AIPS Memo 119

TVSAD: interactive search and destroy

Eric W. Greisen

January 25, 2016

Abstract

TVSAD is a new task in AIPS, which first appeared in November 2014. It is an interactive version of the automatic source finding task, SAD (or "search-and-destroy") which has been in AIPS for a long time. SAD finds a list of Gaussian components and writes a residual image with the components removed. However, any components for which the fit appears bad are left in the image. TVSAD is an attempt to allow the user to avoid these left-behind components.

1 Introduction

Sources seen in interferometric images that are well resolved are almost certainly complex and not described by any simple mathematical form. However, source components that are unresolved or slightly resolved may be described as a Gaussian, for example, with a particular peak brightness, spatial location, and spatial extent. Since Gaussian "Clean beams" are used in typical deconvolution algorithms, the use of Gaussian fitting to the final images in order to make a list of the unresolved or weakly resolved components of the image seems appropriate.

"Manual" fitting of Gaussian source components is implemented in AIPS tasks IMFIT and JMFIT. These two tasks are nearly identical except for the mathematical algorithms used in the least-squares minimization. IMFIT uses the Levenberg-Marquardt algorithm while JMFIT uses Davidon's optimally conditioned variable metric (quasi-Newton) method for function minimization. The user is required to specify a window into the input image over which the fit is conducted, plus initial guesses for the peak brightness, pixel position in the image, major and minor axis dimensions, and position angle for each component to be fit. These tasks also allow the fitting of a planar "zero-level" as well. In November 2014, a verb called MFITSET was added to AIPS to allow the user to set these adverbs interactively.

The AIPS task "search-and-destroy" or SAD was initially written and contributed by Walter Jaffe. It has been used extensively in the years since and has fared reasonably well in studies of the completeness of the source lists produced. SAD functions by finding connected groups of pixels, called "islands," above a user-selected brightness limit. It then works through the list of islands in order of peak brightness, finding one or two maxima within the island, guessing component widths, and then attempting a Gaussian fit using the mathematical algorithm of JMFIT. The resulting Gaussian parameters are compared against a number of criteria and, if all seems well, the fit components are subtracted from the input image and added to the output component list and table. The criteria include peak brightness and total flux cutoffs, maximum rms in the residual image of the island, excessive component widths, peaks found outside the island or image, and total residual flux in the island. SAD can then loop to find more islands using a lower brightness limit than the previous one. Such looping allows the brightest sources to be fit first, ignoring low-level emission surrounding them, and then to fit weaker sources elsewhere in the field. Note, SAD deliberately leaves those components it did not fit well in the residual image and out of the component lists and output MF ("model-fit") table. The user is then required to handle these regions by hand with IMFIT, JMFIT, or other means.

The code for TVSAD was based on the code for SAD, but it was converted into a more "modern" memory-based *AIPS* style, rather than the older, disk-based style of SAD. The converted code was developed until it produced identical results to SAD before the interactive enhancements were added.

2 Interactive search and destroy: TVSAD

2.1 Inputs

The adverbs for SAD and TVSAD are identical, so all comments here apply more or less equally to both. The INNAME *et al.* adverbs select the image to be fit, while BLC and TRC select the sub-image plane to be fit. These tasks fit two-dimensional Gaussians in celestial coordinates, recording to which plane they apply, but not fitting anything in the third spectral or other axis domain. The tasks are restartable. If INVERS specifies an existing MF ("model fit") table, the components listed in that table will be subtracted from the input image before any islands and Gaussians are found in the remainder. Any new Gaussians found will be added to this table. Adverbs DORESID and OUTNAME *et al.* request the cataloging of the residual image and specify its name parameters. Since the tasks may be restarted, saving the residual image is not necessary until the fit is essentially complete.

