Very Long Baseline Interferometry and the VLBA
ASP Conference Series, Vol. 82, 1995
J. A. Zensus, P. J. Diamond, and P. J. Napier (eds.)

Printed June 13, 1995

CHAPTER 16

Multi-Frequency Synthesis

J. E. CONWAY

NATIONAL RADIO ASTRONOMY OBSERVATORY

R. J. SAULT

AUSTRALIA TELESCOPE NATIONAL FACILITY

Abstract Theory and practice of the technique of multi-frequency synthesis (MFS) are described, with emphasis on its application to VLBA observations.

Contents

16.1 Introduction

Multi-frequency synthesis (MFS) is a method of dramatically improving (u, v) coverages, and hence improving image fidelity, by observing at multiple frequencies with a bandspread up to about 30%. We analyze the effects of the frequency dependence of source structure expected within VLBI images and show that in many cases spectral effects will be comparable or less than the thermal noise. More generally in MFS we must correct for such spectral effects and we present several algorithms capable of carrying out such corrections. We discuss the unique capabilities of the VLBA for MFS, including its wide frequency coverage, its ability to sample several frequencies over a wide band simultaneously and its ability to rapidly change observing frequency. These capabilities, combined with better amplitude calibration, will allow multi-frequency synthesis and multi-frequency self-calibration techniques to be fully exploited for VLBI imaging.

16.2 Image Fidelity

In any interferometric array containing a finite number of elements the (u, v) plane will only be partially sampled. Particularly when there are relatively few antennas, as in most VLBI experiments, there will be large gaps in the (u, v) coverage (see figure 16.1). In order to produce astronomically useful images we must invariably use non-linear *deconvolution* methods such as CLEAN, MEM, etc. (see chapter 3); these algorithms attempt to interpolate from the measured (u, v) points across gaps in the (u, v) coverage (and also extrapolate somewhat beyond the outer edge of the sampled (u, v) coverage).

The success of the interpolation in the (u, v) plane, and hence the quality of the resulting image depends, in general, on the ratio of the source complexity to the quality of the (u, v) coverage. If the number of non-zero beam areas in the images is less than the number of independent (u, v) points then the resulting image is likely to be a good representation of the sky (see chapter 3). However, as the image gets more complex it is increasingly likely that the deconvolved image will contain significant *reconstruction errors.* Hence deconvolved images have a finite fidelity which depends on the quality of the (u, v) coverage. This limited fidelity has a significant effect on the science that can be done with VLBI because we are often not sure which features in an image are real. Furthermore even subtle errors in an image which are astronomically insignificant by themselves will become significant if we attempt to look for changes between two epochs (a common goal of VLBI experiments).

16.3 Multi-Frequency Synthesis

To improve image fidelity we clearly need to improve the (u, v) coverage. This might at first appear impossible without increasing the number of antennas.

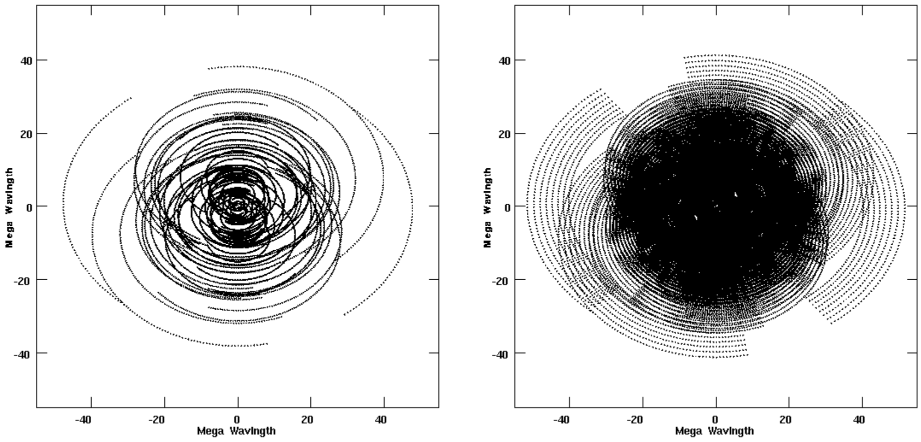

Figure 16.1: *Left (a):* VLBA (u, v) coverage for a full track at $\delta = 50°$. *Right (b):* Using MFS observations with 8 frequencies spread over 25%.

However, we can note that the dimensions of the (u, v) plane are measured in wavelengths. Therefore by observing at several frequencies a single baseline can sample several ellipses in the (u, v) plane, each with a different size. By using several frequencies we can therefore fill in the gaps in the single frequency (u, v) coverage. To this technique we give the name multi-frequency synthesis (MFS). For the VLBA using 8 frequencies over a 15%–30% frequency range MFS can produce a dramatic improvement in (u, v) coverage as shown in figure 16.1*a*, *b*.

The principle behind MFS has been known since the earliest days of radio interferometry (McCready, Pawsey and Payne-Scott 1947). However, it has not been much utilized because of an obvious problem—in general the intensity distribution of the source is itself frequency dependent. In this chapter we discuss the effects of this frequency dependence on MFS imaging. We will find that, at least for the 15%-30% frequency ranges required in order to fill in the VLBA (u, v) coverage, a surprising amount of frequency dependent structure can be tolerated.

