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

On-The-Fly Calibration at STScI

Stephen Lubow, Joseph Pollizzi

Space Telescope Institute, 3700 San Martin Dr, Baltimore MD 21218

Abstract. An on-the-fly calibration system (OTFC) is being developed at STScI. OTFC will allow users to retrieve data that is calibrated at STScI with the latest parameters, reference files, and software. In general, the calibration will be carried out as part of the process of satisfying the user retrieval request. OTFC will allow for a considerable savings in storage space, since calibrated data will (eventually) not be stored in the HST archive. The system allows changes to be made in calibration parameters for individual datasets through the use of a database system. It has features for handling datasets with known problems. The OTFC system must be highly reliable. The calibration processing will be managed by the OPUS system. A prototype system has been developed for WFPC2.

1. Introduction

Currently, Hubble Space Telescope (HST) data are calibrated as they are received at the Space Telescope Science Institute (STScI). Raw and calibrated data are stored in the HST archive (DADS). Frequently, users recalibrate data at their home sites to take advantage of better calibration files or software. To recalibrate, users can obtain the raw data and the recommended reference files from the Starview system. The user must then update the raw data headers with the names of the recommended reference files. The appropriate STSDAS calibration software is then run.

An on-the-fly calibration (OTFC) system at STScI will allow users to obtain data from the DADS archive which is calibrated with the most up-to-date calibration parameters, files, and software. In most cases, this capability requires that data be automatically calibrated at STScI when the data is requested by a user. An OTFC system has been successfully developed for HST data by the CADC/ST-ECF (Crabtree et al. 1995). The requirements for an OTFC system at STScI are somewhat different from the CADC/ST-ECF. More details about the OTFC system can be found on the web at

`http://www.dpt.stsci.edu/otfc/otfc_index.html`

2. Motivation

There are two main motivations for the OTFC system: saving data storage and providing users with improved calibration. These motivations were discussed in the HARP report (Hanisch et al. 1996). HST data are stored in an optical disk jukebox as part of the DADS system. The jukebox system has an online capacity of about 3 TB. The system is essentially full now. We have investigated the possibility of using DVD storage technology, but found that this is not yet ready for our purposes. Furthermore, the Advanced Camera for Surveys is scheduled for installation in HST within the next two years. It will substantially increase the rate at which data are received at STScI and would worsen the storage problem. With OTFC, only the raw data would need to be stored. Since raw data compresses better than calibrated data, there are substantial savings (of about a factor of ten in some cases) by storing only raw data in the archive in compressed form.

The other advantage of OTFC is in providing improved calibration to users. Nearly all (more than 90% of) datasets can benefit from recalibration (although the level of improvement is not always significant). Cameras such as WFPC2 produce dark files that typically should be retroactively used to calibrate observations taken in the prior week. Unfortunately, this means that the original calibration carried out in the STScI pipeline is not the optimal. Furthermore, this less than optimal calibration is stored in the DADS archive. An OTFC system would help solve this problem by providing users with the calibration having the appropriate dark, as soon as it is available.

Currently, instruments that undergo evolution of calibration files or calibration software require users to carry out their own recalibrations at their home sites. It may well be easier for both the users and the STScI support groups to have STScI carry out the recalibration.

Apart from reference files and software, improvements to calibration also can come about through changes to calibration switches and other keywords. Another goal of OTFC is to make such keyword changes possible without rearchiving the data in the optical disk archive.

3. Design Goals

We expect that the OTFC system will calibrate a large number of datasets each day, many times more than are calibrated by the current calibration pipeline for new data. For WFPC2, about 300 datasets are currently requested from DADS each day. Consequently, the system needs to be highly automated, while at the same time be robust. The system should provide a reasonable level of performance. It should not impose a substantial delay beyond the current DADS retrieval times.

Some users have established their own calibration pipelines, while others will sometimes customize the calibrations. To support such users, tools should be provided to determine the latest information about calibration parameters.

Another requirement is that the OTFC system should be compatible with other STScI systems, such as Starview and DADS. In fact, these existing systems will play a role in the OTFC system.

