JOURNAL OF THE KOREAN ASTRONOMICAL SOCIETY
27: 191 ~ 201, 1994

MAXIMUM POWER ENTROPY METHOD FOR LOW CONTRAST IMAGES *

CHAE, JONG-CHUL AND YUN, HONG SIK
Department of Astronomy, Seoul National University
(Received Oct. 7, 1994;Accepted Oct. 20, 1994)

ABSTRACT

We propose to use the entropy of power spectra defined in the frequency domain for the deconvolution of extended images. Spatial correlations requisite for extended sources may be insured by increasing the role of power entropy because the power is just a representation of spatial correlations in the frequency domain. We have derived a semi-analytical solution which is found to severely reduce computing time compared with other iteration schemes. Even though the solution is very similar to the well-known Wiener filter, the regularizingng term in the new expression is so insensitive to the noise characteristics as to assure a stable solution. Applications have been made to the IRAS $60\mu m$ and $100\mu m$ images of the dark cloud B34 and the optical CCD image of a solar active region containing a circular sunspot and a small pore.

Key Words : techniques:image processing, infrared:interstellar, sun:photosphere

I. INTRODUCTION

Deconvolution of images is one of many inference problems frequently encountered in astronomy. It aims to seek an object image from observed blurred image data. As often found in most inference problems, we encounter a difficulty of setting a unique and well established solution in the image deconvolution. Any direct method such as the matrix inversion usually fails to give a physically meaningful result. Thus, statistical apporach is commonly employed.

The statistical approach seeks a solution which maximizes a posteriori probability, $P(\mathbf{x}|\mathbf{d})$, which is the conditional probability of the solution $\mathbf{x}$ when the data $\mathbf{y}$ are given. According to Bayes' theorem, the *a posteriori* probability can be wirtten as

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{x})P(\mathbf{y}|\mathbf{x})}{P(\mathbf{y})}, \tag{1}$$

where $P(\mathbf{y}|\mathbf{x})$ is the likelihood function and $P(\mathbf{x})$ is the *a priori* probability of the restored image. $P(\mathbf{d}|\mathbf{x})$ expresses the fittness to the data with a chosen solution, while $P(\mathbf{x})$ contains *a priori* information on the true solution. If the number of data points is much greater than that of object points or/and the quality of data is sufficiently good, it is possible to find a satisfactory solution by maximizing only the likelihood function, $P(\mathbf{d}|\mathbf{x})$. In most cases, however, the solution is very unstable and so noisy that an additional regulariziging term is required to retain physically plausible features while reducing artifacts. The prior probability plays the role of this regularizing term.

The total entropy of intensity distribution in an object image S is defined as

$$S = \sum_i s_i \tag{2}$$

where the specific entropy of ith pixel is given by a function of the intensity value at ith pixel

* This research was supported by the Basic Science Research Institue Program, Ministry of Education. BSRI-93-591

$$s_i = f(x_i). \tag{3}$$

This total entropy S has been used to specify the prior probability $P(\mathbf{x})$ preferentially in the form of $P(\mathbf{x}) \propto \exp(\alpha S)$ with a regularizing parameter α. The specific entropy s_i is often taken as, following Skilling(1990),

$$f(x) = x - m + x\log(x/m) \tag{4}$$

which has been proved to be adaquate for positive and additive distributions, where m is the default value which is usually taken to be constant over pixels. In their comprehensive discussion on the general properties of the maximum entropy images, Narayan and Nitayanda(1986) demonstrated that if the second derivative of the entropy function $f(x)$ varies significantly over the range of x, its Fourier extrapolation and interpolation cause the peaks in the image to be very sharp and the osillations present at the base line of x to be very weak because of the non-linearity of the function $f(x)$. These two characteristics, namely the superresolution and the ripple suppression are the main attraction of using the maximum entropy method for image reconstructions. *The maximum entropy method is best suited for images composed of isolated point or compact sources against the low background level.* However, in the case that the background level is high, the nonlinearity of the entropy function $f(x)$ breaks down so that these characteristics disappear. The ripples near a point source emebedded in an extended source region are not well suppressed. The maximum entropy method does not work well in the presence of extended sources or high level background since the entropy does not have any spatial correlation between the pixels.

In order introduce the spatial correlations into image reconstructions, Gull(1989) and Skilling(1989) interpreted the object image $\mathbf{x}$ as a convolution of some hidden, spatially uncorrelated, image $\mathbf{h}$ with a blurring function, namely the intrinsic correlation function. Even though an additional smoothing is brought into the object image, there remains a problem of determining how many correlation length scales are needed and how large they should be taken.

In the present study we propose to use the *entropy of power spectra* defined in the frequency domain instead of the most frequently employed entropy of intensity in the spatial domain, especially when low contrast extended images are deconvolved. In view of the fact that the power spectra is just a representation of spatial correlations in the frequency domain at virtually all length scales, the power entropy should be very well suited for extended sources, since these are characterized primarily by the spatial correlations. Contrary to the approach which makes use of intrinsic correlation functions, the power entropy approach does not need for particular scale lengths. Noting that the Fourier transform of the extended sources defined in the spatial domain becomes compact in the frequency domain and that the intensity entropy has been successful in compact sources, we may expect the power entropy should work well for the case of the low contrast extended images. Namely, *the power entropy favors low contrast extended sources.* One of the benefits of using the power entropy resides in the fact that both the entropy and the likelihood are dealt only in the frequency domain so that a semi-analytic solution can be found very easily.

