The Duchamp Source Finder

Matthew Whiting
CSIRO Astronomy & Space Science

April 29, 2010

Philosophy of Duchamp

- Provide a generic source-finding tool capable of searching 3D astronomical data

- Emphasis on HI data, but not limited to that
- No assumptions made about the nature of the sources
- Aim is to provide locations of significant sources
- Make it user-friendly
  - Straight-forward inputs via simple parameter file
  - Flexibility to account for different data types and search types
  - Graphical feedback showing location, images & spectra of detected sources
  - Stand-alone package, not tied to a larger data-reduction suite

Program flow

- Read parameters
- Read image
- Prepare image for searching
  - Baseline removal, smoothing, wavelet reconstruction, inversion
- Segmentation
- Source aggregation
- Undo baseline removal etc
- Source parametrisation
- Report results

Segmentation

- Is a given voxel part of a source, or part of the background?

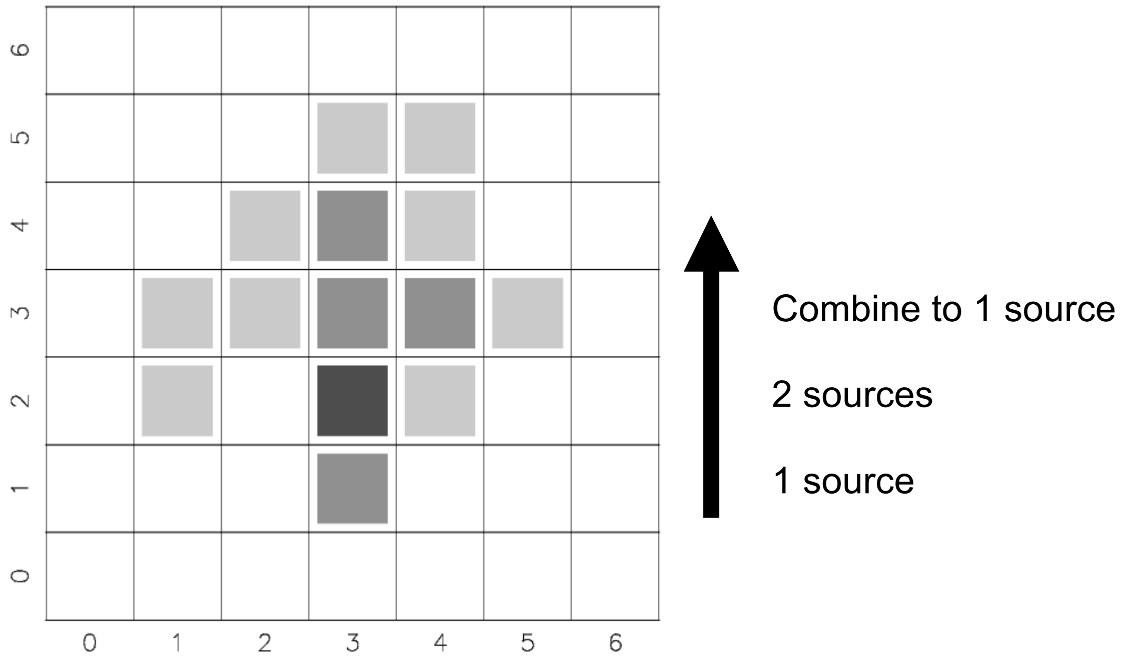
- Determine threshold, then apply voxel-by-voxel
- Threshold can from one of:
  - Flux level (e.g. 100 uJy/beam)
  - Signal-to-noise level (e.g. 5-sigma)
  - Determined by False Discovery Rate method (Miller et al 2001, Hopkins et al 2002)

- If S/N or FDR threshold, need to determine statistics
- This is done with full dataset, or on specified subsection

- Can use robust methods, with the median estimating the mean, and median absolute deviation from median estimating the rms
  - Need to correct, assuming Gaussian statistics
  - rms = MADFM / 0.6744888

Source Aggregation

- To identify sources, need to aggregate detected pixels
- Done in 2D by raster-scanning through image and identifying adjoining groups of pixels



