

The Harvard Plate Stack Scanning Project



Doug Mink	Software and archive	<i>Smithsonian Astrophysical Observatory</i>
Alison Doane	Plate Curator	<i>Harvard College Observatory</i>
Bob Simcoe	Digitizer Design	<i>Harvard College Observatory affiliate</i>
Ed Los	Digitizer Interface	<i>Harvard College Observatory affiliate</i>
Josh Grindlay	Principal Investigator	<i>Harvard University Astronomy Department</i>

Presented at the Virtual Observatory Workshop, Sofia, Bulgaria, 27 April 2005

Presentation Outline

- What We have
500,000 astronomical plates
- What We've Done
Digitizing catalogs, trying scanners
- What We're Doing
3-year NSF grant to build a scanner and digitize 10,000 plates
- What We Want To Do
Digitize them all and serve them on the Web

What We Have

- Harvard's plates contain the most complete sky coverage of both the northern and southern sky over the longest time period - 1880 to 1989
- Harvard Observatory now has 500,000+ photographs, by far the largest collection and 25% of the world's total.
- Of those, between 250,000 and 350,000 will be useful for photometry and astrometry.

What We Have

These are the plates which we think are worth scanning

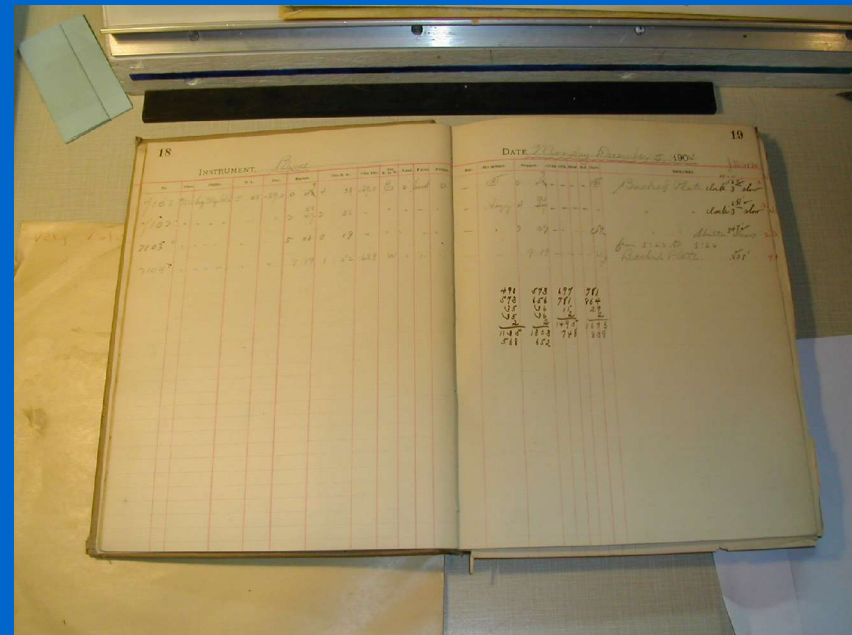
Series	Total	Aperture (in.)	Scale "/mm	N/S	Years	Mag. Limit	Q
A	27504	24	60	S	1893-1950	18	5
ADH	7067	30	68	S	1950-1963	18-19	4
AM,AC, ...	75000	1.5	600	N/S	1898-1957	13-14	3
B	76874	8	179	S	1885-1954	17	4
BR	4176	8	209	S	1938-1944	17	4
DNB,DSB	9000	1.6	580	N/S	1962-1989	15	5
I	59246	8	163	N	1889-1946	17	4
MA	11737	12	97	N	1905-1983	17-18	5
MC	40596	16	98	N	1909-1992	17-18	5
MF	40897	10	167	S	1915-1955	17	4
RH,RB	33000	3	391	N/S	1928-1963	15	3
Total	385097						

What We Have Done

- Created digital catalogs of about 100,000 plates
- Tested scanners on 8x10-inch plates
- Created a web site with access to plate stack catalogs and current scanned plates

