Astron. Astrophys. 202, 316–321 (1988)

ASTRONOMY
AND
ASTROPHYSICS

Radio-interferometric imaging of very large objects

T.J. Cornwell

National Radio Astronomy Observatory*, P.O. Box 0, Socorro, NM 87801, USA

Received November 16, 1987; accepted March 1, 1988

Summary. Imaging of very large objects with high resolution at radio wavelengths cannot be performed by either a single dish or a radio interferometer alone, but requires the combination of observations made with both types of instruments. I describe the difficulties with schemes based upon simple linear combinations, and propose, instead, a scheme based upon joint deconvolution of all data, both interferometer and single dish. The Maximum Entropy method is particularly well-suited to this type of problem and results in a fast, simple algorithm. I demonstrate the performance of this algorithm on a VLA observation of a very extended object. I also discuss the relation of the proposed method to that of Ekers and Rots for obtaining short-spacing information from a scanned interferometer.

Key words: image processing – numerical methods – observational methods

1. Introduction

Radio-interferometry is a very powerful technique for providing high resolution images of stronomical objects. One of the most versatile radio-interferometric arrays, the Very Large Array (Napier et al., 1983) can produce images of the entire primary beam, about 1 degree when operating at 1420 MHz, at a resolution of about 1″. Arrays operating at millimeter wavelengths can produce images of comparable resolution, but only over a field of view about 1′. The field of view of all such interferometers is limited by the primary power reception pattern of the elements in the interferometer, which is in turn determined directly by the element diameter and the operating wavelength. The desire to increase the signal to noise ratio drives up the element size, and thus drives down the field of view. At millimeter wavelengths, this dilemma is particularly severe since the signals are weak, the receiver temperatures are as yet high, and some of the objects of interest, such as molecular clouds in our Galaxy, and nearby galaxies, tend to be large. Single dish techniques allow the imaging of very large objects, but to obtain high resolution one must combine single dish data with radio-interferometric data. Furthermore, to thus image objects spanning many primary beams of the interferometer elements, it is necessary to point the interferometric array in many different directions and combine the data. This general problem is addressed here.

Simple weighted addition of images of different parts of the sky to form a mosaic of a large field has been performed before (e.g. Brinks, 1984), and seems to present no insurmountable problems. Since sampling of the Fourier plane in radio imaging (see e.g. Thompson et al., 1986) is very often incomplete, the deconvolution procedures used must be non-linear. Therefore we might expect to obtain superior results by combining all the data *before or during* the deconvolution phase. I will concentrate upon the latter aspect. I will describe and demonstrate an imaging algorithm, based upon the Maximum Entropy principle (see e.g. Narayan and Nityananda, 1986), which can produce an image of a large region of the sky from interferometer and single-dish observations taken at a number of different pointing centers. Although other deconvolution algorithms exist, Maximum Entropy-based algorithms have important advantages for our application (see Pearson and Readhead, 1984, or Cornwell, 1986 for a review of deconvolution algorithms in radio-astronomy).

I will also demonstrate the connection between this new approach and the proposal of Ekers and Rots (1978) for recovering short spacing information from a single interferometer. Specifically, I will show that the new approach allows implicit estimation of short spacings in a rather more-general way.

The organization of this paper is as follows: in Sect. 2, I provide a unified description of single-dish and interferometer imaging in simple heuristic form and in mathematical form. In Sect. 3, I give a brief account of the Maximum Entropy principle, and its prior applications to radio astronomical imaging. Section 3 also shows how to apply the MEM to mosaicing, and includes discussion of a number of theoretical and practical complications. Finally, an application to real data is shown in Sect. 4 and Sect. 5 contains a discussion of possible future work.

2. Unified description of interferometer and single-dish imaging

Although the practical details are quite different, single dishes and interferometers perform the same basic operation: they provide measurements of the Fourier transform of some region of the sky. The finite element size in each forces a measurement to be a linear weighted sum of a range of spatial frequencies, each of which can pass through the aperture. Indeed, perhaps the best way to understand the information collected by a single dish is to split the aperture into sub-apertures, each of which forms an interferometer with every other sub-aperture. The spacings of the

* The National Radio Astronomy Observatory (NRAO) is operated by Associated Universities, Inc., under contract with the National Science Foundation

interferometers range in length from zero up to the dish diameter. The output power from a single dish is then simply a linear combination of these various spatial frequencies; the actual linear coefficients being a function of the position of the receiver feed with respect to the optical axis, and of the optical axis with respect to the sky. Thus by varying either the feed position or the optical axis, it is possible to obtain different linear combinations of the values for these notional sub-interferometers. Given enough of these linear combinations, one can untangle the individual values and thus image explicitly the sky over a region corresponding to the offsets used. Of course, this is just a convoluted way of expressing the usual practice of single dish imaging, but it does reveal the intimate connection to interferometry.