3D Merging

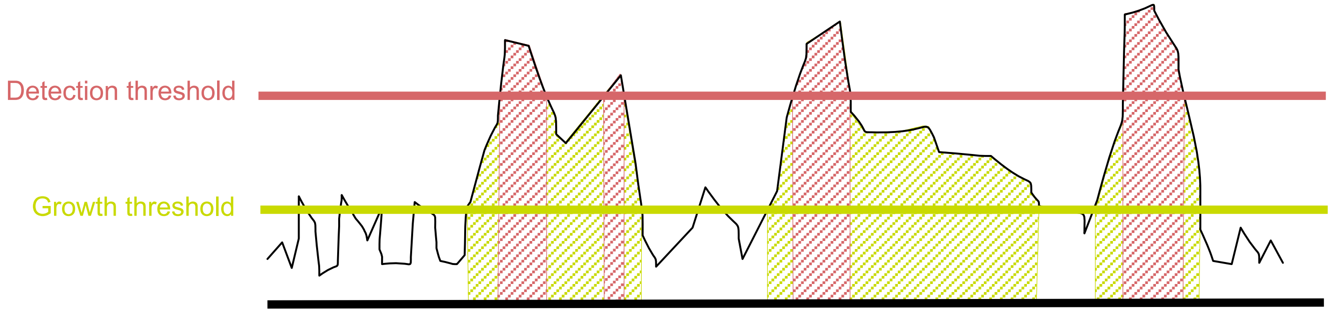
- Combine simple 2D objects to form larger 3D (or even 2D) sources

- As sources are identified, combine with “connected” previously-detected objects
- Also have separate complete merging phase, where all possible combinations are examined.

- Can use different criteria to judge “connectedness”
  - Can combine objects that are not adjacent but lie within predefined separation
  - Can have different separations in spatial & spectral directions
- No assumptions made in software as to nature of sources
  - No requirement to be particular shape
  - Only minimum number of pixels & channels

Growing objects

- Ability exists to specify two thresholds
- One is the detection threshold to build up the source list
- Then “grow” each object down to the second threshold
- Examine neighbouring voxels and add to source if above 2nd threshold
- A way of expanding the size of objects to a low flux level, without filling the source list with many faint (spurious?) sources
- Objects may grow into each other - they will be merged



The THIRD Dimension

- Real spectral sources have extent in three dimensions
- How to emphasise this structure over spurious signals?

