

MASSEE International Congress on Mathematics MICOM'2009

Workshop Astrominformatics: Computer Science and Mathematical Methods in Astronomy

Ohrid, September 16-20, 2009

**Direct access to digitized logbooks of astronomical Wide-Field Plate
Collections**

N. Kirov*, M. Tsvetkov**

*Institute of Mathematics and Informatics, BAS *and* Department Informatics, NBU

*Institute of Astronomy, Bulgarian Academy of Sciences

Introduction

Logbooks for astronomical observations

Logbooks contain metadata for plates, written by the observers at the time of observation.

Tasks:

- Link from a record of WFPDB to the corresponding logbook page image – valuable information in **Remark** section.
- Check the data in WFPDB for correctness – the data are inserted manually (from the keyboard)
- Insert new data in WFPDB directly (from logbook files).

Harvard College Observatory

22

INSTRUMENT, 8" ROSS-REI

No.	Class	Object	R. A.	Dec.	Started	Obs. H. A.	Obs. Dec.	Tel. E. or W.	Load	Focus	Prisms
✓ 1917	L	-	19 00	0	17 30	30E	0	-	88	-	-
✓ 1918	"	-	19 30	-60	19 02	28E	-60	-	110	-	-
✓ 1919	"	-	21 00	0	20 34	26E	0	-	90	-	-
✓ 1920	"	-	22 30	-60	22 06	24E	-60	-	110	-	-
✓ 1921	L	-	20 00	-45	20 15	15W	-45	-	104	-	-
✓ 1922	"	-	22 00	-45	21 47	13E	-45	-	105	-	-
✓ 1923	L	-	14 00	-35	14 09	09W	-35	-	125	-	-
✓ 1924	L	-	16 30	-60	15 41	0	49E	-60	-	110	-
✓ 1925	L	-	17 00	0	17 16	0	16W	0	-	88	-
✓ 1926	L	-	14 00	-45	14 55	0	55W	-45	-	103	-
✓ 1927	"	-	16 00	-45	16 27	0	27W	-45	-	104	-
✓ 1928	"	-	18 00	-45	17 59	0	01E	-45	-	104	-
✓ 1929	"	-	19 00	0	19 31	0	31W	0	-	88	-
✓ 1930	"	-	21 00	0	21 03	0	03W	0	-	88	-
✓ 1931	"	-	0 00	-35	22 30	1	25E	-75	-	118	-

23

DATE, July 17-18 (Fri.-Sat.) 1931

Sky at Start	Stopped	Exp. C'ds	C'k	Dew	Ref.	Obs'r	REMARKS
0	19 00	90	-	-	-	V	
0	20 32	90	-	-	-	V & S	
0	22 04	90	-	-	-	S	
0	23 36	90	-	-	-	S	Dew on lens.
Sun. Moon.							
July 19-20 (Sat.-Sun.) 1931							
1	21 45	90	-	-	-	V	Lightning right round in E.
0	23 17	90	-	-	-	V	Damp. lens wet with dew.
July 20-21 (Sun.-Mon.) 1931							
1	15 39	90	-	-	-	D	
2	17 11	90	-	-	-	D	Circi in South and West.
5	18 16	90	-	-	-	D	Circi forming, passing over, disappearing, forming again etc.
→ Wellen & continuous - they practically overcast with circi.							
July 21-22 (Tues.-Wed.) 1931							
2	16 25	90	-	-	-	V	blonds in N. Moon 6 ^d .
2	17 57	90	-	-	-	V	" " " "
1	19 29	90	-	-	-	V	" " " "
0	21 01	90	-	-	-	V	blonds 6 ^d Moon set 20.02
0	22 33	90	-	-	-	F	
0	24 06	90	-	-	-	F	