We shall also find that even where *spectral effects* are significant there exist algorithms for removing them from the data and producing high fidelity MFS images (see section 16.5). Such algorithms work because for continuum emission mechanisms, such as synchrotron emission, physics guarantees that the spectrum at each point on the source is relatively smooth over a 15%-30% frequency range. For a synchrotron emission mechanism a single electron of energy E will radiate over at least an octave in frequency centered at a frequency which depends on E. Furthermore shock acceleration mechanisms give smooth distributions of electron energy. Because of the spectral smoothness we can often adequately characterize the spectral behavior at each point on the source by an intensity at the reference frequency plus a gradient with frequency. MFS algorithms must therefore solve for two unknowns per pixel. However, in using MFS we have also increased the number of observables by a factor of n_f, where n_f is the number of frequencies. If $n_f > 2$ the MFS image is still better constrained than the single frequency image and MFS images can be of higher fidelity than single frequency images.

16.4 Spectral Effects

16.4.1 The Spectral Expansion

Here we discuss the effects that arise if we combine all of the multi-frequency visibility data in the (u, v) plane, to produce enhanced (u, v) coverage, and then process the data conventionally. The combined (u, v) data $V_T(u, v)$ is then

$$V_T(u,v) = \sum_i S\left(u,v,\nu_i\right)\, G\left(\nu_i\right)^{-1} \overline{I}\left(u,v,\nu_i\right) . \tag{16.1}$$

Where the overline indicates a Fourier Transform (FT), so that $\overline{I}(u, v, \nu_i)$ is the FT of the intensity distribution at frequency ν_i, $I(x, y, \nu_i)$. $S(u, v.\nu_i)$ is the sampling function, the (u, v) coverage, at frequency ν_i (i.e., consisting of a δ function of area unity at a sampled (u, v) point and zero elsewhere), and $G(\nu_i)$ is an optional scaling factor at frequency ν_i which can be chosen to account for an *overall* spectral dependence $G(\nu)$ in the source. The above equation can be rewritten

$$V_T(u,v) = \sum_i S(u,v,\nu_i)\, \overline{I'}(u,v,\nu_i)\,, \tag{16.2}$$

where $I'(x, y, \nu) = G^{-1}(\nu)I(x, y, \nu)$ is the effective frequency dependent source brightness.

We can substitute a Taylor expansion for intensity in terms of its derivatives with respect to some function $f(\nu)$ of frequency. Conway, Cornwell and Wilkinson (1990) showed that the choice $f(\nu) = \log_e(\nu/\nu_o)$ produced the most compact expansion if the spectrum was power law at every point on the source and had spectral index $\alpha < 0$. Although this is a good approximation for the optically thin synchrotron emission which occurs on arcsecond scales this is not generally true on VLBI scales. On these scales the simple choice $f(\nu) = (\nu/\nu_o)$ often gives the most compact expansion which is

$$\begin{aligned} I'(x,y,\nu) = I'(x,y,\nu_o) &+ \nu_o \left.\frac{dI'(x,y,\nu)}{d\nu}\right|_{\nu=\nu_o} \left(\frac{\Delta\nu}{\nu_o}\right) \\ &+ \frac{1}{2}\nu_o^2 \left.\frac{d^2I'(x,y,\nu)}{d\nu^2}\right|_{\nu=\nu_o} \left(\frac{\Delta\nu}{\nu_o}\right)^2 + \ldots , \end{aligned} \tag{16.3}$$

where $\Delta\nu = \nu - \nu_o$ (Cornwell 1984). Then

$$\begin{aligned} V_T(u,v) = {} & \overline{I'(x,y,\nu_o)} \sum_i S(u,v,\nu_i) \\ & + \overline{\nu_o \frac{dI'(x,y,\nu)}{d\nu}}\bigg|_{\nu=\nu_o} \sum_i \left(\frac{\Delta\nu_i}{\nu_o}\right) S(u,v,\nu_i) \\ & + \overline{\frac{1}{2}\nu_o^2 \frac{d^2 I'(x,y,\nu)}{d\nu^2}}\bigg|_{\nu=\nu_o} \sum_i \left(\frac{\Delta\nu_i}{\nu_o}\right)^2 S(u,v,\nu_i) + \ldots. \end{aligned} \quad (16.4)$$

The *dirty map* is the FT of $V_T(u,v)$

$$= I' * B_o + \nu_o \frac{dI'}{d\nu} * B_1 + \frac{1}{2}{\nu_o}^2 \frac{d^2 I'}{d\nu^2} * B_2 + \ldots \quad (16.5)$$

Where * indicates convolution, where each of the I', $\nu_o dI'/d\nu$ and ${\nu_o}^2 d^2 I'/d\nu^2$ are evaluated at $\nu = \nu_o$ and are functions of position (x,y) and where the $B_n(x,y)$ are generalized dirty beams such that

$$B_n(x,y) = \overline{\sum_i S(u,v,\nu_i) \left(\frac{\Delta\nu}{\nu_o}\right)^n}. \quad (16.6)$$

16.4.2 Spectral Dirty Beams

Figure 16.2 illustrates graphically the nature of the spectral dirty beams, B_o, B_1, B_2, etc. Note that there is a significant difference between even and odd order spectral dirty beams. The former have a central peak while the later do not. This follows from the fact that the value of a beam at its center is equal to the integral over its (u,v) coverage, which for even beams is always positive but is almost equally positive and negative for odd beams.

