1997 HST Calibration Workshop
Space Telescope Science Institute, 1997
S. Casertano, et al., eds.

Associations of WFPC2 Exposures

A. Micol

Space Telescope European Coordinating Facility, ESA/SSD/SA, ESO

P. Bristow

NATO SACLANT, Undersea Research Centre

B. Pirenne

Space Telescope European Coordinating Facility, ESO

Abstract. Unlike NICMOS and STIS, where a dataset might be constituted of a set of exposures, the WFPC2 data set's structure was thought to be a repository of all the files belonging to a single exposure. Lacking higher level of abstraction, there is no simple way of tracing back which observation strategy (e.g., cosmic ray rejection, dithering) a WFPC2 PI decided to employ.

Building *associations* of WFPC2 exposures can therefore be considered an important step towards a comprehensive description of the HST archive contents.

ST-ECF has started a project to build such associations and to store them into a database. Knowing what are the shifts among the exposures belonging to an association will allow the ST-ECF Archive to provide its users not only with on-the-fly recalibrated exposures, but also with cosmic ray-rejected and resolution-enhanced mosaics of WFPC2 images.

1. Introduction

The ST-ECF (Space Telescope European Coordinating Facility) HST archive stores each of the files generated by individual exposures (raw data, calibration files, calibrated data, auxiliary data), but does not store cosmic ray-cleaned or co-added exposures, nor does it currently have the capacity to handle these problems. The process of cosmic ray (hereafter CR) removal can be rather time consuming, requiring the identification of the relevant exposures, application of CR removal algorithms to only those exposures which are appropriately aligned and further combination possibly involving complex techniques such as "dithering" and "drizzling" (see Dickinson & Fosbury 1995, Hook 1995, and the Web page by Fruchter at URL `http://www.stsci.edu/~fruchter/dither/`).

Ideally one would prefer a more user-friendly situation whereby users can request the "product" of any given observation. A similar system is currently being prepared for the upcoming set of HST instruments (NICMOS and STIS) where the ground system will actually produce "observation products" rather than individual exposures in a semi-finished state as is currently the case.

To tackle this problem and provide a more complete service to archive researchers, the ST-ECF embarked upon a project which aimed to firstly group together into "associations" those WFPC2 images which "belonged" together for CR removal and co-addition and secondly combine these exposures and make the resultant products available, thereby retrofitting existing and future WFPC2 exposures into observation products. Determining (in an automatic way) which exposures belonged together turned out to be a rather intractable problem for three reasons.

- The primary parameter influencing the grouping is obviously the pointing. Unfortunately existing tabulated values were either not sufficiently accurate or (more worryingly) unreliably recorded. The same applied to the values appearing in the headers of the exposure fits files.
- A variety of observing strategies have been adopted by PIs so that we must consider associations which contain any number of exactly registered exposures and any number of shifted exposures.
- Ultimately, the decision of whether to include or reject an exposure from a group depends upon the scientific demands upon the resultant image. Some users may want good photometry, some good astrometry, in some cases low fluxes or sky variations are vitally important, in others high resolution.

Section 2 below discusses the restrictions automated data reduction techniques place upon this grouping, Section 3 explains how accurate pointing information was obtained, Section 4 describes the grouping criteria, Section 5 describes the relational scheme, and the automated CR removal and stacking is discussed in Section 6.

2. Requirements of Automation

Whilst the various steps in the "on the fly recalibration" are well suited to automation, this is not the case with CR rejection and co-addition. An interactive user may select and reject exposures for inclusion in co-addition according to his judgement and scientific goals, while here we must find an automatic way of grouping exposures. This involves taking into account the likely observing strategy of the PI whilst only grouping suitable exposures for CR removal.

2.1. Observing Strategies

In order to reproduce the CR rejection and co-adding procedures of real observers in an automatic way, we need to consider the strategy they are likely to adopt for their exposures. Of course, if the standard practice were to make several integrations with precisely the same pointing, then the co-addition and CR removal would be trivial. However, the real situation is more complicated for two reasons.

Firstly, in order to reduce the influence of hot pixels or bad columns, a shift of several pixels is made between the pointing of exposures of one target. Secondly, a further subpixel shift allows the true spatial resolution of HST, which is under-sampled by even the PC pixels, to be substantially recovered (Hook 1995). Consequently, a non-integer pixel shift, for example 5.5 pixels (along each axis), is likely to be present between the exposures of one target. In principle there may be just one exposure at each pointing, as tasks that can remove cosmic rays from exposures shifted by non-integral pixel offsets are now being investigated, but their eventual reliability and date of availability are not yet known. At the moment therefore the recommended strategy is to obtain multiple exposures at a single pointing (Leitherer et al. 1995). CR-cleaned exposures can then be shifted and added or "drizzled" together for higher resolution.

The problem we face here therefore is to recover the intentions of the PI by finding how many exposures are exactly registered so that we can combine them with a CR rejection algorithm, and how many of these "subgroups" belong together as part of a dithering strategy.

