

RAYLEIGH TASK PERFORMANCE AS A METHOD TO EVALUATE IMAGE RECONSTRUCTION ALGORITHMS

Kenneth M. Hanson
Los Alamos National Laboratory
MS P940
Los Alamos, New Mexico 87545 USA

Kyle J. Myers
Center for Devices and Radiological Health, FDA
12720 Twinbrook Parkway
Rockville, Maryland 20857 USA

ABSTRACT. We evaluate several tomographic reconstruction algorithms on the basis of how well one can perform the Rayleigh discrimination task using the reconstructed images. The Rayleigh task is defined here as deciding whether a perceived object is either a pair of neighboring points or a line, both convolved with a 2D Gaussian. The method of evaluation is based on the results of a numerical testing procedure in which the stated discrimination task is carried out on reconstructions of a randomly generated sequence of images. The ability to perform the Rayleigh task is summarized in terms of a discriminability index that is derived from the area under the receiver-operating characteristic (ROC) curve. Reconstruction algorithms employing a nonnegativity constraint are compared, including maximum *a posteriori* (MAP) estimation based on the Bayesian method with entropy and Gaussian priors as well as the additive and multiplicative versions of the algebraic reconstruction technique (ART and MART). The performance of all four algorithms tested is found to be similar for complete noisy data. However, for sparse noiseless data, the MAP algorithm based on the Gaussian prior does not perform as well as the others.

1. Introduction

Appropriate measures of the quality of reconstructed images should ultimately be related to how well the images can be used to make decisions or estimate relevant physical quantities. The performance of imaging tasks is basically a statistical matter, not only when the data are degraded by random noise, but also when many randomly placed objects are reconstructed from limited, noiseless data. As such, a reconstruction technique cannot be judged on the basis of a reconstruction of only one scene. We use a method recently introduced to judge the quality of reconstruction by numerically evaluating task performance for a specified imaging situation (Hanson, 1987b; 1988a; 1990a). The method consists of a Monte Carlo simulation of the entire imaging process including random scene generation, data taking, reconstruction, and performance of the specified task. An essential aspect of this method is that the evaluation is based on many randomly generated scenes, leading to

a statistically significant estimate of performance. This method overcomes the restrictions of the standard calculation for the propagation of errors, which is applicable only to image noise that is stationary and approximately Gaussian distributed. Nonlinear reconstruction algorithms lead to image noise that violates these conditions. We use the above technique to evaluate the performance of several reconstruction algorithms, all employing the non-negativity constraint. Such a constraint has been shown to be of great value in improving task performance when dealing with a paucity of data (Hanson, 1988a; 1990a) for as simple a reconstruction algorithm as the algebraic reconstruction technique (ART) (Gordon *et al.*, 1970). This work expands on the previous findings by consideration of algorithms founded in the Bayesian method and by investigating the much more complex task of Rayleigh discrimination. Besides using the now-familiar entropy prior, we also reintroduce the classic Gaussian or normal prior probability distribution, but with the twist of an enforced nonnegativity constraint.

2. Bayesian Reconstruction

If measurements are linearly related to an image and are degraded by additive Gaussian noise, the measurements can be expressed as: $\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}$, where \mathbf{g} represents the vector comprising all measurements, \mathbf{f} the image vector, \mathbf{n} the random noise vector, and \mathbf{H} is the matrix that specifies the linear relationship between the image and the measurements. In the tomographic problems we are considering, the \mathbf{H} matrix corresponds to the calculation of line integrals through the image, that is, the projection calculation. The Bayesian approach to image reconstruction leads to the conclusion that the best reconstruction is the image that maximizes the posterior probability. The Bayesian solution thus minimizes

$$\phi(\mathbf{f}) \sim (\mathbf{g} - \mathbf{H}\mathbf{f})^T \mathbf{R}_n^{-1} (\mathbf{g} - \mathbf{H}\mathbf{f}) + \Pi(\mathbf{f}), \tag{1}$$