The interferometer can be analysed in just the same way: the aperture can be split conceptually into sub-apertures and all possible sub-interferometers formed. The output is then just a linear combination of all the sub-interferometers, where, in this case, the weights are determined by three variables: the positions of the receiver feeds relative to the optical axes of the elements of the interferometer, the pointing centers of the elements of the interferometer, and the phase center of the interferometer. Similarly to the single-dish case, the values appropriate to the sub-interferometers can be found by observing with a number of different arrangements, each of which yields a different linear combination. In essence, this is the scheme of Ekers and Rots (1978), in which the interferometer is scanned over some region, and the visibility recorded as a function of scan offset. The sub-interferometer visibilities are then obtained by inverse Fourier transforming the observed visibilities as a function of the offset.

A simple mathematical description of the operation of both an interferometer and a single dish can be given in terms of the pointing-position-dependent visibility function measured. Let $\boldsymbol{x}_p$ be the pointing position of the elements of the telescope on the sky, and $\boldsymbol{u}$ be the vector separation of the telesopes measured in wavelengths, as seen from some direction. We describe the "primary beam" or sensitivity to emission of an element by the function $A(\boldsymbol{x}-\boldsymbol{x}_p)$, which tapers off to zero for large offsets from the pointing center. The visibility function sampled at pointing position $\boldsymbol{x}_p$ is then basically the Fourier transform of the true sky brightness $I(\boldsymbol{x})$ weighted by the primary beam A (see e.g. Thompson et al., 1986).

$$V(\boldsymbol{u},\boldsymbol{x}_p)=\int A(\boldsymbol{x}-\boldsymbol{x}_p)\,I(\boldsymbol{x})\,e^{j2\pi\boldsymbol{u}\cdot\boldsymbol{x}}\,d\boldsymbol{x}\,. \tag{1}$$

For a single dish, $\boldsymbol{u}$ is zero and then:

$$V(0,\boldsymbol{x}_p)=\int A(\boldsymbol{x}-\boldsymbol{x}_p)\,I(\boldsymbol{x})\,d\boldsymbol{x}\,. \tag{1'}$$

The basic measurable quantity is just $V(\boldsymbol{u},\boldsymbol{x}_p)$, which is four dimensional, whereas the required sky brightness, $I(\boldsymbol{x})$, is two dimensional. This reflects the fact that an image can be formed in many different ways: either mainly by single dish observations taken at many different $\boldsymbol{x}_p$ with a small primary beam or mainly be interferometer observations at a fixed $\boldsymbol{x}_p$ but with an extended primary beam A, or, indeed, by any combination of these two extremes. Practicality is the determining factor.

If complete coverage of the $\boldsymbol{u}$ plane is obtained for a given pointing then Eq. (1) can be inverted to yield an estimate of the true sky brightness:

$$\int V(\boldsymbol{u},\boldsymbol{x}_p)\,e^{-j2\pi\boldsymbol{u}\cdot\boldsymbol{x}}\,d\boldsymbol{u}=A(\boldsymbol{x}-\boldsymbol{x}_p)\,I(\boldsymbol{x})\,. \tag{2}$$

In practice, however, one cannot simply invert the equation because only limited sampling is possible. Let the sampling for a given pointing center be described by a function $s(\boldsymbol{u},\boldsymbol{x}_p)$, which is usually representable by a collection of Dirac δ-functions, and let its Fourier transform with respect to $\boldsymbol{u}$ be $S(\boldsymbol{x},\boldsymbol{x}_p)$. Then it is simple to show that Fourier inversion of the sampled visibility function yields:

$$\int s(\boldsymbol{u},\boldsymbol{x}_p)\,V(\boldsymbol{u},\boldsymbol{x}_p)\,e^{-j2\pi\boldsymbol{u}\cdot\boldsymbol{x}}\,d\boldsymbol{u}=S(\boldsymbol{x},\boldsymbol{x}_p)*[A(\boldsymbol{x}-\boldsymbol{x}_p)\,I(\boldsymbol{x})]\,, \tag{3}$$

where $*$ denotes convolution with respect to $\boldsymbol{x}$. This is the well-known convolution relationship: the Fourier inversion of the sampled visibility function yields the answer which would be obtained with complete coverage, but convolved with a point spread function. The effect of the convolution is to filter out the non-sampled spatial frequencies. Linear methods, such as Wiener filtering (e.g. Andrews and Hunt, 1977), cannot recover these non-sampled spatial frequencies. Instead, one must use a non-linear deconvolution algorithm, examples of which are the CLEAN algorithm (Högbom, 1974) and the Maximum Entropy Method (Narayan and Nityananda, 1986). The simplest method of imaging a large field proceeds as follows:

