

HiPS Catalogue Generation



Why HiPS catalogues ?

- **Explore** large tables (SDSS DR9, 2MASS-PSC, USNO-B1.0, GAIA DR1, etc) by **panning and zooming**.
- Clients (Aladin, Aladin Lite, etc) automatically load the minimum required data : less risk of misuse by overloading the client.

How ?

- HEALPix based multi-scale tessellation of the sky :
 - zoom level → HEALPix norder
 - FOV around a position → load the few covered tiles
- Most relevant data appear first :
 - Brightest (or more cited, etc) sources loaded at large scales.
 - Fainter (or less cited, etc) sources loaded when zooming.
- Pre-defined upper limit on the maximum number of sources per tile.

Standard HTTP API (defined in IVOA documents)

- Get the metadata of a table in VOTable (example with 2MASS):
<http://baseURL/2MASS/metadata.xml>
- Get the allsky files at level 1 and 2 :
[http://baseURL/2MASS/Norder\[12\]/allsky.tsv](http://baseURL/2MASS/Norder[12]/allsky.tsv)
- Get the tile number N from level 3 to max level :
[http://baseURL/2MASS/Noerder\[3-12\]/Dir\[1000*\(N%1000\)\]/NpixN.tsv](http://baseURL/2MASS/Noerder[3-12]/Dir[1000*(N%1000)]/NpixN.tsv)

Design choices at CDS, beyond the basic algorithm

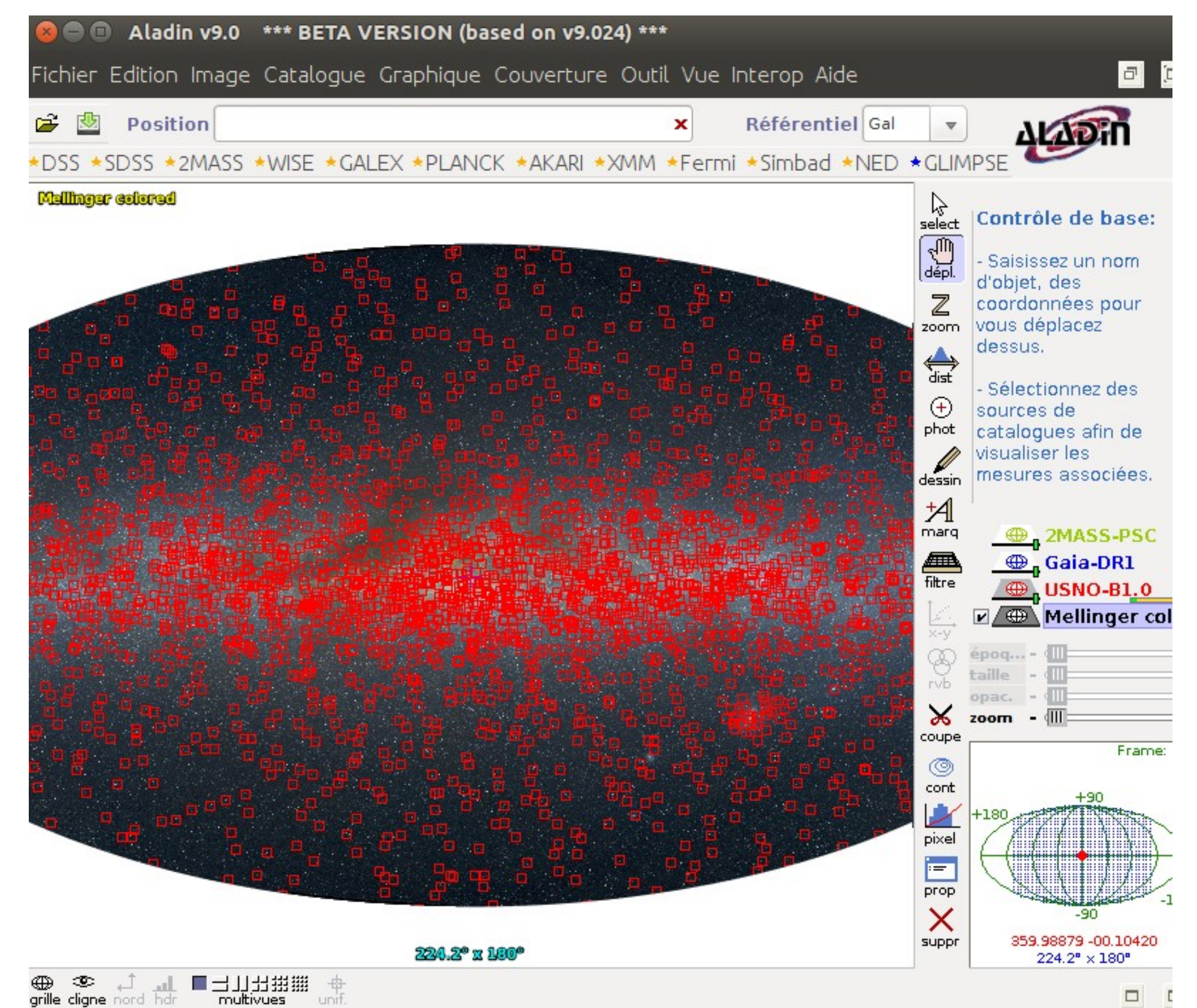
- Allsky views (norder 1 and 2) : show brightest sources BUT with a spatial distribution mimicking the global density of sources
 - at a given norder, heterogeneous limit magnitude (tile dependent)
- From level 3 : fixed *min* and *max* number of sources per tile
 - #srcs (N) in a tile of given norder depends on :
 - lowest and highest #srcs per HEALPix cell at the given norder
 - #srcs in the HEALPix cell wich corresponds to the tile
 - fraction of the tile covered by the catalogue
 - put in the tile the N brightest (or more cited, etc) sources in the corresponding HEALPix cell

