Optimization for Object Localization of the Constrained Algebraic Reconstruction Technique

K. M. Hanson*
Los Alamos National Laboratory, MS P940
Los Alamos, NM 87545 USA

Abstract

A method for optimizing image-recovery algorithms is presented that is based on how well the specified task of object localization can be performed using the reconstructed images. The task performance is numerically assessed by a Monte Carlo simulation of the complete imaging process including the generation of scenes appropriate to the desired application, subsequent data taking, image recovery, and performance of the stated task based on the final image. This method is used to optimize the constrained Algebraic Reconstruction Technique (ART), which reconstructs images from their projections under a nonnegativity constraint by means of an iterative updating procedure. The optimization is performed by finding the the relaxation factor, which is employed in the updating procedure, that yields the minimum rms error in estimating the position of discs in the reconstructed images. It is found that the optimum operating points for the best object localization are essentially the same as those obtained earlier when the performance of simple object detection is to be optimized.

Introduction

Previously we showed how the evaluation of image-recovery algorithms could be based on how well the resulting reconstructions allow one to perform the tasks set forth for the imaging system [1]. A technique that permits one to numerically evaluate a task performance index for a specified imaging situation was proposed. This technique consists of a Monte Carlo simulation of the entire imaging process including random scene generation, data taking, reconstruction, and task performance. Accuracy is judged by comparison of the results with the original scene. Repetition of this process for many possible scenes provides a statistically significant estimate of the performance index that has been chosen to summarize the accuracy of the task performance. Averaging over many scenes is important because artifacts in reconstructed images depend on the scene being reconstructed. Thus a single realization of a simple scene is completely inadequate to judge a reconstruction algorithm. The advantage of this numerical approach is that it readily handles complex imaging situations, nonstationary imaging characteristics, and nonlinear reconstruction algorithms. Its major disadvantage is that it provides an evaluation that is valid only for the specific imaging situation investigated. More detail about this method can be found in [1].

An advantage of the Monte Carlo method of performance evaluation is that the reconstruction algorithm may be optimized for any fixed number of iterations. Alternatively, the number of iterations may be varied to achieve the optimum performance for algorithms that tend to diverge after many iterations. Such behavior is observed in the EM (estimation maximization) algorithm [2,3] and, when data are inconsistent, some implementations of the ART algorithm [4].

The above method to evaluate task performance has been used [5] to optimize the tomographic reconstruction algorithm, constrained ART (Algebraic Reconstruction Technique) [4] with respect to detectability. The object of that study was the relaxation factor, about which there is very little theoretical guidance. Two parameters were used to control the behaviour of the relaxation factor as a function of iteration number. The optimum operating point was found by searching for the combination that yields the largest detectability index d'. This optimization procedure demanded much higher relaxation factors than suggested by theory for unconstrained ART. In [1] we demonstrated that in certain imaging situations the use of the nonnegativity

*This work was supported by the United States Department of Energy under contract number W-7405-ENG-36.

constraint in the reconstruction process improved d' by almost a factor of three when nominal relaxation factors were used. It was found that with optimization, d' could be increased by another factor of ten. The improvement in d' correlated well with visually estimated image quality. If the optimization of the algorithm with respect to the common measure of reconstruction faithfulness, the rms difference between the reconstruction and the original scene, were followed, considerably poorer detectability and perceived image quality resulted.

We report on an extension of the previous work to the consideration of the performance of a higher order task, namely estimation of object location. This extension is pertinent because the accuracy of object localization depends on the higher spatial frequencies in the reconstruction, which are believed to dominate the performance of other high-order tasks such as medical diagnosis [6].