The maximum number of Gaussians fit in any one island is eight. Adverb NGAUSS specifies the maximum number to be found over the entire image in all islands. At "major cycle i", the tasks look over the entire residual image to find "islands" (connected regions of pixels) above the brightness level set by CPARM(i). Then they sort the islands into descending order of peak brightness, make guesses for the component(s) in each island, and attempt fits. You should set DOWIDTH greater than zero to fit widths as well as positions and fluxes. If the peak residual in the island is greater than ICUT, a second fit with two components is attempted (if DOALL > 0) and the result with the lower peak residual is taken. The fit is then subjected to a variety of tests, controlled by adverbs GAIN and DPARM(1) through DPARM(7). The defaults for these adverbs are useful, but you may wish to adjust them. If the fit meets all tests, then the component(s) are added to the list of components and subtracted from the residual image. You may examine in detail the reasons for failures by setting the PRTLEV adverb. The tasks use the image rms to set a variety of default values and cutoff levels. You may read in an image of position-dependent rms under control of adverbs DPARM(9) and IN2NAME *et al.* to make these defaults be position dependent. Task RMSD can create such an image.

After all islands are fit for all non-zero CPARM(i), the two tasks apply position-dependent corrections for bandwidth smearing and primary beam under control of adverbs BWSMEAR and PBPARM. For Clean images, they also attempt to deconvolve the components. Adverb EFACTOR is used at this stage. Then they write a variety of outputs. The most important of these is the MF (model fit) table containing all possible parameters of each component. This table will be discussed in some detail below. The tasks write a printer display under control of adverb DOCRT, with the text file being written to FITOUT if desired. The SORT adverb determines the order of the printing with flux and component coordinate as the main choices. The deconvolved fit components may also be written to a CC ("Clean components") table, under control of OUTVERS, for use in modeling for calibration. Note that such CC tables may only be used with CMETHOD = 'DFT' since they contain a variety of component widths. Furthermore, the components may also be written to a ST ("stars") table, under control of STVERS, for use in display tasks.

2.2 Fitting

While TVSAD may be run for a time in a non-interactive mode, exactly like SAD, TVSAD always begins in an interactive mode and will resume that mode whenever an attempted fit does not meet the acceptance criteria. In the interactive mode, before a fit is attempted on the next island, the image of the island is displayed on the TV. The image will be linearly interpolated to make a reasonable display on the screen. Using data provided by Emmanuel Momjian, the display is as shown in Figure 1 with an island containing two maxima and an initial guess of two Gaussians. The first column of the menu offers the options:

