AutoFib2+WYFFOS (AF2)

Data Reduction Pipeline

Version 3.0 (Stable version)

September 2015

Isaac Newton Group of Telescopes

www.ing.iac.es

Pipeline installation

The pipeline directory contains the IDL-based AF2 pipeline procedures and some third-party IDL routines which are modified or customized to be used by the pipeline .

In addition to pipeline directory, user must install the **ASTRON** (version 61: Aug 2013 or later) and **Coyote** IDL libraries, in addition to built-in and **default IDL libraries** that are usually in IDL installation directory.

1- Add all this libraries to your IDL_PATH

NOTE: We strongly recommend you to set your IDL_PATH in your startup file (.cshrc, .bashrc, .bash_profile, etc).

csh users:

```
setenv IDL_PATH
"+/home/user/pipeline_directory:${IDL_PATH}:+/home/user/astron/:+/home/user/coyote/"
```

bash users:

```
IDL_PATH=+/home/user/pipeline_directory:$IDL_PATH:+/home/user/astron/:+/home/user/coyo
te/
```

NOTE: The plus sign (+) refers to all subdirectories within these paths and " : " acts to join different paths.

Then reset your terminal (or source your startup file) or log-out user to make sure that all changes are implemented correctly.

NOTE: In order to avoid versions-conflict, we strongly recommend to add the pipeline directory at top of other routines, before IDL_PATH, as you might set older version of pipeline in your IDL_PATH, earlier.

2-Check the environmental variable IDL_PATH to make sure that all changes are implemented correctly.

If you are using a c shell (csh, tcsh) type:

```
printenv IDL_PATH
```

If you are using a bourne shell (sh, ksh, bash):

```
echo $IDL_PATH
```

Setting project directory

A separate project directory is required for each distinct project. The top level directory contains the control files associated with a set of observations (one for each reduction block, RB) plus a subdirectory contains input data. The name of the project directory is unique and is held in the control file. Default names of the subdirectory is “data”.

What is New:

In version 3, all working subdirectories (int, temp, spec, log) will be created automatically based on the information in the control file.

NOTE: Please note that you can also set the PATH for subdirectories, so it is not necessary to keep control_file and subdirectories in the same folder, however it makes the life easier!

The subdirectory data/ contains all the science and calibration files associated with the RBs within a project. These include copies of bias or dark files, recorded prior to the observing run, which are referenced by the RB. Data files follow the standard format and naming convention of AF2/WYFFOS data files.

Pipeline installation test *NEW Feature*

There is a new routine (af2_version.pro) in current version of pipeline, which checks IDL/pipeline installation and returns a detailed report of your OS/IDL/Pipeline status.

After installing the pipeline and other necessary libraries, run the following routine:

IDL> af2_version

and check the output log (which is also saved as version_check.txt in the same directory).

NOTE: if you can not run this command or received an IDL error while running this routine, it means that you have not installed the new version of pipeline or you are not set the IDL_PATH correctly.

Edit the control file

Parameters controlling the operation of the pipeline are contained in the control data file (`control_data.txt`). This is an ASCII text file unique to a particular RB. The control file comprises a list of parameter values followed by a semi colon and the variable name. Appendix 2 shows a typical control file together with a description of the function of each parameter in the pipeline software.

There are three types of data:

- Data that is expected to vary between RBs. This comprises key parameters, options and the data file numbers.
- Data that varies according to the AF2/WYFFOS set up. These are principally parameters relating to wavelength calibration.
- Quasi-standard parameters that do not generally vary from the default values given in the reference control file provided by the ING.

What is New:

In the current version of pipeline, some of the parameters are eliminated, as they are fixed to an optimal values. If using an old version of the control file, the pipeline automatically remove those parameters and edit the structure of control file to a more user-friendly and human readable structure.

A full description of the parameters in control file is given in Appendix 3.

Running pipeline

A) Calibration modules

There are four calibration modules in the pipeline process which are run prior to extracting spectra (see Fig, 2). These are **Bias**, **Mask**, **Flat** and **Circ**. In principle the Bias and Mask files need only be run if the relevant control parameters and/or data file numbers have changed between RBs. However, their run times are short, so it is good practice to rerun all modules for each RB.

What is New:

In current version of pipeline, we have implemented an improved algorithm for fiber tracing on the CCD in CIRC routine. It contains 4 iterations (Basic peak finder, Extrapolation in low S/N parts using low order polynomials, Gaussian Peak fitting and Extrapolation over the whole range of CCD in Y direction). It also presents the result of fiber tracing on a Master-flat frame using an ATV window and user can check the accuracy of fiber tracing result for each fiber.

