

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

**Nicolás Cardiel López¹, Enrique Alejandro Galcerán García²,
Jaime Alonso Hernández¹, Sergio Pascual Ramírez¹,
José Manuel Alacid Polo³, Enrique Solano Márquez³, Miriam Cortés Contreras³,
M^a Teresa Ceballos Merino⁴, Jesús Aceituno Castro⁵, Santos Pedraz Marcos⁵**

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

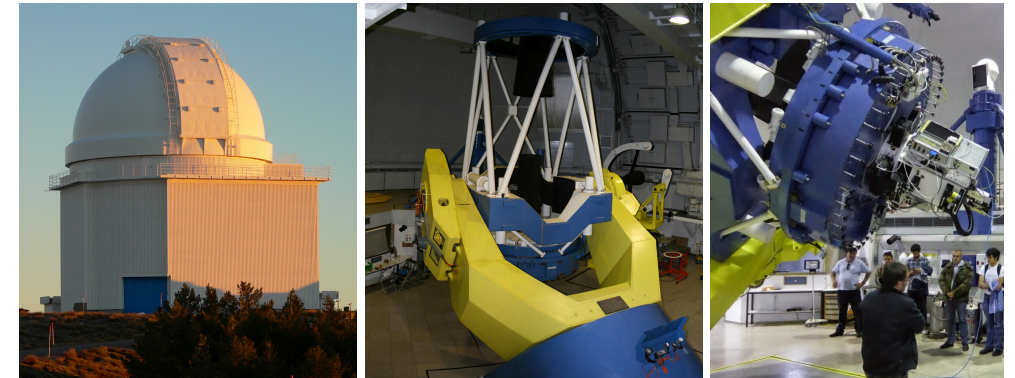
An automatic data pipeline is required to reduce large data sets, where the interactive approach is typically unaffordable in terms of human resources and/or time.

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/>.



The screenshot shows the GitHub repository for Filabres. The main heading is "Reduction of science-imaging frames". A note states: "The reduction of science images is carried out individually, i.e., the scientific images are not combined within any time span prior to the reduction process (in fact, the value of `maxlinespan_hours` is set to zero in the file `configuration_cafos.yaml`).". Below this, there is a terminal output showing the command `$ filabres -lc science-imaging` resulting in 4931 files. Another terminal output shows the command `$ filabres -lc science-imaging -n 170225*` resulting in 48 files. The page also includes a section for "Only in the first night we have 48 images:" with a list of file paths and their timestamps.