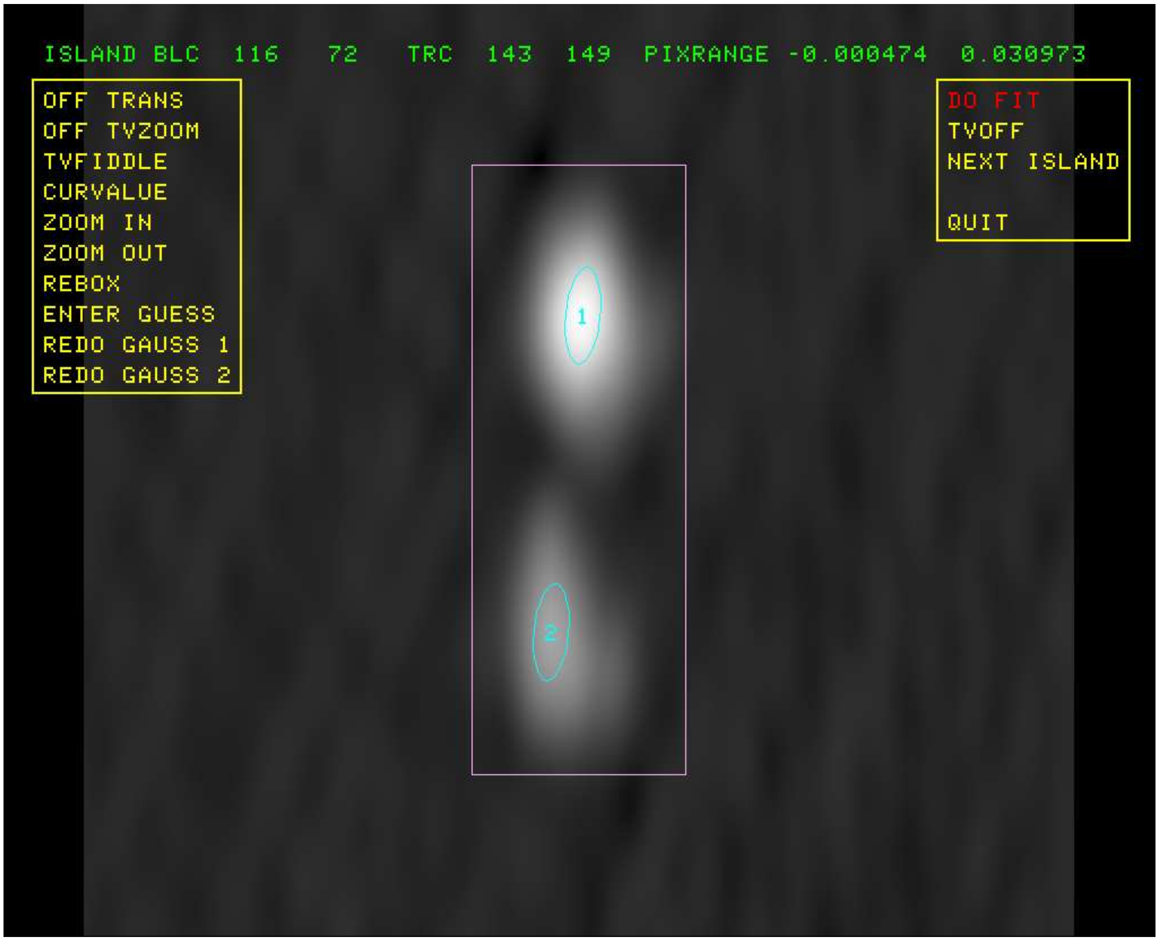

Figure 1: Island image with automatically generated initial guess.

`OFF TRANS`	Turn off black & white and color TV enhancements
`OFF TVZOOM`	Turn off TV zoom
`TVFIDDLE`	Interactive image enhancement and zoom
`CURVALUE`	Display image value selected by TV cursor
`ZOOM IN`	Reload image interpolated by one more step
`ZOOM OUT`	Reload image interpolated by one less step
`REBOX`	Change the island `BLC` and `TRC`
`ENTER GUESS`	Select number of Gaussians and their initial guess
`REDO GAUSS 1`	Change the initial guess parameters for the Gaussian labeled 1
`REDO GAUSS 2`	Change the initial guess parameters for the Gaussian labeled 2

The first four options are familiar AIPS functions to turn adjust black and white image contrast or color enhancement with pixel replication zoom, to display image values under the cursor, and to reset the black and white and color enhancements and the pixel replication zoom. The `ZOOM IN` and `ZOOM OUT` functions cause the image to be re-displayed with the degree of interpolation increased or decreased by one. `REBOX` allows you to adjust the rectangular island boundaries. `ENTER GUESS` allows you to set the initial guess for up to eight Gaussians within this island. You are directed to position the cursor to the peak of component one and press buttons `A`, `B`. or `C`. This sets the maximum and the x and y positions of the guess. Then you are directed to position the cursor at the half-power point of the component along the major axis and press buttons `A`, `B`. or `C`. Finally you are directed to position at a half-power point on the minor axis and press

