J. Astrophys. Astr. (1986) 7, 275–292

Minimum-Relative-Entropy Method—Solution to Missing Short-Baseline Problem

R. K. Shevgaonkar *Indian Institute of Astrophysics, Bangalore 560034*

Received 1986 March 7; accepted 1986 August 6

Abstract. Minimum-relative-entropy method (MREM) has been presented as a solution to the missing short-baseline problem in the synthesis observations. It is shown that a measure of distance between the prior model and the image in the plane of pixel brightness is an adequate measure of relative entropy. The method has been further extended for polarization observations and the potential of the method against the conventional MEM has been demonstrated by simulated examples.

Key words: image reconstruction—minimum-relative-entropy method—radio astronomy, reduction techniques

1. Introduction

In radio astronomy, looking at the need of large-field mapping, the maximum-entropy method (MEM) has received a new impetus over the last decade (Gull & Daniell 1978; Cornwell & Evans 1985). Application of the method to a variety of images in astronomy as well as in other fields like picture processing and medical science has clearly demonstrated the potential of the technique. However, one must accept the fact that to make the method work the problem should be well-posed. In practice one observes that it becomes difficult to obtain useful reconstruction of an image if the measurements are too noisy or too sparse. For example, we see that the extended sources are not very faithfully reconstructed if the spatial Fourier coverage (also called *uv*-coverage) is not compact (Nityananda & Narayan 1982; Narayan & Nityananda 1984). Here, by compact we mean a *uv*-coverage which is uniformly sampled with the Nyquist rate around the origin of the *uv*-plane.

In the past, the maximum-entropy-method has been presented as a general technique devised for estimating the unmeasured visibility coefficients from the knowledge of the measured ones (Nityananda & Narayan 1982; Shevgaonkar 1986b). The method does not make any assumption regarding the spatial domains of the measured and unmeasured visibilities. It is assumed that the measurements are performed over an arbitrary spatial region and we are interested in predicting the visibility coefficients over another region using entropy of the image as a measure of goodness of the image. However, if we look at the visibility function carefully, we note that the method cannot provide faithful reconstruction for any arbitrary kind of data. For an observation where the visibility coefficients are measured over a compact *uv*-coverage, all extended as well as localized sources have their respective contribution in the measured data. On the contrary, if the observations do not contain shorter baselines, the very extended sources are completely missing in the observed data. The extended distributions, which are neither totally missing nor completely sampled, are mainly the source of degradation of the observed image. The presence of improperly sampled largescale structures is manifested in large negative bowls in the observed image. The very extended features whose visibility function lies completely within the unmeasured central region of the *uv*-plane, do not give any indication of their existence and one would never know what is actually missing from the observed data. Here we are not interested in reconstruction of these largescale structures, as their presence or absence does not make any difference in the quality of the reconstruction. In this paper we are concerned about those extended distributions whose visibility function is partly sampled and which manifest themselves in the negative bowls. The MEM, as presented in the past, due to its inherent nature of reconstructing flat background and sharp peaks, although a useful technique for improving resolution, faces serious difficulties in reconstructing extended sources observed with non-compact *uv*-coverages.

Aperture synthesis telescopes, while generally providing a uniform sampling of the visibility function, often leave an unsampled hole near the origin of the *uv*-plane. The zero-lag visibility coefficient or the integrated power of the brightness distribution can be obtained (although not very easily) from any of the synthesizing elements, but the other short spacings still remain unmeasured. A method for estimating the visibility function over short baselines is highly desired from the view-point of mapping largescale structures with synthesis telescopes.

Although the conventional deconvolution method CLEAN (Högbom 1974) is capable of removing the effect of sidelobes from the observed maps, it finds difficult to take away the negative bowls around the extended sources. A modified version of CLEAN called 'window CLEAN' (Schwarz 1978), by putting a CLEANing box around the positive source within the negative bowl, may relatively improve the quality of the reconstruction by forcibly discarding the presence of the extended negative features. However, practitioners of CLEAN would agree that the method does not do a very good job for reconstructing largescale structures.

In the past, few proposals have been made towards estimation of the short-baseline visibilities from synthesis data (Rots 1979; Ekers & Rots 1979; Braun & Walterbos 1985). The method suggested by Rots (1979) is a data fitting technique using the positivity constraint on the image. This technique requires a precise knowledge of the source parameters which could possibly be available only for simple distributions, and therefore the method is not quite suited for complex distributions. The method presented by Braun & Walterbos (1985) is based upon a direct nonlinear fit of the missing Fourier coefficients to the isolated map plane response. In any case, to obtain an equally good quality image of the compact sources along with the extended ones, the map ultimately has to be deconvolved with the point spread function. Therefore it appears that the most efficient way would be to obtain a proper deconvolution method which itself is capable of estimating lower as well as higher spatial frequency visibilities.

We present here a possible application of the minimum-relative-entropy method for the estimation of the visibilities over short baselines. We show that the distance between the image and the prior model in the plane of pixel brightness can provide a good measure of relative entropy *i.e.*, entropy of the image with respect to the entropy of the prior model. We further generalize this definition of relative entropy for the polarization images and by simulated examples we show that the method works very promisingly for complex extended images observed with limited information over short baselines.

2. Concept of relative entropy

The concept of relative entropy (also called 'cross entropy') was first proposed by Kullback in 1959. The concept has been subsequently promoted in various forms by others (Hobson & Cheng 1973; Johnson 1979) and has been successfully utilized by radio astronomers (Gull & Skilling 1984; Cornwell & Evans 1985) to obtain reliable deconvolution of real synthesized images. The relative entropy is a generalization of the entropy that applies in cases where a prior image B_0 that estimates the image B is known in addition to the measurements and other constraints like positivity.

Shore & Johnson (1980), using consistency arguments, have shown that in the presence of prior knowledge about the distribution one must minimize the relative entropy defined by a unique function given as

$$E_r = \iint B \ln (B/B_0) \, dx \, dy \tag{1}$$

where x, y are the image coordinates. On the other hand, Cornwell & Evans (1985) treat relative entropy as a measure of distance of reconstructed image B from an *a priori* expected image B_0. Following this concept it is possible to obtain a variety of functions which will define an equally good measure of distance between B and B_0. However, a function $F(B, B_0)$ which defines a good measure of distance between two images B and B_0, and can be used as a relative-entropy function must possess certain basic characteristics.

(1) Firstly, since the true brightness distribution is positive definite, the entropy function should impose positivity on the reconstructed image. In other words the entropy function should explicitly be definable for positive brightness only.

(2) Secondly, as has been argued in the past (Nityananda & Narayan 1982), for any kind of extrapolation of the visibility function and for obtaining a translation-invariant reconstruction, the entropy function should provide a nonlinear image transfer function, or in other words the first derivative of the function should be a nonlinear function of the brightness.

(3) Thirdly, it is desired that the entropy function should be such that one obtains a unique reconstruction for a given measurement and a given biasing image *i.e.*, the entropy function should possess a single minimum in the acceptable range of brightness.

It is immediately clear that the scalar distance between two images defined by the sum of the squares of the pixel differences cannot be taken as a measure of relative entropy, as it obtains only a linear-image-transfer function which, as pointed out above, is not adequate for extrapolation of the visibility function. Also, the scalar distance minimization does not impose the positivity on the reconstructed image.

