The ChiVO Library: Advanced Computational Methods for Astronomy

MAURICIO ARAYA[1], MAURICIO SOLAR [1], DIEGO MARDONES [2], MARCELO MENDOZA [1], CAMILO VALENZUELA[1], TEODORO HOCHFARBER[1], MARTÍN VILLANUEVA[1], MARCELO JARA [1], AXEL SIMONSEN[1]

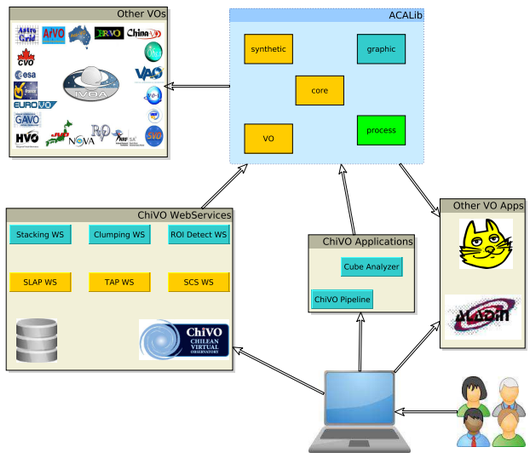
[1]UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA, VALPARAÍSO, CHILE

[2]UNIVERSIDAD DE CHILE, SANTIAGO, CHILE

ABSTRACT

The main objective of the Advanced Computational Astronomy Library (ACALib) is to ensemble a coherent software package with the research on computational methods for astronomy performed by the first phase of the Chilean Virtual Observatory between years 2013 and 2015. During this period, researchers and students developed functional prototypes, implementing state of the art computational methods and proposing new algorithms and techniques. This research was mainly focused on spectroscopic data cubes, as they strongly require computational methods to reduce, visualize and infer astrophysical quantities from them, and because most of the techniques are directly applicable either to images or to spectra.

ACALIB



Color notation:

- **yellow**: element already implemented and integrated
- **green**: element already prototyped, but not integrated
- **blue/cyan**: element under development

Key points:

- Library that implements several algorithms and tools for analyzing spectroscopic data cubes, images and spectra.
- Coherent framework for developing novel *webservices* for ChiVO
- But also has an API for stand-alone applications (`python` interface)
- Strongly grounded in `astropy` and `numpy`
- Reuse some algorithms from `scipy`, `scikit-learn` and `astroML`.
- Algorithms automatically search into VO services
- Compatible with SAMP (e.g., connects with Topcat and Aladin)

Modules:

- **core**: main classes to manipulate astronomical data
- **vo**: workspace abstraction and VO comm interfaces
- **synthetic**: generates synthetic spectroscopic cubes
- **process**: algorithms developed so far by ChiVO
- **graphic**: widgets and tools for 3D visualization

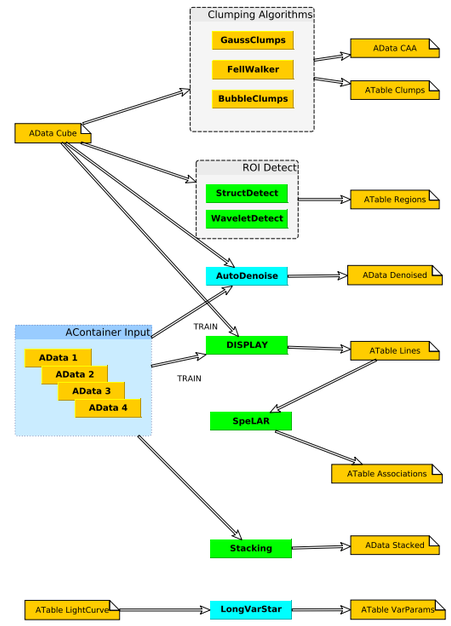
PROCESS MODULE

Algorithms (OA = Original Algorithms)

- **Clumping**: detect clumps in an AData. Similar to CUPID package [1].
  - GaussClumps: Mixture of Gaussians fitting (Stutzki & Gusten, 1990, ApJ)
  - FellWalker: Agregation of hill-climbing paths (Berry, 2015, A&C)
  - BubbleClumps: Clustering of small Gaussians (in preparation) (OA)
- **ROI Detection**: index multi-resolution regions of interest in an AData.
  - StructDetect: Morphological processing (Mendoza et al., 2015, A&C) [2] (OA)
  - WaveletDetect: Multi-scale detection in Wavelets space (Gregorio et al., 2015, SPIE) (OA)
- DISPLAY: Learn dictionaries for line detection (Riveros, 2015, Thesis) (OA)
- SpeLAR: Compute association rules for spectral lines (Miranda, 2015, Thesis) (OA)
- Stacking: Automatic stacking of images (Jara, 2015, Thesis) (OA)

