A&A manuscript no.
(will be inserted by hand later)

Your thesaurus codes are:
13(03.20.2; 03.13.2)

ASTRONOMY
AND
ASTROPHYSICS
17.7.1996

An approach to interferometric mosaicing

R.J. Sault[1,2], L. Staveley-Smith[1], and W.N. Brouw[1]

[1] Australia Telescope National Facility, CSIRO, P.O. Box 76, Epping, N.S.W. 2121, Australia
[2] Netherlands Foundation for Research in Astronomy, Postbus 2, 7990 AA Dwingeloo, The Netherlands

Received date; accepted date

Abstract. In radio astronomy, mosaicing is the practice of combining multiple pointings of a telescope to produce an image significantly larger than the telescope's primary beam. We consider an approach to imaging and deconvolving interferometric mosaic observations. This approach directly deconvolves a linear mosaic of dirty images. We consider the advantages and limitations of this technique. Projection issues for mosaicing are also discussed. We present an example from a 320-pointing, spectral-line, mosaiced observation of the Small Magellanic Cloud.

Key words: techniques:interferometric – methods:data analysis

1. Introduction

With a single pointing, an interferometric array is only sensitive to a patch of the sky determined by the primary beam of each antenna. Consequently, to image objects larger than the primary beam, multiple pointings are required. Here we consider interferometric mosaicing, i.e. combining interferometric data from different pointings to form a single image of part of the sky.

Although mosaicing techniques have proven advantageous in survey observations (e.g. Bremer 1994; Condon et al. 1994), here we consider mosaicing to be more than a 'cut and paste' merging of images from different pointings. Mosaicing can recover information inaccessible from a single-pointing observation. This is because an interferometer with projected baseline d does not measure just a single spacing in the $u-v$ (Fourier) plane. For antenna size D, an interferometer measures an integral of the spacings from $d-D$ to $d+D$. Whereas the information in this range of spacings is generally not accessible in a single pointing, it can be recovered in a mosaic experiment (i.e. the effective $u-v$ coverage can be improved). This was first noted by Ekers & Rots (1979), where they suggested a scheme of scanning antennas. A much firmer theoretical framework was developed by Cornwell (1988), where he showed that scanning was not required provided the pointing sampling grid is sufficiently fine (the fineness required is given by a sky-plane Nyquist sampling theorem). Cornwell presented a practical algorithm for combining data from multiple pointings, and showed that this approach retrieved the same information as the Ekers & Rots technique. The extra information retrieved from a mosaic experiment is usually most noticeable with extended emission. This is because the relative improvement in effective $u-v$ coverage is greatest at the short spacings. Indeed, measuring short spacings was the intention of Ekers & Rots.

Because the $u-v$ coverage achieved in an interferometric observation is often incomplete, a non-linear algorithm is needed to produce good deconvolutions; only a non-linear algorithm can estimate the missing information (e.g. Cornwell & Braun 1989). In this case, Cornwell (1988) stresses that it is possible to produce superior results by combining all the data *before or during* the deconvolution phase, not afterwards. Such approaches are called 'joint' deconvolution schemes. His approach involved combining the data during the deconvolution: an image is formed at each pointing, and a maximum entropy method is used to find a model of the entire field consistent with these data. Cornwell's scheme also includes single-dish data in a natural way.

In this paper we consider an approach to imaging and deconvolving observations where the image data are combined before deconvolution. This is a joint deconvolution approach where we consider interferometric data only. The technique shares a number of commonalities with work by Gueth et al. (1995). The advantages and limitations of our approach are considered. Finally we give an example, and consider scope for further work.

2. An imaging and deconvolving approach

Our approach to mosaicing proceeds in three steps, which we have implemented as separate imaging, deconvolving and restoring tasks. First a dirty mosaiced image dataset

Send offprint requests to: R.J. Sault

is formed from the visibility data. A dataset used to determine the point-spread function (PSF) is also produced. Secondly a "deconvolution" task is used to deconvolve this mosaiced dirty image to produce a model of the sky. Finally we convolve this model with a Gaussian beam and add the residuals of the deconvolution process. The mosaic schemes described here are implemented in the Miriad system (Sault et al. 1995).

2.1. The imaging step

The imaging step is fairly standard. Dirty sub-images from the individual pointings are made using a conventional grid-and-FFT scheme. The visibility weighting scheme and the problem of making the sub-images on the same pixel grid is addressed in the following sections. The sub-images are then mosaiced together using a standard linear mosaicing scheme (e.g. see Cornwell et al. 1993). For linear mosaicing, a pixel in the output is a weighted sum of the corresponding pixels in the input sub-images. The weights are determined both to correct for primary beam attenuation and to minimize the noise level in the output. For simplicity, we use one-dimensional notation and assume that the primary beam for all pointings is the same. In this case, the linearly mosaiced image is given by

$$I_{\rm LM}(\ell) = W(\ell)\frac{\sum_p A(\ell-\ell_p)I_p(\ell)/\sigma_p^2}{\sum_p A^2(\ell-\ell_p)/\sigma_p^2}. \quad (1)$$

The sums run over the pointings, p. Here ℓ_p is a pointing centre, $I_p(\ell)$ is a dirty sub-image and $A(\ell)$ is the primary beam response of an antenna. The weighting in the numerator and denominator by the noise variance of a sub-image, σ_p^2, causes the noise variance in the resultant mosaic to be minimized. We use the theoretical noise variance for σ_p^2, which is computed from on-line measurements of the system temperature, as well as other observation parameters and the weights used for the visibilities in the imaging.

The weighting factor, $W(\ell)$, requires further explanation. We introduce this somewhat arbitrary factor to suppress excessive noise amplification at the edge of the mosaic. We set this weight so that the noise at a pixel in the mosaic never exceeds a particular threshold value, which is set at the maximum noise variance in the individual sub-images. Specifically, for a threshold $\sigma_{\rm T}^2$,

$$W(\ell) = \min(1, (\sigma_{\rm T}^2 \sum_p A^2(\ell-\ell_p)/\sigma_p^2)^{1/2}). \quad (2)$$

We find that this weighting has a number of good characteristics. It produces an output mosaic which has a roughly constant noise variance (constant within a factor of a few, assuming the noise variances in the sub-images are fairly uniform). $W(\ell)$ will be unity within the perimeter of the outermost pointings (assuming there are no holes in the pointing grid), and will taper to zero beyond this. In the extreme case of 'linearly mosaicing' a single pointing, this weighting exactly counteracts the primary beam correction introduced by the other terms, and so the output image is simply the conventional dirty image. Heuristically, the weighting produces aesthetically pleasing images, and will improve some deconvolution processes (e.g. inverse or Wiener filtering as suggested by Cornwell et al. 1993). However, as for the primary beam of a single pointing observation, the attentuation left at the edges may need to be considered during the astronomical analysis phase.