As a next immediate choice the scalar distance between some function of the pixel brightness of the image and the prior model as given below

$$F(B, B_0) = [f(B) - f(B_0)]^2 \tag{2}$$

can be tried for defining relative entropy. The choice of f should be such that F satisfies the three above-mentioned conditions. The first two conditions can be easily achieved by choosing f as one of the simple entropy functions like $\ln B$, $-B \ln B$, or in general B^s. The third, minimality condition requires that the second derivative of $F(B, B_0)$ with respect to B must be greater than zero for all positive values of B giving

$$\frac{\partial^2}{\partial B^2}\{[f(B)-f(B_0)]^2\} > 0 \tag{3a}$$

or

$$\{f(B)-f(B_0)\}\frac{\partial^2 f}{\partial B^2}+\left(\frac{\partial f}{\partial B}\right)^2 > 0. \tag{3b}$$

Now since all conventional entropy functions f have $\partial^2 f/\partial B^2 < 0$ (Burg 1975; Nityananda & Narayan 1982), Equation (3b) can be rewritten as

$$f(B)+\frac{(\partial f/\partial B)^2}{\partial^2 f/\partial B^2} < f(B_0). \tag{4}$$

Therefore for a given prior model a unique reconstruction is possible only if Equation (4) is satisfied for all possible values of B.

Let us now verify which of the conventional entropy functions satisfy Equation (4)!

For entropy function $f(B) = \ln B$ the uniqueness condition (4) reduces to

$$\ln B - 1 < \ln B_0 \tag{5a}$$

or

$$B < eB_0 \tag{5b}$$

where e is the exponential constant. Equation (5b) indicates that only those points which have $B < eB_0$ in the image will be biased towards B_0 whereas the others will not. In other words this means that the *a priori* known model distribution should not be too different from the observed distribution. This is quite an undesirable requirement of the entropy function $\ln B$. On many occasions when the prior model is not available, a choice of uniform B_0 will violate condition (5b) at many locations in the image and one will not get a good reconstruction. The relative entropy using $f(B) = \ln B$ can be made to work by choosing a prior model which satisfies Equation (5b) but has features of required B_0. The model image B_0 can then slowly be modified towards the required B_0 as the iteration progresses. One of the simplest schemes is to add a suitable constant to the image as well as to the required prior model such that Equation (5b) is satisfied at all points of the image. As the iteration progresses the value of the constant is gradually reduced and ultimately a reconstructed image is obtained which is closest to the required prior model.

So, it appears that $f = \ln B$ can be a correct choice (although with little modifications) to define the relative entropy as in Equation (2).

Other well-known function $-B \ln B$ when substituted in Equation (2) gives two solutions namely $B = e^{-1}$ and $B = B_0$. Since both solutions could very well be acceptable, the uniqueness criterion is not fulfilled and therefore $f(B) = -B\ln B$ is not a correct function to define a relative entropy of kind given by Equation (2). It should be noted that in the case of $f(B) = \ln B$, although mathematically there are two solutions (*i.e.*, at $B = B_0$ and $B = \infty$), for all practical purposes there is only one solution at $B = B_0$. Solution $B = \infty$ is automatically discarded since no observational data would support it.

Another entropy function is a power-law function of kind $f(B) = B^s$, where s is a suitable power-law index. For this function the uniqueness condition (4) can be written as

$$B_0/B > \left(\frac{2S-1}{S-1}\right)^{1/s}. \tag{6}$$

This function also gives one solution at $B = B_0$ regardless of the choice of s. However, one may get an additional solution at $B = 0$ or $B = \infty$ depending upon the value of s. If we choose the power-law index s such that the second solution lies at $B = \infty$, for all practical purposes the function has single solution at $B = B_0$. This, in other words, puts a condition on s to be less than unity. Further, to satisfy the condition given by Equation (6) for all positive values of B, the point at which the second derivative of $F(B, B_0)$ with respect to B is zero, should be pushed to infinity giving $s = 1/2$. These arguments clearly indicate that the relative entropy defined by Equation (2) using $f(B) = B^{1/2}$ is always concave for any positive value of B and therefore, always biases the solution toward $B = B_0$. It appears that this form of relative entropy should be as good as the one obtained from information-theoretic arguments (Shore & Johnson 1980; Gull & Skilling 1983). This point should be verified by actually reconstructing images with two types of entropy functions.

3. Minimum relative entropy method

We are given a set of measured visibility points $\rho(u_j, v_j)$; $j = -K$ to K and an *a priori* known model distribution B_0. We want to estimate the unmeasured visibility coefficients such that the measurements are unchanged and the relative entropy integral given by Equation (2) is minimized. The measured visibility coefficients are related to the true brightness distribution $B_{\text{true}}(x, y)$ through a Fourier-transform relationship, *i.e.*,

$$\rho(u_j, v_j) = \iint_{\text{field}} B_{\text{true}}(x, y)\exp[2\pi i(u_j x + v_j y)]\,\mathrm{d}x\,\mathrm{d}y, \qquad j = -K \text{ to } K \tag{7}$$

and the observed image

$$B(x, y) = \sum \rho(u_j, v_j)\exp[-2\pi i(u_j x + v_j y)] \qquad j = -K \text{ to } K. \tag{8}$$

Substituting Equation (8) in (2) and differentiating with respect to the unmeasured visibility coefficients $\rho(u_j, v_j)$ we get the gradient of the entropy with respect to the unmeasured visibility coefficients as

$$g(u_j, v_j) = \frac{\partial E_r}{\partial \rho(u_j, v_j)} = \iint [f(B) - f(B_0)]\frac{\partial f}{\partial B}\exp[-2\pi i(u_j x + v_j y)]\,\mathrm{d}x\,\mathrm{d}y. \tag{9}$$

Knowing the gradient of the entropy through a Fourier transform, the conjugate gradient method can be implemented quite easily as has been done in the past (Shevgaonkar 1986a, b; Nityananda & Narayan 1982). However, the choice of model distribution has to be made before we proceed for gradient computation.

Simple considerations tell us that at least the integrated power of the model distribution should be equal to that of the observed image. Also, if we do not have any *a priori* knowledge about the distribution, an obvious choice would be to distribute the total power uniformly over the entire field of view (see also Cornwell & Evans 1985).

One can also use a low-resolution image of the brightness distribution as a biasing model. However, it should be mentioned that none of these choices work satisfactorily for estimating the visibilities over short baselines and one has to provide a model which is neither present in the data nor present in the built-in nature of the MEM.

In the simulations presented here it is assumed that some *a priori* information is available only for the part of the sources in the field of view. First, a model distribution is generated for the partially known sources. The integrated power in the generated distribution $\rho_0(0,0)$ is computed. The difference between the actually measured integrated power $\rho(0,0)$ and $\rho_0(0,0)$ is then uniformly distributed over the field of view of the model map.

Since entropy can be defined only for positive images, during initial stages of iterations, when there are large negative bowls, the entropy of the image becomes imaginary. For computational convenience a constant intensity is temporarily added to all pixels of the image such that the image becomes positive definite. This operation is called the FLOATing of the image and it has been discussed and used successfully in the past (see Nityananda & Narayan 1982; Shevgaonkar 1986a, b). The value of the floating constant decides (although not very crucially) the degree of nonlinearity in the reconstruction.