which is -2 times the logarithm of the posterior probability, where the first term is χ^2 arising from the likelihood of that particular data set given the image \mathbf{f} . This term measures the disagreement between the image and the data. The second term comes from the assumed prior probability distribution. From the point of view of regularization theory (Nashed, 1981; Titterton, 1985), this term can control the bad behavior of the least-squares solution that minimizes the first term. In the maximum entropy method, the second term is proportional to the entropy of the image

$$\Pi(\mathbf{f}) = -2\alpha S(\mathbf{f}) = -2\alpha \sum_i [f_i - \tilde{f}_i - \ln(f_i/\tilde{f}_i)] . \tag{2}$$

The sum is over all pixels in the image and \tilde{f}_i is aptly called the default value for the i th pixel representing the prior model for the reconstruction. The maximum-entropy prior implies a solution that is nonnegative.

We wish to consider an alternative prior with a significant history in image recovery (Andrews and Hunt, 1977) that is based on a Gaussian probability distribution leading to

$$\Pi(\mathbf{f}) = (\bar{\mathbf{f}} - \mathbf{f})^T \mathbf{R}_f^{-1} (\bar{\mathbf{f}} - \mathbf{f}) , \tag{3}$$

where $\bar{\mathbf{f}}$ is the mean and \mathbf{R}_f the covariance matrix of the prior probability distribution that is assumed to characterize the ensemble of images with which one is dealing. In the standard Bayesian approach introduced to image reconstruction by Hunt (1977), these quantities are assumed to be known *a priori*. Hunt called this method, based on a Gaussian prior,

simply MAP for maximum *a posteriori* reconstruction. We make a distinction between this method and the one based on the entropy prior by calling it GAUSS-MAP. In the present implementation, a nonnegativity constraint is invoked that effectively eliminates the tails of the Gaussian distribution for $f_i < 0$. In comparing the entropy and the Gaussian priors, we note that the former excludes zero as a possible reconstructed value, whereas the latter does not. Also, the entropic probability distribution has a long tail that drops slowly as f_i increases (for small α), while the Gaussian possesses a tail that cuts off more rapidly. The Gaussian prior explicitly encourages correlations between image values to be specified by \mathbf{R}_f . It also provides some additional flexibility in controlling the prior as $\bar{\mathbf{f}}$ can be set arbitrarily (*e.g.*, above or below 0) and independently from \mathbf{R}_f . In contrast, a basic axiom invoked by Skilling (1989) in the derivation of the entropy prior states that no correlation between different pixels should exist. However, this tenet is discarded in the ‘New MaxEnt’ formulation (Gull, 1989).

We use an iterative procedure that is related to the method presented by Butler, Reeds, and Dawson (1981). The reconstruction is taken to be the positive part of a dual functional that is obtained by adding $-\lambda^k \gamma^k \nabla \phi(\mathbf{f}^k)$ to the k th iterate of the reconstruction \mathbf{f}^k . As the present GAUSS-MAP algorithm is closely related to the method of steepest descent, it does not converge to the solution very quickly. Although this technique is satisfactory for the present study, a more sophisticated algorithm (Skilling and Bryan, 1984) might reduce the large number of iterations sometimes required to achieve complete convergence. The CPU time is dominated the number of needed projection operations \mathbf{H} or, equivalently, back-projections \mathbf{H}^T . For GAUSS-MAP the scalar γ^k must be determined for each iteration by multiplication of the incremental change in \mathbf{f}^k by the curvature matrix $[\mathbf{R}_f^{-1} + \mathbf{H}^T \mathbf{R}_n^{-1} \mathbf{H}]$. Therefore each iteration requires four equivalent projection calculations.

