CALIBRATION OF WFPC2 IMAGES
FOR THE HST MEDIUM DEEP SURVEY

K. U. RATNATUNGA, R. E. GRIFFITHS, L. W. NEUSCHAEFER and
E. J. OSTRANDER
Johns Hopkins University,
Baltimore, Maryland 21218

Abstract. We discuss some specific aspects of the calibration of HST WFPC2 data for analysis of faint galaxy and stellar images in the Medium Deep Survey (a Key Project). We comment on WFPC2 flat fields, discuss the removal of cosmic rays from two images which may not be in the same filter, the removal of Earth-light contamination patterns in the sky background, and the generation of an error image for a stacked data set. We also discuss the aspect jitter data and justify the need for this information as an adjunct to parallel WFPC2 observations since HST refurbishment.

Key words: WFPC2 – calibration – jitter

1. MDS PIPELINE

The Hubble Space Telescope (HST) Medium Deep Survey (MDS) Key Project (Griffiths *et. al.* 1994), includes parallel observations of a large number of random fields extending over multiple years and focuses on the statistical properties of all measurable objects (numbering tens of thousands). Much effort in the MDS project has been devoted to: (i) improving the calibration, (ii) selecting the optimum set of procedures for the maintainance of a uniform calibration for the survey and (iii) to minimizing observational errors. These improvements are particularly important for quantitative analysis of the faintest extended sources with low signal-to-noise ratios.

During the workshop held to discuss calibration of pre-refurbishment WF/PC data (see also Ratnatunga *et. al.* 1994), the major concern was that the pipeline flat fields had significant gradients. When skyflats were made from WFPC2 MDS data in July '94 it was clear that, except for a small error close to the edge of the CCDs ($\sim$20 pixels), there were no significant gradients in the WFPC2-IDT flat fields. This minor error is probably caused by the treatment of the boundary conditions in the algorithm used to combine low frequency information from Earth-observation flats with the pixel-to-pixel variations from pre-launch thermal-vacuum flats. The MDS has done the same kind of combination, using our skyflats to define the very large scale flatfield structure. The small number of independent survey fields available to us in July '94 did not generate a sufficiently good skyflat. Now, almost a year later, there is a much larger data base of independent survey fields available and a new skyflat may be useful for the WFPC2 Hubble Deep Field (HDF) to be observed in Dec '95.

Details concerning the bias, darks and hot pixels are discussed elsewhere in these workshop proceedings (Biretta *et. al.* this volume). The Medium Deep Survey pipeline uses bias and dark images which are super-stacks of all available calibration images. Hot pixels detected during each decontamination period are recorded on mask images and flagged. We start from raw MDS observational data in UNIX compressed format. This raw data then takes only about 1.5 MB per WFPC2 dataset and is easily stored online. To ensure uniformity in the survey and simplify the calibration procedure, we have developed a single fully automated shell script to calibrate and stack any set of MDS observations for subsequent image analysis. A number of calibration issues needed to be resolved for the Medium Deep Survey, since the available parallel data is most often not optimal in terms of multiplicity of exposures for each filter. Some of these issues are discussed below.

2. Stacking data

As an example, consider the stacking of a dataset consisting of two exposures in the same filter taken along the same pointing vector and in the same orientation. To illustrate a complicated case, we choose a pair of F814W observations, where one exposure has a significant scattered Earth light pattern. The raw grayscale image of WF4 from this observation in Figure 1a, clearly shows the cosmic rays and scattered Earthlight pattern.

2.1 Earth Light

A simple way to estimate the scattered Earth light in Figure 1a is to subtract from this image another image along the same line-of sight which is free of scattered Earth-light. After such subtraction we use a σ-clip to remove cosmic rays. Convolution by a filter having a width of 25 pixels (FWHM) results in a sufficiently smooth Earth-light pattern (Figure 1b) that it can be subtracted from the contaminated image before combination with the uncontaminated image, using a 3σ clip to remove cosmic rays.