It is obvious that the floating of the image in turn modifies the integrated power or the zero-lag visibility coefficient temporarily. As the integrated power of the image is modified the integrated power of the model image should also be changed accordingly. To equate the integrated power of the model image to that of the brightness distribution, the same floating constant is added to the model map also. Apart from this floating, the maps have to be further floated to satisfy the condition given by Equation (5) if $\ln B$ function is used to define the relative entropy.

As an example we have taken an extended source (Fig. 1) as the true image. If this source is observed with a *uv*-coverage as in Fig. 2 one obtains the observed image as in

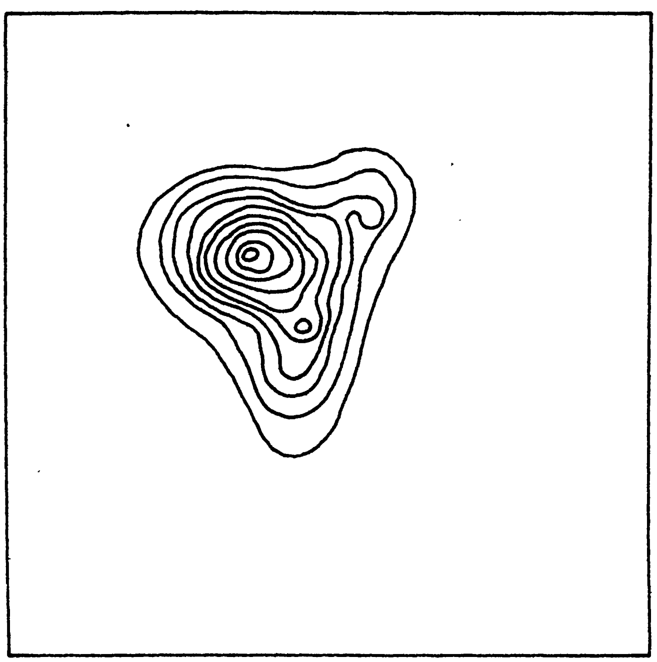

Figure 1. True image of an extended brightness distribution. Contour interval = 13.8 units.

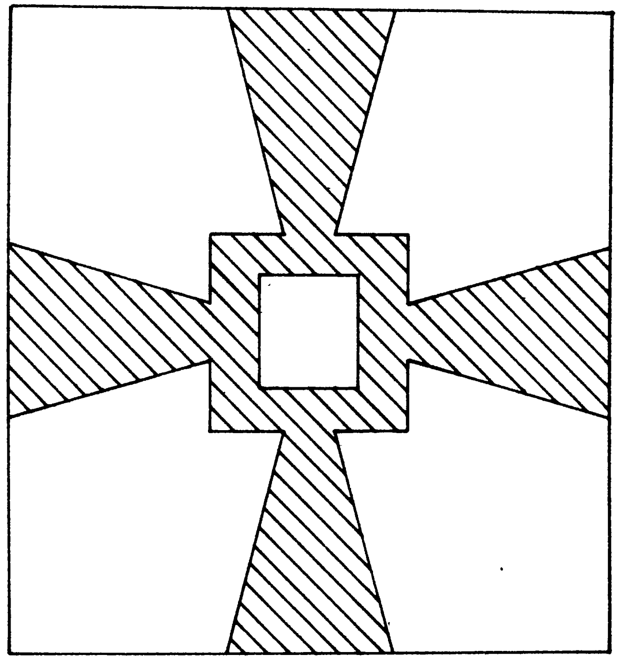

Figure 2. Simulated *uv*-coverage. Visibilities have been sampled in the hatched region. In the empty region the visibility coefficients are completely unknown. The origin of the *uv*-coverage or the integrated power of the distribution has also been measured.

Fig. 3. One should note that apart from the four missing sectors in the *uv*-plane the measurements leave a big hole around the center of the *uv*-coverage. From Fig. 3. it can be seen clearly that due to missing short baselines the map has a large negative bowl around the true position of the source. Just to emphasize the inadequacy of the existing MEM or use of relative entropy with flat model for estimating short baselines, the MEM with a uniform B_0 has been tried on the observed image. The reconstructed image has been shown in Fig. 4. It is clear that the MEM without a prior model or a flat default image could not give a faithful reconstruction.

Minimum-relative entropy method using relative entropy function $F(B, B_0) = (\ln B - \ln B_0)^2$ has been tried on the image shown in Fig. 3 and found to give quite satisfactory results. A simple circular Gaussian source (Fig. 5) is chosen as a model image. It should be noted that the model distribution does not have any direct similarity to the true source except that it tells that the source is of extended nature. The size of the gaussian is roughly decided by the expected source dimension and the height of the gaussian is decided by the equality of the total integrated powers in the model and the observed distribution. The reconstructed image as shown in Fig. 6 is a remarkable improvement over the image obtained by simple MEM (Fig. 4). The stability of the reconstruction against the choice of the model is tested by choosing model gaussian sources of different sizes and intensities located at slightly shifted locations. The reconstruction seems to be fairly stable against the choice of the initial model. However, following scheme may be useful in deciding the first-order biasing model.

As we have already mentioned in the introduction, our main objective is to reconstruct those distributions which are responsible for the large negative bowls in the

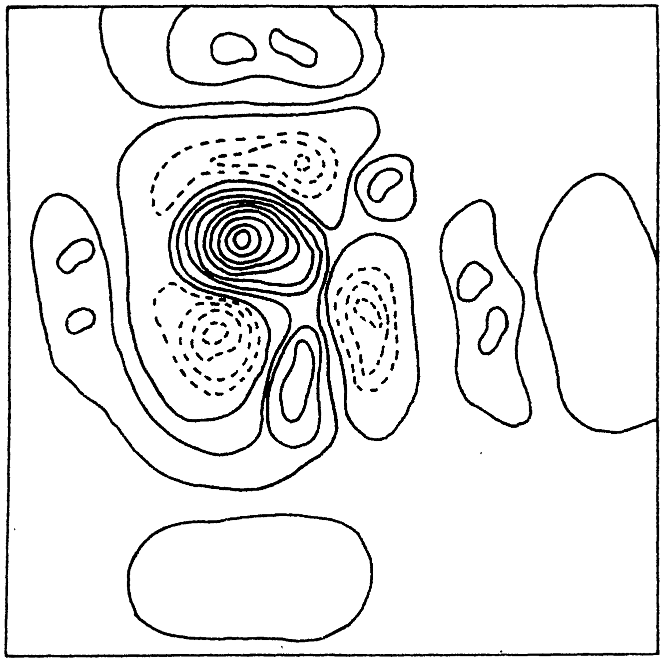

Figure 3. Image mapped by the *uv*-coverage in Fig. 2. Contour interval = 5.9 units.