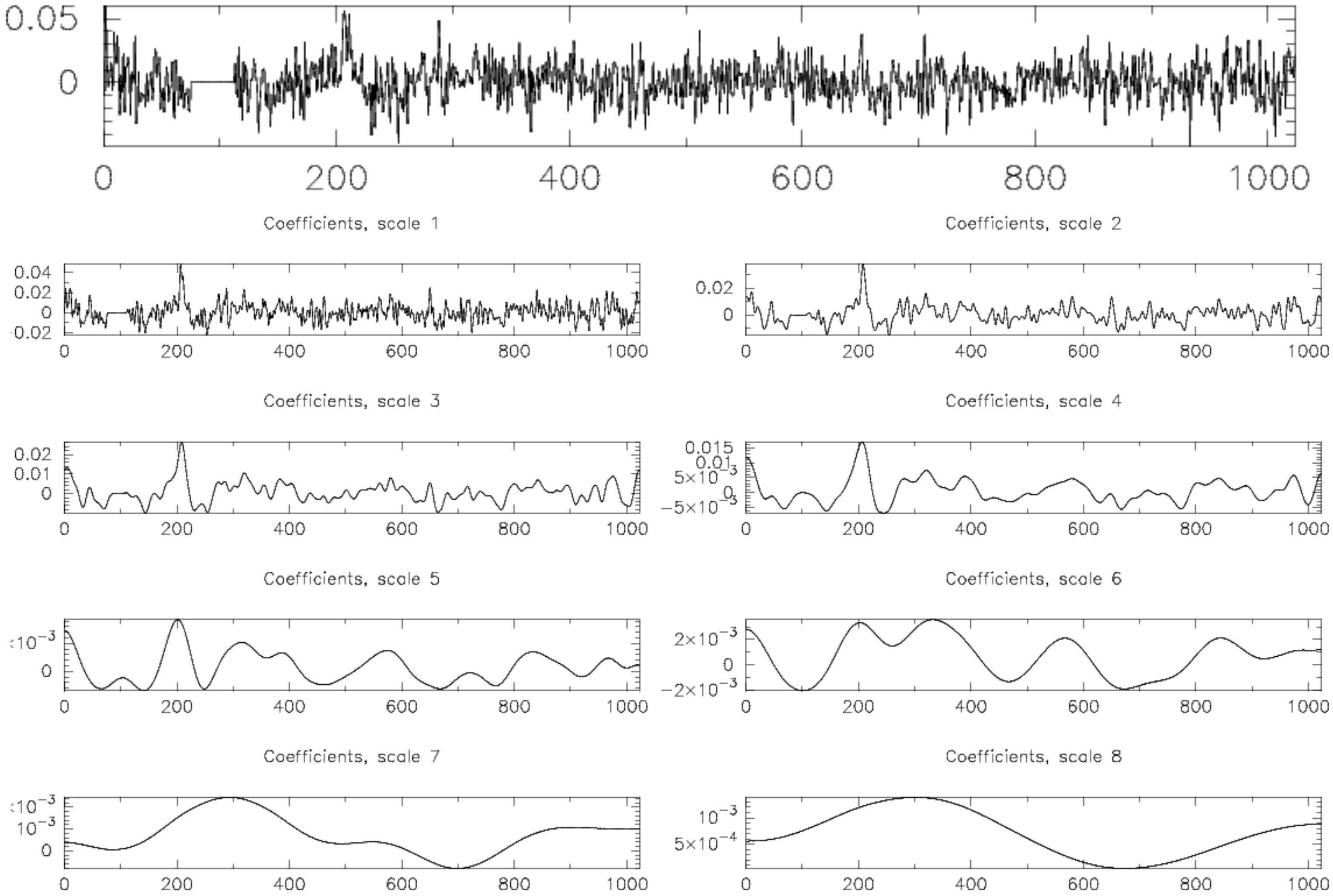
- Smooth data in one direction while searching in an orthogonal one
  - Pick a "typical" scale and smooth by that
  - E.g. typical spectral width of signals, then search channel maps
  - Duchamp: Hanning smooth in spectral direction or Gaussian smooth in spatial directions

- Use a range of scales via multi-resolution wavelet transforms
  - Logarithmically increase the scale of smoothing
  - Keep only the pixels with significant signal and restore
  - Do source-detection on reconstructed cube