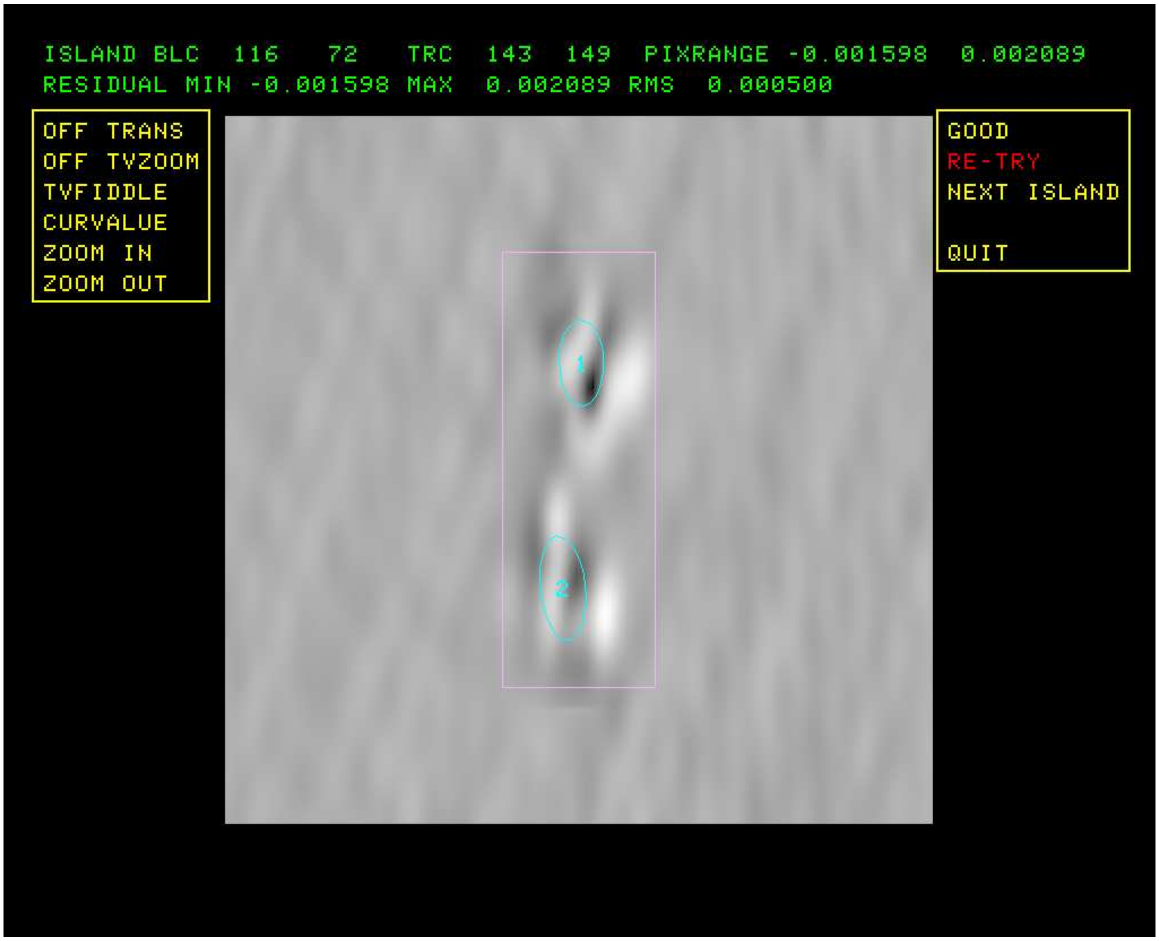

Figure 2: Island image with fit from automatically generated initial guess.

one of the 3 buttons. While you are moving the cursor around, a `CURVALUE`-like display will be given at the upper left corner to assist you. When the minor axis point is selected, an ellipse representing the component will appear on the screen in a reddish color (graphics plane 5). Then you are directed to point at the peak of component 2, and so on. To stop at any time, press button `D`. The number of Gaussians to be fit will then be set to the number fully specified.