In the following sections we present the maximum power entropy method and its applications to the low contrast IRAS images of dark cloud B34 and the optical image of an solar active region containing a sunspot and a pore. Detailed discussions will be made on the role of the power entropy in the image reconstruction.

II. MAXIMUM POWER ENTROPY METHOD

The deconvolution is the process of finding out $\mathbf{x}$ from a set of data $\mathbf{y}$ given by

$$\mathbf{y} = \mathbf{x} * \mathbf{f} + \mathbf{n} \tag{5}$$

or, in a dicrete form,

$$y_j = \sum_i x_i f_{ji} + n_j \tag{6}$$

with $f_{ji} = f(\mathbf{r}_j - \mathbf{r}_i)$ where $f(\mathbf{r})$ is the point spread function and n_j is the random noise with a standard deviation σ_j associated with the pixel j. If the noise is spatially uncorrelated, the likelihood function becomes

$$P(\mathbf{y}|\mathbf{x}) = \prod_j \frac{1}{\sqrt{2\pi}\sigma_j} \exp(-\frac{(y_j - z_j)^2}{2\sigma_j^2}) \tag{7}$$

for the Gaussian noise and

$$P(\mathbf{y}|\mathbf{x}) = \prod_j \exp(-z_j)\frac{z_j^{y_j}}{y_j!} \tag{8}$$

for the Poisson noise. Here z_j is the expectation value of the data y_j which is equal to $\sum_i x_i f_{ji}$. The prior probabilty can be defined as

$$P(\mathbf{x}) \propto \exp(\alpha \sum_n s_n) \tag{9}$$

with the power entropy

$$s_n = P_n - m_n - P_n \log(P_n/m_n) \tag{10}$$

where P_n is the power spectra equal to $X_n X_n^*$. We note that the discrete Fourier transform of x_i is given by

$$X_n = \sum_i x_i \Theta_{ni} \tag{11}$$

where $\Theta_{ni} = \exp(j\mathbf{k}_n \cdot \mathbf{r}_i)$ is the Fourier Kernel.

In the case of Gaussian noise, it can be shown from equation (1) that the maximization of the *a posteriori* probability $P(\mathbf{x}|\mathbf{d})$ is equivalent to the minimization of the following functional

$$H = \alpha \sum_n [P_n \log(P_n/m_n) - P_n + m_n] + \frac{1}{2}\sum_j \frac{(y_j - z_j)^2}{\sigma_j^2}. \tag{12}$$

If the standard deviation is assumed to be constant over pixels, we can specify this functional purely in the frequency domain as

$$H = \alpha \sum_n [|X_n|^2 \log(|X_n|^2/m_n) - |X_n|^2 + m_n] + \frac{1}{2}\sum_n |Y_n - F_n X_n|^2/N\sigma^2 \tag{13}$$

when the convoultion theorem and Parseval's relation are employed. In this equation, N is the total number of pixels. The last term of equation (13) can be also derived directly by assuming that the distance between Y_n and $F_n X_n$ follows Gaussian statistics with the standard deviation σ_n at the nth pixel in the frequency domain. Differentiating H with repect to X_n^* one finds that

$$2\alpha \log(|X_n|^2/m_n)X_n - \frac{(Y_n - F_n X_n)F_n^*}{N\sigma^2} = 0, \tag{14}$$

which yields

$$X_n = \frac{Y_n F_n^*}{|F_n|^2 + \alpha \log(|X_n|^2/m_n)N\sigma^2}. \tag{15}$$

This equation is the main result of the present work. Successive substitution of X_n with an initial guess $X_n = Y_n$ usually converges to a solution within several iterations because the regularizing term $\alpha \log(|X_n|^2/m_n)N\sigma^2$ is smaller than unity and it is rather insensitive to the actual value of $|X_n|$. The solution happens to be very similar to the well-known Wiener filter. In the case of the Wiener filter, the regularizing term is given by $N\sigma^2/|X_n|^2$ which is relatively sensitive to $|X_n|$, especially at the high frequency region. One of the great advantages of the maximum power entropy method over other iterative algorithms lies in the fact that it does not take much in computations because only three Fourier transforms and a few complex operations will do for the reconstruction of an image, while the conventional maximum entropy method requires over fifty interations where four Fourier transforms are