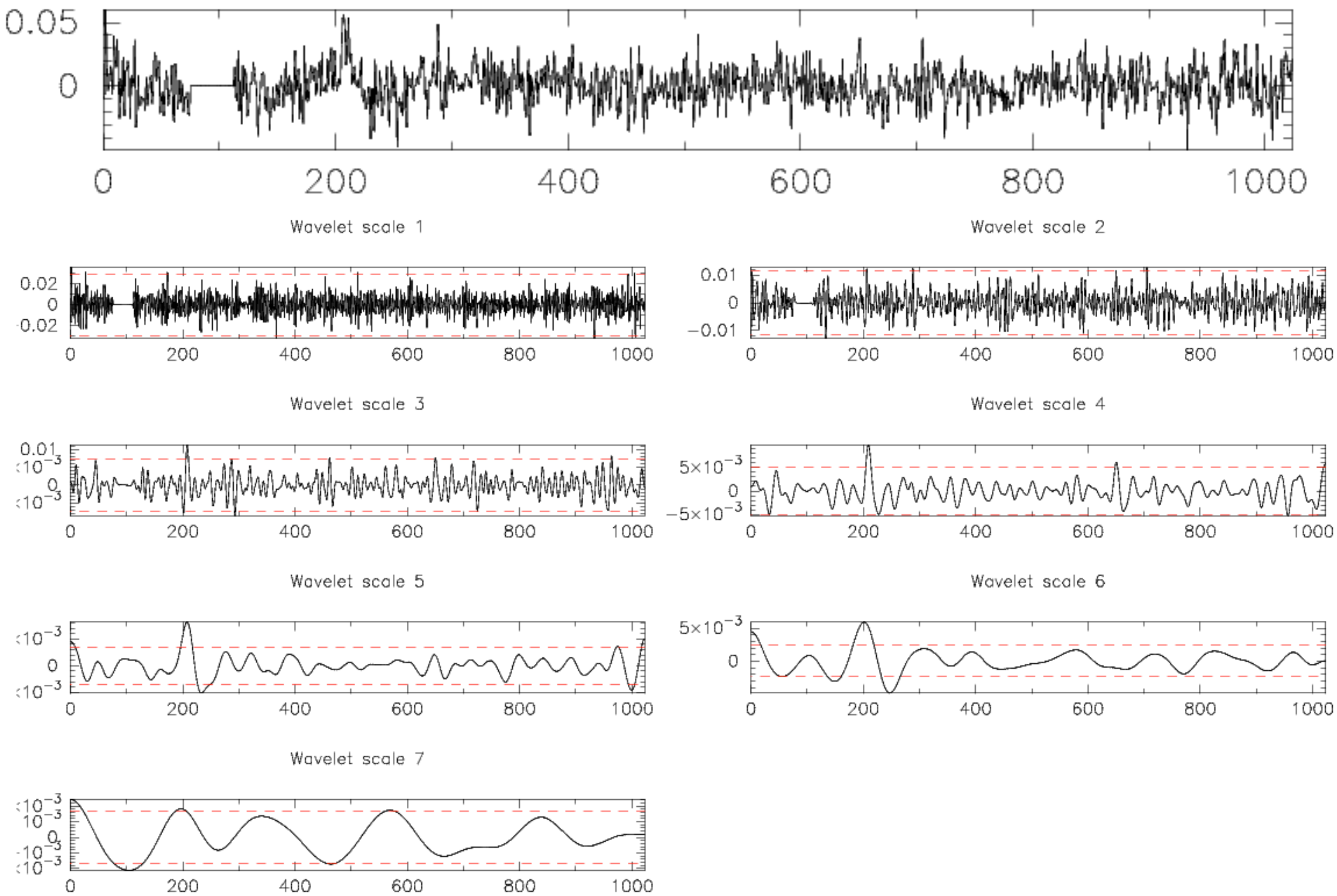
Wavelet reconstruction: algorithm

- Start with a spectrum (the input data): $S^0 = \{S_i^0\}, \forall i \in [1, N]$
- Also have a filter, used to smooth the data: $F^1 = \{F_j^1\}, \forall j \in [1, f]$
- Convolve the spectrum with the filter to produce first smoothed array $S^1 = \{S_i^1\} = S^0 \otimes F^1$
- Subtract the coefficients from the spectrum to produce the wavelet array $W_i^1 = S_i^0 - S_i^1$
- Apply some threshold to the wavelet array, so that only pixels with signal are kept. $$\hat{W}_i^1 = \begin{cases} W_i^1 & |W_i^1| \geq T^1 \\ 0 & |W_i^1| < T^1 \end{cases}$$
- Double the spacing between the filter coefficients
- Convolve the smoothed array with the filter
- Produce the wavelet array and apply threshold.
- Continue until size of filter ~ size of spectrum
- Reconstruction: add thresholded wavelet arrays, plus final smoothed spectrum. $R_i = \sum \hat{W}_i^k + S_i^n$

