

Youpi, a Web-based Astronomical Image Processing Pipeline

M. Monnerville & G. Sémah

Terapix, Institut d'Astrophysique de Paris, CNRS, University of Pierre et Marie Curie, UMR 7095, Paris, F-75014, France

Abstract. Youpi⁴ stands for “YOUpi is your processing Pipeline”. It is a portable, easy to use web application providing high level functionalities to perform data reduction on scientific FITS images. It is built on top of open source processing tools that are released to the community by Terapix, in order to organize your data on a computer cluster, to manage your processing jobs in real time and to facilitate teamwork by allowing fine-grain sharing of results and data. On the server side, Youpi is written in the *Python* programming language and uses the *Django*⁵ web framework. On the client side, Ajax techniques are used along with the *Prototype* and *script.aculo.us* Javascript libraries⁶.

1. Introduction

Youpi is a new generic and versatile web-based pipeline that is suitable to handle both large surveys, like the CFHTLS (Goranova et al 2009), and smaller, less-well organized sets of observations.

Youpi is designed to run and supervise data management by keeping track of all data locations and parameters involved during the processing life cycle. It starts by ingesting FITS image information into the database and manages the processing all the way to the end products (image stacks, catalogues).

2. Handling FITS Data Ingestion

Ingesting images is the main entry point for Youpi users. Thanks to a dedicated selection widget, the user chooses one or several data paths holding FITS images (available either locally or on the network by using network shares such as NFS), which prompts Youpi to read the headers and ingest the relevant information in the database. Youpi reports back to the user by e-mail as soon as ingestion is completed. Ingestion can be carried out once for all, so that every ingested image is available for processing at any time. Youpi provides user-friendly tools to query images and data (see section 4.).

⁴<http://youpi.terapix.fr/>

⁵<http://www.djangoproject.com/>

⁶<http://www.prototypejs.org/>, <http://script.aculo.us/>

3. Data Processing

Youpi is a high-level tool acting as a wrapper for freely available, low-level software packages: it currently supports the *QualityFITS* image quality assessment software, the *SCAMP* astrometric and photometric calibration tool (Bertin 2006), the *SWarp* image resampling and stacking software (Bertin et al. 2002) and the *SExtractor* source extraction and catalogue production tool (Bertin&Arnouts 1996).

All these processings modules are internally available as plug-ins. Some of them may be disabled to perform any specific or fragmented processing sequence requested by the user. The plug-in architecture is modular and versatile. A detailed guide on how to write custom plug-ins and extend processing capabilities is available on the Youpi website.

Youpi manages data access rights and protections by implementing a Unix-like file permission system. Each user has a personal password-protected account, where data sets and results can have user-defined permissions for the owner, a group of owners and the rest of the world. Permissions can be changed to share images, data products or even pipeline configuration files.

4. Data Management

The various widgets and tools of the Youpi user interface have been painstakingly designed to simplify the selection and organisation of data, as well as the management and set up of configuration and result files.

The Data Path Selector The path selector may be invoked from any processing plug-in to select and save data paths in real time (e.g. to provide a path to external `.ahead` files for calibrating a bunch of images using the *SCAMP* plug-in). The path selector widget can browse the user’s local – and remote, if network mounts are used – filesystems, and save those directories for later reuse.

The Image Selector The image selector (Fig. 1) sets up image lists based on previously ingested images. Multiple search criteria can be combined together, including observing run IDs, filters/channels, grades (Youpi provides an integrated interface for image quality assessment), . . .

Once images are selected, basic editing commands can be used to save the current selection. Saved selections can also be deleted or merged with others. In order to increase the image selector’s flexibility, options have been added to upload a plain text file describing the selection (file names and optional checksums) or to batch-import all text files within a directory. These two features allow Youpi to create and save selections in a semi-automated way, rather than manually. This comes in particularly handy when processing surveys with a very large number of images as all image selections may be prepared beforehand.

Organizing Data With Tags A tag is a non-hierarchical keyword that can be assigned to a piece of information. Youpi allows tagging ingested images in order to add some kind of metadata useful for describing items. Tagged images



Figure 1. A selection of 1450 images is built by retrieving all images from the *CFHTLS-T0006-W3_Scamp* saved selection.

can later be searched by tag name using the image selector. Applying tags on a selection of images is straightforward (Fig. 2): all it takes is to create a new tag – or use an existing one – and drag-and-drop it over a dedicated drop zone to mark or unmark the current selection. The number of tags which may be applied to an image is practically unlimited.