First: Digitize Metadata

From hand-written cards and logbooks



Web Access to Digital Plate Stack Catalogs



Telescope
SMITHSONIAN ASTROPHYSICAL OBSERVATORY
Data Center



Harvard Plate Stacks

A Series Search

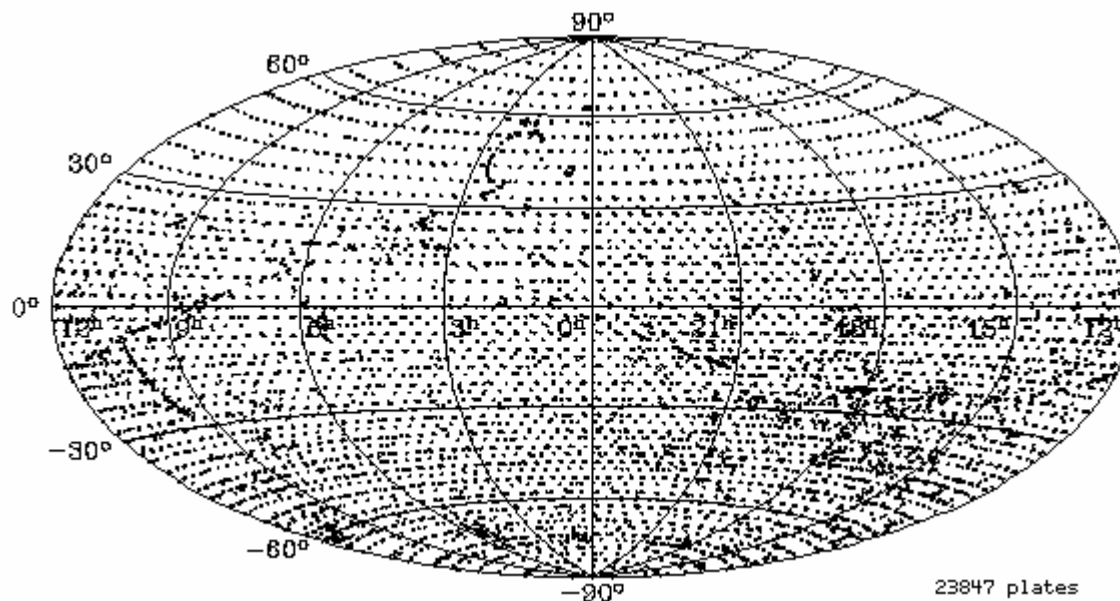
Name 5-digit Sequence
or Object Name

J2000 Right Ascension Declination
(hh:mm:ss.sss) (dd:mm:ss.sss)


Search Starting date Ending date
(year or ISO date) (year or ISO date)

Click on map to see plates from 1890's


· Harvard Observatory A plate centers (aplates)



Search on sky position and time



Telescope Data Center
SMITHSONIAN ASTROPHYSICAL OBSERVATORY



Harvard Plate Stacks
A Series Search

Name 5-digit Sequence or Object Name

J2000 Right Ascension (hh:mm:ss.sss) Declination (dd:mm:ss.sss)

Starting date (year or ISO date) Ending date (year or ISO date)

Searching for plates containing 08:40:22.198 +19:40:19.43 J2000 from 1890 to 1910

Plate	RA2000	Dec2000	Exp	Epoch	Arcsec
0045	08:34:52.827	+19:55:38.21	18.00	1893-11-26	4738.41
0147	08:33:53.101	+20:07:41.64	59.00	1894-01-02	5728.35
0194	08:33:53.101	+20:07:41.64	16.00	1894-01-26	5728.35
0196	08:33:52.997	+20:01:41.64	60.00	1894-01-31	5638.83
0207	08:32:53.272	+20:13:45.08	15.00	1894-02-02	6639.87
0244	08:32:53.272	+20:13:45.08	10.00	1894-02-17	6639.87
0263	08:33:52.789	+19:49:41.65	14.00	1894-02-25	5526.18
0265	08:49:53.978	+22:06:48.14	21.00	1894-02-25	11892.47
0326	08:52:49.101	+17:18:38.58	13.00	1894-03-08	13606.37
0537	08:32:50.291	+17:19:45.16	10.00	1894-04-18	10604.32
0549	08:35:53.070	+20:13:34.78	60.00	1894-04-26	4287.29
1252	08:47:54.341	+22:18:54.64	20.00	1894-12-24	11429.00
1280	08:31:53.338	+20:13:48.54	68.00	1895-01-30	7450.71
1285	08:35:53.070	+20:13:34.78	60.00	1895-02-15	4287.29
6639	08:32:50.291	+17:19:45.16	10.00	1904-04-05	10604.32
6661	08:52:49.101	+17:18:38.58	10.00	1904-04-16	13606.37
6667	08:52:53.935	+22:18:38.47	10.00	1904-04-17	14178.60
6708	08:32:50.291	+17:19:45.16	10.00	1904-05-03	10604.32
7203	08:32:55.499	+22:19:45.02	60.00	1905-03-07	11428.92
7205	08:32:55.499	+22:19:45.02	10.00	1905-03-07	11428.92
9236	08:52:49.101	+17:18:38.58	62.00	1909-05-13	13606.37

