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 13

Practical VLBI Imaging

R. C. WALKER
NATIONAL RADIO ASTRONOMY OBSERVATORY

Abstract Practical aspects of VLBI imaging are discussed, with emphasis on the choices of parameters. Particular attention is paid to the interaction between the parameter choices and both the image appearance and the fit of the model (image) to the data. The process is illustrated using an image made from 13 cm data from a VLBA geodesy experiment.

Contents

13.1 Introduction

An ideal interferometer would deliver perfectly calibrated amplitudes and phases on a regularly sampled, rectangular grid. An image would then be made by simply applying a Fourier transform. Such an ideal cannot be achieved in practice. For real interferometers, the Fourier or (u, v) plane is incompletely and irregularly sampled. This requires the use of some sort of interpolation procedure, usually in the form of a deconvolution in the image plane. The amplitude calibration is not perfect because of weather, pointing, imperfectly measured gains, and system variations. The phases have errors because of the troposphere, ionosphere, imperfect clocks/frequency standards, and imperfect knowledge of the array geometry and of phase shifts in the electronics. For most radio interferometers, the amplitude calibration is adequate for moderate dynamic range imaging. For linked interferometers, the phase calibration is also adequate for imaging most of the time. Self-calibration is only needed in order to reach high dynamic ranges. For VLBI this is not usually the case. The phase errors are so severe that, without self-calibration, they are essentially random and no meaningful image can be made. (See chapter 17 for exceptions.)

In the early days of VLBI, it was thought hopeless to extract any meaningful information from the phases. Model fitting to the correlated amplitudes was the main "imaging" method. Eventually it was realized that even uncalibrated phases contained powerful constraints on source structure if one was careful to maintain the antenna dependence of any instrumental phase offsets—not a trivial proposition at the time. For example, it really matters that the earth rotates a bit between the times of arrival of the wavefront at different stations and this was not accounted for in early correlators. Once the effort was expended to make the calibration offsets antenna dependent, the phases began to be used. At first this was done by including the closure phases in the model fitting algorithms. In 1978 a major breakthrough was made with the invention of "hybrid mapping" (Readhead and Wilkinson 1978; Cotton 1979). These procedures, with refinements to include amplitudes and to deal with station gains explicitly, are still used for most VLBI imaging applications.

The basic idea of hybrid mapping is to solve simultaneously for both the image of the target radio source and for the antenna calibration parameters. This is done by alternately making a deconvolved image using assumed calibration parameters and then using that image to improve the calibration parameters. In the good cases, the procedure converges in a few iterations. In the most troublesome cases, usually involving poor (u, v) coverage and calibration, a large number of iterations must be tried, many leading to dead ends, before a satisfactory image is obtained. Hopefully the latter cases will be less common in the era of the VLBA.

There are a number of ways that an observer can influence the progress of the hybrid mapping procedure by adjusting the parameters controlling both the self-calibration and deconvolution steps. The problem is to guide the process to a good solution while avoiding situations analogous to finding a local, but not global, minimum in a least-squares fitting procedure. This chapter covers the various ways in which the observer can control hybrid mapping. It also covers how to recognize the difference between a finished image and one that still must, or at least can, be improved.

Hybrid mapping is not fundamentally different from using self-calibration on

VLA data. In fact, for some of the highest dynamic range VLA images, no attempt was made to do absolute phase calibration and of order 10 iterations of hybrid mapping were used following exactly the VLBI procedure. The "details" that make it seem so different are the quality of the starting image and the great power of the closure constraints on the VLA with so many antennas. For the VLA, the starting model is probably an image with a dynamic range of a thousand or more that already looks a lot like the final source. For VLBI, the starting model is usually just a point source or a simple model consisting of a few gaussians derived from a model fit.

13.2 Editing and Calibration

Prior to beginning the hybrid mapping process, the data need to be calibrated, fringe fitted, and edited. All of these steps have been discussed in earlier chapters in this volume. Actually, an image could be produced without doing any of these steps, but it would take longer, require a lot more computer resources, and probably would not be as good. The amplitude calibration insures that the flux scale is correct. It also greatly improves the early iterations of the hybrid procedure where the self-calibration step is done on phases only. In cases of a very small number of baselines or a complex source, good amplitude calibration, especially if it is good enough that no amplitude self-calibration is needed, may make the difference between whether or not a reasonable image is produced. The amplitude calibration should be done with the best available monitor data from the stations and should be checked using a few scans on very compact sources and using baseline crossing points in the (u, v) plane.

The purpose of the fringe fitting step for an imaging experiment is to allow self-calibration to be done over longer time intervals and greater bandwidths for greater sensitivity. It also allows the data to be averaged in time and frequency to reduce the data set size to manageable proportions (see chapter 10). The smaller the data set, the faster the hybrid loops will go and the faster the image can be made. Also, the more the data can be averaged, the higher the SNR on each point will be and the easier it will be to compare data and models. Fringe fitting is just a variant on self-calibration in which the first order rates of change of phase in time and frequency are determined. It is not fundamentally different from what is done at later steps of the mapping procedure.

The fringe rates need only be good enough to allow averaging and, since the rates introduced by source structure will usually be small, a solution based on a point source or other simple model is usually adequate. In fact, with the high quality geometric model used on the VLBA correlator, fringe fitting is not even needed in some cases. However, just like self-calibration, when it is needed, fringe fitting is improved by having a good starting model. Therefore, in cases where the initial fringe fitting failed for many intervals or where there are large rates introduced by large scale source structure, it may be profitable to redo it after a good source model is obtained.