2.2 rms error image

The resulting combined image (Figure 1c) does not have the scattered Earth light pattern and excludes most of the cosmic rays. Note however, that photon noise from the Earth light is reflected in the error image (Figure 1d) which represents the root-mean-square (rms) of every pixel in the calibrated stack. The error image is important for quantitative image analysis of the observation, and is based on the noise model. It reflects both the cosmic ray pixels rejected in the stacking procedure, the flat field response, and any subtracted scattered Earth-light. When we combine images along the same line of sight (no shifts) we stack the image before flat fielding and then multiply both the stacked image and the rms error image with the inverse flat field. If the exposures comprising the field have been shifted relative to each other, limitations in the current software require us to first flat field and multiply the rms error image with a similarly shifted and weighted sum of flat fields.

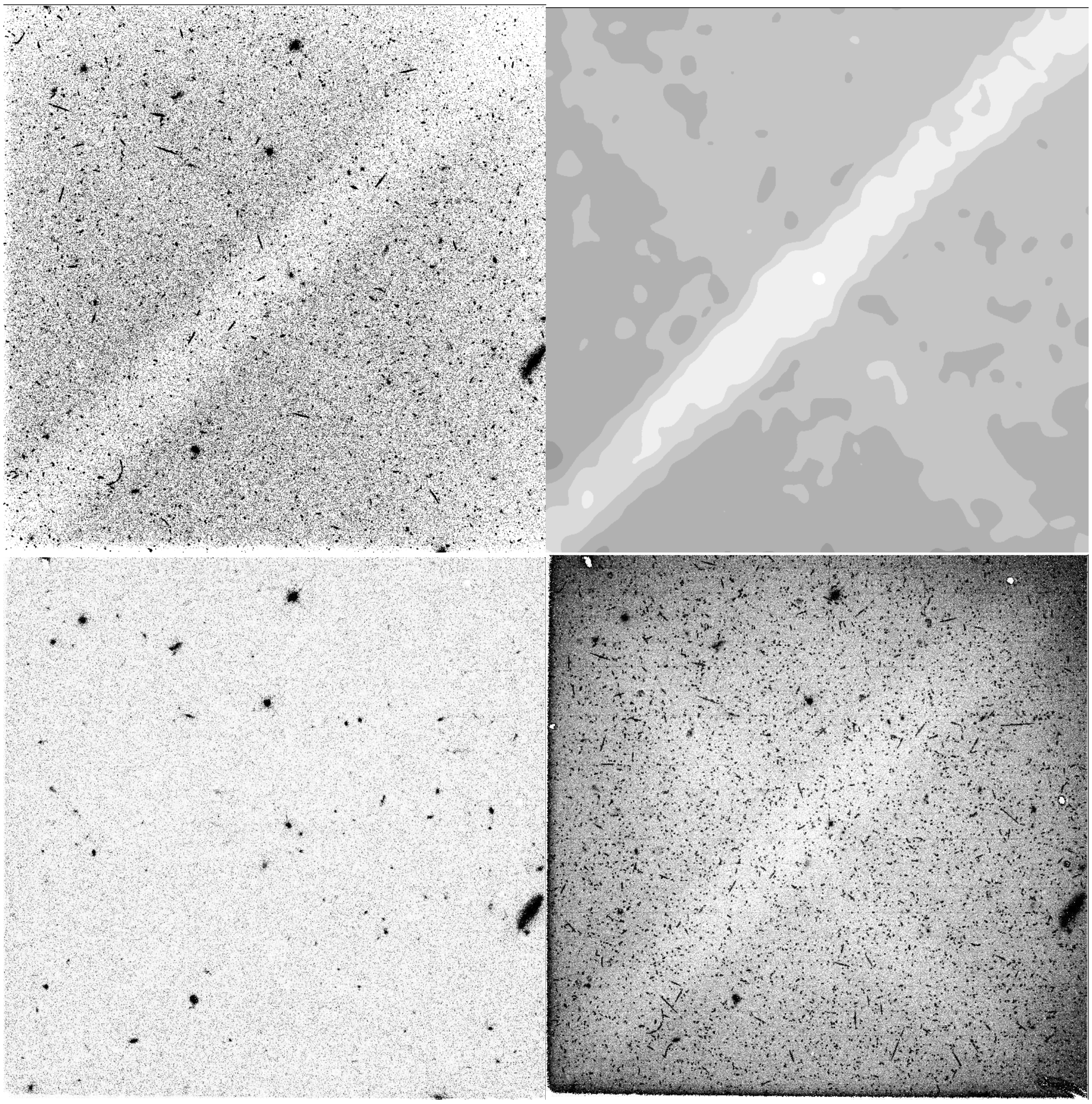

Fig. 1. (a) upper-left : grayscale image of raw u26x0o03p[4], (b) upper-right : estimate of scattered Earth-light (smooth) (c) lower-left : calibrated stack image (d) lower-right : rms error image for calibrated stack

