EPOXI – Hartley 2 Calibration Pipeline Summary

1.0 Calibration Pipeline Processing

This report summarizes the pipeline as of June 2011 used to calibrate version 1.0 of the EPOXI Hartley 2 data currently in the NASA Planetary Data System (PDS). For a full discussion refer to Klaasen, et al., *EPOXI Instrument Calibration* (2011, in preparation).

1.1 Standard Steps

For each image acquired during the EPOXI encounter at comet Hartley 2, a standard set of procedures and settings is applied in our pipeline processing in order to calibrate the images automatically (see Figure PIPELINE). In general, these default settings are the best the science team has been able to derive for the data set as a whole and thus do not necessarily reflect the best possible processing for any particular image. For special cases, the automated pipeline has the ability to specify special settings for particular observations.

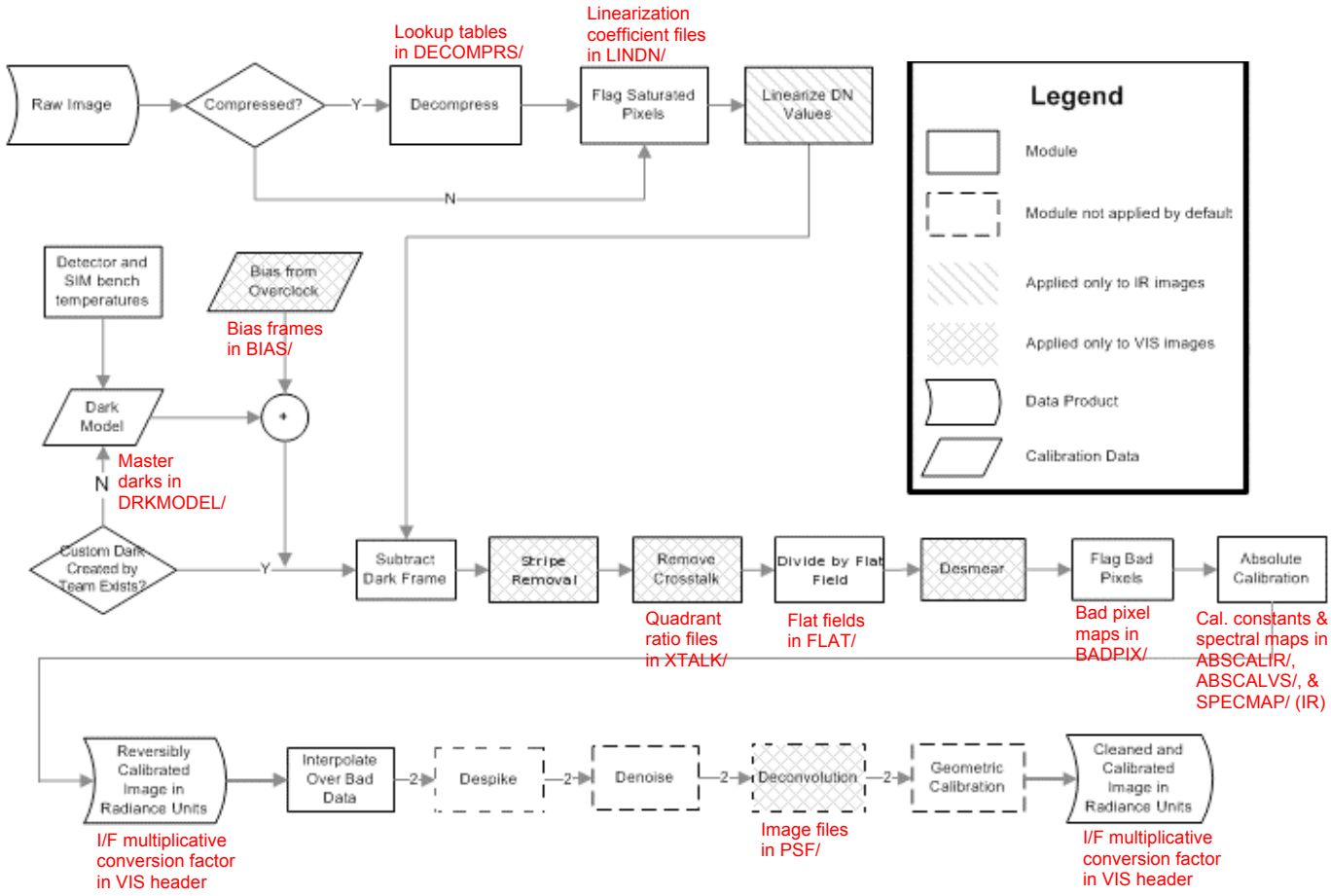

Figure PIPELINE - This flowchart describes the pipeline of as June 2011 used to calibrate EPOXI Hartley 2 data. Some modules are not applied to all instruments. Modules not applied by default were turned off for Hartley 2 data. Input calibration files, such as the look-up tables for decompression, are identified by red text with the CALIB/ subdirectory within a PDS dataset. The pipeline is based on the one used for Deep Impact as discussed in Klaasen, et al., *Deep Impact Instrument Calibration*, (2008).

The standard pipeline begins by decompressing the image if it was compressed on the spacecraft. Images can be compressed using one of four 14-bit to 8-bit lookup tables optimized for different types of exposures. To uncompress the images, a reverse lookup table is used which maps each 8-bit value to the average of all corresponding 14-bit values.

All saturated and missing pixels are flagged in the quality map. Then an IR image is linearized using the derived correction function. Linearization is the first step for IR data because the IR output data represent a read-minus-reset DN value causing all output DN values to represent signal collected in the detector (any fixed bias is subtracted out) and to be subject to the non-linearity of its response. A VIS image does not need this step because the instrument responds linearly.

Next, a dark frame is subtracted from the image. If a dark frame has been created by the science team for the specific observation, then it is subtracted. Otherwise, a dark model is used to generate the frame (for a thorough discussion of dark pattern removal, please see the published version of the EPOXI Instrument Calibration paper).

