J. Astrophys. Astr. (1990) **11**, 311–322

A Modified Algorithm for CLEANing Wide-Field Maps with Extended Structures

K. S. Dwarakanath, A. A. Deshpande & N. Udaya Shankar *Raman Research Institute, Bangalore 560080*

Received 1990 April 5; accepted 1990 May 29

Abstract. A simple but effective modification to the conventional CLEAN algorithm is suggested. This modification ensures both stability and speed when CLEAN is applied to maps containing a mixture of point sources and extended structures. The method has been successfully applied to the recently-completed sky survey at 34.5 MHz (Dwarakanath & Udaya Shankar 1990). This survey was made using the Gauribidanur T array (GEETEE)[1] in 1-D aperture synthesis mode. Since in this case the 'dirty beam' (point spread function) cannot be directly computed, a method to obtain this is discussed in detail. The results of this deconvolution procedure have been encouraging in terms of reduced computing time and improved dynamic range in our maps. This algorithm should find wider application in deconvolving maps which have both extended structures and point sources.

Key words: CLEAN—wide-field map—extended structures—point spread function

1. Introduction

A wide-field synthesis map of the sky at 34.5 MHz has been completed recently by Dwarakanath & Udaya Shankar (1990; DU). In synthesis mapping, some form of deconvolution method is invariably employed to obtain high dynamic range. The conventional CLEAN algorithm (Högbom 1974; Schwarz 1978) is one of the most successful deconvolution methods when the angular extents of sources are not too large compared with the size of the synthesized beam. However, the conventional CLEAN is not particularly well-suited to applications at low radio frequencies, where large-angular-scale Galactic background emission dominates. A number of modified versions of CLEAN (Cornwell 1983; Steer, Dewdney & Ito 1984; Wakker & Schwarz 1988) have been suggested to treat situations where the performance of the conventional CLEAN is not satisfactory. However, in the present case, one is dealing for the first time with wide-field maps ($-50° \leqslant \delta \leqslant +70°$ and all the 24 hours of α) for which it is desired to compensate for the artefacts produced by the sharp cutoff in the aperture coverage without degrading the angular resolution ($26' \times 42' \sec(\delta - 14°.1)$). Unfortunately, the existing procedures do not fulfil our requirements.

[1] This telescope is jointly operated by the Indian Institute of Astrophysics, Bangalore & the Raman Research Institute, Bangalore.

In this paper, we first describe the 34.5-MHz survey (Section 2) and then discuss the difficulties (Section 3) in using the existing deconvolution techniques when dealing with wide-field maps such as result from the survey. In Section 4, we present a modification to the conventional CLEAN algorithm which effectively applies the CLEAN process only to source structures comparable in size to the resolution of the dirty beam. This aspect makes our modified CLEAN a stable algorithm, even with large loop gains which increase considerably the speed of operation. In Section 5, we detail the procedure adopted to generate a point spread function (PSF or dirty beam). Its use in the modified algorithm to CLEAN the wide-field maps is discussed in Section 6.

2. The survey

The celestial brightness distribution, $B(l, m)$, is a function of the direction cosines (l, m) and is related to the visibility $V(u, v)$ through the Fourier transform relation:

$$B(l, m)=\iint V(u, v)\, e^{-j2\pi(ul+vm)}\, du\, dv. \qquad (1)$$

The dirty map, $B'(l_0, m_0)$, at any position in the sky is the convolution of the brightness distribution with the antenna response $P(l, m)$:

$$B'(l_0, m_0)=\iint P(l-l_0, m-m_0)\, B(l, m)\, d\Omega \qquad (2)$$

(Christiansen & Högbom 1985).

The procedure for obtaining the dirty map is discussed in detail by DU.

Using the Gauribidanur telescope (GEETEE) in the transit mode, the visibilities corresponding to the 88 baselines which the East-West (EW) array forms with the rows of the South (S) array are recorded continuously. Assuming Hermitian symmetry (see the discussion concerning 'Tee' and 'Cross' arrays in Christiansen & Högbom 1985) a 1024-point one-dimensional FFT is performed on each set of 88 complex visibilities (resulting from an integration of 24 s) to obtain a strip of dirty map that contains the entire observable sky in the NS direction along the meridian. Successive strips are obtained from successive integrations. The Fourier transform produces dirty maps in the $m=\sin$(zenith angle) coordinate. The 1024 points are uniformly spaced in m and correspond to sufficiently fine sampling of the observed distribution in declination. Since the successive baselines are separated by 5 m and the wavelength is 8.69 m, the FFT produces maps in the range of -0.869 to $+0.869$ in m. The dirty maps are produced in sizes of 1 hr × full range of m for all 24 hours of Right Ascension.

We emphasize that the antenna response in the NS direction is shift-invariant when expressed in sin(zenith angle) coordinates (neglecting the effects of bandwidth decorrelation). Use of this property simplifies the CLEAN procedure.

