Astronomical Data Analysis Software and Systems VII
ASP Conference Series, Vol. 145, 1998
R. Albrecht, R. N. Hook and H. A. Bushouse, eds.

HST Paper Products: A New Way to Look at HST Data

W. Hack
Space Telescope Science Institute, 3700 San Martin Dr., Baltimore, MD, 21218, Email: hack@stsci.edu

J.-C. Hsu
Space Telescope Science Institute, 3700 San Martin Dr., Baltimore, MD, 21218, Email: hsu@stsci.edu

Abstract. Two new instruments, STIS and NICMOS, were added to HST during the SM97 servicing mission, each with new types of data that need to be processed. As a result, a completely new system for generating paper products was created which provides a detailed summary of the observations and of the calibrations applied to the observations in the HST pipeline. This new version of the paper products generating software produces a totally redesigned summary for the older instruments (WFPC2, FOS, GHRS, and FOC) as well as creating paper products for the newly installed STIS and NICMOS. In this paper, we discuss the design and algorithms used to handle the extremely large amounts of data produced by the instruments on board HST, as well as the new task's limitations. This new version of the HST paper product software is now available in STSDAS Release V2.0.

1. Introduction

The Hubble Space Telescope (HST) not only generates a great deal of data in a single day, but a wide variety of data as well - from 1-D and echelle spectra to images and time-tag data. All of this data is processed on the ground to apply basic instrumental calibrations and prepare the data to be ingested into the HST archive and then analyzed by the observer. One final step in this processing is to create a summary for initially judging the quality of the observations and for determining the status of the calibrations applied to the data. This summary, however, must be efficient in its reporting and yet comprehensive enough in its results to be useful. Generating such a report for each instrument's data (WFPC2, FOC, FOS, GHRS, STIS and NICMOS) while retaining a similar enough format for easy comprehension was the primary goal for the software. In addition, the software must run quickly enough to prevent backlogs when ingesting into the archive. These two primary problems guided our effort to create a single software system to provide summaries known as 'paper products' for all HST data. This paper will summarize the major features of this software system.

2. Overview of Data Processing

The paper products software performs the following steps in creating the summaries for each set of observations:

- **Sorts input data**: Input data can be a combination of one or more instrument's data from any number of proposals. The input data can either be in FITS format, as if obtained from the HST DADS system, or GEIS format. All of the observations, or only selected observations, can be processed at one time depending on the user's input. A separate output file with its own cover page and summary will then be created for each logical unit defined by data from each instrument used in each proposal.
- **Summarizes each observation**: The basic exposure information for all observations in the proposal is provided in a tabular form. It includes target data, instrument configuration (such as filters used), and basic data statistics. Observations for each instrument are handled by their own function, allowing the summary to reflect the unique nature of each instrument.
- **Displays the data for each observation or observation set**: Grey scale images or plots of the data are provided for each exposure. These displays are created using SPP functions or CL scripts tailored specifically for the type of data produced by each instrument. In addition to the graphical display of the data, a summary of the exposure information is also provided. The basic forms of displaying the data are given for each instrument:

Table 1.

FOS	Target acquisition position plots, spectral plots, total flux plots
GHRS	spectral plots, jitter plots
FOC	image display
WFPC2	mosaic of CCD images, single chip display
NICMOS	thumbnail index of images, mosaiced images, target acquisition images
STIS	image display, spectral plots, jitter images, target acquisition images

- **Provides a summary of the observations data quality**: The last page(s) of the paper products is a summary of the calibration(s). A summary of pointing information (jitter data), explanations regarding quality flags and some basic image statistics are all reported here with an exhaustive list of calibration files and the origin of each file.
- **Process the IGI scripts**: The results from all the processing are generated in the SPP procedures and CL scripts as IGI commands, then processed by IGI to produce the final output.

This processing will produce a report tailored to each instruments data and contain the following estimated number of pages:

Table 2.

