Filabres, a new pipeline for the automatic data reduction of CAFOS direct imaging

Nicolás Cardiel López[1], Enrique Alejandro Galcerán García[2], Jaime Alonso Hernández[1], Sergio Pascual Ramírez[1], José Manuel Alacid Polo[3], Enrique Solano Márquez[3], Miriam Cortés Contreras[3], Mª Teresa Ceballos Merino[4], Jesús Aceituno Castro[5], Santos Pedraz Marcos[5]

[1]Departamento de Física de la Tierra y Astrofísica, Universidad Complutense de Madrid, Spain / [2]GMV Innovating Solutions, Spain / [3]Centro de Astrobiología (CSIC-INTA), Spain / [4]Instituto de Física de Cantabria, Spain / [5]Observatorio de Calar Alto, Spain

Filabres is a new Python pipeline created with the idea of performing the automatic reduction of direct images obtained with the instrument CAFOS, placed at the 2.2 m telescope of the Calar Alto Observatory. The goal is to provide useful reduced images through the Calar Alto Archive hosted by the Spanish Virtual Observatory. The typical workflow with **Filabres** consists of the following steps: (1) Image classification (bias, flat-imaging, arc, science-imaging, etc.); (2) Reduction of calibration images (bias, flat-imaging) and generation of combined master calibrations as a function of the modified Julian Date; (3) Basic reduction of individual science images, making use of the corresponding master calibrations (closest in time to the observation of the science target). The main reduction steps considered here are: bias subtraction, flatfielding of the images, and astrometric calibration (performed with the help of additional software tools provided by Astrometry.net and by AstrOmatic.net). The behaviour of the data reduction is easily defined through a set of reduction rules set in a configuration YAML file, specifically built for the considered instrument and observation mode. Note, however, that the software has been designed to allow the future inclusion of additional observing modes and instruments. The software is publicly available through GitHub at https://github.com/nicocardiel/filabres, and its documentation in https://filabres.readthedocs.io/.

The need of an automatic data pipeline

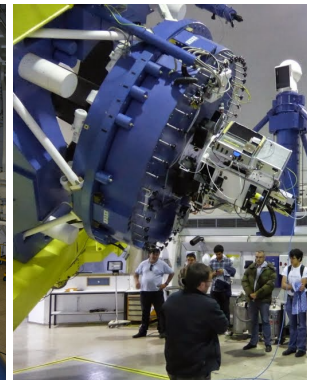
An automatic data pipeline is required to reduce large data sets, where the interactive approach is typically unaffordable in terms of human resources and/or time.

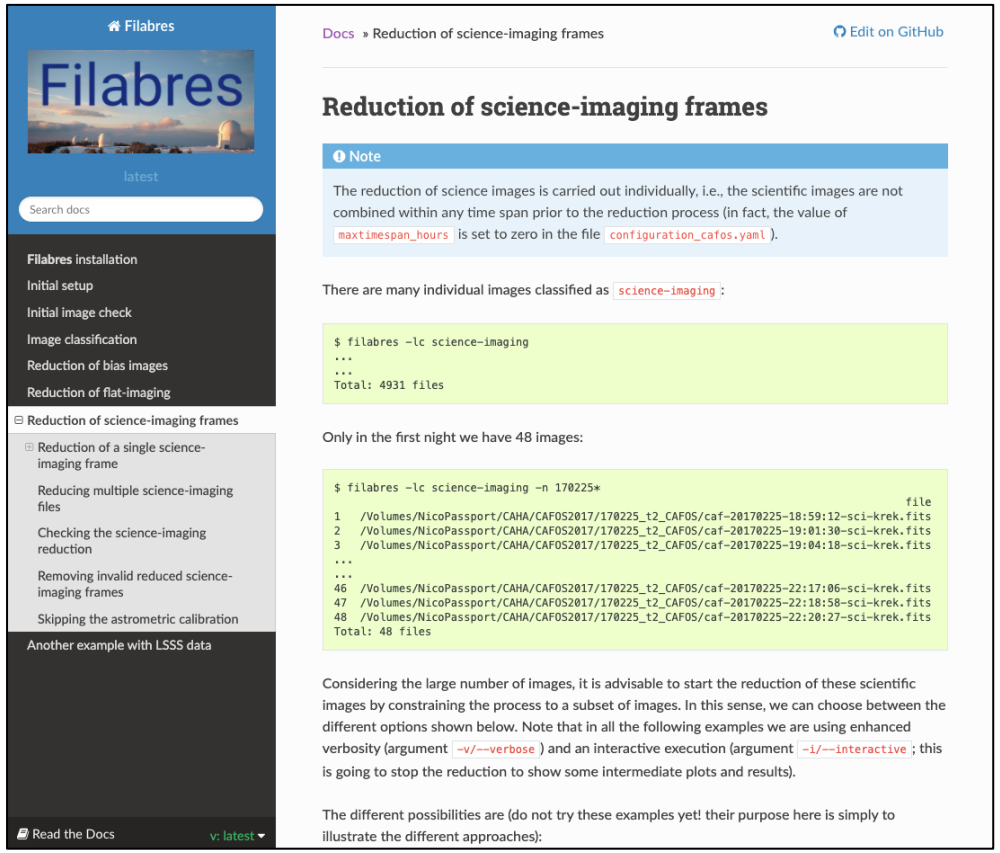
Important benefits of an automatic data pipeline are:

