

## MissFITS and WeightWatcher: two optimised tools for managing FITS data.

Chiara Marmo<sup>1</sup>, Emmanuel Bertin<sup>1</sup>

*Institut d'Astrophysique de Paris, Paris, France*

**Abstract.** We present the latest versions of MissFITS and WeightWatcher, two software packages dedicated to handling very large amounts of FITS data in a fast and optimised way. Metadata are generated as VOTables and can easily be accessed as web pages using the provided XSLT interfaces.

### 1. Introduction

Large programs are a feature of modern astronomy, making use of wide fields of view as well as extended wavelength and time domains. Scientific information must be extracted from huge amounts of data, often inhomogeneous in both format and quality. Optimising I/O intensive procedures requires stable data formats and changes in image quality have to be characterised and kept under control to allow reliable scientific measurements over large survey areas. We present the latest versions of MissFITS and WeightWatcher, two software packages dedicated to handling very large amounts of FITS data in a fast and optimised way. Metadata are generated as VOTables and can easily be accessed as web pages using the provided XSLT interfaces.

### 2. MissFITS

The number of standard header keywords specified by the FITS format convention is somewhat limited. Hence each observatory has defined its own set of keywords, which can be a problem for data processing in a generic pipeline.

Data acquired by mosaic cameras are distributed as Multi-Extension FITS (MEF) files or not depending on the instrument. Moreover, some telescopes choose data cubes (FITS files with  $NAXIS = 3$ ) to archive micro-dithered exposures more efficiently.

A processing pipeline may be optimised for single extension FITS files but not for MEFs, or it may be incompatible with data cubes. Some header keywords may be confusing or have unusual names, making them incompatible with the processing chain.

MissFITS is a program that performs basic maintenance and packaging tasks on FITS files using an optimised, house-made FITS library. MissFITS can:

- add/edit/remove FITS header keywords,
- split/join Multi-Extension-FITS (MEF) files,
- unpile/pile FITS data-cubes,

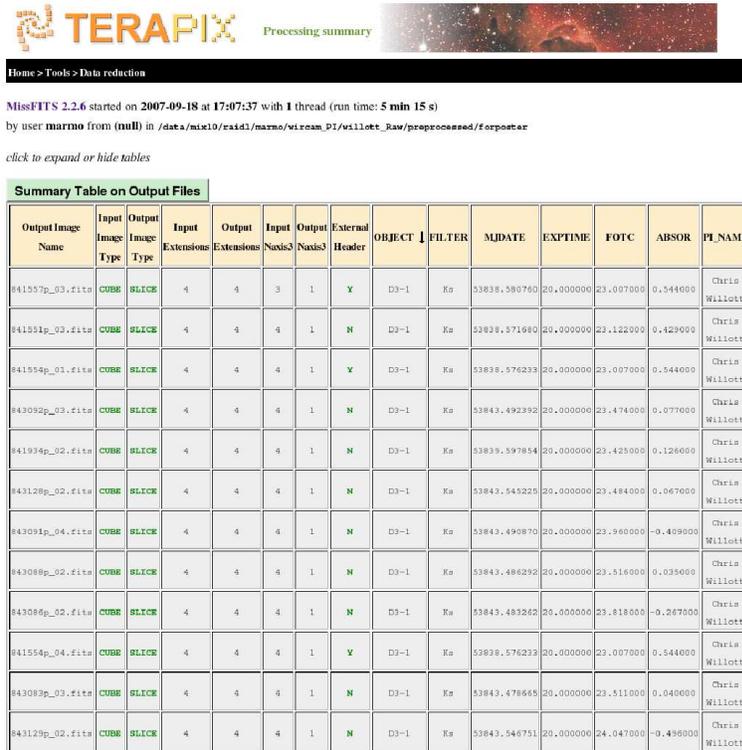
- create/check/update FITS checksums, using R. Seaman's protocol.  
MissFITS is easy to install, for it is packaged conforming to GNU Autoconf and Automake requirements. It can change keyword names, add arbitrary keywords, remove useless or confusing keywords incompatible with some processing chains. It produces XML VOTable compliant meta-data on output files.

