Announcements

- You should be finishing Lab 2...
- Lab 2 is now due on Thursday Oct. 1 (as is Lab 3)
- Follow instructions on turning in assigned materials (workfile file, Report.tar.gz, datafarm...)
- This week we will be starting LAB 3
- Paul is available via appointment to discuss Lab 2.

Reduction/Calibration I

- A bit more about flatfields...
- Working with calibration data
  - Image histograms
  - Calculating gain and read noise
  - Calculating SNR
  - Alternate approach: photon transfer curve
Flatfield Paranoia

• There are many potential sources of pixel to pixel nonuniformity variations in CCD images
  – Intrinsic variations in pixel responsivity
  – Gradients induced by optics
  – Gradients, imperfections in filters
• These can be wavelength dependent
• These can change with time

Acquiring Flatfields

• At the telescope:
  – Can look at the inside of the dome
  – Can look at the twilight (or daytime) sky
• In the lab
  – Look into an integrating sphere
• During a space flight mission
  – Onboard "flat" reflecting targets
  – Smeared-out images of the object of interest
  – Martian sky

Wavelength Dependence

(Howell’s "color terms")

• Pixel responsivity will vary depending on the wavelength because of variations in the depth of photon absorption
• Therefore, flatfields should always be taken using light that covers the same wavelengths as what you expect to measure
Example

Working With Calibration Data

- Image histograms
- Calculation of Gain
- Calculation of Read Noise
- Calculation of Signal to Noise Ratio (SNR)
- "Photon Transfer" Curve

Image Histograms

- **LOOK** at every image before you work with it!
- But it is not enough to just look at the image; you must also understand the **distribution** of values in that image
- Most common approach is to plot a histogram of DN vs. the number of occurrences of that DN
- Provides insight into the noise level and hints on how best to stretch the contrast!
Calculating Gain and Read Noise

- Using bias (B) and flatfield (F) images
- Use the standard deviation (σ) of the difference between "identical" images with (F) and without (B) illumination to derive gain and read noise (Howell, p. 53):

\[
\text{Gain} = \frac{[(F_1 - B_1) + (F_2 - B_2)]}{(\sigma_{F_1 - F_2})^2}
\]

\[
\text{Read Noise} = \text{Gain} \times \frac{(\sigma_{B_1 - B_2})}{\sqrt{2}}
\]

Calculating SNR

- SNR = Signal to Noise Ratio
- Quick metric of what is possible with an instrument (though will differ from what is practical or achievable)
- "Signal" is the number of photons detected
- "Noise" is a combination of Poisson (continuous counting) and shot noise (unique event) components

\[
\text{SNR} = \frac{N_p}{\sqrt{N_p + N_s + N_d + N_r}}
\]

Where
- \(N_p\) = Number of object photons
- \(N_s\) = Number of sky/background photons
- \(N_d\) = Number of dark current electrons
- \(N_r\) = Read noise, in electrons (shot noise)

Another Approach

- "Photon Transfer" method
- aka "Light Transfer" or "Electron Transfer"
- Uses measurements of signal vs. noise to determine gain, read noise, full well capacity, and other CCD parameters
Signal (DN) = $P_I \cdot QE \cdot S_V \cdot A_1 \cdot A_2$

where $P_I$ = Incident photon flux
QE = QE $\cdot \eta_E$ = average quantum efficiency
$S_V$ = Sensitivity of CCD on-chip circuitry
$A_1$ = electronic gain of the amplifier circuitry
$A_2$ = transfer function of the ADC

note where read noise ($\sigma_R^2$) enters the chain...

$$S (DN) = P_I \cdot QE \cdot S_V \cdot A_1 \cdot A_2$$

or
$$S = P_M \cdot K^{-1}$$

where $P_M$ = $P_I \cdot QE$ = number of measured (interacting) photons
$K = (S_V \cdot A_1 \cdot A_2)^{-1}$ = Gain in electrons/DN