The Royal Observatory of Belgium, located in Uccle

A. Zone 1, 33

NUMERO du cliché.	ASCENSION DROITE approchée du centre.	B. D.	G.	α pour 1900,0.	δ pour 1900,0.	PROCESS
34	$5^h 1^m$	32. 290	9.1	$5^h 0^m 56.6$	$32^\circ 46' 7''$	+ 3.905
		33. 261	2.5	$2^m 36.7$	$33^\circ 7' 19''$	+ 3.912
35	$5^h 10^m$	32. 931	8.8	$9^m 57.4$	$33^\circ 3' 3''$	+ 3.922
36	$5^h 19^m$	33. 1040	2.9	$19^m 36.1$	$33^\circ 5' 52''$	+ 3.930
37	$5^h 28^m$	32. 1028	6.7	$26^m 54.2$	$32^\circ 43' 51''$	+ 3.923
38	$5^h 37^m$	33. 1134	7.7	$37^m 37.8$	$33^\circ 5' 16''$	+ 3.939
39	$5^h 46^m$	33. 1171	2.2	$45^m 9.4$	$33^\circ 19' 10''$	+ 3.950
40	$5^h 55^m$	33. 1209	6.9	$53^m 39.9$	$33^\circ 7' 47''$	+ 3.945
41	$6^h 4^m$	33. 1265	2.2	$6^h 4^m 12.2$	$33^\circ 0' 56''$	+ 3.941
42	$6^h 13^m$	33. 1292	2.8	$11^m 44.7$	$33^\circ 12' 23''$	+ 3.947
43	$6^h 22^m$	33. 1326 33. 1329	2.5 3.5	$21^m 28.3$ $21^m 50.38$	$33^\circ 0' 57''$ $33^\circ 13' 16''$	+ 3.932 + 3.944
44	$6^h 31^m$	33. 1367	2.8	$29^m 56.0$	$32^\circ 58' 56''$	+ 3.933

Leander Mc Cormick Observatory, Charlottesville, VA

					Setting	Mid Expt	Hand-Aux	Expt											
					E.S.T.														
38	May 14	Vega	18° 30'	+38° 42'	23	14 06	0 36 E	10	M	2-3	2-3	clear	66	20	-Blue Sky	g			
39	"	"	"	"	22	14 17	25	10	M							g			
40	"	β Cygni	19 34	+50 4	22	14 31	11 E	10	M							g	Trail β Cygni		
41	"	DM 6-4351 β Cygni	19 50	+60 58	23	15 42	0 0	10	M							g	Wrong field		
42	May 24	α Comae Ber	13 6	+17 57	12	16 17	0 26 E	10	M							g			
43	"	γ Virginis	12 38	-1 1	13	32	13 E	10	M							g			
44	May 25	"	"	"	16	6	6 W	12	M							g			
45	"	α Comae Ber	13 6	+17 57	17	16 19	0 19	12	M	1-2	2	Thin Clouds	68	20	" "	g			
46	"	γ Bootis	13 51	+18 48	20	3 37	36 E	24	M							N.G.	Cloud driving poorly		
47	"	Vega	18 34	+38 42	24	9 0	0 13	20	M	2	2					N.G.	Get new guiding star		
48	"	β Cygni	19 34	+50 4	20	9 37	0 28 W	15	M							g	Trail α Comae Ber		
49	"	"	"	"	21	9 52	0 43	15	M	1-2	2	Hazy	74	20	" "	g	Trail γ Bootis		
50	"	α Comae Ber	13 6	+17 57	21	22	0 49	20	M							g	Trail α Comae Ber		
51	May 30	γ Virginis	12 38	-1 1	23	10 38	1 28	20	M							p	Trail γ Bootis		
52	"	β Cygni	19 34	+50 4	5	11 22	29 W	15	M							g			
53	"	Vega	18 34	+38 42	6	37	1 44	15	M	2	2	Clear	68	20		g			
54	"	β Cygni	19 34	+50 4	22	14 3	0 32 E	10	M							g			
55	"	"	"	"	23	14 16	21	10	M							g			
56	"	α Comae Ber	13 6	+17 57	12	46	0 47	12	M							g			
57	"	γ Virginis	12 38	-1 1	13	14 59	34 E	12	M							g			
58	"	"	"	"	13	16	0 17	12	M							g			
59	"	β Cygni	19 34	+50 4	12	15 29	4 E	12	M							g			
60	"	α Comae Ber	13 6	+17 57	0	16 6	0 35 E	12	M							g			
61	May 31	γ Virginis	12 38	-1 1	1	18	23	12	M	1-2	2	Clouds	75	20	" "	f	Lamp burned out.		
62	"	β Cygni	19 34	+50 4	5	8 53	33	15	M							g			
63	"	Vega	18 34	+38 42	6	9 10	0 50 W	15	M							g			
64	"	α Comae Ber	13 6	+17 57	19	13 55	0 57 E	10	M	3	2	Clear	70	"		g			
65	"	γ Virginis	12 38	-1 1	20	14 7	45	10	M							g			
66	"	β Cygni	19 34	+50 4	20	14 25	0 27 E	10	M	3	3					g	Trail β Cygni		
67	"	"	"	"	19	35	17	10	M							g			
68	"	α Comae Ber	13 6	+17 57	12	15 2	0 15 E	10	M							g			
69	"	γ Virginis	12 38	-1 1	13	12	5	10	M							g			
70	"	β Cygni	19 34	+50 4	5	8 35	0 16 W	15	M	2	2	Thin Clouds	72	20	" "	f	Trail γ Virginis		
71	May 31	γ Virginis	12 38	-1 1	6	8 50	31	15	M							f			
72	"	"	"	"	6	9 12	0 53	15	M							f			
73	"	"	"	"	6	9 27	1 8 W	15	M							f			

