

Technology for digitization of astronomical photographic plates

Nikolay Kirov¹, Milcho Tsvetkov² and Katya Tsvetkova³

Abstract. We present a new technology for digitization of astronomical photographic plates. This technology is applied to include new plates into the Wide-Field Plate Database (WFPDB, [9]). The need of well-formatted plate metadata put requirements to the acceptance of standard formats for plate/archive description, as well as to the software used for data processing applied to original plate catalogues.

The Virtual Observatory [7] requires that the digital image has to be in Flexible Image Transport System (FITS, [6]) file format and equipped with a complete header. The WFPDB offers the most of the metadata needed for creation of the FITS header. The presented here software gives the user appropriate tools for connecting the WFPDB plate catalogues and the plate images.

The professional flatbed scanners are the most appropriate devices for scanning astronomical photographic plates. Usually it is imposed to change image files from the inner format of the scanner (or driver) to the standard FITS format. Our software converts the VueScan [8] output TIFF file to FITS file, and in case of scanning with grayscale wedge, divides the plate image and the wedge image.

Keywords: Astroinformatics, astronomy, digitization.

INTRODUCTION

The Wide-Field Plate Database (WFPDB) is a basic source of data for the wide-field (larger or equal of one square degree) astronomical photographic plates obtained with professional telescopes world-wide (see [5]). The WFPDB represents an instrument in the astronomical research, which allows obtaining information on existing astronomical observations of celestial objects over the past 130 years. Briefly, this is a unique virtual telescope working as a "Time Machine" for obtaining data on historical observations of minor planets, comets, stars and galaxies (for more details see [2], [4]).

The largest plate collection in the world, maintained in one observatory, is Harvard College Observatory plate collection [1]. The digitization of the plates was started several years ago with an especially designed for this purpose scanner. The European plates are distributed in many observatories and other institutions and the Harvard's approach is not applicable in Europe as well as in the small observatories in the USA. Our technology involves first-stage description of the plates and the inclusion of their metadata into the WFPDB. The second step is scanning the

¹ Informatics Department, New Bulgarian University, 21 Montevideo str., 618 Sofia, Bulgaria, nkirov@nbu.bg

² Institute of Astronomy and National Astronomical Observatory, Bulgarian Academy of Sciences, 72, Tsarigradsko Chaussee Blvd., 1784, Sofia, Bulgaria, milcho.tsvetkov@gmail.com

³ Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Akad. G. Bonchev str., block 8, 1113 Sofia, Bulgaria, katya.tsvetkova09@gmail.com

plates with professional flatbed scanners and further processing for standardization of the output images for easy handling with astronomical software and easy access via Internet.

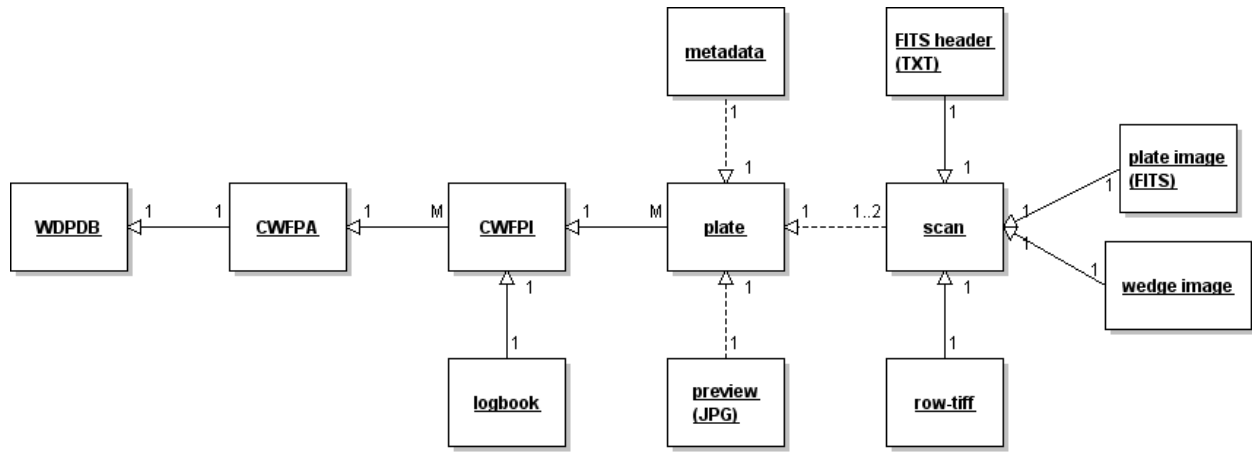


Fig. 1. Main elements of WFPDB with their links

The diagram in Fig. 1 shows the main elements of WFPDB and the links among them. The elements are:

