Astronomical Data Analysis Software and Systems VIII
ASP Conference Series, Vol. 172, 1999
D. M. Mehringer, R. L. Plante, and D. A. Roberts, eds.

The STScI HST Astrometry Pipeline

John J. Schultz, Philip Goldstein, Pete Hyde, Mary Alice Rose, Keneth Steuerman, John Baum[1], Rick Perrine[1], Daryl A. Swade[1]

Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218

Abstract. An Astrometry Data Processing Pipeline has been developed and deployed at the Space Telescope Science Institute (STScI). This pipeline is responsible for producing Hubble Space Telescope (HST) Astrometry Data Sets used to analyze astrometry observations and the operation of the HST Fine Guidance Sensors (FGS). Prior to the STScI Astrometry Data Processing Pipeline becoming operational, HST Astrometry Data Sets were produced by the Astrometry and Engineering Data Processing (AEDP) System at Goddard Spaceflight Center (GSFC).

This paper describes the processing performed by the STScI Astrometry Pipeline and how the pipeline was designed and developed to reuse existing software components from the STScI OPUS system. OPUS is an STScI developed automated data pipeline system providing a distributed processing environment used to control and monitor applications executing in a sequential order. OPUS is described in detail in other papers[2].

1. Introduction

Prior to early 1998, GEIS format HST Astrometry Data Sets were produced and provided to STScI by the AEDP System at Goddard Spaceflight Center. In early 1998, the HST Astrometry portion of AEDP was discontinued and the STScI Astrometry Pipeline, which produces FITS format datasets, became operational.

The nominal automated Astrometry Pipeline consists of seven sequential processing stages. Each stage, contingent upon successful completion of the prior stage, invokes a specific application to perform a specific function. There are also two stages that can be manually invoked to provide additional functionality. The name and functionality of each stage application is as follows:

1. **ASTPOL** – ingest astrometry observations into the pipeline
2. **SSAST** – produce astrometry observation support schedule files
3. **OMSAST** – produce FITS format Astrometry Data Sets
4. **HTPAST** – save data sets for evaluation by operations personnel
5. **ARCINS** – generate requests to archive incomplete data sets

[1] Computer Sciences Corporation

[2] `http://www.dpt.stsci.edu/dpt_papers/opus_bib.html`

6. **GENREQ** – generate requests to archive complete data sets
7. **INGRSP** – process archive responses from the HST Archive System
8. **ARCTAPE** – archive data sets to tape
9. **ARCCLN** – delete successfully archived data sets from disk

As mentioned earlier, the Astrometry Pipeline was designed and developed to utilize existing components of the STScI OPUS System. In fact, of the above nine Astrometry Pipeline applications, only applications 1, 3, and 4 had to be newly developed. The other six applications are OPUS generic applications that were used as-is with minor modifications to their configuration set-up files. In addition, development of the three new applications incorporated numerous modules from existing OPUS Support libraries containing over 250 modules. This reuse of existing software enabled development of the astrometry pipeline to occur in a minimal amount of time, with a minimal amount of risk, and with considerable confidence that the resultant pipeline would be robust and very reliable.

2. The Astrometry Pipeline

The main responsibility of the STScI Astrometry Pipeline is to produce an Astrometry Data Set for each astrometry observation executed onboard HST. Each Data Set is comprised of four FITS format files; one Science FITS file for each of the three HST FGS, and a Support Schedule FITS file. The Science FITS files contain planned observation information extracted from Science Mission Schedules (SMS) as well as telemetry data accumulated during observation execution time. SMS related information includes data such as observation ids, predicted observation start and stop times, FGS reference positions, and predicted spacecraft orientation. As-executed telemetry data consists of items such as FGS detector photon counts, Star Selector Positions, Star Position Coordinates, and FGS Flag Words. The Support Schedule FITS file contains planned observation information extracted from the STScI Proposal Management Database. This information includes items such as the official target name, the observation proposal id, the name of the observation proposer, information about the instrument performing the observation, and various predicted positional information about the spacecraft.

As described in the Introduction Section, the Astrometry Pipeline is composed of nine applications; seven of which constitute the nominal automated pipeline. Of the seven automated applications, only the first three are involved in producing Astrometry Data Sets; the remaining four deal only with archiving of the data sets once they have been produced. Processing descriptions and additional information pertaining to code reuse for each of the nine applications is provided in following subsections.