3. Problems with existing deconvolution procedures

The dirty map is the convolution of the celestial brightness distribution with the antenna response (Equation 2). Unfortunately, the antenna response ('dirty beam' or 'point spread function (PSF)') has undesirable features, *viz.* sidelobes, if the corresponding aperture has any sharp cut-offs. It is desirable to obtain the brightness distribution of the sky corresponding to an antenna whose response is free of such sidelobes but has a similar resolution to that of the true response. Hence, to obtain the brightness distribution in the sky, the observed distribution (the 'dirty map') has to be deconvolved (CLEANed) for the dirty beam and the derived CLEAN components are convolved by a 'clean' (usually Gaussian) beam. To achieve this a two-dimensional CLEAN of the dirty maps can be made. However, we will first examine some simpler algorithms, as these are likely to be computationally more efficient. There are many methods of deconvolution/convolution that may be considered. We discuss some of them here and examine the problems in adopting them for the present survey.

3.1 *Two 1-D CLEAN or Convolution*

Given the dirty beams along hour angle and dec., it is possible to adopt CLEAN (Högbom 1974) or its variant (discussed in Section 4) to deconvolve the dirty maps. However, in our case this procedure cannot be used due to the complex nature of the beams along hour angle and dec., unless CLEAN is performed with complex dirty beams on the complex dirty map. This was not pursued further due to this 'complex' nature.

Methods of deconvolution exist to reduce the sidelobe levels of the dirty beam at the expense of resolution. Despite this loss of resolution, these methods are sometimes used due to their computational ease and efficiency. In our case, this advantage does not exist due to the complexities just mentioned.

3.2 *Difficulties with the Conventional 2-D CLEANs*

In the conventional method of CLEANing maps (Högbom 1974), successive maxima are located in the dirty map and at each stage a scaled version of the dirty beam placed at the location of the maximum is subtracted. The scale depends on the height of the maximum and the fraction of the dirty beam to be subtracted called the loop gain (typically 0.2). Each time a maximum is located, a delta function whose height equals the fraction of the maximum that has been subtracted is placed at the position of the maximum in a separate 'CLEAN map'. This procedure is repeated until a predetermined level is reached in the residual dirty map (usually 5 times the r.m.s. value of noise). At this stage, the CLEAN components are convolved with a well-behaved (or CLEAN) beam, usually a Gaussian whose full-width-half-maximum equals that of the dirty beam. Finally, this distribution is added to the map of the residuals. The implicit assumption in this method is that the sky is empty except for point sources. In addition, the 'emptiness' is a function of the sidelobe level in the dirty beam (Schwarz 1978). In the present case, the dirty map contains both point sources and the Galactic background emission. In addition, the background level varies over the region being CLEANed. This poses problems if one uses CLEAN in its original form. Consider a case such as that shown in Fig. 1 where three sources A, B and C appear against a background. The source A has a height of A_{max} above zero while its height above the local background is only A'_{max}. Conventional CLEAN will pick up A_{max} and proceed. It

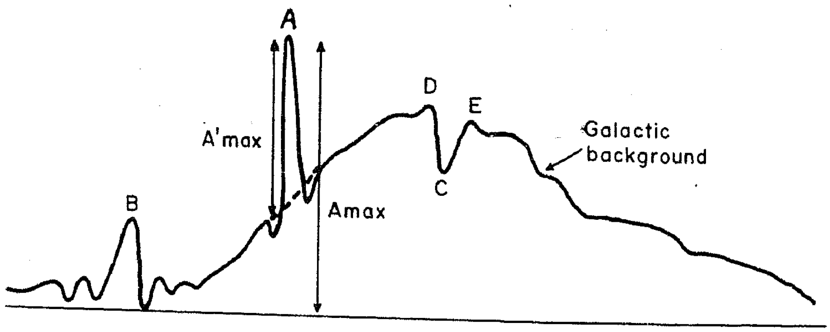

Figure 1. This illustrates the problems in CLEANing maps which contain both point sources and absorptions in the presence of a varying background. The sidelobe level in the map due to source A corresponds to a height A'_{max} and not to A_{max}. It will take considerable time before the dirty map can be deconvolved for responses from sources such as B since the background level is higher than the source height. In addition, discrete absorptions such as C will not be CLEANed—instead, the side lobes at D and E will be treated as sources. The present modification to the conventional CLEAN takes care of all of these problems.

may be seen, however, that the sidelobes due to A do not correspond to A_{max} but to A'_{max}. In order not to oversubtract the source response, a very small loop gain could be used. However, this slows down the CLEAN process considerably. In addition, if the background were to vary over the region being CLEANed, it would take much longer before CLEAN can deconvolve the responses from sources like B or from absorptions like C (Fig. 1). Meanwhile, the algorithm would have selected maxima that represent the variations in the background and carried out the process of CLEAN on these. This is not a desirable situation.

