Jean Ballet, CEA Saclay

SLAC , 23 May 2005

All-sky source search

Aim: Look for a fast method to find sources over the whole sky

- ✓ Selection criteria
- ✓ Algorithms
- ✓ Iteration
- ✓ Simulations
- ✓ Energy information
- ✓ Variable and extended sources

All-sky source search. Selection

Aim: Define criteria for selecting a source search algorithm

Difficulty is that criteria are best applied on final output (after maximum likelihood). Because of computing requirements, can be done only after most parameters of a given method have been optimized (projection, pixel size, energy bands, …)

Criteria in approximate order of importance for source search itself:

1. Detection power (no new source will be found after that)
2. Resolving power (sources close to each other)
3. Flux and position estimate (to start later steps)
4. False detection rate (can be controlled later)
5. Computing time (of entire pipeline. Many false detections bear on this)

All-sky source search. Algorithms

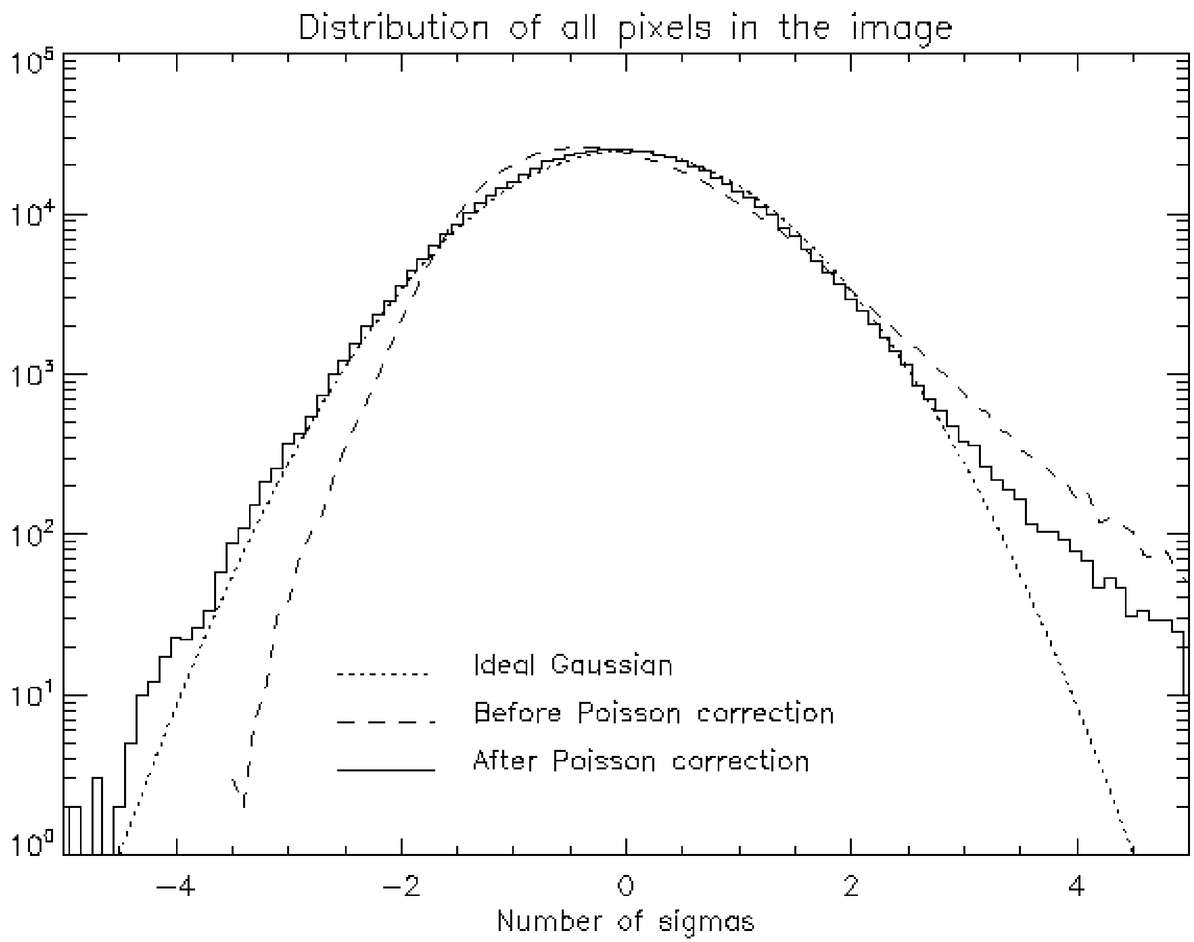
1. Wavelet algorithm developed at Perugia (next talk)
2. Wavelet algorithm developed at Saclay (MR_filter by J.L. Starck, presented in September 2004). On DC1, was quite comparable to 1.
3. Optimal filter PSF/(1+BKG) in Fourier space (myself , presented in September 2004). On DC1, did not find as many sources as 1 or 2 but returns source significance.
4. Multichromatic wavelet method developed by S. Robinson and T. Burnett. Tries using the PSF at the energy of each photon. On DC1, did not find as many sources as 1 or 2 but still interesting for localization.
5. Aperture photometry tried by T. Stephens as a reference. On DC1, did not find as many sources as the other methods but useful exercise.
6. Voronoi algorithm proposed by J. Scargle (later today). Was not ready for DC1.