Estimation of Object Position

The task of estimating the position of an object with known shape is performed using a minimum chi-squared (χ^2) fitting procedure. It is known that this procedure is equivalent to maximum likelihood estimation [7] when the noise is gaussian distributed and uncorrelated. For a set of measurements f_i, χ^2 is given by

$$\chi^2 = \sum \frac{(f_i - p_i(\alpha))^2}{\sigma_i^2} \,, \tag{1}$$

where $p_i(\alpha)$ is the predicted value of the ith measurement for the parameter vector α and σ_i is the rms noise in the ith measurement. The sum is over all the measurements that are to be included in estimating the unknown parameters. The best fit to the data is obtained by finding the set of parameters α that minimizes χ^2. The fitting algorithm we have used is essentially identical to the CHIFIT program presented by Bevington [8] for fitting a nonlinear function of the parameters. For the task at hand, the estimation of the location of a known signal, the predicted function p_i is nonlinearly related to the position.

Algebraic Reconstruction Technique

The Algebraic Reconstruction Technique (ART) [4] is an iterative algorithm that reconstructs a function from its projections. It has proven to be a very successful tomographic reconstruction algorithm, particularly when there is a limited number of projections available. Assume that N projection measurements are made of the unknown function f, which is considered a vector. The ith measurement is writeen as

$$g_i = H_i f, \quad i = 1, ..., N, \tag{2}$$

where H_i is the corresponding row of the measurement matrix. The ART algorithm proceeds as follows. An initial guess is made; for example, $f^0 = 0$. Then the estimate is updated by iterating on the individual measurements taken in turn:

$$f^{k+1} = f^k + \lambda^k H_i^{\mathrm{T}} \left[\frac{g_i - H_i f^k}{H_i^{\mathrm{T}} H_i} \right] , \tag{3}$$

where f^k is the kth estimate of the image vector f, $i = k \bmod(N)+1$, and λ^k is a relaxation factor for the kth update. In constrained ART a nonnegativity constraint is enforced by setting any component of f^{k+1} to zero that has been made negative by the above updating procedure. We use the index K to indicate the iteration number ($K = \mathrm{int}(k/N)$), which in the standard nomenclature corresponds to one pass through all N measurements. We express the relaxation factor as

$$\lambda^K = \lambda_0 (r_\lambda)^{K-1}. \tag{4}$$

The proper choice of the relaxation factor is the issue at hand. There is very little guidance on this choice in the literature. A value of unity is often suggested and used. It is known [9] that if a unique solution to the measurement equations exists, the ART algorithm converges to it in the limit of an infinite number of iterations provided that $2 > \lambda > 0$. If many solutions exist, ART converges to the one with minimum norm. Censor *et al.* [10] have shown that unconstrained ART ultimately converges to a minimum-norm least-squares solution if the relaxation factor approaches zero slowly enough. However, λ^K asymptotically approaches zero for any value of $r_\lambda < 1$. The value appropriate for a finite number of iterations remains uncertain. In previous work the author has assumed for λ_0 and r_λ the nominal values of 1.0 and 0.8 for problems involving a limited number of projections, and 0.2 and 0.8 for problems involving many (~100) views [1]. Next we discuss a way to find the best choice for the relaxation parameters for a specific problem.

Optimization of ART

Several classes of measures have been employed in the past on which to base the optimization of image-recovery algorithms [11]. Some are based on how close the reconstructed images are to the original image, such as the conventional measure of the rms difference between the reconstruction and the original image, simply called the rms error. This figure of merit may be convenient from a mathematical standpoint, but it does not correlate well with the usefulness of reconstructed images. There are alternative measures based on how closely the estimated reconstruction reproduces the measurement data, for example, the mean-square residual. Unfortunately, without further constraints, reconstruction based on minimizing the mean-square residual is known to be ill-conditioned or even worse, ill-posed [11]. We have proposed [5] that the most meaningful measure upon which to optimize reconstruction algorithms is the ability to perform the kind of task for which the imaging system was intended. We will use the following example to demonstrate how this can be accomplished.

The numerically calculated task performance can be used to search for the optimum choice of λ_0 and r_λ for the ART algorithm. For the present purpose, the scene is assumed to consist of a number of non-overlapping discs placed on a zero background. For this example, each scene contains 10 high-contrast discs of amplitude 1.0 and 10 low-contrast discs with amplitude 0.1. The discs are randomly placed within a circle of reconstruction, which has a diameter of 128 pixels in the reconstructed image. The diameter of each disc is 8 pixels. In this computed tomographic (CT) problem, the measurements are assumed to consist of a specified number of parallel projections, each containing 128 samples. Ten iterations of ART are used in all of the present examples. It is assumed that the task to be performed is the estimation of the positions of the discs. To produce noisy data, random noise is added to the projection measurements using a Gaussian-distributed random number generator. For a display of the kinds of scenes used in this study, please refer to [1,5].