HiPS catalogue generation

- Tool publicly available at CDS (jar file). Please visit:
<http://aladin.u-strasbg.fr/hips/HipsCat.gml>
- *Two tools in one* sharing the same core algorithm:
 - One for general user input (CSV, FITS, VOTable): works on TSV files
 - Generates a directory/file hierarchy: each tile is in a TSV file
 - Directly loadable in Aladin or through an HTTP sever
 - One dedicated to CDS large catalogues: works on sorted binary files
 - Generates one indexed binary file per norder
 - Accessible through an Apache Tomcat servlet
- Allows creation of a HiPS catalogue in one command line

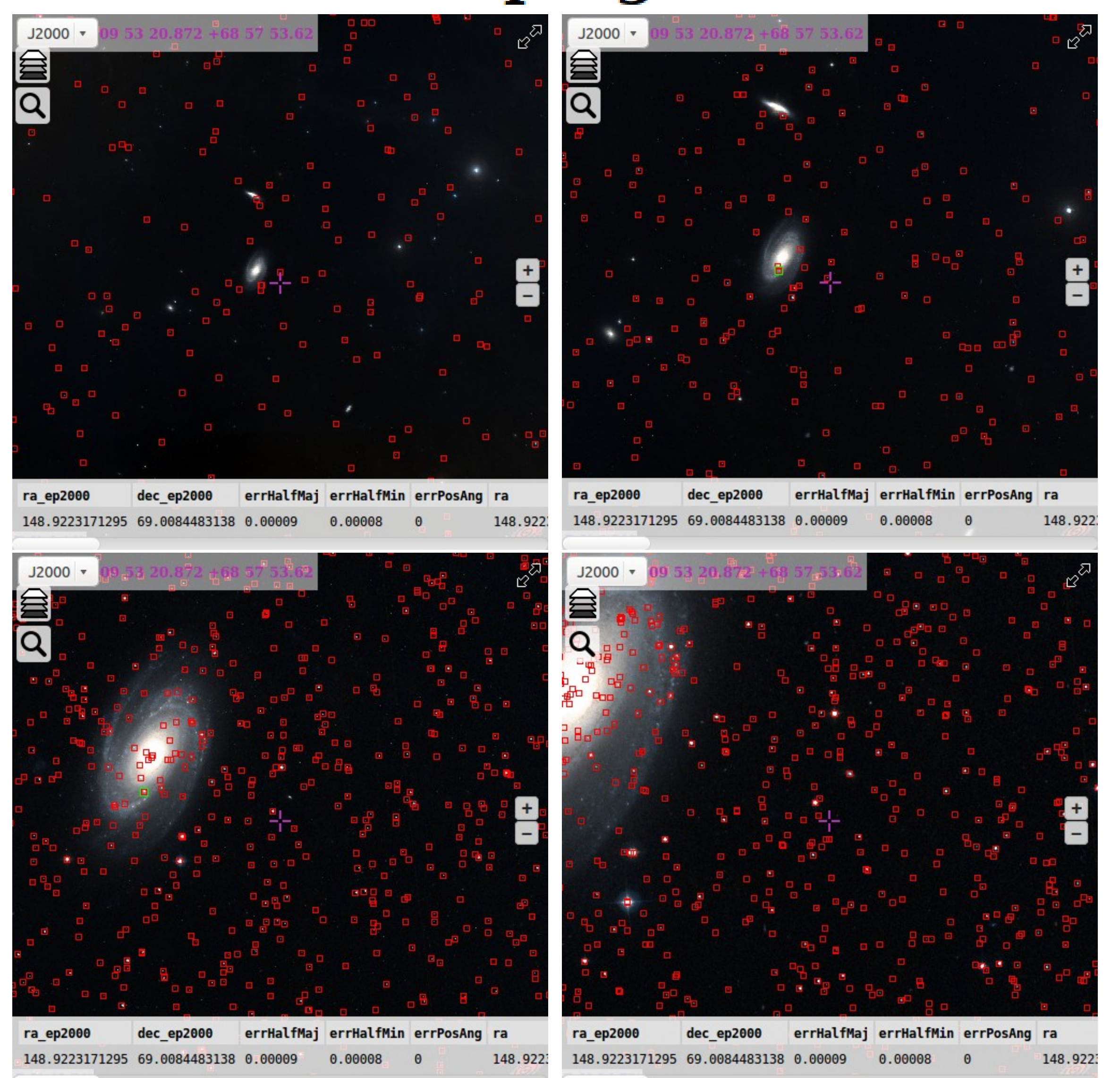
Algorithm used by the tool

- Read data and convert in multiple TSV files that are put in a simple directory/file hierarchy that we call "naive hierarchical hash tree" (unnecessary steps for large catalogues binary files).
- For each cell of norder 3 (768 cells):
 - Compute the number N of sources to keep at level 1, 2 and 3.
 - Read all sources of the pixel and perform a Heapselect (see Numerical Recipes) to keep the N brightest from the data stream
 - Keep the limit value and recursively select the sources of the 4 sub-tiles
 - Remark: after a given level, the recursive process ensures that the data stay in the HDD cache



Visualisation in **Aladin** of the USNO-B1.0 HiPS catalogue on top of the Mellinger HiPS image

The Gaia-DR1 progressive table



Extract of the web page (using **Aladin Lite**) located at the Gaia DR1 HiPS catalogue base URL:
<http://axel.u-strasbg.fr/HIPSCatService/Gaia-DR1>

Your HiPS in one command line:

```
java -Xmx5012m \
-jar cds.hips.cat.jar \
-in mycat.csv -out myhips \
-cat CATNAME \
-ra RA -dec DEC -IM 4 \
-score Bmag -desc
```

In this example, the key used to sort sources is the Bmag column.



François-Xavier Pineau
(francois-xavier.pineau@astro.unistra.fr)
Thomas Boch
Pierre Fernique

