ST-ECF Participation in the GSC-II Generation Project

B. PIRENNE, ST-ECF; B. MCLEAN, B. LASKER, STScI

1. Background and Rationale for the GSC-II Project

There are a number of motivations for scanning and cataloguing photographic plates: first of all, the release of a 2-billion-object catalogue would generate uncountable science and mission-support projects which would be more than welcome for HST and future space missions as well as ground-based telescopes. The significance of this project has already been acknowledged with the formation of an international consortium to proceed with its construction with currently available resources. These partners include STScI, CRA (Osservatorio di Torino), ESO and GEMINI. Additional support for the digitisation of the plates (DSS-II) is also being received from ESO, CDS, CADC, NAOJ and the AAO.

The HST Advanced Camera (ACS), poised to be launched at the end of 1999, has in particular "bright" object avoidance scheduling constraints. The Space Telescope Science Institute is currently performing similar bright object checks for observations using the STIS MAMA detectors, but this is being done manually, and is laborious and time consuming. These could be greatly simplified and automated if a homogeneous, deep, complete catalogue down to magnitude 18 were available. Whilst the final requirements for the operation of the future HST COS (the selected Servicing Mission SM-4 replacement instrument scheduled for 2002) are still being determined, it is known that it has a small 2″ acquisition aperture which may be affected by the proper motion distribution of the guide stars. Hence the availability of guide star motions, or even the positions from the second epoch plates, can simplify the process and reduce the acquisition failure rate. The planned GSC-II would meet all these requirements but its currently scheduled delivery time of 12/2000 will not meet the current ACS launch constraint. We feel that the GSC-II is one essential component in reducing the operations cost of HST.

In order to expedite GSC-II availability, we have put together an action plan with extra resources contributed to the project so that a preliminary deep all-sky catalogue (without proper motions) can be completed around Q4 1999 to support bright-object protection checks.

2. Generation of the GSC-II

The generation of the GSC-II catalogue consists of a number of steps, simple in concept, but heavy in operational burden.

2.1 Scanning

2.1.1 Scan all 3576 plates from the second-generation Schmidt surveys

(POSS-II in the Northern Hemisphere and the AAO-SES in the south). The sampling is 15 micron (1 arcsec) which results in 23,000 pixel square raster images of 1.1 GB byte each. This represents almost 4 Terabytes of data in total. This first step has been done at the Space Telescope Science Institute and is now virtually complete. A lightly compressed version is now being distributed to a small set of selected sites, including ESO (Garching), where the scanned plates are available via the WWW: http://archive.eso.org/dss/dss

2.1.2 Combination of old and new plate scans

Combine with the first-generation survey plates that STScI has already scanned with a resolution of 25 microns (1.7 arcsec pixels) for the GSC-I and DSS-I projects. [Note that eventually STScI will re-scan some of these plates at 15 microns but this is of lower priority]. These 2390 plates represent another 1TB of data.

2.1.3 Scan old plates

Scan at lower priority the 894 POSS-I O plates that were not used in the GSC-I project.

2.2 Plate Processing

The second step involves a pipeline processing on single digitised plates where object detection (with de-blending) and preliminary calibration is taking place on-line. Results are stored in an object-oriented database management system which will reside on an HSM (Hierarchical Storage Management) controlled mass-storage system. Quality control happens throughout this phase as well. This is the most resource intensive part of the production work: The massive amounts of scan data are retrieved from the CASB image archive, processed extensively and the results saved and loaded into the COMPASS database. Estimates for the size of the database range from 4 to 8TB.

2.3 Catalogue Construction

2.3.1 Analysis of the calibrated object parameters

This step involves an analysis of the calibrated object parameters in the database to quantify the systematic errors and to recalibrate the derived parameters without a major reprocessing of the original scan data.

2.3.2 Export Catalogue

Derive exported catalogue from the large object database after merging overlapping plates and plates of the same field but of different colours. The new export catalogue will be compatible with the ESO SKYCAT interface (see http://www.archive.org/ skycat), and many other (web) interfaces can be expected to be available to retrieve catalogue data. The storage possibilities of a large table have been examined in [Wicenec 1996].

3. Participation of the ST-ECF

The ST-ECF has extensive experience in manipulating and processing large volumes of scientific data. It has strong connections with the STScI thanks to its HST support mandate in Europe. In addition, the location of the group inside ESO (another patron of the GSC-II project) as well as its proximity to the other European partner has motivated the ST-ECF to provide direct help to the existing collaboration.

At this stage in the project, the software development is essentially over. Production has already started in March. The best use of ST-ECF resources is to assist with the operation of the massive pipeline processing in order to accelerate the availability of the data for HST operations after Servicing Mission 3.

3.1 Interfaces

The geographical separation of the three sites where processing is taking place (STScI, ST-ECF and Osservatorio di Torino) implies the definition of reliable data transfer interfaces. The volume of data involved immediately rules out any on-line electronic data transfer of the image data using such means as FTP. Therefore, airmail shipment of media with the actual data to be processed and their results has been organised.

Procedures for problem reporting and bug fixes have been set up in such a way that the pipeline does not stay idle for long periods of time. Proper training of the operation and quality control staff took place.

3.2 Resources required

In order to meet the target date for a preliminary single-epoch catalogue by HST SM-3, it is necessary to approximately double the plate-processing rate that STScI had previously planned. This implies that ST-ECF has effectively duplicated the hardware and manpower that STScI currently has dedicated to the operation of the plate pipeline.

If the above production rate can be sustained, it means a nominal production time of about one year. To this, a large amount of time for manual re-processing of plates failing the processing step for one reason or another and extra quality control/bug fixes has to be added.