- Availability of calibrations that for any reason were not acquired in a given observation program, but which are available in other programs.
- Homogeneous data treatment and reduction: this facilitates the reuse of archived data.
- Allows the study of the temporal evolution of the instrument performance.

Filabres has born as an initial effort to incorporate such automatic data reduction for direct images obtained with CAFOS, the Calar Alto Faint Object Spectrograph, located at the 2.2 m telescope of the Calar Alto Observatory. The goal is to provide those reduced images through the Calar Alto Archive hosted by the Spanish Virtual Observatory.

Filabres source code is available through GitHub at https://github.com/nicocardiel/filabres, and its documentation in https://filabres.readthedocs.io/.





How filabres works

The typical workflow with **Filabres** consists of the following steps:

1. Image classification: bias, flat-imaging, arc, science-imaging, etc., are classified within each observing night. The derived classification is stored in an independent JSON file for each night that works as a database for the subsequent work.

This is a critical step. It must be robust (i.e. able to handle inconsistencies in header keywords)

2. Reduction of calibration images (bias, flat-imaging) and generation of combined master calibrations as a function of the modified Julian Date. This information is stored in additional JSON files that allows the quick retrieval of the required calibrations for the reduction of the science images.

A **signature** is assigned to each reduced image, according to CCD geometry, filter, etc.

3. Basic reduction of individual science images, making use of the corresponding master calibrations. The main reduction steps considered here are:
   - bias subtraction
   - flatfielding of the images
   - astrometric calibration, performed with the help of additional software tools provided by Astrometry.net and by AstrOmatic.net

The **calibrations** with the requested signature and closest in time to the observation of the science target are selected. If any calibration is not available within the considered night, a valid calibration from neighbouring observing nights is selected.

Rules for the automatic classification of the images

The automatic image classification is carried out following the rules provided in the **instrument configuration file** (written in YAML format). This file defines the rules using a hierarchical strategy, based on header keyword values and statistical measurements (predefined image quantiles) performed on the image data themselves.

Although the software has been developed to reduce science images obtained with the instrument CAFOS in imaging mode, the flexibility of the devised **instrument configuration file** should allow the reduction of images obtained with different observation modes and other instruments.

configuration_cafos.yaml

```
instname: cafos
version: 2.3
requirements:
  INSTRUME: 'CAFOS 2.2'
masterkeywords:
  - NAXIS      # number of data axes
  - NAXIS1     # length of data axis 1
  - NAXIS2     # length of data axis 2
  - OBJECT     # Target description
  ...
  ...
imagetypes:
  bias:
    ...
    ...
  flat-imaging:
    ...
    ...
  flat-imaging-wollaston:
    ...
    ...
  flat-spectroscopy:
    ...
    ...
  arc:
    ...
    ...
  science-imaging:
    ...
    ...
  science-imaging-wollaston:
    ...
    ...
  science-spectroscopy:
    ...
    ...
```

Mandatory keyword(s) in order to classify the image as having been obtained with the expected instrument

List of keywords that will be used in the subsequent classification and analysis of the images

Definition of the different image categories

```
imagetype:
  bias:
    executable: True
    classification: calibration
    requirements:
      IMAGETYP: 'bias'
    requirementx:
      EXPTIME: 0.0
      QUANT975.LT.: 1000
    signature:
      - CCDNAME
      - NAXIS1
      - NAXIS2
      - DATASEC
      - CCDBINX
      - CCDBINY
    maxtimespan_hours: 1
```

Mandatory keyword(s) to set the image type

Additional requirements (non-conforming images are flagged as suspicious)