Cover Page and Help Page	1 Page each
Help Page	1 Page
Exposure Summary	1+ Pages
Data Display	1 Page/observation (FOC)
	1- 2 Pages/external observation (WFPC2)
	1-3 Pages/observation (FOS)
	1+ Page/observation (STIS)
	32 dithered obs/page + 1 page/image (NICMOS)
Observation Summary	1-2 pages

HST archive operations currently used Helvetica narrow fonts for all paper products. Use of different fonts may cause overlapping of text in the report.

3. Operation of the Software

The paper products software was designed to be run within the IRAF CL environment. The package consists of one primary task, pp_dads, and many hidden tasks to control each instrument's output. The pp_dads task only uses a small number of parameters; namely, input, device, fits_conv.

The parameter input controls which files are read in for processing, including using wildcards, list files or simply one image name. The parameter device can be set to a PostScript device to create PostScript files or a printer queue to send the output directly to a printer. This relies on using IRAF's printer queues, graphics kernels, and STSDAS's PostScript kernel.

4. Use of FITSIO

All HST data is routinely delivered to observers on tape in FITS format, and archived in the DADS system in FITS format as well. The paper products software, therefore, relies on the latest FITS kernel in IRAF to work with STIS and NICMOS data whilst also working directly with FOC, WFPC2, FOS, and GHRS data that are stored in the GEIS format.

5. SPP, CL and IGI

Most data processing is handled by routines written specifically for each instrument's data. These functions are written in SPP for NICMOS, FOC, WFPC2, and FOS, while CL scripts are used to process STIS and GHRS observations. Modular design can be readily modified for each instrument or even have new

code added to support future new instruments. The use of SPP provides efficient access to image headers, the ability to perform statistical calculations while the image remains open, and great control of the output as IGI commands. The CL scripts used for STIS and GHRS are readily edited and can be used without compiling, however, they lack the efficient file I/O that SPP offers.

The results of the calculations and instructions for displaying the data are sent to an output file to be interpreted by IGI, the Interactive Graphics Interface. This provides several advantages for creating the paper products:

- **manages image display**: graphics and images can be displayed with image resizing being done as needed on the fly, enabling creation of thumbnail indices and annotated image
- **machine independent and flexible output**: IGI uses IRAF's graphics kernels and PostScript kernel to create PostScript output for any platform.

6. Computational Efficiency

The paper products software must run fast enough to keep up with the data stream coming from HST in order to avoid backlog. This led to the use of SPP code to generate the paper products for most instruments. In general, the computation of the paper products takes from 30-60 seconds per observation on a Sparc 5 workstation with only one user, and about 2-3 minutes per observation on a DEC Alpha workstation with several users. This is quick enough to avoid most backlog situations. The real bottleneck is the printing of the paper products. A local Printserver 17 printer can print about 15-20 Mb of PostScript output in an hour, however, the paper products code can generate 24Mb of PostScript in about 6 minutes on a Sparc 5, or about 20 minutes on an Alpha workstation. Therefore, a great deal of emphasis has been placed on setting up ways to print these jobs in parallel on multiple printers to keep up with the stream of data coming from HST. A drawback to this is the need for large temporary disk space for block averaged WFPC2/STIS/NICMOS images, PostScript files, and temporary IGI files and an effort is being made to reduce this need.

7. Summary

The new paper products package within STSDAS can process observations from a mix of instruments in one pass resulting in clear and concise documentation. It relies on many of IRAF's built-in functionality as well as many separate IRAF/STSDAS tasks to produce the paper products. The following packages are necessary for the operation of the paper products software: IRAF Version 2.10.4p2, or newer (preferably Version 2.11) and STSDAS/TABLES Version 1.3.5, Version 2.0, or newer. The simple, one task interface makes it easy to run, both interactively by a user after extracting data from DADS and in batch mode from a script, as done by the HST archive. This experience with processing daily observations has proved that this task can function efficiently and effectively with minimal intervention.