Bonn Observatory

Platten Nr.	Datum	Objekt	Aufnahmesel. (WZ)	Stunden winkel	Fokus	Plattenart Filter	Exposition	Himmelsinstanz	Temp. Beob.	Bemerkungen
A	1970									
2439	November 23/24	NGC 7023	20 ¹³ -20 ⁴³	+3 ^h 50 ^m	4.6	103a-0 16x16 mm	Red. 1:20 7 min	Bo. Ci DS 2, R 3-4	+5°C	Flackung
2440	"	Per OB 2	21 ⁴⁷ -22 ¹⁷	-1 ^h 20 ^m	"	"	"	Bo. 0 DS 1, R 3	"	"
2441	"	NGC 225	22 ²⁰ -22 ⁵⁰	+2 ^h 09 ^m	"	"	"	"	"	"
2442	"	IC 1848	22 ⁵⁴ -23 ²⁴	+0 ^h 43 ^m	"	"	"	"	"	"
2443	November 24/25	NGC 225	23 ¹⁵ -23 ⁴⁵	+3 ^h 08 ^m	"	"	"	Bo. 0 (Ci?) DS 3-4, R 3	+6°C	"
2444	"	IC 1848	23 ⁵⁰ -00 ²⁰	+1 ^h 43 ^m	"	"	"	DS 3	"	"
2445	"	Per OB 2	00 ²⁵ -00 ⁵⁵	+1 ^h 24 ^m	"	"	"	"	"	"
2446	"	NGC 225	02 ¹⁶ -02 ³¹	+6 ^h 09 ^m	4.7!	"	"	Bo. 0 (Ci?) DS 1-2, R 3	+5°C	"
2447	"	NGC 225	02 ³⁴ -03 ⁰⁴	+6 ^h 27 ^m	4.6	"	"	" R 3-4	"	"
2448	"	Per OB 2	03 ⁰² -03 ³⁸	+4 ^h 07 ^m	"	"	"	"	"	"
2449	"	IC 1848	03 ⁴³ -04 ¹³	+5 ^h 37 ^m	"	"	"	" R 3	"	"
2450	November 25/26	Orion T1	17 ⁴⁰ -18 ¹⁰	+1 ^h 36 ^m	"	"	"	Bo. 0 DS 2 R 2	+9°C	"
2451	"	Lac 81	18 ¹³ -18 ⁴³	+0 ^h 20 ^m	"	"	"	"	"	"
2452	"	Cep OB 2	18 ⁴⁷ -19 ¹⁷	+1 ^h 50 ^m	"	"	"	"	"	"

abgebrochen wegen
falscher Fokussierung!

Overview of accessible logbooks

Observatory	WFPDB ID	Plates	Plates in WFPDB	Logbooks No	Data GB	Digitized Logbooks
Harvard	HAR	484668	111665	75	44	Yes
Potsdam	POT	8415	1661	11	1	Yes
Jena	JEN	1100	0	1	4	Yes
Bonn	BON	10530	0	9	0	No
Rozhen	ROZ	9152	9332	2	–	Partially
Bamberg	BAM	30535	24914	32	–	Partially
Vatican	VAT	7936	0	4	0	No
Copenhagen	COP	1364	0	1	0	No

