Archive, discover and match compact and diffuse objects on the galactic plane in the VIALACTEA project

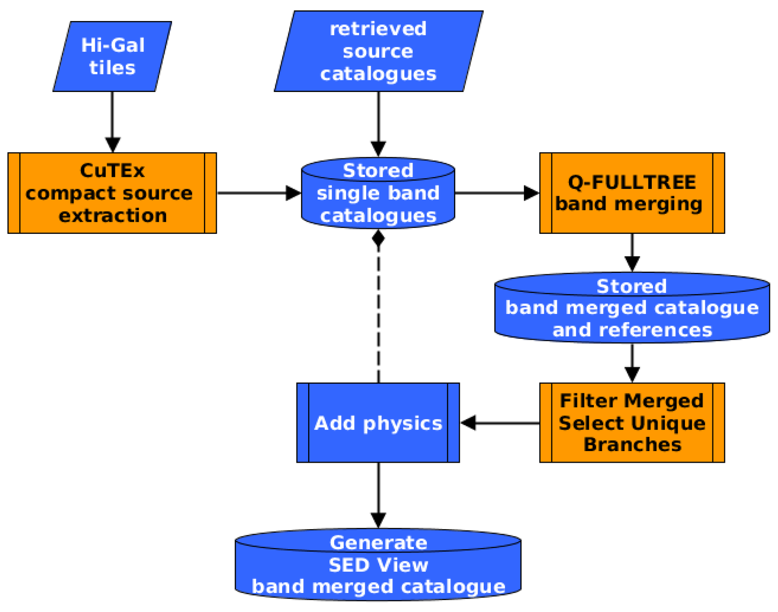
R. Smareglia[a], M. Molinaro[a], R. Butora[a], E. Schisano[b], D. Carey[c], D. Elia[b], S. Molinari[b]

a – INAF – Osservatorio Astronomico di Trieste, Trieste, Italy

b – INAF – IAPS, Tor Vergata (Roma), Italy

c – University of Leeds, Leeds, UK

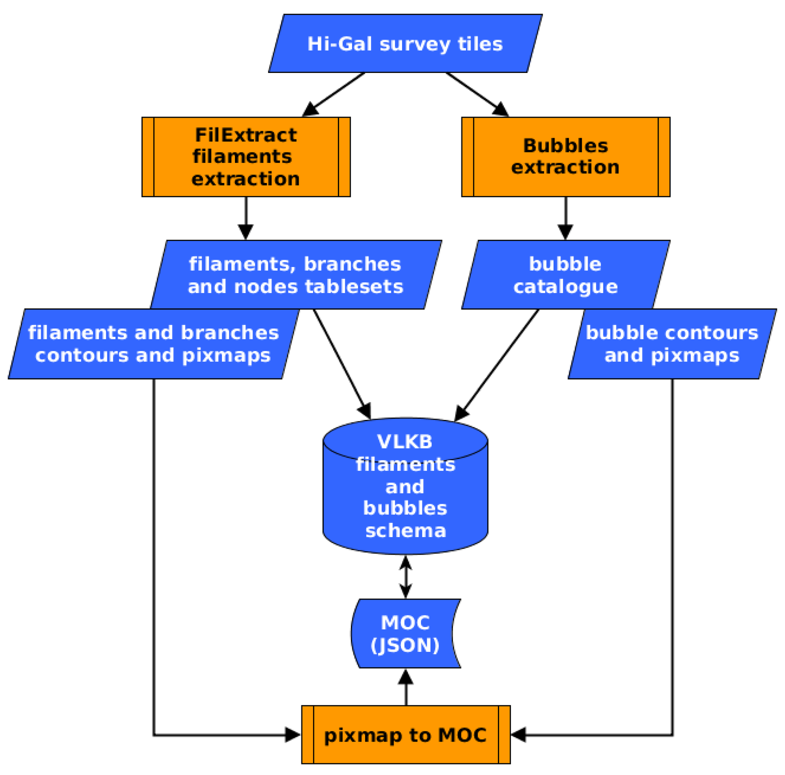
ABSTRACT Among the scientific goals of the EU-FP7 VIALACTEA project, there are some that involve tasks like archiving, discovering and accessing catalogues of infrared and sub-mm objects that are not simple point-like sources on the sky, but compact or diffuse structures. These structures may have more general shapes (bubble-like or filament-like the two categories identified by the project). Also, alongside catalogue search and access, a service to match compact sources that are located inside a more diffuse structure like a bubble or a filament, is needed. In this contribution we describe the scenario and catalogues related to the project scientific tasks and the solutions put in place to grant the VIALACTEA (and astrophysical community) users interaction with these compact and diffuse object catalogue data. Shape and contour archiving techniques inside a catalogue as well as catalogue cross-matching using tessellation solutions are here described.



