Cross Log Entropy Maximization and Its Application to Ringing Suppression in Image Reconstruction

Yu Cao, Paul P. B. Eggermont, Susan Terebey

Abstract— **We present a multiplicative algorithm for image reconstruction, together with a partial convergence proof. The iterative scheme aims to maximize cross log entropy between modeled and measured data. Its application to *IRAS* data shows reduced ringing around point sources, compared to the EM (Richardson–Lucy) algorithm.**

Keywords— **astronomical image reconstruction, *IRAS*, log entropy, Burg entropy, ringing suppression.**

I. Introduction

FOR many astronomical image reconstruction problems, ringing around bright point sources is a prominent artifact when there is a non-zero background. In Fourier language, the reconstruction process tries to make the image agree with the true scene in the low spatial frequency components (data constraint), without access to the infinitely high spatial frequencies inherent in the point source scene. The result is known as the Gibbs ringing.

Astronomical images are often taken with the intent of making photometric measurements of objects in the field. The ringing artifact hinders the increase of photometric accuracy with smaller aperture, and numerous approaches have been tried in the field of astronomical image reconstruction to overcome the difficulty [3], [13].

While tackling the ringing problem in *IRAS* (*Infrared Astronomical Satellite*, [2]) image reconstruction, it was found using the image space reconstruction algorithm (ISRA, [5]) gives more severe ringing than the Richardson–Lucy algorithm ([14], [12]; known as EM in medical imaging, [16], [11]; "EM" will be used hereafter). In light of the fact that ISRA and EM can be integrated into the same mathematical framework [7], it is natural to ask whether there is an iterative scheme in the same family, which gives even less ringing than EM.

The following sections present such an algorithm, together with a partial convergence proof and application to *IRAS* data. The algorithm is a multiplicative one that aims to maximize the cross log-entropy between observed and modeled data.

Y. Cao is with Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, California 91125, USA. E-mail: yucao@srl.caltech.edu.

P. P. B. Eggermont is with Department of Mathematical Sciences, University of Delaware, Newark, Delaware 19716, USA.

S. Terebey is with Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, California 91125, USA.

II. Algorithm

Consider the system of equations $Ax = b$, where A is a non-negative $n \times m$ matrix, and $b \in R^n$ is the noisy non-negative data vector. If the modeling and the data were exact, the system would have a non-negative solution. In light of the noisy data and the modeling, the sense in which the system $Ax = b$ is to be solved is that of minimum negative "cross log entropy" (negative cross Burg entropy, or Itakura–Saito distance), i.e. by solving

$$\begin{aligned} &\text{minimize} \quad L(x) \stackrel{\text{def}}{=} \sum_{i=1}^{n} -\log \frac{b_i}{[Ax]_i} + \frac{b_i}{[Ax]_i} \\ &\text{subject to} \quad x \geq 0. \end{aligned} \tag{1}$$

The first thing to note is that if x^* satisfies $Ax^* = b$, then $x = x^*$ solves the problem. The second thing to note is that $L(x)$ is not a convex functional of x, since the logarithmic term is concave. However, it is strictly convex on the set

$$C = \{x \in R^m : [Ax]_i \leq 2b_i, i = 1, 2, \ldots, m\}, \tag{2}$$

since $-\log(a/t) + a/t$ is convex on $0 < t \leq 2a$, for $a > 0$, which is non-empty.

We look at two related algorithms

Algorithm I.

$$\begin{aligned} x_j^{k+1} &= x_j^k \sqrt{\frac{[A^T p^k]_j}{[A^T q^k]_j}}, \\ j &= 1, 2, \ldots, m, \end{aligned} \tag{3}$$

and

Algorithm II.

$$\begin{aligned} x_j^{k+1} &= x_j^k \cdot \frac{[A^T p^k]_j}{[A^T q^k]_j}, \\ j &= 1, 2, \ldots, m, \end{aligned} \tag{4}$$

where

$$\begin{aligned} p_i^k &= \frac{b_i}{([Ax^k]_i)^2}, \\ q_i^k &= \frac{1}{[Ax^k]_i}, \\ i &= 1, 2, \ldots, n. \end{aligned} \tag{5}$$

If x^1 has positive components, then all the future x^k have all positive components as well. So expressions like x_j^{k+1}/x_j^k are always meaningful. We denote the vector with components x_j^{k+1}/x_j^k simply by x^{k+1}/x^k.

Cross log entropy maximization has been suggested by Herman et al. [10] in tomography, together with the algorithm

$$\begin{aligned} x_j^{k+1} &= x_j^k\{1-\alpha([A^T q^k]_j + [A^T p^k]_j)\}, \\ j &= 1,2,\ldots,m, \end{aligned} \tag{6}$$

where α is a relaxation parameter with the same dimension as x_j. The above algorithm's convergence is still a subject of study.