- CWFPA – The Catalogue of Wide-Field Plate Archives contains data for archives – the set of plates which are obtained with one telescope or camera at one place. In the actual version 5.7 there are 476 archive descriptions [9]. For every archive the unique identifier is constructed from abbreviation of observatory name and instrument (telescope/camera) aperture.
- CWFPI – The Catalogue of Wide-Field Plate Indexes contains meta-data for plates with regularly updated version (now more than 500 thousand plates). For every plate many details are stored in the database: the coordinates of the plate center, the date and time of the observation, object name and type, method of observation, duration of exposures, type of emulsion, the size of the plate, the quality of the plate, the name of the observer, etc. [4].
- logbook – The scanned astronomical logbooks allow direct access to the notices of the observer, which are written in different languages and it is difficult to convert them to computer readable form. If the logbook exists and can be found in the observatory archives, it is the main source of plate metadata.
- plate – The plate is the based object in WFPDB – with its unique identifier every plate can be found in the corresponding CWFPI. It consists of archive identifier and the original plate number. For example the identifier ROZ200_000417 is for a plate with original serial number 417 from the plate archive of the Rozhen Observatory obtained by the 2 m RCC telescope. The plate as a physical object is represented in the WFPDB with preview image, scan image and metadata. That is why these three components are mentioned in Fig. 1 with dotted arrows.
- preview image – Digitized plate color image (JPG file) with low resolution for quick plate visualization and easy on-line access.

- scan image – For every plate there are one or two scan images done with high resolution. The choice of the needed resolution depends on the astronomical task for plate use, and on the telescope used. So the plate with good astrometric accuracy needs higher resolution for scanning. In order to avoid possible distortion during the scanning the plate is scanned twice with rotation at 90 degree – for increasing precision of the astrometric measurements.
- row-tiff – The image, produced by flatbed scanner as a result of plate scanning.
- fits header – Plain text file containing FITS file header for easy on-line access to the plate meta-data.
- plate image – Digitized plate (16-bit grayscale) image (FITS file), with high resolution aiming photometric and astrometric measurements.
- wedge image – Some of the plates are scanned with a grayscale step wedge for calibrating the scanning process – we separate the wedge image because some astronomical software work correctly only with the sky images.

The technology is supported by five software tools: timetool, epochtool, header, tif2fits and updatefitshdr programs. The software is written in C++ using Qt cross platform application (see also [3]).

WFPDB METADATA FORMAT PREPARATION

There are different sources for gathering plate metadata:

- telescope logbooks;
- photographic plates;
- plate envelopes.

As a result the original plate catalogues very often have their own format, which requires to be converted to the WFPDB format. The data which need to be converted in the most cases are the coordinate's transformation to epoch 2000.0 and the time transformation. Epochtool software transforms equatorial coordinates (RA and DEC) from arbitrary equinox to J2000. The time of observation, which is written in the astronomical logbooks, is often in local time or sidereal time. Timetool software transforms the time from local sidereal time (LST) or local time (LT) to universal time (UT). The both tools read the input data in the WFPDB format and output the data in the same format, following some strong conventions about file names.

The accepted WFPDB format for plate metadata gives possibility to select plates among all included plate catalogues. For this purpose the WFPDB has a WEB-based search engine [9].

