master file

CAL/008

TO: Distribution DATE: July 16, 1986
FROM: Keith Horne

SUBJECT: Proposal to Implement a Dynamic Throughput Generator in CDBS

This document proposes changes to CDBS requirements in the areas of "absolute flux calibration" and "photometric transformations". Formal changes to the CDBS requirements document SO-11 are in preparation. The proposed changes to CDBS will implement a dynamic throughput generator that uses a database of component throughput functions, a graph description of the HST observatory, and synthetic photometry techniques to calculate sensitivities of HST observing modes. All science instruments are affected. Please review and send comments to Keith Horne. Instrument scientists should take note of section 5.2.

1 Introduction

The Calibration Data Base Software (CDBS) must perform two main functions in the area of throughput calibration.

1. Supply users (*e.g.*, RSDP, Pepsi, simulators, ISB) with the throughput calibration for any requested observation by the Hubble Space Telescope (HST).
2. Provide tools for maintaining and updating of the throughput calibration information on the basis of observations of calibration targets.

The main purpose of this document is to define requirements for implementing these functions in CDBS.

The basic philosophy of the calibration program based on synthetic photometry techniques is described in Koornneef, *et al.* (1986). The first of the two functions needed in CDBS is fulfilled by a software tool known as a dynamic throughput generator, which is described in Horne, Burrows and Koornneef (1986). Section 2 defines at a requirements level the procedures needed to implement the dynamic throughput generator in CDBS. The required CDBS databases are discussed in Section 3.

The process to be used to update the throughput calibration information on the basis of calibration target observations is well understood, but some details remain to be defined. Section 4 outlines the calibration updating process and defines requirements for those steps that can be implemented in CDBS at this time. Section 5 examines areas outside of CDBS where further work is needed. Section 6 summarizes by giving benchmarks for the complete implementation of the throughput calibration program.

2 Implementation of the Dynamic Throughput Generator

2.1 Overview

Because the HST observatory has a vast number of interrelated observing modes, it is impractical and inefficient to maintain an independent calibration for every possible instrument configuration. When this problem was first widely appreciated in the ISB, the initial reaction was to identify and restrict calibrations to a smaller number of "core" modes. The ISB has since adopted an alternative solution based on a software tool, the dynamic throughput generator, with the capability to generate the throughput function for *any* HST observing mode. Throughput functions and calibration data files for specific observing modes are to be generated *dynamically*, whenever they are needed. This approach reduces to a manageable level the number of calibration data files that must be created and maintained, and ultimately saves considerable observing time, since information from calibration observations in one mode can be easily "transferred" to other closely related modes.

The basic concepts, data structures and software needed for dynamic throughput generation are discussed by Horne, Burrows and Koornneef (1986) and are reviewed in this section. The throughput calibration of the HST observatory is represented in a framework consisting of

1. *component throughput functions*, $P_i(\lambda)$, for every optical component (*e.g.*, mirror, filter, polarizer, disperser, detector)
2. *configuration graph* describing the allowed combinations of the components.

A particular observing *mode* is specified by a list of *keywords*, which might be generated from parameters in the data header of an observation or familiar names of filters, detectors, etc. The keywords are used to trace a path through the *observatory configuration graph*, thereby translating the keyword list into a list of pointers to data files that contain the component throughput functions. The grand throughput function, $P(\lambda)$, of the requested observing mode is formed by multiplying together the individual component throughputs at each wavelength.

Some users, in particular RSDP, require the *sensitivity*, S, or *inverse sensitivity*, U, rather than the full throughput function of the observing mode. These quantities are used to convert an observed count rate, C, to a flux density, f_λ in $erg\ cm^{-2}s^{-1}\mathring{A}^{-1}$, via

$$f_\lambda = C/S = CU.$$

U is the flux density required to produce a unit signal of 1 count per second. The sensitivity is *calculated* as an integral of the throughput function over the passband,

$$S = \frac{Area}{hc}\int P(\lambda)\lambda d\lambda,$$

where $h = 6.6262 \times 10^{-27} erg\ s$ is Planck's constant, $c = 2.997925 \times 10^{10} cm\ s^{-1}$ is the speed of light, and $Area = \pi(120\ cm)^2$ is the nominal area of the HST entrance aperture (the obscuration factor 0.86 due to the secondary mirror is taken into account in the Optical Telescope Assembly throughput).