A. Derivation

As a first step in deriving the algorithms we show LEMMA 1.

$$L(y)-L(x) \geq \sum_{j=1}^{m}\left\{[A^T q]_j - [A^T p]_j \frac{y_j}{x_j}\right\}(y_j - x_j), \tag{7}$$

where

$$\begin{aligned} p_i &= \frac{b_i}{([Ay]_i)^2}, \\ q_i &= \frac{1}{[Ay]_i}, \\ i &= 1,2,\ldots,n. \end{aligned} \tag{8}$$

for any vectors x and y with positive components.

PROOF. First of all we note that by the strict concavity of the logarithm

$$\begin{aligned} -\log\frac{b_i}{[Ay]_i} + \log\frac{b_i}{[Ax]_i} &= \log{[Ay]_i} - \log{[Ax]_i} \\ &\geq \frac{[Ay]_i - [Ax]_i}{[Ay]_i}, \end{aligned} \tag{9}$$

and so

$$\begin{aligned} &\sum_{i=1}^{n} -\log\frac{b_i}{[Ay]_i} + \log\frac{b_i}{[Ax]_i} \\ \geq &\sum_{i=1}^{n} \frac{[A(y-x)]_i}{[Ay]_i} \\ = &\sum_{j=1}^{m}[A^T q^k]_j(y_j - x_j). \end{aligned} \tag{10}$$

Also, there is equality here if and only if $Ax = Ay$. Secondly, we write

$$\frac{b_i}{[Ay]_i} - \frac{b_i}{[Ax]_i} = \frac{b_i}{[Ay]_i}\left(1 - \frac{[Ay]_i}{[Ax]_i}\right). \tag{11}$$

Now observe that

$$\frac{[Ay]_i}{[Ax]_i} = \Psi\left(\frac{[Ax]_i}{[Ay]_i}\right) = \Psi\left(\frac{[A\{y(x/y)\}]_i}{[Ay]_i}\right), \tag{12}$$

where $\Psi(t) = 1/t$. Now $\Psi(t)$ is strictly convex for $t > 0$, and so by Jensen's inequality

$$\Psi\left(\frac{[Ay(x/y)]_i}{[Ay]_i}\right) \leq \frac{[A\{y\Psi(x/y)\}]_i}{[Ay]_i}. \tag{13}$$

with equality if and only if x_j/y_j is independent of j, for those indices j for which $a_{ij} > 0$. Consequently

$$\begin{aligned} \sum_{i=1}^{n}\frac{b_i}{[Ay]_i} - \frac{b_i}{[Ax]_i} &\geq \sum_{i=1}^{n}\frac{b_i}{[Ay]_i}\left(1 - \frac{A\{(y)^2/x\}}{[Ay]_i}\right) \\ &= \sum_{i=1}^{n}\frac{b_i}{([Ay]_i)^2}[A(y-(y)^2/x)]_i \\ &= \sum_{j=1}^{m}[A^T p]_j\frac{y_j}{x_j}(x_j - y_j). \end{aligned} \tag{14}$$

with equality if and only if x/y is the constant vector. Adding (10) and (14) gives

$$L(y)-L(x) \geq \sum_{j=1}^{m}\left\{[A^T q]_j - [A^T p]_j\frac{y_j}{x_j}\right\}(y_j - x_j). \tag{15}$$

Q.E.D.

LEMMA 1 can be rewritten as

$$L(x) \leq L(y) + \sum_{j=1}^{m}[A^T p]_j\left\{\frac{(y_j)^2}{x_j} - y_j\right\} - [A^T q]_j(y_j - x_j). \tag{16}$$

Since we want to minimize the left side of (16), but are unable to do this directly, let us minimize the right hand side of (16). Assuming that we can do this by differentiating the right hand side of (16) with respect to x_j, and setting this equal to 0 gives

$$\begin{aligned} -[A^T p]_j\frac{(y_j)^2}{(x_j)^2} + [A^T q]_j = 0, \\ j = 1,2,\ldots,m, \end{aligned} \tag{17}$$

and so (16) suggests a way of achieving a new estimate of vector x from y

$$\begin{aligned} x_j &= y_j\sqrt{\frac{[A^T p]_j}{[A^T q]_j}}, \\ j &= 1,2,\ldots,m. \end{aligned} \tag{18}$$

For this reason we call inequalities like (16) "tendentious inequalities" because they suggest algorithms, and much more. De Pierro [8] calls this approach "majorizing functions algorithm."

An "accelerated" version of (18) can be derived if we choose x_j such that the second term in (16) equals zero

$$\begin{aligned} -[A^T p]_j\frac{y_j}{x_j} + [A^T q]_j = 0, \\ j = 1,2,\ldots,m, \end{aligned} \tag{19}$$

which leads to

$$\begin{aligned} x_j &= y_j \cdot \frac{[A^T p]_j}{[A^T q]_j}, \\ j &= 1,2,\ldots,m. \end{aligned} \tag{20}$$

Applying (18) or (20) iteratively gives rise to ALGORITHM I or ALGORITHM II.