2.2. Visibility weighting schemes

In normal interferometric imaging, weights are generally applied to the visibility data before the Fourier transform step (see e.g. Sramek & Schwab 1989 or Briggs 1995). The most commonly used schemes are uniform weights (which minimizes the sidelobe level) and natural weights (which maximizes the point-source sensitivity).

When mosaicing, maximizing the point-sensitivity will be achieved by using natural weights for each sub-image. However simply using uniform weights for each sub-image does not minimize the sidelobe level in the output mosaic: as the linearly mosaiced image is a weighted sum of pixels in the sub-images (with the weights varying across the sub-images), the contribution to the mosaic of sidelobes in the sub-images varies with position. Also when the field size of the sub-images is large (as is suggested by the sensitivity and deconvolution considerations in Sect. 2.4), uniform and natural weighting become essentially identical. This is because in determining the sampling density function for uniform weights, the $u-v$ grid becomes so fine that generally at most one visibility lies in each cell. In this case uniform weights (like natural weights) result in high sidelobe levels. Another way of interpreting this is that uniform weighting is attempting to minimize the sidelobes over too large a field, and so produces poor sidelobe suppression overall. Super-uniform weights can circumvent this behaviour: these attempt to mimimize the sidelobe level in a region smaller than the full field being imaged.

Optimal weighting schemes which mimimize the sidelobe level in the linearly mosaiced image would be possible using a weighting approach described by Sault (1984). However the computational expense in determining these weights would not justify their modest advantages. To produce low sidelobe levels in a mosaiced image, we suggest using super-uniform weights in each sub-image.

2.3. Projection considerations

When imaging large sections of the sky, attention must be paid to the non-planar nature of the celestial sphere. Ideally the resultant mosaic should represent a single projection operation on this sphere. Our approach is to produce sub-images that use the same projection, and lie on the same pixel grid. We achieve this in two steps. First the (u, v) coordinates of the visibilities are manipulated so that the resultant sub-images have the same projection geometry. The form of projection natural to an interferometer is the so-called "sine" projection about some tangent point. Second we multiply the visibility data by a phase term, corresponding to a shift in the image plane by a fraction of a pixel, to make the resultant images lie on the same pixel grid. We treat the projection-related manipulations in more detail, considering separately the cases of observations where the (u, v, w) coordinates of a visibility either do or do not lie entirely within a plane (so-called coplanar or non-coplanar observations).

We consider non-coplanar observations first. Non-coplanar observations produce images with a smearing or "w-axis" distortion (e.g. Cornwell & Perley 1992). This distortion is minimized by using the delay-tracking centre (which presumably coincides with the pointing centre) as the tangent point for imaging. Using this approach in a mosaiced observation would result in the different pointings using different projection tangent points and so would have different projection geometries. There are several approaches to overcome this. One is to regrid the resultant sub-images to a common tangent point. This, however, introduces interpolation errors (Braun 1988a). Also, if the correction amounts to anything other than a rotation or stretch, then the synthesized beam is no longer constant across the sub-image. Another approach is to rotate the (u, v, w) coordinates, before imaging, to a common tangent point. This will be adequate provided the w-axis distortion is not a significant problem over a field the size of the entire mosaic. A third approach, which we consider in detail in Appendix A, is to apply a linear transformation to the (u, v) coordinates. This transformation is chosen to minimize the projection mismatch between the individual fields and to be exact at the pointing centre of each field. Because only u and v are transformed, no additional w-axis distortion is introduced over that which would be present in an image without any transformation. As the net result of the transformation is to rotate and stretch the resultant image, the convolution relationship between the sky and dirty image is not altered (the synthesized beams must be computed with the same transformed (u, v) coordinates as the dirty image). As it is simple and computationally cheap, it should be the preferred approach. However, this technique, as well as the other two, will be inadequate if the w-axis distortion is significant within an individual sub-image. In this case, a full "3-dimensional" imaging approach will be required (e.g. Cornwell & Perley 1992).

When the baseline vectors are coplanar, it is possible to avoid the above approximations. The most common types of telescopes that produce coplanar observations are east-west arrays. One way of describing the projection geometry of such arrays is to consider the tangent point to be at the normal to the plane of the baselines (the pole for east-west arrays). In this case, w is always zero, and a pure sine projection, with a fixed tangent point, results – there is no w-axis distortion. Although it is common practise to produce images using (u, v, w) coordinates computed relative to a point other than the pole, this is unimportant. This simply introduces a geometric distortion which only makes the mapping between sky and pixel coordinate somewhat more elaborate (e.g. Brouw 1974a). The important point is to compute (u, v) coordinates, for all pointings, relative to a common reference point in the sky. In this case, all sub-images will have the same projection geometry. Although it is normal for the (u, v, w) coordinates to be computed relative to the delay-tracking centre of each pointing, it is a simple matter to rotate them to a common point.

We do not consider as a special case observations where the data for a given pointing are coplanar, but where the set of all data is not coplanar. Such observations result from a physically planar array where the data for each pointing are derived from a single snapshot. We treat such observations as non-coplanar. However, Condon et al. (1994) present an approach which is advantageous for such cases.

2.4. Deconvolution

It is sometimes believed that the 'magic' of a non-linear deconvolution algorithm is in some way responsible for recovering the extra information from a mosaicing experiment. This is not the case. For example Ekers & Rots' (1979) technique uses only linear operations. Cornwell et al. (1993) show that a linear mosaic of dirty sub-images also contain the extra information. In particular, they show that the effective $u-v$ coverage of a linear mosaic is approximately equivalent to convolving the conventional $u-v$ plane sampling pattern with the Fourier transform of the square of $A(\ell)$. In principle, then, deconvolving a linear mosaic of dirty sub-images has the advantages of a 'joint' approach.