Using 8 frequencies over a 25% range we find that the VLBA has a B_1 beam similar to those in figure 16.2 with peak sidelobes of about (1/200). The B_2 beam has a peak of order $\frac{1}{2}(\Delta\nu/\nu_o)^2$ i.e., about (1/200), with sidelobes around (1/2000). The sidelobe levels in the B_3 and B_4 beams are 100 times smaller still. Conway, Cornwell and Wilkinson (1990) show that the peak sidelobes of the B_1 and B_2 beams both scale as $(\Delta\nu/\nu)^2$. This has been approximately confirmed by numerical simulations.

16.4.3 First Order Spectral Errors

The first order spectral errors are those due to the $\nu_o dI'/d\nu * B_1$ term of the spectral expansion. Figure 16.3 illustrates what we expect for the spectra of features on the VLBI scale. Each component within the jet has a synchrotron self-absorbed spectrum which peaks at some turnover frequency but which has a power law form above and below this frequency. Below the spectral peak

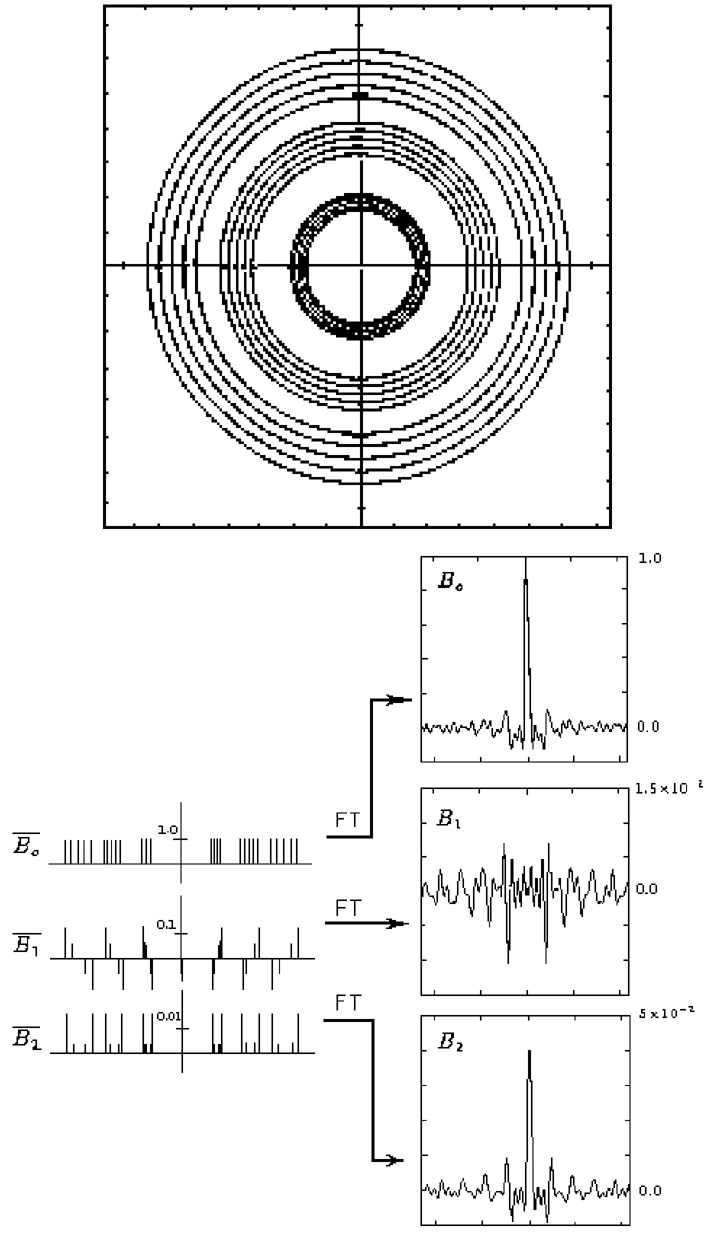

Figure 16.2: The nature of spectral dirty beams. *Top:* Simple MFS (u, v) coverage for 3 antennas and 5 frequencies. *Bottom left:* Slices across (u, v) coverage showing weightings corresponding to the different beams. *Bottom right:* Slices through the beams formed from Fourier transforming the weighted coverages.

$I = I_o(\nu/\nu_o)^{\alpha_{thick}}$ and above it $I = I_o(\nu/\nu_o)^{\alpha_{thin}}$. Generally α_{thick} is between $+1$ and $+2.5$ while α_{thin} is between -0.5 and -1.0.

Let us consider the first order spectral errors contributed by a single synchrotron component in figure 16.3. If the turnover frequency is in the observing band then $dI'(x, y, \nu)/d\nu$ is small and so the first order spectral errors are also small. Larger effects occur if we observe in the optically thin or optically thick parts of the spectrum. Using the fact that for power law spectra $\nu_o dI'(x, y, \nu)/d\nu = I'(x, y, \nu_o)\alpha(x, y)$ at $\nu = \nu_o$, we find that for VLBA MFS over a 25% frequency range the peak first order spectral errors occur at a level of $\alpha/200$ of the peak brightness in the map. In the worst case of an optically thick spectrum we expect that the uncorrected first order spectral errors to occur at about 1% of the peak. However, for such a single component source we can use the multiplicative factors $G(\nu)$ in equation 16.1 to correct the data for the spectrum of the source, make $dI'/d\nu$ and all higher derivatives zero and therefore totally remove all spectral errors. Applying such corrections naturally requires excellent relative amplitude calibration of each frequency. For the VLBA if care is taken sufficiently accurate calibration should be achievable.