Prior to averaging, the data set will consist of anywhere from 4 to 1024 spectral points in each of between 1 and 16 baseband channels (see chapter 7). Each data point will be an average of about 1 to 10 seconds. This can be a very large dataset. After fringe fitting, the amount of averaging that can be done is set by the size of the source, the coherence time, and any concerns about different baseband channels needing separate self-calibration. The impact of source structure is discussed below. Any time averaging should be over a time short compared to any phase fluctuations not due to noise that are left in the data after the fringe fit. It should be for a time short compared to the coherence time. Thirty seconds is a good value for many cases of decent weather and intermediate frequencies. It should nearly always be possible to average all spectral points in a baseband channel. In most cases, it will also be possible to average across baseband channels. However if there is any reason to suspect that the different channels might be behaving differently (they do go through different electronics while still in analog form) it might be best to keep them separate. There is no need to go to extremes to average enough to detect the source in each data point. The self-calibration programs can do solutions on data that have been averaged in time and across channels dividing out a model.

Source structure can limit reasonable averaging because large sources or sources with well separated compact components will have fine scale structure in the (u, v) plane. Excessive averaging will smear out such structure. The finest scale will be approximately the inverse of the maximum separation of features in the source. Spectral points or baseband channels at the same time on the same baseline but at different frequencies constitute measurements at different (u, v) points because the (u, v) values are set by the projected separation divided by the wavelength. The observer should calculate roughly the finest scale of fluctuations expected in the (u, v) plane due to source structure and then make sure that any frequency averaging is over a significantly smaller scale. Likewise, data points taken at different times are measurements at different points in the (u, v) plane because of the rotation of the earth. Again, the observer should make sure that any desired time averaging will not smear out real fluctuations in the (u, v) plane due to source structure.

Editing is another important step. In an ideal world, the on-line systems would know about any bad data and automatic flagging systems would deal with it. Both the VLBA and the VLA come close to this goal. However, the a priori editing on these instruments is not perfect and, for many stations in global experiments, there may be little if any editing information. Therefore some editing must be done by the user. Editing programs exist in various packages that allow some form of "point-and-click" editing on plots of data as a function of time. Most bad data show up as low points where they should not be. The most common are low points or non-detections at the beginning of a scan while an antenna was slewing to the source or while system and/or antenna temperatures were being measured. Bad points will usually be antenna dependent—all baselines to an antenna will show the same problem. This means that station based editing is desirable, although current software does not encourage it.

Two other steps of calibration can help although they have not been done often in the past. The first is bandpass calibration. This requires either autocorrelation spectra or a strong source on which a solution can be made for the complex antenna dependent bandpasses. The latter is preferred since it gets the phases as well as the amplitudes. It is expected that bandpass calibration will be standard procedure for the VLBA. A related step that can help with band shapes is to omit the edge channels from any frequency averaging. Usually the worst mismatches between filters are at the band edges and such data can be avoided entirely. The other calibration step that can be done is closure calibration. The usual method is to use a strong, well modeled source as a standard. Time independent deviations of each baseline from the source model (image) are treated as instrumental in origin and removed from the program source data. Sometimes this is done using data from the program source itself. This works best for large data sets with many baseline crossing points. However the failure mode is insidious—it will produce too "good" a map. Since the baselines are being forced to match the model, there may be too few constraints left and a near perfect image of the input model will be produced. This is certainly what would happen if this method were used on a snapshot observation consisting of one observation on each baseline.

13.3 Controlling the Hybrid Mapping

As mentioned above, hybrid mapping is an iterative procedure that alternates between self-calibration and image deconvolution. Prior to self-calibration, the input model, which usually consists of the CLEAN components from the previous deconvolution, can be edited. After deconvolution, the image can be examined, as can the match between the output model (CLEAN components) and the self-calibrated data, in order to gauge progress. There are a number of ways that the observer can influence the model, the self-calibration, and the imaging in order to guide the procedure. In difficult cases, adept use of these controls is the key to successful imaging. In easy cases, such as simple sources with lots of data, it probably doesn't matter much what is done. In this section, the more important controls are discussed.

One of the most important tasks of the observer while hybrid mapping is to judge progress. Is the output of the last iteration any improvement over the input? If so, it may be best to just do the same thing again to obtain further incremental improvements. However, if there is no change, or if the changes that occur do not constitute an improvement, the observer needs to decide if the image is finished or if some control parameters should be changed to try to provoke further improvements. There are several criteria that can be used to judge the quality of an image. It may not be possible to obtain an image that fully satisfies all criteria. Some judgment needs to be used to decide when the criteria are sufficiently well satisfied to deal with the astronomical questions at hand or are as good as they are ever likely to get.

The appearance of the image itself is the first indicator of quality. If the image matches one's preconceptions of what it should look like, it is likely to be fairly good. If not, the unexpected features need to be analyzed. Some types of feature are clearly signs of trouble, e.g., negative holes (except in absorption experiments or in cases of poorly sampled short spacings). Multiple features going off in various directions from the brightest points are likely to be the result of calibration problems. Excessive symmetry may indicate that the amplitudes are not yet calibrated. Any streaks, bands or other features well away from the source, especially ones that look like patterns in the dirty beam, are cause for concern.

The off-source noise level is a useful quality measure, especially in comparing iterations and checking for progress. If comparing iterations, it is important to measure the rms noise level over the same region since the it can vary over the image. It is useful to know the expected sensitivity of the image. If the off-source rms is close to the expected image noise level, it is unlikely that much improvement can be had by further iterations—the noise limit has been reached. If there is a big difference, then it is likely that progress can still be made. Note that it is possible to have an image that fits most but not quite all of the data well. In such cases, there is very little leverage driving the hybrid process toward a better solution. Only small progress may be made on each iteration and there is a temptation to quit. However if each iteration is making progress, the results can be significant when accumulated. A 10% improvement in rms may seem small, but do that seven times and the total improvement is factor of two! Unless faced with limited computer resources, there is no need to be shy about doing lots of iterations as long as progress is being made.