Spectrographic and broad-band photometric modes are to be treated identically. Thus each pixel of the spectrum is regarded as a narrow passband. For some spectrographic observing modes, *e.g.*, HRS observations, the throughput $P(\lambda)$ does not vary appreciably over the wavelength interval covered by a pixel, and the passband integral can be simplified to

$$S_i = \frac{Area}{hc} P(\lambda_i)\lambda_i \Delta\lambda_i,$$

where S_i is the sensitivity of pixel i, λ_i is the wavelength of the pixel, taken from the wavelength calibration of the spectrograph, and $\Delta\lambda_i = (\lambda_{i+1} - \lambda_{i-1})/2$ is the bandwidth covered by the pixel. However, for other spectroscopic modes, *e.g.*, the red end of FOS PRISM observations, an appreciable wavelength range is spanned by single pixels and the integral must be performed more carefully.

2.2 Basic Requirements for the Dynamic Throughput Generator

INPUT:

- A character string containing a list of keywords that specify the required observing mode.
- A wavelength set, λ_i, on which the throughput is to be evaluated.

PROCEDURE:

1. Use keywords to trace a path through the observatory graph and thereby to generate pointers to the required component throughput data.
2. Interpolate the log-throughput data for each component j onto the required wavelength set to obtain $\log P_j(\lambda_i)$. Linear interpolation should suffice since the component throughput data will be finely-sampled.
3. Interpolate the uncertainty in the log-throughput for each component j onto the required wavelength, λ_i, set to obtain $\sigma(\log P_j(\lambda_i))$.
4. Add the individual component log-throughputs, wavelength by wavelength, to generate a grand log-throughput function for the mode,

$$\log P(\lambda_i) = \sum_j \log P_j(\lambda_i).$$

5. Add in quadrature the uncertainties of the individual component log-throughputs, wavelength by wavelength, to generate the uncertainty in the grand log-throughput function for the mode,

$$\sigma(\log P(\lambda_i)) = \sqrt{\sum_j [\sigma(\log P_j(\lambda_i))]^2}.$$

   This formula assumes that throughput errors for different components are independent.

6. Compute the sensitivity, S, from the following weighted integral of $P(\lambda)$ over the passband:

$$S = \frac{Area}{hc} \int P(\lambda)\lambda d\lambda.$$

   The constant $Area/hc$ has the value $2.2773 \times 10^{12} cm^2\ erg^{-1} \mathring{A}^{-1}$ in units that give f_λ the desired flux density units. The numerical error in the calculation of this and other numerical integrals should be at most 0.1 %.

7. Compute the uncertainty in the sensitivity,

$$\sigma(S) = \frac{Area}{hc} \int \sigma(P(\lambda))\lambda d\lambda.$$

   This formula assumes that throughput errors at different wavelengths are highly correlated. Note that the integrand contains $\sigma(P) = P \ln(10)\sigma(\log P)$.

8. Compute the *pivot wavelength*, λ_p, given by

$$\lambda_p = \sqrt{\frac{\int P(\lambda)\lambda d\lambda}{\int P(\lambda) d\lambda/\lambda}}$$

   The pivot wavelength is to be used for conversions between flux densities based on ν and λ

$$f_\nu = \frac{\lambda_p^2}{c} f_\lambda$$

OUTPUT:

- List of the component throughput files used.
- Grand log-throughput function, $\log P(\lambda_i)$, evaluated on the requested wavelength set.
- Uncertainty in the grand log-throughput function, $\sigma(\log P(\lambda_i))$, evaluated on the requested wavelength set.
- Sensitivity of the mode, S.
- Uncertainty in the sensitivity, $\sigma(S)$.
- Pivot wavelength λ_p in $\mathring{A}$.
- Error code flagging failure of keywords to uniquely specify a valid mode, or inability to retrieve component throughput data.

3 Databases

CDBS must establish, maintain, and provide access to four expanding and evolving databases.

1. component throughput functions
2. observatory graph links
3. calibration target flux distributions
4. observations of the calibration targets

This section describes the purpose and contents of each database and the associated software tools needed for access, maintenance and updating.

3.1 Components Database

Each entry in the components represents a specific optical component of the HST. The entry for each component gives the log-throughput and the uncertainty of the log-throughput at a number of wavelengths. Different wavelength sets may be used for different components. The component throughputs are dimensionless; they represent at each wavelength the number of photons exiting from the component (or counts detected in the case of a detector) per photon incident on the component. Laboratory measurements or manufacturers specifications are available for most of the components. Our knowledge of the component throughput functions will evolve with time. It may be necessary to add additional components to the database.

Software should be developed in CDBS to do the following:

- add a new component
- modify (create new version of) an existing component
- make log-log plots of throughput data and uncertainties for selected components

3.2 Observatory Graph

The observatory graph consists of a set of *nodes* that are connected together by a network of *links*. Each node represents a place along one of the many possible paths that light can take through the HST observatory. Each link represents the possibility that light can pass from one node, the entry node, through a specific optical component, and on to the next node, the exit node. The graph is thus somewhat like a tree, except that different paths through the observatory graph can reconnect, while the branches of a tree, by definition, do not.