1. Sample the visibility function for an object at a number of pointing centers.

2. For each pointing center, perform the necessary deconvolution to remove the effects of the point spread function B.

3. Form a full image by taking a sum of the images for the individual pointing centers, weighted by the corresponding primary beam, $A(\boldsymbol{x}-\boldsymbol{x}_p)$.

This method is perfectly satisfactory to the extent that step 2 is linear. In reality, it may be highly non-linear. The deconvolution of each pointing center can thus be improved by taking account of the deconvolution of the neighbouring images. In other words, the deconvolution of all the pointing centers should be performed *jointly* rather than separately. As an example of this, consider the deconvolution of a nearly unresolved source lying within the overlap region of a number of pointings. If the source is sufficiently small that the gradients of the primary beams can be neglected then the optimum strategy is to add the data together before deconvolution. The sidelobes of the combined beam would probably be substantially lower, and so the deconvolution would be easier.

I can now state the problem rather precisely: we wish to find an estimate of the sky brightness $I(\boldsymbol{x})$ which is consistent with the visibility data from many pointing centers. There will be many such feasible images, so we will have to use some criterion to select just one of the many. In the next two sections I will describe how the Maximum Entropy method can perform this selection, and how it can be used for the joint deconvolution of all the data.

Before proceeding, I will discuss the connection to the proposal of Ekers and Rots (1979), which naturally leads on to an examination of the sampling requirements. Suppose that we perform a Fourier transform of Eq. (1) with respect to the pointing center position $\boldsymbol{x}_p$. Let $a(\boldsymbol{u})$ be the Fourier transform of the primary beam, $A(\boldsymbol{x})$, and $i(\boldsymbol{u})$ be the Fourier transform of the true sky brightness. Using the fact that $a(\boldsymbol{u})$ is even, it is easy to show that the Fourier transform of $V(\boldsymbol{u},\boldsymbol{x}_p)$ with respect to $\boldsymbol{x}_p$ yields:

$$V(\boldsymbol{u},\Delta\boldsymbol{u})=\frac{\int V(\boldsymbol{u},\boldsymbol{x}_p)\,e^{-j2\pi\Delta\boldsymbol{u}\cdot\boldsymbol{x}_p}\,d\boldsymbol{x}_p}{\int A(\boldsymbol{x}_p)\,d\boldsymbol{x}_p}=\frac{a(\Delta\boldsymbol{u})}{a(0)}\,i(\boldsymbol{u}+\Delta\boldsymbol{u})\,, \tag{4}$$

where $\Delta\boldsymbol{u}$ is the variable conjugate to $\boldsymbol{x}_p$. Therefore, by scanning continuously over some region, and then Fourier inverting with respect to the scan position, one can obtain an estimate of the true visibility function weighted by the Fourier transform of the primary beam of the elements. This is exactly the proposed scheme of Ekers and Rots.

Scanning is not actually required for complete sampling of all the spatial frequencies: we only need samples of $V(\boldsymbol{u}, \boldsymbol{x}_p)$ spaced in $\boldsymbol{x}_p$ by an increment $\Delta \boldsymbol{x}_p$ equal to $\lambda/2D$ where λ is the observing wavelength, and D is the diameter of the elements. To formalize this, we can introduce a sampling function in $\boldsymbol{x}_p$-space, similar to the sampling functions in $\boldsymbol{u}$-space: let $P(\boldsymbol{x}_p)$ be this sampling function, which would typically be a collection of Dirac δ-functions. The normalized Fourier transform of the sampled visibility function yields:

$$V(\boldsymbol{u}, \Delta \boldsymbol{u}) = \frac{\int P(\boldsymbol{x}_p)\, V(\boldsymbol{u}, \boldsymbol{x}_p)\, e^{-j2\pi \Delta \boldsymbol{u} \cdot \boldsymbol{x}_p}\, d\boldsymbol{x}_p}{\int A(\boldsymbol{x}_p)\, d\boldsymbol{x}_p}$$

$$= p(\Delta \boldsymbol{u}) * \left[\frac{a(\Delta \boldsymbol{u})}{a(0)}\, i(\boldsymbol{u} + \Delta \boldsymbol{u}) \right] \tag{5}$$

