



LAWRENCE  
LIVERMORE  
NATIONAL  
LABORATORY

# Ray-weighted constrained-conjugate-gradient tomographic reconstruction for security applications

J. S. Kallman, S. G. Azevedo

June 8, 2012

The Second International Conference on Image Formation in  
X-Ray Computed Tomography

## **Disclaimer**

---

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

# Ray-weighted constrained-conjugate-gradient tomographic reconstruction for security applications

Jeffrey S. Kallman and Stephen G. Azevedo

**Abstract**—Computed tomography (CT) reconstruction of objects in luggage is affected by surrounding clutter, which can contribute artifacts such as streaking and beam hardening. We have been investigating a constrained conjugate gradient (CCG) algorithm for CT reconstruction. We have found that the time required to perform each iteration of the CCG can be reduced by a factor of 10 or more by using an approximation to the error in the minimization line search. We have also found that ray weighting can alleviate streak artifacts.

**Index Terms**—Gradient methods, Least squares methods, Reconstruction algorithms, X-ray tomography

## I. INTRODUCTION

COMPUTED tomography (CT) reconstruction of many objects of interest is affected by surrounding clutter. The clutter may contribute scatter, beam hardening and photon starvation [1]. This case is particularly important in security applications such as checked baggage scanning, where quantitative data are used to decide whether a bag contains a threat. Clutter can significantly alter the reconstructed attenuation of a material or create false alarms. Iterative reconstruction techniques offer a number of possible ways to alleviate many of these environmental effects, including incorporation of prior knowledge, regularization and weighting of rays.

An example of environmental effects is shielding, which occurs when a highly-attenuating material blocks many of the projections of a less attenuating material. An example of this effect is a steel bar lying across a bottle of jelly in an airport bin (Fig 1). If reconstructed using filtered backprojection [2] (as in 1a), beam hardening and photon starvation make the bar appear hollow and some of the streaks in air have higher image intensity than the jelly.

Manuscript (LLNL-JRNL-560413) received June 12, 2012. This work performed under the auspices of the U. S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. This work was funded under DHS Interagency Agreement HSHQPM-10-X-00034/P00002.

J. S. Kallman is with the Lawrence Livermore National Laboratory, Livermore, CA 94551 USA: 925-423-2447; (e-mail: kallman1@llnl.gov).

S. G. Azevedo is with the Lawrence Livermore National Laboratory, Livermore, CA 94551 USA (e-mail: azevedo3@llnl.gov).

The streaks in the jelly change its mean attenuation. Many of the effects of shielding can be alleviated by using a ray-weighted iterative reconstruction (as in 1b).

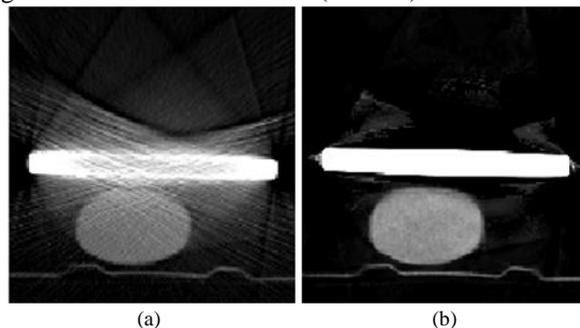


Fig. 1. (a) Filtered backprojection reconstruction of a steel bar lying over a bottle of jelly in an airport bin. Note that the streaks in empty air have points of higher intensity than the jelly, blur the space between the bar and the jelly, and that the bar appears to be hollow. (b) Ray-weighted iterative reconstruction. Note that the streaks have been diminished, there is a clear separation between the jelly and the bar, and the jelly is more uniform.

Scatter, beam hardening and photon starvation act to change the x-ray features of materials in luggage. If a particular set of features is used for explosive detection, these changes require enlarging the regions of feature space where it is necessary to raise an alarm and thus increase the false alarm rate (Fig 2). The goal of our iterative reconstruction work is to show that the effects of containers and concealment can be reduced.