This trick can only be used to completely remove the spectral effects of one component in the source. In the worst possible case of two or more equal brightness components with very different spectra (i.e., optically thick and optically thin) even after rescaling according to the mean spectral index we cannot significantly reduce the first order spectral contributions. In this case first order spectral errors will still effect the image at a level of about 1% of the peak brightness of the components. However, for the common case of a source dominated by a single component, then we can rescale the data to remove its spectral effects entirely, the spectral errors will be determined by the *second* brightest component. First order spectral errors will therefore be below (often well below) 1/100 of the brightness of the second brightest component. This level of error will often be comparable to, or less than, the thermal noise in which case all spectral errors can be ignored.

16.4.4 Second Order Spectral Errors

The second-order spectral errors are those due to the spectral expansion term:

$$\frac{1}{2}{\nu_o}^2 \frac{d^2 I'}{d\nu^2} * B_2 \,. \tag{16.7}$$

For a power-law spectrum with spectral index α this term becomes

$$\frac{1}{2} I'(x, y, \nu_o)\alpha(x, y)[\alpha(x, y) - 1] * B_2 \,. \tag{16.8}$$

Therefore there is a second order contribution even for power law spectra which are *straight* (in a Log-Log plot of flux versus frequency). However, the maximum

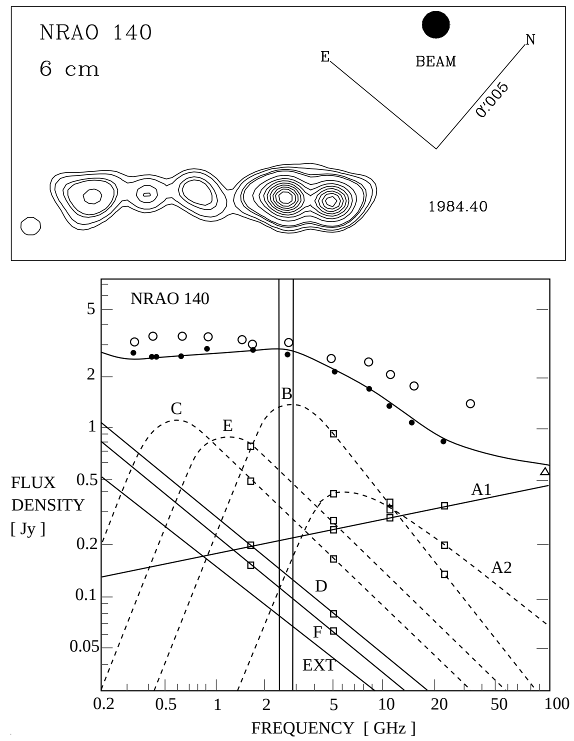

Figure 16.3: *Top:* A VLBI map of NRAO 140 (Marscher 1988) showing individual components. *Bottom:* Estimates (dotted lines) of the spectra of each component. The vertical lines delineate a 25% frequency range around 2.3GHz.

second order effects occur near the peak of the synchrotron spectrum where

$$\frac{1}{2}{\nu_o}^2 \frac{d^2 I'}{d\nu^2} \approx I' \,. \tag{16.9}$$

Again by correcting amplitudes according to the spectrum of the brightest component we can eliminate its second order spectral errors. Hence in many cases the level of the second order errors will be set by the second brightest component. Another method of reducing second order effects (see section 16.5.2) is to image each frequency channel separately and estimate and remove the effect of spectral curvature in small regions. For our standard observing parameters the VLBA second order spectral dirty beam has a central peak with value of about 1/200 and sidelobes at about 1/2000. Note that the effect of the contribution from the central peak is fairly benign, it can be thought of as creating a slight bias in intensity on the source. If we can tolerate such a bias then the sidelobe contribution from second order errors only occurs at a level below about 1/2000 of the second brightest map component, in most cases this will be well below the thermal noise and so can be neglected.

The nature of the *bias* effect introduced by the second order spectral errors can be understood if we rearrange the dirty map expansion (Conway 1991) to give

$$[I' + \frac{b_2}{2}{\nu_o}^2\frac{d^2I'}{d\nu^2} + \frac{b_4}{24}{\nu_o}^4\frac{d^4I'}{d\nu^4} + \ldots] * B_o$$
$$+ \nu_o\frac{dI'}{d\nu} * B_1 + \frac{1}{2}\nu_o^2\frac{d^2I'}{d\nu^2} * (B_2 - b_2B_o) + \ldots, \quad (16.10)$$

where $b_2 = B_2(0,0)$, $b_4 = B_4(0,0)$, etc. Therefore the dirty map can be thought of as

$$I'' * B_o + \nu_o\frac{dI'}{d\nu} * B_1 + \frac{1}{2}{\nu_o}^2\frac{d^2I'}{d\nu^2} * B_2'' + \ldots, \quad (16.11)$$

where the *biased* intensity is

$$I'' = I' + \frac{b_2}{2}{\nu_o}^2\frac{d^2I'}{d\nu^2} + \frac{b_4}{24}{\nu_o}^4\frac{d^4I'}{d\nu^4} + \ldots. \quad (16.12)$$

The modified even order spectral dirty beams, e.g.,

$$B_2'' = (B_2 - b_2B_o),\ B_4'' = (B_4 - b_4B_o),\ \ldots \quad (16.13)$$

now have no central peaks, just sidelobes (at about 1/2000 for B_2'' for 25% frequency range VLBA MFS).

