MILLIMETER ARRAY
MEMO NO. 32

Mosaicing with the mm array

Tim Cornwell NRAO/VLA

Since the mm array is currently conceived as a collection of differing size arrays and dishs, it is important that we consider how the data from the various telescopes will be combined to form an image of the sky. This memo will outline some preliminary arguments that such mosaicing is indeed feasible.

For the purpose of this memo we shall define the mm array as :

(1) A single structure of maximum size 25m supporting about 25 dishs of size D_{min} = 4m which combine to form an interferometer array. This telescope is designed to image large regions of the sky at low resolution. This array will be super-critically sampled i.e. when gridding for the full field of view, each cell has at least one sample (Hjellming, private communication).

(2) A reconfigurable array of about 20 elements, each of size D_{int} = 10m, with maximum baselines of D_{max} = 90m, 300m, and 1km in successive configurations. This array is super-critically sampled in the first two configurations.

The final goal is an image covering a field of view θ_{FOV}, given by λ/D_{min}, at a resolution θ_{res}, given by λ/D_{max}. Worst still, we may wish to mosaic a large field by pointing the 4m dishes at different points on the sky. The heterogeneity of the mm array is the root cause of the mosaicing problem : we wish to observe large fields with high resolution given by the long baselines, using arrays consisting of different size elements.

I believe that the simplest approach to imaging would be to use the observed data sets as direct constraints upon the deconvolution. To illustrate this possibility, consider a variant of MEM : in normal MEM one maximises the entropy of an image of the sky while trying to simultaneously minimise the fit to the observed data i.e. χ^2. The chi-squared term for interferometers is evaluated by Fourier transforming the trial image and subtracting the resulting visibility from that observed, then summing the resulting residuals. The gradient of chi-squared is obtained by inverse transforming the residuals. One can accomodate the primary beam by tapering the trial image by the primary beam <u>before</u> Fourier transformation. The gradient of chi-squared is again given by inverse transforming the residuals but then they must be tapered then by the primary beam. Multiple data sets can be accomodated simply by summing the relevant contributions to chi-squared and its gradient. The net cost in Fourier transformations is thus the cost per field times the number of fields required. Further details are given in appendix 1.

For those of you who prefer good ole CLEAN, it too can be modified to accept multiple pointing centers. In MX, one would treat the map plane subtraction loop for each pointing center independently but couple the fields together in the u,v plane subtraction (see appendix 2). I am dubious as to the effectiveness of such an approach since CLEAN usually works best on a dirty image which is reasonably close to the true sky (remember why uniform weighting is important !). In the multiple pointing centers CLEAN, the equivalent of the dirty map would be quite distorted.

The true virtue of MEM lies in the way all pixels are changed simultaneously rather than in any mystical property of entropy. As a consequence MEM is faster than CLEAN for large, well filled fields. Furthermore, the MEM approach of guessing an answer and then comparing it to whatever data one has may be much more feasible for the heterogeneous observations that the mm array would provide.

Thus, mosaicing the mm array seems to be feasible and not too far in advance of state-of-the-art image processing. Two algorithms for mosaicing have been sketched out. Neither of these involve disproportionate amounts of computer time; the effect involved being approximately that involved in imaging the same number of fields with a homogeneous inteferometer such as the VLA. The dominant spectral line case will, of course, be correspondingly more demanding but one may be able to save time by starting the optimisation from the continuum image or from the previously processed channel image.

Finally, to illustrate that the required sampling is attainable, consider the following example : suppose that we wish to image the 4m field of view with the 90m resolution. Bob Hjellming (private communication) has calculated the u,v coverage of a typical 90m configuration tracking a source for twelve hours. When gridded onto the u,v grid appropriate to the 4m field of view (53*53), 83% of all cells have samples. One would require about $(10m/4m)^2 \simeq 10$ different pointing centers to cover the 4m field of view. Therefore, since the longest baseline crosses a cell in about 24/(3.1415*53) hours ≃ 9 minutes, one must cycle fields at the rate of one per minute. This does not seem excessively fast and should be acheivable with 10m dishs. We should note that the ratio D_{int}/D_{min} determines the cycling time required to cover all pointing centers and thus cannot be two large.

