Astron. Astrophys. 137, 159–165 (1984)

ASTRONOMY AND ASTROPHYSICS

# Enhancements to the deconvolution algorithm "CLEAN"

D. G. Steer[1,*], P. E. Dewdney[1], and M. R. Ito[2]

[1] Dominion Radio Astrophysical Observatory, Herzberg Institute of Astrophysics, P.O. Box 248, Penticton, B.C., Canada, V2A 6K3
[2] Department of Electrical Engineering, University of British Columbia, Vancouver, B.C., Canada, V6T 1W5

Received December 27, 1983; accepted February 16, 1984

**Summary.** We describe changes to the deconvolution algorithm CLEAN, which are effective in preventing the formation of "stripes" or "corrugations" during the processing of images with extended features. The modifications also reduce the arithmetic rounding problems that may be encountered when a very large number of subtractions are performed with limited precision. The major modification is to remove components in groups rather than individually. This is very effective in preventing the formation of the stripes. A minor change to the order in which the operations are performed reduces the loss of accuracy due to the effects of the repeated subtractions. Because the components are identified and removed in groups, there is a significant speed advantage for large maps. The number of iterations is also relatively independent of the complexity of the object distribution, and there are fewer "control" parameters to be set by the user. This algorithm will be of particular interest to those who process large images with extensive extended areas of brightness.

**Key words:** radio astronomy – data analysis – CLEAN technique – image enhancement

## I. Introduction

Aperture synthesis radio astronomy made a significant step forward with the introduction of the CLEAN algorithm by Högbom (1974). This algorithm is a solution to the problem of deconvolving the synthesis telescope's point response function, or beam, from the image. Before CLEAN, the formation of images with interferometers was practical only with (nearly) fully sampled apertures so that the beam sidelobes were well outside the main image area. With apertures that are not fully sampled, the sidelobes may lie inside the image area of interest, and the beam must be deconvolved before the image can be studied. The CLEAN algorithm is a practical solution to this problem and it is now in use with almost every major synthesis instrument. The algorithm has proven to be straightforward to implement, to work effectively, and to be robust in operation with noisy data. Schwarz (1978) has shown that in spite of its simplicity, the algorithm is an optimum least squares fit of sine functions to the visibility data.

At the time of CLEAN's development, aperture synthesis instruments were of limited sensitivity and most of the images were easily represented by an empty field with a few compact bright sources. With the development of better instruments and bigger computers, the algorithm is now being applied to quite different images which include areas of extended brightness. The deconvolution of these images requires a great many ($10^4$) iterations and there may be two types of failure: often the extended areas of brightness are reconstructed as a series of ridges (stripes or corrugations), and there are sometimes numerical failures due to the very large number of subtraction operations being performed with limited precision. While Schwarz (1983) has shown that the stripes correspond to visibilities outside the range of measurement, and thus a striped image is consistent with the data, the stripes are nonetheless inconsistent with the astronomer's (current) view of the universe. The fabric of the celestial sphere was not cut from striped cloth! As the CLEAN algorithm is able to perform the deconvolution successfully for narrow sources, it seems reasonable to expect that it also be consistent with the constraint that extended sources are not striped. Segalovitz and Frieden (1978), Palagi (1982), Cornwell (1983), and Schwarz (1983) have recently discussed the problem and some possible remedies. In this paper we will review the algorithm and the failure mechanism. The methods of correcting the problem will be discussed and we will present a new, very simple, remedy which we believe gives superior results. An example image will illustrate the effectiveness of the modified algorithm.

## II. Cleaning, and stripes

The basic CLEAN algorithm is an iterative solution to the (classic) deconvolution problem

$$D = B_d * O(+\text{noise}), \tag{1}$$

where $O$ is the object brightness distribution (on the celestial sphere), $B_d$ is the instrument point response function or "dirty beam", $D$ is the "dirty image" which is obtained as the Fourier Transform of the observed visibility data, and $*$ denotes the convolution operation. In practice the three functions are all sampled and are of limited extent. An important feature of the beam function, $B_d$, is that it consists of a prominent main lobe surrounded by a pattern of smaller sidelobes. This has the practical consequence that the position of bright features in the object will be preserved in the dirty image. The objective is to recover the object brightness, $O$, given the dirty image and the dirty beam. The

---

*Send offprint requests to:* P. E. Dewdney

* Current address: Bell-Northern Research Ltd., 7M50, Switching Division, P.O. Box 3511, Station "C", Ottawa, Ontario, Canada, K1Y 4H7