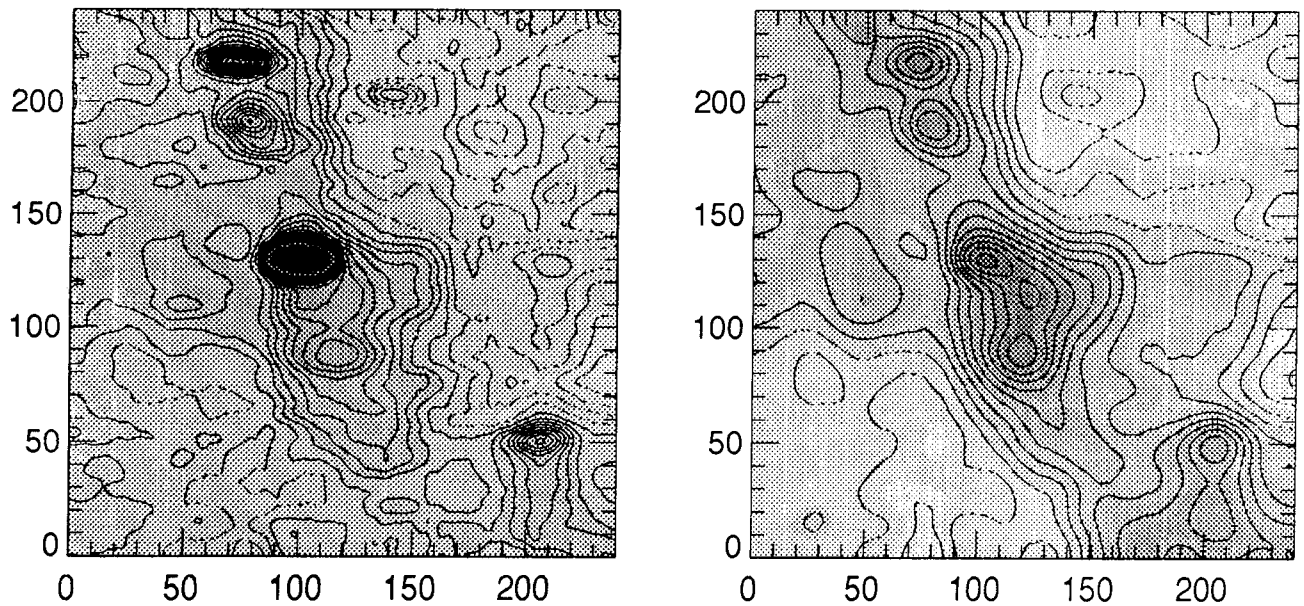

Fig. 1. The IRAS coadded images of the dark cloud B34 at $60\mu m$ (left) and $100\mu m$ (right). Drawn are contour levels with equal intervals 0.2 and 1 MJy/sterad respectively and the negative levels are represented by dotted curves. The field of view is $1^\circ \times 1^\circ$ so that one pixel corresponds to $15''$.

needed per one iteration. The weak point of the maximum power entropy method is found in a lack of correlations between Fourier components in the frequency domain, which makes it difficult to do the Fourier interpolation and extrapolation. It is impossible to reconstruct an image from an incomplete set of Fourier data. It is also difficult to get superresolution. Furthermore, the positivity is not guaranteed. In the case of low contrast images, however, this method works very well to a degree comparable or even superior to that obtained by the conventional maximum entropy method.

III. APPLICATIONS

(a) Deconvolution of IRAS B34 Images

It is now generally accepted that it is very difficult and time consuming to reconstruct high resolution images directly from the IRAS sky survey data owing to the IRAS's peculiar arrangemnet of detectors with different responses and non-uniform scan coverage. Since the sky survey data are not evenly spaced, IPAC(Infrared Processing and Analysis Center) processed the data by means of the so-called coadding algorithm to generate evenly spaced Faint Source Survey images. In this process, various artifacts such as cosmic particle events have been removed. The point sources have been filtered and the base line removed by applying median filters. The processed data are then placed into all pixels covered by the detector size to resample them to pixel centers using linear interpolations.

Figure 1 represents two coadded images of the dark cloud B34 taken at $60\mu m$ and $100\mu m$, respectively. The images are composed of a total of 241 x 241 pixels with a pixel size $0.25'$, and the center of the frame is located at $\alpha = 5^h40^m09^s, \delta = +32^\circ38'18''$. The median detector noises are 0.20 and 0.58 MJy/sterad, respectively. As seen from Figure 1, the zero level of the two images are a little bit different from each other, which could have taken place during their base-line removal.

To reduce the computing time, we have performed 2 x 2 binning on the original coadded images. Thus one pixel in these images corresponds to $0'.5$. Since our main interst lies in the cloud region, the strong point source located near the center of the image(see Figure 2) has been eliminated from the image with the use of the point spread

functions derived by the following method.

The point spread function(PSF) of the 60 μm image was obtained by using the strong point source located near the pixel x=32 and y=49. The extended components extrapolated from adjacent pixels are subtracted from the image so that only the point source component remains. The PSF is then given by this point source image divided by the total flux. The derived PSF is shown at the upper right side of Figure 2. The FWHMs along the major and minor axis are found to be 4.5′ and 2′ respectively, which are very close to the values reported by Moshir *et al.*(1992). The irregularities shown in the lowest level contour may be partly due to the error introduced by the process of separating the point source component from the extended component, which is inevitable when the two components have a strong contrast between them.

In the 100 μm case, it was very difficult to separate the point source component directly from the extended one. Thus we have made use of the analytical point spread function suggested by Diego(1985),

$$f(x,y) = \frac{C}{1 + r^{p(1+q)}} \tag{16}$$

where

$$r = \sqrt{(x/w_x)^2 + (y/w_y)^2}$$

and

$$q = \sqrt{(x/\rho_x)^2 + (y/\rho_y)^2}.$$

Here w_x and w_y are the FWHMs of the point spread function along the major and minor axis, which are found to be 5.8′ and 4.0′, repectively (Moshir *et al.*, 1992). The parameters p, ρ_x and ρ_y may be adjusted to fit the observed flatness near the top of the point spread function. By reguiring the extended component to be as smooth as possible, a set of the parameters has been determined and the resulting point spread function is shown in the right lower side of Figure 2.