The continuum data processing requirement is trivial compared to typical VLA examples, for which the image size can be up to 4096*4096. Hence even if we wished to mosaic together 4m fields at the 90m resolution, only modest total memory requirements would be involved. Imaging a 4m field at 300m resolution would still be easy, but the surface brightness sensitivity may be inadequate. Spectral line observations will, however, dominate the computing requirement. The mapping speed will be limited by sensitivity and thus very long tracks will be necessary. For example, one may require 512 separate channel images in 12 to 24 hours. Rough estimates indicate that the MEM algorithm described here running on a small number of VAXes (1-2) could keep up with real time when mosaicing a 4m field at 90m resolution. Certainly, if we allow for the expected evolution in computer power over the next 5-10 years, the load could be handled quite inexpensively (parallel processing is very well suited to spectral line observations). However, this conclusion is very sensitive to the actual mixture of science on the telescope : spectral line snapshots mosaiced together would drastically increase the computing load. Hence, a careful look at the possible observations and their associated computing requirements would seem necessary.

Development of the MEM based algorithm described here will be pursued independently of the mm array design since it will have applications to exisiting interferometers. Numerous practical details have yet to be settled such as : how far out in the primary beam of the smallest element does one have to image ?, can large pointing errors be corrected by including them in the optimisation ?, is the modified CLEAN faster ?, etc.

Appendix 1 Modification of MEM for many pointing centers

For simplicity we use a discrete representation of the images :

$$\underline{b} = (b_i \mid i=1,N)$$

where the images have N pixels. We should emphasize that no commitment to one dimension is implied in this notation. Restricting our attention to an interferometer, our observed data are samples of the visibility function, one for each pointing center. The k'th sample for the l'th pointing center is :

$$V_{k,l} = \Sigma_i \, q_{i,l} . t_i . \exp(2\pi . j . \underline{u}_{k,l} . \underline{x}_i) + n_{k,l}$$

where $\underline{n}$ represents noise, $\underline{t}$ is the true image, $\underline{u}_{k,l}$ is the position vector of the k'th sample in the u,v plane for the l'th pointing center, $\underline{x}_i$ is the position vector in the image plane of the i'th pixel, and $q_{i,l}$ represents the effects of the primary beam on the i'th pixel. The corresponding data predicted from $\underline{b}$ are :

$$V'_{k,l} = \Sigma_i \, q_{i,l} . b_i . \exp(2\pi . j . \underline{u}_{k,l} . \underline{x}_i)$$

Forming the χ^2 for $\underline{n}$:

$$\chi^2 = \Sigma_{k,l} \, w_{k,l} . \mid V_{k,l} - V'_{k,l} \mid^2$$

where $w_{k,l}$ is the relevant weighting factor for the k'th visibility sample for the l'th pointing center.

Maximising the relative entropy subject to the constraint that χ^2 be equal to the expected value we find that the solution is:

$$b_i = m_i . \exp(-\alpha . (\partial \chi^2 / \partial b_i))$$

where α is a Lagrange undetermined multiplier and in this particular case:

$$\partial\chi^2/\partial b_i = \Sigma_l\, q_{i,l}.2.\mathrm{Re}(\Sigma_k\, w_{k,l}.(V'_{k,l}-V_{k,l}).\exp(-j.2\pi.\,\underline{u}_{k,l}.\,\underline{x}_i))$$

In the image plane :

$$\partial\chi^2/\partial b_i = \Sigma_l\, q_{i,l}.2.(\,\Sigma_j\, p_{i,j,l}.(q_{j,l}.b_j - d_{i,l})$$

where $\underline{d}_l$ is the Fourier transform of $\underline{V}_l$ and p_l is the point spread function, given by the Fourier transform of the weights.

$$p_{i,j,l} = 2.\mathrm{Re}(\Sigma_k\, w_{k,l}.\exp(-j.2\pi.\,\underline{u}_{k,l}.(\,\underline{x}_i - \underline{x}_j)))$$

The gradient would be used to find the MEM image just as described by Cornwell and Evans (1985). Computationally, all this is simple if longwinded. I have programmed a toy FORTRAN version to run under VMS, taking certain shortcuts to ease the coding. Figures 1 and 2 show two images made from fake data sets simulating a SNR. Figure 1 shows a deconvolution using MEM for fairly complete coverage of the u,v plane but only one pointing center and a very large primary beam, while Figure 2 shows a reconstruction from 16 separate pointings with a correspondingly smaller primary beam but with the same number of visibility points. Differences between the two reconstructions are noticeable at the few percent level.

Programming this algorithm in AIPS would be painful unless some suitable higher level routines are first written. Also, including selfcalibration is conceptually easy but slightly messy to implement unless one programs specifically for the small field case assuming a large amount of physical or virtual memory.

Just to emphasize the obvious, different size primary beams can be accomodated simply by changing the appropriate $p_{i,j,l}$. While mosaicing to the 4m field of view, one can also remove the 4m primary beam.