B) Wavelength Calibration

ARC routine

This arc module optimally extracts the arc spectra as a function of y-pixel position then uses these intermediate spectra to determine the wavelength calibration.

The Arc module uses a predefined table of arcline data (part of the control data) to identify the approximate y-pixel location and exact wavelengths for a set of well- separated unsaturated lines in the arc spectra. The precise position of the peaks (in pixels) is then found by fitting Gaussian profiles over a short section of the spectrum (typically ±15 pixels) encompassing each peak. The results are used to determine the coefficients of a cubic polynomial of wavelength against pixel number for each active fiber. These coefficients are then used to provide a wavelength scale for the each fiber which is valid for the current RB.

options:

arc, /ATLAS - running atlas module to create and edit the set of arcline data for wavelength calibration.

Atlas routine

The Atlas module is used to create and edit the set of arcline data for wavelength calibration. It can be run independently or called from the Arc module using the option Arc ,/atlas.

The module displays a GUI that takes the user through the steps required to create a set of arcline data, These are:

1) Select the arcline used to find the y-offsets.

2) Select an arc lamp atlas file and the wavelength range.

3) Select arc lines used for wavelength calibration.

4) Accept the calibration and the update control file(s).

What is New:

In this version, atlas routine or arc, /atlas returns a warning message if the wavelength solution is not working well for all fibers. In this case, we strongly recommend to edit the arclines and repeat the calibration process. Please note that, bad wavelength calibration could strongly affected the sky-subtraction procedure and accuracy of final results.

You can also follow the process of wavelength calibration for fibers, one by one in a semi-interactive mode. You can active this feature by putting a minus sign '-' before the value of ARCORDER parameter.

C) Optimal extraction

The process used for optimal extraction follows the algorithm given by Horne (1986).This applies a non-uniform weight to pixels in the extraction sum to minimize statistical noise whilst preserving photometric accuracy (reference Horne K., 1986, PASP, 98, 609).

What is New:

In this version, we introduced two new options to c.profile, which allow users to use master-science frame to calculate the probability profile for optimal extraction (c.profile=2) or a combination of master-flat and master-science frames, depend on S/N and total magnitude of object to calculate the prob. profile. (c.profile=3). Therefore now we have 4 options for optimal extraction:

c.profile=0 - A top hat profile is used producing non-optimal extraction.
c.profile=1 - The spatial profile is found from the flat image for optimal extraction.
c.profile=2 - The spatial profile is found from the master-star image for optimal extraction.
c.profile=3 - The spatial profile is found from the master-star (bright targets) or master-flat (faint targets) for optimal extraction.

Offset sky

The offset module is used to determine the median sky spectrum per fiber from a set of offset sky observations. The offset sky spectra of target fibers (type P) are optimally extracted and the wavelength calibration is copied from the intermediate arc file.

Options:

Offset, /NOEXTRACT – skips optimal extraction stage

New:

Offset, /PCA – Using the PCA algorithm to reconstructed sky spectra (see PCA section: Appendix 2)

Offset, /fprob – force to use master-flat to evaluate the probability profile for optimal extraction (see optimal extraction section)

Output:

File containing intermediate offset sky spectra per observation (t1234567.fit)

File of median offset sky spectra on a common wavelength range (o1234567.fit)

New:

IRAF compatible file of median offset sky spectra on a common wavelength range (iraf_o1234567.fit)

File of median offset sky spectra on a common wavelength range obtained by PCA method (pca_o1234567.fit)

Science spectra

The Star module first extracts the science spectra, both designated targets (type P) and sky fibers (type S), copies the wavelength calibration from the associated intermediate arc file and saves the intermediate science spectra to file (s1234567.fit).

The module then processes the intermediate spectra to produce sky subtracted output spectra on a common wavelength base.

Options:

Star,/NOEXTRACT – skips optimal extraction stage
Star,/PLOT — saves plots of individual spectra for each fiber
Star,/AUTO —Optimal extraction vs pixel number (i.e. no wavelength calibration)

New:

Star, /PCA- Using the reconstructed sky spectra based on PCA method (see PCA section: Appendix 2)

Outputs:

Starfiles (s1234567.fit) containing intermediate target spectra
Specfiles (p1234567.fit) sky subtracted spectra over common wavelength range

IRAF compatible spec files (iraf_p1234567.fit) sky subtracted spectra over common wavelength range.

ASCII tables (p1234567.txt) fiber no, line, RA, Dec, Mag, Mean, SNR, sky lev,name

Test the star module output *NEW Feature*

A new procedure (star_check) has been implemented in new version of pipeline which allows user to check the raw extracted science and sky spectra, the **quality of sky subtraction**, **quality of wavelength calibration** and the final output of star routine for an specific fiber.