Signature definition: allows the proper assignment of master calibrations to science images

```
science-imaging:
  executable: True
  classification: science
  maxfieldview_arcmin: 16
  requirements:
    IMAGETYP: 'science'
    INSGRID: 'GRISM-11'
    INSPOFPI.NE.: 'Wollaston'
  requirementx:
    EXPTIME.GT.: 0
    QUANT975.LT.: 60000
  signature:
    - CCDNAME
    - NAXIS1
    - NAXIS2
    - DATASEC
    - CCDBINX
    - CCDBINY
    - INSGRID
    - INSFLID
    - INSPOFPI
    - INSPOROT
    - RA
    - DEC
  maxtimespan_hours: 0
  basicreduction: True
```

Additional configuration files

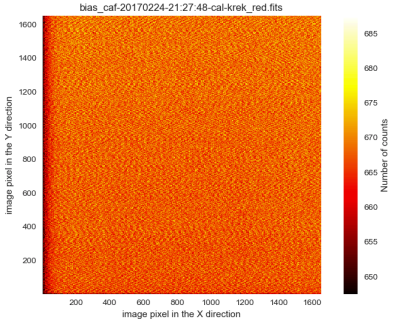
- `setup_filabres.yaml:` instrument, data location, naming of Filabres and AstrOmatic configuration files
- `image_header_corrections.yaml:` rules to replace wrong keywords in specific images
- `ignored_images.yaml:` list of images to be ignored
- `forced_classifications.yaml:` manual classification of problematic images

Reduction results

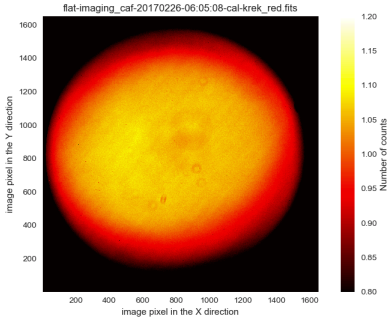
Master calibrations are classified by image signature and sorted by modified Julian Date. Diagnostic diagrams, showing relevant statistical information, can easily be created in order to identify and remove suspicious images.

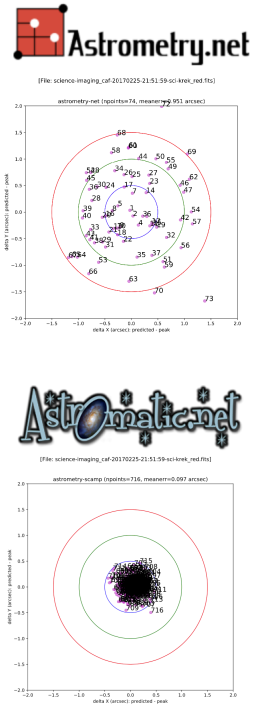
Science images are reduced using the appropriate master calibrations.