2.2. Cosmic Ray Removal

The most straightforward, and therefore most reliably automated, algorithms for CR rejection perform a comparison of corresponding pixels in several well registered exposures. If a pixel from one exposure contains an anomalously high flux then it is disregarded and replaced by a scaled average derived from the other exposure(s). Hence it is essential that the input exposures are very closely aligned to avoid degrading the PSF. Moreover, a small shift in the position of the pixel grid with respect to the centroid of a star may confuse the comparison for the pixels involved. Therefore we need to know which exposures are precisely registered and which are offset for dithering.

For interactive data reduction this does not pose such a problem, a user can select several bright point sources in each exposure, find their centroids and thus establish whether they are registered accurately enough for CR rejection (or measure the offset for dither combination, see Section 6). This problem can be tackled automatically with auto-correlation algorithms of course, however not all exposures have enough features to make auto-correlation reliable. Moreover, we must first establish what the members of the group are by comparing the position of each exposure with any potential companions, auto-correlation would obviously not be a practical solution for such a large number of comparisons. Therefore we need to find pre-recorded pointing information for all exposures.

3. Accurate Pointing Information

Initial attempts to group exposures involved searching the WFPC2 database for exposures well enough registered with one another to allow CR rejection (what constitutes "well enough registered" is discussed in Section 4). It soon became apparent however that associations obtained in this way often in fact contained exposures that were offset from one another, in contradiction to the tabulated co-ordinates[1]. This discovery prompted an examination of the co-ordinates recorded in the FITS file headers. In some cases these were indeed more reliable, but by no means always. The problems appear to arise for two reasons;

- A set of consecutive exposures are carried out for a given target and dithering is requested for some of this set. In some cases, exactly the same co-ordinates are recorded in the headers for all of the exposures and no indication as to the dithering offset is recorded.
- A target is re-acquired at a later epoch for further examinations, the re-acquisition may not be precise, but the recorded co-ordinates reflect the desired position (i.e., that of earlier exposure), not the actual position.

It is important to realise that the problems do not stem from some intrinsic lack of precision, in principle *relative* offsets can be accurately measured to within a fraction of a PC pixel (see STScI on-line Observation Logs documentation, M. Lallo 1996), but from inconsistently recorded positions[2].

The information required must exist, at least in the HST engineering data. Since October 20, 1994, observation logs (often referred to as "jitter" files) for all exposures have been

[1]The are in fact several possible tabulated co-ordinates that can be considered (eg. the V1 axis co-ordinates in the HST science table or the "crval" co-ordinates in the "wfpc2_primary_data" table) but none proved to be reliable

[2]A related problem here is that the CD matrices, which allow, for instance, SAOimage to convert pixel co-ordinates to RA and DEC given a reference pixel, recorded in the headers suffer the same problems as the pointing information. Hence measurements which depend upon the CD matrices are not necessarily reliable when very high precision is required.

produced by the Observatory Monitoring System, an automated software system which interrogates the HST engineering telemetry and correlates the time-tagged engineering stream with the scheduled events data. These logs contain tabulated pointing information as a function of time, from which we may calculate the mean and standard deviation, along with other diagnostics which give an indication as to the quality and reliability of the exposure.

Hence the *jitter co-ordinates* (RA, DEC and ROLL ANGLE) for all exposures were extracted along with other useful information (see associations tables). The provision of this information alone should greatly assist archive users carrying out data reduction. The ST-ECF has also undertaken to generate jitter files for exposures pre-dating October 20, 1994, so that jitter co-ordinates will exist for all WFPC2 exposures.

4. Grouping with Jitter Information

4.1. Member Types

Before an exposure can be considered as a group member it has to be assigned a "member flag". This is necessary because we only want to process those exposures for which we are confident that the pointing information is accurate (and there are no other anomalies). These exposures are flagged "**P**". The remainders are mostly flagged "**G**" (Groupable) except in those cases where there are indications that it is not even meaningful to group the exposures, in which case they are flagged "**B**" (Bad).

There are a number of other flags which indicate whether an exposure is suitable for automatic processing. The following conditions lead to an exposure being considered to be unprocessable (i.e., they have flag "**G**" or "**B**").
A "**G**" flag is assigned when:

- The exposure was taken in parallel mode.
- The exposure was not taken in "FINELOCK" mode.
- There was a loss of lock during the exposure.

A "**B**" flag is assigned when:

- Slewing occurred during the exposure.
- The aperture field is null in the jitter file.
- No mean RA, DEC, or ROLL can be extracted from the jitter file.

A restriction was also placed upon the standard deviation in RA, DEC and ROLL derived from the jitter files (which reflects the movement of the reference pixel during the observation and will effect the PSF. It is required that the RA and DEC have $\sigma < 0.1$ PC pixels ($\sim 0\farcs005$); exposures with larger jitter cannot satisfy the subgrouping criteria and are assigned a "**G**" flag. Similarly for ROLL, $\sigma < 7\farcs2$ is required.

4.2. Grouping Criteria

We have adopted a very conservative approach: where there is any uncertainty in the extracted jitter information we do not set the member flag to "**P**".