There have been earlier attempts to take account of extended structures while CLEANing. For example, Wakker & Schwarz (1988) have proposed a Multi Resolution CLEAN (MRC). In this algorithm, the dirty map is smoothed to a lower resolution and subtracted from the original dirty map. Thus one produces a 'smooth map' and a 'difference map'. The dirty beams corresponding to these two maps are also generated and CLEAN applied to both maps separately. This makes the algorithm faster than the conventional CLEAN. However, in wide-field maps such as ours, where extended structures have varying angular scales, it is necessary to consider a 'Multi-Resolution' approach rather than a 'Double-Resolution' approach. This makes the MRC algorithm complicated. Further, if the extended structures in the map have widely different angular dimensions in α and δ in different portions of the map (*e. g.* the Galactic plane emission), then it is not clear to us whether such an algorithm would be useful. It would have been possible to incorporate the Clark (1980) approach to CLEAN to enhance the speed of either MRC or conventional CLEAN (with very low loop gain) if we had a limited declination coverage. In the absence of such a situation, the speed enhancement due to Clark CLEAN is not obvious. For the same reason, the algorithm proposed by Steer, Dewdney & Ito (1984) may not be very helpful in our case.

4. Modified CLEAN

To make the basic CLEAN algorithm work satisfactorily in the presence of extended structures, we have introduced a simple modification. Instead of locating the maxima in the dirty map (as with the conventional CLEAN and the MRC), we locate the maxima in the absolute value of the 'difference'. The 'difference' at any given point in the dirty map is obtained by subtracting from its value the average value in the map one resolution away. Using this procedure, the more slowly-varying background is essentially not deconvolved by the algorithm, but all sources with sizes of the order of the beam are CLEANed, including absorption features (*e.g.* feature *C* in Fig. 1) which are quite common at low frequencies. In addition, by not modifying the background, this procedure converges rapidly and a uniform level of CLEANing can be achieved over the region. However, for any deconvolution procedure, the PSF is an essential input (Section 5). Given the PSF, deconvolution can be achieved using the modified CLEAN. This procedure has been successfully applied in CLEANing our maps. Even with a loop gain of 0.5 (usually considered high) we had no difficulty with convergence. Compared to the conventional CLEAN, in regions where there is no significant variations in the background, this algorithm gave a reduction in processing time of at least a factor of three (due largely to the larger loop gain). We believe that the increase in speed using this algorithm is substantially larger than three for regions having significant contribution from extended structures.

5. Generation of the point spread function

The PSF required is the antenna response to a celestial point source. When visibilities at different spatial frequencies (u, v) have been measured, the PSF can be obtained directly from the u, v distribution (often graded by some weighting function) through a 2-D Fourier transform. However, in the present case where we use 1-D image synthesis, we can obtain only limited information about the PSF. To obtain the complete information, we need to know the complex weighting function over the EW array and therefore this needs to be measured accurately. In this connection, we note that the dirty map around a point source will be the PSF if the source is extremely strong and there exists a sufficiently wide field around the source which can be considered essentially empty. At low radio frequencies, due to the ubiquitious Galactic background, this situation does not occur. Though sources like Cas A and Cyg A are very strong compared to the noise in the map, the criterion of an empty field around them is not satisfied.

5.1 *The Complex Nature of the Antenna Response*

Since the telescope is used in the correlation mode (*i.e.* $EW \times S$) and the two arrays are expected to be uniformly illuminated, it is tempting to suggest that the PSF is simply the product of two sin(x)/x functions, one along hour angle and the other along declination. However, this does not give us the true PSF for two reasons. Firstly, the hour-angle beam departs from the sin(x)/x function significantly due to the fact that the actual illumination pattern of the *EW* array is not completely uniform. Secondly, the hour-angle beam is in general complex since the corresponding complex illumination pattern is unlikely be Hermitian symmetric.

Given these problems, one may generate the point spread function by measuring the antenna response to a strong source along hour angle (at the source δ) and along declination (at the α of the source) and multiplying them. This should be a reasonable approach since in the correlation mode one produces the 2-D response by multiplying the response of the individual telescopes.

5.2 *Generating the PSF*

For a simple multiplying system, the antenna response, b, to a point source at (l_0, m_0) can be written as

$$b(l-l_0, m-m_0)=b_{EW}\,(l-l_0, m-m_0)\; b_{NS}(l-l_0, m-m_0), \tag{3}$$

where b_{EW} and b_{NS} refer to the far-field voltage radiation patterns of the EW and the S arrays respectively. Since the EW and the S arrays have slowly-varying patterns along NS and EW (*i.e.* along m and l) respectively, this can be approximated to

$$b(l-l_0, m-m_0)=b_{EW}\,(l-l_0, m_0)\; b_{NS}\,(l_0, m-m_0). \tag{4}$$

Both b_{EW} and b_{NS} are in general complex and

$$b_{NS}\,(l_0, m-m_0)=\frac{1}{2}C_1+\sum_{n=2}^{88}\,(C_n+jS_n)\,e^{j\phi_n}, \tag{5}$$

where, C_n and S_n are the cosine and sine correlations obtained using the nth unit of the S array. $\phi_n=2\pi(n-1)d(\sin(\theta)-\sin(\theta_0))/\lambda$, $m=\sin(\theta)$, and $m_0=\sin(\theta_0)$.