master bias

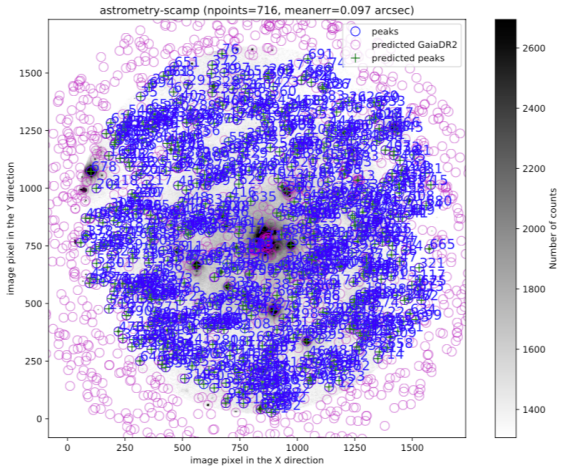


master flat

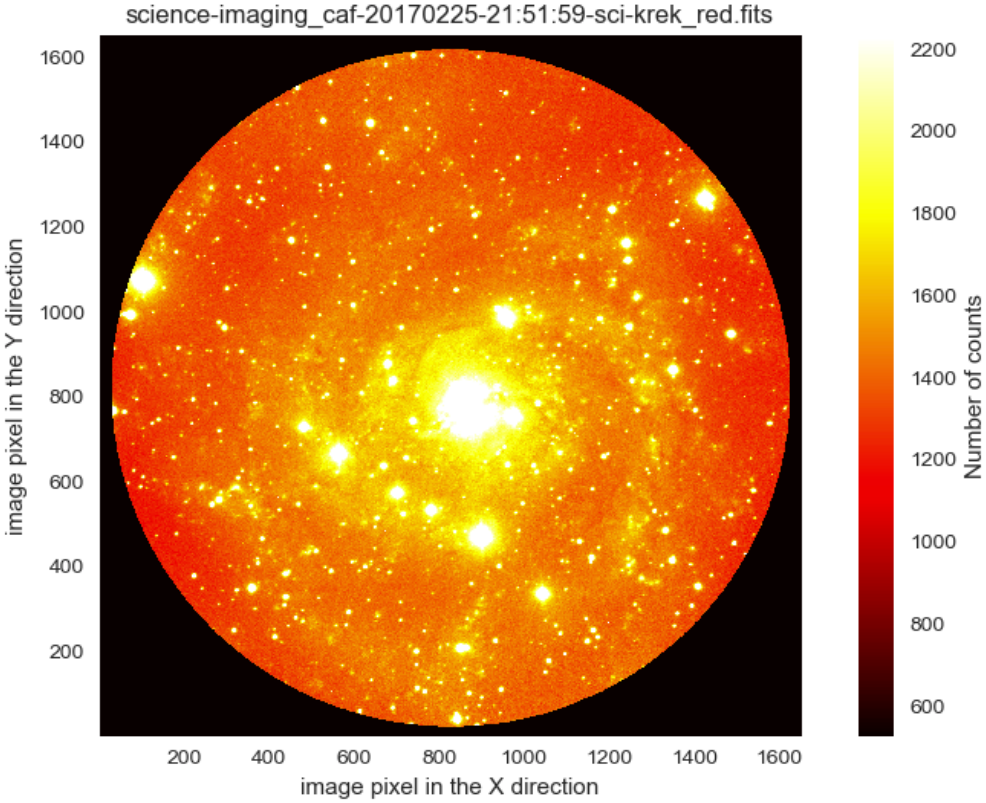




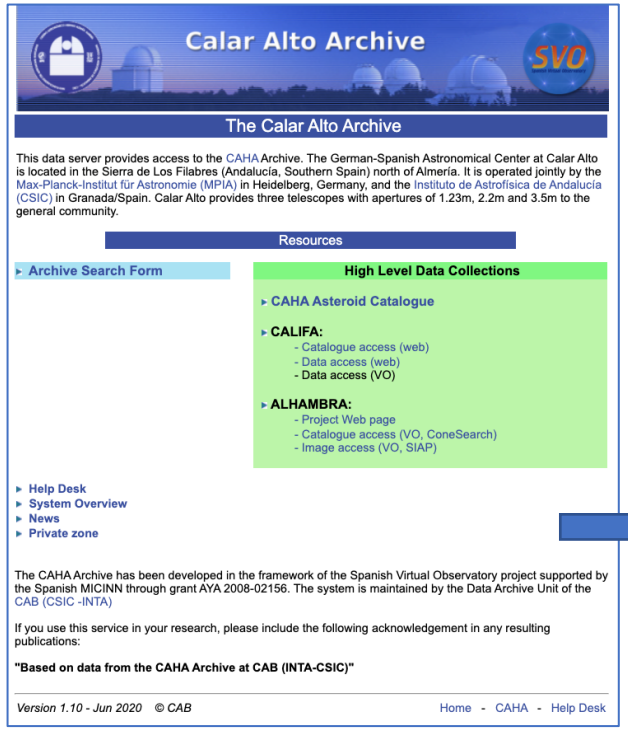
The initial astrometric calibration performed with Astrometry.net tools (downloading Gaia DR2 sources), is refined using the AstrOmatic.net utilities (SExtractor & SCAMP).