A suitable number of `REDO GAUSS` n menu items will appear, reflecting the current number to be attempted in this island. These options follow the same pattern described above but change only the component number selected. The second menu column contains the operations that cause this island to be processed in some way. They are

`DO FIT`	Try to fit the island with the current initial guess
`TVOFF`	Fit this island with the current initial guess and then fit other islands until a failure occurs
`NEXT ISLAND`	Skip this island, leaving it unchanged in the residual image and omitting any components from the lists
`QUIT`	Exit the fitting process at this point

`NEXT ISLAND` tells `TVSAD` to skip this island and simply go on to the next with a new display like that in Figure 1. No components for the current island are kept in the lists or subtracted from the image. `QUIT` is even more drastic, telling the task to stop building the component list and computing further residuals

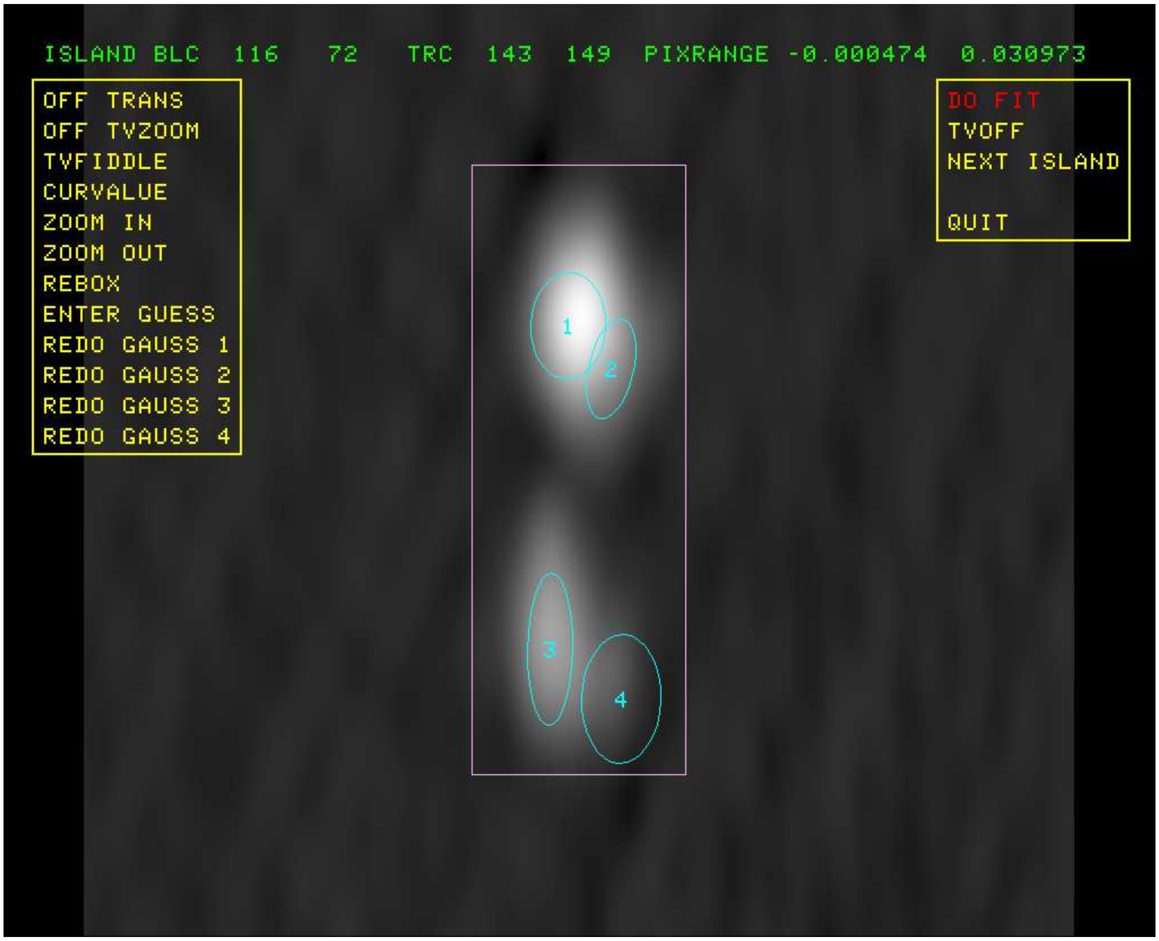

Figure 3: Island image with near 4-component initial guess entered with the TV cursor.

at this point and simply to go on to the output routines to write out whatever components were found previously.