B. Monotonicity Property

For ALGORITHM I, we have the following monotonicity property from (16)

$$\begin{aligned}
& L(x^{k+1}) - L(x^k) \\
\le\ & \sum_{j=1}^{m} 2x_j^k \sqrt{[A^T q^k]_j [A^T p^k]_j} - x_j^k [A^T p^k]_j - x_j^k [A^T q^k]_j = \\
\le\ & -\sum_{j=1}^{m} x_j^k \left| \sqrt{[A^T p^k]_j} - \sqrt{[A^T q^k]_j} \right|^2 = \\
\le\ & -\sum_{j=1}^{m} [A^T q^k]_j \frac{\left| x_j^{k+1} - x_j^k \right|^2}{x_j^k},
\end{aligned} \tag{21}$$

or

$$L(x^k) - L(x^{k+1}) \ge \sum_{j=1}^{m} [A^T q^k]_j \frac{\left| x_j^{k+1} - x_j^k \right|^2}{x_j^k}. \tag{22}$$

It shows that $L(x^k) > L(x^{k+1})$, unless $x^k = x^{k+1}$, and that if x^* solves (1), then it is a fixed point of ALGORITHM I.

For ALGORITHM II, it is obvious from the derivation that

$$L(x^k) \ge L(x^{k+1}) \tag{23}$$

i.e. the negative log entropy never increases.

III. Application to *IRAS* Data

A. Relevant Information About IRAS

The *IRAS* survey was designed for the identification of point sources, rather than as an imaging instrument. The data were taken with rectangular detectors that scanned the sky multiple times in "push-broom" fashion (e.g., see Fig. 2). The satellite data are fundamentally in the form of one-dimensional data streams for each detector, therefore for any given field the data points do not lie on a regular grid and the point spread function is not translation invariant.

The *IRAS* focal plane (shown in Fig. 1) included eight staggered linear arrays subtending $30'$ in width, two in each of four spectral bands at 12, 25, 60, and 100 μm. Data rate considerations forced the detector sizes to be much larger than the diffraction limit of the telescope. The typical detector sizes were 45×267, 45×279, 90×285, and 180×303 arcsec (full width at half maximum response, FWHM) respectively, at the four wavelength bands.

This combination of focal plane, detector size, and scan pattern optimized detection of point sources in areas of the sky where the separation between sources was large compared to the sizes of the detectors. Image reconstruction techniques were later applied to obtain images with resolution $\sim 1'$ from the *IRAS* data [1], [3], [4].

B. Convergence Speed

From test runs using *IRAS* data, it was found ALGORITHM I and ALGORITHM II give quantitatively similar images, with ALGORITHM II converging roughly twice as fast as ALGORITHM I (Table I). This can be explained by the fact that corrections are small except in the first few iterations.

Besides ALGORITHM II's faster convergence speed, it also gives better photometric integrity in the first few iterations compared to ALGORITHM I. The first iteration result from ALGORITHM II is just the usual coadded image, identical to that from EM's first iteration, assuming the zeroth image is flat in both cases. ALGORITHM I however, because of the square root corrections employed, gives images that have absolute scales dependent on the magnitude of the zeroth image (although the effect is washed out quadratically in the later iterations).

These considerations led us to prefer ALGORITHM II in the *IRAS* application, and we restrict our discussion to ALGORITHM II in the following sections (the result from ALGORITHM I being similar at twice the number of iterations anyway).

C. Ringing Suppression

The algorithm was tested on several fields of the *IRAS* data. For all cases, the resulted images showed weaker ringing around point sources than images made with the EM algorithm.

Fig. 3 shows a comparison of reconstructed images using EM, ISRA, and the log entropy algorithm (ALGORITHM II), all at 20 iterations, plus a co-added image. The co-added image in Fig. 3(a) is a simple average of detector fluxes weighted by the response function, equivalent to the first iteration image from EM, ISRA, and ALGORITHM II. The reconstructed images ((b), (c), and (d)) show enhanced resolution compared to the co-added image. The field captures one of the brightest stars in the sky, α Ori, [1] at 60 μm. The arc to the top-left of the star is a bow shock caused by the motion of the star in the interstellar medium. In the EM and ISRA images ((b) and (c)), the severe ringing artifact disturbs the shape of the bow shock (the spurious ring is slightly brighter than the bow shock). It is also apparent that ISRA resulted in more severe ringing than EM. The image from the new log entropy algorithm ((d)) shows great improvement, effectively suppresses the ringing, reconstructs the bow shock cleanly, and even recovers a hint of the diffraction spikes around the bright star. The log entropy image also gives a sharper profile of the star than the EM image at the same number of iterations, with a peak intensity (star) 1.4 times that of the EM image. The maximum pixel intensity of the bow shock (~ 25 MJy/ster) is about 1/150 of that of the star (3596 MJy/ster in the log entropy image).