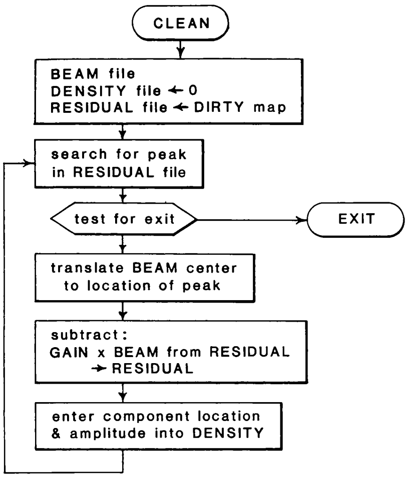


**Fig. 1.** Flow-diagram of the "standard" CLEAN algorithm

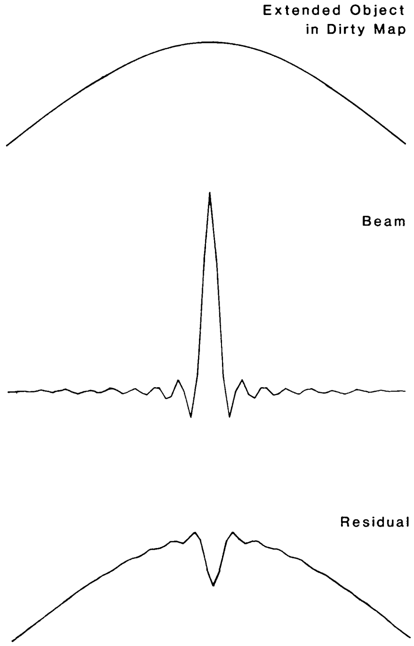


**Fig. 2.** The impression of ripples by the beam into an extended source. The residual after subtraction of the beam (with a loop gain of 0.25) has new false peaks 3 points either side of the center

problem is not straightforward however, as the beam function, $B_d$, does not have a convolutional inverse (due to the many zeros in the sampling function), and the data are noisy. The problem is solved in practice by the inclusion of constraints on the image brightness. The major constraints are that the sky is largely empty (below the noise level) and that any features are compact and few in number. The desired solution is a set of components, or a collection of points of brightness, referred to as the density file, $C$. The requirement is that this density file, when convolved with the dirty beam, fits the dirty image within the noise level. That is:

$$D \approx C * B_d . \qquad (2)$$

The process by which the density file, $C$, is created by the "standard" CLEAN algorithm (Högbom, 1974) is outlined in Fig. 1. This involves the repeated subtraction of the dirty beam from the location of the current brightest element in the "residual map", $R_f$. The residual map is initially set equal to the dirty map and it collects the remainder after each subtraction. The "loop gain" parameter, $g$, is used to scale the beam before each subtraction. The density file components are modified delta-functions which are placed in the corresponding location where each beam pattern is subtracted from the residual map. The amplitude of each delta-function denotes the amplitude of the beam removed, thereby recording the location of each component hidden in the object. The subtraction process continues until the residual map is smooth and contains no more major peaks, or until the noise level is reached, or until the computing budget is depleted.

The accumulated density file is then convolved with an arbitrary restoring beam, $B_r$ (the "clean beam" which has no sidelobes), to yield a restored image, $I_r$, free from the confusing effects of the beam sidelobes and representing a reasonable estimate of the object given the resolving power of the telescope:

$$I_r = C * B_r . \qquad (3)$$

The final residual map is usually added to this restored image to keep the noise level realistic in the final image:

$$I_f = C * B_r + R_f . \qquad (4)$$

The final residual, $R_f$, will be scaled slightly to account for differences in area between the dirty and the restoring beams. The restoration operations of convolution with the restoring beam and addition of the residuals are not shown in Fig. 1.

When confronted with a dirty image which includes broad (extended) features the algorithm does not perform well and often turns the smooth extended areas into a series of ridges or stripes (Schwarz, 1983). The mechanism for the formation of these stripes is shown in Fig. 2. When the beam pattern is subtracted from the peak of a broad feature in the dirty map, the effect is to superimpose the pattern of the sidelobes onto the previously smooth feature. Thus, when looking for the next component in the residual map, the impressed ripples are found as the peak locations, and the result is a density file with components spaced apart at the interval of the sidelobes. It is these separated components in the density file that produce the ridges in the restored map. The problem is especially serious if the beam is heavily oversampled and thus the "sidelobes" (in this case the points adjacent to the beam center) are large and make a significant impression on the residual map. If the image were noise-free, and the CLEANing operations could proceed for a great many iterations without loss of accuracy due to arithmetic precision, then the algorithm would eventually fill in the holes to yield a smooth reconstruction. However, the data are always noisy, and the arithmetic is done to limited precision, and so the process is usually terminated before the holes can be repaired.