The maximum-entropy reconstructions shown here are obtained with the MEMSYS 2 code*, which is referred to later as ENTROPY-MAP. The operation of this algorithm has been described by Skilling and Bryan (1984). We are not using the ‘New MaxEnt’, which is available as MEMSYS 3, but will address its added features in a future publication. In the four-dimensional search mode used here, between 6 and 20 equivalent projection calculations are required per iteration, with the average number being about 10. Results are also shown for ART (Gordon *et al.*, 1970) and its multiplicative cousin MART, which is known to converge to the solution of the measurement equations with maximum entropy (Lent, 1977). Both algorithms require two equivalent projection calculations per iteration. They are noteworthy for their speedy convergence to an acceptable solution.

3. Estimation of Rayleigh Task Performance

A natural imaging task, named after Lord Rayleigh, is to determine whether an observed astronomical object is actually a pair of stars or a single star. Precisely stated, we will consider the task of distinguishing between a pair of point-like objects and a line, both convolved with a 2D Gaussian spread function. By specifying binary pairs that are farther apart than the width of the Gaussian and by specifying the alternative line to be similar in appearance, the decision must be made on the basis of more than an observed slight asymmetry; an honest dip must be observed between the two lobes of a reconstructed binary for it to be distinguished from a line. To accomplish the stated binary decision, it

* Maximum Entropy Data Consultants Ltd., Royston, England

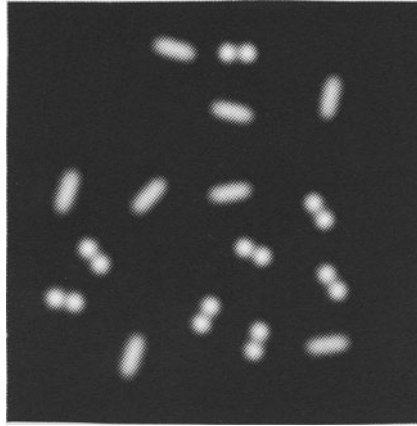


Fig. 1. The first of ten randomly generated scenes used to evaluate performance of the Rayleigh discrimination task. Each scene contains 8 lines and 8 pairs of point-like objects, all convolved with the same 2D Gaussian function.

is necessary to take into account many parameters including signal amplitude, major axis orientation, length of line or point separation, and possibly background level. As these are not specified *a priori*, they need to be estimated as part of the task.

A Monte Carlo technique, one that employs pseudo-random numbers to generate its results, is used to simulate the entire imaging process from scene generation to the final task performance. The general strategy for simulation of the scenes, data taking, and reconstruction is straightforward (Hanson, 1988a; 1990a). What is new here is the Rayleigh task, which affects how the scenes are generated and how the reconstructions are analyzed to carry out the implied binary decision. The scenes used to test task performance are generated each with eight binary pairs and eight lines randomly placed and randomly oriented with the circle of reconstruction inscribed within a 128×128 -pixel image. The binary pairs of points are separated by 6 pixels and the lines are 10.4 pixels long. Each of these types of objects is convolved with a symmetric 2D Gaussian function with a FWHM of 4 pixels. The line length and amplitude are chosen to minimize the mean-square difference between the two possible objects. The purpose is to make the two objects similar in appearance so the decision must be made on the basis of the details of the image, not gross features such as integrated intensity (Hanson, 1983). The first of the series of images generated for testing performance of the Rayleigh task is shown in Figure 1.

The Rayleigh task is carried out using each reconstruction as follows. Under the prevailing assumption that the orientation, amplitude, asymmetry, position, binary separation (or line length), and Gaussian width of the objects are not known beforehand, these must be determined from the reconstruction. The background is also assumed to be unknown. The eight unknown parameters are fit to minimize the mean-square difference between the reconstructed region and the model for the object over a 16-pixel-diam region centered on each object. The optimal decision variable would be the logarithm of the ratio of the likelihoods of the two hypotheses. The mean-square error is the proper likelihood function when the error fluctuations are independent, stationary, and Gaussian distributed. How-