D. Validation Using Simulated Data

To validate the authenticity of the bow shock structure and the result of reduced ringing, simulated data were constructed using the actual *IRAS* scan pattern and detector response functions, and taking the log entropy image in Fig. 3(d) as the sky brightness. The simulated data were

[1] α Orionis, also called Betelgeuse, is the second brightest star in the constellation Orion and marks the eastern shoulder of the hunter. It is a red supergiant star approximately 600 light-years from the Earth.

then processed with the EM and the log entropy algorithm respectively.

The resulted images are shown in Fig. 4. Again, the EM image showed ringing similar to the corresponding image in Fig. 3. The log entropy image is similar to the input from which the simulated data were made (Fig. 3(d)), indicating the log entropy algorithm is more consistent with our preference [2] for this test case.

Also, the fact that the output images made from simulated data are similar to the ones made from real data indicates the two sets of data contain comparable amount of high spatial frequency signal, i.e. the input image that went into the simulated data can be considered as well resolved.

E. Discussion

We use a set of notations more familiar to the field of astronomical image reconstruction in this section (see Table II).

In this set of notations, ALGORITHM II takes the form of

$$f_j^{k+1} = f_j^k \cdot \frac{\sum_i r_{ij} D_i / F_i^2}{\sum_i r_{ij} / F_i}. \tag{24}$$

The ringing artifact can be seen as being caused by the propagation of data misfit at the point source. The new algorithm attenuates the propagation by the $1/F_i$ weighting factor, compared to EM:

$$f_j^{k+1} = f_j^k \cdot \frac{\sum_i r_{ij} D_i / F_i}{\sum_i r_{ij}}. \tag{25}$$

The $1/F_i$ weighting results in better determination of the background, as the data samples less affected by the point source cast more influence on the background intensities.

In the reconstructed image, the point source has a finite width profile due to the finite resolution achieved, while in the true scene the point source mimics a delta function. This causes the modeled data F_i to have a longer "tail" than the measured data D_i. While trying to compensate for the misfit ($F_i \gg D_i$), the correction factors push down the pixels around the point source, giving the first dip in the ripples. The dip then causes misfit further away from the point source, which in turn results in the bright ring, so on and so forth.

The reduction in ringing by the log entropy algorithm can be traced to the concavity of L at large F_i. Similar to M-estimates in robust estimation (e.g., see [15]), the derivative of L decreases in absolute value as F_i increases (at fixed D_i), i.e. the loss function curve flattens out at large values of F_i. Therefore the data points that have F_i mixing the blurred point source and the background ($F_i \gg D_i$) are seen as "outliers", and the corrections incurred on the image pixels by these data points are weighted down.

Fig. 5 shows a plot of the negative log entropy and the negative log Poisson likelihood as a function of F_i. The negative log Poisson likelihood is convex, while the negative log entropy is concave for large values of F_i.

Minimizing the negative log entropy functional corresponds to maximum likelihood estimation from Gamma distributions. The Gamma likelihood has a "fatter tail" than the Poisson or Gaussian, making the log entropy algorithm more tolerant of bright point source scenes than the EM and ISRA, which are suitable for maximum Poisson likelihood and least squares estimates [6], [9]. The choice of Gamma likelihood (which is invariant under scaling of D_i and F_i), has some justification for astronomical scenes, where it is reasonable to regard two similar combinations of model and data on different absolute scales as equally likely *a priori*. It appears natural to use the Gamma likelihood when the misfit between model and data is dominated by the mixing of signal on different magnitude scales in the model (like the bright point source case), instead of photon counting statistics (Poisson) or read-out noise (Gaussian).

Quantifying the ringing magnitude (and the reduction from EM to the log entropy algorithm) is not a trivial task, as the ringing depends on such parameters as the background intensity, the point source strength, and the position of the data samples. First we look at the asymptotic behavior of minimum modeled data (F_i) when the point source strength is large, assuming the point source sits on a flat background of fixed intensity.

Assume an image pixel j is covered by only two data samples, 1 and 2. Both are centered far enough from the point source, so that

$$D_1 = D_2 = f_B \sum_j r_{1j} \stackrel{\text{def}}{=} B, \tag{26}$$

where f_B is the background intensity.

We then look at the dependence of F_1 and F_2 on each other when the iteration has proceeded near convergence.

Under the EM scheme, we would then have

$$r_{1j} D_1 / F_1 + r_{2j} D_2 / F_2 = r_{1j} + r_{2j}. \tag{27}$$

While if the log entropy algorithm was used, we would have

$$r_{1j} D_1 / F_1^2 + r_{2j} D_2 / F_2^2 = r_{1j} / F_1 + r_{2j} / F_2. \tag{28}$$

A schematic comparison of (27) and (28) is shown in Fig. 6 (r_{1j} and r_{2j} were assumed to be equal in the plots).