In actual implementations a special care should be taken for the boundary effect arising from the finite extent of the observed images. The discrete Fourier transform tacitly assumes that an image is spatially periodic. If the intensity of an image is uniform near the boundary, the image may be regarded as a part of an intensity distribution which is extended periodically. One can then use the discrete Fourier transform for the convoultion exactly in the same way as used in the infinitely extended sources.

In the presence of an intensity gradient near the boundary as is in our case, the boundary effect is inevitable. In this case the following two techniques may be exploited. The first is to prepare a region of image which encloses the region of our interest. After deconvolution by using the periodic boundary condition, we take only the inner part of the image which is free from the boundary effect. The other one is to modify the discrete convolution to accommodate the finite extent of the image. Here we may use a modified convolution form of

$$z_j = \frac{\sum_i x_i f_{ji}}{\sum_i f_{ji}} \tag{17}$$

as suggested by Aumann et al(1990) where the summation is over the pixels within the image. The weight factor $w_j = 1/\sum_i f_{ji}$ has the value of unity at the inner region, while it becomes larger near the boundary. This approach may be useful when the images encompassing the region of interest is not available. In the present work, we have chosen the first approach because it is easy to implement and larger images are available to us.

Figure 3 and Figure 4 show the 60μm and 100μm images reconstructed by means of our power entropy scheme with the use of the dimensionless regularizing parameters $\tilde{\alpha} \equiv \alpha N\sigma^2 = 10^{-3}$, 10^{-4}, 10^{-5} and 10^{-6}, where we have taken sufficiently small value $m_n = 10^{-20}$. As can be seen from the figures, the reconstructed images look very smooth without artifacts or noises when $\tilde{\alpha}$ is taken sufficiently large. As $\tilde{\alpha}$ decreases, fine details appear but regretably accompanying artifacts. It still remains a problem of how one determines an optimal value of the regularizing parameter.

Figure 5 illustrates a relation between the regularizing parameter $\tilde{\alpha}$ and the root mean square value of residuals given by