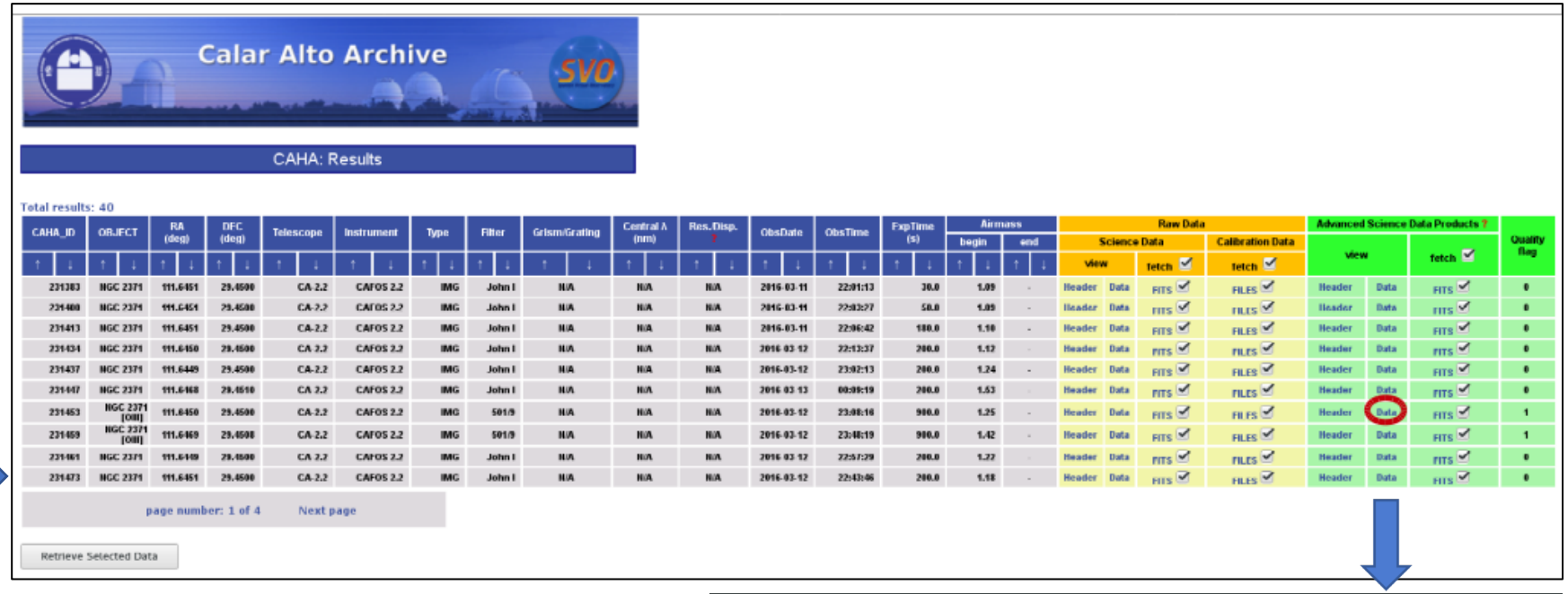


reduced science image



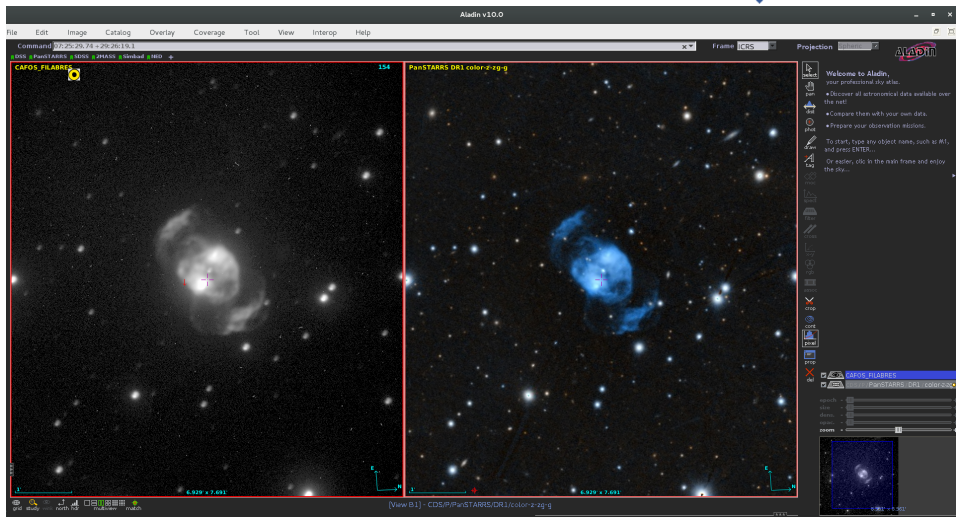
Impact and Prospects for the future





This is an on-going project: reduced CAFOS images corresponding to **2016** and **2017** are already available through the Calar Alto Archive hosted by the Spanish Virtual Observatory.

Since the archive is integrated within the Virtual Observatory, it is possible to quickly visualize the reduced data using, for example, the widely-used VO tool Aladin. This allows the immediate comparison with available data from well-known surveys and databases. In this example we compare the CAFOS observation of the planetary nebula NGC 2371 (left panel) with the PanSTARRS DR1 image (right panel).



This work is funded by the Spanish Programa Estatal de I+D+i Orientada a los Retos de la Sociedad under grant RTI2018-096188-B-I00, which is partly funded by the European Regional Development Fund (ERDF).