In the GEETEE synthesis system, the dirty map at any instant is a summation of Equation (4) and its complex conjugate. Here one assumes Hermitian symmetry to fill in the corresponding unmeasured visibilities. Hence the PSF is generated by obtaining the b_{EW} and b_{NS} from a point source response and obtaining the real part of the product.

5.3 *Producing the PSF in Practice*

There are several points to be borne in mind in constructing the PSF according to the above recipe:

1) Ideally one should use a strong source at the zenith on a flat background with no significant confusing sources. In such a case the response of the source will reflect any north-south asymmetries in the antenna behaviour. The closest approach to this requirement is Cygnus-A, with a flux density of $\approx$25000 Jy and a zenith angle of 26°. Using the calibrated visibilities around the Cygnus-A position, the complex responses in hour angle and declination (b_{EW}, b_{NS}) were obtained.

2) The basic criterion in choosing the extent (in hour angle and declination) of the PSF is that the expected antenna response for Cygnus A should be well above noise on the map. Along declination this is true over the full extent of the map and the natural size for the PSF along NS is $-0.86<m<0.86$. In hour angle, the sidelobes of Cygnus A can be seen up to ± 3 hr from transit. However, for hour angles $> \pm 1$ hr the curvature of the sidelobes (due to the apparent zenith angle change as a function of hour angle) becomes significant and there is no point in going beyond this extent (see Fig. 2 for a schematic illustration of this effect). In theory, this problem can be overcome by generating a PSF which takes into account the change in the apparent zenith angle of a

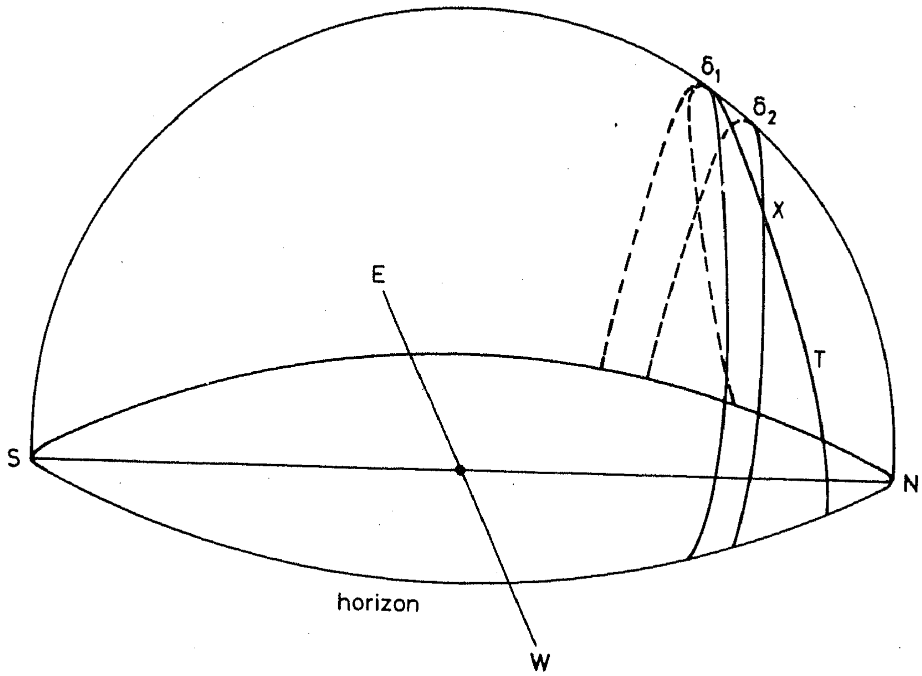

Figure 2. This illustrates the origin of the curved hour angle sidelobes. The response of the array will be in a plane perpendicular to the horizon corresponding to two declinations δ_1 and δ_2 at transit. A source at dec. δ_1 moves in a track (T) tilted at an angle to these. This angle is equal to the latitude of the array. Thus, at an hour angle corresponding to X, the response of the array to the source will be at δ_2.

source as a function of its declination and hour angle. This was not done. Instead, the extent of the complex beam along hour angle was restricted to ± 1 hour to minimize this effect. Of course, the truncation of the PSF along hour angle may affect the dynamic range of the map outside the CLEANed region.

3) The Galactic background also contributes to the estimation of the PSF. One can investigate which baselines are sensitive to the background by looking at the visibility amplitudes as a function of baseline when there is no strong source in the EW beam. This exercise reveals a monotonic decrease in the visibility amplitude from the first to the seventh baseline. The contribution of the background drops well below noise for the longer baselines. To minimise the effects of the background in the generation of the PSF, the measured visibilities in the first *seven* baselines were replaced by suitably extrapolated values.

4) If our estimates of b_{EW} and b_{NS} are affected by ionospheric scintillations then these will affect the derived PSF. To minimise the effects of scintillation, the complex beam along hour angle has been convolved with a $\sin(x)/x$ function of resolution 100 s (the EW resolution of the antenna at $\delta = 0°$) in hour angle. This filters out most of the amplitude scintillations as their time scales are usually shorter than 100 s, but does not affect the resolution of the map.