The match of the self-calibrated data to the model (usually the CLEAN components) is an important quality criteria and should be checked, at least for the "final" image. Any significant deviations are a clear sign of trouble. A deviation is significant if a number of data points in a row (maybe 10 to pick a totally arbitrary number) deviate from the model by more than the scatter in the points. Sometimes spurious features will only cause a mismatch for an hour or two on one or two baselines and still be quite a serious problem for the accuracy of the image. Of course, such deviations could represent some sort of problem with the data rather than an error in the source structure. This possibility should be investigated and the data edited if necessary. However be careful that the data are not just being forced to match a bad model when the deviant points are removed.

The remainder of this section describes the significant controls that the user has over the hybrid mapping procedure. These controls need to be set at the beginning and need to be considered when it seems necessary to make changes in order to make progress.

13.3.1 Self-Calibration Solution Mode

The hybrid mapping procedure starts with data that have amplitudes calibrated using known antenna gains, and that have phases that are either totally uncalibrated or that have had the equivalent of a point-source self-calibration applied by fringe fitting. At this stage, the amplitude calibration is better than anything self-calibration with a poor model can manage so phase-only self-calibration should be done. After a few iterations, the image will start to be limited by the quality of the amplitude calibration and it is time to switch to solving for both phase and amplitude gains. The decision of when to do this should be based on how well the current model (CLEAN components) fits the data and on some perception of the quality of the calibration. It may be useful to do some iterations first in which only one amplitude factor per antenna is found if that option is available as it is in the Caltech programs (Pearson 1991; Shepherd, Pearson and Taylor 1994). Later the amplitude gains can be allowed to vary as fast as the phases.

13.3.2 Windowing

Radioastronomical imaging with incompletely sampled (u, v) data depends on the fact that the sky is mostly dark. The sources being imaged cover a finite area. With uncalibrated phases in the early stages of hybrid mapping, it helps to guide the process by only accepting emission from a finite window. If the CLEAN algorithm is used for the deconvolution, which is the only case that will be discussed here, this can be done in two ways. First, the cleaning program can be told to only accept points within some window or windows. Second, the CLEAN components can be edited before being passed to the self calibration program in the next loop. Both methods can be used. Each has certain advantages.

For a snapshot observation or an observation with a small number of baselines, CLEAN tends to produce a very low noise image with scattered bright spots if run with a wide open window. It is using those scattered points to fit the data well, but they have little to do with reality. Therefore, in these cases, it is important to use as small a window in CLEAN as possible. The noise level in the image will be higher, but the image is much more likely to look like the real source. If there are lots of baselines and long tracks, leaving the CLEAN window wide open does not generally cause problems. Using a very tight CLEAN window while the phase calibration is poor can have undesirable effects. CLEAN can try to match the data by putting clearly spurious features, often of negative amplitude, within the window. In such cases, it is usually more effective to leave the window within CLEAN open and allow the program to put the spurious features wherever it wants, usually fairly far from the real source. Then the CLEAN components can be edited to remove the ones outside the expected region of the source before the next self-calibration.

Windowing can be used to try getting rid of what are thought to be spurious features such as counterjets in core/jet sources. When starting a hybrid mapping sequence with a point source model, the initial image will almost always be much more symmetric than the actual source because the phases have been forced as close to zero as the closure constraints will allow. This means that core/jet type structures will usually show counterjets. Such a counterjet would be very important scientifically if real because it would allow significant constraints to be placed on beaming models. It is important to get rid of such features if they are not required by the data. One way is to use windows that are tucked very tightly against the assumed core feature so that the counterjet is outside the window. My experience is that the post-CLEAN editing of components is much more effective at causing spurious counterjets to go away than is windowing in CLEAN itself. When windowing in CLEAN, the algorithm puts spurious components at various places in the allowed window to try to match the data that, with its calibration errors, is trying to produce the counterjet. To the extent that it succeeds, those calibration problems are left after the next self-calibration. When the windowing is done after CLEAN, no extra spurious features are generated and the self-calibration does not use the presumed spurious counterjet in the next iteration.

All of the above may cause concern about affecting the science that is produced based on preconceptions of what the source should look like. There is some cause for concern, but not as much as it might seem. If, say, a counterjet is real, it will prove very difficult if not impossible to make it go away and still produce a good image that has low noise and lacks obvious sideband patterns. I have witnessed cases where someone was using too tight a CLEAN window for a source and was having a tremendous amount of trouble making a decent image. When the window was opened up to include regions of real emission outside the earlier window, the process converged quickly. In fact, this behavior can be used as a test of the reality of components. If they can be made to go away with windowing, their reality must be suspect. If not, they are probably real.

The windows in use can be varied as the hybrid mapping proceeds. It is probably best to give the self-calibration only features that are real, even if they are not the only real features and even if they are not very well characterized. If spurious features are used, they have a tendency to keep coming back. Therefore, in early hybrid iterations, small windows may be required even if, at low levels, the source covers a large area.

13.3.3 Weights

Associated with each data point is a weight. From this weight, it should be possible, if all calibration steps have been done properly, to determine the error of that measurement. The weight typically does not reflect any calibration errors. The weight is used in the self-calibration step in the same way that the measurement error is typically used in a least-square fit. In the imaging step, it is used in the gridding process to insure that high quality points are given more influence than low quality points.

For VLBI, there is often a very wide range of antenna sensitivities because of the wide range of antenna sizes and of receiver quality. This leads to a wide range of weights for individual points in a data set. The default schemes for using weights are such that the best SNR images would be obtained. This means for VLBI that the weaker baselines, which may be contributing interesting constraints by sampling portions of the (u, v) plane not covered by strong baselines, may hardly be used at all. Also if there is an extremely sensitive antenna such as Arecibo, the weights on the baselines to that antenna will be so high that the effective (u, v) coverage is just that of the baselines to that antenna, even if there are many other baselines. This can lead to very high sidelobes. It is unlikely that any one antenna, no matter how sensitive, will have much better calibration than the others, so such dominant weights are probably not justified. For this reason, it is probably a good idea, if there are one or two very dominant antennas, to adjust the weights on the baselines to those antennas to be only a bit better than the weights on the best other baselines.