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'
    INSPFP1.NE.: 'Wollaston'
  requirementx:
    EXPTIME.GT.: 0
    QUANT975.LT.: 60000
  signature:
    - CCDNAME
    - NAXIS1
    - NAXIS2
    - DATASEC
    - CCDBINX
    - CCDBINY
    - INSGRID
    - INSFLID
    - INSPFP1
    - INSPOR0T
    - RA
    - DEC
    maxtimespan_hours: 0
    basicreduction: True
  
```

Additional configuration files

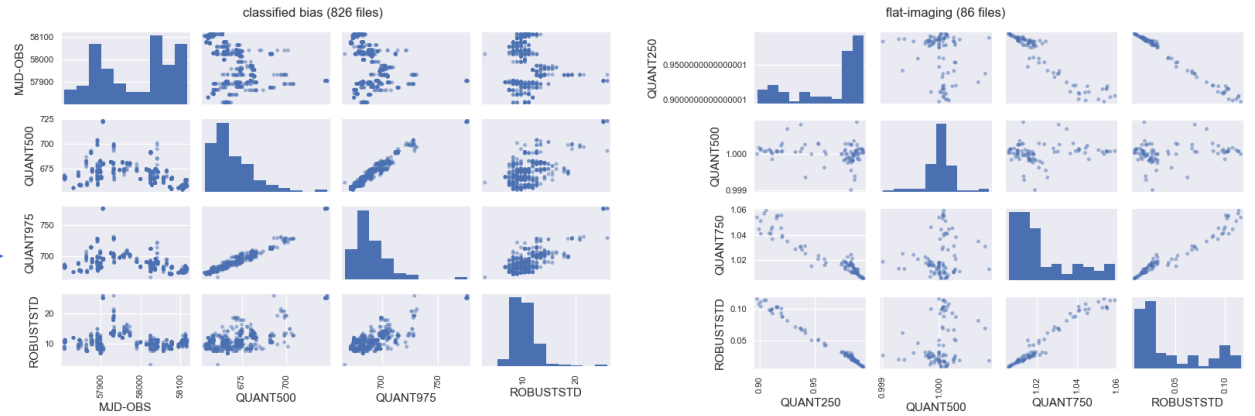
- `setup_filabres.yaml`: instrument, data location, naming of Filabres and AstrOmatic configuration files
- `image_header_corrections.yaml`: rules to replace wrong keywords in specific images
- `ignored_images.yaml`: list of images to be ignored
- `forced_classifications.yaml`: manual classification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



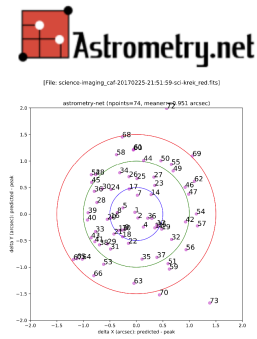
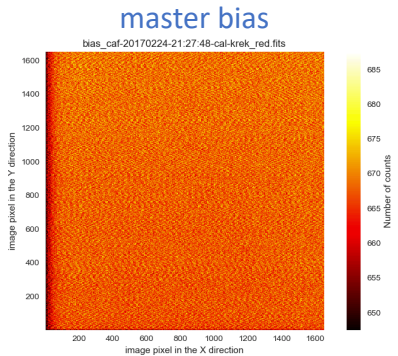
Reduction results

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.

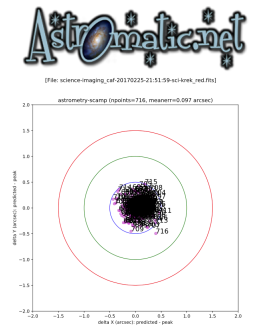
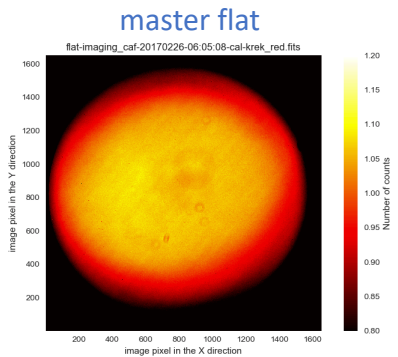
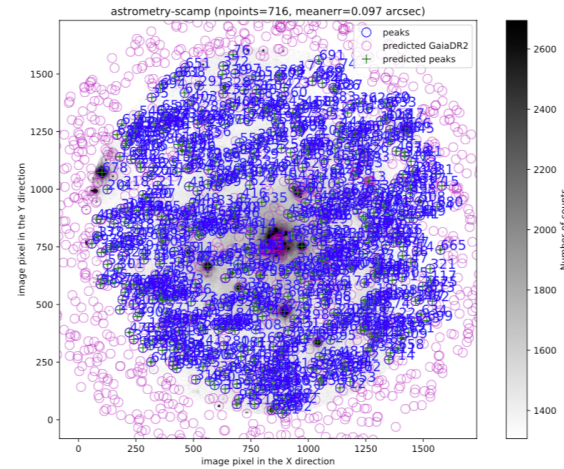


Science images are reduced using the appropriate master calibrations.

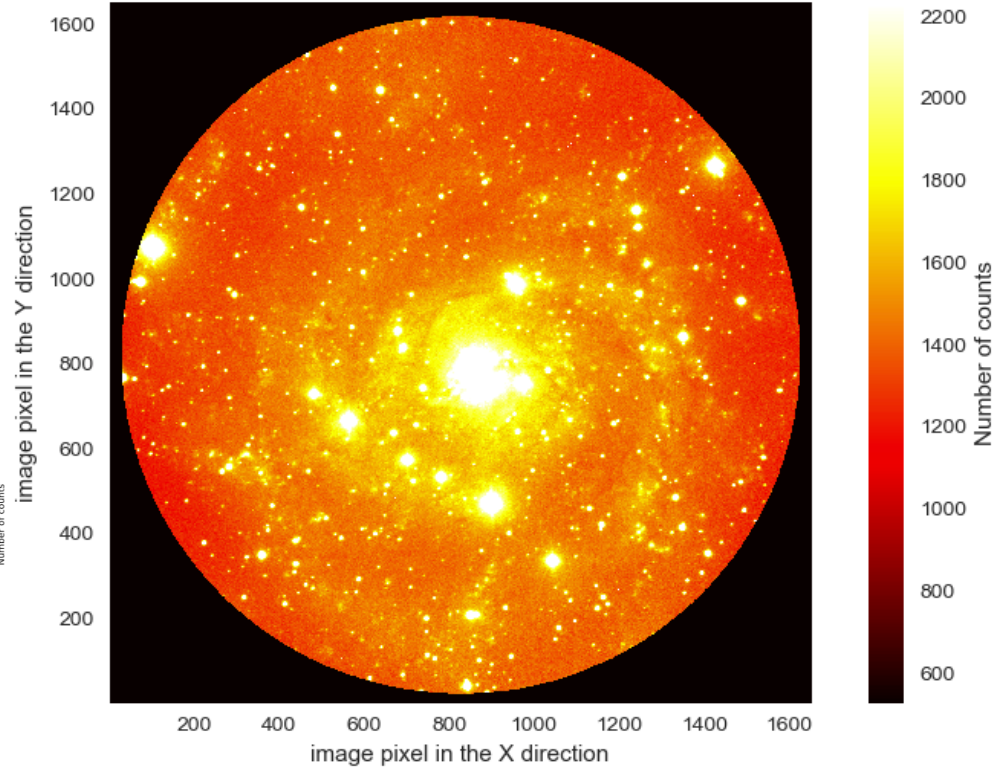


The initial astrometric calibration performed with Astrometry.net tools (downloading Gaia DR2 sources), is refined using the AstrOmatic.net utilities (SExtractor & SCAMP).

[File: science-imaging_caf-20170225-21:51:59-sci-krek_red.fits]



reduced science image
science-imaging_caf-20170225-21:51:59-sci-krek_red.fits



Impact and Prospects for the future

Calar Alto Archive SVO

The Calar Alto Archive