There are a number of ways to address this problem. One approach notes that it is the sidelobes which lead the algorithm astray, and a solution lies in making them less prominent. This can be done by reducing the loop gain or by altering the dirty beam. One way to alter the beam is to add an artificial "spike" or "delta-function" to the center of the main lobe. This has the effect of suppressing the sidelobes relative to the center, and cleaning with such a "spiked" beam does indeed reduce the effect of ripples in the image. Cornwell (1983) provides a relation for determining the

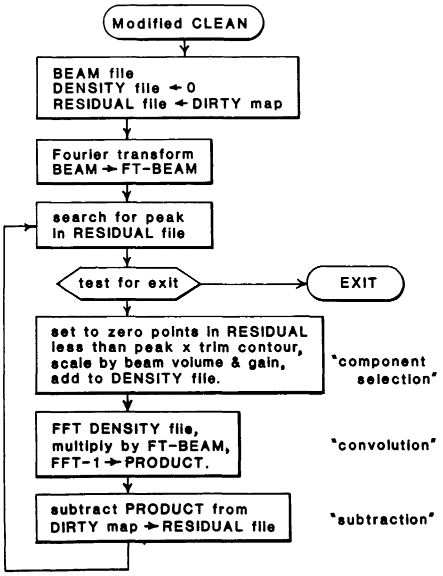


**Fig. 3.** Flow-diagram of the "modified" CLEAN algorithm that prevents the formation of stripes in extended objects. As explained in Sect. IV, the scaling by the beam volume indicated in the "component selection" box involves an additional FFT pair

spike size and illustrates maps successfully deconvolved with this procedure. There are two difficulties with this solution. Because the spike has the effect of reducing the loop gain, more subtraction operations will be required than with an unmodified beam. Also, as the sidelobes are now scaled improperly, there will be a higher residual sidelobe level in the final map, and the restored image may need to be scaled slightly. If the spiked beam is used only to *locate* components in the "minor" iteration cycles of the Clark (1980) CLEAN algorithm but is not used for the subtraction in the "major" cycles, then these difficulties are alleviated.

Cornwell (1983) and Palagi (1982) also discuss other methods of preventing the stripes by incorporating algorithms which directly suppress the stripes. This may take the form of alternating the CLEAN iterations with a least squares model fit or a solution by the maximum entropy method (Gull, 1978). These approaches can be effective, but they spoil the attractive simplicity of the basic CLEAN technique.

## III. Arithmetic errors

The standard CLEAN algorithm repeatedly subtracts the dirty beam from the dirty map, and numerical failures can result from the accumulation of errors which occur due to the finite precision of the arithmetic. For example, given a sidelobe level of 1% of the peak, and a gain of 0.1 with arithmetic performed with five significant figures, then one significant figure of the beam is lost in each subtraction. As the truncation is always in the same direction, after approximately $10^3$ subtractions, the accumulated error will be as large as the sidelobe. This is the danger in CLEAN of choosing a small loop gain and thereby increasing the number of subtraction operations. These errors effectively add a "noise" to the residual map in direct proportion to the number of subtractions performed and will have the greatest effect on a weak, small-angular-size source in the vicinity of a strong source. These errors are cancelled somewhat when positive and negative sidelobes are combined at one spot. However the overall effect is to add a distortion to the residual map which will cause the algorithm to terminate earlier than it would otherwise with low noise data, and to bias the final map when the residuals are added in (Chen, 1981).

## IV. The modified algorithm

The standard assumption, implicit in the CLEAN algorithm, is that the peak element in the residual map represents the location of a single component hidden in the image. For a map consisting mainly of isolated point sources this is a reasonable assumption. For a map with extended sources however, this is not reasonable and the algorithm can be led astray. The "obvious" solution, therefore, is to estimate the area of the source surrounding the peak in the residual map and remove, not a single component, but a feature of corresponding area. Since most of the components in an extended source will be removed at once, the formation of stripes will be prevented.

The hurdle to be overcome to implement this modified algorithm is to estimate the group of components surrounding the peak. Figure 3 shows a diagram of the algorithm we use to perform the modified CLEAN operation. The group of components is chosen by simply including all those points in the residual map that rise above a contour set at some fraction of the peak level. We call this the "trim contour", $T_c$. Note that if the residual map includes a number of separated peaks of comparable brightness, then the trim contour will select several isolated "islands" of components.