16.5 The MFS Deconvolution Problem

16.5.1 Nature of The Problem

In most VLBI sources even for ≈25% *bandspread* the dirty map can be represented by the expression

$$I'' * B_o + \nu_o\frac{dI'}{d\nu} * B_1\,. \quad (16.14)$$

to at least the several thousand to one level, where I'' is the slightly biased intensity at the reference frequency as described by equation 16.12. Therefore the task of MFS algorithms is to remove first order spectral errors and estimate I''. Several algorithms exist and although research is still continuing, progress is very encouraging. We discuss MFS algorithms (in order of increasing complexity) below.

16.5.2 Map and Stack

This is the simplest approach: make individual channel maps at each frequency, from the resulting "data cube" estimate $\nu_o dI'/d\nu$ (and possibly higher derivatives) at each pixel. Having determined the derivatives at each pixel, the spectral variation is subtracted from the visibility data (using Equation 16.4), which is then Fourier inverted, CLEANed, etc.

Unfortunately, because the spectral dependence is ultimately determined from single frequency maps this algorithm feeds back the large single-frequency reconstruction errors into the data. It is possible to show that some forms of reconstruction errors, such as sinusoidal ripples in CLEAN images (i.e., *CLEAN ripples*; see chapter 3) will not be significantly reduced using this algorithm (Conway 1988). We can argue (Conway, Cornwell and Wilkinson 1990) that any algorithm which processes the individual frequency data sets piecemeal must be non-optimum since it does not exploit the non-linearity of deconvolution algorithms. However, the algorithm does have the advantage that higher order derivatives can easily be found. There may therefore be a role for this algorithm in estimating and removing the second order effects in small bright regions of a source prior to using one of the other algorithms.

16.5.3 Direct Assault

We could treat the MFS problem as one of explicit model fitting. We have unknowns I'' and $\nu_o dI'/d\nu$ at each pixel, therefore we should vary these unknowns to better fit the MFS data set i.e., to reduce the χ^2 fit to the data. Optionally we can include a 'regularizing' penalty function P, (i.e., such as the sum of the image entropies for the resulting model images at each frequency) and then minimize $\chi^2 - P$. The function P acts to bias the output image to have the properties that we a priori know or assume it to have, i.e., positivity, smoothness, etc. An efficient means of iteratively finding a global minimum in the above function is required, but this may only require minor modifications to existing MEM algorithms. This approach shows promise and should be investigated further. Although it may be "using a sledgehammer to crack a nut", with improving computing power it might be the best approach.

16.5.4 Data Weighting Methods

Cornwell (1984) discussed methods in which spectral effects are reduced or eliminated by appropriate weighting of the data in the (u, v) plane. If a single (u, v) point were sampled at two different frequencies ν_1 and ν_2 then $V(u, v, \nu_1) = \overline{I} + (\nu_1 - \nu_o)\overline{dI/d\nu}$ and $V(u, v, \nu_2) = \overline{I} + (\nu_2 - \nu_o)\overline{dI/d\nu}$ then $\overline{I} = a_1 V(u, v, \nu_1) + a_2 V(u, v, \nu_2)$ where the constants a_1 and a_2 depend only on ν_1 and ν_2, hence spectral effects can be eliminated.

In general, it is unlikely that any two (u, v) points will exactly overlap, however $\overline{I}$ and $\overline{dI/d\nu}$ will be approximately the same for two visibility points provided they are well within a distance $1/\Delta\theta$ wavelengths in the (u, v) plane where $\Delta\theta$ is the angular size of the field over which there is emission from the source. It is possible to get a large reduction in spectral effects over the field of the source if we incorporate frequency dependent weighting into the (u, v) *gridding* process. This process grids the irregularly sampled (u, v) points on a square grid of cell size of order $1/\Delta\theta$ prior to Fourier transformation.

Algorithms of this type were briefly considered by Conway (1988). The major problem is that only those grid cells with two or more (u, v) points at different frequencies can contribute to the MFS map. For large complex images for which MFS is most needed, $1/\Delta\theta$ becomes small, few cells are double sampled and so the effective MFS (u, v) coverage is relatively poor, For typical image sizes and VLBA MFS observing parameters the effective MFS (u, v) coverage can be comparable or even worse than the single frequency (u, v) coverage, removing the whole point of MFS! However, since the effective MFS (u, v) coverage is a strong function of image size this algorithm may have a role in producing very high reliability images of relatively small objects.

16.5.5 Double Deconvolution

This variant of the CLEAN algorithm (Conway, Cornwell and Wilkinson 1990) is based on the original justification of CLEAN as a pattern recognition algorithm (see Högbom 1974 , and chapter 3). The MFS dirty image is formed from the superposition of I'' and $\nu_o dI'/d\nu$ distributions each convolved with its own beam, i.e. B_o or B_1. Therefore point components in either the I'' or $\nu_o dI'/d\nu$ distributions have unique signatures in the dirty map. We can try to find such point components in I'' or $\nu_o dI'/d\nu$ by alternately convolving ('correlating') the dirty image with B_o or B_1 and recording point components (i.e., CLEAN components) at those parts of the convolved image with high brightness.