Algorithms: Recent developments

- ✓ Few since DC1. At Saclay, we have concentrated on building a pipeline prototype.
- ✓ Optimal filter method now provides an (approximate) all-sky significance map in the Poisson regime.
- ✓ This can be used to combine energy bands.

Variance is easy to get at each point.

Signal / dispersion at each point is converted into an effective gaussian significance depending on the local density of counts and the PSF shape (same principle as is used in wavelet methods for each coefficient).



All-sky source search. Iteration

- ✓ Methods based on filtering have trouble finding weak sources in the wings of bright ones (because they generate a negative ring around sources).
- ✓ The solution is to iterate. After the bright sources have been detected, they can be modelled and entered into the diffuse emission.

- ✓ Issue is whether this may be done internally (independently of the other energy bands) or must be done globally (via likelihood).
- ✓ Internal solution is simpler
- ✓ Global solution is probably more accurate (because the bright source position can be accurately determined at high energy)

All-sky source search. Simulations

The DC1 data has provided a rich resource for testing the algorithms. The DC2 data will improve on it. They have however several limitations :

1. It is only a particular trial. It is useful to be able to simulate the same model many times, to test robustness.
2. The point spread function that the sources follow is representative, but not well reproduced by the Science Analysis Environment (for DC1 at least).
3. It does not allow to study in detail particular aspects, like sources close to each other at varying flux ratio.

We don't actually need a full instrument simulation for testing source detection algorithms. It is more important to control carefully the conditions of the simulation.

- ✓ Points 1 and 2 are best addressed by gtobssim, with a smoothly varying PSF as a function of energy and off-axis angle (like the one used for the 30 days simulation of the recent Science Tools checkout).
- ✓ Point 3 is best addressed by local simulations (with known PSF).

All-sky source search. Energy bands

- ✓ Background limits source detectability by its Poisson fluctuations, even if the diffuse emission model is accurate. Approximate signal to noise (for weak sources) is $S / \sqrt{B}$, where S and B are taken over 1 PSF.
- ✓ All sources do not have the same spectrum. Soft sources will be better seen above the diffuse emission at low energy, hard sources at high energy.
- ✓ PSF varies enormously from low energy (> 4° below 100 MeV) to high energy (< 0.2° above 1 GeV). This means that low energy photons from bright sources act as background to nearby fainter sources.

- ✓ Splitting into several energy bands is better than summing everything. Example for optimal filter method (just from source lists): 105 sources in 0.1-1 GeV band, 109 in 0.1-0.316 (51) + 0.316-1 (89).
- ✓ 4 energy bands (32 MeV / 100 MeV / 316 MeV / 1 GeV / 10 GeV) was all right for DC1. For longer integration time (like 1 year) shift to higher energy (confusion at low energy). Cannot split indefinitely (more degrees of freedom).
- ✓ Pixel size should be adapted to PSF in each band.

Putting energy bands together

- ✓ Simplest solution is to run algorithm over each energy band separately and merge source lists. This is better than using a single band, but not very powerful.
- ✓ A better solution is to add likelihood images (before applying threshold).

Can be done also on a full significance map, as presented before for the optimal filter method. If S is the significance (in sigma units), and i an index for energy bands, then $\Sigma_i S_i^2$ is expected to follow a χ^2 distribution with N (number of energy bands) degrees of freedom. Excesses can be detected on the combination directly.

The broad PSF at low energy adds a complication. A true source can induce false detections in its vicinity (via its low energy wings).

All-sky source search. Special cases

Variable sources (blazars mostly). This covers two different things:

- ✓ Identify as variable sources which have been detected over the entire time period.
- ✓ Detect variable sources which have been missed over the entire time period (because of dilution). Can be done by repeating the source search over shorter time intervals (like one week), or by a specific algorithm (like looking for variability systematically in sky ‘pixels’).

Extended sources (external galaxies, supernova remnants, interstellar structures). This covers two different things:

- ✓ Identify as extended sources which have been detected by the point-source algorithm. Can be done by comparing source shape with PSF convolved with a Gaussian of variable width.
- ✓ Detect extended sources which have been missed by the point-source algorithm. Can be done by wavelet algorithms, or simply by looking for excesses in residual photon map (sources and diffuse emission subtracted).

Source detection studies

We have several viable options !

Several open points:

1. How to deal best with the energy information ? Will be addressed at Saclay.
2. Is cartesian geometry all right (paving the sky with large pieces) ? Is it worth working in spherical geometry ?
3. How best to detect variable sources ?
4. How best to detect extended sources ?