For complete, uniform sampling, this reduces to Eq. (4). The sampling requirements can be deduced from this equation: if $P(\boldsymbol{x}_p)$ is represented by a collection of δ-functions, then so is $p(\Delta \boldsymbol{u})$. Thus aliasing of power along the $\Delta \boldsymbol{u}$ axis will occur if the extent of the sensitivity function $a(\Delta \boldsymbol{u})$, at most twice the element diameter, is greater than the separation of the δ-functions in $p(\Delta \boldsymbol{u})$. Hence, to avoid aliasing, the spacing of pointings must obey:

$$|\Delta x_p| \leqq \lambda/2D\,. \tag{6}$$

If the signal-to-noise is sufficiently weak then poorer sampling may be allowed with little consequent degradation in image quality. Note that this limit applies directly to simple single-dish imaging where it is often violated, but usually the illumination of the aperture is arranged to fall off severely near the edge so that the aliased power is negligible.

3. The maximum entropy method (MEM)

Since the sampling of the Fourier plane is incomplete, there are many images consistent with the observed visibility data (Bracewell and Roberts, 1954). The deconvolution problem is that of selecting one image from the many feasible. In the maximum entropy method (e.g. Jaynes, 1982), the image selected is that which fits the data, to within the noise level, and also has maximum entropy. The use of the term *entropy* has lead to great disagreement over the rationale for the MEM. Narayan and Nityananda (1986) have reviewed the use of the MEM in image processing. They discuss extensively the philosophical issues involved, as well as the practical details of implementing a Maximum Entropy algorithm, and so I will avoid re-stating the various arguments. Instead, I will use a "lowest common denominator" justification and define entropy is something, which when maximized, produces a positive image with a compressed range in pixel values. The compression in pixel values forces the MEM image to be "smooth" and the reconstruction stable, and the positivity forces interpolation of unmeasured Fourier components. MEM also produces super-resolution on bright, isolated objects, but it is often unreliable (see e.g. Pearson and Readhead, 1984), and is of secondary importance in this application. There are many possible forms of this extended definition of entropy, see e.g. Narayan and Nityananda (1986), but one of the most suitable for general purpose use is:

$$\mathscr{H} = -\sum_i I_i \ln\left(\frac{I_i}{M_i e}\right), \tag{7}$$

where I_i is the brightness in the i'th pixel of the MEM image, and M_i is a pixel of a default image incorporated to allow a priori knowledge to be used. For example, a low resolution image of the object can be used to good effect as the default. The base of natural logarithms in the denominator ensures that the unconstrained maximum entropy image is the default image. The utility of this form for the entropy relative to the other forms lies mainly in its more-stable behaviour (Narayan and Nityananda, 1986).

A requirement that each data point be fitted exactly is nearly always incompatible with the positivity of the MEM image. Consequently, data are usually incorporated in a constraint that the fit, χ^2, of the predicted data to that observed, be close to the expected value.

Algorithms for solving this maximization problem have been reviewed by Narayan and Nityananda (1986). Algorithms specifically for radio-interferometry have been developed by Wernecke and D'Addario (1976), by Cornwell and Evans (1985), and by Skilling and Bryan (1984). The Cornwell-Evans algorithm is coded in NRAO's Astronomical Image Processing System (AIPS) as task "VM".

From a purely pragmatic stand-point, the attraction of the MEM is that it defines the best image as the solution to a simple optimization problem: maximizing the entropy while fitting to the data. Thus extra information can simple be added as constraints in the optimization. Thus, in our application of mosaicing, we can simply regard all the separate samples of $V(\boldsymbol{u}, \boldsymbol{x}_p)$ as constraints. The relevant form of χ^2 is:

$$\chi^2 = \sum_p \sum_r \frac{|V(\boldsymbol{u}_r, \boldsymbol{x}_p) - \hat{V}(\boldsymbol{u}_r, \boldsymbol{x}_p)|^2}{\sigma^2_{V,r,p}} \tag{8}$$

where the caret denotes the predicted value, and $\sigma^2_{V,r,p}$ is the variance in the visibility for the r'th visibility sample for the p'th pointing sample.

In the optimization the gradient of χ^2 is required for the updating of the current MEM iterate, and χ^2 is needed for control of the iteration (see Cornwell and Evans, 1985, for further details). The calculation of χ^2 is worth discussing in some detail because it reveals how weighting of the data and suppression of the sidelobes both occur automatically as part of the optimization. It can be described as following:

1. For a given pointing center $\boldsymbol{x}_p$, multiply the current trial MEM image by the corresponding primary beam, $A(\boldsymbol{x} - \boldsymbol{x}_p)$, and Fourier transform to obtain the predicted visibilities.

