Associations for the Wide Field Camera 3 and Preliminary Pipeline Requirements Discussion

M. Stiavelli, R. Henry, H. Bushouse, W. Hack
January 2000 - revised August 2000

ABSTRACT:
The purpose of this ISR is to recommend a structure of associations for WFC3 data and to document the rationale for this recommendation. We will also discuss aspects of the pipeline requirements related to the definition of associations.

1. Introduction
The Wide Field Camera 3 is a two channel instrument for imaging in the optical and near-ultraviolet (UVIS channel) and in the near-infrared (IR channel). The design and operations of the UVIS channel are based on ACS/WFC heritage. The IR channel will operate in a single mode essentially identical to the NICMOS MULTIACCUM mode. As a cost saving measure and in order to allow as much software re-use as possible, changes with respect to the ACS/WFC for the UVIS channel and NICMOS for the IR channel have been kept to a minimum. Within this context, the problem arises of how to design associations for WFC3. In fact, since the two channels of WFC3 derive their heritage from different instruments there are three possible options:

1. reduce software development to a minimum accepting that the two channels will handle associations differently;
2. make the two channels look similar by simplifying them to the lowest common denominator;
3. make the two channels look similar by upgrading them to a common enhanced capability.

Clearly, option 3 would be the most desireable from a science point of view even though it may require more budgetary resources. The purpose of this ISR is to provide an overview of the structure of associations currently implemented for NICMOS and ACS, to
discuss their applicability to WFC3 (respectively in Sections 2 and 3), and to recommend a cost-effective approach for WFC3 (Section 4).

2. The NICMOS Associations in the context of the WFC3 IR channel

The philosophy from which “associations” were first developed applies to WFC3: the combination of data from multiple exposures is necessary for basic data reduction steps such as cosmic ray removal and PSF sampling improvement (by dithers and drizzling) and is also necessary to produce a scientifically viable data product.

Each NICMOS exposure, itself consisting of several detector read-outs, results in a dataset consisting of several files: an SPT file containing the standard header packet and a unique data log file (UDL), a raw data (RAW) and a calibrated data file (CAL). In the second step of the NICMOS calibration pipeline, carried out by CALNICB, one or more mosaiced files (MOS) are produced (see NICMOS ISR 010, Mackenty and Baum, 1996, “Associations of NICMOS and STIS Exposures: an extension of the HST data processing pipeline” for more information). RAW, CAL, and MOS files have a similar structure and contain image data, errors, data quality, samples, and integration time arrays. They are all FITS files with image extensions. The products of CALNICB make use of the description of the association contained in the association table (ASN.) The MOS products of CALNICB may be either combinations of several exposures at the same pointing or mosaiced images where exposures at different pointings are combined to obtain a larger field of view. The combination of these mosaiced exposures is done by using a source-driven cross-correlation technique to identify the shifts between images and then by registering each image to the first in the sequence using a bilinear interpolation.

Let’s consider the technical issues relevant for the application of such a scheme to WFC3:

1. **File Size.** A single WFC3 read-out is 16 times larger than that of NICMOS so that the size of a 16 read-out WFC3 MULTIACCUM sequence is 32 MBytes. The total file size including 16 samples, a 32 bit error array, a 16 bit data quality array, and a 32 bit exposure time array is 42 MBytes. If error, data quality, and exptime arrays are included for each sample image of the MULTIACCUM, the file size grows to 224 MBytes. These file sizes are much larger than the equivalent of NICMOS but do not appear to be excessively large for present day computers. Similarly to ACS, the WFC3 pipeline should probably include a switch to delete the intermediate products.

2. **Combination algorithm.** The bilinear combination algorithm of NICMOS is useful to users who are dithering mainly to mosaic the individual exposures into a larger contiguous field of view. The algorithm cannot take advantage of subpixel stepping in order to improve the PSF sampling. In WFC3 we expect that mosaicing will be comparatively less important than for NICMOS because: i) the camera has larger field of view and ii) WFC3 lacks the higher sampling camera options present...
in NICMOS (NIC1 and NIC2). The latter implies that a larger fraction of users will carry out sub-pixel stepping to improve the PSF sampling. A simple NICMOS-style combination will only be of moderate use for WFC3 unless the code can be modified to combine the images using drizzling.

3. **Geometric distortion.** The geometric distortion in WFC3 is more severe than for NICMOS, partly due to the larger field of view. For this reason any automated combination of images at different pointing should take geometric distortion into account. This could easily be done using drizzle as it is done for, e.g., WFPC2.

4. **Bad pixels.** The presence of bad pixels is another reason for dithering. However, bad pixels can be eliminated through dithers by a small integer number of pixels. The reason why the displacement needs to be small is that geometric distortion would convert any large displacement by an integer number of pixels at the field center into a non-integer displacement far from the field center. The image combination present in CALNICB is probably adequate for bad pixel elimination in WFC3 since small displacement would be sufficient for this purpose.

5. **Subarrays.** WFC3/IR will include subarrays which were not supported in NICMOS. However, through the use of keywords (e.g., offset values), the subarrays should be easily identified by the pipeline.

   Based on the above considerations it would appear that WFC3 IR could make use of the NICMOS association and pipeline processing as is but that it would benefit from improving the combination algorithm to better accommodate sub-pixel stepping and geometric distortion calibration/correction. These steps could be carried out using the drizzle technique. Considering the performance of present day computers drizzle run times should be acceptably short, especially for the 1 K by 1K size of the individual WFC3/IR MULTI-ACCUM products. The best strategy for determining the drizzle parameters should be studied in more detail.