As mentioned above, estimation of the position the discs is performed using a minimum χ^2 fitting technique. For the input measurements to the fitting procedure, we use the set of pixels in the reconstructed image f that fall inside a circle with a radius 1.7 times the radius of the disc whose position is to be determined. This circular fitting region is centered on the position of the disc being fit. The function fitted to the data is a disc of variable amplitude with a linearly tapered edge, which is chosen to approximately match the shape of the reconstructed discs. The background is assumed to be zero. In the fits performed here, the radius and taper of the discs are held constant and the amplitude and the horizontal and vertical position of each disc are allowed to vary. The rms noise in the measurements σ_i is assumed to be constant. For the performance index, we use the rms error in the estimated position of the discs, averaged over both horizontal and vertical positions of all the discs of the same amplitude in all ten reconstructions. We refer to this performance index as localizability and designate it as σ_Δ. An inherent problem arises when the presence of a disc is uncertain. In the present study when the fitted amplitude is less than 20% of the correct value, we assume the disc is not reliably detected and simply replace the estimated position with a randomly generated position located within the range of the data being fitted. However, such a strategy can lead to a lack of continuity in the optimizing function, which can easily play havoc with any routine that is to find the minimum of such a function.

Fig. 1 shows how various choices for optimization functions depend on λ_0 and r_λ for constrained ART in one data-taking situation. The contours for σ_Δ obtained by fitting the high-contrast discs are remarkably similar to those obtained by fitting the low-contrast discs. In fact, these two optimization functions are almost exactly an even factor of ten apart; that is, in the same ratio as their amplitudes. Thus it doesn't