Next: Digital access to image data

Moving the plates out of the 20th century



Scanner Testing

CreoScitex EverSmart \$35,000-\$50,000

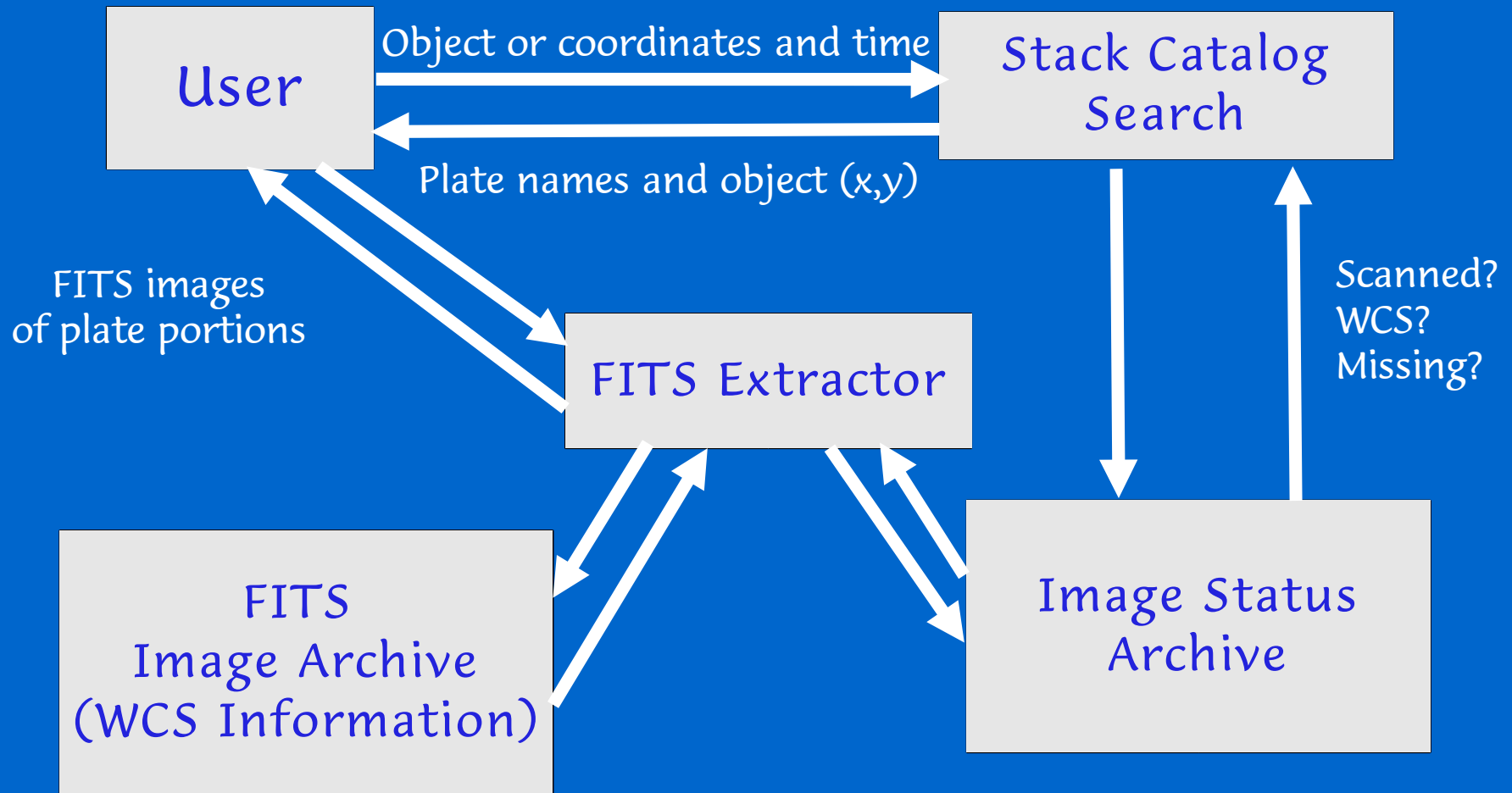
- High resolution (4800 dpi), but crashed a lot
- 20 minutes per scan

UMAX PowerLook 3000 \$5,000

- Adequate resolution (1200 dpi)
- Reliable, for a while
- Under 10 minutes per scan

We bought a Umax and scanned about 100 plates, 40 of them from the MC series containing the open cluster M44 for a project with Leonid Berdnikov, then it broke.

Access to digital images



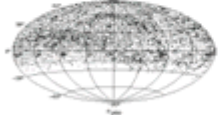
Access to digital images

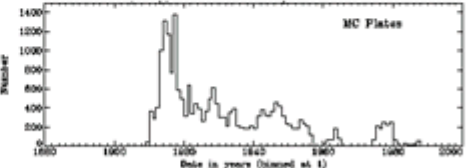
File Edit View Go Bookmarks Tools Help
http://tdc-www.harvard.edu/plates/mcseries.html
Internet Lookup New&Cool Google AstroResources MassBike SAO TDC Weather Harvard Plates

Telescope Data Center
Harvard-Smithsonian Center for Astrophysics
Harvard Observatory Plate Stacks
MC Series Characteristics

Basics
Plate Size Plate Scale Telescope Focal Length
8x10, 4x5 98"/mm 16-inch Metcalf Doublet 83 inches

Observatory
Plate nos. Location Dates
1-26093 Cambridge 1909-06-22 - 1932-05-28
26094-40591 Oak Ridge 1932-09-28 - 1988-12-12

[Online Searchable catalog](#)

[Maps of spatial and temporal distribution](#)


[Histogram by year](#)

ASCH Catalog

- [MC Series tab-separated table \(number-sorted B1900 Starbase MJD\)](#)
- [MC Series tab-separated table \(number-sorted B1900 Starbase ISO date\)](#)
- [MC Series tab-separated table \(number-sorted J2000 Starbase ISO date\)](#)
- [MC Series tab-separated table \(RA-sorted B1900 Starbase MJD\)](#)
- [MC Series tab-separated table \(RA-sorted B1900 Starbase ISO date\)](#)
- [MC Series tab-separated table \(RA-sorted J2000 Starbase ISO date\)](#)

Location in Stacks
Plate Size Location Numbers
8x10 first floor all
4x5 first floor all

Card Catalog: complete; there is also an Approach database of the blue plates.