Finally, I should acknowledge that MEM does not work particularly well for low SNR images since the answer is biased. I have experimented with variants of MEM which do not require positive images (and are thus not biased) but which use the same optimisation scheme. The prelimary results are very promising and indicate that the low SNR case can be thus treated.

Appendix 2 Modification of MX for many pointing centers

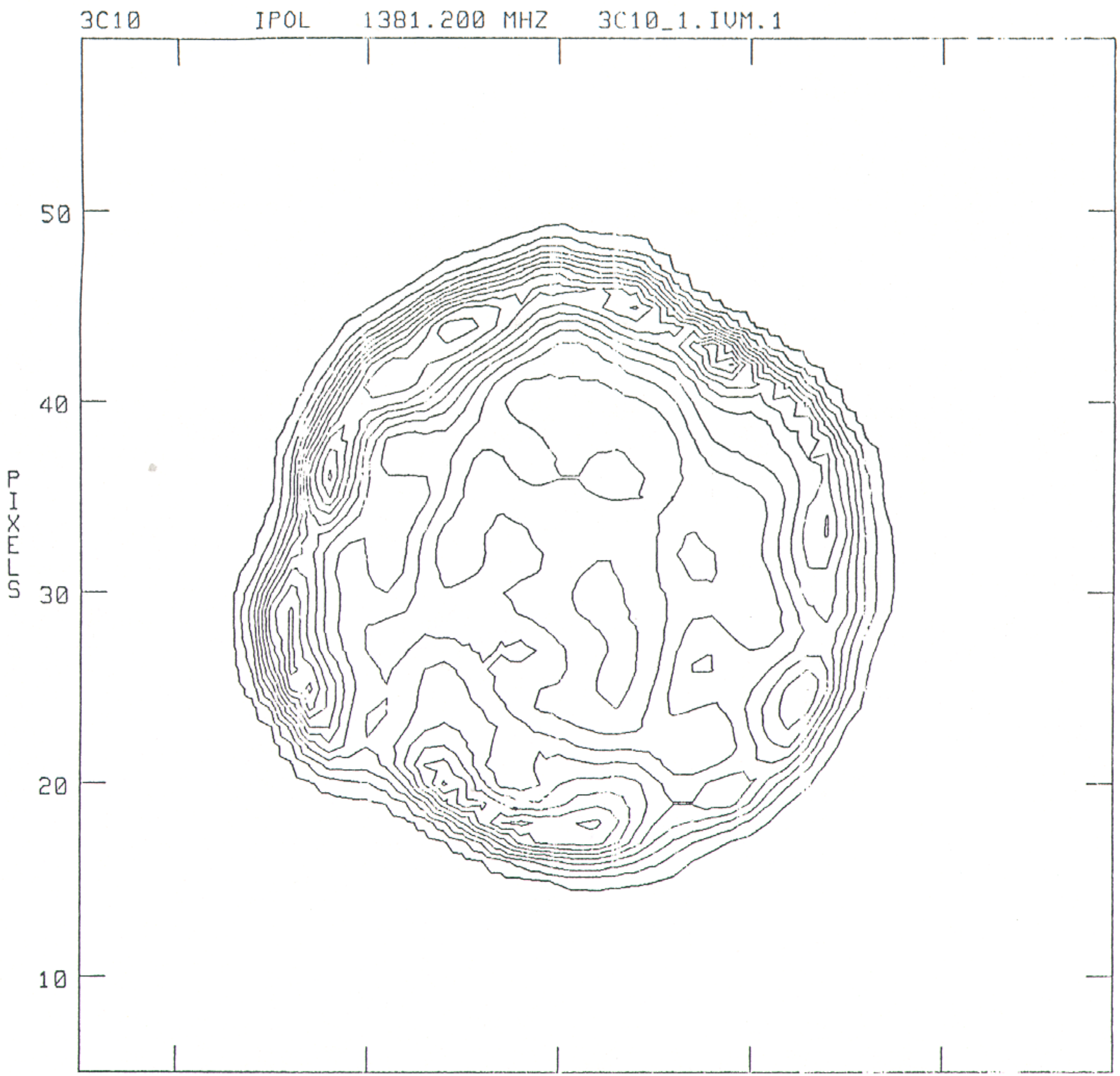
One plausible scheme for adapting MX would be :

(1) In the minor, image plane subtraction loop, treat each pointing center separately : form the residual/dirty map as usual, clean with the appropriate dirty beam. At the end of each minor cycle, correct the clean components by dividing out the primary beam. (Fred Lo has pointed out that a Wiener type linear sum of all possible clean components would be more stable.)