5) When generating the PSF using Cygnus A, the measured visibilities were corrected for the bandwidth decorrelation at its zenith angle.

6) The implicit assumption in generating the PSF by the multiplication of the complex beams along hour angle and δ is that no other sources have contaminated this estimate. We believe this to be reasonable in the present case.

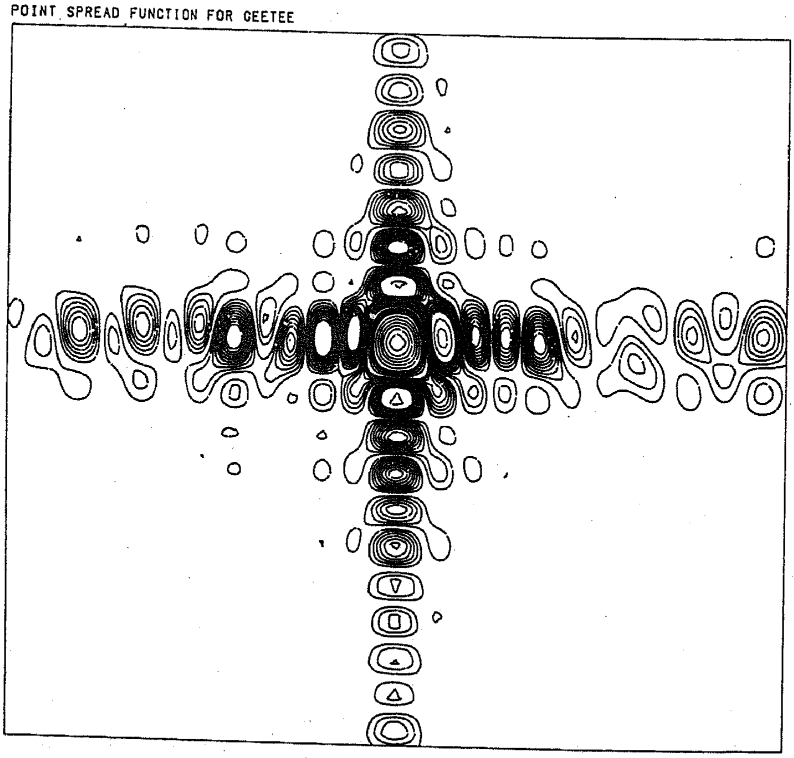

Figure 3. A small section (≈ 10 sq. deg.) of the Point Spread Function which was used for CLEANing. The PSF was derived from Cygnus A. Contours down to the 1% level of the peak have been plotted. The contour interval down to 30% of the peak is in steps of 10%. Between +30% and +10% as also between −30% and −10%, the interval is 5%. Up to ±10% of the peak value, the interval is 1%. Successive sidelobes have opposing signs. The four sidelobes surrounding the peak are each −21% of the peak.

7) Interference could also affect the complex beams. An examination of the map around Cygnus A showed no noticeable effects from interference.

Fig. 3 shows a two-dimensional representation of the PSF derived from Cygnus A.

6. CLEANing in practice

Fig. 4 presents a flow chart for the algorithm used in CLEANing our maps. The individual dirty maps have a size of 1 hr × entire zenith-angle range. Since the PSF has an extent of ±1 hr, we have to have the adjacent maps present to CLEAN the central hour. If the dirty map has non-astronomical features in it, CLEAN can oscillate. To prevent this, the raw data were prefiltered to remove all such features. To achieve this, the raw data were convolved along α with a $\sin(x)/x$ function of resolution ≈ 100 s as described above. This filters out all features sharper than the beam at $\delta = 0°$, although it leaves some at higher declinations. This cannot be avoided without losing resolution. There is no need for smoothing in declination since it is the direction of synthesis.

The PSF covers the entire zenith angle range. The PSF resolution along NS does not change with its position in declination since both the PSF and the dirty maps are generated in the m coordinate. However, the PSF will have to be expanded or contracted along hour angle depending on the declination of the source. In any run of