Record Books: (49) 1-37502 (Catalog)
(2) MC Important Regions
(2) MC Long Period Variables
(1) MC Photometric Work

Record Sheets: 26154-40591

Comments: This instrument, the most powerful refractor in the northern station up to 1936, was especially valuable for faint objects where the large scale was of benefit. When the plates are used for photometric purposes, it is well to take precautions, since the color coefficient is large and the distance correction is both large and variable from one plate to another.

The instrument was refigured after about 3500 plates had been taken; the distance correction was thereby diminished, but the color coefficient was practically unchanged. In general, it is well to avoid plates taken before the refiguring for accurate photometric work. All plates subsequent to 4171 were taken with the curved plate holder, which diminishes the distance correction due to curvature of field.

Done


MC Series description page with spatial and temporal distribution

Access to digital images

File Edit View Go Bookmarks Tools Help

http://tdc-www.harvard.edu/cgi-bin/plates/mcsearch

Internet Lookup New&Cool Google AstroResources MassBike SAO TDC Weather Harvard Plates

 **Telescope Data Center** **Harvard Plate Stacks** [A Series \(about\)](#) [MC Series \(about\)](#) [MF Series \(about\)](#)

MC Series Search

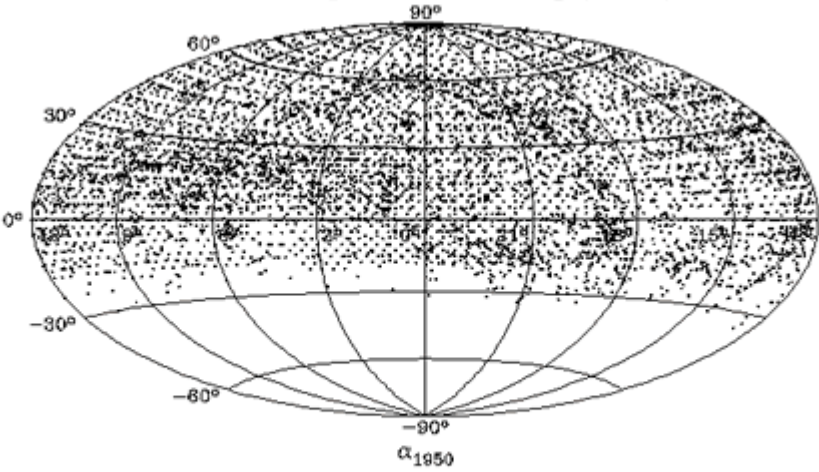
Name 5-digit Sequence or Object Name m44 X pixels to extract Y pixels to extract

J2000 Right Ascension (hh:mm:ss.sss) Declination (dd:mm:ss.sss) Extracted Image Type JPEG

Search Starting Date (yyyy-mm-dd or frac. year) 1950-01-01 Ending Date (yyyy-mm-dd or frac. year) 1980-01-01

Click on map to see MC plates from 1909-1919

· Harvard MC plate stack catalog (mebra)



Done

Search page for MC series

Limits may be set in space and time

Access to digital images

Telescope Data Center
Harvard Plate Stacks
MC Series Search

Name: m44
X pixels to extract:
Y pixels to extract:
J2000 Right Ascension (hh:mm:ss.sss):
Declination (dd:mm:ss.sss):
Extracted Image Type: FITS
Starting Date (yyyy-mm-dd or frac. year): 1950-01-01
Ending Date (yyyy-mm-dd or frac. year): 1980-01-01

Searching for plates containing m44 = 08:40:22.198 +19:40:19.43 J2000 from NED from 1950-01-01 to 1980-01-01

Plate	RA2000	Dec2000	Exp	Epoch	Arcsec	Comment
38526	08:50:51.624	+19:48:44.96	10.50	1961-05-04	8900.71	Praesepe_Grating_#
38527	08:40:52.321	+19:49:17.87	10.00	1961-05-04	686.14	Praesepe_Grating_#
38529	08:40:52.321	+19:49:17.87	12.00	1961-05-05	686.14	M44_Grating_#2301
38531	08:40:52.321	+19:49:17.87	8.00	1961-05-18	686.14	Praesepe_M44
38531	08:32:52.854	+19:49:45.09	8.00	1961-05-18	6368.78	M44_Grating_#2301
38532	08:40:52.321	+19:49:17.87	1.50	1961-05-18	686.14	M44_Grating_#2301
39048	08:40:52.321	+19:49:17.87	7.00	1964-04-12	686.14	Praesepe_NGC_2632 scanned plate
39053	08:40:52.321	+19:49:17.87	5.00	1964-04-17	686.14	Praesepe_NGC_2032
39237	08:38:52.338	+19:42:24.61	15.00	1975-01-18	1275.25	M44_clear scanned plate
39715	08:40:51.000	+18:30:17.91	22.00	1977-05-13	4221.31	Praesepe_no_filter scanned plate
39927	08:50:49.699	+17:48:45.00	30.00	1978-02-28	11146.90	Patrol_no_filter scanned plate
39927	08:50:49.874	+17:59:45.00	30.00	1978-02-28	10761.59	No_filter scanned plate
39951	08:38:52.644	+20:00:24.60	30.00	1978-03-10	1746.16	M44_no_filter
39963	08:40:52.203	+19:42:17.88	29.00	1978-04-10	440.01	M44_no_filter scanned plate