After the dark subtraction, a VIS image undergoes a few extra processing steps not taken by every IR image. First, residual stripes extending along the rows of each image quadrant are examined. If they can be accurately measured, they are removed from the image and stored as an image extension. Then, the electrical crosstalk is removed by subtracting a derived ghost frame. Each quadrant in this frame is a linear combination of rotated versions of the other three quadrants. Next, the image (both VIS and IR) is divided by a flat field in order to account for variable responsivity across the detector. Lastly, VIS CCD transfer smear is removed using the parallel overclock rows if the image was taken in modes one through six or a column averaging approximation if the image is in modes seven or eight.

After bad pixels are flagged, the image is radiometrically calibrated to produce a radiance image in $W/[m^2\,sr\,\mu m]$. For a VIS image, this is simply done by dividing the image by integration time and then multiplying by the appropriate conversion factor for the given filter and desired output. Once the VIS radiance image is created, the pipeline computes a multiplicative factor (pi divided by the solar spectrum at the target's distance from the sun) to convert from radiance to I/F and stores the factor in the VIS header. For an IR image, the procedure is more complicated as the absolute calibration is wavelength dependent, which in turn is temperature dependent. First, the wavelength and bandwidth for each pixel are calculated. Then, each pixel is multiplied by the appropriate wavelength-dependent calibration factor and divided by the pixel's integration time and spectral bandwidth.

At this point, a reversible radiance image data product has been created, and a copy is run through the rest of the pipeline, which performs a series of non-reversible steps. First, the data are interpolated over the bad pixels and gaps. For a VIS image, this interpolation is performed using thin plate splines anchored by the valid data around the edges of each hole. For an IR image, a linear interpolation is performed in the spatial dimension only. Next for a VIS image, the pixel values in the overscan rows and columns that border the active area of the CCD are not preserved as in RADREV but are overwritten with 0.

Next, an optional despiking routine is applied in order to remove cosmic rays. This routine performs a sigma filter by calculating the median of each NxN box, where N is odd, and then replacing the central pixel with the median if it is more than M median deviations from the median. By default, both M and N are set to 3. The median deviation of a set S is defined as: $Med(|S - Med(S)|)$.

Lastly, a VIS image may be deconvolved using the derived point-spread function of the instrument. This optional step is especially important for the HRI-VIS instrument, which is out of focus.