(2) In the major, uv plane subtraction loop, remove all clean components from the data from all pointing centers. Before subtraction from the data from a particular pointing center, taper the clean components by the primary beam centered on the appropriate place.

(3) Glue the various subfields back together carefully.

As with MEM, selfcalibration can be thrown in on top if either the gridding problem can be overcome or the sorting cost accepted.



CENTER AT RA 00 22 28.000 DEC 63 51 40.00
PEAK FLUX = 1.2200E-01 JY/PIXEL
LEVS = 0.1000E-01 * (1.000, 2.000, 3.000,
10.00, 11.00, 12.00, 13.00, 14.00, 15.00)

Figure 1

Only one pointing center, large primary beam

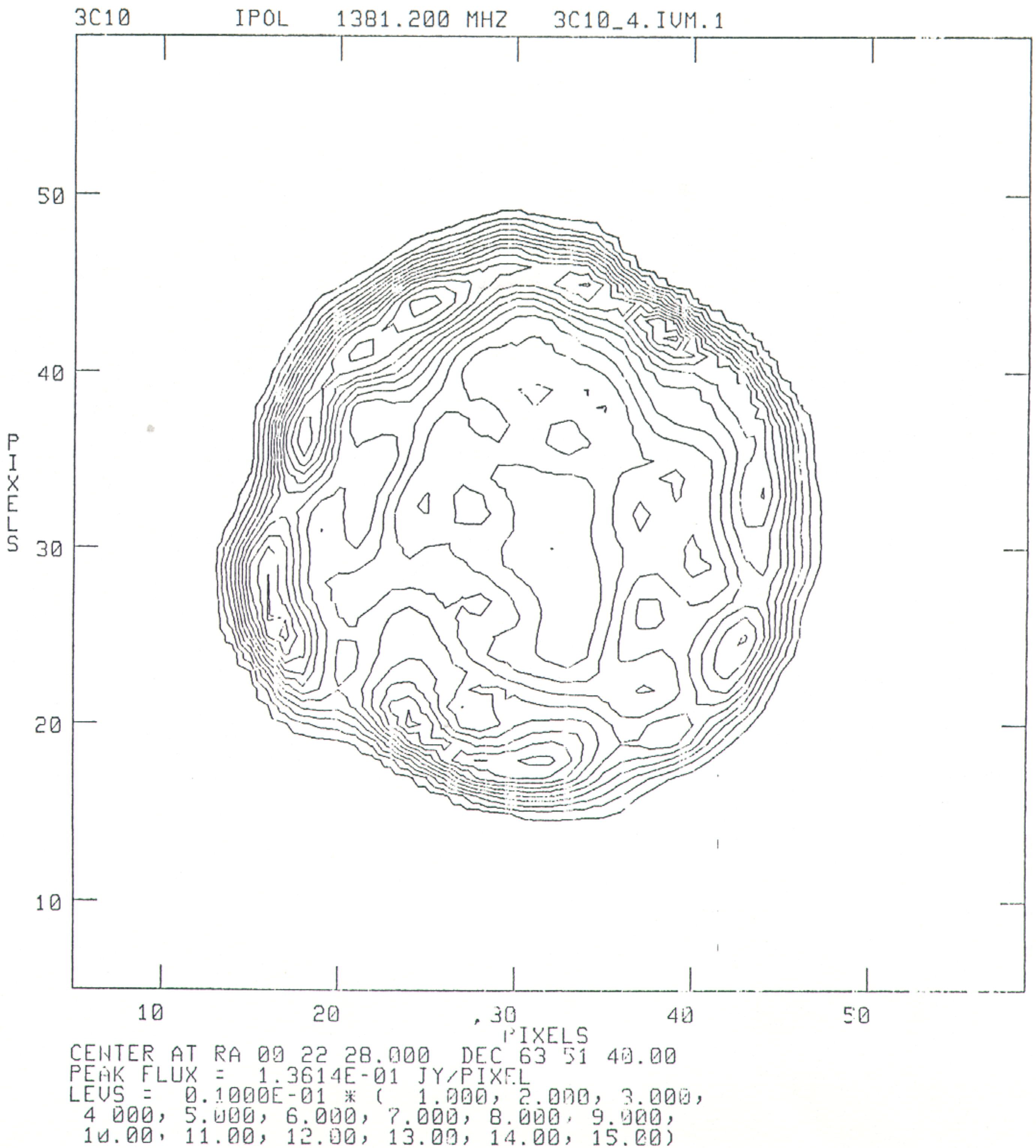


Figure 2

16 pointing centers, small primary beams

FWHM of primary beams