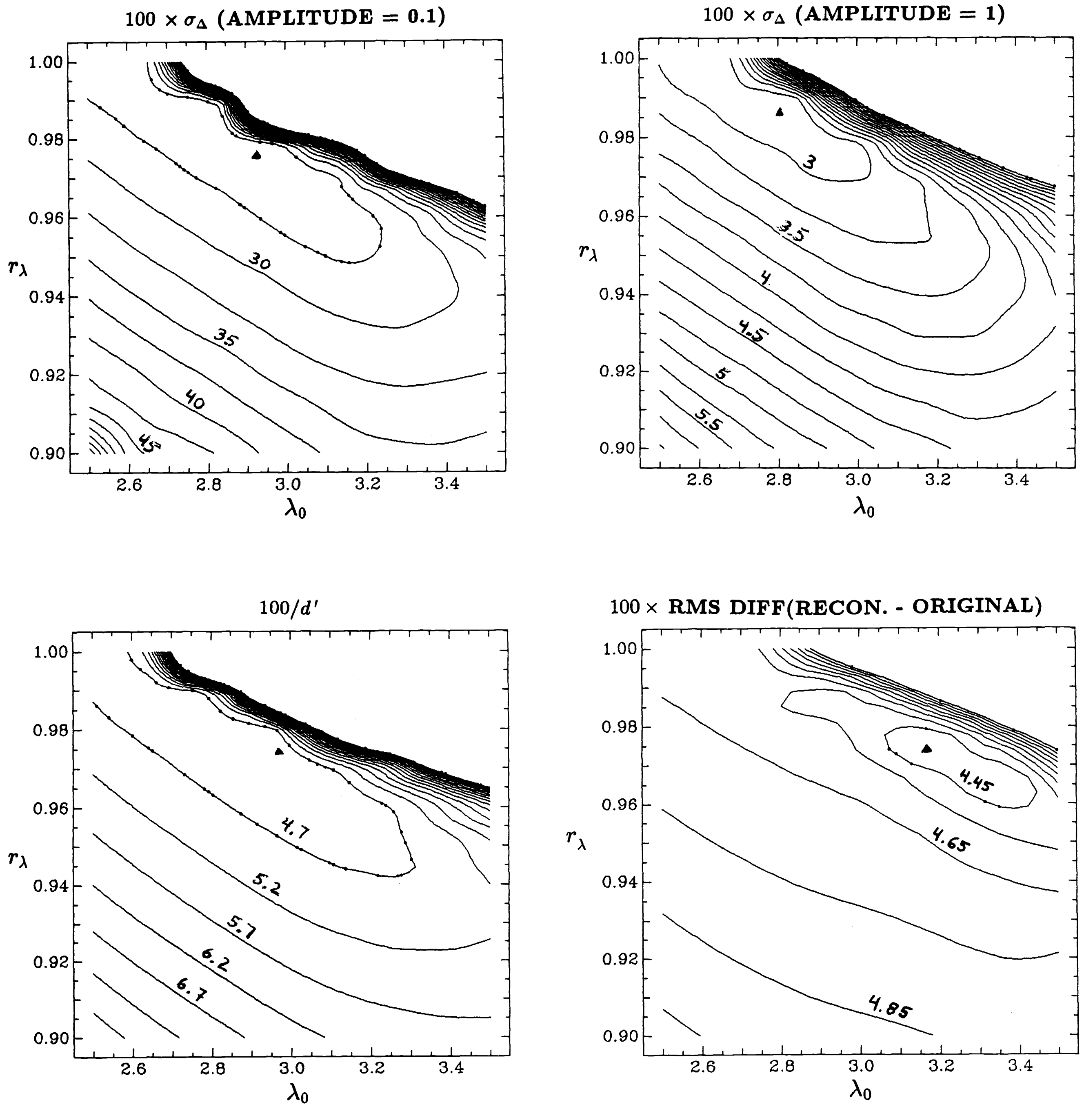

Figure 1: Contour plots of four optimization functions plotted as a function of the relaxation parameters λ_0 and r_λ used in the constrained ART reconstruction algorithm. The measurement data consist of 12 noiseless, parallel projections spanning 180°. The coarse sampling (10 x 10 points) of these functions, necessitated by the lengthy computation time required for each function evaluation, accounts for the scalloping effects.

matter which set of objects is used for position estimation. This fact is important because in some situations the low-contrast discs are difficult to detect, giving rise to the problem described above of knowing how to score such cases. Furthermore, these contours are quite similar to those for the optimization function based on the detectability index, i. e. $100/d'$. Thus optimization with respect to any of these performance indices yields the same operating point. Recall, however, the previous demonstration [5] that optimization based on the rms error in the reconstruction resulted in more artifacts than optimization based on the detectability index d'.

Figures 2 and 3 compare the optimization function based on localizablity to that based on detectability for two other data-taking situations in which random noise is added to the projection data. In these cases the contours of the two optimization functions are not nearly as identical in shape as they are in Fig. 1, but they do show the same general trends. In Fig. 3 both $1/d'$ and σ_Δ show little dependence on λ_0 and similar dependences on r_λ, but with different positions of the minima. In Fig. 2 the two functions have nearly the same minima but demonstrate somewhat different characteristics in their dependence on the two variables.

The optimum values for λ_0 and r_λ are found for various conditions of data collection using a function minimizer from the NAG library[1] called E04JBE. This routine finds the minimum of a function of many parameters after numerous evaluations of the function. From 20 to 100 function evaluations are required for the cases studied here in which just two parameters are varied. Table 1 tabulates the results obtained previously with constrained ART for optimization with respect to the detectability of the low-contrast discs. The nonnegativity constraint was found to be generally useful with the nominal relaxation factors, particularly when the data are limited by the measurement geometry. Optimization produced even further improvements in detectability. Very large relaxation factors are preferred, in fact much larger than might be expected. However, when it is realized that the nonnegativity constraint has the effect of undoing the agreement with each measurement that should result from an update, it seems reasonable that overrelaxation is desirable. Neither the use of nonnegativity nor optimization has much benefit when the data are complete but noisy.