2.3 Cosmic ray removal in two-image stacks

Cosmic rays affect on average about seven pixels per second per CCD between start of exposure and final readout. When three or more exposures in the same filter are available along the same direction and with the same orientation, cosmic rays can be very effectively removed by stacking the images using a 3σ clip. However, for many of the MDS fields we have only one or two images with the same filter along the same line-of-sight. Automated cosmic ray removal in single images is a more complex

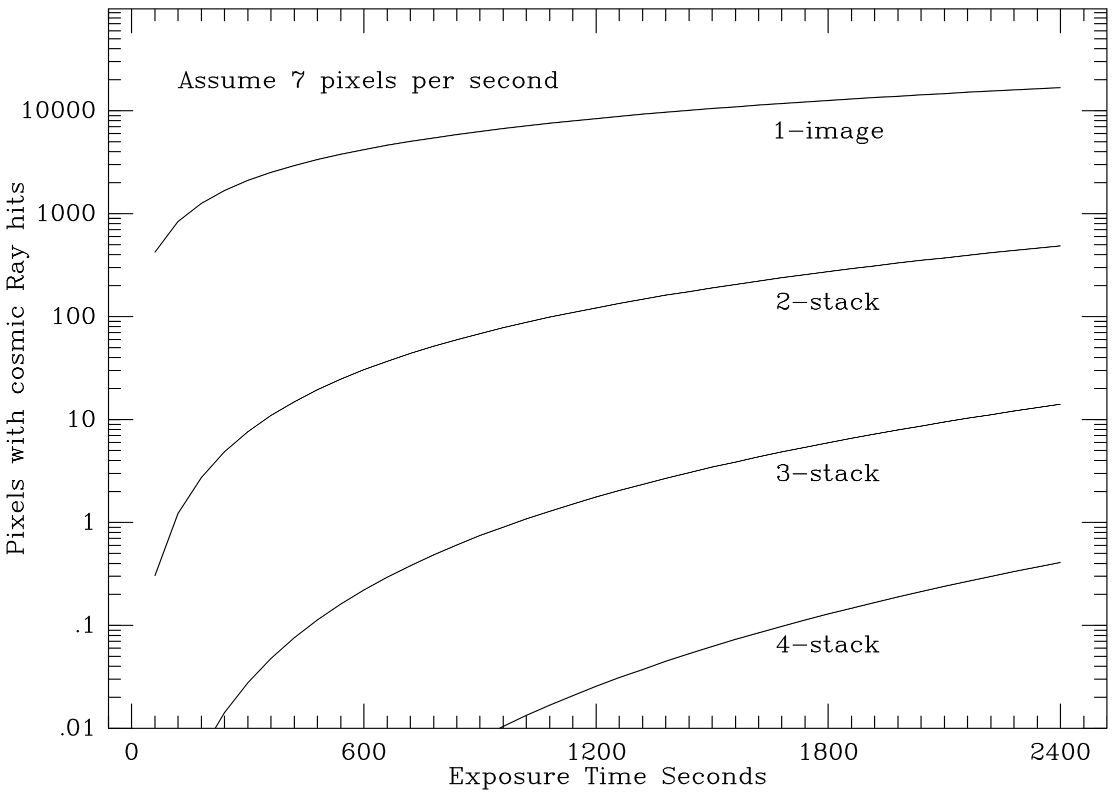

Fig. 2. Expected number of pixels with cosmic ray hits on all images of a multiple stack.

problem (Salzberg *et. al.*). Although single exposures may not be so useful for the statistical analysis of faint point-like objects, they can be used to improve the number statistics of relatively bright field galaxies (I $\sim$ 21 mag) and also in optimizing the search for rare objects such as gravitational lenses or other serendipitous discoveries. These aspects of the survey require us to maximize the number of MDS fields processed.