Each entry in the observatory graph database represents one of the links, and gives (at least) the following four quantities

- entry node

- exit node
- pointer to an entry in the components throughput database
- keyword

The entry and exit nodes specify the two nodes that are connected by this link, and the direction in which the light passes. The pointer directs us to the throughput data for the optical component through which the light passes. The keyword attached to each link is used by the graph searching software to trace the light path through the observatory graph. A path that arrives at the entry node, can proceed to the exit node if the keyword attached to the link matches one of the keywords in the list that specifies the observing mode. It will be possible to invoke a component by any of a number of different keywords; each of these aliases will require a separate link. It will also be possible to specify a mode requiring many components with only a few keywords since the same keyword can be attached to different links.

The following software tools are needed to maintain and update the observatory graph.

- Add and delete nodes from the graph.
- Check that valid throughput data is present for all pointers.
- Check for links with entry nodes that are not attached to the rest of the graph.
- Check for multiple links with identical entry nodes and keywords (these result in ambiguous paths). The process that adds nodes should be designed so this can't happen.
- Make a listing of all links that are "downstream" from a specified starting node.
- Draw a graphical representation of the links that are "downstream" from a specific starting node (useful but not absolutely necessary for verification purposes).

3.3 Calibration Target Flux Distributions

Each entry in the flux distribution database represents the flux density of one of the calibration targets at a number of discrete wavelengths. The wavelength spacing need not be uniform, but should everywhere be fine enough that linear interpolation is adequate to represent the spectrum at intermediate wavelengths. The wavelength set may be different for different targets.

Note that these flux distributions are **not** the same as the spectrophotometric observations that will be available for many of the calibration targets. (The spectrophotometric data on calibration targets reside in the observations database described below.) Our knowledge of the flux distributions will improve as observations accumulate. We may also wish to change the wavelength set of a given target, *e.g.*, if new data with high spectral resolution become available.

3.4 Calibration Target Observations

Each entry in the observations database represents or provides access to an observation of one of the calibration targets. This database will reference ground-based measurements, measurements by IUE, and ultimately measurements by the HST. It will include both spectrophotometric measurements, with high spectral resolution, and broadband photometric measurements, with low spectral resolution. Although the source and format of the observations referenced are diverse, the mode of access to them should be uniform.

Each entry should provide access to at least the following:

- epoch of observation
- engineering parameters that may affect the throughput
- keywords specifying the observing mode
- components list
- throughput function
- observed count rate (counts/sec)
- uncertainty in observed count rate (1-sigma)

It should be possible to select observations from the database by specifying one or more of the following:

- a range of epochs
- a specific target (or wildcard)
- a specific observing mode (or wildcard)
- a specific component

4 Updating the Throughput Calibration

Observations of calibration targets will be used to update the component throughput functions and the target flux distributions. The basic steps of this calibration process are as follows:

1. Collect the relevant observations (fully corrected for background, nonlinearity, and detector sensitivity variations, and integrated over any spatial dimensions).
2. For each observation, form the keyword list that specifies the observing mode.
3. Send the keyword list through the graph searching software, to trace a path through the observatory graph and generate a list of components.
4. Combine the throughput functions for each of these components to form the grand throughput function for the observation.
5. Use the target name to retrieve the flux distribution of the star.
6. Compute predicted count rate for each observation by a numerical integral over the passband, as defined below.
7. Compare the predicted and observed count rates for all the observations.
8. Generate corrected *component* throughput functions and/or target flux distributions to reconcile differences between observed and predicted count rates.

The desired condition is one in which the predicted count rates for the observations are consistent with the observed count rates to within the uncertainties of the measurements. The predicted count rate C for passband $P(\lambda)$ is given by

$$C = S f_\lambda(P),$$

where S is the sensitivity of the passband, as defined above in Section 2, and $f_\lambda(P)$ is the *mean flux density* of the star in the passband, given by

$$f_\lambda(P) = \frac{\int f_\lambda(\lambda) P(\lambda) \lambda d\lambda}{\int P(\lambda) \lambda d\lambda}.$$

Equivalently, the predicted count rate is just

$$C = \frac{Area}{hc} \int P(\lambda) f_\lambda(\lambda) \lambda d\lambda.$$

Again, spectrophotometric and broadband measurements are treated uniformly, each pixel of the spectrograph representing just a narrow passband.