2. Subtract the observed visibilities from those predicted to obtained residual visibilities. Sum appropriately and add to the accumulated χ^2.

3. Inverse Fourier transform the residual visibilities, weighted by the inverse variances, and taper the resulting image by the primary beam. Add this residual image to the accumulated residual image.

4. Repeat steps 1–3 for all pointings. Multiply the final total residual image by 2 to obtain the gradient of χ^2 with respect to the MEM image.

This description applies equally well to both interferometer and single-sish data. For the latter, only the zero spatial-frequency points of the Fourier transforms are required: this simply means that we sum the corresponding image.

Let us now examine steps in more detail. The reasons for steps 1 and 2 are obvious, those for step 3 rather less so. The second tapering, of the residual image, weights down any residual sidelobes by the primary beam for a given pointing. The rationale for this step is that a given pointing cannot tell us anything about

the emission outside the primary beam, and therefore the sidelobes must be weighted down correspondingly. I should emphasize that this result, while quite intuitive, arises naturally from the definition of χ^2. Note also that the final residuals are correctly weighted by both the primary beam of a given pointing, and by the noise on a given visibility sample. Again, this is required by the formalism used.

The algorithm is available as task "VTESS" in AIPS. It is similar to that described by Cornwell and Evans (1985), except for the calculation of χ^2 and the gradient of χ^2 with respect to the MEM image, and some other minor differences concerned with improvements in the control of the iterations. The algorithm will converge in about 20–30 iterations in typical cases. The main computational cost is in calculations of χ^2 and its gradients. As Cornwell and Evans (1985) discuss, the calculation per field can be performed by convolving a trial image with the synthesized point spread function. The total cost per pointing center is then about 60–90 2-dimensional FFTs, each of which has a size determined by the primary beam. Thus, compared to combination by simple linear addition without any deconvolution step, the cost is increased by about 60–90 times. Compared to separate deconvolution of the individual fields followed by linear combination, the incremental cost is small.

There are a number of drawbacks to the use of MEM: first, since an MEM image is constrained to be smooth and positive, the image is biased i.e. the residuals are correlated with the image. Second, MEM works rather poorly in cases where a small object is seen superimposed on an background (Narayan and Nityananda, 1986). Third, the resolution is signal-to-noise dependent. However in radio-interferometry, our major requirement for deconvolution is to interpolate the visibility into *interior* regions of unsampled visibilities, rather than to extrapolate as is often the case in other imaging modalities. I have found that an expedient solution to all of these problems is to pre-taper the visibility data, which effectively converts any point sources into resolved sources. This removes the possibility of super-resolution, but considerably improves the interpolation of missing Fourier samples. Lannes et al. (1987) have discussed this point in some detail, and have proposed and demonstrated a less ad-hoc scheme for Fourier interpolation.

To summarize this section, I have shown that the MEM can be easily adapted to solve the mosaicing problem, and that the resulting algorithm has a number of very nice properties including correct weighting of the data, and suppression of sidelobes outside the primary beam. Given that deconvolution is necessary, the extra cost in performing the deconvolution jointly is small.

4. An example

The most important application of this mosaicing method is to the imaging of objects which span many primary beamwidths. Such an object sould also be imaged *without mosaicing* by an array composed of smaller diameter elements. For example, a mosaiced VLA ($D=25$ m) image of an object could be compared with a conventional Penticton ($D=8.6$ m, Roger et al., 1973) image. I have done exactly this.

The galactic object Simeis 57 was imaged with both the VLA D-configuration (Napier et al., 1983) and the Penticton array. The observing frequencies were 1390 MHz for the VLA and 1428 MHz for Penticton. The Penticton observations were of a single field. I made an image of this single field using the mosaicing algorithm. A gaussian model of the primary beam was used, where the full-