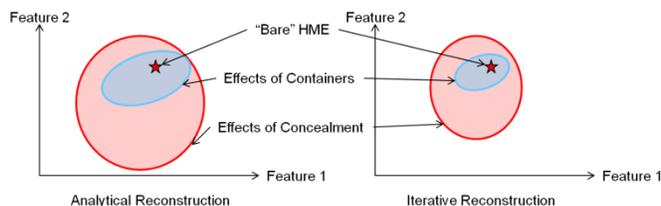


Fig. 2. The goal of our iterative reconstruction work is to show a reduction in the size of the feature space where an alarm must be raised. Features may include x-ray attenuation coefficients, effective atomic number, density, texture, etc. “Bare” is the object scanned as is without clutter.

One of the drawbacks of iterative reconstruction techniques is the amount of computer memory and time they require. We have been examining acceleration techniques for constrained conjugate gradient reconstruction and have found that using an approximation to the error can

accelerate the reconstruction by more than an order of magnitude.

The purpose of this work is to report on our results to date using a ray-weighted constrained-conjugate-gradient iterative reconstruction technique.

## II. RECONSTRUCTION ALGORITHM

### A. High Level Overview

We first generate a model to fill space. This model may be comprised of blobs, regular voxels, or pieces (if we have prior knowledge of the object to be scanned). We determine the interaction of the model with all the rays and prune those parts of the model that are un-attenuating. The reconstruction approach we use is to minimize the mean-square error function of the model for all attenuated rays. We use the adjoint method to find the gradient of the error function for every voxel [3]. This gradient is used in a constrained conjugate gradient algorithm [4] to determine the search direction for error minimization. An iteration of the method consists of:

- For each ray:
  - Execute the forward model,
  - Determine the mismatch between the forward model and the data,
  - Distribute the error gradient to the parts of the model that interact with the ray.
- Generate an appropriate direction given the error gradient, regularization, and the previous descent direction.
- Perform line minimization to find the minimum error in the chosen direction.
- If the error is small enough, exit, otherwise perform another conjugate gradient iteration.

### B. Adjoint Method as a Source of the Gradient

We use the adjoint method on the one dimensional ray equation to determine the error function gradient for the voxels along the ray. We represent position along the ray by  $s$ , modeled intensity along the ray at any point by  $I(s)$ , and modeled attenuation at any point along the ray by  $\mu(s)$ . The initial intensity of the ray is given by  $I(0) = I_0$ . The one dimensional equation for intensity along the ray is

$$\frac{dI}{ds} + \mu(s)I(s) = -I_0\delta(s) \quad (1)$$

where in this case  $\delta(s)$  is a Dirac delta function. We define the error functional of the attenuation distribution as

$$E[\mu(s)] = \frac{1}{2} [I(s_{final}) - I_{obs}(s_{final})]^2 \quad (2)$$

where  $s_{final}$  is at the detector and  $I_{obs}$  is the detected intensity.

We want the gradient of the error with respect to the modeled attenuation distribution. This gradient, when integrated with the variation in the attenuation, gives the variation of the error:

$$\Delta E[\mu(s)] = \int \nabla E(S) \Delta \mu(s) ds \quad (3)$$

here the  $\Delta$  indicates variation. The variation of the ray

equation is given by

$$\frac{d\Delta I}{ds} + \mu(s)\Delta I(s) + I(s)\Delta \mu(s) = 0 \quad (4)$$

and the variation of the error is given by

$$\Delta E[\mu(s)] = [I(s_{final}) - I_{obs}(s_{final})] \Delta I(s_{final}). \quad (5)$$

The forward ray equation represents forward projection through the model. The adjoint ray equation describes the backward projection of the error between the model and the observed data:

$$-\frac{d\tilde{I}}{ds} + \tilde{I}(s)\mu(s) = -\tilde{S}(s) \quad (6)$$

where the source term  $\tilde{S}(s)$  is, in effect, an initial condition:

$$\tilde{S}(s) = [I(s_{final}) - I_{obs}(s_{final})] \delta(s - s_{final}) \quad (7)$$

and thus the variation of the error is given by

$$\Delta E[\mu(s)] = \int \tilde{S} \Delta I ds = \int \left[ -\frac{d\tilde{I}}{ds} + \tilde{I}(s)\mu(s) \right] \Delta I ds. \quad (8)$$