As discussed in section 2.1, we anticipate most exposures to be arranged according to some kind of dithering strategy. Therefore we proceed by firstly grouping together all those exposures -taken within the same HST "visit", in the same filter and by the same PI— which are offset from one another by a distance of less than ~ 25 PC pixels, which should be large enough to account for most dithering strategies (in fact the association may be much larger than 25 PC pixels in diameter, it is only necessary that the nearest neighbour of each exposure is within this distance).

We then identify those exposures which are exactly registered (within ~ 0.10 PC pixels) and place them in subgroups for CR removal. Tests show that CR rejection is reliable and the PSF is not degraded for exposures which are aligned to within 0.1-0.2PC pixels of one another.

The roll angle is restricted so that no two frames in the same association have a difference in roll angle[3] large enough to cause a shift in pixel position of more that 0.1 PC pixels in the pixels most distant from the roll axis (which works out at $\sim 3''\!.6$).

4.3. Association Types

As a result of the above, associations may contain varying numbers of subgroups and "lone" exposures. The existence of lone exposures has implications for the processing of the association as they cannot easily be CR cleaned. Moreover, if there is only one subgroup (i.e., no dither strategy was apparent) then no higher resolution image can be obtained via drizzling (see Section 6). We have therefore identified a number of different association types depending on the grouping:

Table 1. Association types.

Asn Type	Exposures	WFPC2 pipeline
0	a single "**B**" exposure or 1 or more "**G**" exposures	OTF[a] of all the exposures
1	1 Single "**P**" exposure[b]	OTF of all the exposures
2	N perfectly aligned "**P**" exposures[b]	OTF + CR Rejection
3	N groups of M perfectly aligned "**P**" exposures[b]	OTF + CR Rejection and Coaddition (Drizzling) of all the exposures
4	N' groups of M perfectly aligned $+N''$ sparse "**P**" exposures[b]	OTF + CR Rejection and Coaddition (Drizzling) of the $N' \times M$ exposures and OTF of the N'' sparse exposures
5	N'' sparse "**P**" exposures[b]	OTF of all the exposures

[a] OTF stands for On The Fly Recalibration, a service already available at ST-ECF and CADC (Canadian Astronomy Data Center)

[b] plus 0 or more "**G**" exposures logically belonging to the association and for which really accurate pointing information is not available

5. The WFPC2 Associations Relational Scheme

All the jitter information, extracted keywords and computed values (like the RA, DEC and ROLL averages and standard deviations), are stored in a relational database table called *jit_member*.

The resulting associations, built using the information stored both in the *jit_member* table and in other HST database tables, are stored in two tables:

[3] The V3 angle in the jitter files.

- the *hst_associations* table, containing the association identifier, the total exposure time, the total number of "**P**" exposures within the association, the pointing coordinates of the leader[4] as required by the PI, and other fields describing the association (e.g., filter, proposal id, etc.).
- *hst_association_members* table, containing information relevant to the single exposure, the association identifier, the pointing information as computed from the jitter files, and, most importantly, the computed shifts in terms of X and Y PC pixels of the current exposure with respect to the so called "association leader" exposure.

6. The Automatic Pipeline

The WFPC2 pipeline is driven by an association file, which is built on-the-fly at the time of the archive request submission by a simple query to the *hst_association_members* table. The association file is composed of the dataset names, the member flags, the shifts (X and Y), and the sub-group number of all the exposures belonging to the association.

The WFPC2 pipeline consists of the following steps:

1. Run the On-The-Fly Recalibration on all exposures in the association
2. Run the `crrej` STSDAS IRAF task on any sub-group with at least 2 members
3. Run the `drizzle` IRAF task to co-add all the cosmic ray free images obtained with the previous step—if there are at least 2 such images. The shifts to be applied are provided by the association file.

7. Results/Conclusions

The jitter files proved to be a reliable source of pointing information for the Hubble Space Telescope.

It was therefore possible to define and compute associations of WFPC2 exposures. That is, the ST-ECF Archive User no longer needs to download all the exposures in a certain field to be able to work out, after hours of tedious work, what the shifts within such observations are, and which exposures can undergo the cosmic ray rejection algorithm. The shifts and the orientation are pre-computed for him/her.

ST-ECF has now a pipeline in place to automatically produce a final product out of any WFPC2 association. Such a product consists of OTF recalibrated, cosmic ray cleaned exposures co-added in a resolution-enhanced mosaic, according to the type of association selected.

Indeed, the output products of the WFPC2 association pipeline show in all their extent, especially for long total exposure times, the beauty of any high signal to noise ratio image, revealing faint objects otherwise invisible on a single, cosmic ray-covered, image.

References

Dickinson, M., & Fosbury, R., 1995, *ST-ECF Newsletter*, 22, 14

Hook, R. N., 1995, *ST-ECF Newsletter*, 22, 16

Leitherer et al., 1995, *HST Data Handbook*, Version 2.0 (Baltimore: STScI)

---

[4]The leader being defined as the exposure with the maximum number of neighbours within a limiting radius of around 0.15 PC pixels