We can determine K by relating S to its variance ($\sigma_S^2$)

Using standard error propagation (Bevington):

$$\sigma_S^2 = [\frac{\partial S}{\partial P_M}]^2 \sigma_{P_M}^2 + [\frac{\partial S}{\partial K}]^2 \sigma_K^2 + \sigma_R^2$$

Doing the differentiation and assuming that $\sigma_K^2 = 0$ yields:

$$\sigma_S^2 = (\sigma_{P_M}^2 / K^2) + \sigma_R^2$$

But since from Poisson statistics $\sigma_{P_M}^2 = P_M$ then

$$\sigma_S^2 = (P_M / K^2) + \sigma_R^2$$

$$\sigma_S^2 = ((S \cdot K) / K^2) + \sigma_R^2$$

$$\sigma_S^2 = (S / K) + \sigma_R^2$$

and thus $K = S \cdot (\sigma_S^2 - \sigma_R^2)^{-1}$

Note: if $\sigma_S^2 = 0$ then $K = \sqrt{S / K}$, or

$$\log(\sigma_S) = 0.5 \log(S / K),$$

which is a line of slope 0.5 on a log-log plot.

On such a plot the point where $\sigma_S = 1$ is where $K = S$.

Thus there is a graphical solution to finding $K$. 


Example: MER/Pancam LTF Plot

Reduction/Calibration

- Building a data reduction pipeline
- Typical steps
  - Bad pixel replacement
  - Bias subtraction
  - Dark current modeling, subtraction
  - "Electronic shutter effect" subtraction
  - Flat field modeling, removal
  - Background modeling, removal
  - Responsivity modeling, scaling

Pipeline Processing

- Typically in planetary science, we deal with hundreds to thousands of short-exposure images obtained during telescopic observing runs or spacecraft missions
- We usually don’t have the time or the luxury to process images one-by-one, so we are driven to come up with automated methods of processing large volumes of data
- However, we must avoid black box data processing at all costs!

Understand Thy Pipeline!
Pipeline Processing

- The main goal of the data reduction and calibration exercise is to remove instrumental effects from the raw data numbers (DN) so that
  (a) the corrected DN values (DN/sec) are linearly increasing with increasing input radiance (e.g., W/m²/µm/sr), and
  (b) the corrected DN value equals zero DN/sec for an input radiance of zero.

In general [Reid et al., 1999]*, for each pixel $i$,

$$C_i = (I_i \cdot [R(T) \cdot t] \cdot F_i) + [D_i(T) \cdot t] + S_i(I_i) + B_i$$

where

- $C_i$ = the raw DN measured in that pixel
- $I_i$ = the actual radiance of the scene (W/m²/µm/sr)
- $R(T)$ = the radiometric responsivity of that pixel (DN/sec/W/m²/µm/sr) at a temperature $T$ in Kelvin
- $F_i$ = the normalized flatfield responsivity of that pixel
- $t$ = the exposure time (sec)
- $D_i(T)$ = dark current generated during exposure & frame readout (DN/sec)
- $S_i(I_i)$ = any extra scene-dependent signal (e.g., frame transfer smear)
- $B_i$ = the bias or offset value of that pixel (DN)


To calibrate images to "corrected DN" requires that at least four images are acquired:

1. An image of the scene with exposure time $t$
   $$C_{\text{image},t} = I_i \cdot [R(T) \cdot t] \cdot F + D(T) \cdot t + S(I) + B$$