The results of optimizing constrained ART with respect to the accuracy of position estimation are presented in Table 2. As a general observation, optimization with respect to localizability yields very similar operating points for λ_0 and r_λ. Furthermore, the factors by which improvement is made in the optimization function by moving from the nominal relaxation parameters to the optimized ones is nearly the same for each data-taking situation. It appears that the effect of artifacts on $1/d'$ and σ_Δ are similar. Perhaps this fact is a consequence of the randomization produced in the artifacts arising from the randomized placement of the discs in the many scenes that are used to calculate the average performance index. Then the effects of artifacts might be expected to behave similarly to those of additive random noise in the measurements.

The conclusions regarding the optimization with respect to localizability in *unconstrained* ART reconstructions are essentially the same as those previously drawn about detectability. Relatively little improvement in localizability is achieved by optimization compared to that obtained with the nominal relaxation factors. In the noiseless cases, a value of unity for λ^K yields essentially the same results as the optimized values, a choice that is in agreement with common practice. However, for noisy data it seems desirable for r_λ to be less than unity and, when there are many views, λ_0 should be small. These choices are reasonable as they promote significant averaging over all the views. As a rule of thumb, for noisy but complete data, the relaxation factor should be approximately equal to the reciprocal of the number of views for the last few iterations.

[1]Numerical Algorithm Group, 7 Banbury Road, Oxford OX2 6NN, UK

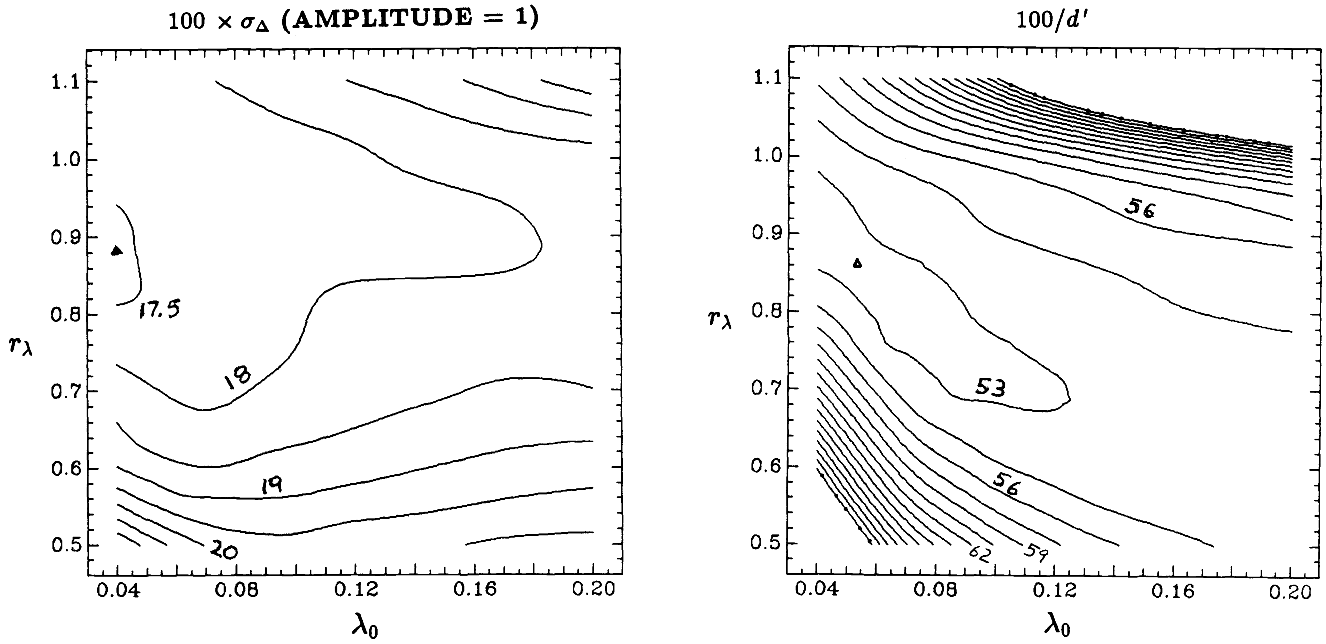

Figure 2: Contour plots obtained with constrained ART for measurement data consisting of 100 parallel projections spanning 180° containing random noise with an rms amplitude of 8.