Using the identity

$$\Delta I \frac{d\tilde{I}}{ds} = -\tilde{I} \frac{d\Delta I}{ds} + \frac{d}{ds} (\tilde{I} \Delta I) \quad (9)$$

and realizing that we can disregard the right hand term of (9) because it is zero at the endpoints, we find (8) becomes

$$\Delta E[\mu(s)] = -\int \left[ \frac{d\Delta I}{ds} + \Delta I(s)\mu(s) \right] \tilde{I} ds. \quad (10)$$

Substituting (4) into (10) yields a form of the variation of the error from which we can easily extract the gradient:

$$\Delta E[\mu(s)] = \int I(s) \tilde{I}(s) \Delta \mu(s) ds = \int \nabla E(s) \Delta \mu(s) ds. \quad (11)$$

The gradient is thus

$$\nabla E(s) = I(s) \tilde{I}(s). \quad (12)$$

### Evaluating the Gradient Along a Ray

Given the gradient of the error as a function of position along the ray, how do we evaluate it? For simplicity, assume a uniform attenuation distribution,  $\mu(s) = \mu$ . Over the course of the forward projection, the intensity at any position is then:

$$I(s) = I_0 e^{-\int \mu(t) dt} = I_0 e^{-\mu s}. \quad (13)$$

Suppose the result of the forward projection is not the same as the observed intensity. The difference is the initial condition on the back projection. Over the course of the back projection, the intensity at any position is:

$$\tilde{I}(s) = [I(s_{final}) - I_{obs}(s_{final})] e^{-\mu(s_{final}-s)} \quad (14)$$

and the resultant product at any position is

$$I(s) \tilde{I}(s) = [I(s_{final}) - I_{obs}(s_{final})] I(s_{final}). \quad (15)$$

Equation (15) holds for any distribution of attenuation along the ray for the simple attenuation model.

### Evaluating the Gradient Along a Ray for a Voxel

If the attenuation distribution to be found is represented by the sum of  $N$  basis functions  $\phi_i(s)$  with weightings  $p_i$

$$\mu(s) = \sum_{i=1}^N p_i \phi_i(s) \quad (16)$$

then the finite-dimensional gradient is given by

$$\frac{\partial E}{\partial p_i} = \int I(s) \tilde{I}(s) \phi_i(s) ds = I(s_{final}) \tilde{I}(s_{final}) P_i \quad (17)$$

where  $P_i$  is the projection of the ray through the basis function.

### Evaluating the Total Gradient for a Voxel

We extend the error for a ray to the error over the entire reconstruction as follows:

$$E[\mu(\vec{r})] = \frac{1}{2} \sum_{m=1}^M w_m [I_m(s_{final}) - I_{m,obs}(s_{final})]^2 + \beta \iiint \sqrt{\nabla\mu(\vec{r}) \cdot \nabla\mu(\vec{r})} d^3r \quad (18)$$

where  $m$  is the ray index,  $\beta$  determines the balance between the pure error term and the total variation regularization term, and  $w_m$  is used to weight the rays. The gradient for the  $i$ th voxel is then

$$\nabla E(\vec{r}_i) = \sum_{m=1}^M w_m I_m(s_{final}) \tilde{I}_m(s_{final}) P_{i,m} - \beta \frac{\nabla^2 \mu(\vec{r}_i)}{\sqrt{\nabla\mu(\vec{r}) \cdot \nabla\mu(\vec{r})}} \quad (19)$$

where  $P_{i,m}$  is the projection of the  $m$ th ray through the  $i$ th voxel.

