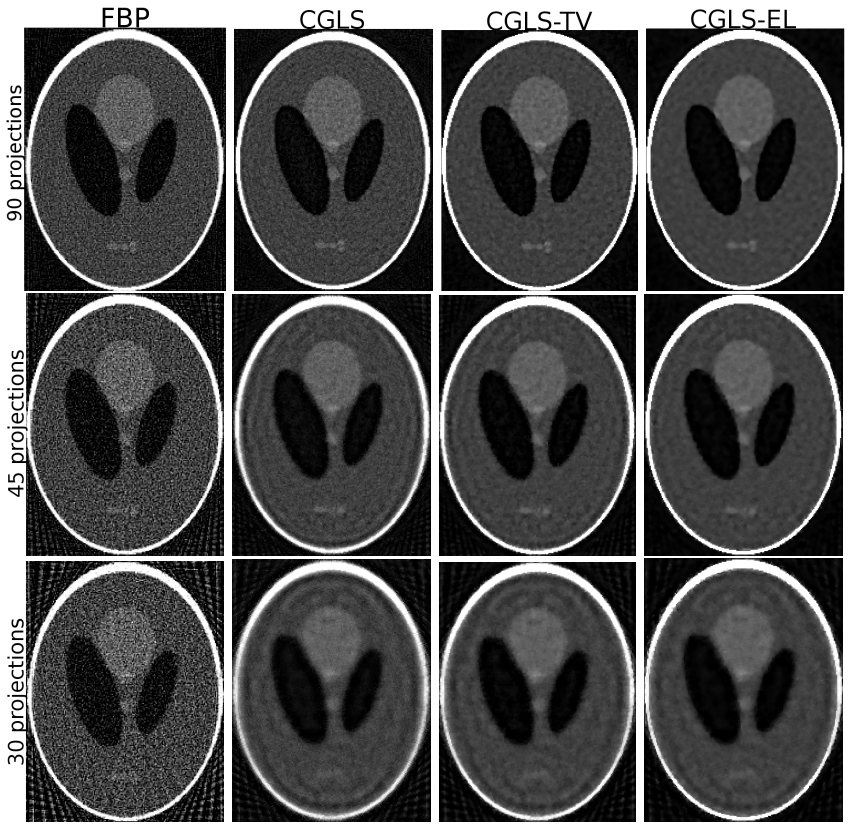

Figure 2: Reconstructions of SL phantom from 90, 45 and 30 projections using FBP, CGLS, CGLS-TV and CGLS-EL methods.

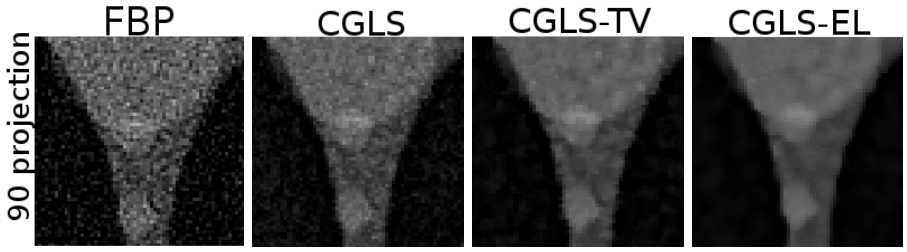

Figure 3: Magnified image of SL phantom reconstructed from 90 projections using FBP, CGLS, CGLS-TV and CGLS-EL methods. Note the better SNR of CGLS-EL reconstruction compared to the CGLS-TV method.

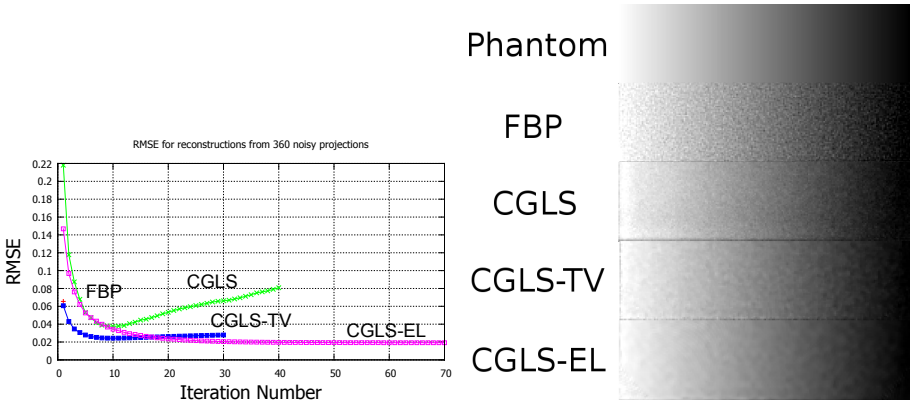

Figure 4: Reconstruction of the ramp phantom from 360 projections using FBP, CGLS, CGLS-TV and CGLS-EL methods; left: RMSE curves for all methods, right: magnified region of the phantom and reconstructed images.