$$RMS = \sqrt{\sum_n |Y_n - F_n X_n|^2/(N\sigma^2)}. \tag{18}$$

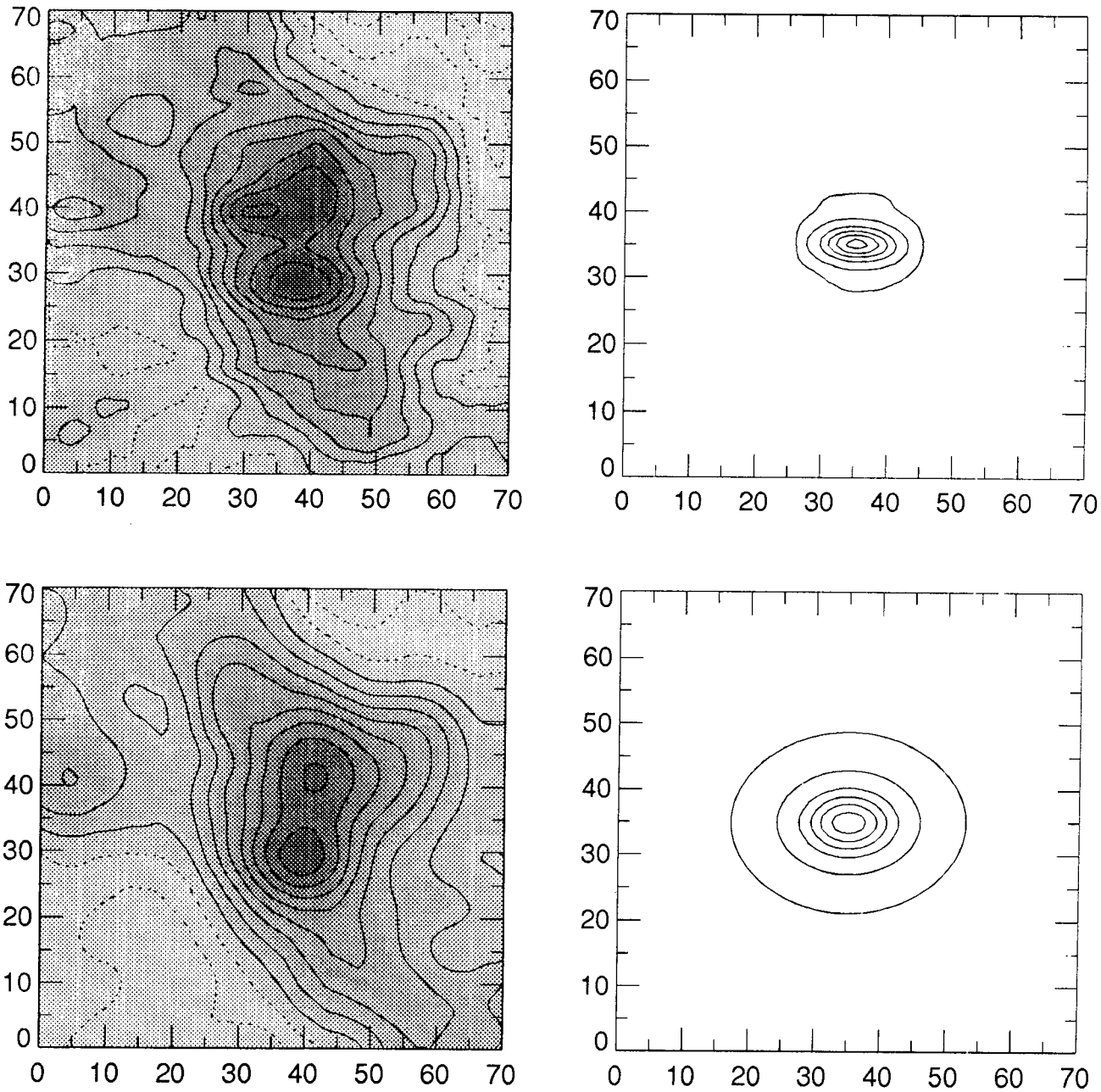

Fig. 2. The regions of our interest along with their point spread functions. From the left upper side are shown $60\mu m$ coadded image, the point spread function associated with it, $100\mu m$ coadded image, and its point spread function. The contour levels in the two coadded images are the same as in Figure 1. Drawn are contours of 0.01, 0.1, 0.3, 0.5, 0.7 and 0.9 of each maximum in the point spread functions.

It is evident that the upper limit of $\tilde{\alpha}$ is 10^{-3}, since any value exceeding this makes the data under-fitted. The most appropriate choice for a set of $\tilde{\alpha}$ values can be found from the qualitative comparison between the two band images. In the present work, we have selected a set of $\tilde{\alpha}$ values in such a way that the resolved feautures should match well between the $60\mu m$ and $100\mu m$ images. They are found to be $\tilde{\alpha} = 10^{-4}$ for the 60 μm image and 10^{-5}, for the 100 μm.

Artifacts seen near the removed point source region such as the fringe pattern in the $60\mu m$ image and the hole in

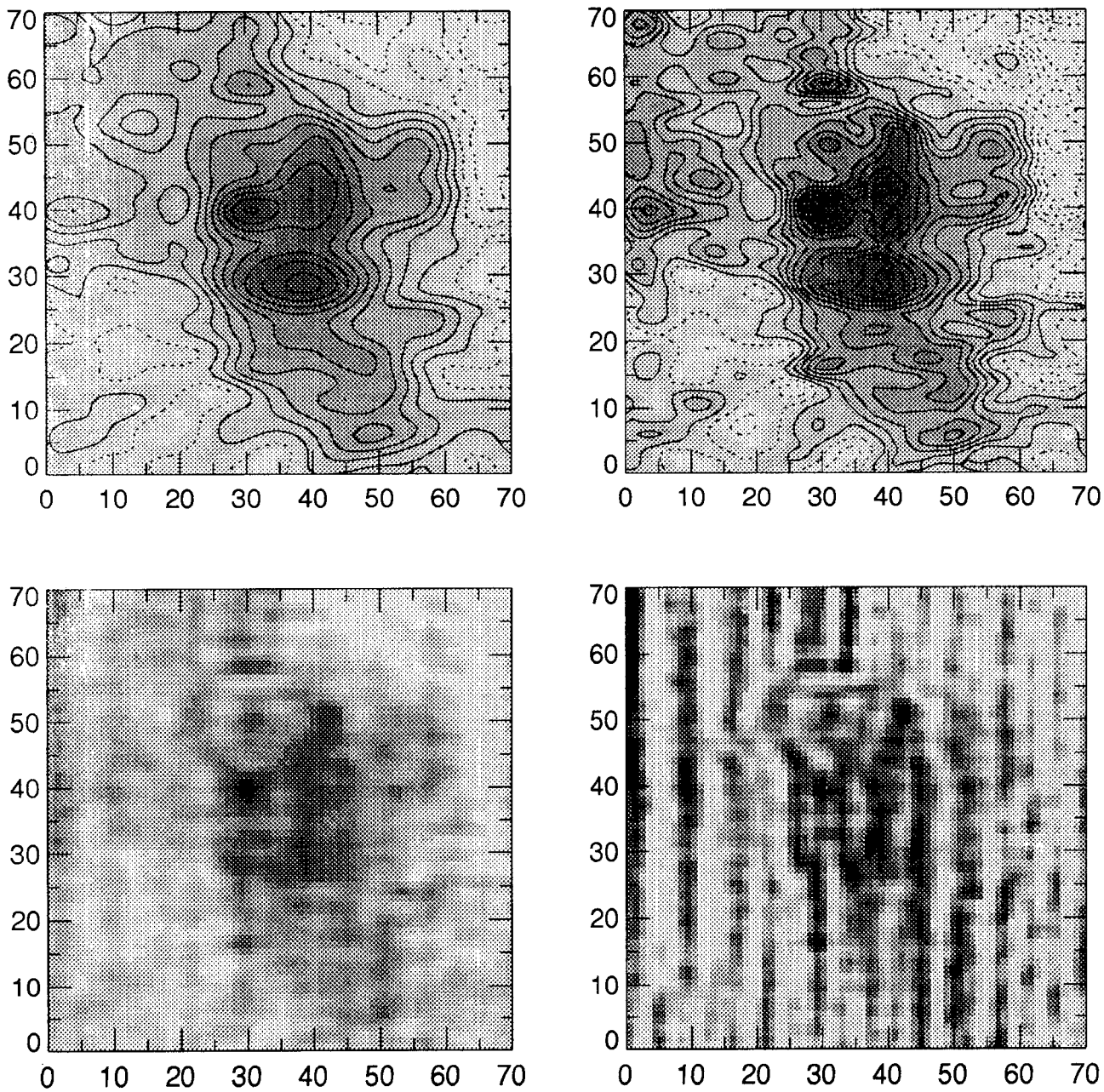

Fig. 3. The reconstructed $60\mu m$ images. From the upper left side are the images reconstructed with the use of $\tilde{\alpha} = 10^{-3}$, 10^{-4}, 10^{-5} and 10^{-6}. The contour levels are the same as in Figure 1. $\tilde{\alpha} = 10^{-4}$ has been chosen as an optimal value. See text.