In Figure 2 we illustrate the expected number of cases where the same pixel is hit by a cosmic ray in multiple images as a function of the exposure time in each. When only two images are available, the stacking operation will leave the fainter cosmic ray on the output image. As a result, spurious objects may be assigned to about 400 pixels per WFPC2 CCD chip in a pair of 40 min exposures. These coincident hits need to be cleaned out since they are significantly larger in number than real objects detected in such an exposure. There are only about 10 expected hits of the same pixel in a stack of three images and less than one pixel in a stack of four images.

Most cosmic ray events in WFPC2 images are spread over several to tens of pixels, so that for a given pixel rejected in both exposures there is a significant probability of its neighboring pixels also being affected by cosmic-rays from both exposures. We therefore look at the distribution of pixels rejected as cosmic rays in each image to predict which pixels have a high probability of being hit in both images. Such an algorithm finds about 1500 such pixels, whereas we expect about 400. We conservatively throw them all away, since this amounts to only 0.25% of the pixels. Many more (~1%) are discarded as hot pixels.

The image that we use for illustration and testing our algorithm is actually a stack of two images from four – the complete stack does not suffer from this kind of contamination. So we compared the removal of spurious objects using the procedure and found the resulting image to be clean and suitable for detection of only real objects. This procedure works well only for images which are coincident or those which are shifted exactly an integer number of pixels relative to each other. Otherwise the stacking could delete a significantly brighter pixel from a point-like image because of the difference in the distribution of the image flux on the CCD pixel grid. The algorithm would then assume these to be cosmic rays in the two-image stack and reject pixels close to the centers of point-like images.

Two exposures in two different filters can also be used for the removal of cosmic rays at the expense of rejecting some of the faint and point-like images. For galaxies of average (V-I) color, the detected flux through both of the primary MDS filters (F606W and F814W) is the same for equal exposure times. The observations can therefore be stacked directly without scaling to detect and reject most of the cosmic rays in each of the individual filter bands. However, to avoid deletion of the cores of bright galaxies with a wide range of color, the algorithm described above could not be used. The removal of these spurious objects is done by filtering out all faint and point-like objects from these catalogs. We must subsequently account for this object selection method in the analysis of these incomplete catalogs.

2.4 Shifted image stacks

Shifted images are useful to reject hot pixels and to average over noise in the calibration. Otherwise the calibration noise is coherent and becomes the dominant term in the error distribution and the critical factor in reaching fainter limiting magnitudes in stacks of multiple images. The tolerable amount that the images can be shifted is limited by the differential change in distortion across the WFPC2 chip. A 4 arcsec shift causes a 0.2 pixel rms error, whereas a 2 arcsec shift causes a 0.1 pixel rms error and can probably be ignored. The latter error is of the same order as the precision in the re-acquisition of a target, which is about 0.07 pixel rms, and is sufficient for stacking with or without shifts of an integer number of pixels. The precision of 'new acquisition' is about 0.15 pixels rms, so that integer shifts should not be expected even if requested.

Non-integer shifts are useful to improve the undersampling of WFPC2 images which are used to reconstruct images at higher resolution than the pixel size (Hook & Adorf, this volume). The relative shifts need to be determined either empirically or from the pointing data provided in the OMS class files extracted from the engineering data. However, optimization of the shifts to minimize noise would be irrelevant for the Medium Deep Survey, since the survey data are taken on a 'no-interference' basis, i.e. parallel observations have no influence on the spacecraft pointing. In general we avoid the use of specialized image reconstruction software at the cost of some increase in the image jitter. When three or more images are available at each individual pointing of a multi-exposure observation, each set is stacked separately and then averaged without σ clipping, or else analyzed independently. The jitter map for the resulting image can be constructed by adding the jitter maps for the individual images allowing for the subpixel shifts.