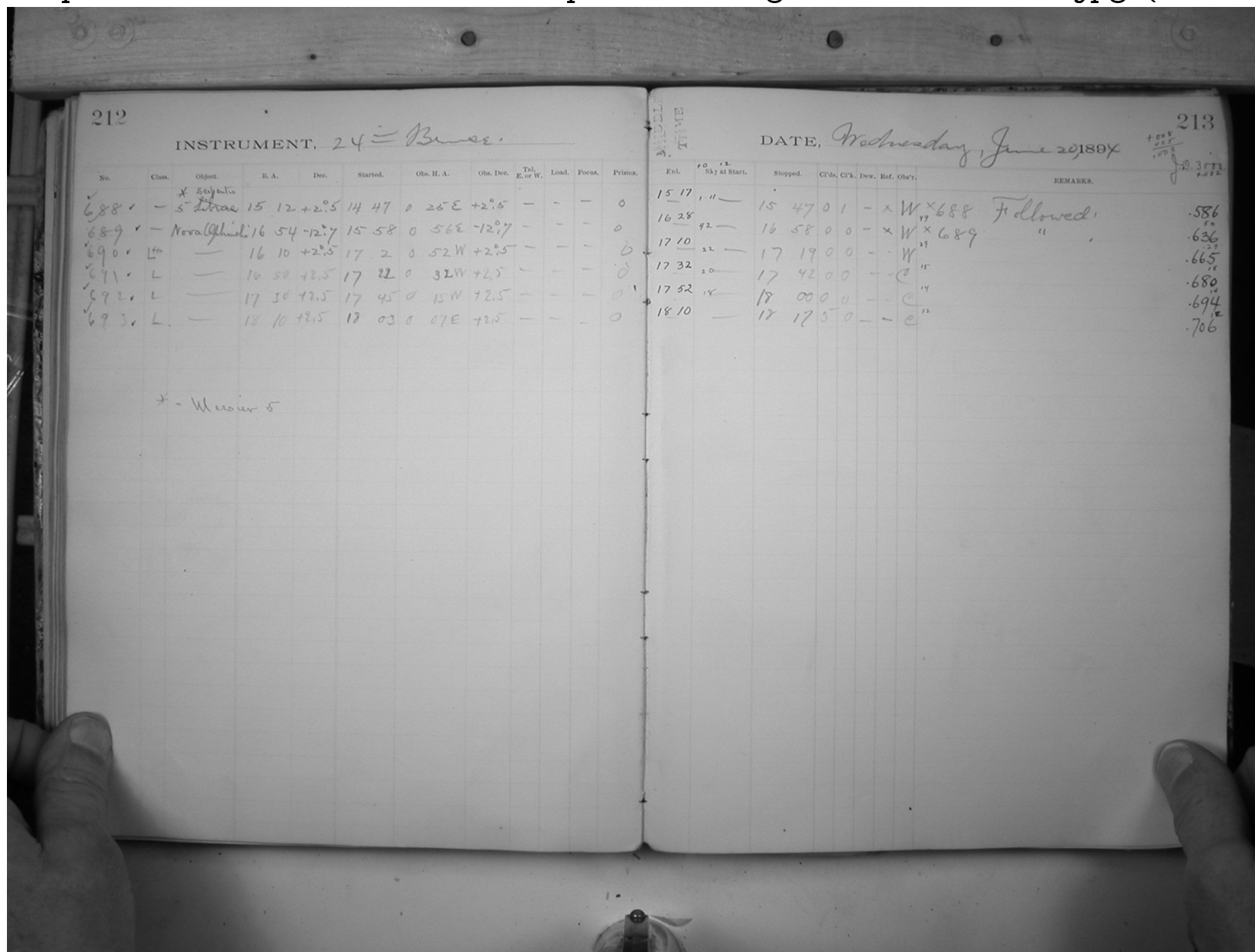
Methods and technologies for data extraction

- Handwritten recognition
 - Low quality of images
 - + Structured pages (tables)
 - + Number recognition
 - + Constraints in columns
-
1. Separation of the text region on the page
 2. Binarization, rotation
 3. Rows detection
 4. Columns and cells detection
 5. Numbers spotting
 6. Numbers recognition

As a result, a digital copy of the logbook will be produced!