1.2 Calibration Quality Map

Along with each calibrated image, a byte map is created that defines the data integrity for every pixel. For each byte in the map, representing one pixel, each bit acts as a flag that is set to 1 if the given criterion is met for that pixel. These flags are:

MSB							LSB
7	6	5	4	3	2	1	0

0. Bad Pixel - This pixel is a known bad pixel.
1. Missing - The datum for this pixel was not received from the spacecraft.
2. Despiked - This pixel was modified by the despiking routine.
3. Interpolated - This pixel has been reclaimed by interpolating from its neighbors.
4. Some Saturated - The raw value for this pixel is above the point where some pixels become saturated. For VIS instruments, this occurs at 11,000 DN, while for the IR spectrometer, this occurs at 8,000 DN.
5. Most Saturated - This raw value for this pixel is above the point where most pixels are full-well saturated. For VIS instruments, this occurs at 15,000 DN, while for the IR spectrometer, this occurs at 11,000 DN.
6. ADC Saturation - The ADC was saturated for this pixel.
7. Ultra Compressed - The pixel was in a compression bin so large that the value contains very little information.

For example, if the pixel is bad and has been reclaimed by interpolation, the decimal value in the quality map will be $2^0 + 2^3 = 9$. In the normal FITS format for the calibrated image, this map exists as the first image extension.

1.3 Signal-to-Noise Ratio Map

In order to provide more information to the end user, the last extension of the image contains a map estimating the signal to noise ratio for each pixel. The signal is taken to be the dark- and bias-subtracted image value in 14-bit DN, while the noise estimate consists of the root-sum-squared of three different noise sources: shot noise, read noise and quantization noise. The shot noise in 14-bit DN is defined as:

$$N_s = \sqrt{\frac{Raw - Bias}{K}}$$

where K is the gain in electrons/14-bit DN and is dependent on the instrument and mode, and *Raw* and *Bias* are in 14-bit DN. For the IR spectrometer, *Bias* is 0 by definition except in Mode 6. The quantization noise is defined as:

$$N_q = \frac{Q}{\sqrt{12}}$$

where Q is the quantization step in 14-bit DN (N_q represents the rms error introduced by quantizing a signal whose true value could lie anywhere within the quantization bin with equal probability, i.e.,

$$\sqrt{\frac{\int_{-\Delta/2}^{\Delta/2} \in^2 d \in}{\Delta}}$$

where Δ = the quantization step size and $\in$ = the quantization error). For uncompressed data, Q depends on the performance of the ADC, while for compressed data, Q is set to the bin size in the decompression lookup table that the pixel used or to the uncompressed Q value, whichever is larger. The parameter values needed for the noise calculation were determined from instrument calibrations and are shown in Table 1.

Instrument	**K (e/DN_{14})**	**Uncompressed Q (DN_{14})**	**Read Noise (DN_{14})**
IR Unbinned	16	1	5.0
IR Binned	64	1	3.0
HRI	27.4	2	0.7
MRI	27.2	2	1.0
ITS	30.5	2	1.2

Table 1 - Noise parameters determined in calibrations of all instruments.

1.4 Spectral Registration Maps

In an IR image product, the second and third extensions are pixel-by-pixel maps of the spectral registration for the image. The second extension contains the effective wavelength of the pixel, while the third extension contains the spectral bandwidth.

1.5 Optional Steps

Beyond the automated calibration pipeline described here, a manual calibration can be performed where the user can specify his/her own settings and calibration files for each step. Also, any processing module can be disabled, and there are two extra ones that can be enabled. The first such module is a noise-reduction module that is applied after the despiking routine. This module applies the BayesShrink wavelet thresholding algorithm with a robust mean noise estimator to remove some of the noise. The other step that can be enabled applies a rubber sheet geometric distortion correction. This function is not normally applied, as the optical distortion though the telescope is minimal.

2. Pipeline Updates

The updates to the instrument calibrations derived during EPOXI required changes to the data processing pipeline software as it existed at the end of the DI mission. In addition, science team internal reviews and peer review of previous DI data slated for archive brought to light some errors in the existing calibration code that were corrected.