As Cornwell et al. (1993) note, a linear mosaic of dirty sub-images results in a position-variant PSF. Part of that paper considers approximate ways of deconvolving such an image (we use 'deconvolve' in a broad sense, rather than the strict sense where it applies only to inverting position-invariant operations). Here we describe an approach which avoids approximations, thus enabling good deconvolutions of a linear mosaic to be achieved. We do, however, assume that the primary beam response is of finite extent. Although this is incorrect, in practice the primary beam is only modelled over a finite region (we consider this approximation later).

If $B_p(\ell)$ is the synthesized beam pattern for pointing p, the PSF at ℓ_0 is

$$B_{\mathrm{LM}}(\ell;\ell_0) = W(\ell)\frac{\sum_p A(\ell_0-\ell_p)B_p(\ell-\ell_0)A(\ell-\ell_p)/\sigma_p^2}{\sum_p A^2(\ell-\ell_p)/\sigma_p^2}. \quad (3)$$

An output dataset from our imager (the so-called 'beam' dataset) consists of the synthesized beams and primary beam model for each pointing, plus other book-keeping information; this information allows the PSF to be computed at any position. For a primary beam model of finite extent, this PSF will also be of finite extent and no more than twice the extent of the primary beam (the actual extent will depend on the relative location of ℓ_0 and the pointing centres). Additionally, the PSF is considerably cleaner than synthesized beams from the individual pointings. This is both because the distant sidelobes are down-weighted by the linear mosaicing process, and because summing together beam patterns is equivalent to filling in the $u-v$ coverage (here we assume a time-shared observing strategy). As an example, Fig. 1 gives a synthesized beam of a pointing and a mosaic PSF for the observation described in Sect. 3.

A basic operation involved in many deconvolution algorithms is to determine residuals by converting a prospective model of the sky into a dirty image and then subtracting it from the observed dirty image. Despite the position-variant PSF of the linear mosaic, this is not excessively computationally expensive, as the process can be decomposed into efficiently implementable operations. The steps are as follows: divide the model sky by $W(\ell)$ and then, for each pointing, multiply by the primary beam and then convolve with the appropriate synthesized beam. This leaves a full set of dirty sub-images, which are then linearly mosaiced together, and subtracted from the observed dirty image.

The convolution operation with the synthesized beam is implemented using FFTs. As with all FFT-based convolutions, care must be taken to avoid the side-effects of their cyclic (as distinct from linear) nature. This is especially so since the sky in a sub-image is full of emission. However, as the models of the primary beam response are of finite extent, an exact treatment is possible provided the sizes of the image being convolved, the sub-image, the synthesized beam and the FFT are decoupled; images can be appropriately zero-padded and post-FFT edges discarded to achieve this. If N_{s} is the size of the sub-image and N_{pb} is the extent of the primary beam (both in pixels), then a synthesized beam of size $N_{\mathrm{s}} + N_{\mathrm{pb}} - 1$ and an FFT of at least this size is required.

Ideally, the sub-image size should be the full extent of the primary beam. Linear mosaicing with a smaller sub-image will result in some lost of sensitivity and will introduce some extra distortions into the dirty mosaic (the PSF will be dirtier, and more strongly position-dependent). With an ideal deconvolution algorithm, the latter distortion can be removed. In practice, results are best if the PSF is as clean as possible.

Using the above method to go from a model sky to a dirty image, a number of deconvolution techniques can be readily implemented. We have implemented a maximum entropy deconvolver (based on the algorithm of Cornwell

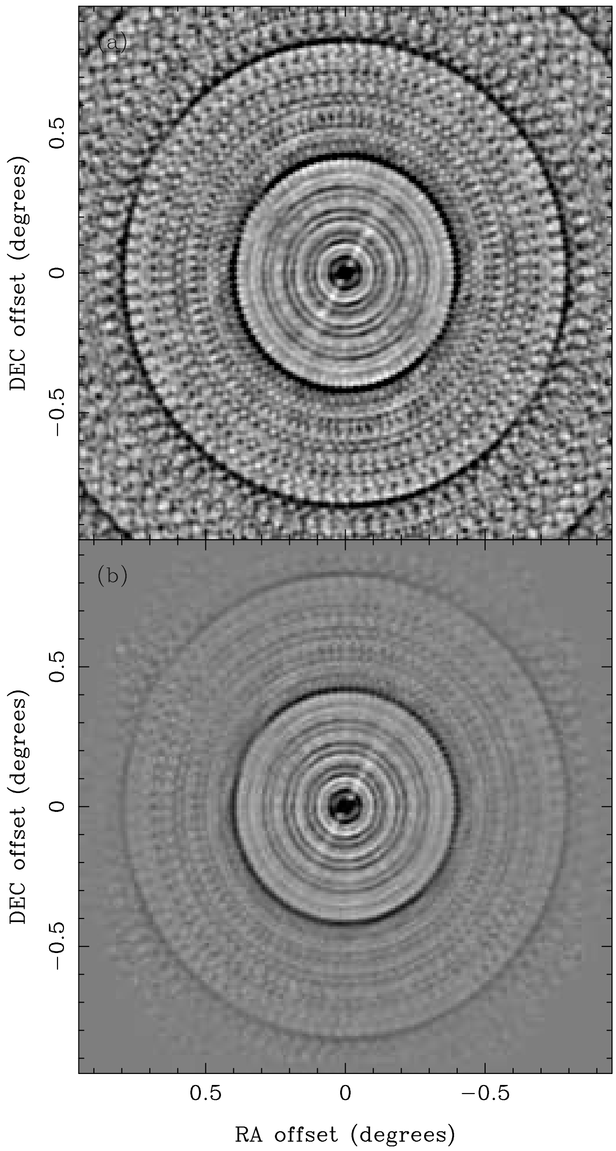

Fig. 1. Point responses from a 320-pointing ATCA mosaic observation. Both displays are saturated at the level ±0.1. a) A conventional synthesized beam from the centre-most pointing; b) The PSF of the mosaic at the same position.