Example: The Harvard College Observatory Astronomical Plate Stacks

http://tdc-www.cfa.harvard.edu/plates/a/logs/01b/a01b_0212.jpg (2048 x 1536 Pixels)



Rows detection

☐ Mark names

7 rows

Directory

File

Threshold

Left m.

Right m.

Bot.m.

Top m.

W.value

W.height

Row height

212

INSTRUMENT, 24 = Buser.

No.	Class.	Object.	R. A.	Dec.	Started.	Obs. H. A.	Obs. Dec.	Tel. E. or W.	Load.	Focus.	Prisms.	Enl.	+0, 1, 2 Sky at Start.	Stopped.	CT'ds.
✓ 688	✓ -	* Serpens	15	12	+2.5	14 47	0 26 E	+2.5	-	-	0	15 17	11	15 47	0
✓ 689	✓ -	Nova (Orion)	16	54	-12.7	15 58	0 56 E	-12.7	-	-	0	16 28	42	16 58	0
✓ 690	✓ L	—	16	10	+2.5	17 2	0 52 W	+2.5	-	-	0	17 10	22	17 19	0
✓ 691	✓ L	—	16	58	+2.5	17 22	0 32 W	+2.5	-	-	0	17 32	20	17 42	0
✓ 692	✓ L	—	17	30	+2.5	17 45	0 15 W	+2.5	-	-	0	17 52	18	18 00	0
✓ 693	✓ L	—	18	10	+2.5	18 03	0 07 E	+2.5	-	-	0	18 10	—	18 17	5

* = Wierier 5

Columns and sells determination

Mark names

Go Directory Harvard File p11112 Next Previous Save All: Go-Save

Threshold 142 Left m. 4 Right m. 1400 Bot.m. 50 Top m. 112 W.value 10 W.height 5 Row height 14

212

INSTRUMENT, 24 = Buser.

No.	Class.	Object	R. A.	Dec.	Start.	Obs. H. A.	Obs. Dec.	Tel. E. or W.	Load.
✓ 688	✓ -	* Scipatus Librae	15	12 + 2.5	14 47	0 26 E	+ 2.5	-	-
✓ 689	✓ -	Nova (Ghuel)	16	54 - 12.7	15 58	0 56 E	- 12.7	-	-
✓ 690	✓ L th	—	16	10 + 2.5	17 2	0 52 W	+ 2.5	-	-
✓ 691	✓ L	—	16	50 + 2.5	17 22	0 32 W	+ 2.5	-	-
✓ 692	✓ L	—	17	30 + 2.5	17 45	0 15 W	+ 2.5	-	-
✓ 693	✓ L	—	18	10 + 2.5	18 03	0 07 E	+ 2.5	-	-
* - Mercier 5									

Digital copy of the page

212

INSTRUMENT, 24 = Bessel.

No.	Class.	Object.	R. A.	Dec.	Started.	Obs. H. A.	Obs. Dec.	E
✓ 688	-	X Serpens 5 Librae	15 12	+2° 5'	14 47	0 25 E	+2° 5'	
✓ 689	-	Nova (Ghiesb.)	16 54	-12° 7'	15 58	0 56 E	-12° 7'	

DATE, Wednesday, June 20, 1894

Enl.	¹⁰ 12 Sky at Start.	Stopped.	Cl's.	Cl'k.	Dew.	Ref.	Obs'r.	REMARKS.
15 17	"	15 47	0	1	-	X	W ¹⁷ X 688	Followed.
16 28	42	16 58	0	0	-	X	W ²¹ X 689	"
17 10	22	17 19	0	0	-	X	W ²¹	

INSTRUMENT 24 = B.... DATE Wednesday, June 20, 1894

No.	Class.	Object.	R.A.	Dec.	Started.	Obs.H.A.	Obs.Dec.	Enl.	Stopped.
688	-	5 ...	15 12	+2^o.5	14 47	0 25e	+2^o.5	15 17	15 47
689	-	Nova...	16 54	-12^o.7	15 58	0 56e	-12^o.7	16 28	16 58

Connection to Wide-Field Plate Database – Sofia

Link to a page

One-to-one correspondence:

WFPDB Instrument $\leftrightarrow$ Logbook (a set of page images)