2. An image of the scene with exposure time zero
   $$C_{\text{image},0} = S(I) + B$$

3. A dark image with exposure time $t$
   $$C_{\text{dark},t} = D(T) \cdot t + S(0) + B$$

4. A dark image with exposure time zero
   $$C_{\text{dark},0} = S(0) + B$$
The four images acquired can then be combined and “flatfielded” so that the corrected DN/sec, $C_{corr}$ is defined as:

$$C_{corr} = \left[ (C_{image,t} - C_{image,0}) - (C_{dark,t} - C_{dark,0}) \right] \cdot F^{-1} \cdot t^{-1}$$

$$C_{corr} = \left[ \{R(T) \cdot I \cdot F \cdot t + D(T) \cdot t + S(I) + B \} - \{D(T) \cdot t + S(0) + B - (S(I) + B) \} \right] \cdot F^{-1} \cdot t^{-1}$$

$$C_{corr} = R(T) \cdot I$$

And therefore, obtaining those four images, plus a flatfield image, plus an image of a scene of known radiance, I, allows the responsivity of the imaging system to be determined

$$R(T) = C_{corr} / I$$

---

**Example: MER/Pancam Pipeline**


---

**Bad Pixel Replacement**

- What does it mean to be “bad”?
  - Usually defined as a value that is some number of standard deviations above or below its neighbors
  - Could also be defined in terms of high dark current or anomalous responsivity
- Common to replace with median of neighbors

<table>
<thead>
<tr>
<th>Value 1</th>
<th>Value 2</th>
<th>Value 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>48</td>
<td>71</td>
</tr>
<tr>
<td>60</td>
<td>3500</td>
<td>68</td>
</tr>
<tr>
<td>55</td>
<td>73</td>
<td>65</td>
</tr>
</tbody>
</table>

- 3500 could be a "bad" value (how? why?)
- Median of 8 neighbors is 65 (why?)
- Why not use the mean instead of median?
Bad Pixel Replacement

- What does it mean to be "bad"?
  - Usually defined as a value that is some number of standard deviations above or below its neighbors.
  - Could also be defined in terms of high dark current or anomalous responsivity.
- Common to replace with median of neighbors.

<table>
<thead>
<tr>
<th></th>
<th>65</th>
<th>48</th>
<th>71</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>3500</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>910</td>
<td>65</td>
<td></td>
</tr>
</tbody>
</table>

- 3500 could be a "bad" value (how? why?)
- Median of 8 neighbors is 65 (why?)
- Why not use the mean instead of median?
  - example: changing lower center to "910"

Bias Subtraction

- Bias is a small voltage offset put there intentionally to bias the device against returning zero values if the responsivity drifts due to temperature, radiation effects, electronics aging, etc. (Lecture 4)
- Bias should be ~ constant from pixel to pixel, but will almost certainly vary with temperature.
- Bias subtraction sometimes occurs "automatically" when other components that contain bias are subtracted from each other.

Dark Current Modeling, Subtraction

- Dark current comes from thermal vibrations freeing electrons from the silicon substrate.
- Strongly temperature dependent
  - \( DN_{\text{dark}} \propto e^{\alpha T} \) (thousands of e/\text{sec} typical at room T)
- Specific temperature dependence needs to be determined by taking many "dark images" over a range of temperatures.
  - Must verify that CCD is completely unilluminated.
- Darks with non-zero exposure time yield dark current that builds up in the imaging area of a CCD during exposure plus dark current that builds up in the frame transfer or interline transfer area during readout.
- Darks with zero exposure time yield dark current that only builds up in the frame/interline transfer area during readout.
Example of Dark Current Data

Bias-subtracted DN

"Electronic Shutter Effect" Subtraction

- aka "Frame Transfer Smear"

Recall (Lecture 4):
- Frame Transfer: Device is split into two equal parts, the active area which detects the signal, and the masked off storage area, where the charge is stored while being clocked out.
- Interline Transfer: Active and masked off storage areas are adjacent and aligned.