In the hybrid mapping procedure it would be best to use weights that reflect both the antenna sensitivities and the current calibration quality. In the early iterations, while the calibration is very poor, this means that the optimum scheme would be to have nearly equal weights on all baselines. This takes best advantage of the (u, v) coverage and allows even weak baselines to help the process converge. As the calibration improves, using equal weights would cause the noise limit to be determined by the weak baselines and the ultimate quality of the image could be quite limited. Therefore, more sensitivity related weights should be used, although no single antenna should be allowed to dominate completely. The imaging programs in both AIPS (MX—see help file for information on UVWTFN) and the Caltech program DIFMAP (Shepherd, Pearson and Taylor 1994) have the ability to give the weights varying degrees of importance.

13.3.4 (u, v) Ranges and Tapers

In the self-calibration procedure, it is possible to restrict the solution to use only baselines in a specified limited (u, v) range or at least to downweight baselines outside of the range. In the imaging step, a (u, v) range can also be used, as can a taper that causes longer baselines to be downweighted. These (u, v) ranges and taper can be helpful in guiding the hybrid procedure.

A case where a limited (u, v) range is useful is when the source contains an unresolved core plus more extended features. If the longer baselines are only detecting the core, they will be described well by a point source model. If the self-calibration is then restricted to the longer spacings only, the resultant antenna gains will be quite good and an image made based on them is likely to also be good—the short baselines will have been well calibrated in the process. Conversely, a compact source may look more like a point source on short baselines but be well resolved on long spacings. In this case, it is best to restrict the self-calibration to the shorter baselines. In most cases, the model (image) will be most prone to errors on the shortest and longest (u, v) scales sampled by the data. It is often best to leave the very shortest and longest baselines out of the self-calibration at essentially all stages of hybrid mapping. If the hybrid procedure has clearly not converged but further iterations are not making any progress, one of the main factors to vary is the (u, v) range in the self-calibration.

The imaging step of the hybrid procedure can also use a (u, v) range. This is less commonly used than the (u, v) range in self-calibration, but might be useful in the early hybrid iterations if the calibration quality is much better on some spacings than others. A (u, v) range can be used to make an image that is more accurate than might otherwise be obtained that can help the next self-calibration. The imaging step also can use a taper in the (u, v) plane. This downweights the longer spacings. For VLA imaging, this is usually used to increase the beam size and to increase the sensitivity to large scale structures. It can have the same use in VLBI, but it can also be used when the long baseline calibrations are poor to get a good low-resolution image that can then be used in self-calibration with a (u, v) range.

13.3.5 Natural vs. Uniform Weighting

The choice of gridding scheme in the imaging step can have some influence on convergence. The two main schemes are natural weighting and uniform weighting. With natural weighting, all points are convolved onto the grid without any attempt to adjust for the local density of points. With uniform weighting, points close together in the (u, v) plane are downweighted. Uniform weighting is an attempt to produce a gridded (u, v) plane that is a close approximation to the actual transform of the image. With natural weighting, cells at (u, v) track crossing points and cells near the center of the grid will have higher amplitude than others. Uniform weighting gives a cleaner and smaller dirty beam (lower sidelobes and more resolution) while natural weighting is optimal for signal-to-noise. Uniform weighting may be the best choice while the calibration is poor in the early hybrid iterations. The best choice later in the process depends on the specific case. Changing weighting scheme is one of the things that can be done if the hybrid procedure is clearly not done but is not making progress. Note that, when using uniform weighting, especially on large images, it may be advisable to use super-uniform weighting in which points more than one cell apart are treated as "close together".

Since this lecture was given, significant advances have been made in the area of weighting. The thesis by Briggs (1995) is highly recommended reading for all serious practitioners of interferometric imaging.

13.3.6 Cell Size and Image Size

In the imaging step of hybrid mapping, the parameters of the image must be chosen. The cell size (the angular interval between adjacent points in the digital image) should be such that there are between 3 and 6 cells across the full width, half maximum (FWHM) of the dirty beam. Larger cells allow larger images to be made, but make it more difficult to represent arbitrary brightness distributions with the discrete points of the digital image. With only three points per beam, the dynamic range that can be achieved is only a few thousand unless the image is dominated by one or more point sources which all are exactly on grid points. With six points per beam, the dynamic range will be limited by other effects. Intermediate numbers of points give intermediate dynamic range limits. For VLBI, four or five points per beam should be ample for most cases.

The overall extent of the image must also be chosen. The image must be large enough to include all detected emission from the source. Otherwise there will be a great deal of difficulty obtaining a decent image. In fact, it will be essentially impossible. It is also wise to have some region around the source and still in the image that is blank to help avoid edge effects, e.g., sidelobe aliasing. On the other hand, larger images take more computing and slow down the imaging procedure.

13.3.7 CLEAN Components

A parameter that can be varied is the number of CLEAN components to either calculate in the imaging step or to keep later. The optimum choice is still a matter of opinion. I typically CLEAN quite deeply (1000 to 5000 components) and then edit the results later. This has the advantage that it doesn't leave a plateau of unCLEANed emission in the residual image (after subtraction of the CLEAN components) in the region of the source. This plateau, in the case of a shallow CLEAN (small number of components) can contain a significant amount of the flux density of the reliable features and it does not contribute to the next self-calibration. Deep CLEANs do, however, pick up lots of noise points which contribute errors to the self-calibration. I try to minimize this by adding all components in each cell, throwing out the negatives, and then throwing out all outside some fairly tight window. Others prefer to stop the CLEAN much earlier. It is probably not wise to stop the CLEAN on the first negative, an option provided in many CLEAN programs. Negatives can indicate that CLEAN is now working on sidelobes or noise, but it can also indicate that CLEAN is trying to fix a mistake among the earlier components. Also, if the number of cells per beam is small and there is a unresolved source that is not on a grid point, negatives in adjacent cells are necessary to produce an image in which the source appears unresolved. In other words, not all negative components indicate errors.