## 2.1. A simple case of figure

63 Multi Extensions FITS data cubes have to be transformed in MEF having NAXIS=2. During this transformation, keywords indexed as NAXIS3 have to be changed to non-indexed keywords. External headers containing information about photometric quality need to be added to the FITS header. Observation Date, Exposure Time, the name of the Principal Investigator and photometric properties need to be written to the XML output file. The command line is

```
% missfits *.fits -OUTFILE_TYPE SLICE -SLICE_SUFFIX _%02d.fits -SAVE_TYPE BACKUP
-DISPLAY_KEYWORD EXPTIME,DATEOBS,FOTC,ABSOR,PI_NAME -HEADER_SUFFIX .head
```

The output XML file can be used as a metadata archive: parsing XML to access keyword values is actually easier and faster than reading FITS headers.



The screenshot shows the TERAPIX web interface with a 'Processing summary' section. Below the header, there is a navigation bar and a summary table titled 'Summary Table on Output Files'. The table contains 15 columns: Output Image Name, Input Image Type, Output Image Type, Input Extensions, Output Extensions, Input Naxis, Output Naxis, External Header, OBJECT, FILTER, MDATE, EXPTIME, FOTC, ABSOR, and PI\_NAME. The table lists 15 rows of data for various FITS files, including their image types, extension counts, NAXIS values, object names, filters, observation dates, exposure times, photometric quality (FOTC), absorption (ABSOR), and the Principal Investigator's name (PI\_NAME).

Output Image Name	Input Image Type	Output Image Type	Input Extensions	Output Extensions	Input Naxis	Output Naxis	External Header	OBJECT	FILTER	MDATE	EXPTIME	FOTC	ABSOR	PI_NAME
841537p_03.fits	CUBE	SLICE	4	4	3	1	Y	D3-1	Ka	53838.580760	20.000000	23.007000	0.544000	Chris Willott
841551p_03.fits	CUBE	SLICE	4	4	4	1	N	D3-1	Ka	53838.571660	20.000000	23.122000	0.429000	Chris Willott
841554p_01.fits	CUBE	SLICE	4	4	4	1	Y	D3-1	Ka	53838.576233	20.000000	23.007000	0.544000	Chris Willott
843092p_03.fits	CUBE	SLICE	4	4	4	1	N	D3-1	Ka	53843.492392	20.000000	23.474000	0.077000	Chris Willott
841934p_02.fits	CUBE	SLICE	4	4	4	1	N	D3-1	Ka	53839.597854	20.000000	23.425000	0.126000	Chris Willott
843128p_02.fits	CUBE	SLICE	4	4	4	1	N	D3-1	Ka	53843.545225	20.000000	23.484000	0.067000	Chris Willott
843091p_04.fits	CUBE	SLICE	4	4	4	1	N	D3-1	Ka	53843.490870	20.000000	23.960000	-0.409000	Chris Willott
843088p_02.fits	CUBE	SLICE	4	4	4	1	N	D3-1	Ka	53843.486292	20.000000	23.516000	0.035000	Chris Willott
843086p_02.fits	CUBE	SLICE	4	4	4	1	N	D3-1	Ka	53843.483262	20.000000	23.818000	-0.247000	Chris Willott
841554p_04.fits	CUBE	SLICE	4	4	4	1	Y	D3-1	Ka	53838.576233	20.000000	23.007000	0.544000	Chris Willott
843083p_03.fits	CUBE	SLICE	4	4	4	1	N	D3-1	Ka	53843.478665	20.000000	23.511000	0.040000	Chris Willott
843129p_02.fits	CUBE	SLICE	4	4	4	1	N	D3-1	Ka	53843.546751	20.000000	24.047000	-0.436000	Chris Willott

Figure 1. XSLT translation of the XML-VOTable output of MissFITS. This XSLT filter is provided in the public distribution.