We further assume sample 1 lies closer to the point source than 2, and sample 1 covers part of the side lobe of the point source in the reconstructed image. When the point source is very bright, $F_1 \gg B$, and as can be seen from Fig. 6, F_2 approaches the correct background flux B with the log entropy algorithm, but only a fraction of B with EM (reflecting the dip around the point source). We take F_2 as an estimate of the minimum F_i.

The different characteristic behavior of minimum F_i was replicated in simulations, using images reconstructed from synthetic data. A point source with varying strength was planted onto a constant background (30 MJy/ster), and the

[2] It should be noted that the data alone do not favor one output image or the other. Reduced ringing is preferred due to our prior knowledge that point sources do not usually come with rings.

actual *IRAS* scan pattern was run through the artificial image, and a set of synthetic data was generated. After that EM and the log entropy algorithm were used to process the data separately. Table IV shows the comparison of minimum modeled data F_i after 100 iterations. For EM, the minimum F_i keeps decreasing with increasing point source strength. And for the log entropy algorithm, the minimum F_i first decreases, then climbs back towards B as the point source strength increases, showing good agreement with the curve shown in Fig. 6.

In these simulations it was found the rings from the log entropy algorithm have smaller sizes than those from EM (which are in turn smaller than those from ISRA). Also, the peak intensity of the reconstructed star is the highest with the log entropy algorithm. These observations are consistent with the fact that the log entropy algorithm requires more high spatial frequency power in the image than EM and ISRA.

IV. Summary

Unlike some other ringing suppression schemes, the log entropy algorithm does not require extra prior information (such as point source position and/or strength) as input, and does not require the fine tuning of parameters. It is also structurally similar to EM and ISRA, making it very easy to incorporate in existing image reconstruction software. These advantages make it likely to be applied to a wide range of problems where ringing artifact is a concern.

Acknowledgments

We thank John Fowler for being a great source of advice, both about ringing suppression and *IRAS* image reconstruction in general. We're also indebted to Dave Van Buren for providing the α Ori data which proved to be a perfect test case for the new algorithm. YC thanks Tom Prince for his guidance. This research received support from the NASA Astrophysics Data Program under contract No. NAS5-32642.

References

[1] H. H. Aumann, J. W. Fowler, and M. Melnyk, "A maximum correlation method for image construction of *IRAS* survey data," *Astron. J.*, vol. 99, no. 5, pp. 1674–1681, 1990.

[2] C. A. Beichman, G. Neugebauer, H. J. Habing, P. E. Clegg, and T. J. Chester, eds. *IRAS Catalogs and Atlases: Explanatory Supplement*, Washington, DC: GPO, 1988.

[3] T. R. Bontekoe, E. Koper, and D. J. M. Kester, "Pyramid maximum-entropy images of *IRAS* survey data," *Astron. & Astrophys.*, vol. 284, no. 3, pp. 1037–1053, Apr. 1994.

[4] Y. Cao, T. A. Prince, S. Terebey, and C. A. Beichman, "Parallelization and algorithmic enhancements of high-resolution *IRAS* image construction," *Pub. Astron. Soc. Pacific*, vol. 108, no. 724, pp. 535–544, Jun. 1996.

[5] M. E. Daube-Witherspoon and G. Muehllehner, "An iterative image space reconstruction algorithm suitable for volume ECT," *IEEE Trans. Med. Imaging*, vol. 5, no. 2, pp. 61–66, 1986.

[6] A. R. De Pierro, "On the convergence of the iterative image space reconstruction algorithm for volume ECT," *IEEE Trans. Med. Imaging*, vol. 6, no. 2, pp. 174–175, 1987.

[7] A. R. De Pierro, "On the relation between the ISRA and the EM algorithm for positron emission tomography," *IEEE Trans. Med. Imaging*, vol. 12, no. 2, pp. 328–333, Jun. 1993.

[8] A. R. De Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," *IEEE Trans. Med. Imaging*, vol. 14, no. 1, pp. 132–137, Mar. 1995.

[9] P. P. B. Eggermont, "Multiplicative iterative algorithms for convex-programming," *Linear Algebra Appl.*, vol. 130, pp. 25–42, Mar. 1990.

[10] G. T. Herman, M. Chan, Y. Censor, E. Levitan, R. M. Lewitt, and T. K. Narayan, "Maximum a posteriori image reconstruction from projections," in *Image Models (and Their Speech Model Cousins)*, S. Levinson and L. A. Shepp, eds. New York, NY: Springer-Verlag, 1996.

[11] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," *J. Comput. Assisted Tomography*, vol. 8, pp. 306-316, 1984.

[12] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astron. J.*, vol. 79, pp. 745–754, 1974.

[13] L. B. Lucy, "Image restorations of high photometric quality," in *The Restoration of HST Images and Spectra – II*, R. J.Hanisch and R. White, eds. pp. 79–85, Baltimore, MD: STScI, 1994.