FITS FORMAT REQUIREMENTS AND PLATE DIGITIZATION SOFTWARE

Flexible Image Transport System or FITS is a digital file format used to store, transmit, and manipulate scientific and other images (see [6]). Each FITS file consists of a header containing

ASCII card images (80 character length strings) that carry keyword/value pairs and an image data block. Creation of complete FITS header is an important part of the plate's digitization. The photographic plates are a specific type of astronomical observations and the FITS header can be done using a standard set of keywords. Our standard consists of 67 keywords describing the file, the plate, the observatory, telescope and scan data (see [3]). The corresponding software implementation for this standard is also created. The most of the metadata for the header is extracted from the corresponding WFPDB catalogs.

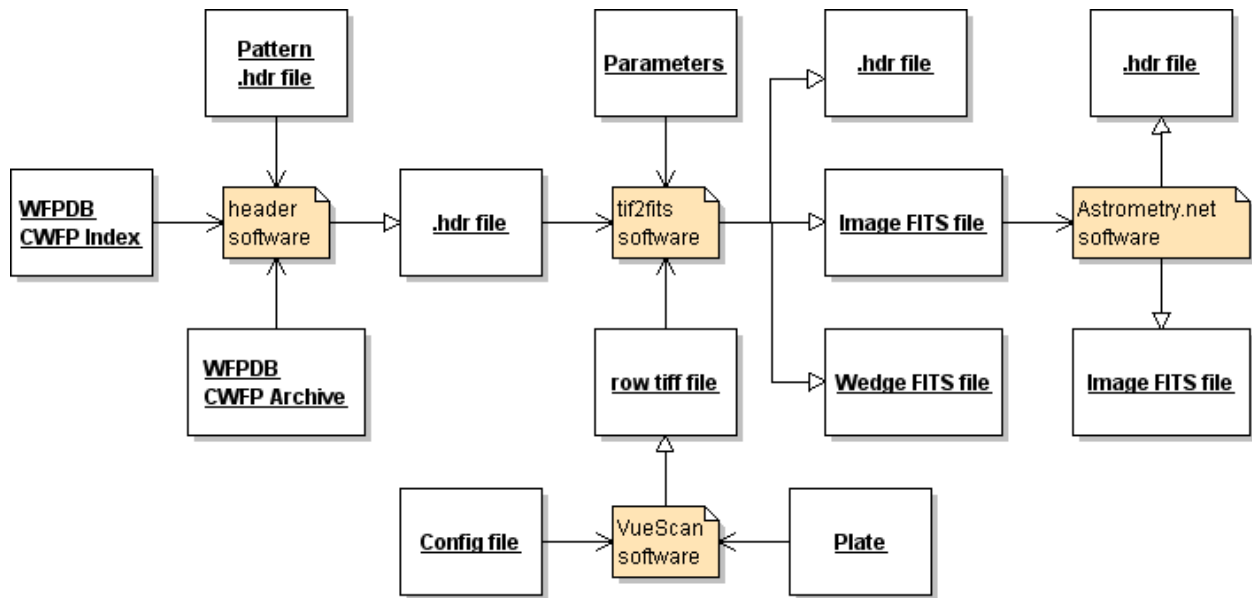


Fig. 2. Data flow diagram for creation of the FITS images

The data flow diagram (Fig. 2) represents the "flow" of data through the information system, modeling its process aspects of creation the FITS format image files. Starting from the WFPDB catalogues we create a text file, which contains the FITS header for a given plate (.hdr file). The input data for the tif2fits software is this file and an image file, obtained as a result of the scanning process with VueScan driver (.row-tiff file). The main functions of tif2fits software are:

- conversion of the row-tiff files, produced by VueScan to FITS format files,
- updating some fields in the FITS header and adding the header to the image to complete the FITS file,
- separation of the wedge part of the image producing two FITS files – one for the plate and one for the wedge,
- rotation and/or flip the plate image,
- splitting up two images in case of scanning two small plates together.

The software tool updatefitshdr is used for replacement of fits header in a FITS file. It is very helpful in the cases with many images (FITS files) of scanned plates with very incomplete FITS headers.

THE TECHNOLOGY IN PRACTICE

The technology has been applied to different plate archives. The software timetool and epochtool have been used for creating the CWFPI of Kvistaberg Observatory, Uppsala, Sweden (UPS100) and National Astronomical Observatory, Rozhen, Bulgaria (ROZ050).