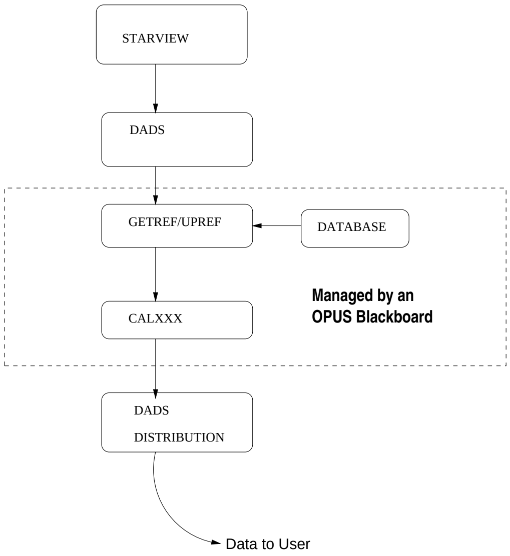

Figure 1. Dataflow for OTFC.

4. Components and Dataflow for On-The-Fly Calibration

Figure 1 shows the dataflow for the OTFC system. First, a user requests data through the Starview user interface. This user interface is currently used by the DADS system and will contain extra options for OTFC. The request is passed to the DADS system which retrieves the appropriate data from the optical disk archive and passes that data to the OTFC pipeline. This pipeline is managed by the OPUS blackboard system (Rose 1996), which handles the process control. An enhanced version of the current getref tool queries a database to determine the names of the currently recommended reference files, the recommended calibration switch values, and any changes to other header keywords for the dataset being calibrated. (The header keyword changes are managed by the OTFC Change System which is described in the next section.) It writes its output to a text file called the delta file. The upref tool applies the delta file to update the data headers. The data is then calibrated by the appropriate STSDAS calibration routine. The data is the returned to the user via the DADS distribution service.

5. Change System

To describe changes to calibration keywords relative to the values in the archived raw data, a set of database tables is used. These tables contain entries for each dataset, keyword, and header (in the case of multiple headers). Each entry contains the new and old values of the keyword, the operation (insert, update, or delete) that should be applied to the header, as well as textual comments that describe the nature of the change. In essence, these tables describe how to edit the header. By using a database, we are able to provide a history of changes to keywords that is readily accessible to support staff. Furthermore, these entries are used to change the values for these keywords in the DADS catalog.

The change system also includes software to determine the calibration switch rules (Bestswitch) and the best reference files for each dataset (Bestref). Changes to parameters in the database for each dataset need to be carefully coordinated. For example, keyword changes need to be applied ahead of the determination of reference files. Furthermore, these operations need to be carried out as an atomic transaction so that the system is not left in an inconsistent state if errors occur.

The keyword changes are applied to the data header by an enhanced version of the getref/upref software. These tools will be provided for remote users to carry out their own calibrations. Getref can be accessed through a web interface, Upref is part of STSDAS.

Some datasets should not be calibrated by the OTFC system. For example, the data may be corrupt or it may take too long to calibrate. The OTFC change system contains a database table that describes such situations. The OTFC will not calibrate such data, but will return the raw data to the user.

Currently, for WFPC2, there are 368 keyword changes in the database and 11 datasets have been identified in the database for noncalibration.

6. Plans

We have developed a prototype system for WFPC2. We will release that system to external users in the near future. We will then extend OTFC to handle STIS, NICMOS, and ACS.

Acknowledgments. We thank Dan Durand for informative discussions.

References

Crabtree, D. R., Durand, D., Gaudet, S., & Hill, N. 1995, in ASP Conf. Ser., Vol. 101, Astronomical Data Analysis Software and Systems V, ed. G. H. Jacoby & J. Barnes (San Francisco: ASP), 505

Hanisch, R.J., et al. 1996, HARP: The Hubble Archive Re-Engineering Project Summary Report, STScI report

Rose, J. 1996, in ASP Conf. Ser., Vol. 125, Astronomical Data Analysis Software and Systems VI, ed. G. Hunt & H. E. Payne (San Francisco: ASP), 38