2.1 New files and constants

The simplest pipeline updates involved adding updated, or removing obsolete, calibration files and calibration constants that track the time-varying performance of the instruments. These were added to the database of previous calibration files along with the applicable time periods for each, and no software changes were usually required for the pipeline to be able to use these files. In addition, the new calibration files, stored in FITS format, were given enhanced labels with comments describing provenance, pixel orientation, and how each file is meant to be used in DI data calibrations. These files are documented in the PDS archive of the calibration, and are summarized here by the first date (YYMMDD: YY=year; MM=month; DD=day of month) of data affected by each set of files:

050112 Revised all MRIVIS calibration constants
071004 Updated IR absolute calibration based on Beta Hyi observations, including 2x correction
050112 Updated IR flat fields that include the ASF and removed previous IR flat fields from calibration database
050112 Updated IR spectral smile equation
071001 Updated IR linearization coefficients
071001 Updated IR master dark file
071004 Updated IR bad pixel maps from 01/2008 calibration
071004 Updated IR read noise values from 2008 calibration
071004 Updated HRIVIS 950-nm filter calibration constant from 2008 calibrations
071004 Updated VIS cross-talk files from 2008 and 2010 calibrations
080528 Updated IR bad pixel maps from 06/2008 calibration
080623 Updated IR linearization coefficients
090601 Updated IR bad pixel maps from 06/2009 calibration
090618 Updated IR linearization coefficients
100101 Updated HRIVIS 950-nm filter calibration constant from 2010 calibrations
100101 Updated all MRIVIS filter calibration constants by 5%
100201 Updated VIS flat fields
100201 Updated IR bad pixel maps from 02/2010 calibration
100216 Updated IR linearization coefficients
100801 Updated IR read noise values from 09/2010 calibration
100901 Updated all HRIVIS filter calibration constants (except 950-nm) by 3%
100928 Updated IR bad pixel maps from 11/2010 calibration using 100216 linearity constants

2.2 IR flat-field update

Improved lunar calibration data acquired during EPOXI allowed construction of an IR Spectrometer flat-field calibration file. The new flat field includes the correction for the anti-saturation filter (ASF) in the flat field. In previous calibrations, the flat-field correction ignored

the change across the ASF boundary, and the correction for the ASF was implemented using two absolute calibration curves, one of which was applied outside the ASF and the other inside the ASF.

2.3 IR ALTFF line-dependent integration time

Ongoing analysis of IR timing revealed that read and reset timings differ in the ALTFF mode, and the effective exposure time varies with line number, i.e., along the slit in the spatial direction. This variation has been modeled in the pipeline software, and line-specific integration times are used to convert from DN to DN/s in the ALTFF mode.

2.4 VIS stripes removal

Several techniques were developed by the Science Team to deal with residual observation-dependent, semi-coherent stripes in the VIS background data. Calibration procedures to deal with these effects were added to the pipeline software, though not always run in the pipeline by default. An additional extension that saves the stripe-removal adjustments for each line of each quad was added to retain the reversibility of the calibration. This extension records the de-stripe correction as a 2xN pixel array, where N is the number of rows in the primary image (e.g., 1024 or 512). The extension has units of DN. The first column of this de-stripe extension is the one-dimensional array of corrections subtracted from image columns <=M, where M is the middle column of the primary image (i.e., 512 for a 1024 column image, 1-based indexing). The second column of this de-stripe extension was subtracted from image columns numbered >M. Note that the serial overclock columns that border the primary image array are not modified by the de-stripe procedure. If the de-stripe correction was not applied, then all values in this extension are set to 0.

2.5 Compressed 0-DN fix

Because of an issue in decompressing 0 DN compressed VIS data, the pipeline software was changed to decompress the VIS cameras' zero LUT entry, and only that entry, to the top of its range, i.e., to 350 DN. Also, because it could not be ascertained whether the original pixel's 14-bit DN value was actually 350 or a lower value, all such pixels were marked as SATURATED in the FLAGS FITS extension, the rationale being that just as the actual value of a saturated pixel at the upper end of the possible DN range (255 for compressed 8-bit DNs; 16383 for uncompressed 14-bit DNs) cannot be determined, the same is true for a pixel at the lower end of the range. The user of the FITS file can easily determine the type of saturation of such a flagged pixel by examining its stored value.

2.6 VIS flat-field file re-orientation

Code was modified to change the orientation of pixels in the flat-field calibration file, including POCs and columns and rows just inside OCs, to match that in the flight images. Earlier calibration files interleaved pixels by quadrant and shifted the overclocked pixels to the end of the rows and columns of each quadrant, which is how calibration images were generated from pre-launch ground calibrations. A keyword in the FITS header designates the pixel orientation.

2.7 Header fixes and elimination of I/F files

The pipeline software changes described above generated changes in the calibrated data FITS headers, such as: whether or not a calibration step was applied; calibration files used; constants used; other parameters; and previous observation history. Where appropriate, new calibration steps are included in the SNR extension.