This data server provides access to the CAHA Archive. The German-Spanish Astronomical Center at Calar Alto is located in the Sierra de Los Filabres (Andalucía, Southern Spain) north of Almería. It is operated jointly by the Max-Planck-Institut für Astronomie (MPIA) in Heidelberg, Germany, and the Instituto de Astrofísica de Andalucía (CSIC) in Granada/Spain. Calar Alto provides three telescopes with apertures of 1.23m, 2.2m and 3.5m to the general community.

Resources

- Archive Search Form
- High Level Data Collections**
 - CAHA Asteroid Catalogue
 - CALIFA:**
 - Catalogue access (web)
 - Data access (web)
 - Data access (VO)
 - ALHAMBRA:**
 - Project Web page
 - Catalogue access (VO, ConeSearch)
 - Image access (VO, SIAP)
- Help Desk
- System Overview
- News
- Private zone

The CAHA Archive has been developed in the framework of the Spanish Virtual Observatory project supported by the Spanish MICINN through grant AYA 2008-02156. The system is maintained by the Data Archive Unit of the CAB (CSIC-INTA)

If you use this service in your research, please include the following acknowledgement in any resulting publications:

"Based on data from the CAHA Archive at CAB (INTA-CSIC)"

Version 1.10 - Jun 2020 © CAB Home - CAHA - Help Desk

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.

Calar Alto Archive SVO

CAHA: Results

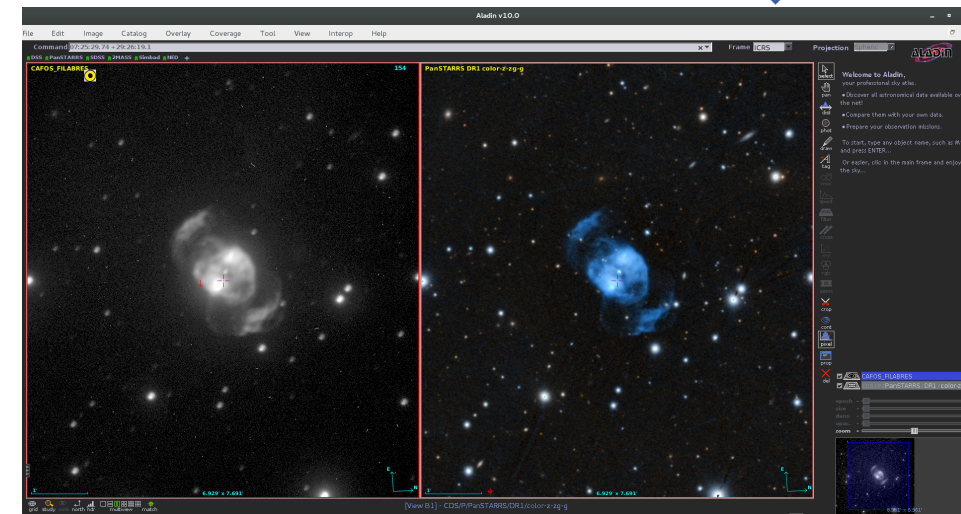
Total results: 40

CAHA_ID	OBJECT	RA (deg)	DEC (deg)	Telescope	Instrument	Type	Filter	Grism/Grating	Central A (nm)	Res. Disp.	ObsDate	ObsTime	ExpTime (s)	Airmass		Raw Data		Advanced Science Data Products		Quality flag			
														begin	end	Science Data		Calibration Data			view		