ever, the error fluctuations in reconstructions obtained under nonnegativity constraints do not conform to these assumptions. Unfortunately, their complete characterization would be nearly impossible to incorporate into the analysis. To surpass such obstacles the decision variable is taken to be the difference between the mean-square error for the two hypotheses. Frequency histograms of the decision variable for each of the two possible classes of objects are formed. A receiver operating characteristic (ROC) curve is generated from these as described previously (Hanson, 1988a; 1990a). The area under the ROC curve is an appropriate measure of how well the distinction between the two hypotheses is made. A useful way to summarize the significance of the area is to quote the discriminability index d_A . An alternative index d' is based on the rms widths and separation of the two histograms. It is equal to d_A when the histograms are Gaussian shaped. However, d' is not used here because the histograms of the decision variable may not be Gaussian. The index d_A is preferable as it is invariant under any monotonic transformation of the decision variable whereas d' is not.

4. Results

We report on the results of carrying out the Rayleigh testing procedure for several different data-acquisition conditions. In the examples each view consists of 128 parallel projections and the projection set covers the full range of angles, 180° . In some circumstances noise is included by adding to each measurement a random number taken from a Gaussian probability distribution. As the projection at each angle is convolved with a triangular waveform with a FWHM of three projection samples before reconstruction, it is the final rms noise value after smoothing that is quoted.

Some ubiquitous features of the analysis are the following. In all cases the operating conditions for each algorithm are chosen to maximize the discriminability index for the Rayleigh task. The two MAP algorithms depend on the value for the rms noise level assumed in the measurements σ_n . The value of α in Eq. (2) is adjusted in ENTROPY-MAP to make the rms residual σ_r the same as σ_n , under the classic assumption that this is appropriate. The GAUSS-MAP algorithm allows no such adjustment; the value of σ_r simply determines the weight given to the regularization term. The number of iterations used in these algorithms does not matter much once convergence has been achieved, a condition that is met in most of the situations studied. For the GAUSS-MAP algorithm, the image \tilde{f} is set to a constant 0.0174 to reflect the known mean of the images in the ensemble. The diagonal elements of the matrix R_f are set to $(0.080)^2$ to match the variance of the ensemble. The off-diagonal elements are set to zero, in effect making no prior assumptions about the correlation between pixels. In the spirit of Bayesian analysis, the ENTROPY-MAP parameter \tilde{f} is also set equal to the mean of the ensemble. We use a fixed relaxation factor of $\lambda = 1$ for GAUSS-MAP and $\lambda = 0.5$ for ENTROPY-MAP.

The number of iterations used in ART and MART can influence the final reconstructions, especially for overdetermined but noisy measurements. Therefore the number of iterations and the relaxation parameters (λ_0 , the initial relaxation factor, and r_λ , which multiplies the relaxation factor after each iteration) are varied in the present study to optimize performance of the Rayleigh task as described by Hanson (1988b). The form of the updating used in the MART algorithm is not truly multiplicative, but amounts to an additive backprojection in the logarithm of the reconstructed image. In addition the effect on the reconstruction in any one update is limited to a factor of ten (increase or decrease).

Figure 2 shows the results of using the algorithms tested to reconstruct Figure 1 from 100 views with an rms noise of 2. For comparison, the peak projection value of one of the objects taken along its long direction is 4.2. In this case the data are essentially complete, but noisy. The four reconstructions possess a fair degree of similarity. In the MART result we see the tendency to accentuate the peaks and suppress the background anomalies more than in the others. The statistical nature of task performance can be appreciated by comparing object-by-object these reconstructions with the actual initial scene shown in Figure 1. The noise fluctuations in the reconstructions make some of the bars look more like binaries and conversely. The only way to quantitatively measure how well the Rayleigh task can be performed in this imaging situation is to obtain the results of analysis of many reconstructed objects of both types. This evaluation is done by randomly generating nine more scenes like the one in Figure 1. All 80 objects of both types are fit, as described above, to yield a decision variable. The results are summarized in Tables 1 and 2 in terms of the discriminability index d_A derived from the area under the ROC curves. The d_A values are very nearly the same for all the algorithms. The estimated statistical uncertainty in the d_A values for this data-collection situation is 0.24. The agreement between the different algorithms is much better than this because the results are highly correlated for the reason that the exact same data are employed for all the algorithms.