Results of search for MC plates from 1950-1989 on which M44 appears

Plates which have been scanned may be displayed

Access to digital images

File Edit View Go Bookmarks Tools Help

http://tdc-www.harvard.edu/cgi-bin/plates/mcplate?rfn=39963&dx=&dy=&mode=FITS&wcs

Internet Lookup New&Cool Google AstroResources MassBike SAO TDC Weather Harvard Plates

 **Telescope Data Center**
SMITHSONIAN ASTROPHYSICAL OBSERVATORY

[Click here for 1/20th scale full FITS image](#)
Click on image for x FITS image.

Harvard Plate Stacks

[A Series \(about\)](#)
[MC Series \(about\)](#)
[MF Series \(about\)](#)

MC Series Plate

MC39963

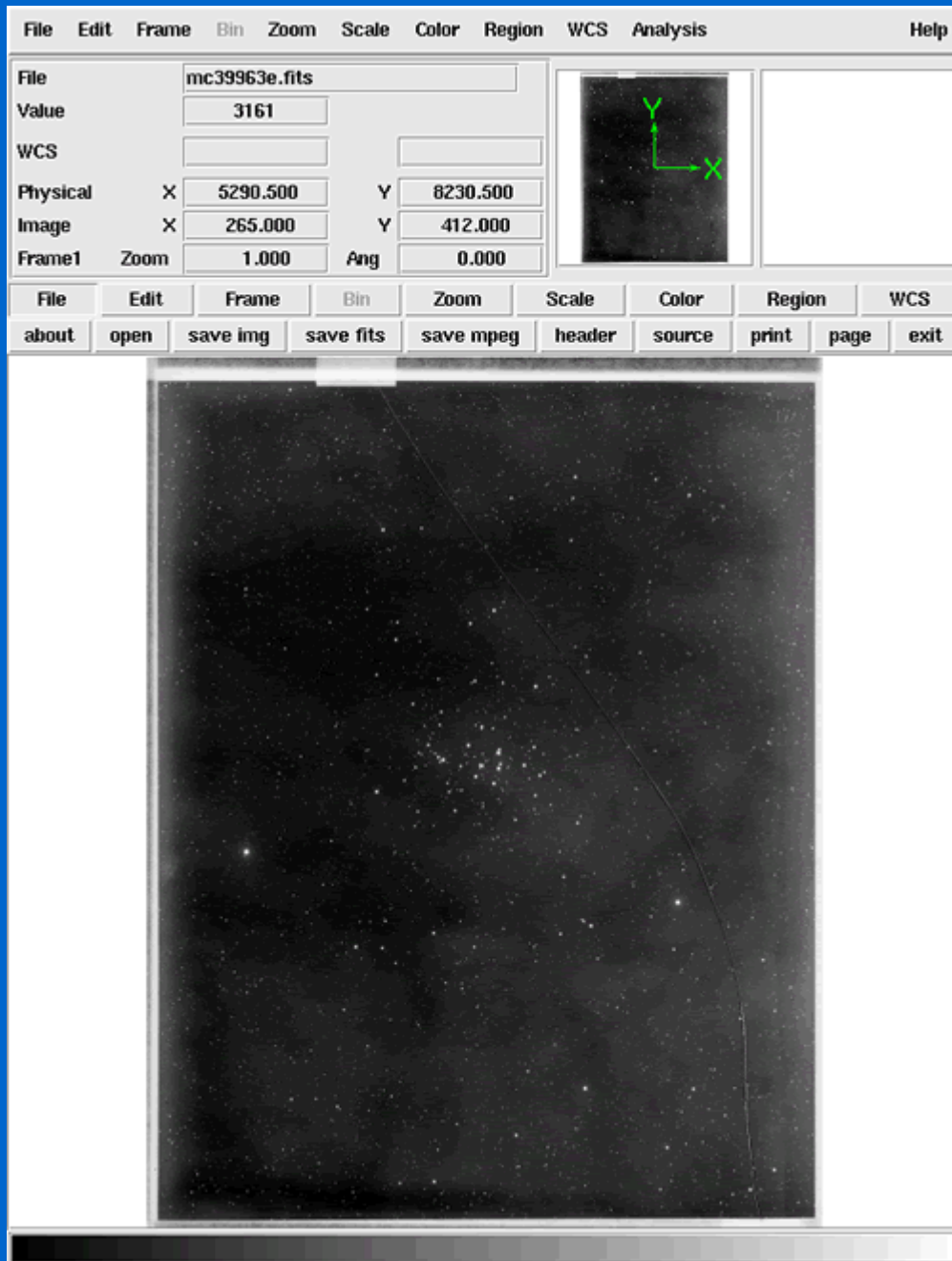
Plate	RA2000	Dec2000	Exp	Epoch	Note
39963	08:40:52.220	+19:43:17.88	29.00	1978-04-10	M44_no_filter