2.5 Jitter image

In Figure 3 we show jitter images for the 64 datasets of the MDS and GTO parallel observations from 28 March to 6 May 1995. This representative set of jitter images for parallel observations shows the complexity and differences in jitter. They have been ordered in increasing rms jitter.

Since the differential aberration corrections are done in a manner appropriate for the primary instrument, parallel observations with WFPC2 can suffer from a drift in aberration over a long exposure. For some observations the change in the differential aberration between the primary and the WFPC2 is significant and causes a drift of over 0.5 WFPC2 pixels. This is particularly true if the primary instrument is the FGS, which has an aperture furthest from WFPC2. In Figure 4 we plot the rms jitter of the major axis and the axis ratio of the jitter ball for all 584 parallel observations since 20th October 1994. The points have been coded with the identity of the primary instrument. Some FGS observations lead to large drifts, while some FOS observations show a significant jitter ball.

If the jitter is sampled uniformly over the observation period, then the time dependence is unimportant for analysis of images which are integrated over time. However, the jitter in parallel observations is also affected by independent events that occur in the primary observation timeline, as is evident in the time sequence of the pointings given in the jitter image. For example, if the primary observation started or stopped during the parallel observation, then the differential aberration correction would have been switched off for some part of the parallel observation, resulting in a slightly different pointing for that period. The jitter in the parallel observation would then be bimodal.

Note however that some of the apparently bimodal jitter images as currently generated are instead created by a loss of jitter information over significant periods of the observation, which leads to an apparent break in the drift. These occasional periods of missing jitter data need to be interpolated using the regression of the drift, and the scatter about this regression, to ensure that the generated jitter image is a reasonable estimate of the jitter over the complete duration of the observation.