A deep CLEAN without any component editing is probably not wise. The CLEAN components will include most of the major sidelobes and noise spikes so they, used as a model, will probably do a fairly good job reproducing the data even though the map is not very good. Then the next self-calibration step will not make much progress.

13.3.8 Self-Calibration Parameters

There are a number of other parameters for the self-calibration that need to be considered. The solution interval needs to be chosen. Longer solution intervals give higher SNR but the gains, especially the phases, can change during a long interval leading to less than optimal results. This is not a problem for the early iterations, as long as the coherence time is not exceeded. However for the later iterations of hybrid mapping, it is probably good to use as short an interval as will obtain good SNR (say above 10 or so, to pick another arbitrary number). In any case, the self-calibration solution interval should be less than or equal to the interval used in fringe fitting.

The choice of reference antenna should not matter much, but if one is chosen that is not there all the time, the programs have to try to smooth out the possible jumps at the times when another antenna has to be used as the reference. It is best to choose an antenna that is there as much of the time as possible. Also, the derived gains will be steadier if one of the more sensitive antennas is chosen.

The self-calibration program may provide options for the solution type. CALIB in AIPS offers least-squares fits and an L1 solution where the absolute values of the deviations, rather than the squares of the deviations, are minimized. The least squares fits would give the best answers if the data are well behaved with gaussian noise. If there is a chance of a few significantly deviant points, the L1 solution is often better since it is less sensitive to such points. The ability to obtain a solution at all may depend on the solution type, especially if there are only a small number of antennas and the model is fairly poor. CALIB, in particular, is subject to this effect, and the prefered solution type depends on whether it is a phase-only or an amplitude-and-phase solution.

If there are too few antennas at some time during the observations, there may be too few constraints to derive a meaningful self calibration solution. In such cases, it is best to not use that time. The minimum allowed number of antennas is one of the possible input parameters for self-calibration. The absolute minimum number that has any information on source structure, after solving for antenna gains, is 3 for phases (one closure phase) and 4 for amplitudes (one closure amplitude). For larger experiments, it may be best to specify 5 or 6 as the minimum number, although 4 is commonly chosen. Note the implication of this constraint for scheduling. If you schedule such that only 2 or 3 antennas are on source for a while, you may not be able to use the data for anything except amplitude model fitting. Also note that it is a bad idea to keep the data in an uncalibrated state if no solution is found. Some older programs would do this. This just adds noise to the final image.

The self-calibration program will make a decision about whether a good solution has been obtained based on the SNR. If a solution has not been obtained, then the data will be flagged. A minimum SNR can be chosen. The usual value used is 5. If the source is weak and a desperate effort is being made to detect it in the noise, a lower limit might be used. For a strong source for which a high dynamic range image is desired, a higher limit might be chosen to help weed out poor data.

There is a choice of input data set for self-calibration at each hybrid iteration. In all cases, the data should have been fringe fitted and averaged. The input data set could have had nothing further done to it, or it could be the output of one or more steps of self-calibration. The self-calibration may work slightly better if the gain changes are small which favors using a previously self-calibrated data set, at least late in the hybrid mapping sequence. It is possible to always use the output of the previous iteration. However there is concern that the flux and position scales could drift slowly through many iterations since they are not really constrained by anything other than the program's attempts to renormalize after obtaining a solution. Also, roundoff errors might accumulate. Attitudes vary on this one, but I would recommend using the fringe fitted output for the first few iterations, and then choose the data from one iteration as the input from then on.

13.3.9 Imaging Parameters

There are also some parameters that need to be set for the imaging step.

If the source is one sided, as is typically the case in VLBI, it is often possible to use smaller images, and hence less computer time, by shifting the image phase center. If this is done, it is a good idea to make the shift by an integral number of cells. As discussed earlier, CLEAN has a much easier time when any bright points are on grid cells.

Sometimes it helps to give a zero spacing flux density. The average flux density over the whole image will equal what is in the center cell of the gridded (u, v) plane. Giving a zero spacing flux density can help remove any offsets in the image. It may also improve the imaging results, especially if there are large scale structures and there is a strong gradient in the (u, v) data as a function of (u, v) distance on the short spacings. The zero spacing flux density can be obtained from single dish measurements, measurements with a shorter baseline interferometer, or from extrapolation of the VLBI data.

The central peak of the synthesized beam for VLBI is often far from a simple gaussian. The beam fitting algorithm in the imaging program can give rather deceptive results. It is wise to make a plot of the dirty beam and measure the size of the 50% contour to see if it is anything like what the program is using. It may be necessary to force the beam size used in reconvolving the CLEAN components onto the image. Note that in such cases, the residuals left after the CLEAN have a rather different response function than the convolved CLEAN components. It may be best to do a very deep CLEAN in such cases so that essentially everything in the image including the noise shares the same response.

If the CLEAN program has an outer loop in which the components are subtracted from the (u, v) data (for example, using MX in AIPS), it will often allow a choice between an FFT subtraction and a direct subtraction of the CLEAN components from the data. Various tests have shown very little difference between quality of the results with the two and a very large difference in computation time. Use the FFT subtraction in most cases.