& Evans 1985, with modifications by Sault 1990), as well as the CLEAN algorithm variant of Steer et al. (1984). Externally and internally, these algorithms are little different to deconvolvers of a single pointing observation. Apart from using a different 'convolution' module, the only significant difference is that the noise variance, $\sigma^2(\ell)$ is not constant. As the image datasets contain the information needed to determine $\sigma^2(\ell)$, the maximum entropy algorithm handles this in a natural way in the χ^2 constraint (we give the user the ability to multiply the theoretical level by a constant 'fudge factor'). The CLEAN-like approach has to be modified slightly: it is no longer correct simply to search for the highest residual at each iteration. However, it is easy to show that the correct approach is to locate the maximum of $|I_{\rm LM}(\ell)/\sigma(\ell)|$. This is correct in a maximum-likelihood sense when the synthesized beams are much smaller than the primary beam.

The above maximum entropy approach differs from that of Cornwell (1988): he advocates calculating the χ^2 statistic from the residual visibilities (although the AIPS implementation of his algorithm calculates χ^2 from the sub-image residuals). We calculate χ^2 from the linear mosaic of the residuals. Calculating χ^2 from the visibilities would be too computationally expensive for our applications. These approaches are distinct. However as Cornwell & Evans (1985) note, how χ^2 is estimated is not excessively important. All approaches still possess the advantages of joint deconvolution.

2.5. Practicality

Our approach to mosaicing has been dictated as much by practicality as theory. Our astronomical interests involve large mosaics of spectral line experiments. We have to be concerned with ease of the data reduction process, as well as minimizing disk space and CPU requirements.

As seen by a user, the steps involved in this reduction process appear fairly conventional. The input visibility dataset contains the multiple pointings. From this dataset, an image is formed, deconvolved and then restored. The user deals with familiar objects at the different steps – single visibility datasets and images (admittedly the beam dataset is not normal). Apart from extra CPU and disk requirements, the user is unaffected by the number of pointings.

In addition to simplifying the user's lot, building the linear mosaicing into the imager is convenient for handling the projection corrections. However, one compelling reason for doing it it is the resulting saving in computer resources. Typically, the linear extent of a sub-image is twice the primary beam FWHM and the pointing grid spacing is half the FWHM. Hence the total number of pixels in the sub-images is typically 16 times more than in the final mosaic. It is best to do the linear mosaicing as soon as possible to save computer memory and disk space. This advantage is not as impressive when the beam dataset is also considered: the beam dataset contains the (uncombined) synthesized beams for the individual pointings. There is no saving in storage of the synthesized beams. As the synthesized beams typically have 4 times as many pixels as the sub-images then, for continuum experiments, the overall fractional saving in disk space is small. However, for spectral line experiments the beam datasets can be made frequency independent. Our software produces image cubes where the image cell size is scaled inversely with frequency (note that the PSF derived from this dataset is frequency dependent, because the pointing grid does not scale with frequency). In this way only one set of synthesized beams is needed for a spectral mosaic cube. Hence savings in disk space of the factor of 16 can be approached when the size of the mosaic cube dominates.

2.6. Limitations

There are a number of shortcomings to our approach. We consider those limitations which are related to our purely image-plane-based deconvolution approach, and those intrinsic to joint deconvolution.

For FFT-based imaging algorithms, any deconvolution scheme that works purely in the image plane will suffer from the aliasing effects, as aliasing violates the assumed convolution relationship between the sky and dirty image. This problem is somewhat more severe for mosaicing experiments because there will always be emission outside each sub-image. In addition to using prolate spheroidal gridding functions (Brouw 1974b; Schwab 1984), we reduce aliasing by discarding the outer portions of the FFTed images when forming sub-images and synthesized beams (typically 10% to 50%, in each dimension, of the FFTed image is discarded). Even so, aliasing will persist at some level. Aliasing errors can be avoided by algorithms that determine residuals in the ungridded visibility plane (and use a direct Fourier transform approach to get there).

A practical limitation to an image-plane-based deconvolution scheme is the large size required for the stored synthesized beams. For proper deconvolution, we require synthesized beams which are $N_{\rm s} + N_{\rm pb} - 1$ in size. Furthermore, to produce better quality deconvolutions, a primary beam model of larger extent (i.e. a larger value for $N_{\rm pb}$) is required, and so a larger synthesized beam is also required. Again, approaches which determine residuals in the ungridded visibility plane do not suffer this limitation. Typically these algorithms require, at most, a beam patch, not the full synthesized beam. In this way, the beam storage requirement can be reduced and decoupled from the extent of the primary beam model (the expense of computing the model visibilities, however, would still increase with the extent of the primary beam model).

The joint deconvolution approach, in general, is not without its drawbacks. With joint deconvolution, the primary beam responses (the assumed model of the primary beam, the pointing centres, etc) are as much a part of the PSF as the synthesized beams. Errors in the primary beam responses result in errors in the PSF. This is quite unlike deconvolving a single pointing. Braun (1988b) and Cornwell et al. (1993) considered the effects of primary beam errors. Braun showed that approaches which combine the sub-images after individual deconvolution can produce higher dynamic ranges for compact sources. Braun (1988a) also showed that a hybrid approach of joint and individual deconvolution of the short and long spacings, respectively, can produce better dynamic ranges while retaining the advantages of joint deconvolution (Holdaway 1992 expanded on these hybrid schemes).

In terms of sensitivity to primary beam errors, our approach does not differ from other joint deconvolution approaches. However, because the sub-images are combined at an early stage, our approach makes post-imaging correction to the primary beam response difficult or impossible. If the primary beam response were to be determined after imaging (e.g. by some form of pointing or primary beam response self-calibration algorithm), then our approach would be less attractive.

3. An example

Here we compare of our method with both the (image-plane-based) algorithm of Cornwell (1988) and individual deconvolution of sub-images. The images presented here have been derived as uniformly as possible. Similarly, the different algorithms have been implemented using as much common code and techniques as possible. Thus the differences between images and computational load are a result of the different algorithms and not implementations. In particular, all images have a residual primary beam weighting given by Eq. (2) and the results of the deconvolutions have been 'restored' (i.e. they have been convolved with a Gaussian beam, and the residuals have been folded in).