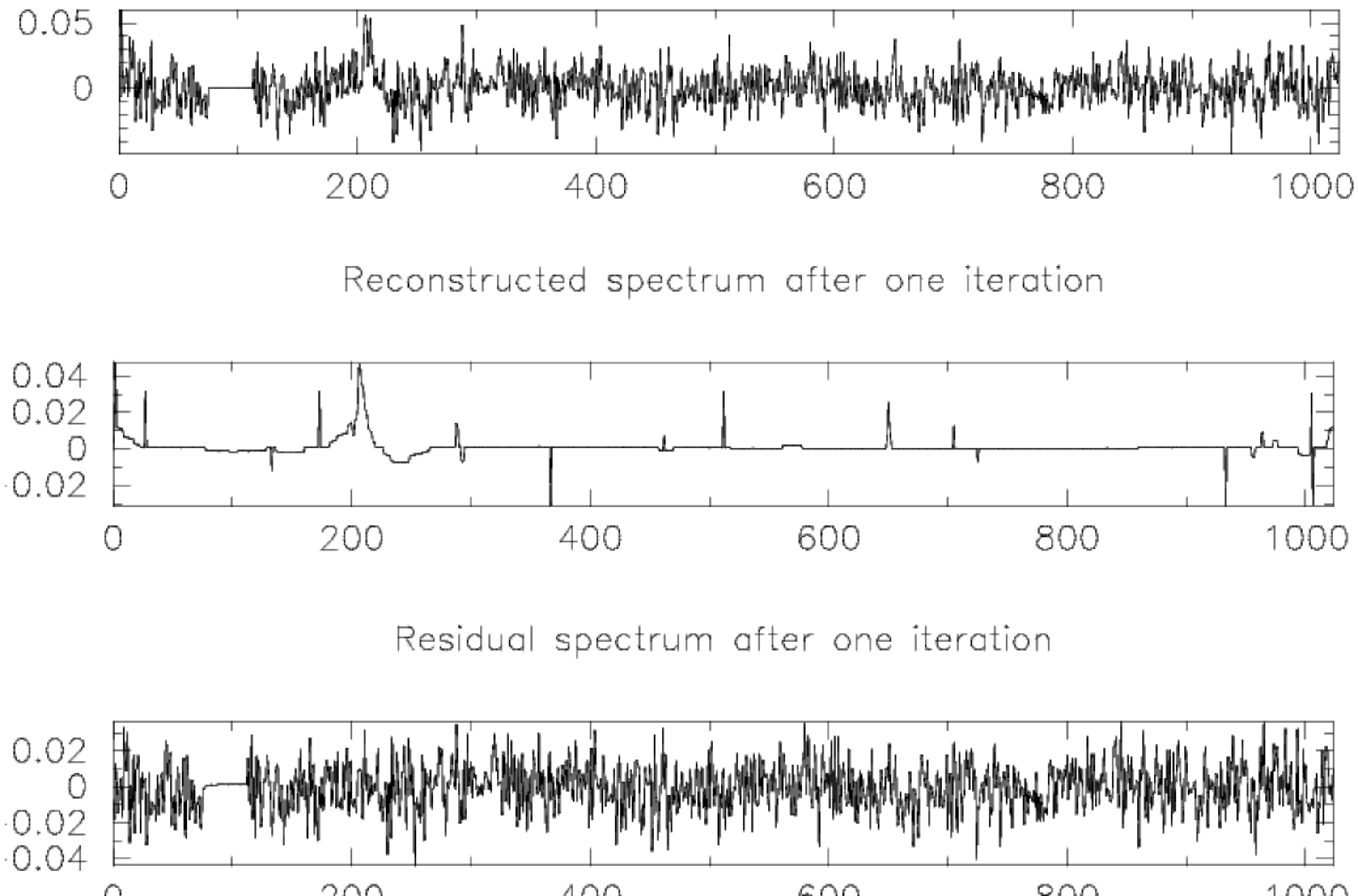
Wavelet reconstruction: example



Wavelet reconstruction: example



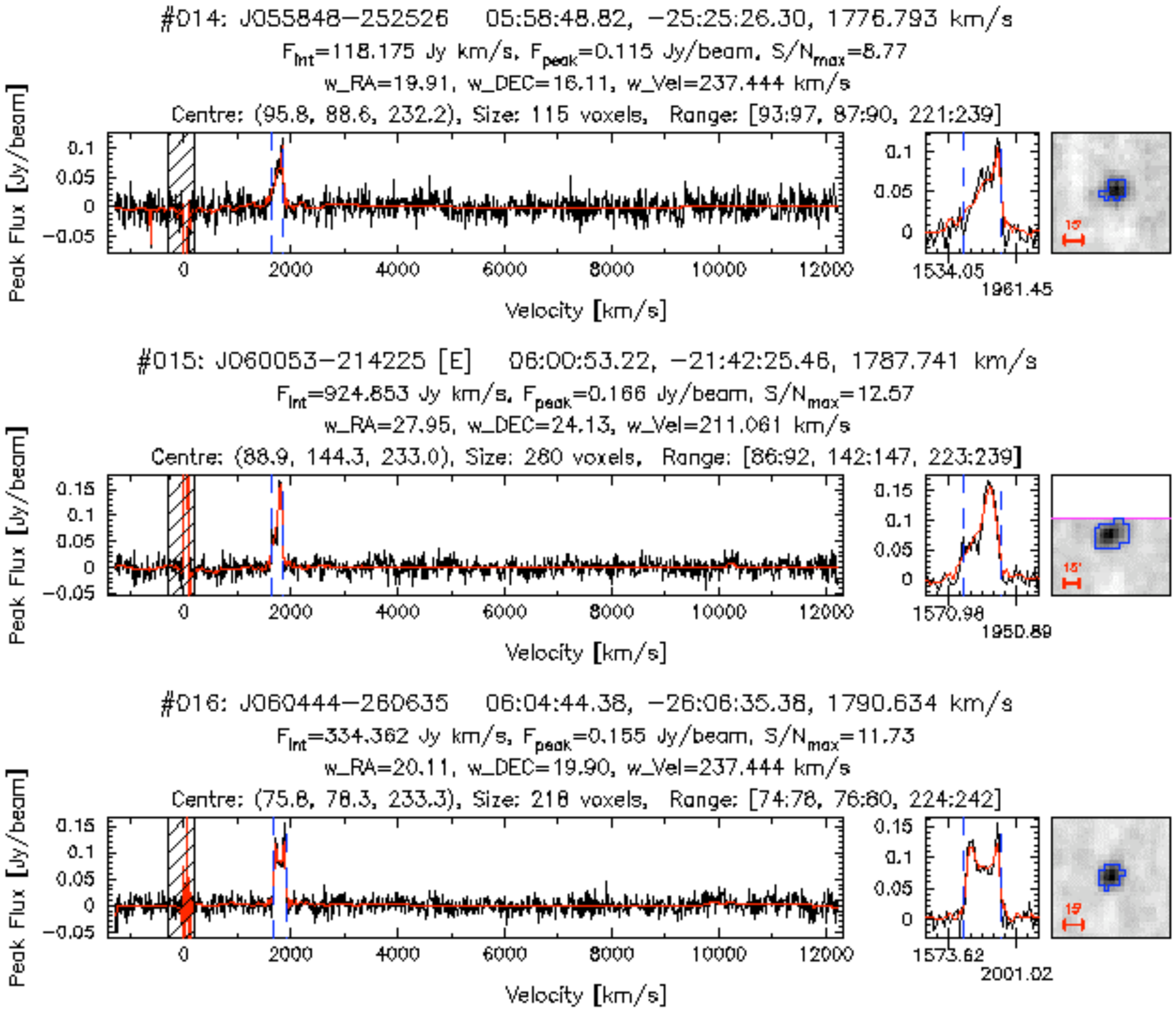
Wavelet reconstruction: example



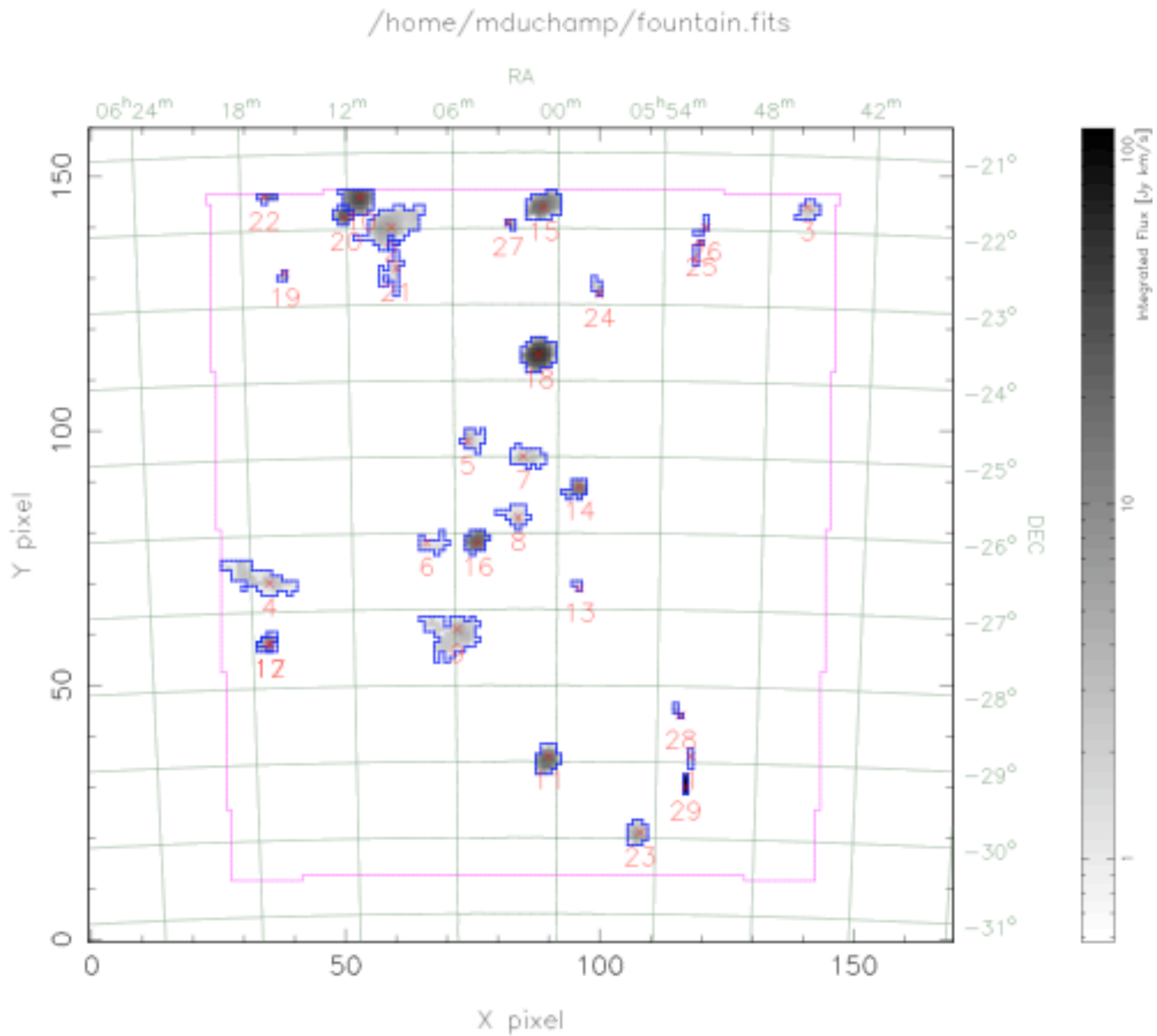
Outputs and feedback

- Duchamp provides a range of graphical outputs to help user understand the detected sources
  - Individual spectra and moment maps
  - Full-field moment map
  - Full-field detection map (which pixels were above detection threshold)
- Range of other output files
  - List of detections in text file
  - VOTable (XML) format
  - Karma annotation file (for overplotting kvis display)
  - Full list of intermediate detections, and pixel-by-pixel breakdown of detected sources