13.3.10 Antenna Subsets

If the best image that can be obtained still does not seem as good as it should be, there is something that can be tried. Sometimes particular antennas or even just baselines have serious closure problems. For example, if one antenna has a baseband filter that is seriously different from those on all other antennas which in turn are all identical, it will not show closure errors. However if two antennas have different filters, the baseline between those two antennas can have very large closure offsets. It might be possible to improve the image substantially by editing out that baseline. If one antenna has special problems, it might be best just to edit that antenna out. There have been quite a few cases where eliminating one problem antenna has significantly improved an image.

13.4 An Example

For an example of real life VLBI imaging, I used the data at 13 cm on 0212+735 from a 24-hour VLBA geodesy experiment made on 28 August 1992. This source is a quasar at a redshift of 2.367. Typical of geodesy experiments, these observations were made simultaneously at 13 and 4 cm and the schedule called for rapid switching between sources to get good sky coverage in a short time. Each source, including 0212+735, was observed about once an hour. The scans were 86 seconds long and, allowing time for the correlator to sync up, provided about 1 minute of data each. The data were recorded in Mark III mode B with 14 baseband channels of 2 MHz (4 Mbps) each. There were 6 channels at 13 cm. After combining the 2 pairs that were upper and lower sidebands from the same baseband converters, there were 4 IF's in the AIPS data set. The IF's were spread over 100 MHz so there is a small multi-frequency synthesis effect. I made no effort to account for possible spectral index variations across the source. There were 6 antennas available at 13 cm but one recorded a bad tape and so was there for only half of the 24 hours. The antennas used were the VLBA antennas at Hancock, North Liberty, Fort Davis, Pie Town, Kitt Peak, and Owens Valley.

This observing style is actually of considerable interest. At least 30 sources were observed in these observations and, as will be seen, quite good images can be made of them. The test source, 0212+735, is especially favorable because of its high declination, but most of the sources had good (u, v) coverage. Figure 13.1 shows the (u, v) coverage that was obtained on 0212+735 at 13 cm for one of the IF's in these observations. It also shows the dirty beam, which is quite good for a "snap shot" observation. Note that single frequency observations with all 10 VLBA antennas have four times the bandwidth and about 3.6 times as many baselines (adjusting for the one that was missing half the time). The VLBA

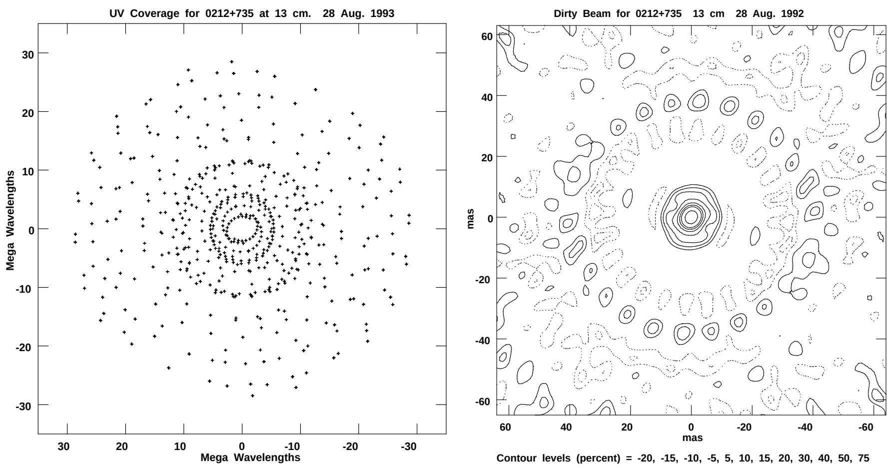

Figure 13.1: The (u, v) coverage and dirty beam for the example data set. This was 13 cm data on 0212+735 taken during a dual band (13 and 4 cm) geodesy experiment on the VLBA on 28 August 1993. The antennas involved were Hancock, North Liberty, Fort Davis, Pie Town, Kitt Peak, and Owens Valley (HN, NL, FD, PT, KP, OV).

correlator also has slightly lower losses so the sensitivity is about four times what was obtained here and as will be seen, it was possible to make a nearly noise limited image with a peak to off-source dynamic range of 1000 from the these data.

The data were processed on the Haystack correlator. All post processing including fringe fitting was done in AIPS. Initial calibration was done based on the VLBA monitor data and the advertised antenna sensitivities. Fringe fitting on the separate IF's was done by FRING. Editing was done using a variety of AIPS programs. This is an area where the software still needs help. Not much editing was required. However IF 3 on all baselines to North Liberty had low amplitudes compared to the other 3 IF's. It might have been possible for self-calibration to fix the problem, but to avoid bad data and to not waste time investigating the nature of the problem, I simply edited all affected data. The data were averaged to one point per IF but the IF's were not averaged together because of the large frequency differences.

Hybrid mapping was done using an AIPS procedure that edits CLEAN components (CCMRG and TABED—since replaced by CCEDT), self-calibrates (CALIB), images (MX), and displays the results (KNTR and LWPLA). This procedure simplifies doing the hybrid mapping considerably just by making it possible to worry only about the significant variables and by running all the programs in sequence.

Figure 13.2 shows the images for each hybrid iteration leading to the final image. All iterations are contoured to the same level, which is clearly too deep for the first two. The first image is based on a point source phase self-calibration and clearly has excess symmetry. The second is also a phase-only self-calibration. In fact, I made several more iterations with phase-only self-calibration while trying to use windows to remove the probably spurious features to the west of the brightest point. This did not work. The images changed as various things were tried, but they did not improve significantly until I allowed the amplitudes to vary. In fact, for one iteration, I used a too restrictive (u, v) range in the self-calibration and the image got worse. In retrospect, I should have been suspicious earlier that the excess symmetry was the result of amplitude problems. After a few more iterations using amplitude self-calibration in this first sequence, I obtained a good image at iteration 16 with an off-source rms of 1.7 mJy.