We have applied our imaging and deconvolution algorithms to a large λ21-cm survey of the Small Magellanic Cloud (SMC) taken with the 375 m configuration of the ATCA. This survey consists of a hexagonal grid of 320 pointings covering an area of $4° \times 4°$, as shown in Fig. 2. The spacing between adjacent pointing centres is $15'$. This is comfortably less than the sky-plane Nyquist sampling interval of $\lambda/\sqrt{3}D = 19'$ required for a hexagonal grid ($D = 22$ m for the ATCA antennas). Other observational details are described in Staveley-Smith et al. (1995).

The $u - v$ coordinates of the visibility data were corrected for geometric effects and Fourier-transformed. Point-source responses at the centre of the mosaic (both the synthesized beam of the central pointing, and the linear mosaic PSF at the same location) are shown in Fig. 1. The linear mosaic of the dirty sub-images for one velocity plane is shown in Fig. 3(a). This plane represents the HI emission at heliocentric velocities from 122.4 to 124.0 km s^{-1}. The bar region of the SMC is prominent but sidelobes in the PSF, mainly because of the first $\lambda/(30\text{ m})$ grating ring, badly confuse the structure. Higher-order grating rings are suppressed by the primary beam of the antennas.

For comparison purposes, we have deconvolved and restored the individual sub-images *before* linearly mosaicing the resultant sub-images together. Here, as with the other images of this section, we used a maximum entropy deconvolver and $1\overset{'}{.}6$ restoring beam. The result, Fig. 3(b), shows that some of the sidelobes have been removed. However, the overall deconvolution is poor – the first grating ring is not removed, and no large-scale structure is recovered. There are also faint 'edge effects' at the boundary of each sub-image, because the artificial flat background that maximum entropy deconvolvers can produce does not taper like the primary beam response.

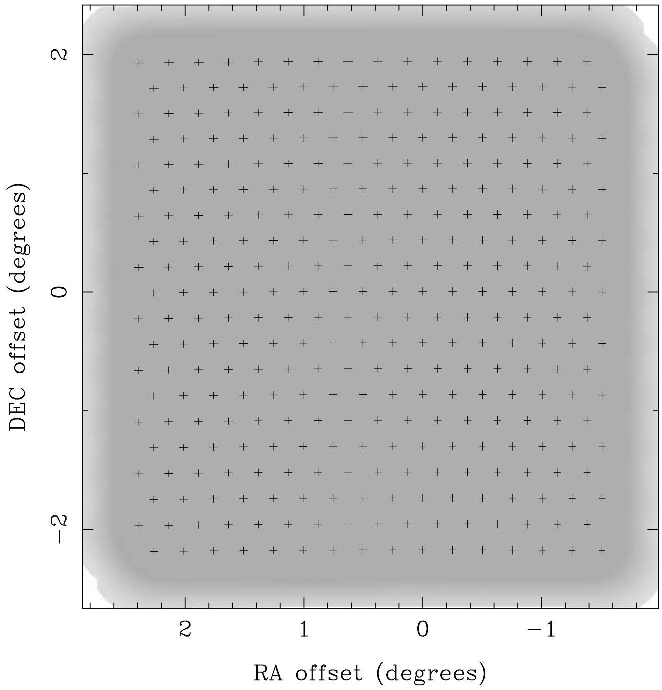

Fig. 2. The locations of ATCA pointings (+) in the Small Magellanic Cloud, superimposed on a grey-scale representation of the overall sensitivity function. The pointing grid is hexagonal with a $15'$ spacing.

The result of our approach (deconvolving the linearly mosaiced dirty cube) is shown in Fig. 4(a). Twenty iterations of a maximum entropy deconvolver were required for the algorithm to converge adequately. Much extended structure has been recovered, showing the advantages of joint deconvolution.

For comparison, we have implemented an image-plane-based version of Cornwell's (1988) algorithm, where the mosaicing is performed *during* the deconvolution (rather than before). Again, 20 iterations of a maximum entropy algorithm were required to reach a comparable convergence. The result is shown in Fig. 4(b). This algorithm applies a full primary beam correction at the edge of the mosaic, which we have removed for the purpose of comparison with Fig. 4(a).

The deconvolved images in Fig. 4 are qualitatively and quantitatively very similar. Both do a good job of deconvolving isolated HI clumps. For the confused structure in the bar region, both algorithms appear to remove the grating rings, recover extra $u - v$ spacing information, and