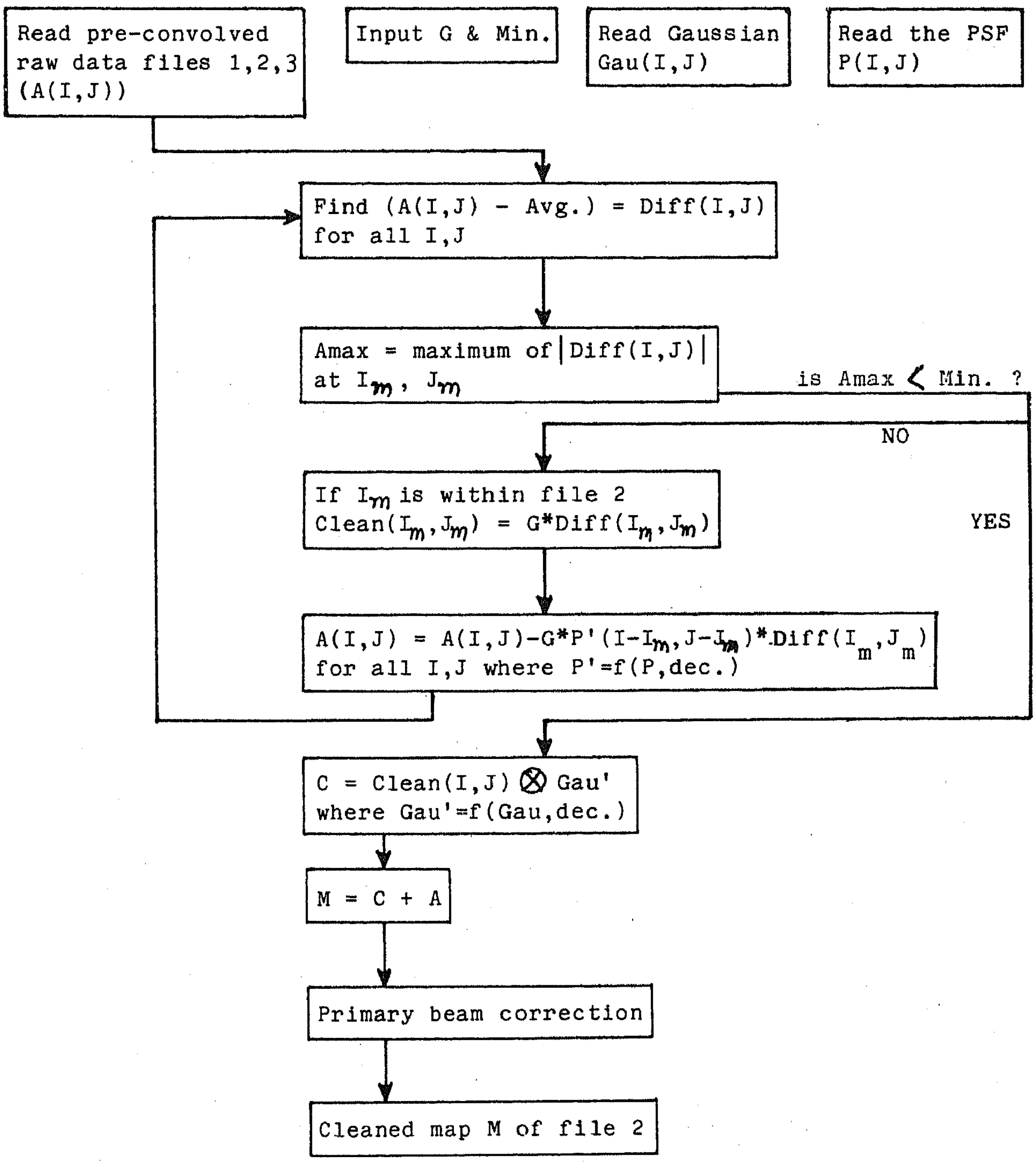

Figure 4. The modified CLEAN algorithm used to deconvolve the maps from the 34.5-MHz survey (DU) which contain both point sources and extended emission.

the program, only the central map of the three read in gets CLEANed and restored. The beam response due to any source in the three hour stretch (say, hours 1, 2, 3) is removed but CLEAN components are written only if the source is located within the central hour. This guarantees that the central hour is CLEANed of the sidelobes of all sources located within an hour of it. Along *NS*, the edges of the map in *N* and *S* are treated as though they are connected to each other. Thus, when the extent of the PSF overshoots any one edge, it continues on to the map from the other edge. In this sense CLEANing is performed as though the map was on a cylinder, with the *NS* extent equal to the circumference of the cylinder. At the end of the run, the middle hour (in this case, hour 2) will be CLEANed. The next run can take dirty maps 2, 3, 4 and so on. The CLEAN components are convolved with a Gaussian whose integral and full widths at half maximum in the two perpendicular directions are equal to those of the PSF and then added to the map of residuals to obtain the final map.

It was possible to CLEAN the whole observed sky down to an almost-uniform level of 5σ, where σ is the r.m.s. fluctuation in the dirty map.

The rows of dipoles in the South array are spaced 5 m apart while the wavelength of observation is 8.70 m. Hence, there is a grating response of beyond zenith angle θ, where,

$$\sin(\theta) = \frac{\lambda}{d} + \sin(\theta_0) \qquad (6)$$

with $d = 5$ m, $\lambda = 8.70$ m and θ_0 = the zenith angle that is aliased with θ. Thus, a grating lobe appears at $\theta_0 = \mp 90°$ when the main beam is at $\theta = \pm 47°$, while at $\theta = \pm 60°$, the two responses are equal. However, for a very large fraction of the sky, we feel that the deconvolution procedure given here has brought out all the relevant features and removed the artefacts to a satisfactory extent.