The components, selected from the dirty map in this way, must be corrected for the presence of the dirty beam before they can be added to the density file as components of the object. The dirty beam, because of its width, acts non-uniformly on the components of the object. A single isolated component (point source) will have its value multiplied by the *peak* of the beam (which for convenience can be normalized to one). A component that is part of a group, however, will have its value multiplied by approximately the *volume* of the beam. The group of components selected by the trim contour must be scaled so that the value of the peak after convolution with the dirty beam is the value of the corresponding point in the residual file. This scale factor is found by convolving the current selection of components with the dirty beam, and thus an additional Fourier transform pair must be performed. This process is part of "component selection" in Fig. 3.

These points are then further scaled by the CLEAN loop gain, and added to the components already accumulated in the density file (Fig. 3). This density file is then convolved with the dirty beam to obtain the pattern to be subtracted from the dirty map. Note that this convolution will restore the scale factor of the beam volume removed when the components were selected. The convolutions with the dirty beam are performed via the Fourier domain. After the subtraction, the (new) residual map is searched for the components surrounding the peak, and the estimation and subtraction operations are repeated.

The numerical problems of repeated subtraction are diminished by always subtracting the *complete* current best estimate of the components from the *original* dirty map. We have chosen to do the subtraction in the image domain for convenience in our programming. It could equally well be done in the Fourier domain from the gridded visibility data.

The trim contour method to estimate the group of components is effective for the simple reason that, given an extended object region and a narrow beam, the convolution of the two resembles the region more than the beam. Thus the central area of the actual extended feature in the dirty map is a reasonable (first) estimate for the shape of the group of components. Any incorrect component amplitudes are corrected in later iterations in the same way as in the standard CLEAN process. Note that if the peaks in the residual map are not in extended regions, then the trim contour will enclose only small areas and once again the effect will be equivalent to the treatment of point sources given by the standard CLEAN algorithm.

At each iteration the modified algorithm must scale the amplitude of a group of components to be subtracted. The group of components is first convolved with the dirty beam, and then the scale factor is set so that the peak in the convolved group is the loop gain times the peak in the residual map. The peak in the group of convolved components may be either a point source or the peak of an extended feature. If the image is made up entirely of extended objects or entirely of point sources, then all the components in the group at each iteration will have the same effective loop gain. However, if there are both point and extended sources of comparable amplitudes, point sources receive a smaller gain. The narrow sources, therefore, are removed quite slowly from the dirty map. The action of the algorithm is thus to remove first the bright, extended areas leaving the point sources and the low-level, extended features for later. We have, therefore, added to the algorithm a method for processing narrow sources with a gain equivalent to the wide ones: the program tests the brightest pixel in the component group at each iteration for the presence of a point source. If a point source is detected, then the gain for that individual pixel is increased to the full loop gain. If several point sources are present in the object, together with extended features of comparable amplitude, then the point sources are processed sequentially, one per iteration. The test for the presence of a point source is done by comparing the location of the peak in the group of components with the peak after convolution with the dirty beam. If the locations are not the same, a point source is assumed, and the gain is adjusted at that location. In this process at no time does the gain exceed the specified CLEAN loop gain and, therefore, over-subtraction does not occur. This addition to the algorithm greatly improves the speed of convergence.

This modified algorithm has some resemblance to the Clark (1980) version of the CLEAN algorithm. In that algorithm the individual components are identified in a similar way to the standard algorithm. An improvement in speed is obtained, however, by subtracting only the central region of the beam ("beam patch") to estimate the residual map in a series of "minor cycles". After a number of components have been estimated in this way, they are removed accurately in a "major cycle" using a convolution and subtraction in the Fourier domain. However, as the "beam patch" includes the major sidelobes, the algorithm will still be led astray when working on extended features. Also, the beam is still being removed incrementally and numerical problems can still occur. Our modified algorithm has the effect equivalent to working with a beam patch one element in size. The trim contour is also a more efficient process to select the components than searching for individual peaks.

Because the new algorithm removes many components in each iteration, the number of operations is reduced significantly. For the standard CLEAN, the number of operations is approximately

$$N_s \approx 2Q_s M, \tag{5}$$

where $Q_s$ is the number of iterations, and $M$ is the (total) number of pixels in the image. For the modified algorithm, the number of operations is approximately

$$N_m \approx Q_m M(4\log(M)/2+4), \tag{6}$$

where $Q_m$ is the number of iterations. [$4M\log(M)/2$ operations are required for the 4 FFT's, (this number will depend on the details of the FFT implementation) and $4M$ operations are required for the add to the density file, the complex multiply, trimming and subtraction.]

For the Clark algorithm, the number of operations is approximately

$$N_c \approx Q_c(2M\log(M)/2+2M+n(a+1)), \tag{7}$$

where $Q_c$ is the number of major iterations, $n$ is the number of minor iterations, and $a$ is the area of the beam patch (in pixels).