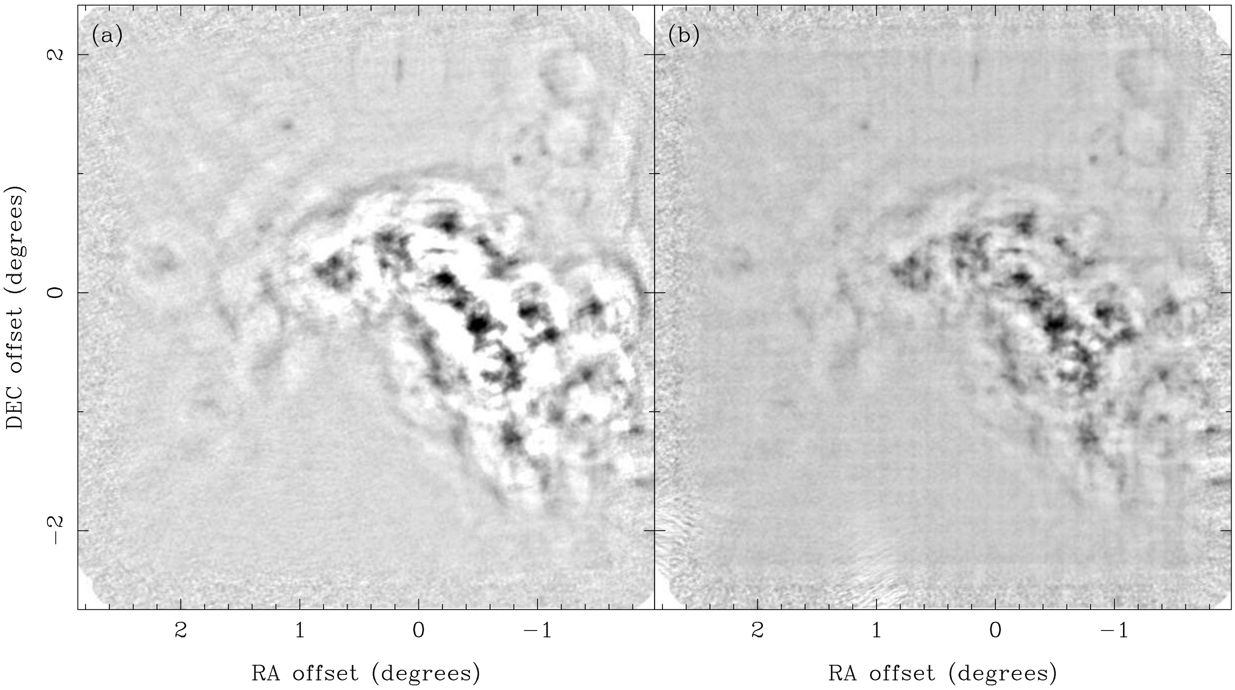

Fig. 3. Mosaiced HI images of the Small Magellanic Cloud at a heliocentric velocity of 123 km s^{-1}: (a) with no deconvolution (i.e. a linear mosaic of dirty images); (b) with deconvolution of the individual sub-images *before* mosaicing (i.e. a linear mosaic of deconvolved images). The sub-images were deconvolved with 20 iterations of a maximum entropy algorithm.

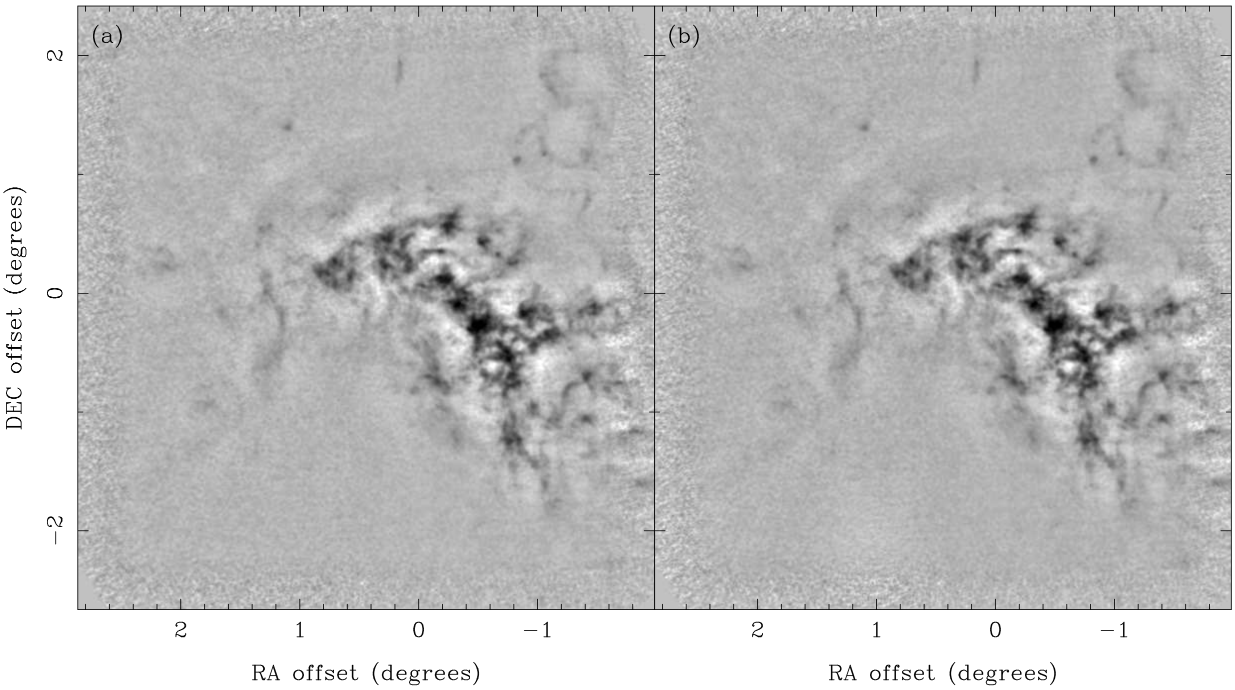

Fig. 4. Mosaiced HI images of the Small Magellanic Cloud at a heliocentric velocity of 123 km s^{-1}: (a) with deconvolution applied *after* mosaicing (i.e. a deconvolved linear mosaic of dirty images). This is the method advocated in this paper. (b) with deconvolution applied *during* mosaicing (i.e. a deconvolved non-linear mosaic of dirty images). This is the approach as advocated in Cornwell (1988). In both cases, 20 iterations of the same maximum entropy algorithm were used.

show structure up to the $34'$ scale of the primary beam. Both also exhibit 'bowls' around the bar region, resulting from missing zero-spacing information in this set of data.

As mentioned in Sect. 2.4, our approach has a significant practical advantage in terms of disk space. There is no need to make individual image cubes for each field, as we produce a linearly mosaiced dirty cube directly from the $u-v$ data. For 100 planes of the SMC cube, the linearly mosaiced cube is about 140 Mbyte in size, compared with about 1800 Mbyte for the 320 individual image cubes. This ratio is typical for any set of observations where the sky plane is sampled at close to the required Nyquist interval for antenna pointings, and is considerable for large spectral-line projects. For continuum projects, as discussed in Sect. 2.4, the overall saving is not as significant because both techniques require the storage of synthesized beams for each field (assuming the the usual case where these differ significantly from pointing to pointing). In this example, the synthesized beams occupy 67 Mbyte.

CPU tests were performed on a single 75 MHz processor on a SGI PowerChallenge computer. The most significant CPU usage occurs at the deconvolution step, where both joint techniques are well matched. Our approach takes 22.5 min per plane for the SMC (20 iterations) and uses 60 Mbyte of memory, while our implementation of the Cornwell method takes 21 min per plane and uses 105 Mbyte of memory. The Cornwell algorithm is marginally faster because of an implementation difference in the convolution modules.