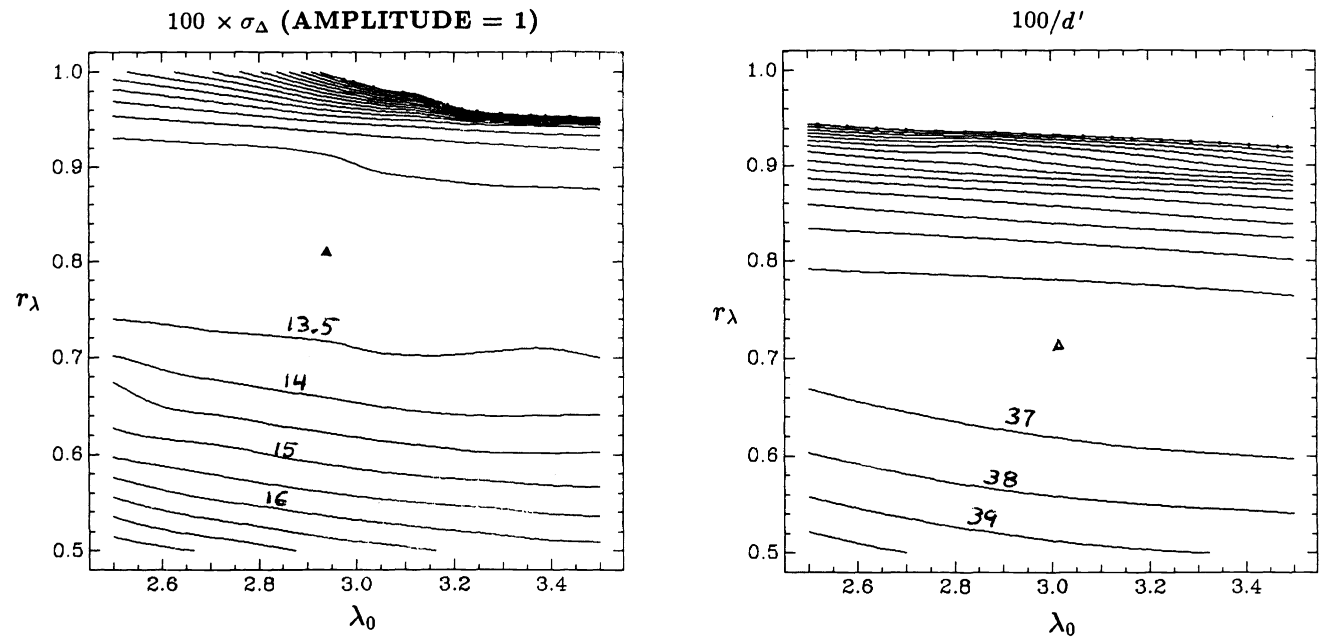

Figure 3: Contour plots obtained with constrained ART for measurement data consisting of 16 parallel projections spanning 180° containing random noise with an rms amplitude of 2.

Table 1: Summary of the effect of optimization with respect to the detectability index d' of the low-constrast discs in reconstructions provided by constrained ART (repeated from Ref. [1]). Dramatic improvement in detectability is seen to be possible when the measurement geometry limits interpretation of the reconstruction rather than noise in the data.

number proj.	$\Delta\theta$ (deg.)	rms noise	nominal λ_0	nominal r_λ	nominal d'	optimized λ_0	optimized r_λ	optimized d'
100	180	8	0.2	0.8	1.825	0.052	0.859	1.908
8	180	0	1.0	0.8	0.653	3.45	0.959	4.91
12	180	0	1.0	0.8	2.054	2.96	0.975	23.46
16	90	0	1.0	0.8	2.050	2.78	0.967	6.30
16	180	2	1.0	0.8	2.372	3.01	0.712	2.747