ingestion workflow
for the compact sources catalogues

COMPACT SOURCES There are two types of compact source catalogues within the VIALACTEA Knowledge Base (VLKB), distinguished upon how they have been retrieved. There's one set of single band catalogues that were generated using the CuTEx tool run on the HI-Gal survey tiles. This set of 5 catalogues were generated as a product of the project itself. Alongside these 5 catalogues a set of other 5 external public catalogues were added: BGPS, ATLASGAL, MIPSGAL, WISE and MSX. These ones were retrieved from various public services and include additional 5 sets of counterparts to the Hi-Gal compact sources and 14 additional fluxes at different wavelengths (since WISE and MSX provide more fluxes for each of their recorded sources). All of these catalogues are ingested in the VLKB and formed the basis to build a multi-band combined catalogue (or multi-mission galactic plane band merged catalogue) using data mining techniques (Q-FullTree package developed within VIALACTEA) to associate the sources at the various observational wavelengths from the single catalogues. The resulting catalogue contains, in this way, SEDs for its various records, including possible branches were multiple sources at one wavelength can be associated to a parent source at another wavelength.

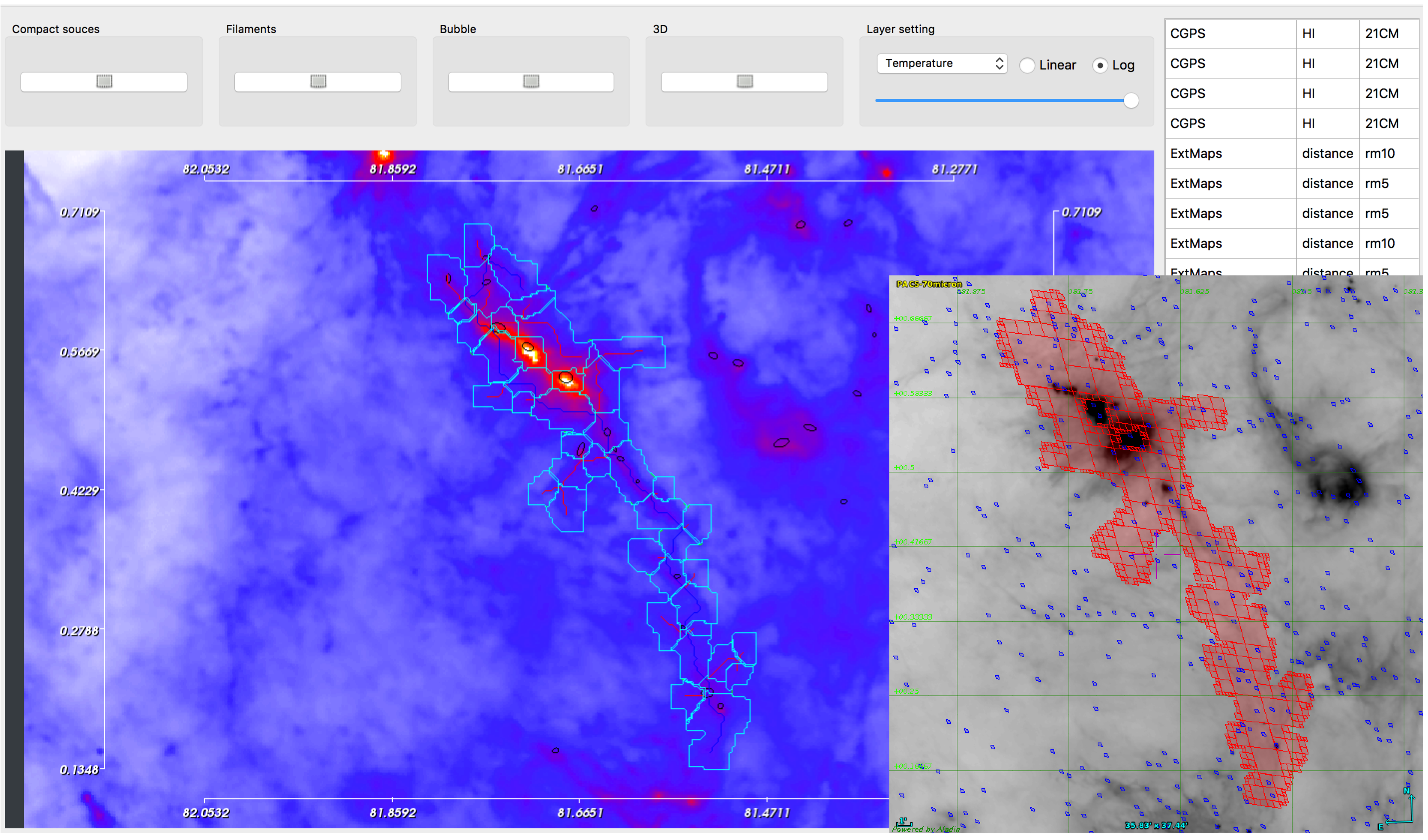
DIFFUSE STRUCTURES Also the diffuse structures stored in the VLKB are of two different types: bubble like and filament like objects. Both the catalogues for these structures were generated as deliverables of the project using extraction tool themselves part of the project and run over the HI-Gal survey tiles. While compact sources can be described with simple geometrical parameters (an oriented ellipse) diffuse structures require more information to help describe their shape and complexity. Bubbles can have these information stored in a quite simple tabular format, their geometry synthesized by a circular and elliptic region but their actual shape defined by a polygonal contour (and included area). Filaments are more complex structures, whose area can be subdivided in branches and connected at nodes. For this reason a simple tabular format is less useful and thus morphological information on contours and areas for the full filaments or the component branches where derived and ingested separately in the VLKB.