### C. Model Pruning

In many cases there are regions of the model that are intersected by rays that are un-attenuated. It is a waste of computational resources, and a potential source of error, to let those parts of the model affect the calculation. The way we determine which parts of the model to eliminate is to count both the number of ray interactions that each part of the model experiences as well as the number of those interactions that occur with un-attenuated rays.

If the number of un-attenuated interactions is greater than a chosen ratio of the total number of interactions (we use 0.25), then that piece of the model is eliminated. One problem that we have encountered is defining the threshold for an un-attenuated ray.

### D. Approximate Error

One of the most time-consuming steps of many descent methods, for example the conjugate gradient algorithm, is the search for the minimum in the descent direction. This can involve tens of evaluations of the forward model and error function for the problem. For even a modest problem involving 1000 rays in 1000 views interacting with approximately 1000 voxels apiece, there are on the order of  $1e9$  multiply-adds per forward modeling and error evaluation.

There is a long history in the conjugate gradient solver community of trying to cut computational effort by using inexact line searches [5,6,7]. In all of these efforts the line search is stopped before it finally converges. Another way of using these results is not to stop the line search before it converges, but to use an approximation to the error. If we take a random sample of the set of all rays for each conjugate gradient iteration and use that to approximate the behavior of the error in the minimization step, we can cut the computational effort significantly. The approximate error we use is the squared error of a random subset of the rays that have at least 0.1% mismatch between the modeled ray intensity and the detected ray intensity. The actual value of the approximate error does not matter as long as the minimum of the approximate error occurs near the minimum of the true error (as defined in [5], [6], and [7]).

One full forward model and error computation must be performed before the line search in order to generate the gradient for the entire problem. This cuts the computational effort of the line search down to less than two full evaluations of the forward model (and resultant error) and consequently speeds the algorithm by a factor of between 10 and 40 times.

The problems we have observed using this technique are that close to the converged solution it becomes difficult to select an appropriate set of rays with which to approximate the error. At this point it is reasonable to switch to the full error conjugate gradient. Depending on when this switch occurs it may significantly reduce the time savings of the method. Another drawback is that the same data will yield different results depending on which random sets of rays are used in the line searches. This drawback can be alleviated by switching to the full error minimization as the problem nears convergence.

### E. Ray Weighting

One of the strengths of iterative methods is that rays can be weighted in importance. A ray that is heavily attenuated will have more Poisson noise and be more sensitive to scatter and beam hardening. By adjusting the weights of the rays we can incorporate our knowledge of these noise sources. In [8] a case is made for weighting by the ray transmission,  $I_{obs}/I_0$ . We examined two ray weighting policies:

$$w_m = \left(\frac{I_{obs}}{I_0}\right)^p \quad (20)$$

which weights a ray by the transmission to a power  $p$ , and

$$w_m = \frac{1}{2} + \frac{1}{\sqrt{\pi}} \int_0^{s\left(\frac{I_{obs}}{I_0} - x\right)} e^{-t^2} dt = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(s\left(\frac{I_{obs}}{I_0} - x\right)\right) \quad (21)$$

which weights rays by a sigmoidal function, erf, centered around  $x$  with a sharpness  $s$ . These weighting policies are illustrated in Fig. 3.

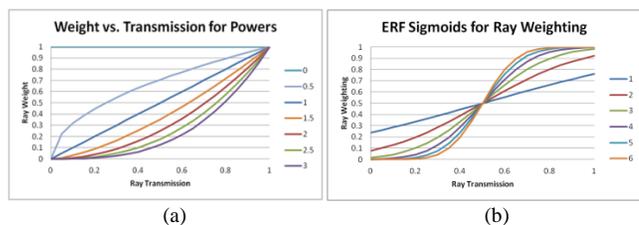


Fig. 3. (a) Ray weighting by powers of transmission. Each curve represents the weighting by transmission to a different power. (b) Ray weighting by a sigmoidal function of ray transmission. Here  $x = 0.5$  is fixed and each curve illustrates the effect of a different  $s$ .