We conclude that even the algorithms of the maximum-likelihood type (ART and MART) can be operated in a fashion to properly average over numerous projections. For this averaging to occur, the relaxation factor for the last few iterations should be approximately equal to the reciprocal of the number of views. The MART algorithm tends to converge to a solution of the measurement equations substantially faster than ART, so fewer iterations are required for MART. If MART is given many more iterations, it tends to 'overfit' the measurements, resulting in spiky looking reconstructions. This overfitting behaviour can be induced in all the algorithms presented here when dealing with noisy data. For the MAP algorithms it occurs when too small an rms value of the noise is assumed.

Figure 3 shows the results of reconstructing Figure 1 with grossly incomplete and noiseless data. In this situation the reconstructions obtained by the four algorithms show some differences. The erroneous fluctuations in the reconstructions have a different character from those in Figure 2 as they arise from the deficit in the number of available measurements and not from random errors in the measurements. These artifacts have a fairly similar structure in all the reconstructions shown. No overfitting phenomenon is observed for any of the algorithms in this case of noiseless data. From Tables 1 and 2 it is seen that the d_A value for GAUSS-MAP is 42% lower than for ENTROPY-MAP and MART. This poorer performance mirrors the fact that the objects in the GAUSS-MAP reconstruction appear to be reconstructed with slightly poorer spatial resolution. The explanation for the decrease in resolution lies in the form of the reconstruction dictated by the prior. The Gaussian prior implies that the reconstruction is a linear combination of backprojections, whereas for the entropy prior, it is the logarithm of the reconstruction that is proportional to the backprojection form. The inherently multiplicative form implied by the entropic prior is advantageous in reconstructing pointlike objects placed on a zero background. The d_A value for ART is about 15% lower than for ENTROPY-MAP and MART reflecting the fact that the ART reconstruction in Figure 3 appears slightly more blurred.

The results for another data-acquisition situation, 8 views with rms noise = 0.5, are also summarized in the tables. The classic condition of $\chi^2 = \text{number of measurements}$ implies