2.1. ASTPOL – Ingest Astrometry Observations Into The Pipeline

The Astrometry Pipeline is activated when astrometry data files are received into a specific pipeline input directory. These data files occur in pairs for each astrometry observation executed; an ASCII file containing planned observation information from an SMS, and a binary file containing astrometry telemetry data from when an observation executed. These files are generated and provided to the astrometry pipeline by the STScI HST Observatory Monitoring System (OMS). The ASTPOL application polls for these files and when a pair is found, invokes the next application in the pipeline.

New Modules	Reuse Of OPUS Modules	Percent Module Reuse
1	147	99%

2.2. SSAST – Produce Astrometry Observation Support Schedule Files

This application queries the STScI Proposal Management Database to extract planned astrometry observation information. The extracted information is inserted into ASCII Support Schedule files.

New Modules	Reuse Of OPUS Modules	Percent Module Reuse
0	169	100%

2.3. OMSAST – Produce Astrometry Data Sets

This application actually produces the four FITS Files that comprise each astrometry data set. To produce the Science FITS files, data are extracted from the file pairs provided by the OMS and converted into FITS format. To produce the Support Schedule FITS file, data are extracted from the ASCII Support Schedule file generated by the SSAST application and converted into FITS format. While the FITS files are being produced, data from the OMS provided files are used to compute other observation information not directly provided. This computed information is also included in the FITS files.

New Modules	Reuse Of OPUS Modules	Percent Module Reuse
27	198	88%

2.4. HTPAST – Save Astrometry Data Sets For Evaluation By Operations Personnel

This application makes copies of astrometry data sets and moves them to holding directories where they reside until being evaluated and then explicitly deleted by operations personnel. This application also creates files containing information on the data quality of each processed astrometry observation.

New Modules	Reuse Of OPUS Modules	Percent Module Reuse
1	0	0%

2.5. ARCINS – Generate Requests To Archive Incomplete Astrometry Data Sets

This is the first of the two manually invoked applications not part of the automated astrometry pipeline. It generates requests to have partial astrometry data sets or data set associated intermediate files inserted into the HST archive. This application is used in scenarios where complete astrometry data sets could not be produced.

New Modules	Reuse Of OPUS Modules	Percent Module Reuse
0	85	100%

2.6. GENREQ – Generate Requests To Archive Complete Astrometry Data Sets

This application is the automated pipeline application to generate requests to have complete astrometry data sets inserted into the HST archive.

New Modules	Reuse Of OPUS Modules	Percent Module Reuse
0	131	100%

2.7. INGRSP – Process Archive Insertion Responses From The HST Archive

This application receives responses from the HST archive system indicating the disposition of requests to archive astrometry data sets or portions thereof.

New Modules	Reuse Of OPUS Modules	Percent Module Reuse
0	151	100%

2.8. ARCTAPE – Manually Archive Data Sets To Tape

This is the second of the two manually invoked applications not part of the automated pipeline. It is used to copy astrometry data sets to tape. This application is used in the rare situation where the HST archive is off-line and there is insufficient space to maintain observation data sets on disk.

New Modules	Reuse Of OPUS Modules	Percent Module Reuse
0	124	100%

2.9. ARCCLN – Delete Successfully Archived Data Sets From Disk

This application deletes astrometry data sets from disk after they have been successfully inserted into the HST archive.

New Modules	Reuse Of OPUS Modules	Percent Module Reuse
0	171	100%

3. Summary

The Astrometry Pipeline has been operational at STScI for over 9 months with no software problems whatsoever. There has also been virtually no concerns from the STScI FGS Team about the astrometry data sets being produced. In fact, the new FITS format astrometry data sets contain desired information not contained in the previous GEIS format data sets and are much preferred. In addition, since the astrometry pipeline's underlying OPUS architecture is easily scalable, there are plans to incorporate the FGS Teams' Astrometry Data Set Calibration Software into the astrometry pipeline as an additional application.

From the outset, the STScI Astrometry Pipeline was designed to extensively reuse existing, time-proven components of the STScI OPUS system. This reuse has been the major contributing factor to the timely implementation and success of the STScI Astrometry Pipeline.

Totals for Astrometry Pipeline Software		
New Modules	Reuse Of OPUS Modules	Percent Module Reuse
29	357	92%