### 3. The ACS associations in the context of the WFC3 UVIS channel

The WFC3 UVIS channel uses the same detector package as the ACS/WFC channel. WFC3/UVIS introduces 2 by 2 and 3 by 3 on chip binning which was not present in ACS. Assuming that the calibration strategy of binned frames can be kept simple, WFC3/UVIS should behave identically to ACS/WFC, leaving open the possibility for significant software re-use. In particular, the ACS associations could be used without change for WFC3. The ACS associations can be thought of as an extension of those of NICMOS. In fact, their logical structure fully supports multiple products (see ISR ACS-99-03, Hack, 1999, “CALACS Operation and Implementation” for more information). In the present implementation of the pipeline, the products corresponding to CR-split exposures (at the same pointing) are combined.

These combined images are logically associated to produce a single dithered product but, for reasons explained below, the present pipeline does not actually carry out this last
step. The ACS patterns make a distinction between **dithering**, which is meant to consist of small shifts to improve the PSF sampling or eliminate bad pixels, and **mosaicing**, which is meant to increase the contiguous field of view imaged in a given observation. The decision not to implement in the pipeline any automatic scheme of combination of dithered exposures was the result of a number of considerations (see e.g. ISR ACS 98-02, Stiavelli et al., 1998, “Dithering Strategies for ACS”), namely:

1. The small pixel size and large image size of ACS makes the straight application of a cross-correlation based algorithm to determine image shifts less desirable than e.g. for NICMOS. The use of jitter-information would be better suited to this end but would require either a change to the flight software or a dependency of the science stream on the engineering data.

2. The significant geometric distortion of ACS requires the use of drizzle for the image combination and thus implies a significant execution time given the size of ACS images.

3. At the time, the drizzle software was not straightforward to apply without human intervention and thus it may have lead to final products unsuitable for a large fraction of science applications.

The above considerations also apply to the UVIS (and IR) channel of WFC3. Some important differences are: i) that our experience with drizzle may make it better suited to automated processing now than it was few years ago; ii) that the constantly improving computing power of modern workstations also makes the excessive execution time concern less relevant.; iii) that the higher blue sensitivity of WFC3/UVIS compared to the ACS/WFC - with a similar pixel size - will probably increase the fraction of WFC3/UVIS proposals making use of sub-pixel stepping compared to those for ACS/WFC. The strongest of the objections to automated combination in the pipeline remains the need of using jitter information for determining the shifts. Further testing may be required to assess whether the jitter data are indeed necessary.

ACS associations make use of the following MEMTYPEs:

- **EXP-CRn** : for the input CR-SPLIT exposures for pointing $n$
- **EXP-CRJ** : special case for $n=1$, i.e. single pointing
- **PROD-CRn** : for the CR-combined exposure at pointing $n$
- **PROD-CRJ** : for the CR-combined output product in the $n=1$ case
- **EXP-RPh** : for the input REPEAT-OBS for the combined image $n$
- **PROD-RPT** : REPEAT-OBS combined output product
- **PROD-DTH** : dither-combined product (in principle obtained by combining the PROD-CRn outputs)

The main difference between CR-SPLIT and REPEAT-OBS in ACS is that REPEAT-OBS observations are simply co-added rather than undergoing a cosmic ray rejection
before combination. In the case of WFC3 there is no strong science case in favor of REPEAT-OBS.

The present IPPPSSOOT file name structure appears to be adequate for our needs. The maximum number of pointings combined into a dithered exposure is set by the index \( n \). This index can be assumed to be in the range 1-i (18 values) without significant loss of generality. In fact, the ACS DITHER patterns of type BOX have only four pointings while those of type LINE can be limited to 18 pointings without much impact. The value \( n=0 \) is reserved for the combination product (PROD-DTH). The maximum number of images in a single intermediate product PROD-CR\( n \) is set by the values of the second character from the right in the IPPPSSOOT name. In principle this character could assume up to 26 different values, i.e. there can be up to 26 images combined into a single PROD-CR\( n \).

4. A proposal for WFC3

The ACS associations appear to be general enough to be used for both channels of the WFC3 camera. These considerations lead to the following proposal:

1. The WFC3 pipeline will use the basic concepts developed for ACS/WFC and NIC-MOS. Two options are possible: (i) have a common CALWF3 task starting a flow similar to ACS for WFC3/UVIS and a modification of CALNIC for WFC3/IR; or (ii) try to integrate some of the processing into modules. In the latter scheme one could e.g. imagine a WF32D module performing DQI, linearity correction, dark subtraction, flat fielding, shutter shading correction on both UVIS and IR data. Investigating which of these two options is the most convenient is not within the scope of this document.

2. The associations will be similar for both channels and based on the ACS associations. No automated combination of dithered exposures will be implemented for WFC3 UVIS while WFC3 IR will initially retain the combination scheme of NIC-MOS based on bi-linear interpolation so as to be able to correct for bad pixels.

3. In order to simplify processing of on-chip binned data in the UVIS channel we can assume that reference files for the binned mode will be present in CDBS. This should not have a large impact on either the total volume of reference files nor the calibration overhead.

4. The automated combination of dithered exposures will be considered as a future expansion. This step will not imply any change to the structure of WFC3 associations. Unless the ACS pipeline is modified in the near future to include such automated combination, the WFC3 group will need to revisit the issue of how to efficiently determine relative image shifts before beginning the implementation of any automated combination.

5. To handle sub-array calibration for IR data, the STIS methods can be adapted to the WFC3 pipeline, especially if “modules” as described in 1) are used.
5. Acknowledgements

We thank Olivia Lupie and Daryl Swade for a careful reading of the manuscript and useful comments and suggestions.