`DO FIT` tells `TVSAD` to leave the current display and attempt to fit the current initial guess in this island. `TVOFF` is the same, except that it also turns off the interactive mode. That mode will be turned back on only after some failure in fitting is found. If the fit is found to be defective by the various component value tests, then the initial screen will be re-displayed with the fit values shown as the new initial guess. If `TVSAD` does not think the fit is defective, and the interactive mode remains on, the screen shown in Figure 2 will appear. This figure contains the result of fitting with the 2 Gaussians initially guessed and has rather high rms and residuals.

This second display of the residual image after the fit offers a menu of two columns. The first column contains some of the functions of the previous display, without the option to change any of the initial guess.

`OFF TRANS`	Turn off black & white and color TV enhancements
`OFF TVZOOM`	Turn off TV zoom
`TVFIDDLE`	Interactive image enhancement and zoom
`CURVALUE`	Display image value selected by TV cursor
`ZOOM IN`	Reload image interpolated by one more step
`ZOOM OUT`	Reload image interpolated by one less step

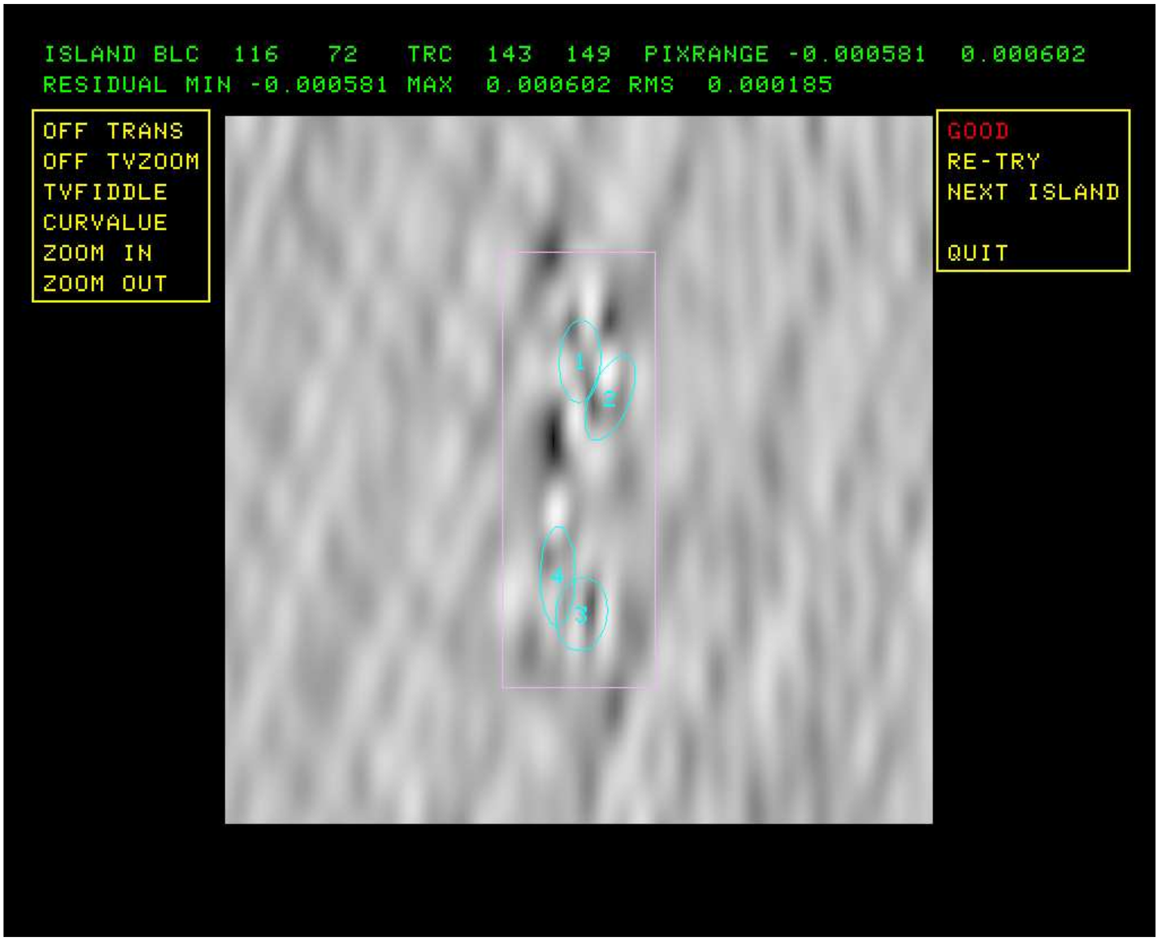

Figure 4: Island image with fit from user-generated initial guess.

The second menu column in the residual display contains functions

`GOOD`	Accept the current fit, adding components to the list and subtracting from the image
`RE-TRY`	Go back to the previous display and enter a new initial guess