In detail the algorithm proceeds in the following way. In the first half of the first cycle we start with the MFS dirty map and attempt to search for the I'' distribution. Strictly we should convolve the dirty map with B_o, but, because for uniformly weighted data $B_o * B_o = B_o$ and $B_o * B_1 = B_1$, we can omit this step. At this stage there will be compact structure due to $I'' * B_o$ and diffuse structure due to $\nu_o dI'/d\nu * B_1$. This follows because even order spectral dirty beams have a central peak but odd order beams do not (see section 16.4.2). In this half of the algorithm we CLEAN for I'' (removing B_o and finding CLEAN components) until the fit no longer improves, giving us an estimate of the I'' distribution I''_m. After removing the effects of this model from the data the residual map is now

$$(I'' - I''_m) * B_o + \nu_o dI'/d\nu * B_1 \,. \tag{16.15}$$

In the second half of the first cycle we recognize $\nu_o dI'/d\nu$ by convolving with B_1. Using $B_n * B_m = B_{n+m}$ the residual map now becomes.

$$(I'' - I''_m) * B_1 + \nu_o dI'/d\nu * B_2 \,. \tag{16.16}$$

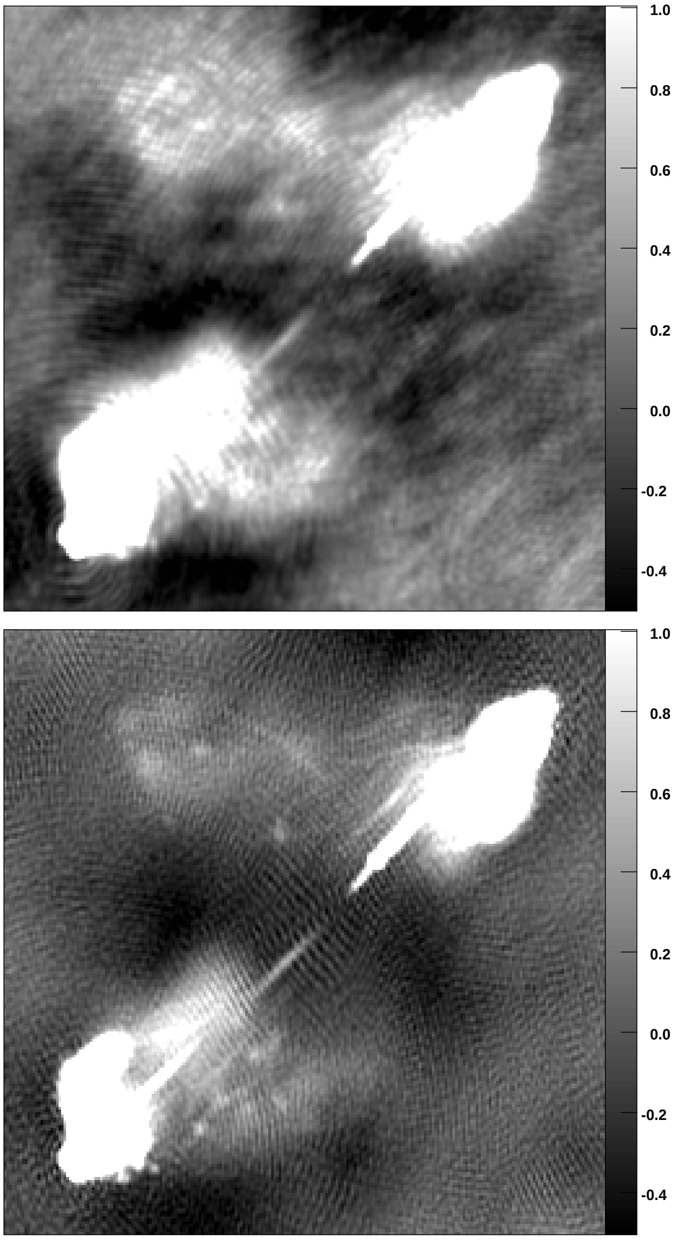

Figure 16.4: VLBA Imaging Simulations. *Top:* Grayscale of a single frequency deconvolution. *Bottom:* Using MFS data from 8 frequencies over a 25% range after two cycles of Double Deconvolution.

Now the I'' portion is convolved with a *diffuse* beam and $\nu_o dI'/d\nu$ with a centrally condensed one. We now CLEAN with B_2 and form an $\nu_o dI'/d\nu$ model. The effects of this model can now be removed from the data, the I'' distribution searched for and a new cycle begun. The process of alternately searching for the I'' and $\nu_o dI'/d\nu$ distributions can be carried on iteratively until the fit no longer improves. Note that at every stage we CLEAN an image consisting of compact structure from our desired distribution in the presence of diffuse sidelobes from the unwanted distribution. It is partly the in-built bias of CLEAN toward finding compact structure which drives convergence.

