ERROR RECOGNITION & IMAGE ANALYSIS

Ed Fomalont (NRAO)
PREMABLE TO ERROR RECOGNITION and IMAGE ANALYSIS

• Why are these two topics in the same lecture?

  -- **Error recognition** is used to determine defects in the data  
    and image during and after the ‘best’ calibration, editing, etc.
  
  -- **Image analysis** describes the almost infinite ways in which  
    useful insight, information and parameters can be  
    extracted from the image.

• Perhaps the two topics are related to the reaction one has  
  when looking at an image after ‘good’ calibration,  
  editing, self-calibration, etc.
OBVIOUS IMAGE PROBLEMS

Rats!!

This can’t be right. This is either the most remarkable radio source ever, or I have made an error in making the image.

Image rms, compared to the expected rms, unnatural features in the image, etc are clear signs of problems.

How can the problems be found and corrected?
Great!!

After lots of work, I can finally analyze this image and get some interesting scientific results.

What were defects?

Two antennas had 10% calibration errors, and one with a 5 deg error, plus a few outlier points.

This part of the lecture.

How to find the errors and remove milliarcsec.
GENERAL PROCEDURE

Assuming that the data have been edited and calibrated reasonably successfully (earlier lectures). Self-calibration is usually necessary.

So, the first serious display of an image leads one--

to inspect again and clean-up the data with repetition of some or all of the previous reduction steps.

to image analysis and obtaining scientific results from the image.

But, first a digression on data and image display.
Numbers are proportional to the intensity.

Good for slow links, i.e., from the Gobi desert to Socorro.
These plots are easy to reproduce and printed

Contour plots give good representation of faint emission.

Profile plots give a good representation of the ‘mosque-like’
TV-based displays are most useful and interactive:

Grey-scale shows faint structure, but not good for high dynamic range and somewhat unbiased view of source.
DATA DISPLAYS (1)

List of u-v Data

<table>
<thead>
<tr>
<th>Source= J0121+11</th>
<th>Freq= 8.434858511</th>
<th>Sort= TB</th>
<th>1</th>
<th>RR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vis #</td>
<td>IAT</td>
<td>Ant</td>
<td>Su</td>
<td>Fq</td>
</tr>
<tr>
<td>2191 0/22:35:08.22</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>94220</td>
</tr>
<tr>
<td>3971 0/22:43:44.34</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>97659</td>
</tr>
<tr>
<td>6431 0/23:03:05.15</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>106307</td>
</tr>
<tr>
<td>6611 0/23:03:14.98</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>106364</td>
</tr>
<tr>
<td>6791 0/23:03:24.81</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>106421</td>
</tr>
<tr>
<td>6971 0/23:03:34.64</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>106477</td>
</tr>
<tr>
<td>7151 0/23:03:44.47</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>106534</td>
</tr>
<tr>
<td>7331 0/23:07:54.30</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>106591</td>
</tr>
<tr>
<td>7511 0/23:15:06.84</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>109027</td>
</tr>
<tr>
<td>7691 0/23:15:15.67</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>109081</td>
</tr>
<tr>
<td>7871 0/23:15:26.50</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>109135</td>
</tr>
<tr>
<td>8051 0/23:15:36.33</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>109189</td>
</tr>
<tr>
<td>8231 0/23:15:46.16</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>109243</td>
</tr>
<tr>
<td>8411 0/23:15:55.99</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>109297</td>
</tr>
<tr>
<td>9701 0/23:31:02.36</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>114020</td>
</tr>
<tr>
<td>9791 0/23:31:06.29</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>114040</td>
</tr>
<tr>
<td>10301 0/23:31:29.88</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>114156</td>
</tr>
<tr>
<td>10861 0/23:39:02.08</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>116320</td>
</tr>
<tr>
<td>10951 0/23:39:06.01</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>116339</td>
</tr>
<tr>
<td>11171 0/23:39:15.84</td>
<td>5-6</td>
<td>1</td>
<td>0</td>
<td>116384</td>
</tr>
</tbody>
</table>

Very primitive display, but sometimes worth-while: e.g.s, can search on Amp > 1.0, for example, or large Wt. Often need precise times in order to flag the data appropriately.
DATA DISPLAYS(2)

Visibility Amplitude

versus

Projected uv

spacing

General trend of data.