Table 1 presents our experience of applying the described scan technology. The FITS headers were composed using the header tool, the plates were scanned with grayscale wedge and the created raw TIFF files were transformed to two FITS format files each one – for the plate image and for the wedge image – by tif2fits software tool.

Table 1. Scanned data for plate catalogues with application of the new technology

Telescope (catalogue)	Scale (arcsec/mm)	Plate size (cm x cm)	Resolution (dpi)	File size (MB)
POT015	138	20 x 20	2400	680
POT080	17	16 x 16	1600	190
POT032	120	16 x 16	2400	440
BAM10C	338	16 x 16	2400	440
BON030	138	16 x 16	2400	440
HAR025C	167	20 x 25	1600	430
ROZ050	120	16 x 16	2400	440
ROZ200	13	16 x 16	2400	440
ROZ200	13	30 x 30	1600	640
KON060	144	16 x 16	2400	440
CLU020	69	9 x 12	2400	186
CLU050	82	6 x 9	2400	92
BUC016	258	13 x 18	2400	400
BUC038	34	24 x 24	1600	430
BUC038	34	13 x 18	1600	120

The plate catalogues are from the Astrophysical Observatory Potsdam now Leibniz-Institute for Astrophysics (POT); Dr. Remeis Observatory Bamberg, Astronomical Institute of the University of Erlangen, Nuremberg (BAM); Argelander-Institute for Astronomy, Bonn (BON); Konkoly Observatory, Hungarian Academy of Sciences (KON); National Astronomical Observatory Rozhen, Bulgarian Academy of Sciences (ROZ); Cluj Observatory of the Institute of Astronomy, Romanian Academy (CLU); Bucharest Observatory of the Institute of Astronomy, Romanian Academy (BUC).

CONCLUSIONS

The technology presented here is an important milestone in the expansion and development of the WFPDB. It closes the cycle from finding information about new plates to including the plates

into the database with their complete metadata and digitized plate images. The new technology speeds up the digitization processing time and decreases the possible errors in FITS header. Software tools are especially designed for WFPDB data format, so they reduce human work in the process of incorporating new archives (catalogues and plates) in the database.

ACKNOWLEDGEMENTS

This work is supported by German DFG grant STE: 710/6-1/2009 and Bulgarian Ministry of Education, Science and Youth grant: DO-02-275/2009.

BIBLIOGRAPHY

- [1] Grindlay, J., et al., *DASCH to Measure (and preserve) the Harvard Plates: Opening the ~100-year Time Domain Astronomy Window*, In: Preserving Astronomy's Photographic Legacy: Current State and the Future of North American Astronomical Plates. ASP Conference Series, Vol. 410. Eds. W.Osborn and L.Robbins. San Francisco: Astronomical Society of the Pacific, 2009, p.101.
- [2] Kirov, N., Kounchev, O., Tsvetkov, M., *Astroinformatics: Image Processing and Analysis of Digitized Astronomical Data with WEB-based Implementation*, Proceedings of the 6th Annual International Conference on Computer Science and Education in Computer Science, 26-27 June 2010 in Fulda and 28-29 June 2010 in Munich, Germany, 176-181.
- [3] Kirov, N., Tsvetkov, M., Tsvetkova, K., *Software tools for digitization of astronomical photographic plates*, Serdica Journal of Computing, vol. 6, No.1 (2012).
- [4] Tsvetkova, K., Tsvetkov, M., *Catalogue of Wide-Field Plate Archives: version 5.0*, In: Virtual Observatory: Plate Content Digitization, Archive Mining and Image Sequence Processing, Eds. M. Tsvetkov, V. Golev, F. Murtagh, R. Molina, Heron Press, Sofia, 2006, 45-53.
- [5] Tsvetkov, M., *Wide-Field Plate Database: a Decade of Development*, *ibid*, 10-41.
- [6] Pence, W. D., et al., *Definition of the Flexible Image Transport System (FITS), version 3.0*, *Astronomy & Astrophysics* 524, A42 (2010).
- [7] The European Virtual Observatory EURO-VO, <http://www.euro-vo.org>
- [8] VueScan scanning software, <http://www.hamrick.com>
- [9] Wide-Field Plate Database (WFPDB), <http://www.wfpdb.org>