"Electronic Shutter" or Frame Transfer Smear

- Shift of the charge from active to storage areas is fast compared to the readout, but still takes a finite amount of time (typically 5-10 msec)
- During this time, for CCDs without a mechanical shutter:
  - Photons keep getting detected in rows before charge is clocked out
  - Dark current keeps building up in rows before charge is clocked out
- The charge being built up during frame transfer is "smeared" down the array and added to the storage area image
- The effect is linearly proportional to the number of rows away from the storage area (or output register for ILT CCDs)
- The effect is extremely scene dependent!
- While analytic solutions exist for correcting frame transfer smear, the simplest correction approach is to take a subsequent image of the same scene with a zero exposure time, and subtract that image from the non-zero exposure scene image
Frame Transfer Smear Example

Mars Pathfinder IMP Dark

Row Number

DN

Frame Transfer Smear Example

MER/Pancam Grid Target: 40 msec exposure

MER/Pancam Grid Target: 0 msec exposure

Zoom of 40 msec image minus 0 msec image, 0 to 500 DN stretch
Flatfield Modeling, Removal

- Flatfield acquisition discussed in Lecture 5...
- Many flats are usually averaged together and divided by the average (or median) value in a nice, clean center region to derive “normalized” flats with mean = 1.0
- Dividing out a normalized flat then corrects for pixel-to-pixel variations without changing the average DN level of an image

Background Modeling, Removal

- Sometimes an additional subtraction might be needed if the source or object of interest is superimposed on or in a nonzero radiance background
  - e.g., faint asteroids amid background galaxies, planets in the infrared against the warm night sky, rocks on Mars illuminated by sky radiance, ...
- And sometimes that background is very nonuniform or might have a strong color dependence
  - e.g., strong scattered light gradient around faint inner satellites of Jupiter & Saturn, Mars sky, ...

Responsivity Modeling, Scaling

- Responsivity determined using observations of calibrated radiance sources
  - Integrating spheres and calibrated lamps
  - Calibrated reflectance standards
- Always verified at the telescope or in flight!
  - Observations of standard calibrator stars
  - Measurements of onboard calibration targets to reflect sunlight or radiate at known temperatures
- Scaling applied at end to convert DN to radiance
Uncertainties

- Temperature variations
- Viewing angle variations
- Compression artifacts
- Groups of bad pixels
- Changes in instruments, environments
  - Launch/landing shocks
  - Radiation damage
  - Mars dust, comet dust, thruster propellant, ...

Summary/Main Points (1 of 2)

- Flatfields are difficult to acquire but at least must be obtained over the range of wavelengths expected for the source
- Always view an image and its histogram to fully evaluate image content, quality, and noise
- Several methods exist to estimate/calculate CCD gain, read noise, SNR, and full well
  - Require multiple bias images to assess read noise component
  - Require multiple flatfield images to assess photon noise component

Summary/Main Points (2 of 2)

- It is usually necessary to build some kind of software data reduction “pipeline” to process planetary images
- Bias, Flatfield, Dark Current, Bad Pixels, Frame Transfer Smear, Background Removal and Responsivity Scaling are all typical components of a pipeline
- No Black Boxes!
- Keep Track of Your Uncertainties!
Thoughts on the Final Project

• Goal is for you to get experience working with a data set or with data processing algorithms that interest you
• For example, your project could be
  a) A worked example of image processing techniques that we have studied this semester applied to a data set that you have identified, or
  b) A report that describes and shows examples of a real application of any of the image processing techniques that we have studied this semester
  c) A fully realized scientific analysis project that uses the skills you have learned this semester (preferred).
• Prof. Hayes or Prof. Lloyd can help you find data, references, and software tools, especially if you start early and plan in advance. NOTE: Topic Email (1 Para.) Due Oct. 15... (or earlier)

Final Project Proposal

• A written report, uploaded to datafarm, 2-3 pages including any figures, tables, programs, or references
• In the report, you must tell me:
  – In general terms, what do you want to work on and why?
  – Specifically, what data/algorithms will you work with, and how?
    • Describe/show the data set, format, calibration/reduction status, programming plans/status, relevant references, etc.
    • Provide a schedule for your work
• Optional draft due 5:00 p.m. 10/15; Final due 5:00 p.m. 10/22
• We will grade these (20 points) and return before 10/29
  – Unsatisfactory or incomplete proposals will be docked 5 points, and returned for re-submittal by 11/3
  – Repeat as needed...