Useful for relatively strong Sources.

Triple source model.

Large component cause rise at short spacings.

Oscillation at longer

Mega Wavelength

Jy
Visibility amplitude and phase versus time for various baselines. Good for determining continuity of the data. Should be relatively smooth with time. Outliers are obvious.
Weights of antennas 4 with 5, 6, 7, 8, 9

All u-v data points have a weight. The weight depends on the antenna sensitivity, measured during the observations. The amplitude calibration values also modify the weights.

Occasionally the weight of the points become very large, often caused by subtle software bugs.

A large discrepant weight causes the same image artifacts as a large discrepant visibility value.

Please check weights to make sure they are reasonable.
**IMAGE PLANE OR DATA (U-V) PLANE INSPECTION?**

**Errors obey Fourier relationship**

Narrow features $\leftrightarrow$ Wide features  
(easier to find narrow features)

Orientations are orthogonal

Data uv amplitude errors $\leftrightarrow$ symmetric image features

Data uv phase errors $\rightarrow$ asymmetric image features

<table>
<thead>
<tr>
<th>Image Plane</th>
<th>uv Plane</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta x = \Delta \theta$</td>
<td>$\Delta u = L^{-1}$</td>
</tr>
<tr>
<td>$L = n\Delta \theta$</td>
<td>$U_{\text{max}} = \Delta \theta^{-1}$</td>
</tr>
</tbody>
</table>
GOLDEN RULE OF FINDING ERRORS

---Obvious outlier data (u-v) points:

100 bad points in 100,000 data points gives an 0.1% image error

(unless the bad data points are 1 million Jy)

LOOK at DATA to find gross problem (but don’t go overboard)

FURTHER OPPORTUNITIES TO FIND BAD DATA!

---Persistent small data errors:

egs a 5% antenna gain calibration error is difficult to see

in (u-v) data (not an obvious outlier), but will produce a 1% effect in image with specific
ERROR RECOGNITION IN THE U-V PLANE

Editing obvious errors in the u-v plane

---Mostly consistency checks assume that the visibility cannot change much over a small change in u-v spacing
---Also, double check gains and phases from calibration processes. These values should be relatively stable.

See Summer school lecture notes in 2002 by Myers
See ASP Vol 180, Ekers, Lecture 15, p321
VISIBILITY AMPLITUDE PLOTS

Amp vs uvdist shows outliers
Amp vs time shows outliers in last scan
Amp vs time without ant 7 should good data

(3C279 VLBA data at 43 GHz)
VISIBILITY AMPLITUDE RASTERS

BASELINE
Ant 1 2 3 4 5 6 7 8

Raster scan of baseline versus time immediately shows where the bad data are

Pixel range is 5 to 20 Jy

Bad data can be flagged with an interactive clipping control

(Last two scans from previous slide)

Use AIPS task TVFLG, CASA viewer
Fourier transform of nearly symmetric Jupiter disk

Butler lecture: Solar System Objects
Drop-outs at Scan Beginnings

Often the first few points of a scan are low. Egs. antenna not on source.

Software can remove these points (aips,casa ‘quack’)

Flag extension:
Should flag all sources in the same manner even though you cannot see dropout for weak sources
Editing Noise-dominated Sources

No source structure information is detected.
Noise dominated.

All you can do is remove outlier points above 0.3 Jy. Precise level not important as long as large outliers removed.
USING TVFLG (VIEWER) DISPLAY on a source

Plot amplitude
rms

ANT-23 problems

quack these!

Baseline-->
RFI environment worse on short baselines

Several 'types': narrow band, wandering, wideband, ...