Done

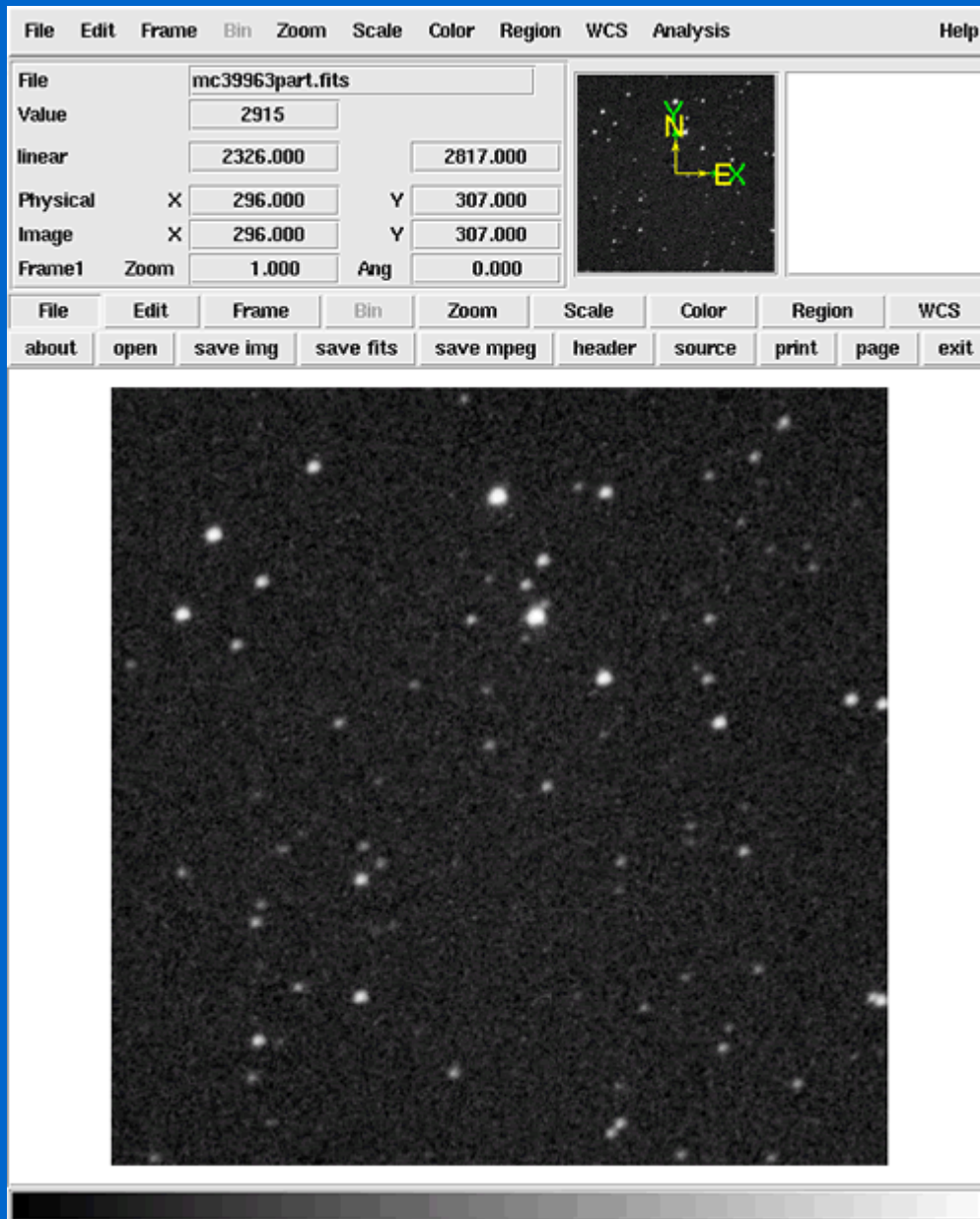
View of entire plate with each pixel representing 400 (20x20) pixels in the scanned image

Access to digital images



The same 1/20 scale image may be downloaded as a FITS image to an image browser such as ds9.

Access to digital images



By clicking on any position in the web display of the entire plate, a portion with dimensions defined on the search page is returned as a FITS file displayable in ds9 or a JPEG file at the full scanned image resolution.