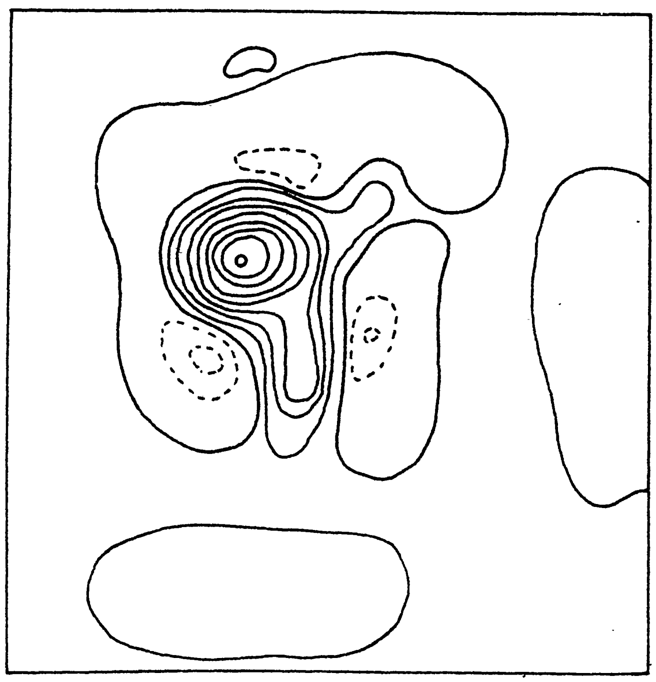

Figure 4. MEM reconstructed image after 30 iterations. Entropy function is $\ln B$. Contour interval = 7.9 units.

image. It can be seen easily that the partly missing largescale structure should have a size which is larger than the inner rim and smaller than the outer rim of the negative bowl. Therefore, a first-order prior model can be taken as a gaussian distribution of width equal to the mean diameter of the negative bowl. As the reconstruction progresses and

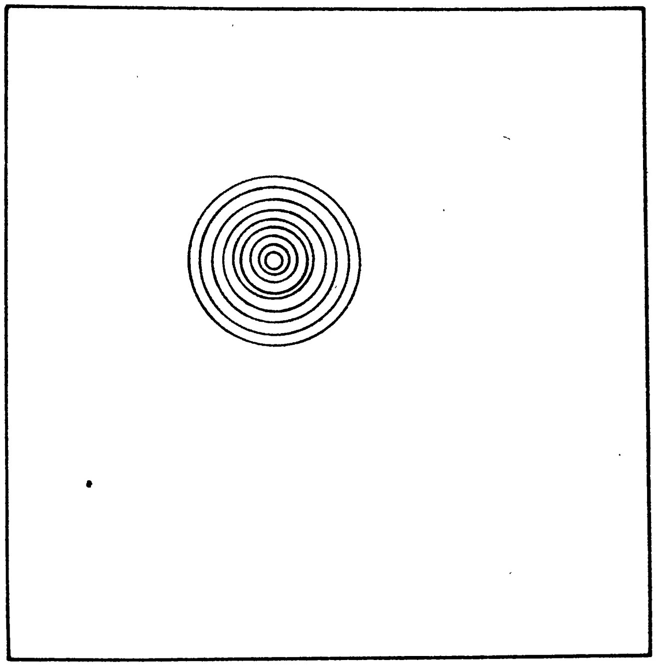

Figure 5. A circular gaussian source used as a default image to define the relative entropy. Contour interval = 16.8 units.

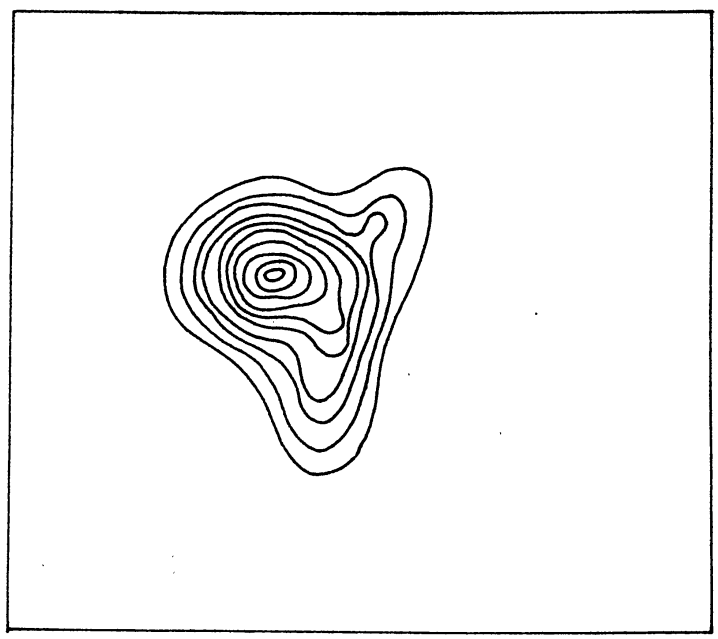

Figure 6. MREM reconstructed image after 20 iterations. Relative entropy function is $(\ln B - \ln B_0)^2$. Contour interval = 12.6 units.

the missing largescale structures are more clearly visible, the default image should subsequently be modified.

Other relative entropies $B \ln (B/B_0)$ and $(B^{1/2} - B_0^{1/2})^2$ have also been tried and found to provide reconstruction of the same quality as per our expectation. A comparison of the three relative-entropy functions is given in the next section.

4. Minimum-relative entropy method for polarized emission

The maximum entropy method has been formulated and successfully applied to the polarization images (Ponsonby 1973; Nityananda & Narayan 1983; Shevgaonkar 1986b). It has been shown that although the *a priori* positivity constraint is applied only to the total-intensity image, the polarized component images also get reconstructed remarkably well. Here we seek a possibility of successful application of the minimum-relative-entropy method to polarization images. We will confine our formulation to three relative entropy functions, namely $B \ln (B/B_0)$, $(\ln B - \ln B_0)^2$ and $(B^{1/2} - B_0^{1/2})^2$.

A partially polarized brightness distribution $\underline{B}$ can be elegantly represented by the Stokes parameters I, Q, U and V. Parameter I represents the total intensity, Q and U together represent linear polarization, and V represents the circular polarization. These Stokes parameters are related to the visibility functions ρ^I, ρ^Q, ρ^U and ρ^V through Fourier-transform relationship similar to that given by Equations (7) and (8). Following previous authors (Ponsonby 1973; Nityananda & Narayan 1983; Gull & Skilling 1984; Shevgaonkar 1986b) the entropy of an arbitrary polarized distribution is equal to the sum of the entropies of the two orthogonal polarization components λ_1 and λ_2 of the polarized image. It is shown in the past (Ponsonby 1973) that the two orthogonally polarized components λ_1 and λ_2 are equal to $I(1+d)/2$ and $I(1-d)/2$ respectively. Here d is the degree of polarization and is defined as

$$d = (Q^2 + U^2 + V^2)^{1/2}/I. \tag{10}$$

Now if $F(B, B_0)$ is the relative entropy function, where B is the image and B_0 is the biasing model, the relative entropy of a polarized brightness distribution $\underline{B}$ with respect to the prior model $\underline{B}_0$ can be written as

$$E_{\rm rp} = \iint F(\underline{B}, \underline{B}_0)\, {\rm d}x\, {\rm d}y = \iint [F(\lambda_1, \lambda_{10}) + F(\lambda_2, \lambda_{20})]\, {\rm d}x\, {\rm d}y. \tag{11}$$