[14] W. H. Richardson, "Bayesian-based iterative method of image restoration," *J. Opt. Soc. Am.*, vol. 62, pp. 55–59, 1972.

[15] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. pp. 699–702, Cambridge: Cambridge U. Press, 1992.

[16] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction in positron emission tomography," *IEEE Trans. Med. Imaging*, vol. 1, pp. 113–122, 1982.

TABLE I

CONVERGENCE SPEED COMPARISON

ALGORITHM I		ALGORITHM II	
Iter.	$L-n$	Iter.	$L-n$
10	201.54	5	205.97
20	191.15	10	192.13
30	187.42	15	187.93
40	185.12	20	185.48
50	183.39	25	183.68

TABLE II

NOTATIONS COMMONLY USED IN ASTRONOMICAL IMAGE RECONSTRUCTION

Quantity	Notation Used in Tomography	Notation Used in Astronomy
image pixel intensity	x_j	f_j
response function	A_{ij}	r_{ij}
flux of measured data	b_i	D_i
flux of modeled data	$[Ax]_i$	$F_i = \sum_j r_{ij} f_j$

TABLE IV

COMPARISON OF MINIMUM MODELED DATA. UNIT OF F_i: $10^{-13}\mathrm{W/cm^2}$. BACKGROUND FLUX $B = 4.76 \times 10^{-13}\mathrm{W/cm^2}$.

Point Source Magnitude (Jy)	Min. F_i with Em	Min. F_i with Log Entropy
10	4.72	4.73
10^2	4.51	4.59
10^3	4.01	4.37
10^4	2.96	4.38
10^5	2.41	4.54

TABLE III

Comparison of Different Loss Functions

Algorithm	Function L	$\partial L/\partial F_i$	$\partial^2 L/\partial F_i^2$
ISRA	$(D_i - F_i)^2$	$2(F_i - D_i)$	$2 > 0$
EM	$-(D_i - F_i) + D_i \log D_i/F_i$	$1 - D_i/F_i$	$D_i/F_i^2 > 0$
Log Entropy	$-\log D_i/F_i + D_i/F_i$	$1/F_i - D_i/F_i^2$	$-1/F_i^2 + 2D_i/F_i^3 < 0$ (F_i large)

TABLE V

Comparison of Maximum Pixel Intensity

Point Source Magnitude (Jy)	Max. f_j with EM (MJy/ster)	Max. f_j with Log Entropy	Ratio
10	4.63×10	4.91×10	1.06
10^2	7.16×10^2	1.08×10^3	1.51
10^3	1.26×10^4	3.01×10^4	2.39
10^4	1.63×10^5	7.39×10^5	4.54
10^5	1.73×10^6	8.12×10^6	4.70

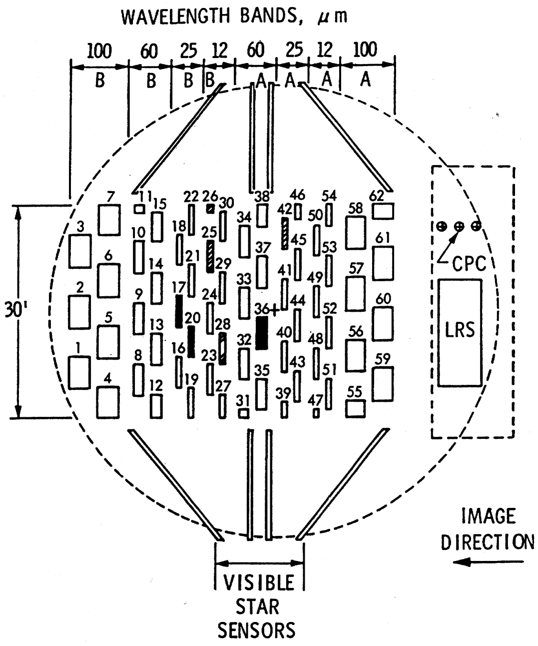

Fig. 1. The *IRAS* Focal Plane. The numbered rectangles in the central portion each represent the field of view of a detector, filter and field lens combination. The filled-in detectors were inoperative while the cross-hatched detectors showed degraded performance during the mission. (Adapted from Figure II.C.6, [2].)

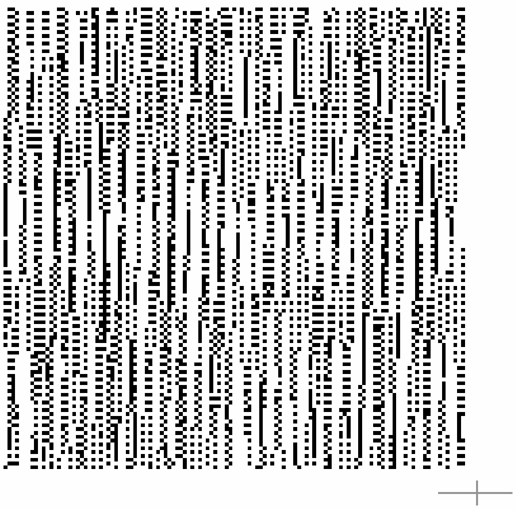

Fig. 2. *IRAS* Scan Pattern Near α Ori at 60 μm. Each black dot represents the center point of a data sample. The cross at the lower-right corner shows the typical FWHM of the detector response function. Angular scale is 0.5° on each side.