- Full reporting of all relevant input parameters and processing steps

Spectral plots



Moment maps



/home/mduchamp/fountain.fits
RA
DEC
X pixel
Y pixel
Integrated Flux [Jy km/s]

Outputs and feedback

- Duchamp provides a range of graphical outputs to help user understand the detected sources
  - Individual spectra and moment maps
  - Full-field moment map
  - Full-field detection map (which pixels were above detection threshold)
- Range of other output files
  - List of detections in text file
  - VOTable (XML) format
  - Karma annotation file (for overplotting kvis display)
  - Full list of intermediate detections, and pixel-by-pixel breakdown of detected sources
- Full reporting of all relevant input parameters and processing steps

Status and future development

- Duchamp is still under development
  - Last release: version 1.1.9 on March 26
- Duchamp API used in ASKAP software development
  - Some feedback of features and bug fixes into the standalone package
- I welcome any feedback, feature requests and bug reports
  - Make use of the Trac wiki

- Duchamp development will continue into the future
- Stand-alone package is separate to the ASKAP reduction pipeline and is (hopefully) more widely useful

Australia Telescope National Facility
Matthew Whiting
ASKAP Computing, Science Applications

Phone: 02 9372 4683
Email: matthew.whiting@csiro.au
Web: http://www.atnf.csiro.au/projects/askap/
http://www.atnf.csiro.au/people/Matthew.Whiting
http://www.atnf.csiro.au/people/Matthew.Whiting/Duchamp

www.csiro.au

Thank you