Wideband interference hard for automated routines

Example using AIPS tasks FLGIT, FLAGR

Unfortunately, still best done by hand!
Some Questions to ask?

Noise properties of image:
Is the rms noise about that expected from integration?
Is the rms noise much larger near bright sources?
Are there non-random noise components (faint wave features)?

Funny looking Structure:
Non-physical features; stripes, rings, symmetric
Negative features well-below 4x rms noise
Does the image have characteristics in the dirty?

Image-making parameters:
Is the image big enough to cover all significant?
Is cell size too large or too small? ~4 points
Is the resolution too high to detect most of the
Data bad over a short period of time

Results for a point source using VLA. 13-5 min observation over 10 hr.
Images shown after editing, calibration and deconvolution.

no errors:
max 3.24 Jy
rms 0.11 mJy

10% amp error for all antennas for 1 time period
rms 2.0 mJy

6-fold symmetric pattern due to VLA "Y". Image has properties of dirty beam.
EXAMPLE 2

Short burst of bad data

Typical effect from one bad u-v point:
Data or weight

10 deg phase error
for one antenna at
one time
rms 0.49 mJy

20% amplitude error
for one antenna at
1 time
rms 0.56 mJy
(self-cal)

anti-symmetric ridges

symmetric ridges
EXAMPLE 3

Persistent errors over most of observations

NOTE: 10 deg phase error to 20% amplitude error cause similar sized artifacts

10 deg phase error for one antenna all times
rms 2.0 mJy

20% amp error for one antenna all times
rms 2.3 mJy

rings – odd symmetry

rings – even symmetry
EXAMPLE 4

Spurious Correlator Offset Signals

Occasionally correlators produce ghost signals or cross talk signals. Occurred last year during change over from VLA to EVLA system.

Symptom: Garbage near phase center, dribbling out into image.

Image with correlator offsets

Image after correlation of offsets

μJy
DECONVOLUTION ERRORS

Even if the data are perfect, image errors and uncertainties will occur because the (u-v) coverage is not adequate to map the source structure.

The extreme rise of visibility at the short spacings makes it impossible to image the extended structure. You are better of imaging the source with a cutoff below about 2 kilo-wavelengths.

Get shorter spacing or single-dish data.
The dirty beam has large, complicated side-lobe structure.

It is often difficult to recognize any details on the dirty image.

An extended source exaggerates the side-lobe structure.
CLEANING WINDOW SENSITIVITY

Spurious emission is always associated with higher sidelobes in dirty-beam.

Tight Box
Big Box
One small clean box
(interactive clean shown next)

Middle Box
Dirty Beam
One clean box
around all emission

Clean entire inner map
How Deep to Clean?

Under-cleaned

- Residual sidelobes dominate the noise
- Emission from second source sits atop a negative "bowl"

Over-cleaned

- Regions within clean boxes appear "mottled"

Properly cleaned

- Background is thermal noise-dominated; no "bowls" around sources.
FINDING HIDDEN BAD DATA

Chandra Deep Field South
Peak = 45 mJy, rms = 0.02 mJy

Source to NE in first Primary beam sidelobe

See Lectures Perley on Wide-field Imaging, and Uson on High dynamic Range Imaging

Center of Field
Fourier Transform Dirty Image

Shows the u-v data as gridded just before imaging.

Diagonal lines caused by structure in field.

A few odd points are not very noticeable.
Fourier Transform Clean Image

Shows the u-v data from clean image.

Diagonal lines still present. Notice that clean does an interpolation in the u-v plane between u-v tracks.

The odd points are smeared, but still present. These produce the low level ripples.
Bad weighting of a few u-v points

After a long search through the data, about 30 points out of 300,000 points were found to have too high of a weight by a factor of 100. Effect is <1% in image.

Cause??
Sometimes in applying calibration produced an incorrect weight in the data. Not present in the original data.