Suffix 0 indicates the corresponding value for the model distribution. Differentiating Equation (11) with respect to the unmeasured visibility coefficients $\rho^k(u_j^k, v_j^k)$ ($k = I, Q, U, V$) we get the gradient of the entropy as

$$g^k(u_j^k, v_j^k) = \frac{\partial E_{\rm rp}}{\partial \rho^k(u_j^k, v_j^k)} = {\rm FT}\left[\frac{\partial F(\underline{B}, \underline{B}_0)}{\partial k}\right]. \tag{12}$$

Therefore, for computation of the entropy gradient we require derivatives of $F(B, B_0)$ with respect to the four Stokes parameters. The term inside the bracket in Equation (12) can be split into two parts as

$$\frac{\partial F(\underline{B}, \underline{B}_0)}{\partial k} = \frac{\partial F(\lambda_1, \lambda_{10})}{\partial k} + \frac{\partial F(\lambda_2, \lambda_{20})}{\partial k}. \tag{13}$$

Equation (13) can be re-written for different Stokes parameters as

$$\frac{\partial F(\underline{B}, \underline{B}_0)}{\partial I} = \frac{\partial F(\lambda_1, \lambda_{10})}{\partial \lambda_1} + \frac{\partial F(\lambda_2, \lambda_{20})}{\partial \lambda_2} \tag{14a}$$

and

$$\frac{\partial F(\underline{B}, \underline{B}_0)}{\partial (Q, U, V)} = \frac{(Q, U, V)}{\Sigma}\left[\frac{\partial F(\lambda_1, \lambda_{10})}{\partial \lambda_1} - \frac{\partial F(\lambda_2, \lambda_{20})}{\partial \lambda_2}\right], \tag{14b}$$

where $\Sigma = (Q^2 + U^2 + V^2)^{1/2} = Id$.

Now, substituting for $\lambda_1, \lambda_2, \lambda_{10}, \lambda_{20}$ and desired relative entropy function $F(\underline{B}, \underline{B}_0)$ in Equations (14a, b) we can obtain elegant expressions for $\delta F/\delta k$.

(i) For $F(\underline{B}, \underline{B}_0) = \underline{B} \ln(\underline{B}/\underline{B}_0)$ we get

$$\frac{\partial F}{\partial I} = 2 + \ln(D/D_0) \tag{15a}$$

and

$$\frac{\partial F}{\partial(Q, U, V)} = \frac{(Q, U, V)}{\Sigma} \ln(\chi/\chi_0) \tag{15b}$$

where $D \equiv I^2 - \Sigma^2 = I^2 - Q^2 - U^2 - V^2$ and $\chi \equiv (I + \Sigma)/(I - \Sigma)$, and suffix 0 indicates the corresponding value for the model distribution $\underline{B}_0$.

(ii) For $F(\underline{B}, \underline{B}_0) = (\ln \underline{B} - \ln \underline{B}_0)^2$ we get

$$\frac{\partial F}{\partial I} = \frac{I}{D}[\ln(D/D_0) - \mathrm{d}\ln(\chi/\chi_0)] \tag{16a}$$

and

$$\frac{\partial F}{\partial(Q, U, V)} = -\frac{(Q, U, V)}{D}\left[\ln(D/D_0) - \frac{1}{d}\ln(\chi/\chi_0)\right]. \tag{16b}$$

(iii) Finally, for $F(\underline{B}, \underline{B}_0) = (\underline{B}^{1/2} - \underline{B}_0^{1/2})^2$ we derive

$$\frac{\partial F}{\partial I} = -\frac{(I_0 - \Sigma_0)^{1/2}}{(I + \Sigma)^{1/2}}(\chi^{1/2} + \chi_0^{1/2}) \tag{17a}$$

and

$$\frac{\partial F}{\partial(Q, U, V)} = \frac{(I_0 - \Sigma_0)^{1/2}}{(I + \Sigma)^{1/2}}(\chi^{1/2} - \chi_0^{1/2})\frac{(Q, U, V)}{\Sigma}. \tag{17b}$$

By Fourier transforming Equations (15a, b), (16a, b) or (17a, b), we can compute the entropy gradient $(g^k(u_j^k, v_j^k))$ with respect to the unmeasured visibility coefficients $\rho^k(u_j^k, v_j^k)$. Knowing the gradient of the entropy, a simple-gradient or conjugate-gradient method can be implemented easily (for details see Shevgaonkar 1986b) for minimizing the relative entropy. To test the algorithm let us take the total intensity (I) and circular polarization (V) maps of a solar active region at 6 cm wavelength (Fig. 7a, b) (Shevgaonkar & Kundu 1985) observed with the Very Large Array. To demonstrate the strength of the method we have assumed here that the measured *uv*-coverages are not identical for the two component images I and V. The total intensity map has a *uv*-coverage with four sectors and a central annular rectangle as shown in Fig. 2. We assume that the circular polarization measurement could be performed reliably only over the central annular rectangle, and for long baselines in the four sectors the polarization measurement is heavily affected by the instrumental errors and therefore it has been treated as unmeasured. It should be noted that the *uv*-coverages for both I and V have a large hole around the origin of the *uv*-plane. The synthesized beams corresponding to the two *uv*-coverages are convolved respectively with true total intensity and circular polarization maps (Fig. 7a, b) to get the observed dirty maps as shown in Fig. 8a, b.

Before we go to the minimization of relative entropy the choice of model polarized brightness distribution has to be made. A prior model for the intensity distribution could be obtained from other independent observations. However, the polarization