`NEXT ISLAND`	Skip this island, leaving it unchanged in the residual image and omitting any components from the lists
`QUIT`	Exit the fitting process at this point

`NEXT ISLAND` and `EXIT` function as they did for the first display screen. `GOOD` tells the task to accept this solution, to save the components to the lists, to remove them from the image, and then to go on to the next island, if any. `RE-TRY`, selected here, tells `TVSAD` to return to the first display to allow the user to alter the initial guess. That alteration, to 4 Gaussians, is shown in Figure 3. The result of a `DO FIT` with much improved residual rms and extrema is shown in Figure 4. The user accepts this solution and goes on to the next island or, in this case, the output routines.

2.3 Outputs

All of the details of each component fit are written in the model fit (MF) table specified by INVERS. The contents and use of this table will be discussed below. If OUTVERS ≥ 0, a Clean components (CC) table will be written containing the components found. The deconvolved widths are written to the table unless DOWIDTH was false. In that case, widths of zero are written. This non-point CC table could be used as a model in, for example, CALIB to do a self-calibration. Note that CMETHOD = 'DFT' will be required since there will be a variety of fit widths in the table. If STVERS ≥ 0, a stars (ST) table will also be written (or appended). Tasks like KNTR can over-plot "star" positions and sizes on top of their contour and grey-scale displays. Verb TVSTAR allows such over-plots on the TV screen.