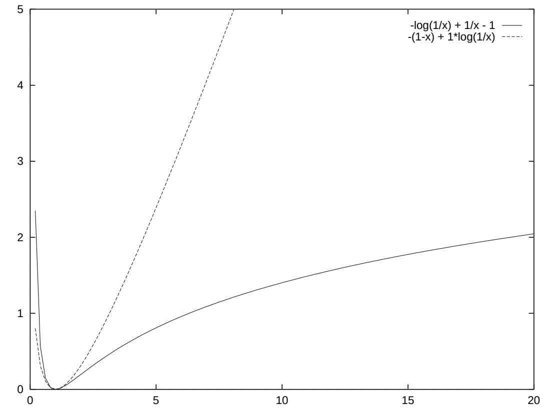

Fig. 5. Comparison of the Negative Log Entropy and the Negative Log Poission Likelihood.

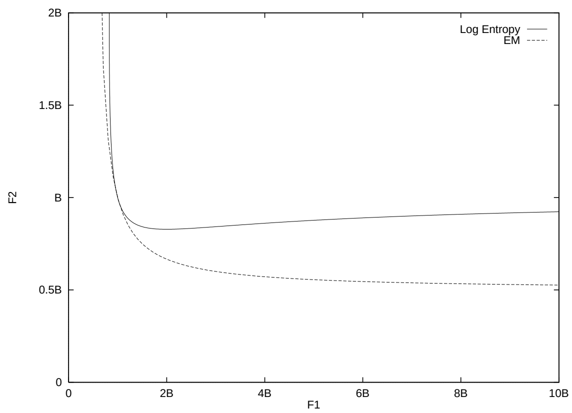

Fig. 6. Comparison of the Dependence of Modeled Data. Solid curve: log entropy case, F_2 approaches B when F_1 is large; Dashed curve: EM, F_2 approaches $B/2$.