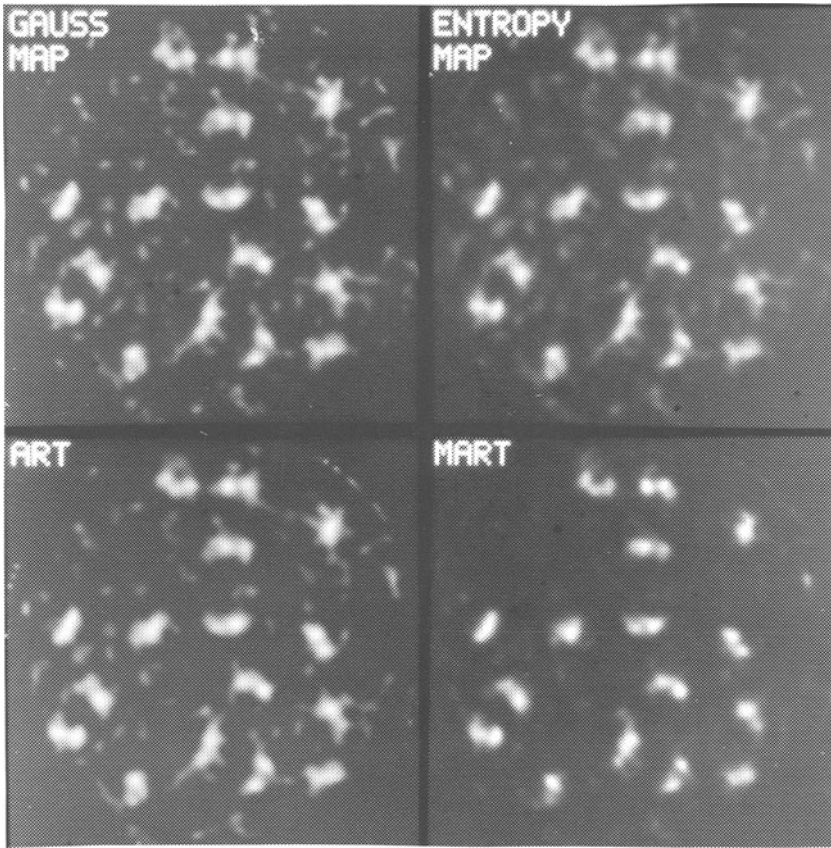


Fig. 2. Reconstructions of Figure 1 from 100 noisy projections obtained with four different reconstruction algorithms that employ a nonnegativity constraint.

$\sigma_n = \sigma_r$. For this condition we find the ENTROPY-MAP algorithm yields $d_A = 0.926$, about 20% lower than for the optimized σ_r . This classic condition on χ^2 is not optimal for task performance, a conclusion previously arrived at by Myers and Hanson (1990) on the basis of object detection and also suggested by others, including Gull (1989).

5. Discussion

We have seen that the constrained ART, MART, GAUSS-MAP, and ENTROPY-MAP algorithms yield similar Rayleigh discriminabilities for some conditions of data acquisition. The similarity between reconstructions obtained by different algorithms might have been anticipated on the basis of the relationship between different methods that is the essence of regularization theory (Nashed, 1981; Titterton, 1985). However, for very incomplete noiseless data, the MART and ENTROPY-MAP algorithms achieve better performance of

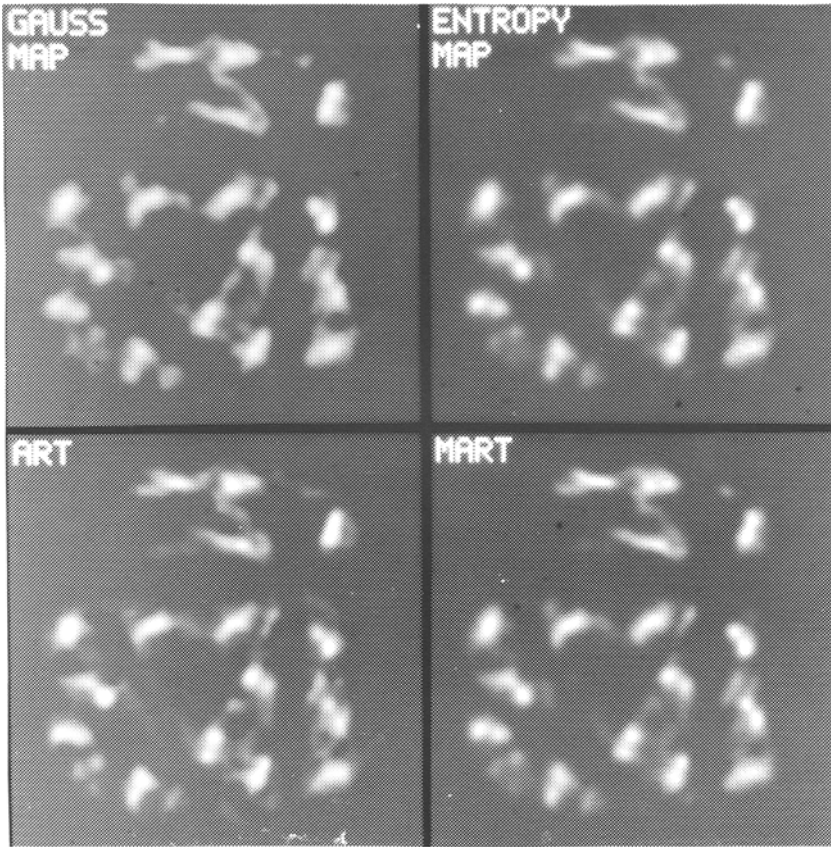


Fig. 3. Reconstructions of Figure 1 from 6 noiseless projections obtained with the nonnegativity constraint. x

the Rayleigh task than the ART and GAUSS-MAP algorithms. The reason for this improved performance may lie in the multiplicative updating inherent in the former algorithms in contrast to the additive updating of the latter.