the $100\mu m$ have been originated from the error introduced by the process of separating the point source component from the extended source component. The reconstructed images clearly show that the cloud is composed of four or more clumps. The clump located just left from the image center (x=40, y=40) displays rather a large value of the intensity ratio, I_{60}/I_{100}, suggesting that it may be a hot point-like source such as H II regions. To clarify their detailed structures, it may be desirable to study optical depth and temperature maps deduced from these two band images.

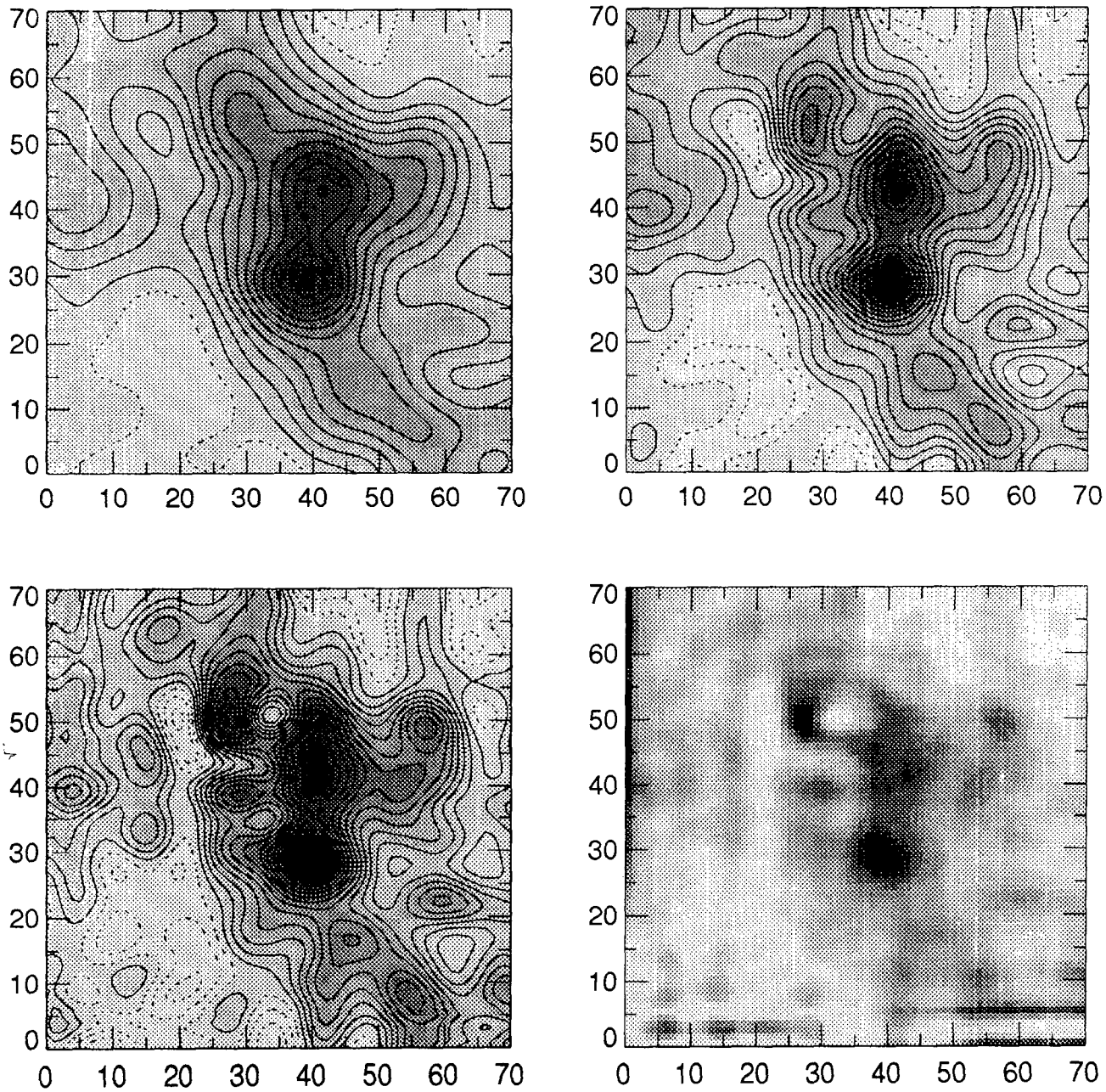

Fig. 4. The reconstructed $100\mu m$ images. From the upper left side are the images reconstructed with the use of $\tilde{\alpha} = 10^{-3}$, 10^{-4}, 10^{-5} and 10^{-6}. The contour levels are the same as in Figure 1. $\tilde{\alpha} = 10^{-5}$ has been chosen as an optimal value. See text.