														view	Tetch	view	Tetch	view	Tetch				
231383	HGC 2371	111.6451	29.4590	CA-2.2	CAFOS 2.2	IMG	John I	N/A	N/A	N/A	2016-03-11	22:01:13	30.0	1.09	-	Header	Data	FITS	FILES	Header	Data	FITS	0
231400	HGC 2371	111.6451	29.4590	CA-2.2	CAFOS 2.2	IMG	John I	N/A	N/A	N/A	2016-03-11	22:03:27	50.0	1.09	-	Header	Data	FITS	FILES	Header	Data	FITS	0
231413	HGC 2371	111.6451	29.4590	CA-2.2	CAFOS 2.2	IMG	John I	N/A	N/A	N/A	2016-03-11	22:06:42	100.0	1.10	-	Header	Data	FITS	FILES	Header	Data	FITS	0
231431	HGC 2371	111.6450	29.4590	CA-2.2	CAFOS 2.2	IMG	John I	N/A	N/A	N/A	2016-03-12	22:13:37	200.0	1.12	-	Header	Data	FITS	FILES	Header	Data	FITS	0
231437	HGC 2371	111.6449	29.4590	CA-2.2	CAFOS 2.2	IMG	John I	N/A	N/A	N/A	2016-03-12	23:02:13	200.0	1.24	-	Header	Data	FITS	FILES	Header	Data	FITS	0
231447	HGC 2371	111.6450	29.4590	CA-2.2	CAFOS 2.2	IMG	John I	N/A	N/A	N/A	2016-03-12	00:08:19	200.0	1.53	-	Header	Data	FITS	FILES	Header	Data	FITS	0
231453	HGC 2371 [OH]	111.6450	29.4590	CA-2.2	CAFOS 2.2	IMG	501/5	N/A	N/A	N/A	2016-03-12	23:00:16	900.0	1.25	-	Header	Data	FITS	FILES	Header	Data	FITS	1
231459	HGC 2371 [OH]	111.6450	29.4590	CA-2.2	CAFOS 2.2	IMG	501/9	N/A	N/A	N/A	2016-03-12	23:10:19	900.0	1.42	-	Header	Data	FITS	FILES	Header	Data	FITS	1
231461	HGC 2371	111.6449	29.4590	CA-2.2	CAFOS 2.2	IMG	John I	N/A	N/A	N/A	2016-03-12	22:57:29	200.0	1.22	-	Header	Data	FITS	FILES	Header	Data	FITS	0
231473	HGC 2371	111.6451	29.4590	CA-2.2	CAFOS 2.2	IMG	John I	N/A	N/A	N/A	2016-03-12	22:43:06	200.0	1.18	-	Header	Data	FITS	FILES	Header	Data	FITS	0

page number: 1 of 4 [Next page](#)

Retrieve Selected Data

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).