To achieve the best results, the operating conditions must be carefully chosen (Hanson, 1988b; 1989; 1990b). An advantage of the MAP algorithms is that regularization is explicitly incorporated into their formulation, obviating the necessity to stop the algorithm at a critical number of iterations. However, when the MAP formulations employed in this study have to cope with noisy data, it is still necessary to pick one or more parameters, which can dramatically affect the results. It will be interesting to see whether the 'classic' MaxEnt formulation (Gull, 1989) with its intrinsic determination of α overcomes this difficulty.

In many reports concerning the use of MAXENT, the parameter \tilde{f} is not set equal to the mean of the ensemble of expected images. Rather, it is often set to a very small

Number Proj.	$\Delta\theta$ (deg.)	RMS Noise	GAUSS-MAP				ENTROPY-MAP			
			N	σ_n	σ_r	d_A	N	σ_n	σ_r	d_A
100	180	2	30	0.5	1.89	1.364	15	1.90	1.90	1.325
8	180	0.5	100	0.075	0.23	1.071	30	0.25	0.25	1.135
6	180	0	300	10^{-5}	0.003	1.095	30	0.006	0.006	1.877

Table 1. Summary of the performance of the Rayleigh task based on reconstructions by two Bayesian algorithms employing a nonnegativity constraint. The operating conditions are chosen by adjusting the assumed rms noise σ_n to maximize the discriminability index d_A . The resulting rms residual σ_r is presented.

Number Proj.	$\Delta\theta$ (deg.)	RMS Noise	ART					MART				
			N	λ_0	r_λ	σ_r	d_A	N	λ_0	r_λ	σ_r	d_A
100	180	2	20	0.2	0.85	1.90	1.341	7	0.05	0.8	2.02	1.384
8	180	0.5	100	3.0	0.96	0.24	1.099	30	1.0	0.9	0.29	1.134
6	180	0	100	3.0	0.98	0.008	1.572	40	1.0	1.0	0.013	1.880

Table 2. Summary of the performance of the Rayleigh task based on constrained reconstructions obtained with two forms of the Algebraic Reconstruction Technique, ART and MART, which employ additive and multiplicative updating, respectively. The number of iterations N and the relaxation parameters are chosen to maximize the discriminability index d_A as described by Hanson (1988b).

value, perhaps indicative of the background value. This practice is perhaps understandable because \vec{f} is the default value and will tend to be favored in the absence of good information about the scene being imaged. But this practice seems to be at odds with the historical interpretation of the Bayesian approach in which the prior typically represents the properties of the ensemble of images being studied (Hanson, 1987). Nevertheless, when we use 0.001 for \vec{f} , instead of 0.0174, we find the Rayleigh task performance is essentially unchanged for all circumstances reported in the tables. We note that the choice of \vec{f} or \bar{f} affects the bias in the amplitudes of low-contrast objects in reconstructed images (Myers and Hanson, 1990; Hanson, 1990a).

Auxiliary runs of the GAUSS-MAP algorithm without the nonnegativity constraint show the same trends as previously observed by Hanson (1988a; 1989, 1990a; 1990c) for ART reconstructions obtained with and without the nonnegativity constraint. The non-negativity constraint is found to be very useful when data are meager and noiseless, but, for complete and noisy data, the constraint does not improve the performance of vision tasks.

ACKNOWLEDGMENTS. The authors wish to acknowledge many provocative conversations with Robert F. Wagner. We have also benefitted a great deal from conversations with Stephen F. Gull, John Skilling, and Charles E. Metz. This work was supported in part by the U. S. Department of Energy under contract number W-7405-ENG-36.

REFERENCES

Andrews, H.C. and B.R. Hunt: 1977, *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs.