4. Conclusions

We have presented an approach to imaging and deconvolving a mosaic experiment which has proven useful in processing a number of observations. The deconvolution process is such that the advantages of a 'joint deconvolution' are preserved. If nothing else, the packaging of the software has greatly reduced the burden on the user processing a mosaic observation. We have processed mosaiced experiments varying from a few to more that one thousand pointings.

Several avenues of possible further work are apparent. One obvious application is single-pointing, wide-field imaging with an array containing several different primary beam types (e.g. the Penticton array). Our 'mosaicing' techniques would handle such an observation in a natural fashion.

Our approach to deconvolution may be useful in other imaging systems with position-variant PSFs. It may be applicable if the PSF at a particular location, is, or can be approximated by, a weighted sum of some set of position-independent PSFs (i.e. as it is in the mosaic case). It would be most advantageous if the PSFs were also of finite extent. Our approach has some conceptual similarities to that of Waldram & McGilchrist (1990), where they deconvolve an image with position-variant PSF (caused by bandwidth smearing and the w-axis effect) using a set of beams. As a final possible application, combining our approaches to deconvolution and projection geometry with a number of the elements of the polyhedron method of Cornwell & Perley (1992) should produce a useful algorithm for eliminating the w-axis distortion problem. As the PSFs for this problem often cannot be approximated as being of finite extent, a deconvolution algorithm which determined residuals in the ungridded visibility domain (as the polyhedron method does) would be attractive.

Acknowledgements. We thank R.D. Ekers for his stimulating discussions and probing questions. D. McConnell and M.J. Kesteven developed the on-line software that makes the ATCA a good mosaicing instrument. We are grateful to the referee, F. Viallefond, for pointing out similarities with his work. RJS acknowledges support for part of this work from the Netherlands Organisation for Scientific Research under their visiting scientist scheme.

A. Projection geometry for non-coplanar baselines

To avoid w-axis distortions when mosaicing using an array with non-coplanar baselines, it is best to delay-track at different positions for different pointing. This normally results in different projections for different sub-images. Our aim is to convert (at least approximately) these different projections into a single sine projection by a linear transformation of the (u,v) coordinates before imaging. We choose a linear transformation because this preserves the convolution relationship between the dirty image, the synthesized beam, and the sky. An equivalent way of viewing this is as a manipulation of the phase term in the van Cittert-Zernicke integral.

We consider the problem of converting a single sub-image to a given projection geometry. Here we revert to two- or three-dimensional notation. Note that there are two relevant direction cosine coordinate systems in the image plane: one relative to the tangent point of the final projection (the global tangent point), and the other relative to the delay-tracking centre of the sub-image (for simplicity we assume that the delay-tracking and pointing centres are the same). Let (ℓ',m',n') be the direction cosines in the sub-image relative to the delay-tracking centre, and let (ℓ,m,n) be the direction cosines of the same point on the sky, but relative to the global tangent point. These are related by a three-dimensional rotation matrix:

$$\begin{pmatrix} \ell \\ m \\ n \end{pmatrix} = \boldsymbol{R} \begin{pmatrix} \ell' \\ m' \\ n' \end{pmatrix} . \tag{A1}$$

As $n' = \sqrt{1-\ell'^2-m'^2}$, we can linearize the above equation about the delay-tracking centre, $(\ell',m')=(0,0)$, which gives (after simplification)

$$\begin{pmatrix} \ell \\ m \end{pmatrix} \approx \boldsymbol{R_2} \begin{pmatrix} \ell' \\ m' \end{pmatrix} + \begin{pmatrix} \ell_p \\ m_p \end{pmatrix} . \tag{A2}$$

Here (ℓ_p,m_p) are the direction cosines of the delay-tracking centre of interest relative to the global tangent point, and $\boldsymbol{R_2}$ is upper-left 2×2 submatrix of $\boldsymbol{R}$. The intuitively reasonable shift by (ℓ_p,m_p) represents a shift of the final sub-image, and simply means that we do not have to rephase the data in any way before imaging. We now wish to express these coordinate manipulations in the visibility plane. We state the Fourier theorem that for vectors $\boldsymbol{\ell}$ and $\boldsymbol{u}$ and Fourier transform pairs $f(\boldsymbol{\ell})$ and $F(\boldsymbol{u})$, a linear transformation, $\boldsymbol{A}$, of the coordinate vector $\boldsymbol{\ell}$ gives transform pairs

$$f(\boldsymbol{A}\boldsymbol{\ell}) \rightleftharpoons |\det(\boldsymbol{A})|^{-1} F(\boldsymbol{A}^{-1\mathrm{T}}\boldsymbol{u}). \tag{A3}$$

Thus the coordinate transformation in the image plane can be expressed in the visibility plane as

$$\begin{pmatrix} u \\ v \end{pmatrix} = \boldsymbol{R_2}^{-1\mathrm{T}} \begin{pmatrix} u' \\ v' \end{pmatrix} \tag{A4}$$

where

$$\boldsymbol{R_2}^{-1\mathrm{T}} = \frac{1}{n}\begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \tag{A5}$$

$$a_{11} = \cos(\alpha_p - \alpha_0)\sin\delta_p \sin\delta_0 + \cos\delta_p \cos\delta_0 \tag{A6}$$

$$a_{21} = -\sin(\alpha_p - \alpha_0)\sin\delta_0 \tag{A7}$$

$$a_{12} = \sin(\alpha_p - \alpha_0)\sin\delta_p \tag{A8}$$

$$a_{22} = \cos(\alpha_p - \alpha_0) \tag{A9}$$

$$n = \sin\delta_0 \sin\delta_p + \cos(\alpha_p - \alpha_0)\cos\delta_p \cos\delta_0. \tag{A10}$$

Here (u', v') are the visibility coordinates computed relative to the delay-tracking centre, (α_p, δ_p) are the right ascension and declination of the delay-tracking centre, and (α_0, δ_0) are those of the global tangent point. Note that the term corresponding to $|\det(\boldsymbol{A})|^{-1}$ in the above Fourier theorem is just the $1/n$ factor that normally appears in the van Cittert-Zernicke equation.

A more intuitive approach to this transformation is depicted in Fig. 5. Points A, B, C and T are points on the celestial sphere. T is also the tangent point of the main projection plane – plane 1. Point A is some arbitrary point, and A' is its sine projection.