The penultimate step of the calibration process yields the ratio of predicted to observed counts for each observing mode and target. These results will be reviewed and analyzed by ISB personnel, who will perform the final step of updating component throughput and target flux distribution databases to reconcile any discrepancies. This final step may be automated at some point in the future when a suitably robust procedure is developed.

The following tasks are needed in CDBS to support this activity.

1. Utilities to access, update, verify and select information from the four databases.
2. Computational tools that perform the required numerical integrals consistently, accurately and efficiently.
3. Display programs and scripts for comparing observations with predictions. These should include the following:
    - For each target, a plot of the flux distribution overlain with crosses giving observed flux density ± uncertainty and pivot wavelength or effective wavelength ± rms bandwidth.
    - For each target, a plot of the ratio of observed to predicted count rate vs wavelength, with error bars as above.
    - Tabular presentation of the above information.
4. A program to perform a least-squares fit of a general parameterized flux distribution to a set of observations in various but fixed passbands (spectrophotometric and broadband). This will be needed to generate and tweak flux distributions for the calibration targets.

5 Work Areas Outside of CDBS

This section identifies areas where work outside the context of CDBS is needed to implement the throughput calibration program.

5.1 RSDP

The Routine Science Data Processor (RSDP) should be modified to place throughput calibration information in the data headers of *all* HST observations. The format of this information should be uniform across all the science instruments. This header information should include the parameters ST_0, K, and λ_p as defined below.

The sensitivity, S, is provided in the headers in the form of a magnitude zero-point, ST_0, defined by

$$ST_0 = +2.5 \log S - 21.10$$

This allows a user to place the observed count rate C on the ST magnitude system via

$$ST = -2.5 \log C + ST_0,$$

and to compute the corresponding flux density f_λ via

$$ST = -2.5 \log f_\lambda + K.$$

K initially will have the value -21.10 but may change to a wavelength-dependent value in the future, *e.g.*, if the adopted fundamental calibration of Vega is improved.

The pivot wavelength λ_p should be given in the headers to allow users to convert the flux density to f_ν via

$$f_\nu = \frac{\lambda_p^2}{c} f_\lambda.$$

According to current plans, the dynamic throughput generator in CDBS will be used to create data files that will be referenced by the RSDP pipeline. Thus someone (or some process) must anticipate every HST observation and explicitly create the required throughput calibration files shortly before the data enter the pipeline. This scenario probably can be made to work in the short term, but it would be better in the long run if RSDP used the throughput generator explicitly, rather than accessing data files generated by it.

5.2 Instrument Graphs

Graph descriptions must be developed for each of the instruments in the HST observatory. As a guideline, these instrument graphs should be kept as simple as possible, yet with sufficient realism to accurately describe the observations of calibration targets. Graph development is best done by or under supervision and approval of the instrument scientists, since it requires intimate knowledge of the instrumental configuration and intuition as to which components or other parameters can appreciably affect the throughput. The simple preliminary graphs developed for testing purposes by the ISB calibration target data group can serve as starting points for several of the instruments.

5.3 Components Updating Strategy

Although the best available information on the throughput functions of individual components will be placed into the components database before launch, the throughputs will need to be adjusted to bring about agreement between observations of calibration targets and the corresponding predictions. We must therefore develop a general strategy for adjusting the individual component throughput functions.

We can choose among many different ways to assign the throughput corrections, since each observing mode involves many components. For example, we may initially apply throughput corrections to filters, thereby confining the the calibration information to those modes in which calibration targets have been observed. Later, if it is established that all the filters require similar throughput corrections, the calibration information can be easily "transferred" to related modes by shifting the mean correction from the filters to the detector. This updates the calibration of all modes involving the same detector, including modes for which calibration observations have not been explictly undertaken. Similarly, if similar throughput corrections are found for the detectors of the individual science instruments, the mean correction may be shifted to the OTA throughput.

The process of systematically comparing observed and calculated data, adjusting component throughputs, and iterating until the agreement is satisfactory, can be laborious and prone to error if carried out entirely by hand. Some automation of the process is therefore desirable, and this can be done only after the adoption of general policies for distributing corrections among the components.

An automatic and objective procedure to decide how to distribute throughput corrections among the components could be based on the maximum entropy method (MEM). The envisioned MEM procedure would fit (at a specified level of χ^2) any given collection of calibration observations by making small multiplicative throughput corrections to the affected components while leaving the throughputs of all other components at their default values. The MEM throughput corrections would be as uniformly distributed as possible, assuming that each component is given equal weight in the definition of the entropy. For example, if an observation involving a telescope, a filter and a detector yields a count rate 9 % higher than predicted, MEM would increase the throughput of all three components by 3 %. Alternatively, we could choose to apply most of the 9 % correction to the detector simply by assigning a high weight to the telescope and filter, thereby forcing them to remain close to their default values, and a low weight to the detector. Thus the distribution of corrections is easily controlled by the assignment of low weights to components whose throughputs we believe are most likely to be in error, or to be changing rapidly.