- Butler, J.P., J.A. Reeds, and S.V. Dawson: 1981, 'Estimating Solutions for First Kind Integral Equations with Nonnegative Constraints and Optimal Smoothing', *SIAM J. Numer. Anal.* **18**, 381–397.
- Gull, S.F.: 1989, 'Developments in Maximum Entropy Data Analysis', in *Maximum Entropy Bayesian Methods*, J. Skilling (ed.), Kluwer, Dordrecht, 53–71.
- Gordon, R., R. Bender, and G. Herman: 1970, 'Algebraic Reconstruction Techniques for Three-Dimensional Electron Microscopy and X-Ray Photography', *J. Theor. Biol.* **29**, 471–481.
- Hanson, K.M.: 1983, 'Variations in Task and the Ideal Observer', *Proc. SPIE* **419**, 60–67.
- Hanson, K.M.: 1987a, 'Bayesian and Related Methods in Image Reconstruction from Incomplete Data', in *Image Recovery: Theory and Application*, H. Stark (ed.), Academic, Orlando, 79–125.
- Hanson, K.M.: 1987b, 'Evaluation of Image-Recovery Algorithms on the Basis of Task Performance', *Proc. 11eme Colloque sur le Traitement du Signal et des Images*, 547–550.
- Hanson, K.M.: 1988a, 'Method to Evaluate Image-Recovery Algorithms Based on Task Performance', *Proc. SPIE* **914**, 336–343.
- Hanson, K.M.: 1988b, 'POPART – Performance OPTimized Algebraic Reconstruction Technique', *Proc. SPIE* **1001**, 318–325.
- Hanson, K.M.: 1989, 'Optimization for Object Localization of the Constrained Algebraic Reconstruction Technique', *Proc. SPIE* **1090**, 146–153.
- Hanson, K.M.: 1990a, 'Method to Evaluate Image-Recovery Algorithms Based on Task Performance', *J. Opt. Soc.* **7A**, 1294–1304.
- Hanson, K.M.: 1990b, 'Object Detection and Amplitude Estimation Based on Maximum *a Posteriori* Reconstructions', *Proc. SPIE* **1231**, 164–175.
- Hanson, K.M.: 1990c, 'Optimization of the Constrained Algebraic Reconstruction Technique for a Variety of Visual Tasks', in *Proc. Information Processing in Medical Imaging*, D. A. Ortendahl and J. Llacer (eds.), Wiley-Liss, New York, 45–57.
- Hunt, B.R.: 1977, 'Bayesian Methods in Nonlinear Digital Image Restoration', *IEEE Trans. Comput.* **C-26**, 219–229.
- Lent, A.: 1977, 'A Convergent Algorithm for Maximum Entropy Image Restoration, with a Medical X-Ray Application', in *Image Analysis and Evaluation*, R. Shaw (ed.), Society of Photographic Scientists and Engineers, New York, 45–57.
- Myers, K.J. and K.M. Hanson: 1990, 'Comparison of the Algebraic Reconstruction Technique with the Maximum Entropy Reconstruction Technique for a Variety of Detection Tasks', *Proc. SPIE* **1231**, 176–187.
- Nashed, M.Z.: 1981, 'Operator-Theoretic and Computational Approaches to Ill-Posed Problems with Applications to Antenna Theory', *IEEE Trans. Antennas Propagat.* **AP-29**, 220–231.
- Skilling, J. and R.K. Bryan: 1984, 'Maximum Entropy Image Reconstruction: General Algorithm', *Mon. Not. R. Ast. Soc.* **211**, 111–124.
- Skilling, J.: 1989, 'Classic Maximum Entropy', in *Maximum Entropy Bayesian Methods*, J. Skilling (ed.), Kluwer, Dordrecht, 45–52.
- Titterton, D.M.: 1985, 'General Structure of Regularization Procedures in Image Reconstruction', *Astron. Astrophys.* **144**, 381–387.