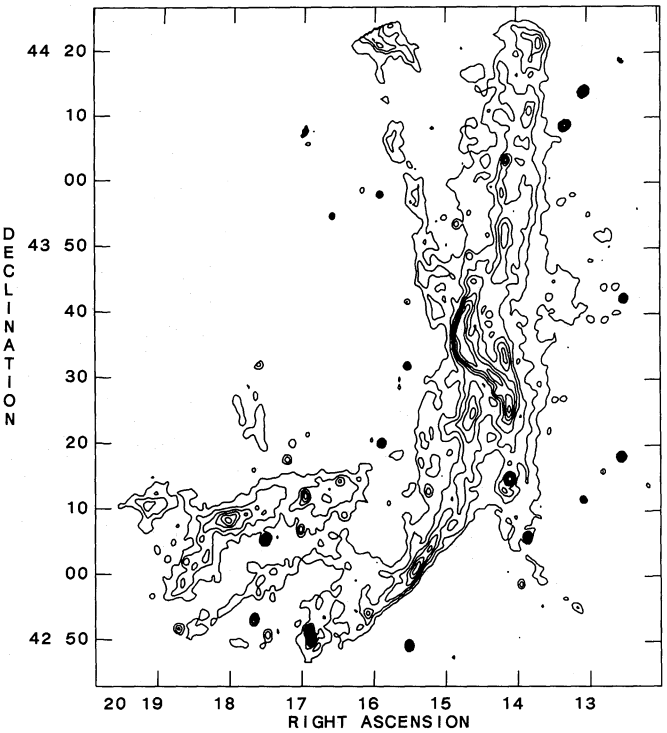

Fig. 1. Penticton image of Simeis 57 at $\lambda 20$ cm. The contour levels are 0.3, to 6.0 mJy/pixel in steps of 0.3 mJy/pixel, where the pixel size is 15″ by 15″. The tick marks are at intervals of minutes of time in right ascension, and minutes of arc in declination

width-to-half-maximum was 104′. As discussed in the previous section, to improve the final image quality, the visibility data was tapered with a circular gaussian corresponding to an image-plane gaussian of FWHM 45″. The algorithm converged in 12 iterations, or about 30 2-D FFTS, to a fit of 900 μJy/beam. The image is shown in Fig. 1. For the VLA observations, 22 pointing centers were chosen to span the brightest part of the emission shown by the Penticton image, and placed on a grid of spacing 15′ so that the sampling requirement, Eq. (6), is obeyed. The primary beam was modeled by a polynomial (Napier and Rots, 1982), truncated at the 7% level. The visibility data were tapered with the same circular gaussian as the Penticton data. The algorithm converged in 22 iterations, or about 1000 2-D FFTs, to a fit of 1.3 mJy/beam for each field. The mosaiced image is shown in Fig. 2, along with crosses at the pointing centers. For comparison, a non-mosaiced image is shown in Fig. 3. This was formed by deconvolving each field individually, and then combining pixel-by-pixel with a weight given by the primary beam divided by the noise level:

$$I_{\text{linear}}(\boldsymbol{x}) = \frac{\sum_p \frac{A(\boldsymbol{x}-\boldsymbol{x}_p)}{\sigma^2_{I,p}} I_{\text{separate}}(\boldsymbol{x})}{\sum_p \frac{A(\boldsymbol{x}-\boldsymbol{x}_p)^2}{\sigma^2_{I,p}}}, \tag{9}$$

where $\sigma^2_{I,p}$ is the noise variance for the p'the pointing center image. The main results of this comparison are:

– The mosaiced VLA and the Penticton image show substantial agreement in both fine and broad detail. A Fourier transform of the difference of the mosaiced region shows that the disagreement is largest for spacings below about 18 meters. Since the shortest measured VLA spacing is about 28 meters, the mosaicing algorithm has allowed interpolation inwards by about 10 meters,

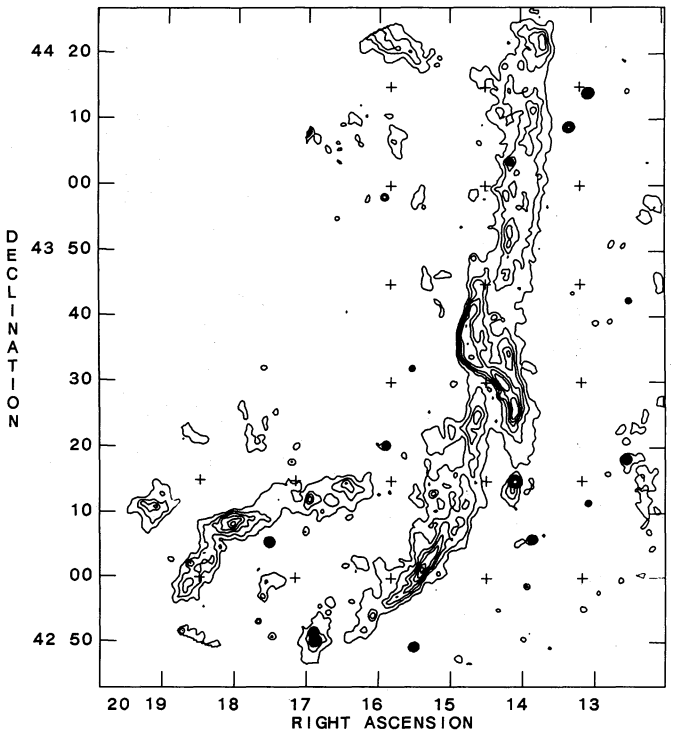

Fig. 2. Mosaiced VLA image of Simeis 57 at $\lambda = 20$ cm. The crosses refer to the 22 pointing centers used. The contour levels are those used in Fig. 1

