Proceedings of the VIII Serbian-Bulgarian Astronomical Conference (VIII SBAC)
Leskovac, Serbia, May 8-12, 2012,
Editors: M. S. Dimitrijević and M. K. Tsvetkov
Publ. Astron. Soc. "Rudjer Bošković" No 12, 2013, 349-357

WFPDB METADATA FORMAT PREPARATION

KATYA TSVETKOVA[1] and MILCHO TSVETKOV[2]

[1] *Institute of Mathematics and Informatics, Bulgarian Academy of Sciences*
[2] *Institute of Astronomy, Bulgarian Academy of Sciences*

Abstract. We present an overview of the preparation of plate metadata for the Wide-Field Plate Database (WFPDB, http://www.wfpdb.org), of the accepted formats as standards for plate/archive description, the data reduction applied to original plate catalogues, as well as the work of making the existed logbooks accessible online

1. INTRODUCTION

The re-use of the old astronomical photographic plates needs a selection of worthy observations, which is dependent on the available plate metadata, i.e. the plate metadata have to contain all necessary information for future plate identification, selection and processing. In order to index and make as much as possible informative plate search and selection, the Wide-Field Plate Database (WFPDB, http://www.wfpdb.org) requires well-formatted plate metadata and data reduction pipeline. The first established standards of plate metadata were given in the standardized description of the ReadMe file (http://cdsarc.u-strasbg.fr/viz-bin/Cat?VI/90) of the WFPDB (Tsvetkov et al. 1997) which can be found through the Catalog VI/90 Selection Page in the Strasbourg Astronomical Data Center (CDS). Since the installation of the WFPDB in CDS the plate metadata standards have been developed. Here we present this development of the accepted formats as standards for plate/archive description, the used data reduction applied to original plate catalogues, as well as the work of making the existed logbooks accessible online.

2. WFPDB ARCHIVES AND PLATE INDEX METADATA

Informational sources for the plate archives included in the Catalogue of the Wide-Field Plate Archives (CWFPAs, Tsvetkova and Tsvetkov 2006, 2008) and for the individual plates in the Catalogue of the Wide-Field Plate Indexes

(CWFPIs, Tsvetkov 2006) as parts of the WFPDB, are the observatories or other astronomical institutions with plate observations. In preparation of the list of such organizations we used The Strasbourg Astronomical Institutes Directory, The Nautical Astronomical Almanac, Annual Reports of different observatories, their Web pages, private correspondence and contacts, published plate data files and manual data entries or applying Optical Character Recognition tool.

The formats of the CWFPAs and CWFPIs are given in Table 1 and Table 2.

Table 1: Byte-by-byte Description of file: archives

Bytes	Format	Units	Label	Explanations
1-3	A3	---	IDobs	[1]WFPDB observatory identifier
4-6	I3	cm	IDins	[1]Instrument aperture
7	A1	---	IDsuf1	[1] [A-Z]Suffix to the instrument identifier
9-25	A17	---	LOCs	Location of the plate archive, town (site)
27-40	A14	---	LOCc	Location of the plate archive, country
42-64	A23	---	OBSn	Observatory, name
66-83	A18	---	OBSs	Observatory, site
85-95	A11	---	OBSc	Observatory, country
97-99	I3	---	MNo	Marsden's number
101	A1	---	TZ-	Time zone, sign
102-103	I2	h	TZ	Time zone
105	A1	---	LON-	Observatory longitude, sign
106-108	I3	deg	LONd	Observatory longitude, deg
110-113	F4.1	arcmin	LONm	Observatory longitude, arcmin
115	A1	---	LAT-	Observatory latitude, sign
116-117	I2	deg	LATd	Observatory latitude, deg
119-122	F4.1	arcmin	LATm	Observatory latitude, arcmin
124-127	I4	m	ALT	Observatory altitude
129-130	I2	---	MULT	Multiplicity of telescope cameras
131	A1	---	---	[x] Sign 'x'
132-135	F4.2	m	APR	Clear aperture of the telescope
137-140	F4.2	m	MD	Diameter of telescope mirror
142-146	F5.2	m	FL	Focal length of the telescope
148-151	I4	arcsec/mm	SCL	Plate scale
153-155	A3	---	ITYPE	[2]Instrument type
157-161	F5.1	deg	FIELD	Field angular dimension
163-166	I4	yr	YEAR1	Year of beginning of telescope operation
168-171	I4	yr	YEAR2	Year of end of telescope operation
173	A1	---	PF	[F] Indication 'F' for 'film'
175-180	I6	---	NPd	Number of direct plates