To be able to relate the number of operations in the algorithms, the iteration counts $Q_s$, $Q_m$, and $Q_c$ must be determined. In the basic CLEAN subtraction operation, the residual peak is

$$P_{n+1} = (1-g)P_n, \tag{8}$$

where $g$ is the CLEAN gain and $P_n$ is the peak after the $n^{\rm th}$ subtraction. After $n$ iterations at one point, the fraction, $f$, of the original peak remaining is

$$f = (1-g)^n. \tag{9}$$

From this we have determined that for the standard algorithm the number of iterations is approximately

$$Q_s \approx B\log(fP_{\max}/G)/\log(1-g), \tag{10}$$

where $B$ is the area of the source in units of beam areas, $P_{\max}$ is the initial peak value, and $G$ is the geometric mean amplitude of the peaks to be cleaned. For the Clark algorithm, $Q_c$ is just $Q_s/n$. For the modified algorithm $Q_m$ is approximately

$$Q_m \approx \log(f)/\log(1-g_v). \tag{11}$$

where $g_v$ is the gain reduced by the effect of the beam volume. This is valid only when $g < 1 - T_c$. Thus the approximate ratio of the number of iterations is

$$\frac{Q_m}{Q_c} \approx \frac{n}{B}\frac{\log(f)}{\log(fG/P_{\max})}, \tag{12}$$

which indicates that the new algorithm will be significantly faster for sources covering a large fraction of the image.

The new algorithm requires the trim contour level to be set as a fraction of the beam peak. Like the loop gain in the CLEAN algorithm, the results are not particularly sensitive to the parameter setting. It is clearly desirable to choose a trim level as low as possible, as this will increase the number of points to be subtracted in one iteration. Clearly also, a very high trim contour will be equivalent to the standard CLEAN. However, choosing too low a trim contour will result in sidelobe features being incorrectly included as components. As in the standard CLEAN, sidelobe locations incorrectly interpreted as true components are almost always irrecoverable. We recommend the setting of the trim contour, $T_c$, at a level slightly above the largest "sidelobe" in the dirty beam. Note that this "sidelobe" is the largest value of the beam other than the central peak and if, as is often the case, the beam is oversampled, this point may still be on the main lobe. A typical case is shown in Fig. 4. This trim level will allow narrow sources to be deconvolved correctly, and will also allow extended areas to be handled without the development of stripes. The