Index of /plates/a/logs/01b

	Name	Last modified	Size
[DIR]	Parent Directory	07-Apr-2007 09:24	-
[IMG]	a01b_0000.jpg	05-Apr-2007 16:33	675k
[IMG]	a01b_0001.jpg	05-Apr-2007 16:33	557k
[IMG]	a01b_0002.jpg	05-Apr-2007 16:33	557k
[IMG]	a01b_0004.jpg	05-Apr-2007 16:33	546k
[IMG]	a01b_0006.jpg	05-Apr-2007 16:33	546k
[IMG]	a01b_0008.jpg	05-Apr-2007 16:33	560k
[IMG]	a01b_0010.jpg	05-Apr-2007 16:33	573k
[IMG]	a01b_0012.jpg	05-Apr-2007 16:33	566k
[IMG]	a01b_0014.jpg	05-Apr-2007 16:33	578k

...

Instrument $\equiv$ (observatory, telescope, site)

<http://tdc-www.cfa.harvard.edu/plates/a/>



**Telescope
Data Center**
SMITHSONIAN ASTROPHYSICAL OBSERVATORY



Harvard-Smithsonian Center for Astrophysics

Harvard Observatory Plate Stacks A Series Characteristics

Basics

Plate Size	Plate Scale	Telescope	Focal Length
8x10, 14x17	60"/mm	24-inch Bruce Doublet	135 inches

Observatory

Plate nos.	Location	Dates
1- 1730	Cambridge, Mass.	1893-11-09 - 1895-10-21
1731-14110	Arequipa, Peru	1896-04-02 - 1926-11-06
14111-27504	Bloemfontein, South Africa	1929-07-15 - 1950-10-11

Instrument $\equiv$ (observatory, telescope, site)

Wide-Field Plate Database – Sofia

Details for archive: HAR061A

<i>Location of the Archive:</i>	<i>Clear aperture:</i> 0.61 m
Site: Cambridge	<i>Mirror diameter:</i>
Country: USA MA	<i>Focal length:</i> 3.34 m
<i>Observatory:</i>	<i>Scale:</i> 60 "/mm
Name: Harvard	<i>Type:</i> Rfr
Site: Cambridge	<i>Field size:</i> 7.2°
Country: USA MA	<i>Years of operation:</i>
<i>Time zone:</i> -5 h	From: 1893
<i>East longitude:</i> -71° 52.2'	To: 1895
<i>Latitude:</i> 42° 22.8'	<i>P/F:</i>
<i>Altitude:</i> 24 m	

Many-to-one: WFPDB record (plate) → Logbook page

INSTRUMENT 24 = B.... DATE Wednesday, June 20, 1894

No.	Class.	Object.	R.A.	Dec.	Started.	Obs.H.A.	Obs.Dec.	Enl.	Stopped.
688	-	5 ...	15 12	+2 ⁰ .5	14 47	0 25e	+2 ⁰ .5	15 17	15 47
689	-	Nova...	16 54	-12 ⁰ .7	15 58	0 56e	-12 ⁰ .7	16 28	16 58

Add a list of plate numbers to a page (manually or as a result of first column recognition)

page 212 → 688-693 (6 records)

catalog Harvard A Series Plates, equinox 1900, radecsyz B1900

id	ra	dec	exp	epoch
00688	15:12	+02 : 50	60	1894-06-21T02:03:50.4
00689	16:54	-12 : 70	60	1894-06-21T03:15:50.4

Wide-Field Plate Database – Sofia

IDobs	IDins	IDSuf1	IDno	RAJ2000	DECJ2000	Date	Time
HAR	061	A	000688	15 17 02	02 27 55	1894 06 21	02 03 50
HAR	061	A	000689	16 59 37	-13 19 05	1894 06 21	03 15 50

ID of a plate in WFPDB: [IDobs IDins IDSuf1 IDno]

page number is contained in the image filename

[HAR 061 A 000688] → http://tdc-www.cfa.harvard.edu/plates/a/logs/01b/a01b_0212.jpg

Numbers recognition

Interactive software system specialized for logbook recognition.

Conclusions

- Direct access to digitized logbooks of astronomical Wide-Field Plate Database – Sofia
plate → image of logbook page: manually or automatic (first column numbers recognition)
- OCR of the numbers in logbooks allows automatic insertion of new records in WFPDB

This talk has been supported by Grant No. DO02-275/2008
Astroinformatics, Bulgarian NSF, Ministry of Education and
Science.

Thank you for your attention.