### 3. WeightWatcher

Control maps are a mean to quantify the quality of pixel data for error determination and source properties reliability. TERAPIX software can manage two types of control maps: weight maps and flag maps. In weight maps, pixel values are inversely proportional to the local variance of the image; image weights affect error estimates in source measurements and the relative weighting of pixels in image combinations. In flag maps, bit values label specific issues with pixels in the image: source detection is not affected but objects detected in flagged regions are flagged accordingly. WeightWatcher is a program that combines weight-maps, flag-maps and polygon data (produced with ds9) in order to produce control maps which can directly be used in astronomical image-processing packages like Drizzle (Fruchter & Hook 2002), Swarp (Bertin et al. 2002) or SExtractor (Bertin & Arnouts 1996). Weight-thresholding and/or specific flag selections are applied by WeightWatcher through a configuration file: this alleviates other programs from such interpretation work. WeightWatcher is easy to

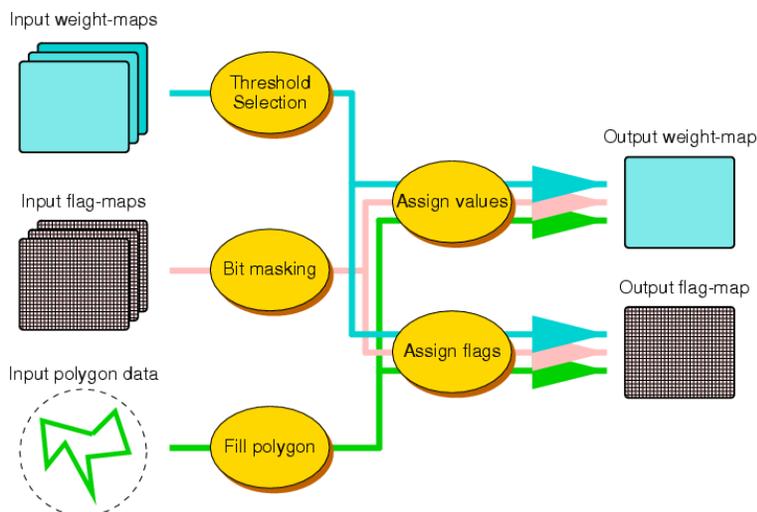


Figure 2. the WeightWatcher layout.

install, for it is packaged conforming to GNU Autoconf and Automake requirements. Its processing speed is limited by the I/O performances of the machine (typically 50 Mpixel/s on a workstation). It is able to work with very large images (up to, say,  $10^8$ – $10^9$  pixels on a 64 bit system). It can handle simultaneously up to 30 weight-maps, 30 flag-maps, and rasterise thousands of polygons at once. It performs statistics on flagged and weighted areas. It produces XML VOTable compliant meta-data.

#### 3.1. A simple case of figure

A flat-field is used to build a weight map for an image in the context of source extraction. One generic bad pixel mask and one specific to the image itself are

used to remove spurious detections. Pixels with a value of zero in the image are flagged and weighted to zero. The command line is

```
% ww -WEIGHT_NAMES domeflat_Ks.fits,cosmic.fits,badpix.fits,841511p_01.fits
-WEIGHT_MIN 0.7,-1.e9,0.5,1.e-15 -WEIGHT_MAX 1.3,0.1,1.5,1.e16 -WEIGHT_OUTFLAGS 2,8,1,16
-POLY_NAMES flat.reg -POLY_OUTFLAGS 4 POLY_OUTWEIGHTS 0 -GETAREA Y -GETAREA_WEIGHT 0.0
-GETAREA_FLAGS 255 -OUTWEIGHT_NAME 841511p_01_weight.fits -OUTFLAG_NAME 841511p_01_flag.fits
```

The unflagged area useful for science is computed and printed to the XML output file.



Figure 3. XSLT translation of the XML-VOTable output of Weight-Watcher. This XSLT filter is provided in the public distribution.

MissFITS and WeightWatcher are available for download at <http://terapix.iap.fr/soft/missfits> and <http://terapix.iap.fr/soft/weightwatcher> respectively. User's guides are also provided. For any question or suggestion users can visit the TERAPIX Forums<sup>1</sup>.

## References

- Bertin, E. & Arnouts, S. 1996, *â*, 117, 393.  
 Bertin, E., Mellier, Y., Radovich, M. Missonnier, G., Didelon, P., & Morin, B. 2002, in ASP Conf. Ser. 281, ADASS XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley (San Francisco: ASP), 228.  
 Fruchter, A. S. Hook, R. N. 2002 PASP, 114, 144.

<sup>1</sup><http://terapix.iap.fr/forum>