In an effort to provide a shorter sequence for this demonstration, I returned to iteration 2 and used it as the input for an amplitude self calibration to get the iteration 3 image that is shown in figure 13.2. That image is fairly good, but the noise is well above the noise limit and there are features at the east end of the source that look spurious based on experience. Iterations 4 through 9 were a good case of slow but steady progress. In each case, the off source rms improved by just enough that I kept going. Here I had the additional motivation that I knew from the first sequence that a better image was possible. Finally for iteration 10, I used an "L1" solution for the self-calibration. There was a significant improvement. Another iteration produced no further change so I threw it out and quit. The final off source rms is 1.6 mJy per beam which is very close to the expected noise limit. The final image of this sequence was essentially identical to the final image of the first sequence with differences on the peaks of the components of less than 0.5%. Note however that those differences are larger than the off source noise level which is a demonstration of the fact mentioned in previous chapters that the dynamic range and the uniqueness are not the same.

As mentioned earlier, it is a good idea to check how well the "model", consisting of the CLEAN components, fits the data. Figures 13.3 and 13.4 show the fit of the model and data on a few representative baselines for one IF. All other baselines and IF's were as good. These plots are for iteration 16 of the first sequence but are essentially identical to what would be obtained with iteration 10 of figure 13.2. The fit is very good for all data, giving no cause for concern. Plotting the model against amplitudes that have only had initial calibration applied does show that the self-calibration has lowered the amplitudes on the shortest baseline (Pie Town to Kitt Peak) by about 20 percent. This may indicate that the initial calibration on one or both of those stations was off, but may also indicate that there is unmodeled large scale structure and that the self-calibration on this few antennas has managed to absorb it into the gains. Data on shorter baselines, such as the VLBA baseline between Pie Town and Los Alamos, will be required to settle the issue.

The example hybrid mapping sequence described above did not require some of the tricks that can be useful with other data sets. The antennas were all nearly identical so there were no problems with weights. The (u, v) coverage was fairly good despite the fact that there was not all that much data. And, the

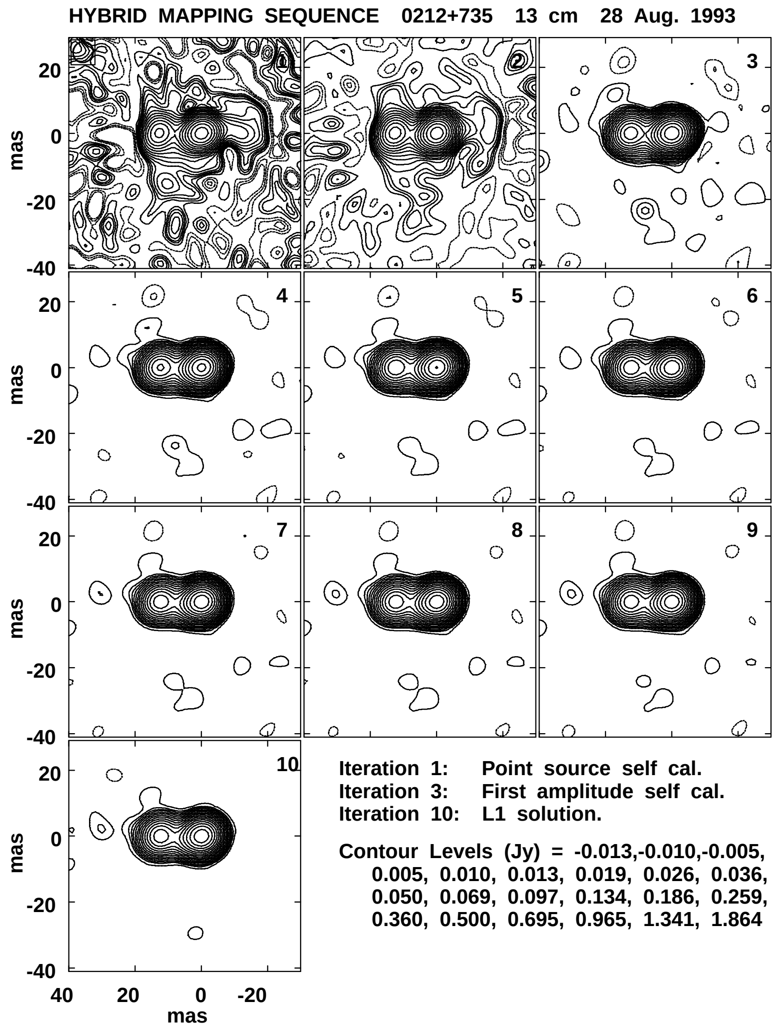

Figure 13.2: The hybrid mapping sequence on 0212+735. Good convergence, to an off-source rms noise level of 1.6 mJy was obtained in 10 iterations. The peak flux density in the final image is 1.75 Jy.