Less qualitatively, it can be shown (Conway 1988) that the algorithm always converges to fit the data, therefore in cases in which the MFS (u, v) plane is oversampled the algorithm will give a unique solution for the two distributions. More generally, it can be shown mathematically that convergence towards a correct separation is guaranteed as long as we switch half-cycles when the peak residual is larger than the largest sidelobe due to the unwanted distribution (Conway 1988). We can attempt to fulfill this condition by monitoring the rms. sidelobe level in a distant region of the image thought a priori to be free of radio emission. However, we find in practice that the exact point at which we switch half cycles is not very critical because it takes a very large number of CLEAN iterations to incorporate a lot of the emission from the diffuse unwanted distribution.

Figure 16.4 shows the result of applying the Double Deconvolution algorithm to simulated L-band VLBA data. After two full cycles the residual spectral errors are reduced below the noise, the resulting MFS map shows none of the reconstruction errors present in the conventional single frequency reconstruction.

16.5.6 The Sault Algorithm

Another CLEAN-based algorithm (the *Sault* algorithm, see Sault and Wieringa 1994) has been developed and used at the Australia Telescope (an instrument similar to the VLBA in frequency agility and lack of antennas). Although in some respects similar to Double Deconvolution, it searches for the two beam patterns simultaneously, and is so a more natural analog of normal CLEAN. The chief advantage of this algorithm over Double Deconvolution is that there is less possibility of confusion between the responses due to the B_0 and B_1 beams, as the data are fitted to both simultaneously.

For each iteration, normal CLEAN can be viewed as finding a location, j, and flux, a_0, which minimizes

$$\epsilon^2 = \sum_i (R(i+j) - a_o B_o(i))^2 , \qquad (16.17)$$

(where $R(j)$ is a vector representing the residual image). Provided uniform weighting is used, this amounts to simply finding the location and intensity of the peak residual. For the Sault algorithm the optimum location, flux and flux derivative are determined for a point source. In more detail, a location, j, and coefficients, a_0 and a_1, are found which minimize.

$$\epsilon^2 = \sum_i (R(i+j) - a_0 B_0(i) - a_1 B_1(i))^2 \,. \tag{16.18}$$

Unlike double deconvolution, there is an obvious generalization to cope with higher order spectral effects.

Although the process of finding the optimum point source is not as intuitively pleasing as locating the peak residual, some manipulation shows that the optimum occurs at the location which maximizes

$$R_0(j)^2 A_{11}(0) + R_1(j)^2 A_{00}(0) - 2R_0(j)R_1(j)A_{01}(0) \,. \tag{16.19}$$

where we define the various images

$$\begin{array}{rlrl} R_0 = & R \star B_0 & R_1 = & R \star B_1 \\ A_{00} = & B_0 \star B_0 & A_{11} = & B_1 \star B_1 \\ A_{10} = & B_1 \star B_0 & A_{01} = & B_0 \star B_1 \,. \end{array} \tag{16.20}$$

Here $\star$ denotes correlation. Similarly, the coefficients a_0 and a_1 are found to be simple functions of the A, R_0 and R_1 images at location j.

Whereas the A images are auto- and cross-correlations of the beams (and hence constant), R_0 and R_1 are images which are correlations of the residuals with the beams. Rather than perform this correlation process each iteration, R_0 and R_1 can be computed once at the start of the CLEAN, and then updated at each iteration, by subtracting off the appropriate weightings of the A images. Indeed, the residual image never needs to be stored—the algorithm CLEANs the residuals correlated with beams with beams correlated with beams.

This algorithm is a factor of a few slower than conventional CLEAN. As spectral effects will be unimportant during the early stages of CLEANing (particularly if the visibilities have been compensated for the spectral index of the dominant component), a speed advantage can be achieved by starting with a normal CLEAN, and then switching to the Sault algorithm.

16.6 Multi-Frequency Self-Calibration

If we self-calibrate each frequency independently, image quality is ultimately restricted by the self-calibration step (Conway, Cornwell and Wilkinson 1990). Some improvement results if the model image for self-calibration is determined by using the I'' and $\nu_o dI'/d\nu$ images from the MFS imaging step to generate estimates of image intensity at each frequency. These different frequency models can then be used to improve the quality of the self-calibration solutions and so on iteratively.

We can improve on this scheme—the antenna based errors are not independent at each frequency. In effect there are only three unknowns per antenna, a frequency constant phase offset (due to the independent electronics, etc.) a *delay-like* term $\phi \propto \nu$ (electronics and troposphere) and (at low ν) an ionospheric delay with $\phi \propto \nu^{-1}$. Therefore over each solution interval there are only two or three unknowns to solve for instead of n_f. If we make use of the fact that the antenna-based phase errors are linked at the different frequencies we can show that the ratio of observables to unknowns can be greater for MFS than for a conventional array of equivalent (u, v) coverage (Conway 1991). In order to do multi-frequency self-calibration successfully we must be able to observe at all the frequencies simultaneously or be able to rapidly switch frequency (on a timescale less than that for atmospheric/ionospheric phase fluctuations). Some Australia Telescope software (solving for a delay term only) for MF self-calibration has been developed. This software still needs to be adapted for the VLBI case, since on these much longer baselines antenna based delays can be many wavelengths.