For a sub-image in the mosaic (the region of the sky beneath plane 2) we should normally image with a tangent point at the delay-tracking centre – point B. Without the transformation applied to the (u, v) coordinates, the resultant image will be a sine projection onto plane 2. The transformation of (u, v), which amounts to rotation and stretching operations in the image plane, is equivalent to further projecting from plane 2 to plane 1. Point B is correctly projected to B'. We can see, however, the error involved in the technique. Point C, at the edge of the sub-image, is projected to C', whereas the sine projection of C is C''. This projection error (the distance between C' and C'') will be

$$\epsilon = \sin\theta_1(1 - \cos\theta_2) \tag{A11}$$

$$\approx \frac{1}{2}\theta_1\theta_2^2. \tag{A12}$$

Here θ_1 is the angle between the delay-tracking centre and the global tangent point (the angle between B and T), and θ_2 is the angle between a point in a sub-image and its delay-tracking centre (the angle between C and B). Note that this error is no worse than the w-axis distortion present in a sub-image (see Perley 1989). If the w-axis distortion is tolerable in a sub-image, then so should the error in the above projection.

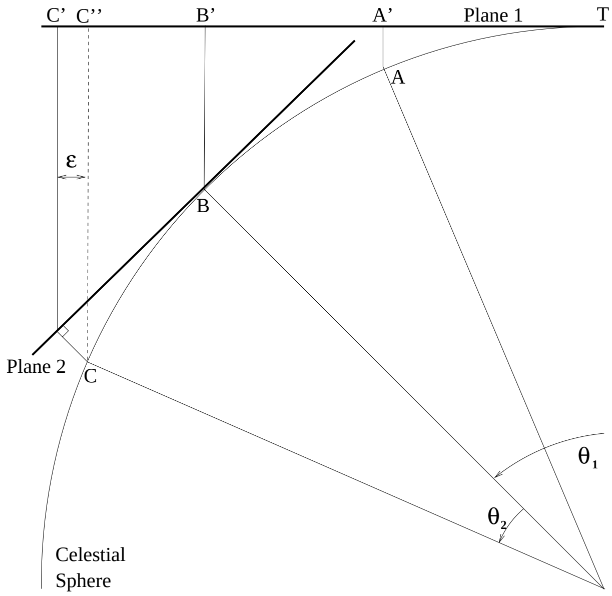

Fig. 5. Geometry of various projections involved in a mosaic observation with non-coplanar baselines.

The above error will be considerably smaller than the w-axis distortion that would result if we imaged using (u, v, w) coordinates rotated to a single tangent point (the worst error for this would be quadratic in the size of the mosaic, whereas the above technique is quadratic in the size of a sub-image).

References

Braun R. 1988a, NRAO millimeter array memo 46, Socorro, NM

Braun R. 1988b, NRAO millimeter array memo 54, Socorro, NM

Bremer M.A.R. 1994, in D.R. Crabtree, R.J. Hanisch, J. Barnes (eds.), Astronomical Data Analysis Software and Systems III, A.S.P., Provo, Utah, p. 175

Briggs D.S. 1995, Ph.D. thesis, New Mexico Institute of Mining and Technology, Socorro, NM

Brouw W.N. 1974a, in B. Adler, S. Fernach, M. Rotenberg (eds.), Methods in Computational Physics, Academic Press, New York, 14, p. 131

Brouw W.N. 1974b, NFRA Note 186, NFRA, Dwingeloo

Condon J.J., Cotton W.D., Greisen E.W., Yin Q.F., Perley R.A., Broderick J.J. 1994, in D.R. Crabtree, R.J. Hanisch, J. Barnes (eds.), Astronomical Data Analysis Software and Systems III, A.S.P., Provo, Utah, p. 155

Cornwell T.J. 1988, A&A, 202, 316

Cornwell T.J., Evans K.F. 1985, A&A, 143, 77

Cornwell T.J., Braun R. 1989, in R.A. Perley, F.R. Schwab, A.H. Bridle (eds), Synthesis Imaging in Radio Astronomy, A.S.P., Provo, Utah, p. 167

Cornwell T.J., Perley R.A. 1992, A&A, 261, 353

Cornwell T.J., Holdaway M.A., Uson J.M. 1993, A&A, 271, 697

Ekers R.D., Rots A.H. 1979, in C. van Schooneveld (ed.), Proc. IAU Col. 49, Image Formation from Coherence Functions in Astronomy, D. Reidel, p. 61

Gueth F., Guilloteau S., Viallefond F., 1995, in D.A. Green, W. Steffen (eds.), YERAC 95: Proc. XXVIIth Young European Radio Astronomers Conf., Cambridge Univ. Press

Holdaway M.A. 1992, NRAO millimeter array memo 73, Socorro, NM

Perley R.A. 1989, in R.A. Perley, F.R. Schwab, A.H. Bridle (eds), Synthesis Imaging in Radio Astronomy, A.S.P., Provo, Utah, p. 259

Sault R.J. 1984, VLA Scientific Memo 154, NRAO, Socorro, NM

Sault R.J. 1990, ApJ, 354, L61

Sault R.J., Teuben P.J., Wright M.C.H. 1995, in R. Shaw, H.E. Payne, J.E.E. Hayes (eds), Astronomical Data Analysis Software and Systems IV, A.S.P., Provo, Utah, p. 433

Schwab F.R. 1984, in J.A. Roberts (ed), Proc. IAU/URSI Conference on Indirect Imaging, Sydney, Cambridge University Press, p. 330

Sramek R.A., Schwab F.R., 1989, in R.A. Perley, F.R. Schwab, A.H. Bridle (eds), Synthesis Imaging in Radio Astronomy, A.S.P., Provo, Utah, p. 117

Staveley-Smith L., Sault R.J., McConnell D., Kesteven M.J., Hatzidimirtiou D., Freeman K.C., Dopita M.A. 1995, Publ. Astron. Soc. Aust., 12, 13

Steer D.G., Dewdney P.E., Ito M.R. 1984, A&A, 137, 159

Waldram E.M., McGilchrist M.M. (1990), MNRAS, 245, 532

This article was processed by the author using Springer-Verlag LaTeX A&A style file *L-AA* version 3.