Note: User must run the star routine (or offset and star) before running this module.

This routine, takes the fiber number in IDL console and create a set of informative plots and related information in IDL console (Press -1 to exit this routine)

Options:

star_check, wscl=n — Rescale the graphical window using “wscl” parameter

Example:

star_check, wscl=2

The graphical window size will be change to 2*(default size which is 2400*900) which allows user to precisely check the plots.

Then, the routine asks for the fiber number in IDL console and create a set of informative plots for that fiber. (Press -1 to exit this routine or enter another fiber number to continue).

Median of spectra

This module reads the output spectra for each science observation in an RB and evaluates the median spectra for each fiber. Spectra are normalized to their mean value before the median is calculated to compensate for variations in seeing between frames.

Note: The output spectra are the product of this median spectra times the number of observations. The rms deviation is the median rms deviation of individual spectra multiplied by the square root of the number of observations.

Options:

Median,/AUTO - plots data against pixel value (i.e. no wavelength calibration)

inputs:

Specfiles (p1234567.fit) containing spectra for each frame
Note Starfile (s1234567.fit) used in case of /AUTO mode

outputs:

Median file (q1234567.fit) containing spectra over common wavelength range
Median file (u1234567.fit) containing spectra versus pixel no. (for auto mode)
Plots showing individual spectra and resulting median spectra for each active fiber .

Appendix 1 A quick user manual

Below an example of how to run the pipeline to reduce a set of data:

idl> af2_version

*** Make sure that you are using the current version (version 3.0) of pipeline and all necessary routines and libraries are installed.

idl> bias

idl> mask

mask, /AUTO - Mask a fraction of pixels rather than a fixed dark current per pixel

idl> flat

idl> circ

circ, /AUTO identifies fibers located in gap and overwrites values in control file

idl > arc

arc, /NOEXTRACT - skips optimal extraction stage

arc, /ATLAS - runs procedure Atlas to create/edit arcline data

idl> offset

offset, /NOEXTRACT – skips optimal extraction stage.

offset, /PCA – Using the PCA algorithm to reconstructed sky spectra.

offset, /fprob – force to use master-flat to evaluate the probability profile for optimal extraction.

idl> star

star, /NOEXTRACT – skips optimal extraction stage
star, /PLOT - saves plots of individual spectra for each fiber
star, /AUTO - Optimal extraction vs pixel number (i.e. no wavelength calibration)
star, /PCA- Using the reconstructed sky spectra based on PCA method

idl> star_check

star_check, wscl=wscl — rescale the graphical window, based on wscl parameter.

Note: This module runs interactively and user should enter a fiber number to see the results or enter -1 to exit.

idl> medan

medan ,/AUTO - plots data against pixel value (i.e. no wavelength calibration)

Appendix 2

Sky subtraction based on Principle component analysis (PCA)

Principal component analysis (PCA) is a robust statistical technique, based on the statistical description of random variables that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. In this way, the number of principal components is less than or equal to the number of original variables (For more details please refer to : Budavari et al. 2009, MNRAS, 394, 1496).

What is New:

In current version of pipeline, we have applied the PCA approach to the sky subtraction routines, for both offset and sky fiber modes. Our primary results show a significant improvement in quality of sky subtraction procedure, especially in OH sky-line forest.

Please note, implementation of the PCA method in our pipeline is still under development and has not been thoroughly tested on all available configurations.

In order to test this new feature, you can add '/pca' option to offset and star routines, while running them:

offset, /pca (if the offset sky is available)

star, /pca

Note: The primitive sky subtraction method is still available and you can use them as before.

Appendix 3 Control_data.txt

This table, shows the parameters in structure c, typical values and a descriptor. Note structure c contains additional working variables used to transfer data between modules and GUIs.

What is New:

1- In current version of pipeline, some of the parameters are eliminated, as they are fixed to an optimal values. These parameters are:

c.circlev, c.nwl, c.flattype , c.flag#, c.biasfile, c.maskfile, c.flatfile, c.circfile, c.arcfile, c.osetfile, c.starfile

2- In this version of pipeline, the c.profile parameter accepts new values:

c.profile=0 - A top hat profile is used producing non-optimal extraction.

c.profile=1 - The spatial profile is found from the flat image for optimal extraction.

c.profile=2 - The spatial profile is found from the master-star image for optimal extraction.

c.profile=3 - The spatial profile is found from the master-star (bright targets) or master-flat (faint targets) for optimal extraction.

3- In current version of pipeline, the definition of c.ymax has been changed. Now, Ymax refer to the absolute position of MAX-Y cut.