Table 1: (continuation)

Bytes	Format	Units	Label	Explanations
181	A1	---	NPUNd	[:] Uncertainty of the number of plates
183-184	A2	---	CFORMd	[3]Plate catalog form (direct plates)
186-190	I5	---	NPs	Number of objective prism plates
191	A1	---	NPUNs	[:] Uncertainty of the number of plates
193-194	A2	---	CFORMs	[3]Plate catalog form (obj. prism plates)
196-197	A2	---	QUAL	[4][ABD] Code for archive quality
199-212	A14	---	ANAME	Astronomer in charge

Notes to Table 1:
[1] Fields from byte 1 to byte 7, taken together, constitute the WFPDB instrument identifier.
[2] Ast - astrograph, Cam - camera, FEC - fish eye camera, Men - meniscus, RCr - Ritchey-Chretien, Rfl -reflector, Rfr - refractor, Sch – Schmidt
[3] C - computer-readable form, T - printed table form, TC - computer-readable form in preparation
[4] A - very good, B - good, D – distributed

Table 2: Byte-by-byte description of file CWFPIs maindata

Bytes	Format	Units	Label	Explanations
1-3	A3	---	IDobs	[1]WFPDB observatory identifier
4-6	I3	cm	IDins	[1]Instrument aperture
7	A1	---	IDsuf1	[1] [A-Z] Suffix to the instrument identifier
8-13	I6	---	IDno	[1]Original plate number
14	A1	---	IDsuf2	[1] [A-Z] Suffix to the original plate number
15-16	I2	h	RAh	Right ascension (hours) (J2000.0)
17-18	I2	min	RAm	Right ascension (minutes)
19-20	I2	s	RAs	Right ascension (seconds)
21	A1	---	DE-	Declination sign (J2000.0)
22-23	I2	deg	DEd	Declination, degrees
24-25	I2	arcmin	DEm	Declination, arcminutes
26-27	I2	arcsec	DEs	Declination, arcseconds
28	A1	---	CCOD	[EMU] Code for Error, Missing data, or Uncertainty of coordinates
29-32	I4	yr	DATEy	Date of observation, year (UT)
33-34	I2	month	DATEm	Date of observation, month
35-36	I2	d	DATEd	Date of observation, day
37-38	I2	h	UTh	Observation time (hour) (UT)
39-40	I2	min	UTm	Observation time (min)
41-42	I2	s	UTs	Observation time (sec)

Table 2: (continuation)

Bytes	Format	Units	Label	Explanations
43	A1	---	TCOD	[EMU] Code for Error, Missing data, or Uncertainty of observation time
44-63	A20	---	OBJNAM	Object or field designation
64-65	A2	---	OBJTYP	2Object type code
66-67	I2	---	METHOD	3Method of observation code
68-69	I2	---	MULTEX	Multiplicity of exposure
70-75	F6.1	min	EXP	4Exposure time
76-86	A11	---	EMULS	Emulsion type
87-93	A7	---	FILT	Filter type
94-95	A2	---	SPEC	Spectral band
96-97	I2	cm	DIMx	X dimension of plate
98-99	I2	cm	DIMy	Y dimension of plate
100	I1	---	PQUAL	[0,1] Pointer to file 'quality'
101	I1	---	PNOT	[0,1] Pointer to file 'notes'
102	I1	---	POBS	[0,1] Pointer to file 'observer'
103	I1	---	PAVA	[0,1] Pointer to file 'availability'
104	I1	---	PDIG	[0,1] Pointer to file 'digitization'

Notes to Table 2:

[1] Fields from byte 1 to byte 14, taken together, constitute the WFPDB plate identifier. Fields from byte 1 to byte 7 constitute the WFPDB instrument identifier.

[2] Object type in WFPDB is coded as is shown in Table 3.

[3] Method of observation is coded as is shown in Table 4.

[4] For multi-exposures with different duration of the separate exposures the 2nd, 3rd,… exposures are given in file Notes, if available in the original plate catalogues.