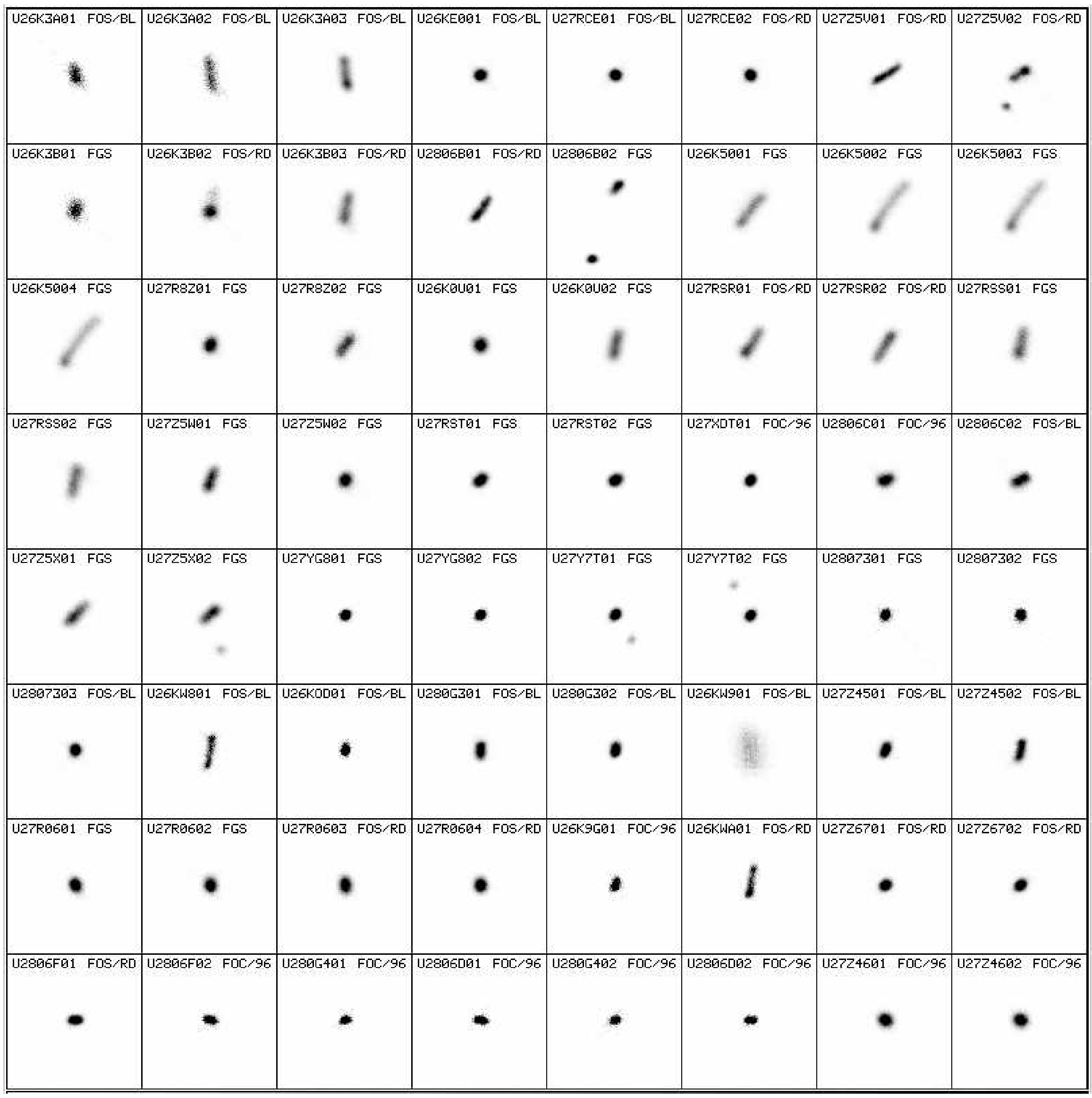

Fig. 3. Jitter images for all pure parallel WFPC2 observations in a 40 day period. Each box represents 100 mas or 1 WFC pixel. Dataset name and primary instrument is indicated.

3. Discussion

We have shown how Earth-light patterns can be subtracted from contaminated images. We observe that although the relative normalizations of each branch of the shadow across each CCD changes between exposures, the pattern remains the same. If an uncontaminated image is not available along the same line-of-sight for the pattern estimate, it may be possible to select an Earth light pattern from another field and subtract it with suitable normalizations. This possibility has not yet been fully explored. To enable such a subtraction, it would be useful to have a model estimate