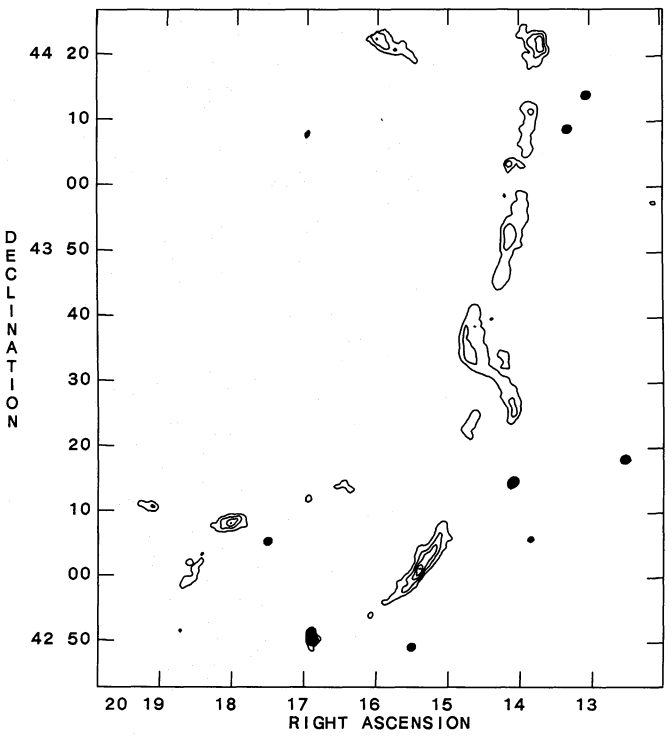

Fig. 3. VLA image of Simeis 57 formed by linear combination of separately deconvolved images. The contour levels are those used in Fig. 1

or just less than half a dish diameter. This is consistent with the limitations of the model of the primary beam, which is truncated at about twice the FWHM. To confirm this results, the Penticton image was remade after removal of visibilities corresponding to the inner portion of the u, v plane. The minimum discrepancy between the mosaiced VLA image and the revised Penticton image is indeed obtained for a cutoff in the range between 16 m and 20 m.

– The mosaiced and non-mosaiced VLA images show the same fine detail, but the broad emission is much attentuated in the latter. This demonstrates the virtue of performing the deconvolution of the fields jointly rather than separately.

– When deconvolved separately, each VLA field converged in about 10–15 iterations, as compared to 22 iterations when deconvolved jointly. Thus the extra cost in the mosaicing operation is no more than a factor of two for the number of fields processed here.

This test is described in more detail elsewhere (Cornwell, 1987b). An application to the combination of interferometer and single dish data is given by Mundy et al. (1987). Simple computer simulations are presented by Cornwell (1984).

5. Discussion

I have shown that mosaicing can be performed by a relatively simple Maximum Entropy Method. The MEM is well-suited to this application for a number of reasons: first, the formalism is very flexible, and is easily extended to mosaicing. Second, it is most efficient on large objects. Third, fast, simple algorithms for the basic optimization problem exist.

Other algorithms could also be adapted to the same purpose. Best suited would be algorithms based upon optimization of some measure such as the various "entropies" (Narayan and Nityananda, 1986). Ironically, the most-often used radio-astronomical deconvolution algorithm, CLEAN (Högbom, 1974; Schwarz, 1978, 1979), may not be very useful, since mosaicing will be most important for very extended objects, spanning many resolution elements, and since CLEAN is notoriously slow and inefficient on such objects. However, the modification of CLEAN by Steer et al. (1983) does not suffer this drawback, and may be a good candidate.