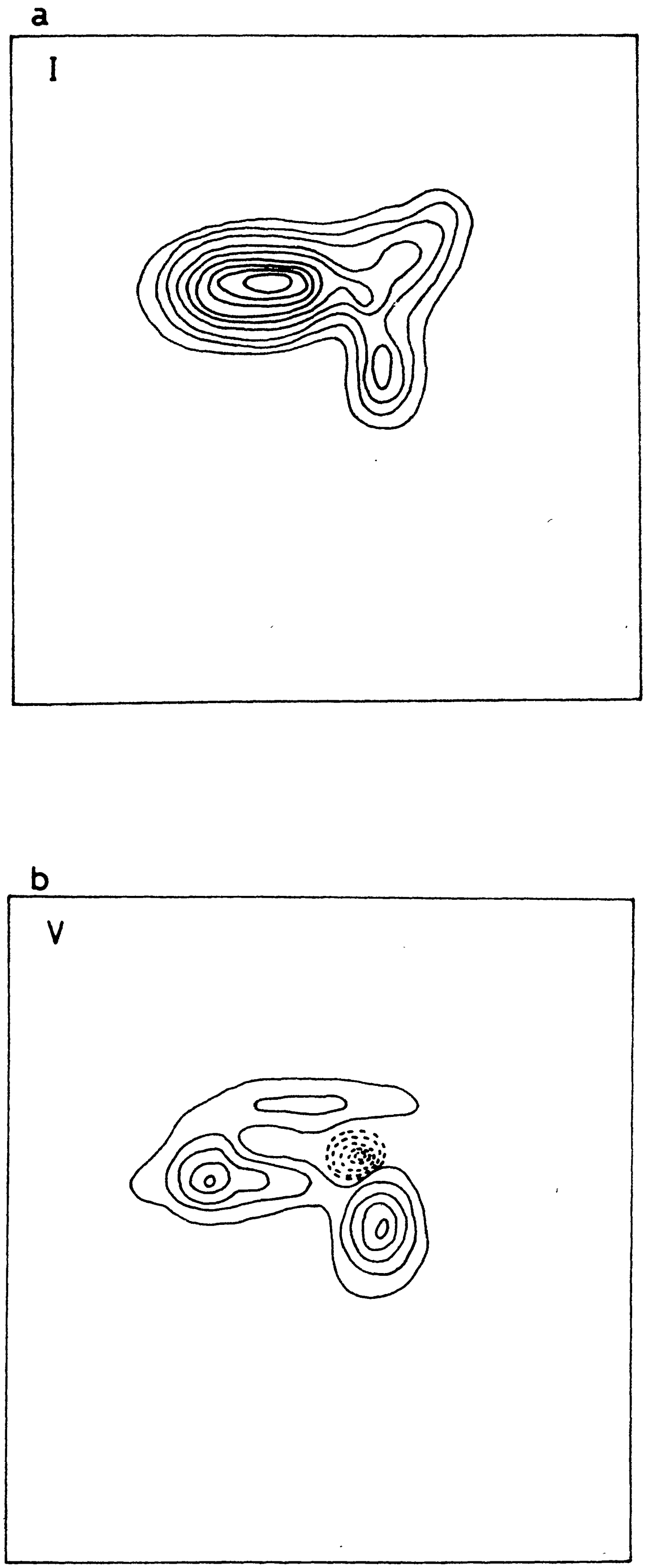

Figure 7. (a) True total intensity *I* and (b) circular polarization *V* images of a solar active region at 6 cm wavelength. Contour interval = 10.0 and 6.0 units respectively for *I* and *V* maps.

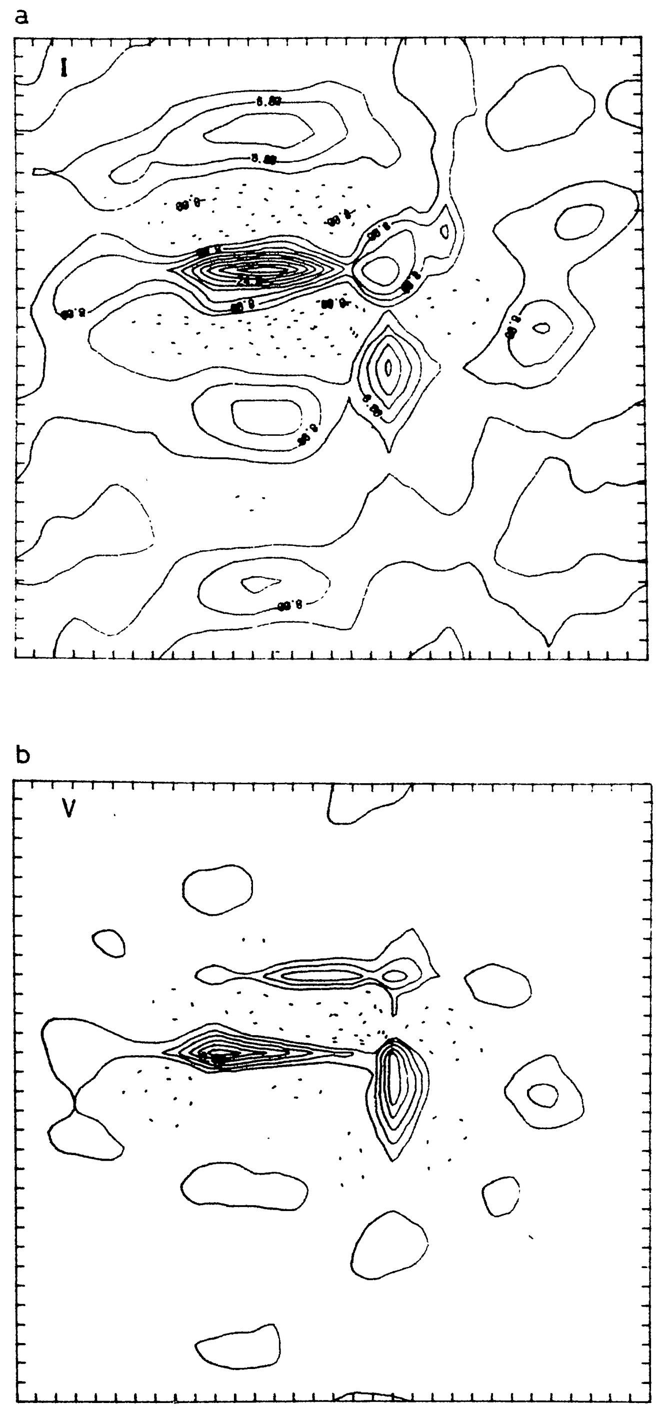

Figure 8. (a) Observed total intensity *I* and (b) circular polarization *V* maps.

model is not that readily available. To justify this statement let us take a concrete example.

Suppose we are observing a flaring event on the Sun. The event is visible over a wide range of frequencies *i.e.*, from X-rays to the radio. The radio data is capable of showing polarization whereas the optical or X-ray emission does not contain any information regarding the polarization. One can use the X-ray or optical image as a model for total intensity image but the polarized emission is still unmodelled. If we use one of the synthesizing elements to obtain zero-lag visibility coefficient, at most we can get the integrated power in each polarization component.

As discussed in Section 3, in the absence of any prior knowledge about the image, the integrated power in each polarization component should be distributed uniformly over the entire field of view. This could be a suitable scheme provided all the four Stokes parameters are assumed to have flat default images. However, if we take a non-uniform total-intensity model and a flat polarization model, it is quite likely, especially on the edges of the source, to encounter an awkward situation of having a degree of polarization greater than unity in the model distribution. We have found that the best way is to give the polarization model as a scaled version of the total-intensity model. The scaling constant is decided by the relative integrated powers in the polarization components and the total-intensity distribution.

A distribution as in Fig. 9 has been chosen as a default model to define the relative entropy. The reconstructions for three entropy functions after 20 iterations are shown in Figs 10–12. It is clear that all the three relative entropy functions give more or less identical reconstructions. The strength of the method is quite apparent from the choice of non-identical *uv*-coverages for different polarization components and from the selection of polarization model as a scaled version of the total-intensity model. From the variety of examples given above it is convincing that the relative-entropy functions which do not have their origin in the information theory are also capable of providing good image reconstruction.