16.7 Practical MFS

The Australia Telescope currently uses MFS routinely. Bandspreads of up to 25% are used. For the higher dynamic range cases, the Sault algorithm has been used to eliminate the spectral effects (Morganti, Killeen and Tadhunter 1993; Simkin, Sadler and Sault 1993). Until now the use of MFS for VLBI has been limited by the restricted capabilities of VLBI arrays. Pioneering experiments include (1) An 18 cm MERLIN 6-station, 5-frequency 7% range observation, of 3C 179 (Conway 1988; Conway 1991); (2) a 2 frequency 5% Global VLBI observation 18cm of 3C 84 (Biretta, Bartel and Deng 1991) which gave a factor of 2 improvement in fidelity; (3) a 3 frequency 8% Global VLBI 1.3cm of 3C 84. (Conway 1991). The restricted frequency ranges available in these experiments have been insufficient to cross most of the gaps in the single frequency (u, v) coverage. Simulations show that only when this happens can we expect a really significant improvement in image fidelity. Despite this these experiments have demonstrated some definite improvements in fidelity (see the MFS map of 3C 179 in Conway 1991).

The advent of the VLBA (and improvements to the EVN) will revolutionize the potential of MFS. Great attention is being paid to accurate amplitude calibration. Furthermore the VLBA hardware has been designed with MFS in mind. Feeds, receivers, etc., are wide bandwidth (generally 15% to 30%); enough to fill most (u, v) holes (see table 16.1). For a single local oscillator (LO) setting it can observe at 8 frequencies simultaneously, set anywhere over 500 MHz (or optionally 1000 MHz, single polarization at X-Band). Furthermore the LO can be switched (taking only a few seconds to settle) allowing > 8 frequencies, and coverage of the whole frequency range covered by feed and receivers. This wide-band frequency agile system should, for the first time, allow multi-frequency self-calibration techniques to be employed. The second column in table 16.1 gives

Table 16.1: Frequency coverage of the VLBA

Wavelength band (cm)	Total ν range %	Instantaneous ν range %
90	10	10
50	5	5
20	35	28
13	9	9
6	13	10
4	12	12
2	25	3
1	15	2
0.7	3	1

the total fractional frequency range over which the noise $< 1.4.$ of the minimum in the band (Biretta 1991). In some cases, by allowing more degradation in noise (to a factor of 2), the total frequency range can be increased by an extra 5%. The third column gives the *instantaneous* range reachable without switching the LO frequency. In contrast to the VLBA most European antennas are restricted to fractional frequency ranges of $< 10\%$. However, because of the larger number of antennas the fractional holes in the EVN+VLBA (u, v) coverage are usually $\ll 10\%$. Hence MFS will have an important role for global as well as VLBA only observations.

16.8 Conclusions

Multi-frequency synthesis (MFS) provides a way of dramatically improving (u, v) coverage and hence image fidelity for VLBI observations. The effects of the spectral dependence of structure have been analyzed and we find that in many cases we expect them to be less than the thermal noise. In other cases there are several promising algorithms for solving for and removing spectral effects. The further development of these algorithms, improvements in multi-frequency capabilities of VLBI arrays and in amplitude calibration will allow MFS to become a standard observing technique in the years to come.

References

Biretta, J. A. 1991. "Frequency dependence of VLBA system temperatures and gains". VLBA Test Memorandum 27, National Radio Astronomy Observatory.

Biretta, J. A., Bartel, N., & Deng, J. 1991. "Dual-frequency synthesis VLBI observations of 3C84 at 18 cm wavelength". See Cornwell and Perley (1991), p. 334.

Conway, J. E. 1988. *Multifrequency Synthesis With MERLIN, a New Technique in Aperture Synthesis Imaging*. Ph. D. thesis, University of Manchester.

Conway, J. E. 1991. "Multi-frequency synthesis". See Cornwell and Perley (1991), p. 171.

Conway, J. E., Cornwell, T. J., & Wilkinson, P. N. 1990. "Multi-frequency synthesis: a new technique in radio interferometric imaging". *Monthly Notices Roy. Astron. Soc.* **246**, 490.

Cornwell, T. J. 1984. "Broadband mapping". VLB Array Memorandum 324, National Radio Astronomy Observatory.

Cornwell, T. J., & Perley, R. A. (Eds.) 1991. *Radio Interferometry: Theory, Techniques, and Applications*, Volume 19 of *ASP Conference Series*. San Francisco: Astron. Soc. of the Pacific.

Högbom, J. 1974. "Aperture synthesis with a non-regular distribution of interferometer baselines". *Astron. Astrophys. Suppl.* **15**, 417.

Marscher, A. P. 1988. "Contemporaneous X-Ray and VLBI observations of the quasar NRAO 140". *Astrophys. J.* **334**, 552.

McCready, L. L., Pawsey, J. L., & Payne-Scott, R. 1947. "Solar radiation at radio frequencies and its relation to sunspots". *Proc. Roy. Soc. London* **A-190**, 357.

Morganti, R., Killeen, N. E. B., & Tadhunter, C. N. 1993. "The radio structures of southern 2-Jy radio sources". *Monthly Notices Roy. Astron. Soc.* **263**, 1023.

Sault, R. J., & Wieringa, M. H. 1994. "Multi-frequency synthesis techniques in radio interferometric imaging". *Astron. Astrophys. Suppl.* **108**, 585.

Simkin, S. M., Sadler, E. M., & Sault, R. 1993. "The structure and kinematics of Pictor A". In G. V. Bicknell, M. A. Dopita, and P. J. Quinn (Eds.), *Proceedings of the First Stromlo Symp.—The Physics of Active Galaxies*. Astronomical Society of The Pacific, p. 347.