DOCRT controls a printer-like display of the results. Two groups of numbers are shown. In the first, the component peak brightness, flux, position, and width parameters are shown along with their uncertainties. In the second, the attempts at deconvolution (for Clean images only) are shown. The first three columns are the deconvolution at the formal solution. The second and third groups of 3 columns show the extrema in the possible deconvolved widths and position angles after trying all possible variations of $\pm$ EFACTOR times the various width uncertainties. A single character expresses the opinion that the components are probably resolved (R), probably unresolved (U), or its anybody's guess (?). The output from the run illustrated in the figures is shown below.

```
Test TVSAD  .SUBIM .      1     Disk  2     Plane    1     User     208

Window BLC  102   63   1   1  1  1  1 TRC  157  157   1   1  1  1  1
Sources found down to  0.000305 in JY/BEAM
Retry level  0.100000 (JY/BEAM ) plus gain 0.100
Reject components peak <  0.00000 in JY/BEAM
Reject components flux <  0.00000
Reject components outside window >   0.0 cells
Reject components outside image >    0.0 cells
Reject residual flux >   0.00000 with gain 0.100
Fluxes expressed in units of milliJY/BEAM
NOTE: Fluxes marked by * have been divided by 1000.
Errors determined by theory from RMS 101.58 microJy
Reference Center: 09 42 21.9911  06 23 34.960
All source widths and coordinates and their errors are in arc seconds
NO corrections for bandwidth smearing have been made
Source peaks and fluxes NOT corrected for primary beam

    #      Peak      Flux     RA---SIN   DEC--SIN    Maj     Min     PA
    1 L    3.582     7.417    -0.12100    0.16825   0.01605 0.00692 157
        (  0.096) (  0.279) (  0.00010    0.00017) (0.00043 0.00019   1)
    2      4.857    10.804    -0.11595    0.12998   0.01305 0.00914 173
        (  0.096) (  0.293) (  0.00008    0.00011) (0.00026 0.00018   2)
    3     31.107    62.556    -0.11573    0.17477   0.01461 0.00738 178
        (  0.096) (  0.274) (  0.00001    0.00002) (0.00005 0.00002   0)
    4     10.340    21.091    -0.11178    0.13672   0.01753 0.00624 179
        (  0.096) (  0.276) (  0.00002    0.00007) (0.00016 0.00006   0)

Component widths & PA: deconvolved at fit &  1.30 sigma low & high from fit
   #  MAJ-dec MIN-dec   PA   MAJ-low MIN-low   PA  MAJ-hi  MIN-hi   PA

    1 0.01145 0.00201  139 R 0.01057 0.00000  134 0.01230 0.00346  144
    2 0.00805 0.00433   89 R 0.00777 0.00298   79 0.00842 0.00528   95
    3 0.00799 0.00581   10 R 0.00787 0.00574    8 0.00811 0.00588   12
    4 0.01255 0.00430    3 R 0.01225 0.00415    2 0.01286 0.00444    4
```

2.4 The MF table

The model fit (MF) table is written by SAD, TVSAD, IMFIT, and JMFIT Gaussian-fitting tasks. It specifies all fit parameters, all uncertainties, the deconvolved widths and their extrema, bandwidth and primary beam correction factors, and certain of the parameters in pixel units to aid in restarting TVSAD and SAD. One row of the table is devoted to each fit component and the table may be sorted in a wide variety of ways.

The MF table has 41 columns and is supported for software by the parameter include file $INC/PMFC.INC which specifies the number of columns and a symbolic name for each column. There is also an initialization routine $APLSUB/MFINI.FOR to define and open such tables. However, due to the extreme number of columns, no table I/O routine has been defined. Instead, every task that accesses the table either reads or writes full table rows with the fundamental TABIO routine or gets selected column values with the GETCOL routine. Two tasks access MF tables. MF2ST simply translates flux-selected portions of the MF table to write a stars table of selected type. MFPRT is more complicated. Its basic mode is to print a selected set of the table columns onto the terminal or into a text file. It also has special formats for *AIPS* tasks STARS, BOXES, SETFC, and FACES.

3 Image data models

There are three tasks designed to modify an image adding a specified model and noise. Of these, IMMOD is the one of most interest here since it is intended for continuum images. The existing image may be scaled (or eliminated) and noise added. It adds NGAUSS components up to four using adverbs OPCODE to specify type, and FMAX, FPOS, and FWIDTH to specify peak brightness and pixel position and size. Alternatively, adverb INLIST may be used to specify up to 9999 components, one per line in the text file. Each non-comment line specifies

1.	Peak brightness (Jy/beam)
2.	Component X center (pixels)
3.	Component Y center (pixels)
4.	Major axis (pixels)
5.	Minor axis (pixels)
6.	Position angle (degrees CCW from North)

To make realistic test images, it might be best to eliminate the input values (FACTOR = 0). set a sensible noise level (FLUX), and then put in components of very small (but > 0) diameters. Then use CONVL to convolve the image, both noise and components, with a Gaussian "Clean beam." *AIPS* has a random number generator (RANDOM) which could be used to generate a collection of more or less random source components, but there are probably better ways to generate proper component lists to test SAD and TVSAD.