(b) Deconvolution of Solar Images

Solar images are usually characterized by the presence of the bright background mixed with dark absorption features as can be seen in Figure 6. The image shown in the figure has been obtained from the CCD observations made with the use of Fe I 6302.5 filter(bandwidth 125 $m\AA$) by Japanese Solar Flare Telescope. One pixel size is $0''.67$. The image appears to be highly blurred due to atmospheric turbulence, guiding errors, off-focusing *etc.*. Since there are no point sources in solar images, the point spread function(PSF) can not be found directly from the observed images. In the present work, we have exploited the power spectra of the solar granulations observed near the disk center to derive the PSF. Since the convolution of any two functions is presented by the product of their

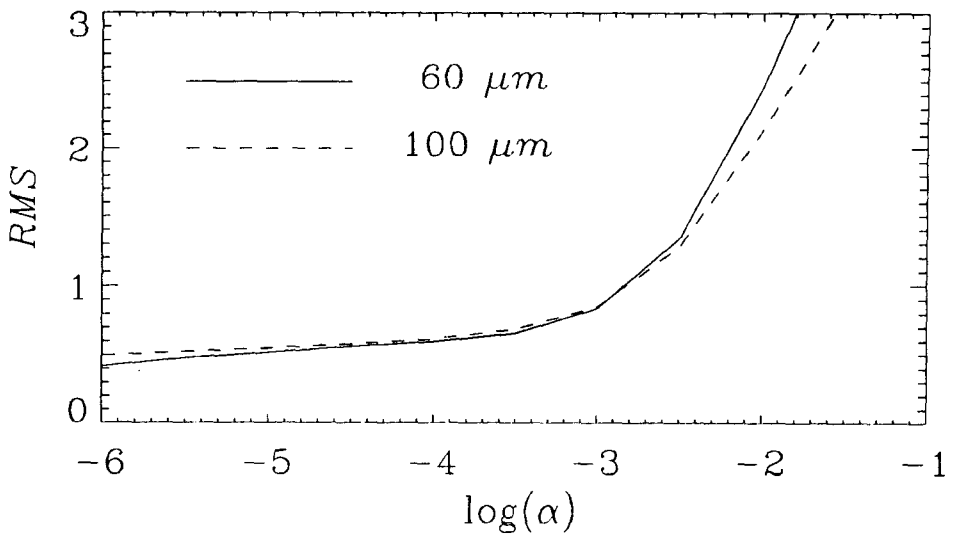

Fig. 5. A relation between $\tilde{\alpha}$ and RMS of residual.

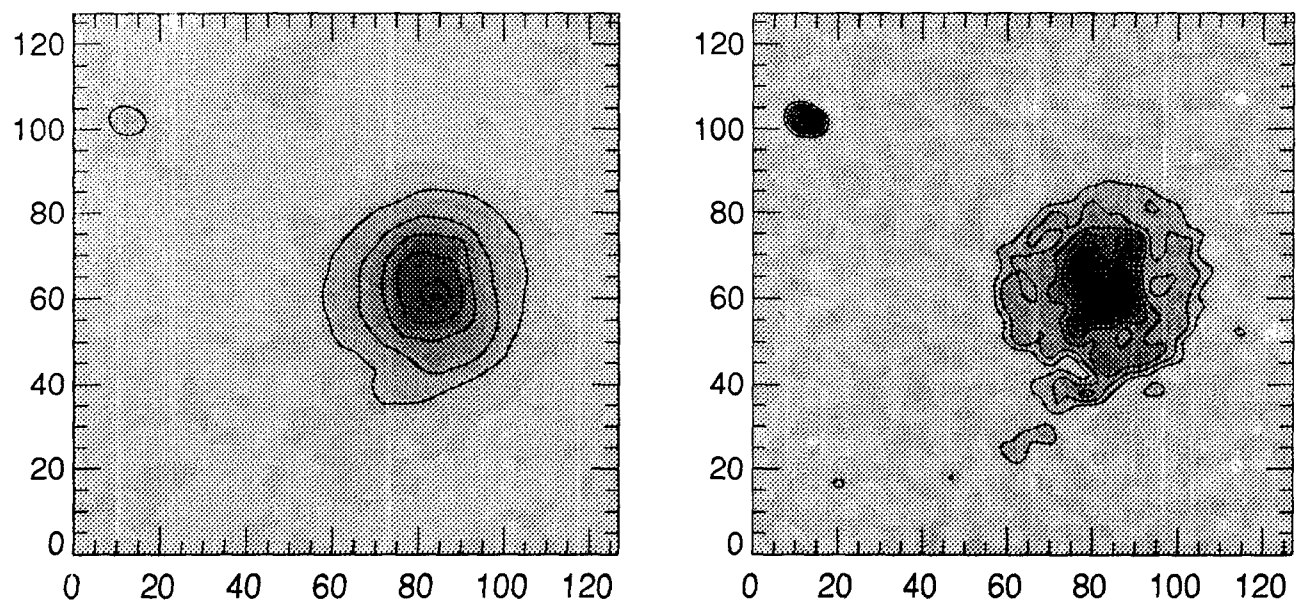

Fig. 6. Observed and reconstructed images of a solar active region containing a circular spot and a small pore. Contour levels 0.1, 0.2, .., 0.8, and 0.9 of the photospheric intensity are drawm.