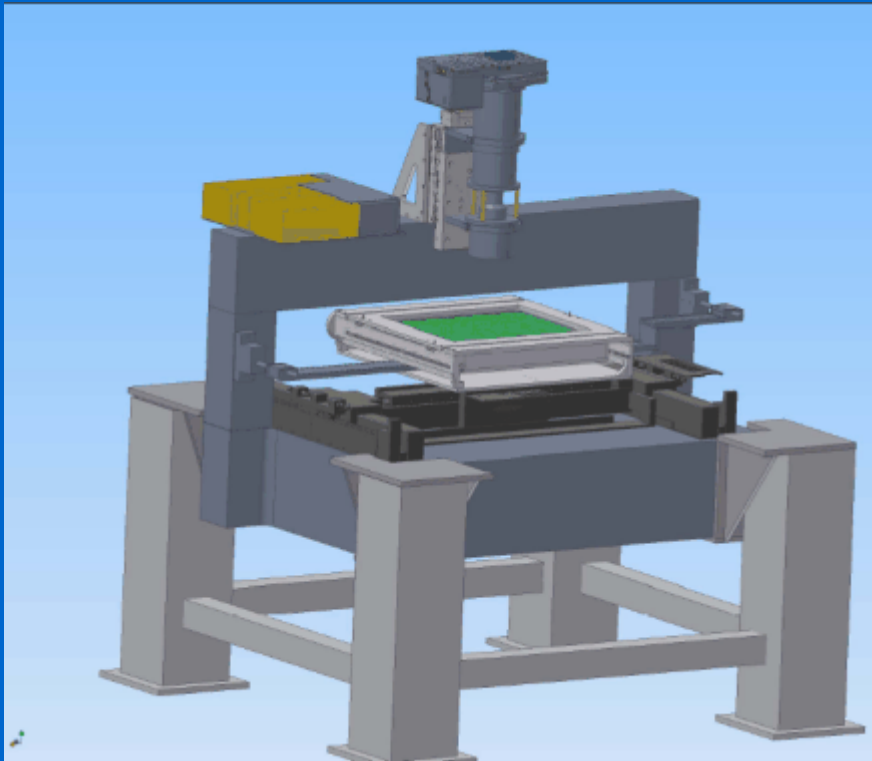
What We Are Doing

- To scan Harvard's library of historic plates in a 3-5 year time frame, we needed a machine that can scan 200 times faster than machines designed 20+ years ago, such as the USNO PMM, which took ~1- 4 hours to scan a single 14 x 14 inch plate.
- To meet astrometric, photometric, and archival goals, the machine needs sub-micron positional accuracy, at least 12 bits of photometric density range, and a scan speed that allows human handling to limit the average plate processing time.

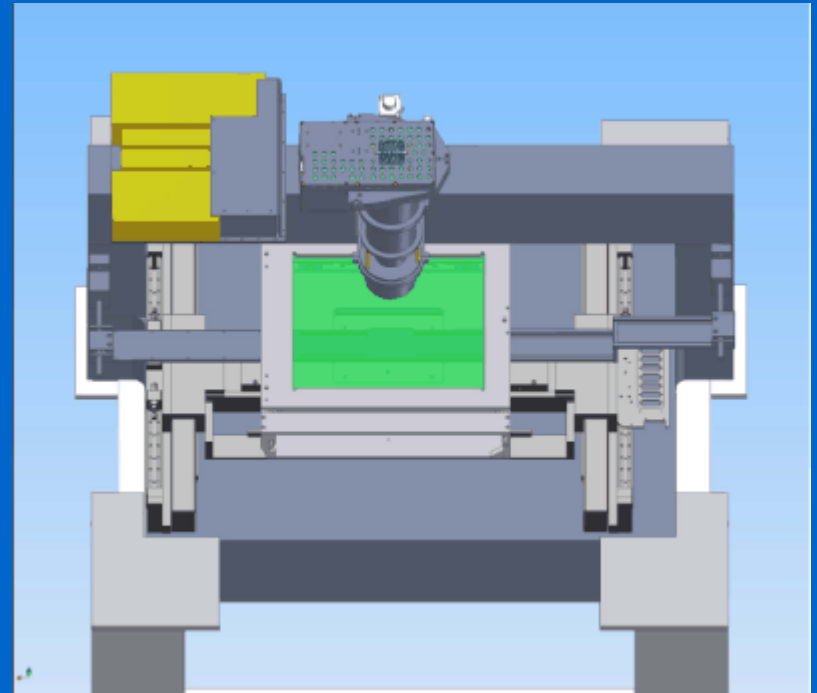
What We Are Doing

- Using technology common to semiconductor wafer and flat panel display inspection stations, a machine can be built today that can do ultra fast, ultra precise scanning.
- The 4Kx4K Summit digitizing camera with 12-bit 11- μ m pixels will be moved across the image, taking slightly overlapping exposures.
- An 8 x 10 inch plate will be scanned in about 20 seconds (though we will probably do 2 at once).
- A 14 x 17 inch plate will be scanned in a little over a minute, generating enough data in that time to fill a DVD (3 Gigabytes).