Also, since the VIS I/F calibrated data differ from the radiance data by only a constant factor, the pipeline no longer generates I/F files. It instead generates only the radiance data file and includes the conversion factor from radiance to I/F in the FITS header. For the IR spectrometer, the I/F version was judged not useful due to the presence of thermal emission from the source.

3. Limitations of the Pipeline

The pipeline as of June 2011 used to calibrate version 1.0 EPOXI Hartley 2 data currently in the PDS has several limitations that must be taken into account when analyzing the data.

3.1 HRI-IR Instrument

Correction of the non-linear response function of HRI-IR is based on quadrant average response functions from a linearity dark frame data collection in February 2010. A given pixel might have a somewhat different non-linearity than the quadrant levels. We have investigated pixel to pixel responses and know that the non-linear response function is more accurately reproduced with a pixel to pixel derivation (improvement by a factor of 10 which will be done in future deliveries), however, the current quadrant response function is still reliable and repeatable and results in errors of 1-2% in signal level across the scene. We also have seen some variation in the quadrant averaged non-linearity functions over the span of the ground and in-flight calibrations, although changes during the in-flight cruise and encounter periods appear to have been minimal. For this version of the calibration, the data from February 2010 are well behaved across all modes and were acquired at a nominal operating temperature of the instrument. We also see effects of using subframe modes on the response nonlinearity. While these effects have been taken into account to first order in our calibrations, not all of the second-order effects may be understood or characterized.

The dark frame background level is strongly dependent on the exposure time, temperature of the instrument (bench, electronics, and detector), mode, and the recent history of detector resets and readouts. A careful analysis of the dark level was performed to correct for these effects with an extensive dark frame set and instrument temperature models, but the result is still only good to within ±15 DN in the best case. Master dark frames have been created for all modes from unbinned full frame data. The dark frame scaling has been optimized for closest approach data to Hartley 2 where the signal is high and fills the field of view. Be aware that dark levels are known to be incorrectly too high or too low, especially in coma data in October and post-encounter data, resulting in dark subtracted signal levels that are artificially too low (negative values) or too high (erroneous positive signal), respectively, across the entire spectral range. Efforts are ongoing to improve the dark level correction and to create mode specific master dark frames. While the residual dark level error is not normally an issue in the 2.0-4.5 μm wavelength range where there is a high sensitivity of the detector, it becomes a real issue at short

(<2.0 μm) and long (>4.5 μm) wavelengths where the sensitivity of the detector drops rapidly. As a consequence, one should be extremely cautious when interpreting any data below 2.0 μm or above 4.5 μm.

Flat-field correction of HRI-IR data is quite reliable and reproduces the transmission function of the anti-saturation filter quite well. However, the flat field will be improved with a pixel-to-pixel non-linearity response function.

Our bad pixel map criteria allow pixels to pass that have residuals to a linear fit of up to 2% rms and up to 15% maximum. We have observed some variation in the locations of bad pixels during flight, both during Deep Impact and EPOXI. For this delivery, we used bad pixel maps based on dark frames acquired in November 2010 and linearized with the quadrant response functions derived from the February 2010 data. There may be some uncharacterized bad pixels that have changed throughout the duration of the Hartley 2 data collection. No bad-pixel reclamation has been performed on pipeline calibrated data to date.

The absolute and relative spectral sensitivity calibrations are limited to the ~10% level. This uncertainty is even more pronounced below 1.5 μm, where there is higher spectral resolution and therefore any residuals from poor wavelength matching is more pronounced, and above 4.5 μm where reference stellar spectra have low signal and the absolute calibration is poor. In addition, the effect of the anti-saturation filter introduces uncertainties in the radiometric calibration. This filter reduces the signal at longer wavelengths so that the thermal contribution of the nucleus does not saturate the detector. As a result, the sensitivity of the detector above 4.3 μm in the anti-saturation zone drops very rapidly to zero at 4.6 μm, and one should be very cautious when interpreting data in this region.

The HRI-IR exhibits scattered light effects that are not corrected in the calibration pipeline. In previous mission phases, response at the 1-2% level was seen 10 pixels off the bright limb of the Moon. And a ghost image at the 3-4% level is seen about 35 slit widths away from the primary image in the cross-slit direction.