Fourier transforms, the observed power is given by

$$P^{ob} = P^{true}|\mathrm{FT}(PSF)|^2 \tag{19}$$

from which we can get the PSF

$$PSF = \mathrm{FT}^{-1}(\sqrt{P^{ob}/P^{true}}) \tag{20}$$

where we assumed that the PSF is symmetric about the origin. Since the observed power is noisy, an analytical model of the PSF has been used to fit the data. We adopted a linear combination of two elliptical Gaussian functions. The FWHM of the PSF for this image is found to be 3.5″.

In the reconstructed solar image, Granulations are clearly visible and the boundary between the sunspot and surroundings becomes sharper. Penumbral region is rather inhomogeneous as it should be, noting the existence of the penumbral fibrils. The umbral region turns out to be composed of several dark fragments. One of the most important findings is the severe reduction of the brightness of the pore as seen from the figure. This suggests that the image of the small sized pore have been highly degraded by blurings. It is noted that the brightness of the pore is almost comparable to that of the sunspot after the reconstruction. This leads us to believe that the maximum power entropy method is quite powerful for the scattered light correction which is crucial for quantatitive works in solar physics.

IV. DISCUSSION

We have introduced the entropy of power spectra to define the prior probability for the deconvolution of low contrast extended images. By maximizing the *a posteriori* probability which is proportional to the product of the priori probability and the likelihood function, we were able to derive a semi-analytic solution of deconvolution. The computing time of this maximum power entropy solution becomes much smaller than that required by the interation schemes employed by the conventional maximum entropy method. The smoothness requisite for the extended images is found to be insured by increasing the regularizing parameter. But if we want to get a smooth image, we cannot but suffer from the loss of real structures as can be seen from Figure 3 and Figure 5. Alternately, if we want to get an image having many real structures, we are bound to have many artifacts. An ideal deconvolution method, however, should be able to reveal real structures and reduce artifacts.

The maximizaion of the entropy in the image reconstruction is equivallent to the minimization of structures. Therefore the negative value of the entropy (negentropy) may be considered as a measure of (structural) information. If it is intended to reduce artifacts by minimizing the negentropy, the negentropy should increase only with the amount of the artifact structural information. In addition, the real structures should contribute to the negentropy as little as possible.

Let us consider two images with a high and a low constant background. We hope to make the high and low background structures contribute to the negentropy as little as possible Then the difference in the negentropy between the two images also should be small as compared with the total negentropy. The conventional intensity entropy defined in the spatial domain fails to satisfy this requirement, because all the pixels in the spatial domain contribute to the total negentropy. The total negentropy of the image with the high background is much greater than that with the low background. This means that it becomes more difficult to discriminate true structures from artifacts. This is the reason why the conventional intensity entropy fails to suppress the ripples around a point source embedded in the high background. However, the power entropy defined in the frequency domain satisfies this requirement. In this case only the pixel located at the origin of the frequency domain contributes to the total negentropy because the power of constant level image is given by a delta function located at the center. Since the real structures contribute little to the negentropy, we can reduce artifacts only by maximizing the total power entropy. The same conclusion may be drawn for more general cases of extended sources. However, if the image is composed of point sources, all pixels in the frequency domain contributes to the total power negentropy so that the power entropy can no longer discriminate between true structures and artifacts. Thus we suggest to employ the combination of the power and the intensity entropy for the reconstruction of images when the images are made of both point sources and extended sources.

Almost all of the deconvolution schemes with a regularizing term have a common problem of determining the regularizing parameter. The optimal value is taken as the value which makes RMS of residulas equal to or a little smaller than unity. However, in many cases either the noise characteristics or the standard deviation of noises is not known exactly so that it is not always possible to estimate for RMS . In such cases subjective judgments should be made for the choice of the parameter $\tilde{\alpha}$ based on any *a priori* information. We have obtained several deconvolved images using different values of the regularizing parameter. Since images taken at several wavelength bands are available, the values of the regularizing parameter can be easily found by imposing that the revealed features in the two reconstructed images should be similar. If the images containing the same signals but with different random noises are available through successive observations, the cross-validation test developed by Nunez and Lacer(1993)

may also be utilized to fix the parameter.

ACKNOWLEDGEMENT

One (Chae) of us is indebted to Prefessor S.S. Hong, Professor M.G. Lee and Mr. H.G. Kim for helpful discussions on the image deconvolution process. He also wish to thank all members of Solar Division, Japanese National Astronomical Observatory for providing us the Solar Flare Telescope data and their hospitality during his stay in Mitaka.

REFERENCES

Aumann, H.H. Fowler, J.H. Melnyk, M. 1990, *A.J.*, **99**, 1674.
Deiego, F. 1985, *PASP*, **97**, 1209.
Gull, S. F. 1989, in *Maximum Entropy and Bayesian Methods*, pp 53-71, eds. J. Skilling.
Moshir, M. et al. 1992, *Explanatory Supplement to the IRAS Faint Sorce Survey*, v.2.
Narayan, R. Nitayanada, R. 1986, *Ann. Rev. A. Ap.*, **24**, 127.
Nunez, J. Llacer, J. 1993, *PASP*, **105**, 1192.
Skilling, J. 1989, in *Maximum Entropy and Bayesian Methods*, pp 45-52, eds. J. Skilling.
Skilling, J. 1990, in *Maximum Entropy and Bayesian Methods*, pp 341-350, eds. P.F. Fougere.