ingestion workflow
for the diffuse structures catalogue

DISCOVERY & ACCESS INTERFACES On top of the above described astrophysical resources an IVOA TAP interface has been deployed to let the user have the capability to search within the full content of the compact and diffuse objects catalogues. However, given the complex morphological characteristics of the diffuse objects (mainly) a different search interface was also required to dynamically match compact sources against diffuse structures and vice versa. Since doing this analytically was not a feasible choice (e.g. because the MySQL backend is not aware of arrays to store polygonal shapes or more complex data for filtering) a discrete solution based upon sky tessellation has been setup.

SOURCE MATCHING USING TESSELLATION A tessellation solution based on the IVOA MOC recommendation (direct derivation of the HEALPix tessellation) has been taken into account for two main reasons: pre-existing simple libraries to work on to produce both the information to be added to the database content and the HTTP search interface needed, compliance with a standard format for serializing sky tessels that is clean and compact (MOC is indeed an HEALPix nested solution fixed to live in ICRS sky coordinates frame, with the addition of including the HEALPix order and npix value in the same integer element). MOCs have been used in both the compact sources part and the diffuse structure part of the VLKB. For the catalogues of compact sources MOC files have been generated for the full catalogue (to be used against MOC coverage of diffuse objects) and single npix values have been stored alongside coordinates of the single compact objects for source identification after the overlap step has been performed. For the diffuse structures single MOC tessellation has been computed for each bubble and filament and stored in their JSON serialization in a database field to be used for overlap against the catalogue full MOCs. Since the number of filaments and bubbles (up to a few tens of thousands) would require a too long loop to overlap each single structure against an input sky position, an helper table was designed to store lower order npix values for pre-filtering of objects to be overlapped.



STATUS AND FUTURE PERSPECTIVE Here on the left the VLKB data (one sample filament with compact source overlap) as available through the Visual Analytics tool developed by the team of INAF – OACt [1]. This tool consumes the VLKB content through its TAP, cube search and access [2] and the MOC based matching interface. The main picture shows a filament and its branches through contours, plus ellipses for the compact sources that fall in the area. The bottom-right box shows the tessellation representation of the same filament (sources are not exactly comparable). Currently the matching interface is quite raw (with CSV output only for identified overlapping objects), plans are to refine it. Further improvement is required also by the DB tessellation content, to be more adherent to source positional errors and observational resolution.

[1] – see poster P2.4 (Becciani & al.)

[2] – see poster P4.4 (Butora & al.)

Contact email: molinaro@oats.inaf.it - This work has benefited from EU-FP7 grant agreement 607380 VIALACTEA