We use the situation of the bar of steel suspended over the container of jelly (Fig. 1.) as our sample problem. One metric for performance is the difference between the mean jelly attenuation in this situation as compared to the mean jelly attenuation when it is scanned alone (bare) in a bin. Another metric is the flatness of the attenuation along the length of the steel bar. For this problem we are not using

regularization, so  $\beta = 0$ . In this situation we have found that ray weighting by powers of transmission yields better results than ray weighting by sigmoidal functions. This is illustrated by Figs. 4-9. In Figs. 5-9 attenuation is measured in Livermore Modified Hounsfield Units (LMHU) which assigns the attenuation of air to 0 and water to 1000.

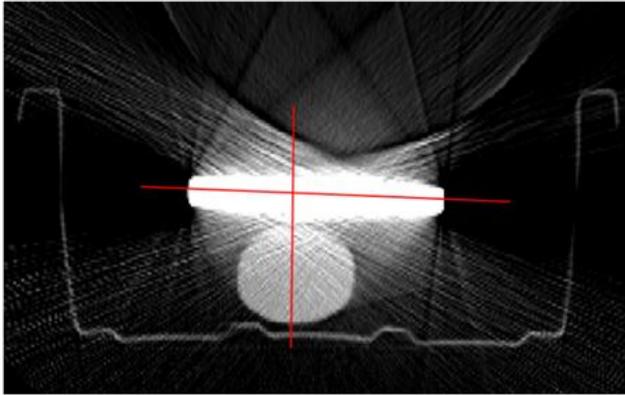


Fig. 4. Lineouts taken along steel bar and through bar and jelly to determine the performance of the reconstruction as a function of ray weighting.

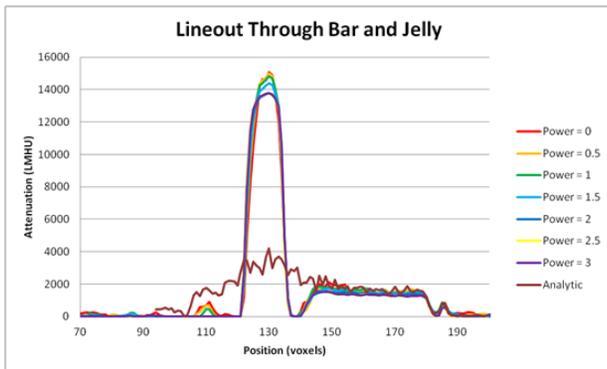


Fig. 5. Lineouts through steel bar and jelly for analytic and iterative reconstruction with transmission power ray weighting.

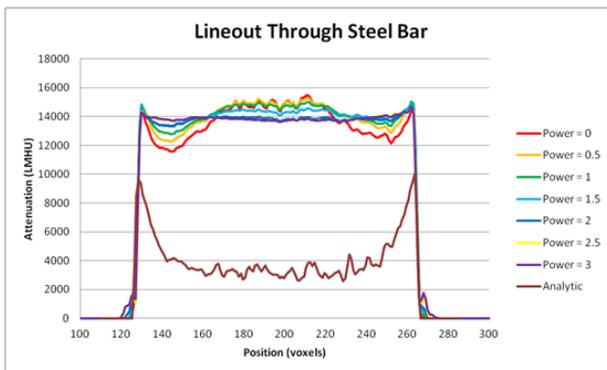


Fig. 6. Lineouts through steel bar for analytic and iterative reconstruction with transmission power ray weighting.

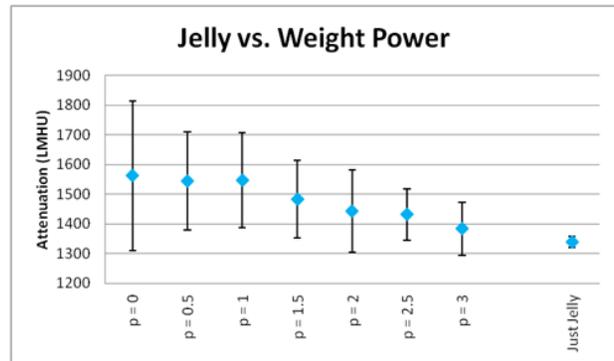


Fig. 7. Mean and standard deviation of jelly as a function of transmission power ray weighting compared to jelly alone (bare) in an airport bin.