Fig. 5 shows a dirty map of 1 hr × 15° in extent observed in this survey. We present in Fig. 6 the CLEANed map of the same region. A weak source which is buried in the

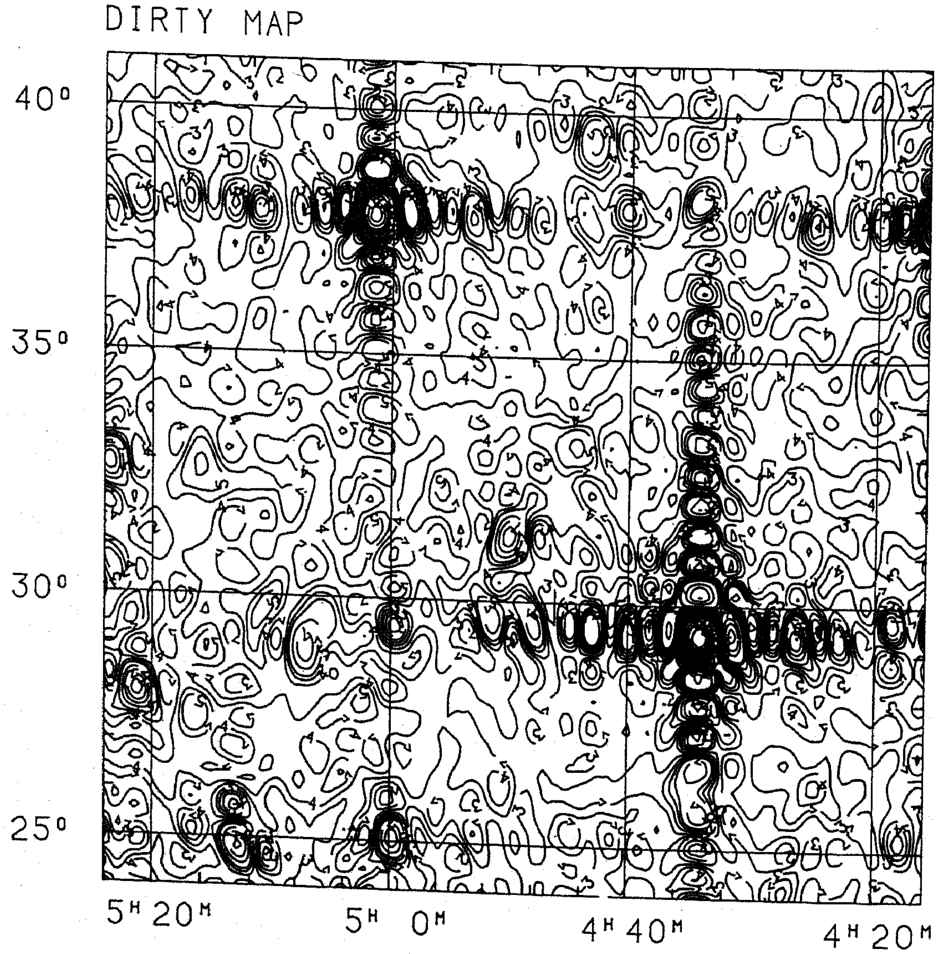

Figure 5. Dirty map of 1 hr × 15° in extent observed in the present survey. The two strong sources 3C 123 ($\alpha = 04^h34^m$; $\delta = 29°23'$) and 3C 134 ($\alpha = 05^h01^m$; $\delta = 37°51'$) and their sidelobes are clearly seen.