These problems can sneak up on you. Beware.
Improvement of Image

Removal of low level ripple improves detectability of faint sources
SUMMARY OF ERROR RECOGNITION

Source structure should be ‘reasonable’, the rms image as expected, and the background featureless.

UV data
Look for outliers in u-v data using several plotting methods.
Check calibration gains and phases for instability.
Look at residual data (uv-data - clean component).

IMAGE plane
Do defects resemble the dirty beam?
Are defect properties related to possible data error?
Are defects related to possible deconvolution problems?
IMAGE ANALYSIS

Ed Fomalont

Eleventh Synthesis Imaging Workshop
Socorro, June 10-17, 2008
IMAGE ANALYSIS

• Input: Well-calibrated data-base producing a high quality image

• Output: Parameterization and interpretation of image or a set of images

This is very open-ended

  Depends on source emission complexity

  Depends on the scientific goals
IMAGE ANALYSIS OUTLINE

• Multi-Resolution of radio source.
• Parameter Estimation of Discrete Components
• Polarization Data
• Image Comparisons
• Positional Registration
Different aspect of source structure can be seen at various resolutions, shown by the ellipse in the lower left corner of each box.

SAME DATA USED FOR ALL IMAGES

For example, Outer components are small from SU resolution There is no extended emission from low resolution
Imaging and Deconvolution of Spectral Line Data:

Type of weighting in imaging

HI contours overlaid on optical images of an edge-on galaxy
PARAMETER ESTIMATION

Parameters associated with discrete components

• **Fitting in the image**
  - Assume source components are Gaussian-shaped
  - Deep cleaning restores image intensity with Gaussian-beam
  - True size * Beam size = Image size, if Gaussian-shaped. Hence, estimate of true size is relatively simple.

• **Fitting in (u-v) plane**
  - Better estimates for small-diameter sources
  - Can fit to any source model (e.g., ring, disk)
  (see non-imaging analysis)

• **Error estimates of parameters**
  - Simple ad-hoc error estimates
  - Estimates from fitting programs
**IMAGE FITTING**

### Component 1-Gaussian

- **Peak intensity**: $0.300 \pm 0.005$ JY/BEAM
- **Integral intensity**: $0.302 \pm 0.008$ JANSKYS
- **X-position**: $270.991 \pm 0.001$ pixels
- **Y-position**: $267.018 \pm 0.001$ pixels
- **Major axis**: $0.53 \pm 0.01$ pixels
- **Minor axis**: $0.00 \pm 0.05$ pixels
- **Pos. angle**: $21.6 \pm 1.1$ deg

### Component 2-Gaussian

- **Peak intensity**: $0.104 \pm 0.005$ JY/BEAM
- **Integral intensity**: $0.998 \pm 9.47$ JANSKYS
- **X-position**: $255.986 \pm 0.0029$ pixels
- **Y-position**: $257.033 \pm 0.0032$ pixels
- **Major axis**: $19.99 \pm 0.02$ pixels
- **Minor axis**: $9.98 \pm 0.03$ pixels
- **Pos. angle**: $135.3 \pm 0.1$ deg

### Component 3-Gaussian

- **Peak intensity**: $0.393 \pm 0.004$ JY/BEAM
- **Integral intensity**: $0.403 \pm 0.008$ JANSKYS
- **X-position**: $241.007 \pm 0.001$ pixels
- **Y-position**: $241.988 \pm 0.001$ pixels
- **Major axis**: $1.54 \pm 0.01$ pixels
- **Minor axis**: $0.21 \pm 0.01$ pixels
- **Pos. angle**: $3.6 \pm 0.2$ deg

---

AIPS task: JMFIT

Casa tool: imfit
(U−V) DATA FITTING

DIFMAP has good u−v fitting algorithm

Fit model directly to (u−v) data
Contour display of image
Compare mode to data
Ellipses show true component