The pipeline software currently runs on DEC Alpha Open VMS systems and is composed of C, C++, FORTRAN and IDL code. Therefore, the ST-ECF has acquired an up-to-date DEC Alpha server 1200 with 2 $\times$ 500 MHz processors, 1 GB of memory and about 100 GB of disk space. A large amount of cassettes for the data transfer has also been acquired.

Operating the pipeline requires manpower to do operations such as loading the plate tapes into readers, write the output on media but also – and more importantly – perform quality control of the object extraction results. The quality control flags bad/doubtful plates, which are forwarded to the science team in Torino for investigation. For this resource, two full-time employees have joined the ST-ECF archive for two years (Nathalie Fourniol, previously in Strasbourg and Roberto Mignani formerly at the Max- Plank Institut für Extraterrestrische Physik).

4. Conclusion

In being involved in the GSC-II project, the ST-ECF is actively taking part in one of the major astronomy achievements of the decade. In a two-year effort, our contribution will bring a more timely delivery of the (first version) of an all-sky, 2-billion-object catalogue complete to beyond magnitude 18, available just in time for the next millennium.

B. Pirenne
bpirenne@eso.org

HST Archive Services Implemented in Java

M. DOLENSKY, A. MICOL, B. PIRENNE, ST-ECF

Abstract

In order to facilitate archive data selection and basic data analysis, a number of Java Applets only requiring a common web browser are now complementing the HST Archive [3].

This article discusses various applets which are already part of the archive web interface. These applets display and manipulate FITS images as well as spectra taken with HST. A generic plot utility is also used to present a set of pointing and specialised engineering data, called jitter files [4].

Spectral

This applet is a previewer for HST spectra. It is integrated into the WDB web interface (Fig. 1) and offers various options to inspect spectra with the mouse and by means of hotkeys. The screenshots in this paper show, that a standard web browser like Netscape 3 or Internet Explorer 3 is sufficient to run this applet. Micol et al. (1996) [1] discussed this issue in more detail.

Java Image Preview Application (JIPA)

While Spectral presents plots of spectra, JIPA's task is to visualise FITS images, i.e. HST preview image collection, and to allow basic image manipulation. The input data format is compressed FITS. There are several options for contrast enhancement, zooming and displaying header keywords (Fig. 2). Another feature is the conversion of mouse locations from pixel space to RA and Dec. JIPA is written in pure Java like the other applets presented in this article and therefore platform independent.

JPlot

JPlot was developed to support the WFPC2 Association Project [2]. It's original task is quality control of HST observation log files (= jitter files). In the meantime it became an integral part of the web interface. It visualises ASCII tables and

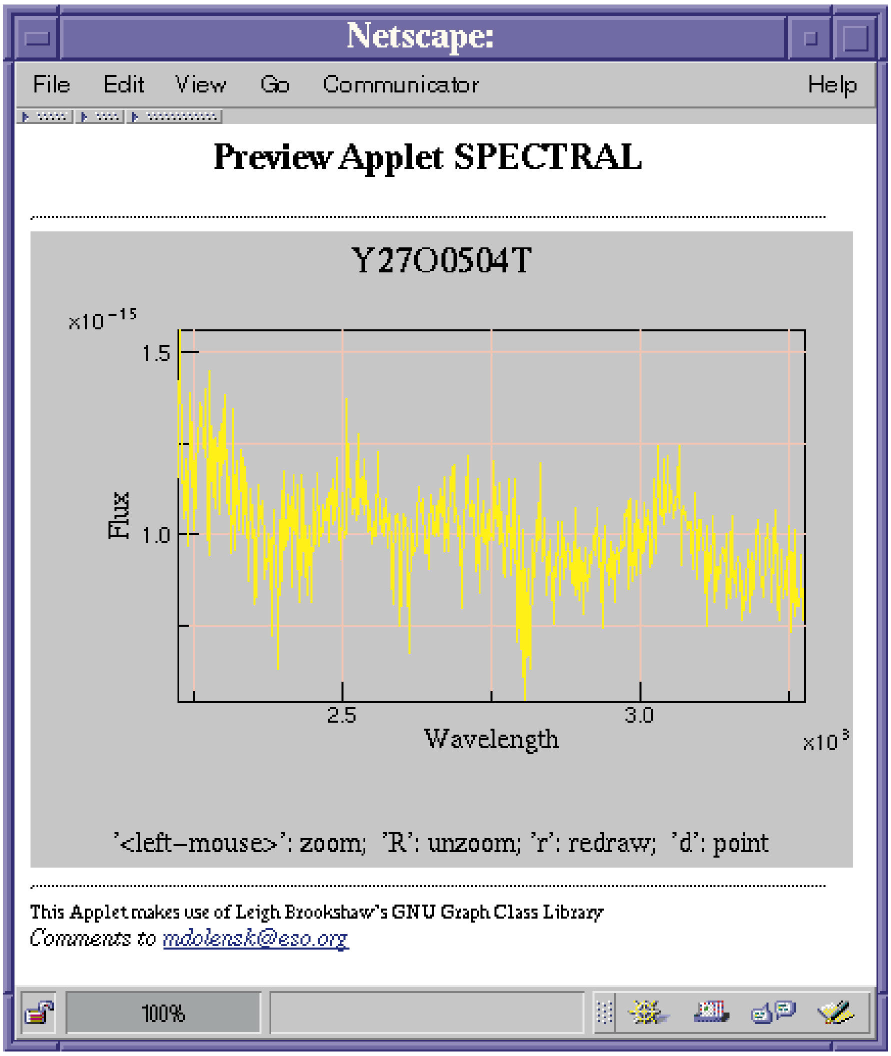

Figure 1.