Future Harvard Scanner with 14x17-inch photographic plate



Harvard Plate Digitizer side view
(massive legs are air-bearing shock
isolation support system)



Harvard Plate Digitizer top view

Harvard Digitizer Reality

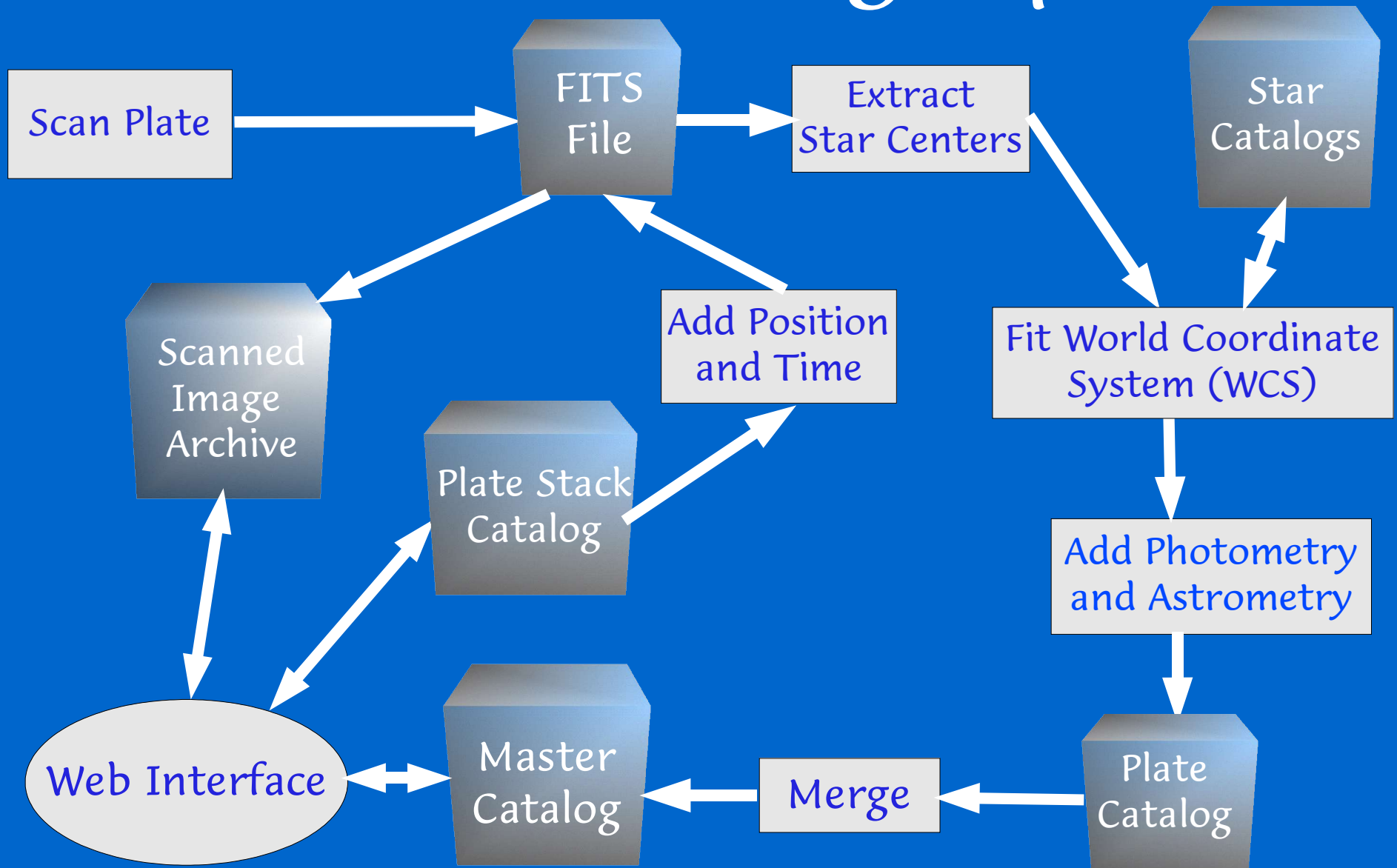


Harvard Plate Digitizer table
at factory last week



Harvard Plate Digitizer room
awaiting digitizer delivery

Plate Scanning Pipeline



But the plates aren't all

- To ease plate processing, we need to have digital catalogs before we scan plates. We have a >100,000 plate head start, but the rest of the plate stacks need to be digitally catalogued in a more timely way.
- To outsource catalog transcription, we need to convert existing observing log books to digital images, a fairly major project
- So two additional sub-projects need funding:
Telescope log imaging and catalog transcription

Catalog Digitization

- Existing digital catalogs have been created very slowly over the past 15 years by part-time plate stack staff.
- Digital catalogs are necessary for the plate digitizing pipeline, so the process must be accelerated
- We are applying for grants to fund catalog transcription, but to make it easier, we also need:

Telescope Log Digitization

- Since many of the older log books are too fragile to travel and transcription in Cambridge is likely to be too expensive, they should be imaged into a digital format
- Log books contain information which may not fit or have been processed into existing digital catalogs.
- While we need funding to image older, fragile logbooks, we can have the looseleaf logbooks of the last half-century scanned at the CfA.

What We Will Do

- Integrate and test our newly-developed high speed digitizer
- Digitize ~10,000 plates for demonstration science (e.g. PG Quasar variability survey)
- Complete digitization of card catalogs and scan telescope log books to enable optimum scan strategy for complete (~500,000) plate collection
- Develop plan for both local and remote serving the entire ~300Tb dataset
- Seek funding and/or donors to support the complete digitization project.