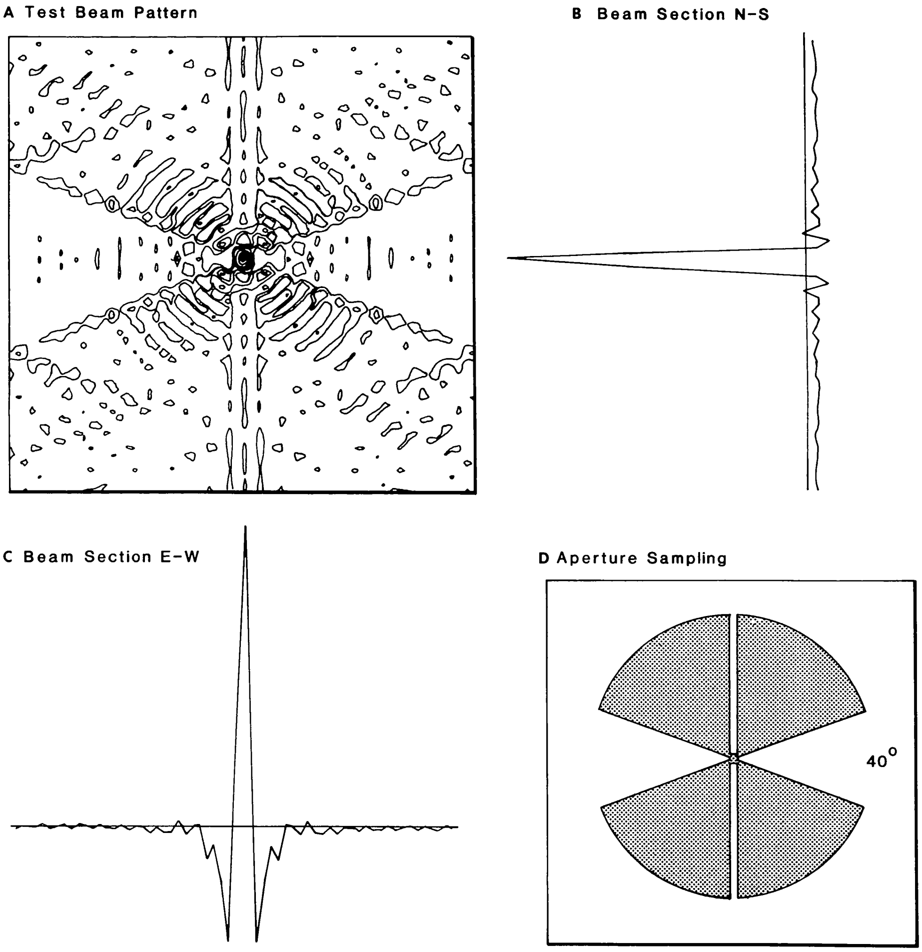


**Fig. 4.** BEAM Test pattern used with the CLEAN algorithms. The cross-sections indicate a positive level outside the peak of 56%, and a negative sidelobe level of 38%. The aperture sampling is circular with a cut-out section of 40°

optimum value of the gain, $g$, for extended sources is slightly less than $1-T_c$. A larger value does not increase the speed of the algorithm, and it may lead to the formation of negative features (except if the image consists entirely of separated, unresolved sources). Since the trim contour may be determined automatically from the sidelobe level, only the CLEAN stopping criterion need be specified by the user.

## V. Example images

Figures 4 and 5 show the results of processing by several CLEAN techniques. The dirty beam is shown in Fig. 4. The cross-sections of the beam in the North-South and the East-West directions indicate that the largest positive point not at the centre is 56% of the peak, and the largest negative sidelobe is 38% of the peak. The aperture sampling pattern corresponds to a circular area with a diameter about two-thirds the extent of the full rectangular aperture. A sector 40° in extent has been zeroed to simulate an unsampled period of observation. The central, or DC, term is included, and the aperture has not been tapered to suppress the sidelobes. Figure 5A shows the test object brightness distribution. All of the points in this subject are positive. This image is $64 \times 64$ pixels in extent and the lowest contour level is about 1% of the highest contour. The dirty map is shown in Fig. 5B. Notice that, while there is considerable confusion due to the beam sidelobes, the extended areas retain much of their general outline, and hence the trim contour is effective in estimating their size. Figure 5E shows the density file resulting from the standard CLEAN, and separated components show up quite clearly in the extended regions. As this is the density file, these would normally be smoothed somewhat in the formation of the restored image by the convolution with the clean beam. Figure 5C shows the density file from the modified CLEAN algorithm, and no separated components are visible. This result was made with a trim contour of 0.55, and with 17 iterations. For reference we also illustrate in Fig. 5D the density file from a CLEAN using a "spiked" beam. This was made with a beam "spike" of 15%, and with 2000 iterations. Neither result has stripes evident, and the images are comparable, although the result from the modified algorithm has a smaller RMS error than the result from the spiked algorithm when compared with the original. This image has no noise.

Figure 6 shows the results of testing with a larger object which also included some noise. Figure 6A shows the object distribution which includes both point sources and an extended area. The object field was $256 \times 256$ pixels in size, however only the centre $64 \times 64$ region is illustrated. Shown in Fig. 6B is the dirty map which was formed when the object was convolved with a beam consisting of a circular "Sinc" function. One percent standard deviation noise was also added to this dirty map. The lowest contour shown is 3% of the peak. Figure 6C shows the density file resulting from the modified CLEAN algorithm while Fig. 6D shows the result from