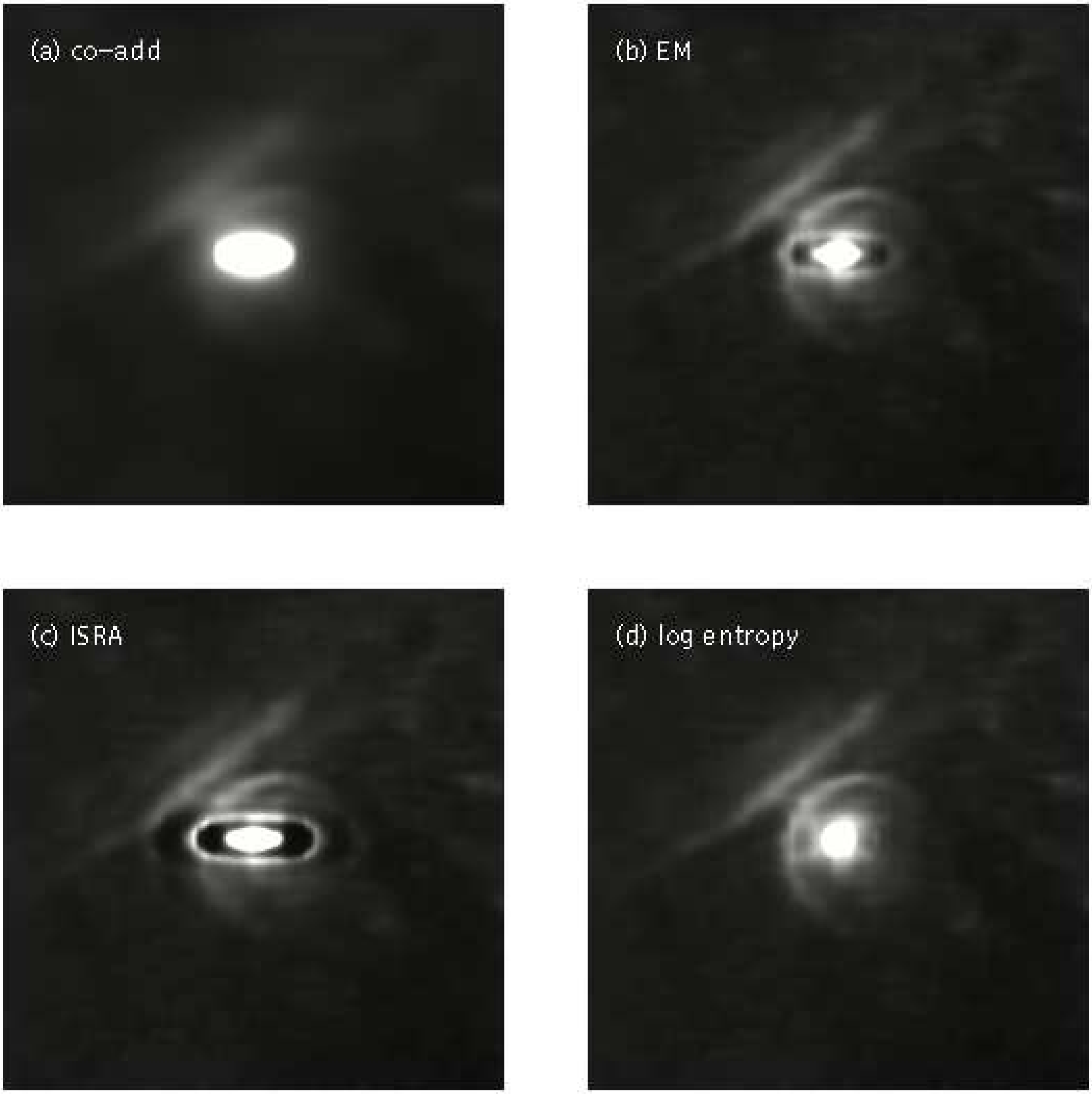

Fig. 3. Comparison of Reconstructed Images Using EM, ISRA, and the Log Entropy Algorithm. (a): co-add (or weighted average) of detector fluxes; (b), (c): EM and ISRA, severe ringing around the bright star is present, and the arc (bow shock) to the top-left of the star is disrupted; (d): the log entropy image suppresses the ringing and reconstructs the bow shock cleanly. Angular scale is 1° on each side. The brightest pixel in the bow shock has intensity $\sim 1/150$ of the peak intensity of the star. The spurious ring in the EM image is slightly brighter than the bow shock.

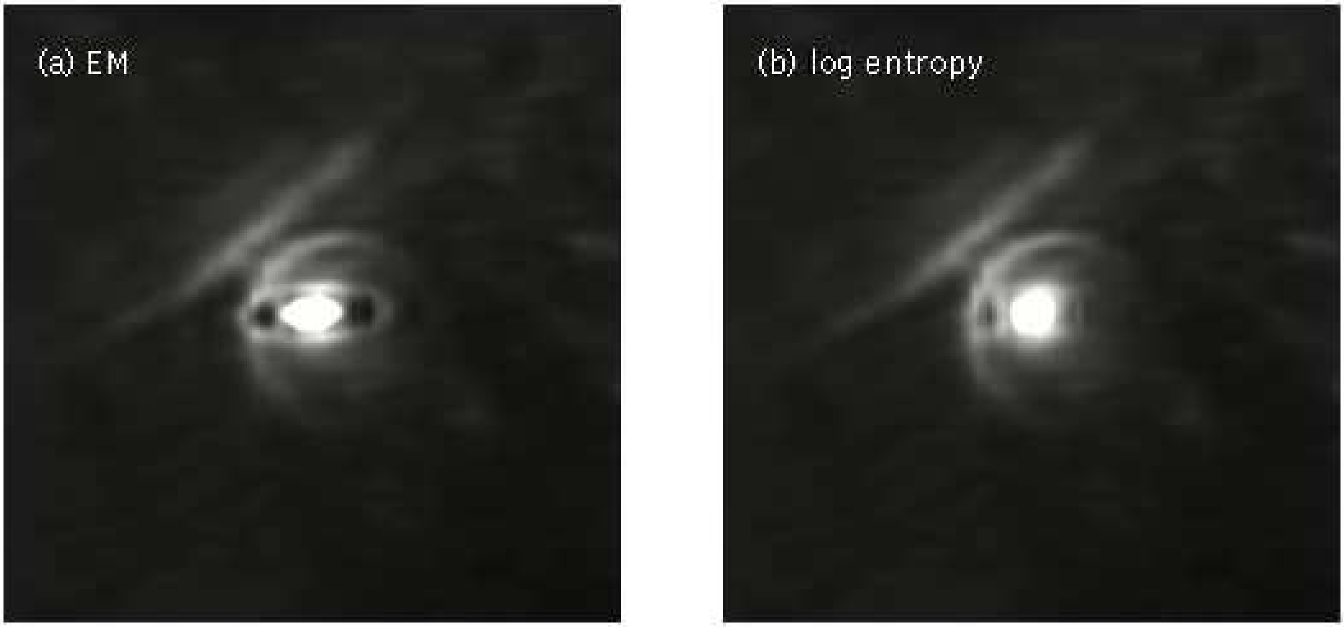

Fig. 4. Comparison of Reconstructed Images From Simulated Data. (a): EM reconstruction; (b): log entropy reconstruction. The results verify the consistency of the log entropy algorithm.