The plate metadata is structured in a multi-file system with conventional file names – Maindata (byte-by-byte description is given in Table 2), Quality (Table 5), Notes (Table 6), Observer (Table 7), Availability (Table 8), and Digitization (Table 9). All files belonging to this multi-file system follow certain file formats given in the respective byte-by-byte description of the files.

Table 3 and Table 4 present the assigned codes for object type and method of observation. It is common that in the plate logbooks the object type is not given. The used method of observations can be found usually in the notes of the respective logbooks. The coded information for the object type and method of observation allows making selection by these searchable constrains quickly. To code this information is a laboring work done manually up to the moment and this is a reason that very often it is missing in the WFPDB catalogues.

Table 3: WFPDB object type code

Code	Object
A1	Planet
A2	Moon
A3	Sun
A4	Asteroid
A5	Comet
A6	Meteor
A7	Artificial satellite
S1	Star
S2	Double star, or multiple star
S3	Variable star
S4	Stellar cluster
S5	HII region
S6	Nebula
S7	Planetary nebula
S8	Supernova + SN remnants
S9	Fundamental star
SR	Reference star around a radio source
SA	Stellar association
SD	Dark nebula
SH	Herbig-Haro object
SM	Molecular cloud
SP	Pulsar
G1	Galaxy
G2	QSO
G3	Group of galaxies
G4	Cluster of galaxies
G5	Supercluster
G6	Void
F	Field
G7	Radio galaxy
GR	Gamma-ray source
XR	X-ray source
RS	Radio source
IR	Infrared source
U	Object of unknown nature

Table 4: Code for method of observation

Code	Method of Observation
1	Direct photograph
2	Direct photograph, multi-exposure
3	Stellar tracks
4	Objective prism
5	Objective prism
6	Metcalf's method
7	Proper motions
8	No guiding
9	Out of focus
10	Test plate
11	Hartmann test
12	With mask
14	Sub-beam (Pickering) prism
13	Focusing
15	Raster scan/trail
24	Objective grating
25	Objective grating, multi-exposure

Table 5: Byte-by-byte description of file Quality

Bytes	Format	Units	Label	Explanations
1-3	A3	---	IDobs	WFPDB observatory identifier
4-6	I3	cm	IDins	Instrument aperture
7	A1	---	IDsuf1	[A-Z] Suffix to the instrument identifier
8-13	I6	---	IDno	Original plate number
14	A1	---	IDsuf2	[A-Z] Suffix to the original plate number
15	A1	---	CONT	Continuation sign (1, 2, ...) or blank
16-80	A65	---	QTEXT	Text of quality information

Table 6: Byte-by-byte description of file Notes

Bytes	Format	Units	Label	Explanations
1-3	A3	---	IDobs	WFPDB observatory identifier
4-6	I3	cm	IDins	Instrument aperture
7	A1	---	IDsuf1	[A-Z] Suffix to the instrument identifier
8-13	I6	---	IDno	Original plate number
14	A1	---	IDsuf2	[A-Z] Suffix to the original plate number
15	A1	---	CONT	Continuation sign (1, 2, ...) or blank
16-80	A65	---	NTEXT	Text of note

Table 7: Byte-by-byte Description of file Observer

Bytes	Format	Units	Label	Explanations
1- 3	A3	---	IDobs	WFPDB observatory identifier
4-6	I3	cm	IDins	Instrument aperture
7	A1	---	IDsuf1	[A-Z] Suffix to the instrument identifier
8-13	I6	---	IDno	Original plate number
14	A1	---	IDsuf2	[A-Z] Suffix to the original plate number
15	A1	---	CONT	Continuation sign (1, 2, ...) or blank
16-57	A42	---	OBSNAM	Observer's name(s)

Table 8: Byte-by-byte Description of file Availability

Bytes	Format	Units	Label	Explanations
1-3	A3	---	IDobs	WFPDB observatory identifier
4 6	I3	cm	IDins	Instrument aperture
7	A1	---	IDsuf1	[A-Z] Suffix to the instrument identifier
8-13	I6	---	IDno	Original plate number
14	A1	---	IDsuf2	[A-Z] Suffix to the original plate number
15	A1	---	CONT	Continuation sign (1, 2, ...) or blank
16-57	A42!	---	ATEXT	Text of availability