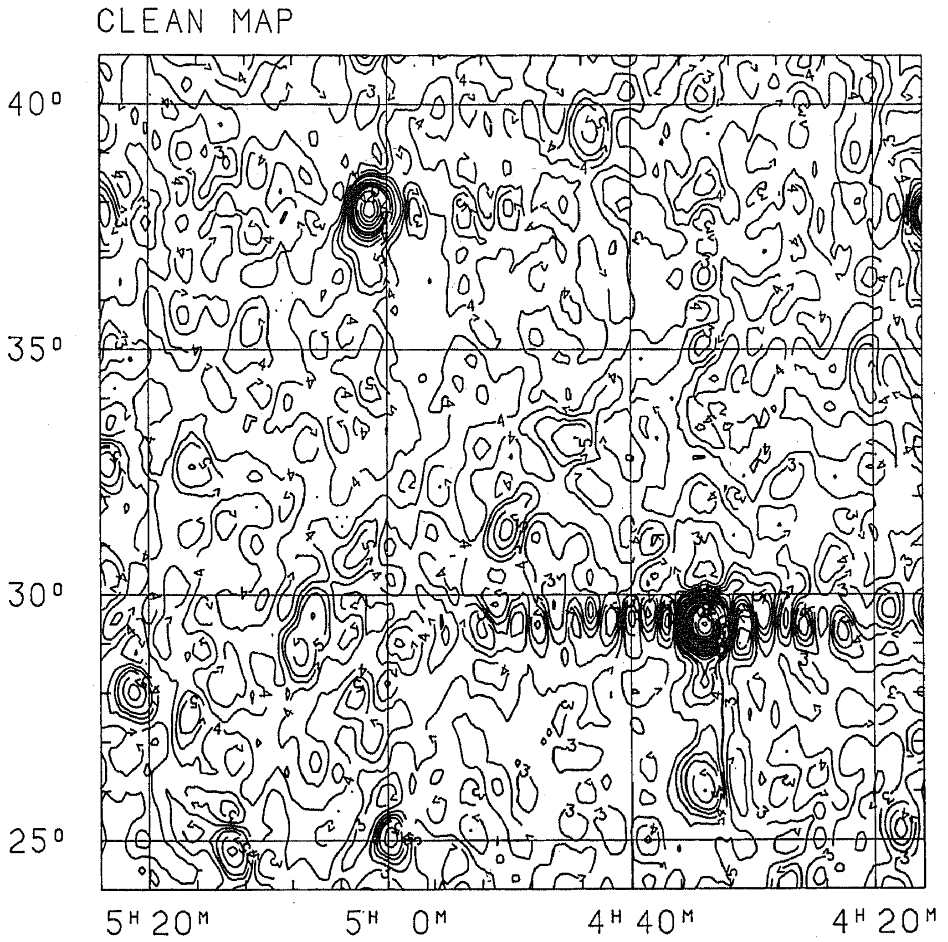

Figure 6. CLEANed map of the region shown in Fig. 5 obtained through the algorithm indicated in Fig. 4. Note that a weak source ($\alpha = 04^h34^m$; $\delta = 26°04'$) which was buried in the sidelobes of 3C 123 (Fig. 5) is clearly seen here.

sidelobes of 3C123 ($\alpha = 04^h34^m$; $\delta = 29°23'$) can be easily seen in the CLEANed map. However, there were some complications near strong sources like Tau A, Vir A, Cas A and Cyg A. Since the PSF extent was only ± 1 hr, sidelobes beyond ± 1 hr from a source cannot be CLEANed. So, when CLEANing any region, a strong source may not be in the field of analysis but the far off sidelobes from such a source could be. To avoid picking up such sidelobes as sources, these regions are flagged beforehand. While the sidelobes remain as they are in the final map, they will be restricted to a narrow range of declination around the source, albeit with a wider range in α.

7. Discussion

The simple modification to the conventional CLEAN proposed here makes the deconvolution algorithm both stable and faster. A method for estimating the PSF reliably from the observed visibilities has been described. These procedures have improved the dynamic range of the maps presented by DU to an estimated > 20 dB. It was not possible to achieve further improvement due to the limitations imposed by the ionosphere and other factors discussed below:

The residual sidelobes seen in the maps of DU are mainly due to the departure of the antenna response (at different α and zenith angles) from the PSF used for CLEANing. These differences are due to changing ionospheric phase perturbations (and amplitude scintillations) as well as due to the changes in the antenna response at different times and zenith angles. Some of the residual sidelobes are due to the following systematic effects: uncompensated bandwidth decorrelation that is significant at large zenith angles, neglecting the curvature of the side lobes in hour angle and limiting the extent of the PSF in hour angle.

Since the final maps were obtained after correcting for the primary-beam gain in zenith angle, the level of the residual sidelobes will be exagerrated at large zenith angles.

All the above effects are easily noticeable for stronger sources, *e.g.* Cas A. Considerable improvements might be possible here if CLEANing were to be performed taking care of the curvature of the sidelobes in α and by correcting for the bandwidth decorrelation.

We believe that our modified CLEAN can be used (with a reliable estimate of the PSF) as a stable, rapid algorithm to deconvolve maps which contain both point sources and extended emission. We note that in our observations there were no 'holes' in the aperture and the purpose of CLEAN was only to remove the undesired sidelobe responses due to the uniform weighting of the aperture. It was not desirable to CLEAN the extended structures. If there are missing short spacings, the conventional CLEAN is expected to fill these in to some extent. It will be interesting to see how the modified CLEAN behaves under such circumstances.

Acknowledgements

We thank R. Nityananda for useful discussions during data analysis and for his many valuable suggestions which contributed significantly to the present work. We thank C. J. Salter for his valuable suggestions which have helped to improve the clarity and readability of this paper.

References

Christiansen, W. N., Högbom, J. A. 1985. *Radiotelescopes*, Cambridge Univ. Press.
Cornwell, T. J. 1983, *Astr. Astrophys.*, **121**, 281.
Clark, B. G. 1980, *Astr. Astrophys.*, **89**, 377.
Dwarakanath, K. S., Udaya Shankar, N. 1990, *J. Astrophys. Astr.*, **11**, 323 (DU).
Högbom, J. A. 1974, *Astr. Astrophys. Suppl.*, **15**, 417.
Schwarz, U. J. 1978, *Astr. Astrophys.*, **65**, 345.
Steer, D. G., Dewdney, P. E., Ito, M. R. 1984, *Astr. Astrophys.*, **137**, 159.
Wakker, B. P., Schwarz, U. J. 1988, *Astr. Astrophys.*, **200**, 312.