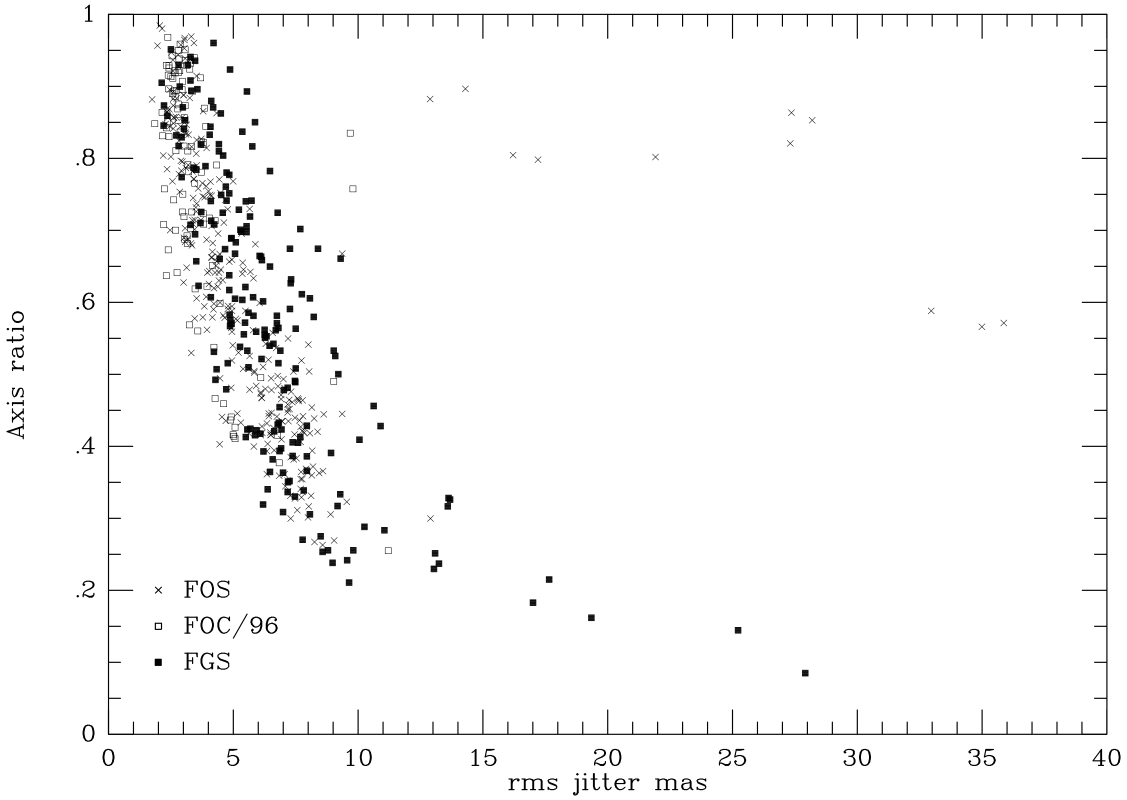

Fig. 4. Axis ratio and major axis rms for all available pure parallel WFPC2 observations. Points have been coded with and primary instrument as indicated.

for the integrated earth light contamination rather than just the peak given in the OMS file (having the extension .cmh).

A proper rms error image is not created by the STSCI calibration pipeline or the STSDAS combine software. A modified version of the STSDAS combine software called mdscombine can be picked up via anonymous ftp from mds.pha.jhu.edu and used to stack and create a rms error image. This version also corrects some software errors, which are important when stacking images of unequal exposure times. The old STSDAS task combine, which had many errors, was replaced in July '94 by a new STSDAS task called gcombine which was based on the corrections made by us in Nov '93. We have not used the new task and are unable to comment on it.

Jitter information has been extracted from the HST engineering data for all observations taken since 20th October 1994, and made available in the HST archive as OMS files. However, to incorporate them consistently into our pipeline, these files should be created for all survey data since the start of WFPC2 observations in January 1994. We expect STScI to generate and make available the OMS files for all parallel WFPC2 observations, or else provide the automated software that has been written so that interested users can extract the engineering data from the archive and create the OMS files with jitter information. This is essential to ensure a uniform HST database for all observations since refurbishment.

The header keywords of the .cmh file include most of the relevant information, such as Earth limb angle, and rms jitter along the v2 and v3 axes. It would be very useful if the cross term in the rms jitter (needed to define the orientation) were also computed, and all this information digested into the HST database so that it could be retrieved via StarView to identify, for example, those datasets with excessive jitter.

The Medium Deep Survey has been used to measure sub-pixel half light radii for thousands of galaxies at very faint apparent magnitudes. In order to quantitatively separate faint stars from galaxies with sub-pixel scale lengths, and to minimize any bias in the measurement of the orientation and axis ratios of these galaxies (e.g. to detect weak gravitational lensing effects), jitter information is very important.

The software and scripts for the calibration procedures discussed above can be obtained from the authors, and will shortly be made available on the web.

Acknowledgements

This paper is based on observations with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. The Medium-Deep Survey is funded by STScI grant GO2684 and daughters.

References

Griffiths, R. E. et al.: 1994, *ApJ*, **437**, 67.

Ratnatunga, K. U., Griffiths, R. E., Casertano, S., Neuschaefer, L. W., & Wyckoff, E. W.: 1994, *AJ*, **108**, 2362.

Salzberg, S., Chandra, R., Ford, H., Murthy, S. K. and White, R.: 1995, *PASP* **107**, 279.