Table 2: Summary of the effect of optimization with respect to the localizability of the high-constrast discs in reconstructions obtained with constrained ART. The object localizability σ_Δ is given in terms of pixels.

number proj.	$\Delta\theta$ (deg.)	rms noise	nominal λ_0	nominal r_λ	nominal σ_Δ	optimized λ_0	optimized r_λ	optimized σ_Δ
100	180	8	0.2	0.8	0.182	0.046	0.920	0.174
8	180	0	1.0	0.8	0.472	3.24	0.977	0.104
12	180	0	1.0	0.8	0.236	2.80	0.989	0.0277
16	90	0	1.0	0.8	0.426	2.41	0.998	0.149
16	180	2	1.0	0.8	0.160	2.93	0.811	0.130

Discussion

In many of the imaging situations studied, the optimization of constrained ART realized through a judicious selection of the relaxation factor can significantly increase the localizability of objects, especially when the data consist of a limited number of noiseless projections. For unconstrained ART, little improvement can be achieved through optimization.

The accuracy of object localization for constrained and unconstrained ART, with and without optimization, follows the same pattern found earlier for detectability. The optimization functions for the performance of the tasks of position estimation and detection of low-contrast objects show similar trends as a function of the two relaxation parameters λ_0 and r_λ. The optimum operating points in terms of these parameters vary with the data-taking situation but they are nearly the same for both of these tasks. This conclusion is perhaps a little surprising because the task of object localization is more dependent on the high-frequency content of the image than is simple detection [6]. If the parameters being varied in the optimization had separate effects on the modulation transfer function (MTF) and the correlation of the noise in the final images, a different result might have been anticipated. Part of the explanation for the observed similarity in the results is that, in the present case of image reconstruction, the resolution of the final images is not affected much by the relaxation factors. On the other hand, the realization in the reconstruction of random noise present in the projection measurements can be affected. Thus it is in the situations in which noise is added to the measurements that we observe some differences in the optimization functions based on the performance of these two tasks.

It is possible that one may desire to optimize an imaging system with respect to the performance of more than one task and that the individual optimization functions might not have the same minima. This type of behavior is seen, for example, in Fig. 3. In such a case one can combine the various optimization functions into a single grand optimization function by weighting each individual function appropriately. The optimum operating point would strike a balance between the operating points that are best for each of the constituent tasks.

References

[1] K. M. Hanson. Method to evaluate image-recovery algorithms based on task performance. *Proc. SPIE*, 914:336–343, 1988.

[2] L. A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imaging*, MI-1:113–122, 1982.

[3] E. Veklerov and J. Llacer. Stopping rule for the MLE algorithm based on statistical hypothesis testing. *IEEE Trans. Med. Imaging*, MI-6:313–319, 1987.

[4] R. Gordon, R. Bender, and G. Herman. Algebraic reconstruction techniques for three-dimensional electron microscopy and x-ray photography. *J. Theor. Biol.*, 29:471–481, 1970.

[5] K. M. Hanson. POPART - Performance OPtimized Algebraic Reconstruction Technique. *Proc. SPIE*, 1001:318–325, 1988.

[6] K. M. Hanson. Variations in task and the ideal observer. *Proc. SPIE*, 419:60–67, 1983.

[7] K. M. Hanson. Optimal object and edge localization in the presence of correlated noise. *Proc. SPIE*, 454:9–17, 1984.

[8] R. R. Bevington. *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill, New York, 1969.

[9] G. T. Herman, A. Lent, and P. H. Lutz. Relaxation methods for image reconstruction. *Commun. Assoc. Comput. Mach.*, 21:152–158, 1978.

[10] Y. Censor, P. P. B. Eggermont, and D. Gordon. Strong underrelaxation in Kaczmarz's method for inconsistent systems. *Numer. Math.*, 41:83–92, 1983.

[11] H. C. Andrews and B. R. Hunt. *Digital Image Restoration*. Prentice-Hall, Englewood Cliffs, New Jersey, 1977.