5. Conclusion

In synthesis observations one commonly encounters situations where the short baselines are inadequately sampled. As a result, the extended distributions are poorly

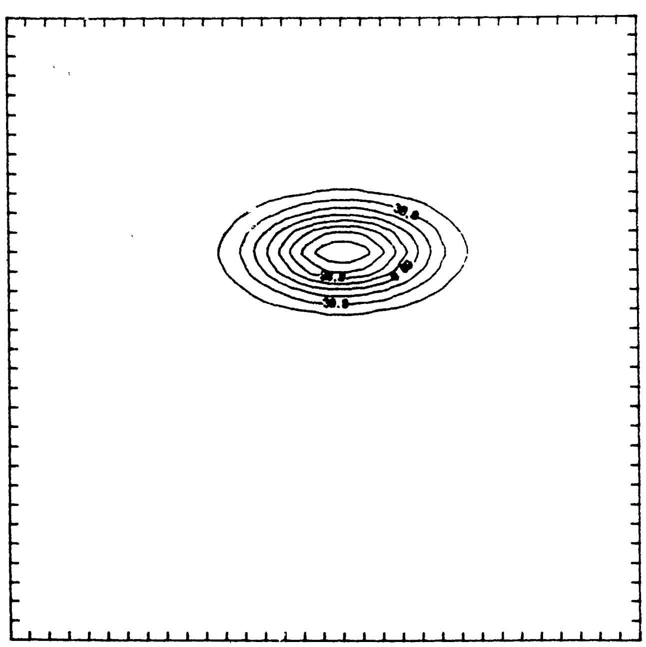

Figure 9. An elliptic gaussian source used as a default image to define the relative entropy. Contour interval = 15.0 units.

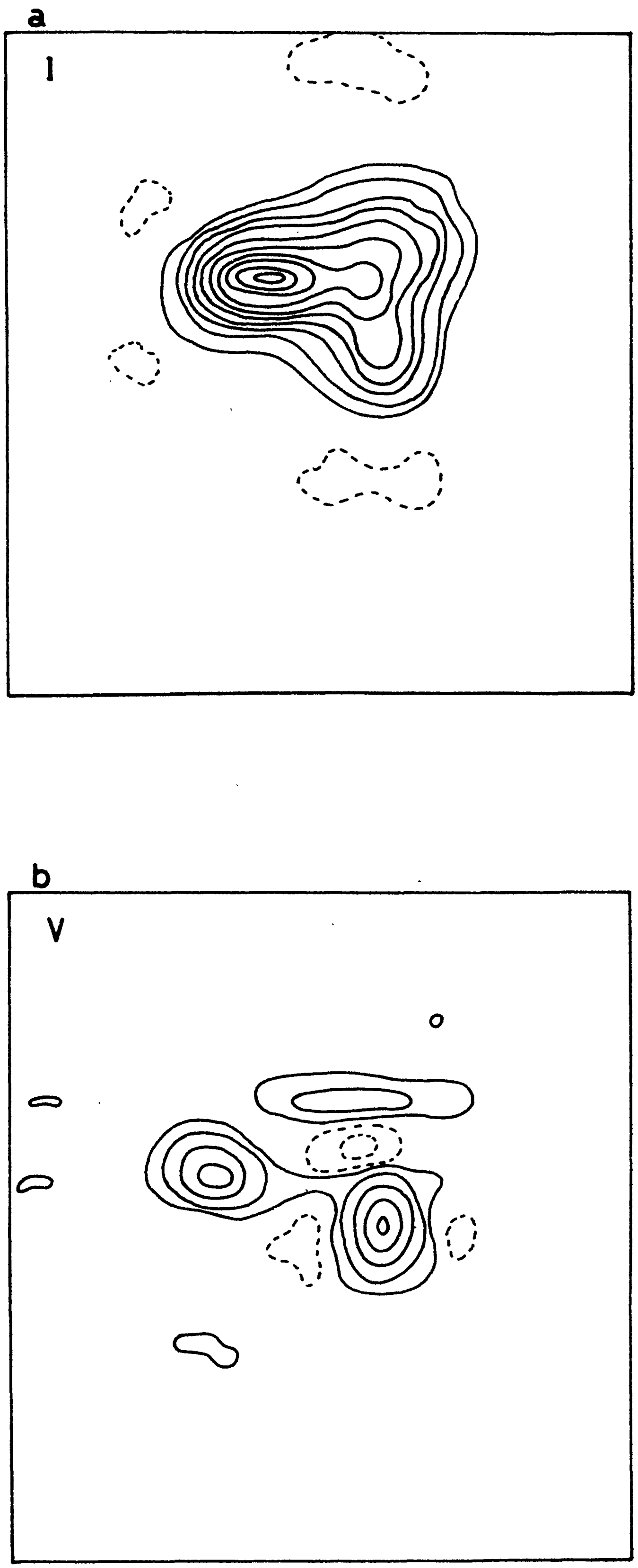

Figure 10. MREM reconstructed images after 20 iterations. Relative entropy function is $B \ln (B/B_0)$. Contour interval = 11.8 and 2.4 units for I and V maps respectively.

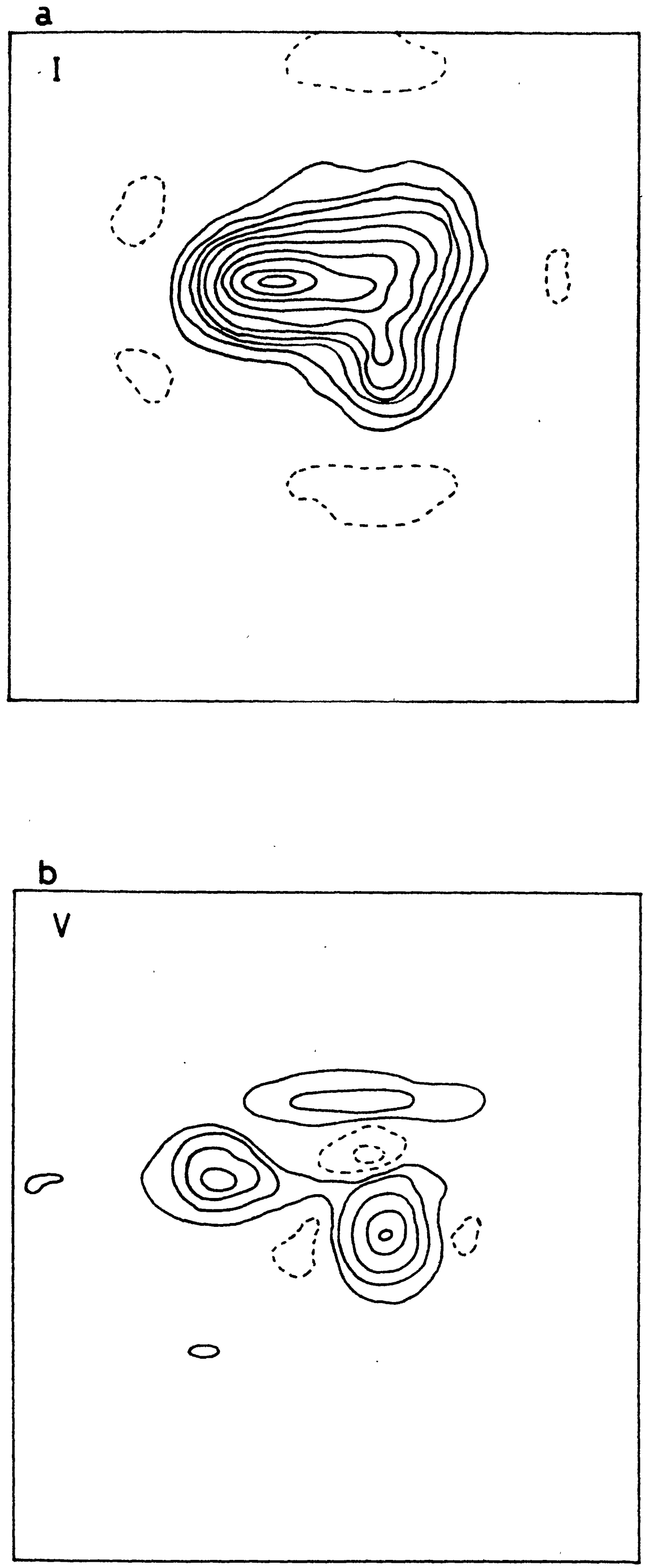

Figure 11. MREM reconstructed images after 20 iterations. Relative entropy function is $(B^{1/2} - B_0^{1/2})^2$. Contour interval = 11.8 and 2.4 units for I and V maps respectively.