Under development

- AutoDenoise: Denoising of images using deep autoencoders
- LongVarStar: Detection of long-term variable stars using ML



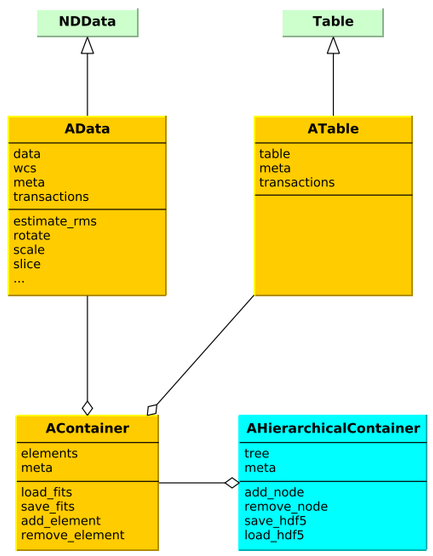
CORE MODULE

Members

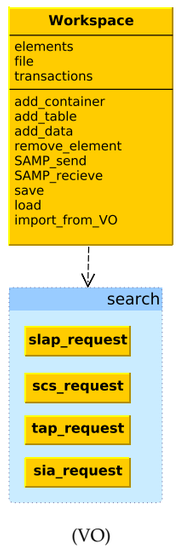
- **AData** is an extension of astropy NDData
  - Vectorized masked arrays: fast computations with missing values
  - Metadata and WCS support: from astropy
  - Self-operations: rotate, scale, slice, stack, statistics, search, etc.
  - Transactions: deferred and online WCS/Meta consistency
- **ATable** is an extension of astropy Table
  - Interface: simpler than its parent
  - Metadata support: from astropy
  - Self-operations: statistics
- **AContainer** is composed by a list of AData and ATable objects
  - Namespace: it works as a namespace for astronomical data
  - Self-operations: Load and save FITS files

Planned features

- Cython integration for less memory consuming operations
- MPI integration for multiprocessing environments
- Hierarchical containers for supporting HDF5 format

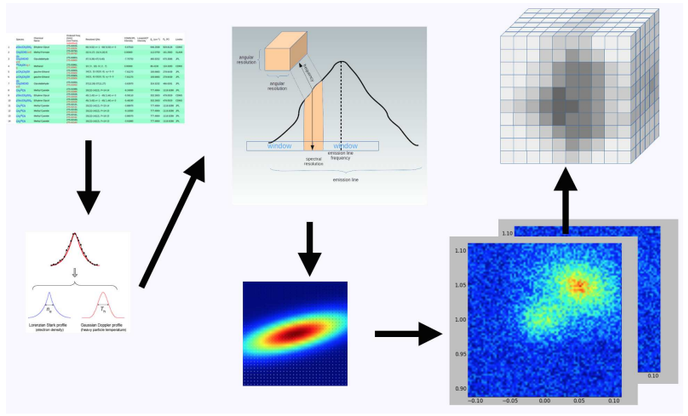


THE OTHER MODULES



(VO)

- **vo** : the workspace can host elements of the core, send them through SAMP and obtain/export data from/to the VO.
- **synthetic** : module is an integrated version of ASYDO [3], it can generate data for testing, training and validation.
- **graphic**: this is a key module that we are implementing to validate the results of each algorithm.



(SYNTHETIC)

Acknowledgements

This work was partially financed by CONICYT through the FONDEF D11I1060 project and the ICHAA 79130008 project.

References

[1] Berry et al. *CUPID: Clump Identification and Analysis Package*. ADASS 2013

[2] Mendoza et al. *Indexing data cubes for content-based searches in radio astronomy*. Astronomy & Computing 2015 (to appear)

[3] Araya et al. *Exorcising the Ghost in the Machine: Synthetic Spectral Data Cubes for Assessing Big Data Algorithms*. ADASS 2014

Chilean Virtual Observatory

Contact e-mail: maray@inf.utfsm.cl / For more information about Chilean Virtual Observatory, please visit our web site http://www.chivo.cl