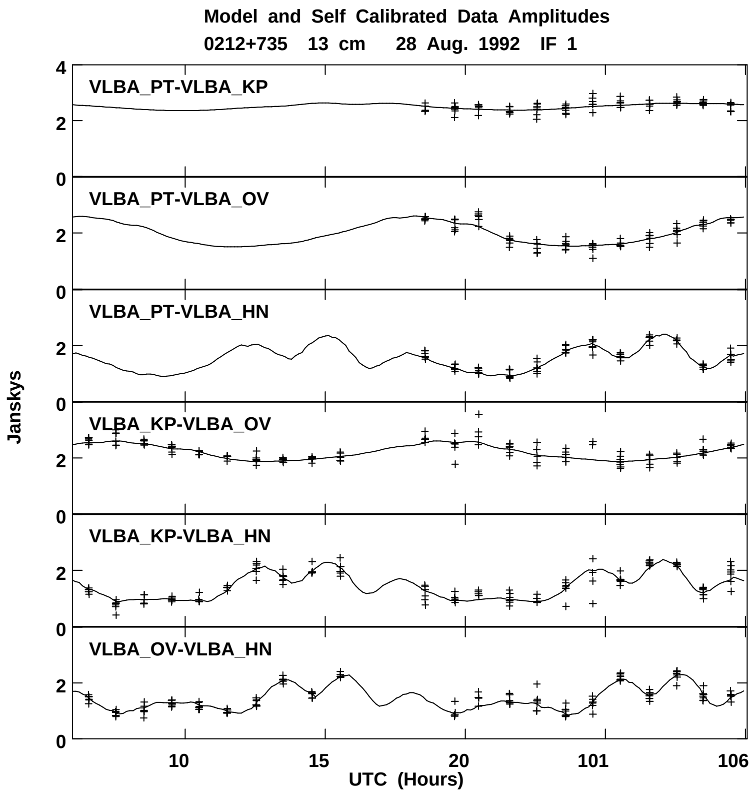

Figure 13.3: The match of data and model for the amplitudes for the hybrid mapping sequence on 0212+735. Note that a scan averages were used for the self-calibration so there was no opportunity to narrow the data spread within each 1 minute scan.

source was fairly simple. More complex sources can require more effort.

13.5 Difference Mapping

The imaging scheme discussed in this chapter is the full hybrid mapping sequence in which the image is completely remade and reCLEANed in each iteration. There is a scheme that was pioneered at Jodrell Bank called "difference mapping" in which the operations are done incrementally and in an integrated program. This saves significant computational effort and has been implemented with very nice, interactive software both at Jodrell Bank and in the Caltech program DIFMAP (Shepherd, Pearson and Taylor 1994). With DIFMAP, it is very easy to examine and edit the data during the imaging process and it is possible to get convergence very quickly. The latest version of this program is able to handle multiple IFs. For many applications, especially with smaller data

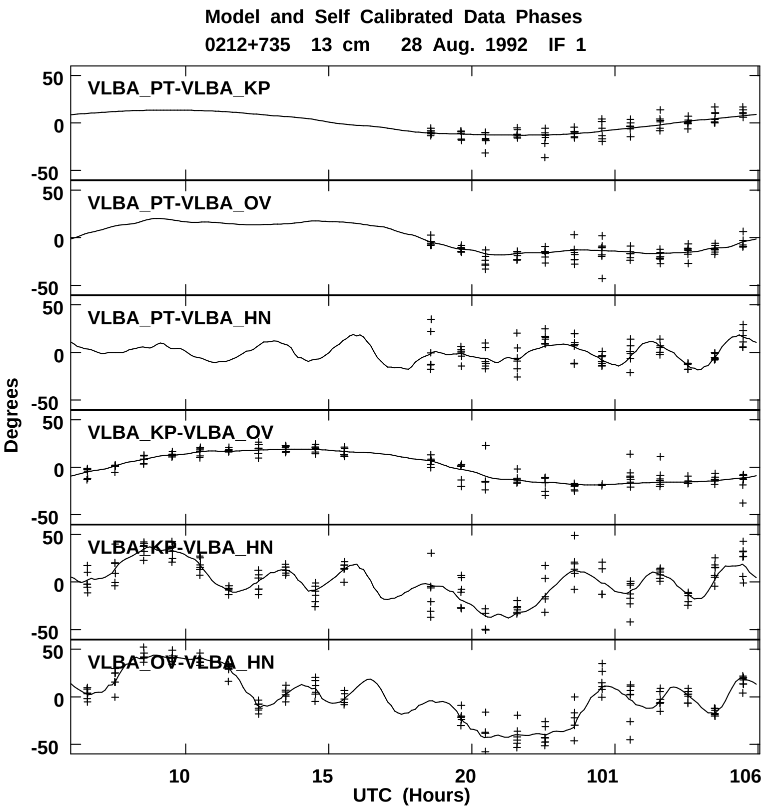

Figure 13.4: The match of data and model for the phases for the hybrid mapping sequence on 0212+735. Again the fit is quite good.

sets, this may be the fastest way to do the editing and imaging operations.

13.6 Advanced Techniques

The hybrid mapping scheme discussed in this chapter is meant for use on sources that are bright enough for self-calibration to work. It is also meant for data taken at a single frequency. It is still possible to image sources that are too weak for self-calibration if *phase connection* is used (see chapter 17). This is just the terminology used in VLBI for the usual phase calibration used on linked interferometers—using a nearby calibrator to calibrate the instrumental and atmospheric phases. For VLBI, this is difficult because the phases change rapidly in both time and angle, but it is possible for at least some frequencies some of the time.

It is also possible to image data using a range of frequencies. Since the values of u and v are dependent on frequency, this is a way of extending the

(u, v) coverage. However, the source structure is not necessarily the same at different frequencies because of spectral index variations. This must be taken into account. The method of using multiple frequencies for continuum source imaging is known as multi-frequency-synthesis or MFS (see chapter 16).

References

Briggs, D. S. 1995. *High Fidelity Deconvolution of Moderately Resolved Sources.* Ph. D. thesis, New Mexico Institute of Mining and Technology.

Cotton, W. D. 1979. "A method of measuring compact structure in radio sources using VLBI observations". *Astron. J.* **84**, 1122.

Pearson, T. J. 1991. "Caltech VLBI analysis programs". *Bull. Am. Astron. Soc.* **23**, 991.

Readhead, A. C. S., & Wilkinson, P. N. 1978. "The mapping of compact radio sources from VLBI data". *Astrophys. J.* **223**, 25.

Shepherd, M. C., Pearson, T. J., & Taylor, G. B. 1994. "DIFMAP: an interactive program for synthesis imaging". *Bull. Am. Astron. Soc.* **26**, 987.