Greg Taylor, Tuesday June 17, “Non-image Data Analysis”
Size. (super-resolution?)
COMPONENT ERROR ESTIMATES

\[ P = \text{Component Peak Flux Density} \]
\[ \sigma = \text{Image rms noise} \quad P/\sigma = \text{signal/noise} = S \]
\[ B = \text{Synthesized beam size} \]
\[ \theta_i = \text{Component image size} \]
\[ \Delta P = \text{Peak error} = \sigma \]
\[ \Delta X = \text{Position error} = B / 2S \]
\[ \Delta \theta_i = \text{Component image size error} = B / 2S \]
\[ \theta_t = \text{True component size} = (\theta_i^2 - B^2)^{1/2} \]
\[ \Delta \theta_t = \text{Minimum component size} = B / S^{1/2} \]

eg. \( S=100 \) means can determine size of \( B/10 \)
Comparison and Combination of Images of Many Types

FORNAX–A Radio/Optical field

Radio is red
  Faint radio core
  in center of
  NGC1316

Optical in
  blue-white

Frame size is
  60’ x 40’
**LINEAR POLARIZATION**

Multi-purpose plot

Contour - I, Q, U Pol

Grey scale - P Pol

sqrt(Q^2 + U^2) - noise

Line segments - P angle

atan2(0.5*Q/U)
COMPARISON OF RADIO/X-RAY IMAGES

Contours of radio intensity at 5 GHz
Dots represent X-ray Intensity (photons) between 0.7 and 11.0 KeV

Contours of radio intensity at 5 GHz
Color intensity represents X-ray intensity smooth to radio resolution
Color represents hardness of X-ray (average weighted frequency)
Blue - soft (thermal)
SPECTRAL LINE REPRESENTATIONS

Intensity Image
Sum of velocity
Amount of HI
Red high, Blue low

Average velocity
Red low vel
Blue high vel
Rotation

Second moment
Velocity width
Turbulence?
Solid-body Rotation in Inner parts of a galaxy

QuickTime™ and a YUV420 codec decompressor are needed to see this picture.
Visualizing Spectral Line Data: Channel Images

Greyscale+contour representations of individual channel images
Visualizing Spectral Line Data: Channel Images

- Velocity
- Right Ascension
- Declination

Images showing channel images at velocities of 471.9 km/s and 456.4 km/s.
IMAGE REGISTRATION AND ACCURACY

• Separation Accuracy of Components on One Image:
  Limited by signal to noise to 1% of resolution.
  Position errors of 1:10000 for wide fields, i.e. 0.1" over 1.4 GHz PB

• Images at Different Frequencies:
  Multi-frequency. Use same calibrator for all frequencies.
  Watch out at frequencies < 2 GHz when ionosphere can
  produce displacement. Minimize calibrator-target separation

• Images at Different Times (different configuration):
  Use same calibrator for all observations. Daily
troposphere changes
  can produce position changes up to 25% of the
DEEP RADIO / OPTICAL COMPARISON

Grey-Scale:
Optical emission faintest is 26-mag

Contours:
Radio Emission faintest is 10 $\mu$Jy
Radio Source Alignment at Different Frequencies

Self-calibration at each frequency aligns maximum at (0,0) point
Frequency-dependent structure causes relative position of maximum to change
Fitting of image with components can often lead to proper registration

43 GHz: res = 0.3 mas
23 GHz: res = 0.6 mas
15 GHz: res = 0.8 mas

(Reid Lecture on Astrometry, Walker Lecture on VLBA Upgrade)
IMAGE ANALYSIS: SUMMARY

• Analyze and display data in several ways
  Adjust resolution to illuminate desired interpretation, analysis

• Parameter fitting useful, but try to obtain error estimate
  Fitting in u-v plane, image plane

• Comparison of multi-plane images tricky (Polarization and Spectral Line)
  Use different graphics packages, methods, analysis tools

• Registration of a field at different frequencies or wave-bands and be subtle.
  Often use adhoc methods by aligning ‘known’ counterparts