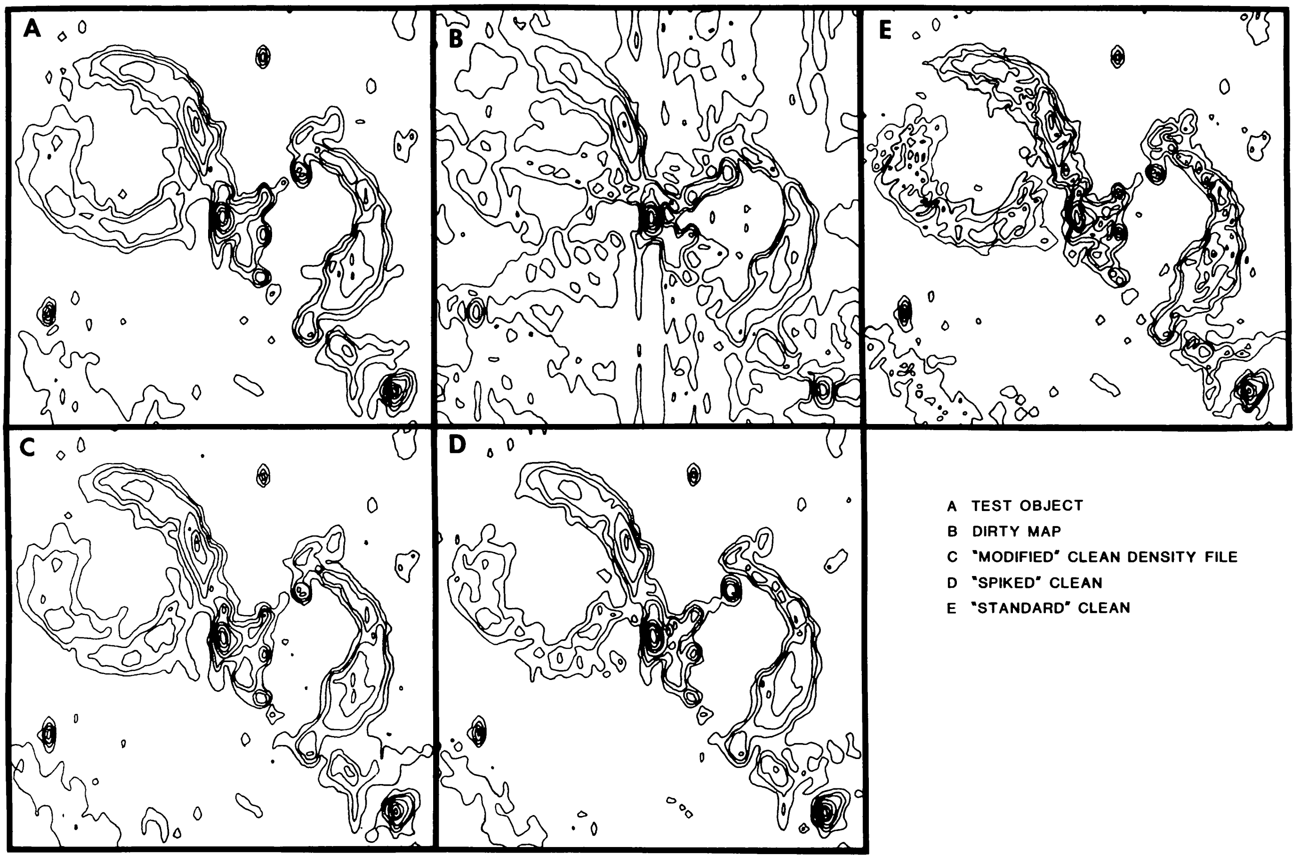


**Fig. 5A–E.** Results of test of CLEAN algorithms: **A** The object distribution, O. **B** The dirty map, *D*. **C** The density file from the "Modified" CLEAN algorithm. The extended areas have been reproduced well. (17 iterations, trim contour 0.55.) **D** The density file from CLEAN with a "spiked" beam. The extended areas have been reproduced smoothly. (2000 subtractions, spike = 15%.) **E** The density file from the "standard" CLEAN algorithm. Note the many separated components. In the case of the standard CLEAN, the density files (deconvolved images) in **C**, **D**, and **E** would be convolved with a "clean beam", free of sidelobes. However, since we are comparing *deconvolution algorithms*, it is appropriate to compare the deconvolved images rather than ones which have been smoothed with a clean beam

the standard algorithm. Although the standard algorithm has correctly reproduced the narrow sources, the extended area is very poorly reproduced.

## VI. Conclusions

The revised CLEAN algorithm has proved to be effective at preventing the formation of stripes and reducing numerical errors in the deconvolution of extended images. The new algorithm also gives good performance with narrow sources. The process is based on the concept of removing components in groups, chosen by means of a trim contour technique. Because the algorithm is stable for extended sources, and components are identified in groups, there is a significant improvement in speed. The new algorithm also has fewer "control" parameters to be set by the user. We believe the modified algorithm will prove to be especially useful to those who process large images with considerable areas of extended brightness.

It is important to realize that a deconvolution process such as CLEAN is actually an interpolation in the Fourier domain. In order to remove the effects of the sidelobes, it is necessary to "fill-in" the unsampled areas of the aperture. This can only be done if there is additional "a-priori" information available about the object. In the case of CLEAN, this is the constraint that the object consists of a limited number of compact areas of brightness. The standard CLEAN algorithm contains the implicit assumption that each peak in the residual is the location of a *single* component in the object. This is equivalent to a directive to the algorithm to produce the narrowest possible features in the image. The result is the generation of stripes. The modified algorithm allows for variable area features at each peak in the residual, and thus the algorithm is not constrained to produce only narrow features and the stripes are avoided. The modified algorithm is able to fill in the unsampled areas of the aperture with more attention paid to the data and less to the implicit assumption in the algorithm.