As with many other sophisticated data reduction techniques, mosaicing requires that the data be well-understood. Most important are the primary beams, the pointing of the elements, and the noise levels of the samples (Cornwell, 1987a). In current telescopes, the first two may be so poorly understood that mosaicing is impracticable. In such cases, a hybrid scheme is preferable, in which only the shortest spacings are processed by mosaicing (Braun, 1988). For future instruments, however, such as the NRAO Millimeter Array, it will be possible to optimize for these factors. For example, for mosaicing to work well, it may be preferable to emphasize stability of the illumination of an element over efficiency, and over the sidelobe level. Much more work is required to investigate the effect of all the possible systematic errors on the quality of images obtained by mosaicing.

Finally, while the computational cost of the joint deconvolution scheme described here is not much more than that of a more-conventional scheme, any form of mosaicing will place increased requirements on the computational resources used for radio-astronomical imaging. The increased requirements arise in two ways: from the larger image sizes required, and from the necessity to take account of the three-dimensional sampling of the visibility function performed by many interferometers. However, provided that we can solve this and the other possible obstacles,

mosaicing should prove to be a very powerful technique for imaging very large objects at high resolution.

Acknowledgements. Although it was apparent for a long time that MEM could usefully be applied to problems such as mosaicing, the final impetus was provided by the NRAO Millimeter Array project, and in particular by Frazer Owen. I thank him for this encouragement, and for subsequent discussions. I also acknowledge very useful discussions with Ron Ekers, Robert Braun, Bob Hjellming, Lee Mundy, and many other radio astronomers, both here at NRAO and elsewhere. I thank Rick Perley for help with the VLA observations of Simeis 57, and John Bally and Frank Israel for providing a copy of the Penticton data. Finally, I am very grateful to the AIPS group.

References

Andrews, H.C., Hunt, B.R.: 1977, *Digital Image Restoration*, Prentice-Hall
Bracewell, R.N., Roberts, J.A.: 1954, *Australian J. Phys.* **7**, 615
Braun, R.: 1988, NRAO millimeter array memo 46
Brinks, E.: 1984, Ph.D. thesis, Leiden
Cornwell, T.J.: 1984, NRAO millimeter array memo 24
Cornwell, T.J.: 1987, NRAO millimeter array memo 42
Cornwell, T.J.: 1987, NRAO millimeter array memo 43
Cornwell, T.J., Evans, K.F.: 1985, *Astron. Astrophys.* **143**, 77
Cornwell, T.J.: 1986, Proc. Second NRAO-VLA workshop on *Synthesis Imaging*, Socorro, NM, 1985, eds. R.A. Perley, F.R. Schwab, A.H. Bridle, NRAO Publ., Greenbank, West Virginia
Ekers, R.D., Rots, A.H.: 1979, *Image Formation from Coherence Functions in Astronomy*, Proc. IAU Coll. **49**, ed. C. van Schooneveld, Reidel, Dordrecht, p. 61
Högbom, J.: 1974, *Astrophys. J. Suppl.* **15**, 417
Jaynes, E.T.: 1982, *Proc. IEEE* **70**, 939
Lannes, A., Roques, S., Casanove, M.J.: 1987, *J. Opt. Soc. Am.* A**3**, 1908
Mundy, L.G., Cornwell, T.J., Masson, C.R., Scoville, N.Z., Baath, L.B., Johansson, L.: 1987, *Astrophys. J.* (in press)
Napier, P.J., Rots, A.H.: 1982, NRAO test memo 134
Napier, P.J., Thompson, R.T., Ekers, R.D.: 1983, *Proc. IEEE* **71**, 1295
Narayan, R., Nityananda, R.: 1986, *Ann. Rev. Astron. Astrophys.* **24**, 127
Pearson, T.L., Readhead, A.C.S.: 1984, *Ann. Rev. Astron. Astrophys.* **22**, 97
Roger, R.S., Costain, C.H., Lacey, J.D., Landecker, T.L., Bowers, F.K.: 1973, *Proc. IEEE* **61**, 127
Schwarz, U.J.: 1978, *Astron. Astrophys.* **65**, 345
Schwarz, U.J.: 1979, in *Image Formation from Coherence Functions in Astronomy*, Proc. IAU Coll. 49, ed. C. van Schooneveld, Reidel, Dordrecht, p. 261
Skilling, J., Bryan, R.K.: 1984, *Monthly Notices Roy. Astron. Soc.* **211**, 111
Steer, D.G., Dewdney, P.E., Ito, M.R.: 1984, *Astron. Astrophys.* **137**, 159
Thompson, A.R., Moran, J.M., Swenson, G.W.: 1986, *Interferometry and Synthesis in Radio Astronomy*, Wiley, New York
Wernecke, S.J., D'Addario, L.R.: 1976, *IEEE Trans.* **C-26**, 351