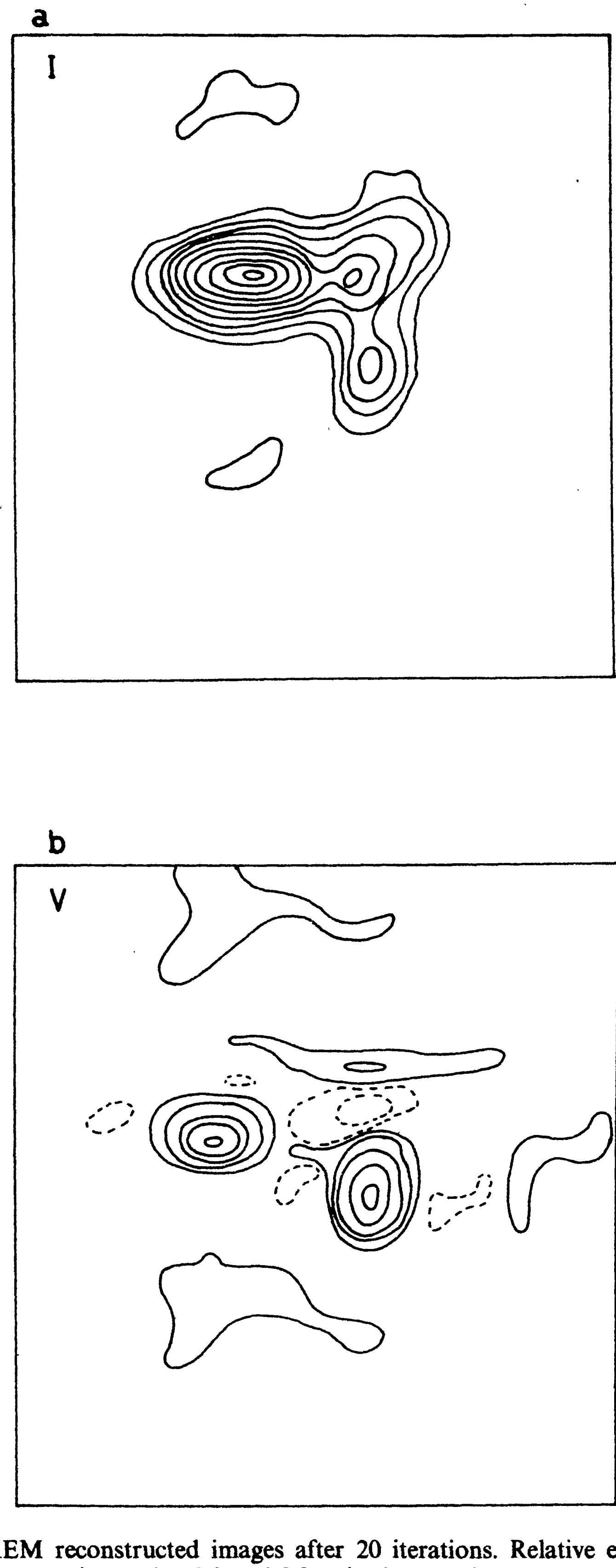

Figure 12. MREM reconstructed images after 20 iterations. Relative entropy function is $(\ln B - \ln B_0)^2$. Contour interval = 9.2 and 2.2 units for I and V maps respectively.

mapped. Minimum-relative-entropy method (MREM) has been presented as a possible scheme for reconstructing extended sources mapped with sparsely sampled short baselines. It is shown that as far as image reconstruction is concerned a measure of distance between the image and the prior model in the plane of pixel brightness is an adequate measure of relative entropy. The reconstructions obtained using this non-information-theoretic definition of the relative entropy have been compared with that obtained by minimizing the entropy $B \ln (B/B_0)$ which has a firm information-theoretic base. It has been argued that for estimation of the short baseline visibilities, a flat default image is not sufficient and one must provide a prior model which is neither present in the measurements nor in the inherent properties of the MEM.

The method has been generalized for partially polarized images. It is argued that in the absence of a prior model for the polarization components, one must choose a default image which has constant degree of polarization over the field of view. The potential of the method has been demonstrated by choosing non-identical *uv*-coverages for different polarization components.

Acknowledgements

I thank Raman Research Institute for providing excellent computing and other facilities.

References

Braun, R., Walterbos, R. A. M. 1985, *Astr. Astrophys.*, **143**, 307.

Burg, J. P. 1975, Ph D Thesis, Stanford Univ.

Cornwell, T. J., Evans, K. F. 1985, *Astr. Astrophys.*, **143**, 77.

Ekers, R. D., Rots, A. H. 1979, in *IAU Coll.* 49: *Image Formation from Coherence Function in Astronomy*, Ed. C. van Schooneveld, D. Reidel, Dordrecht, p. 61.

Gull, S. F., Daniell, G. J. 1978, *Nature*, **272**, 686.

Gull, S. F., Skilling, J. 1984, in *Indirect Imaging*, Ed. J. A. Roberts, Cambridge Univ. Press, p. 267.

Hobson, A., Cheng, B. 1973, *J. Stat. Phys.*, **7**, 301.

Högbom, J. A. 1974, *Astr. Astrophys. Suppl.*, **15**, 417.

Johnson, R. W. 1979, *IEEE Trans. Inf. Theory*, IT-25, 709.

Narayan, R., Nityananda, R. 1984, in *Indirect Imaging*, Ed. J. A. Roberts, Cambridge Univ. Press p. 281.

Nityananda, R., Narayan, R. 1982, *J. Astrophys. Astr.*, **3**, 419.

Nityananda, R., Narayan, R. 1983, *Astr. Astrophys.*, **118**, 194.

Ponsonby, J. E. B. 1973, *Mon. Not. R. astr. Soc.*, **163**, 369.

Rots, A. H. 1979, in *IAU Coll.* 49: *Image Formation from Coherence Functions in Astronomy*, Ed. C. van Schooneveld, D. Reidel, Dordrecht, p. 55.

Schwarz, U. J. 1978, *Astr. Astrophys.*, **65**, 345.

Shevgaonkar, R. K. 1986a, *Astr. Astrophys.*, **162**, 349.

Shevgaonkar, R. K. 1986b, *Astr. Astrophys.*, (in press).

Shevgaonkar, R. K., Kundu M. R. 1985, *Astrophys. J.*, **292**, 733.

Shore, J. E., Johnson, R. W. 1980, *IEEE Trans. Inf. Theory*, IT-26, 26.