*Acknowledgements.* This study was made possible through the cooperation of the University of British Columbia and the Herzberg Institute of Astrophysics. The work was performed at the Dominion Radio Astrophysical Observatory near Pentiction. During the period this work was performed, D. G. Steer was on leave from Bell-Northern Research at the University of British Columbia, and he gratefully acknowledges a scholarship from the Natural Science and Engineering Research Council of Canada and

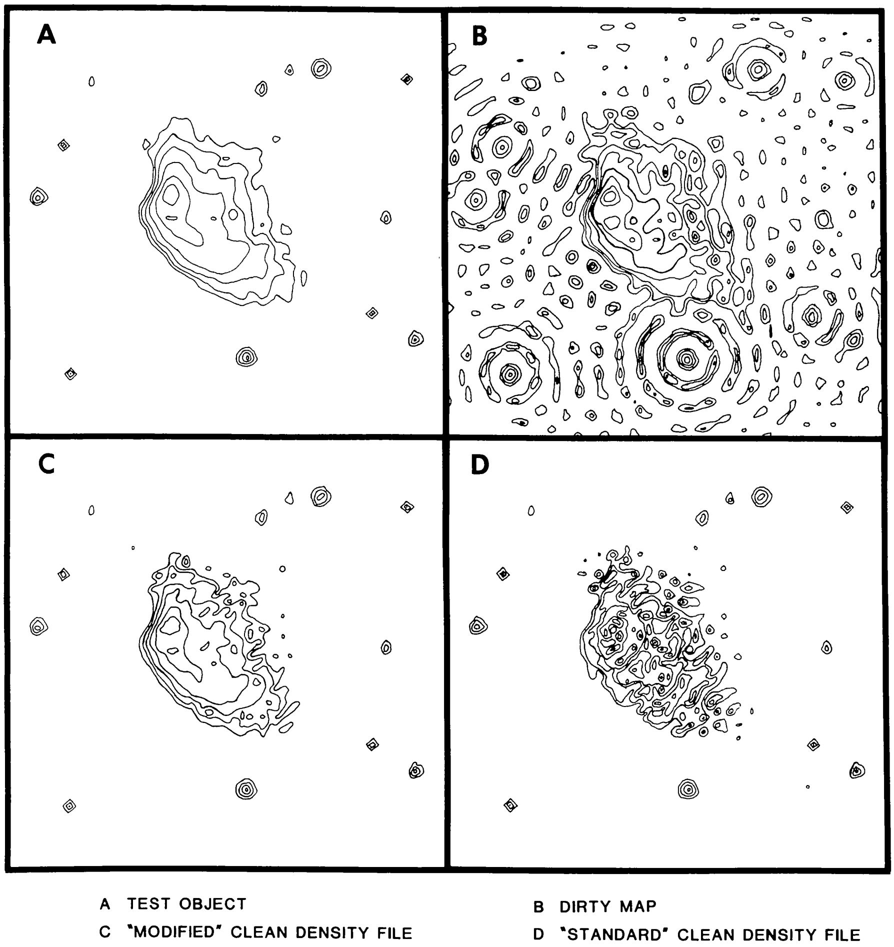


**Fig. 6A–D.** Results of test of CLEAN algorithms: **A** The object distribution. **B** The dirty map with "Sinc" function beam and 1% standard deviation noise. **C** The density file from the "modified" CLEAN algorithm. **D** The density file from the "standard" CLEAN algorithm

the H. R. MacMillan Family Fellowship. Special thanks to L. Higgs, director of D.R.A.O., for providing an efficient version of the standard CLEAN algorithm and to J. Caswell for discussions and for providing the object shown in Fig. 5.

## References

Chen, C.F.: 1981, *Proc. Astron. Soc. of Australia* **4**, 256
Clark, B.G.: 1980, *Astron. Astrophys.* **89**, 377
Cornwell, T.J.: 1983, *Astron. Astrophys.* **121**, 281
Gull, S.F., Daniell, G.J.: 1978, *Nature* **272**, 686
Högbom, J.A.: 1974, *Astron. Astrophys. Suppl.* **15**, 417
Högbom, J.A.: 1983, in *Measurement and Processing for Indirect Imaging, URSI/IAU* Symp., ed. J.A. Roberts Cambridge University Press (in press)
Palagi, F.: 1982, *Astron. Astrophys. Suppl. Ser.* **49**, 101
Schwarz, U.J.: 1978, *Astron. Astrophys.* **65**, 345
Schwarz, U.J.: 1984, in *Indirect Imaging, URSI/IAU Symp.*, ed. J.A. Roberts, Cambridge University Press
Segalovitz, A., Frieden, B.R.: 1978, *Astron. Astrophys.* **70**, 335