The MEM framework is attractive for this problem because it handles straightforwardly a complex combination of modes involving many components, it ensures that individual throughputs are always positive, and it takes into account uncertainties in the observations. Linear programming techniques offer an alternative approach, but constraints would then have to be implemented to ensure that the multiplicative corrections are always positive.

5.4 Trend Analysis

In the steady state era, discrete updates of the throughput calibration will probably be made whenever the accumulating calibration observations indicate that a change is necessary. RSDP will use whatever throughput calibration is in place at the time of an observation. However, when the highest accuracy is required, the calibration should vary smoothly in time and make use of calibration observations both before and after the epoch, t_0, of the observation being calibrated. The MEM approach allows for this possibility as follows. If a calibration observation at time t is weighted in proportion to a function such as $e^{-|t-t_0|/T}$, then that observation affects the calibration with a time constant of order T.

5.5 Extraction Procedures for Panoramic Detectors

The throughput calibration is derived from point-source calibration targets and will therefore depend on the exact algorithm that is used to integrate over spatial dimensions of the data. A robust extraction procedure should be defined and consistently followed for calibration target observations with panoramic detectors, such as WF/PC, FOC. This standard algorithm must be carefully described so that users of other extraction methods can derive empirical corrections.

The point spread function has a very broad component due *e.g.*, to scattering off rough mirror surfaces. Consequently, some fraction of the stellar light will be excluded by the "software aperture" that is implicit in the standard extraction procedure. This wavelength-dependent loss must be included as a component in the graphs of panoramic detectors.

5.6 Extended Sources

Extended sources have a higher effective throughput than point sources because the detector registers photons from a much wider region of the point-spread function. This difference can be treated by using a different and wider "software aperture" in computing the throughput, or by simply assuming that the throughput of the "software aperture" is unity for extended sources. A keyword UNIFORM or POINT could be used to indicate the kind of calibration that was desired.

The calculation of surface brightnesses requires no special treatment. The user simply adds up the counts in any desired spatial region, divides by the exposure time to get a count rate, uses the sensitivity to convert the count rate to a flux density, and divides by the area of the chosen region in arc-seconds to obtain a surface brightness.

5.7 Polarization

Polarization can be treated exactly in our framework by using a wavelength-dependent 2×2 transfer matrix to describe the effect of each optical component on the wavelength-dependent 2×2 polarization matrix of the photon field. However, this treatment is probably too cumbersome to adopt for general use. It might be best to treat the polarizing components of the observatory graph as perfect transmitters when calculating the throughput, and delegate analysis of the polarization to a subsequent step. Alternatively, we could develop separate calibrations of the ordinary and extraordinary beams. Additional thought and discussion is needed to see how polarization fits in.

6 Benchmarks for Complete Implementation of the Throughput Calibration Program

The following items must be completed before launch of the HST.

- **CDBS databases**
  1. component throughput functions
  2. observatory graph
  3. flux distributions for calibration targets
  4. observations of calibration targets
- **CDBS software**
  1. dynamic throughput generator
  2. graph maintenance tools
  3. spectrum fitting tools
- **RSDP software changes**
  1. throughput information to science headers
  2. direct use of dynamic throughput generator (or a process to ensure that the CDBS throughput generator is used to create reference data files for RSDP shortly before every observation takes place).
- **Data**
  1. component throughput functions
  2. graphs for all science instruments
  3. calibration target observations on ST system
- **Tests**
  1. installation of component throughput data
  2. validation of components database (atlas of plots and sign off by instrument scientists)
  3. installation of instrument graphs
  4. testing of observatory graph and components database against TV test data; first update of components database
  5. installation of calibration target observations (ground-based and IUE)
  6. generation of flux distributions for calibration targets

References

Horne,K., Burrows,C. & Koornneef,J. 1986 "Dynamic Generation of Throughput Functions: A Unified Approach", STScI presentation at ISB meeting and STScI memo distributed 16 May 1986.

Koornneef,J., Bohlin,R., Buser,R., Horne,K. & Turnshek,D. 1986, "Synthetic Photometry and the Calibration of the Hubble Space Telescope", to appear in *Highlights of Astronomy*, Vol 7, ed. J.P.Swings (STScI preprint No.91; also distributed at the 16 May 1986 ISB meeting).