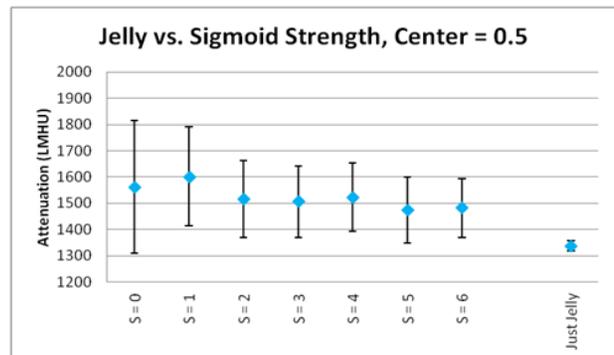


Fig. 8. Mean and standard deviation of jelly as a function of the strength of sigmoidal ray weighting. The center of the sigmoid was chosen to be 0.5.

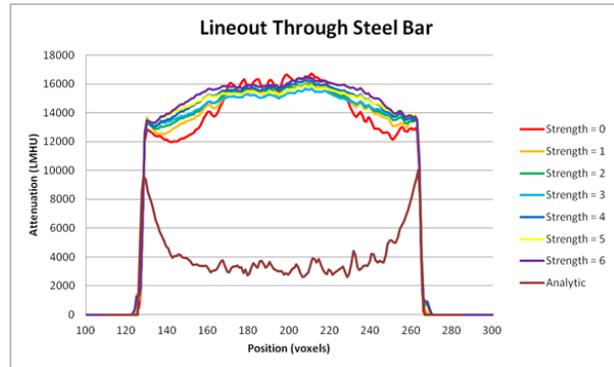


Fig. 9. Lineouts through steel bar for analytic and iterative reconstruction with sigmoidal power ray weighting. The sigmoidal center was set at 0.5 while the strength was varied.

### III. FUTURE WORK

There are several directions for future work. We now need to apply these techniques to a wide variety of data to determine when they are applicable and whether the effects of containment, clutter and concealment can be reduced in the feature spaces that automatic threat detection are performed in.

### ACKNOWLEDGMENT

We are grateful for the flexibility of the conference organizers in allowing us to get our paper in post-deadline.

## REFERENCES

- [1] G. Wang, D. L. Snyder, J. A. O'Sullivan, and M. W. Vannier, "Iterative Deblurring for CT Metal Artifact Reduction," *IEEE Trans. Med. Imaging.*, vol. 15, no. 5, pp. 657-664, October 1996.
- [2] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, New York 1988.
- [3] S. J. Norton, "Iterative inverse scattering algorithms: Methods of computing Frechet derivatives," *J. Acoust. Soc. Am.*, vol. 106, no. 5, pp. 2653-2660, Nov. 1999.
- [4] D. M. Goodman, E. M. Johansson and T. W. Lawrence, *Multivariate Analysis: Future Directions*, C. R. Rao, editor Elsevier Science Publishers, 1993, pp. 209-232.
- [5] P. Wolfe, "Convergence Conditions for Ascent Methods," *SIAM Review*, vol. 11, no. 2, pp. 226-235, April 1969.
- [6] M. L. Lenard, "convergence Conditions for Restarted Conjugate Gradient Methods with Inaccurate Line Searches," *Mathematical Programming*, vol. 10, pp. 32-51, 1976.
- [7] D. F. Shanno, "On the convergence of a new conjugate gradient algorithm," *SIAM J. Numer. Anal.*, vol. 15, no. 6, pp. 1247-1257, Dec. 1978.
- [8] K. Sauer and C. Bouman, "A Local Update Strategy for Iterative Reconstruction from Projections," *IEEE Trans. Sig. Proc.*, vol. 41, no. 2, pp. 534-548, Feb. 1993.