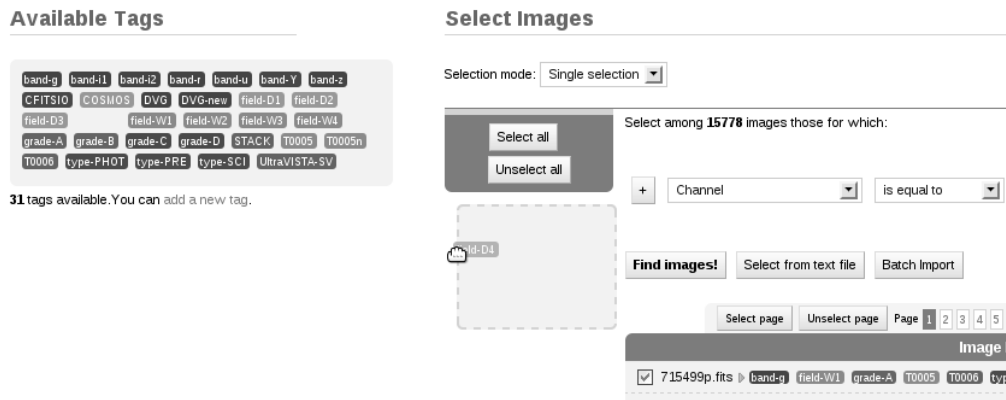


Figure 2. Dragging the *field-D4* tag over a selection of images.

The Processing Cart Before running any processing with Youpi, all input data, configuration parameters and processing tasks are first bundled together in one processing cart item. The processing cart acts like shopping carts on commercial websites. This is a convenient mechanism for sharing processings among users – processing items can be saved for later use and loaded back to the processing cart at anytime by others – or for setting run-time options before submitting jobs on the cluster.

5. Cluster Computing With Condor

Youpi relies on the *Condor*⁷ software framework for distributing jobs on the local machine, a cluster of computers, or a compute grid. Ingestions as well as processing cart items are submitted as Condor jobs. Youpi handles all input parameters and generates a *Condor submission file* – a plain text file specifying job requirements, environments and complete command lines – ready to be executed on the system.

Depending on the requirements, it may be useful to target a specific subset of cluster nodes. Youpi has a page dedicated to Condor requirements setup, from which custom policies or selections may be defined. Custom policies are *dynamic rules*: several criteria can be defined using regular expressions; they will only be computed just before the job is submitted. In contrast, custom selections are made of *static nodes* selected among the available cluster nodes displayed on screen.

Jobs submitted to the Condor cluster can be monitored directly from your browser on the Youpi *Active Monitoring* page. Important information such as the job description, remote host, running time, owner and current status are displayed and refreshed in real time, without reloading the page.

6. Summary

Youpi is a versatile pipeline which can be used on a single host – a desktop computer or even a laptop – or on larger installations such as computer clusters. The software is actively maintained at Terapix⁸ and published under an open source license. Youpi was first released to the general community on September 2009. The current version offers native support for data from the CFHT-MegaCam, CFHT-WIRCam and ESO-VISTA-VIRCAM instruments, and we plan to support more data sources in the future.

Acknowledgments. We are grateful to E. Bertin, Y. Goranova, P. Hudelot, F. Magnard, H. McCracken, Y. Mellier and M. Schultheis for active support and testing. We would also like to thank Y. Kakazu for her Japanese translation of the Youpi abstract.

References

- Bertin E., 2006, ASPC, 351, 112
Bertin E., Arnouts S., 1996, A&AS 317,393
Bertin et al. 2002, ASP Conference Series, 281, 228
Goranova, Y. Hudelot, P. Magnard, F., McCracken, H., Mellier, Y., Monnerville, M., Schultheis, M., Sémah, G., Cuillandre, J.-C., Aussel, H., 2009, *The CFHTLS T0006 Release*, <http://terapix.iap.fr/cpl1/T0006/T0006-doc.pdf>

⁷<http://www.cs.wisc.edu/condor/>

⁸<http://terapix.iap.fr/>