Table 9: Byte-by-byte Description of file Digitization

Bytes	Format	Units	Label	Explanations
1-3	A3	---	IDobs	WFPDB observatory identifier
4-6	I3	cm	IDins	Instrument aperture
7	A1	---	IDsuf1	[A-Z] Suffix to the instrument identifier
8-13	I6	---	IDno	Original plate number
14	A1	---	IDsuf2	[A-Z] Suffix to the original plate number
15	A1	---	CONT	Continuation sign (1, 2, ...) or blank
16-57	A42?	---	DTEXT	Text of digitization

3. WFPDB CATALOGUE PREPARATION

The original plate catalogues have been prepared in the computer-readable form from the existing logbooks by the staff of the observatories or in many cases by the WFPDB team. In some cases, there are even no diaries of observations and then the only option is taking the information directly from the photographic plates or plate envelopes. Another problems occurred during the catalogue preparation are missing coordinates (relying only on the given object name) and no given time of the exposure beginning. Such problems are mentioned in the

maindata file with code “M” for missing information. There are also typo-mistakes, which can not be corrected logically; such cases are mentioned with code "E" standing for errors.

The logbooks are an important part of every plate archive – they contain the original plate records and sometimes are the only source of information for the plate metadata, very often they contain also additional information, which facilitates the possible future plate use. There is a new opportunity for extended logbook search in the WFPDB and access to the scanned logbooks of different plate collections. For the preparation in computer readable form of the existed scanned logbooks (in typed table form) and creation of tables in the WFPDB format the Optical Character Recognition (OCR) tool (Package CuneiForm) for converting the typed original plate catalogues in table form to electronic form was applied last years (Kirov et al. 2012). Another new possibility is the linkage of the astronomical photographic plates scanned images from the WFPDB and the page images from the original astronomical journals using the segmentation of the images from the logbooks.

A comparison of the used OCR software (http://www.cuneiform.ru/) for extracting data from the printed tables with original tables was made for the WFPDB TOK016 and TOK020 catalogues (Kirov et al. 2012). The OCR tool gave 90% correct recognition of the table columns.

4. WFPDB CATALOGUE DATA REDUCTION

According to our experience the most needed data conversion comprises the coordinates and the given time of the beginning of the observations. The plate centre coordinates transformation has to give the equatorial coordinates (R.A. and DEC) in the accepted epoch 2000.0. Since 2010 we used a data pipeline routine developed by N.Kirov and applied to the maindata file of the accepted WFPDB data format. Another data pipeline routine is used for the transformation of local sidereal time (LST) used very often for recording the observing data, to the required universal time (UT). For the conversion of the local standard time to UT the data for the given time zone in the CWFPAs are used, as well the information for the Daylight Saving Time (DST) for the different countries was taken in view.

.

Acknowledgements

This work is supported by the Bulgarian National Science Fund (grants BG NSF DO-02-273 and BG NSF DO-02-275).

References

Kirov, N., Tsvetkov, M., Tsvetkova, K., Kalaglarsky, D.: 2012, *Proceedings of VII BSAC, Chepelare, Bulgaria, June 1-4, 2010,* Eds. M. K. Tsvetkov, M. S. Dimitrijević, K. Tsvetkova, O. Kounchev, Ž. Mijajlović, *Publ. Astron. Soc. "Rudjer Bošković",* No 11, 147-152.

Tsvetkov, M. K., Stavrev, K. Y., Tsvetkova, K. P., Mutafov, A. S., Semkov, E. H.: 1997, *Standardized description of the catalogue ReadMe (http://cdsarc.u-strasbg.fr/viz-bin/Cat?VI/90*

Tsvetkov, M.: 2006, *Wide-Field Plate Database: a Decade of Development,* In: *Virtual Observatory: Plate Content Digitization, Archive Mining and Image Sequence, Processing*, iAstro workshop, Sofia, Bulgaria, Eds. M. Tsvetkov, F. Murtagh, R. Molina.

Tsvetkova, K., Tsvetkov, M.: 2006, In: *Virtual Observatory, Plate Content Digitization, Archive Mining, Image Sequence Processing,* Eds. M. Tsvetkov, V. Golev, F. Murtagh, R. Molina, Heron Press Science Series, Sofia, 45-53.

Tsvetkova, K., Tsvetkov, M.: 2008*, VizieR Online Data Catalog:* VI/126, 10/2008.