Detection and correction of cosmic ray signatures in HRI-IR frames is not yet reliable. Such signatures are best detected by differencing pairs of successive frames. But no pipeline process is currently being applied to detect or correct cosmic rays in individual frames.

3.2 HRI-VIS and MRI-VIS Instruments

The small amount of geometric distortion in the VIS images is not normally being corrected. Radial distortion is ≤0.1 pixel in the MRI and likely less for the HRI (but it has not been measured for HRI). The two central rows of the CCDs in each camera are 1/6 of a pixel smaller vertically than a normal row. Therefore, reconstructed images, which have uniform row spacing, have a 1/3-pixel extension introduced at the center of the array. Thus for two features on either side of the midpoint line, the vertical component of the *actual* angular separation between those features is one-third of a pixel less than their measured difference in vertical pixels in the image. As for all geometric distortions, correction of this distortion will require resampling of the image and an attendant loss in spatial resolution. The inflight focal length measurements result in pixel IFOVs that are slightly different than the canonical 10 μrad for MRI and 2 μrad for HRI. The

actual MRI IFOV is about 0.08% smaller, and the actual HRI IFOV varies from 0.002% larger to 0.02% smaller, depending on the filter. The relative boresight alignments between the instruments are uncertain to about ±1 MRI pixel.

The spatial resolution of the MRI degrades slightly from the center to the edges of the FOV with the PSF width growing by about 50% at the corners of the FOV. The HRIVIS is out of focus with a PSF FWHM of ~9 pixels; the detailed PSF varies with filter.

Several limitations on the accuracy of the radiometric calibrations exist. The calibration pipeline assumes that the VIS CCD response is perfectly linear. In reality, departures from linearity exist, with the actual response being up to 1% high relative to perfect linearity in the signal range between 100 and 12000 DN and up to 2% low relative to perfect linearity below 100 DN and above 12000 DN. The bias correction in Modes 7 and 8 cannot be derived from serial overclocked (SOC) pixels, since SOCs do not exist in these modes. The fixed bias estimate used in these modes can be off by ±5 DN. The residual horizontal striping that remains after quad-level bias subtraction can range from 1 – 10 DN if left uncorrected. For those cases where the stripe-removal algorithm can be applied, the residual striping is reduced to ~0.2 DN. Frame transfer smear correction is accurate to only about 1 DN.

The uncertainty in conversion to absolute radiometric units is estimated to be 5% for HRI except for the 950-nm filter, where the uncertainty is ~10%. For MRI, the uncertainties are about 10% for all but the UV filters, which have an uncertainty of 20%. The actual central wavelength of the MRI 387-nm filter remains uncertain; it could be up to 2 nm longward of its nominal value. Several filters have red leaks that may produce non-negligible signal – they include the 350, 550, 650, and 850-nm HRI filters and the 309-nm MRI filter.

The quality and uncertainty in the radiometric calibration for each pixel is recorded in the quality and SNR extensions. Analyses that involve photometric precision should refer to those extensions to determine the error bars on their measurements. In addition, flat-field calibrations leave uncorrected pixel-to-pixel variations of ~0.5%. The 1/6-pixel narrower central rows collect only 5/6 of the charge of a normal row. This effect is corrected by the flat-field division so that the pixels in these rows have the correct scene radiance assigned to them. However, point-source or disk-integrated photometric measurements using aperture photometry areas that include these central rows will be slightly distorted unless special adjustments are made (e.g., an extra 1/6-pixel worth of signal is added to the pixels in each of these two rows in the reconstructed images, as described in Appendix A of Belton, et. al., Icarus, 213, 345-368, 2011). The ADC quantization error is ~1 DN (this effect is included in the SNR map). Data compression, if used, can result in quantization errors that dominate the pixel-level noise (this effect is also included in the SNR map). Cross talk between CCD quadrants is corrected but may leave residual uncorrected signal of up to 2 DN from a saturated scene area within a quadrant. Internal scattered light can produce measurable diffuse scatter within the FOV. This scatter is normally small enough to be ignored, but should not be neglected when doing high-contrast photometric measurements. The scatter is about 6x worse in the MRI than in the HRI, is 5x worse in the IR filters than in typical filters for both cameras, and is about 3x worse in the HRI violet filter than in the other HRI filters. In addition, a 6-pronged diffraction spike ray pattern is produced at a level of up to 2×10^{-8}/source pixel. Finally, cosmic ray signatures are not reliably removed.