Charge Coupled Devices

C. A. Griffith, Class Notes, PTYS 521, 2016

Not for distribution.
1. Introduction

While telescopes are able to gather more light from a distance source than does the naked eye, its main advantage is its capability to integrate the incident photons, using long time exposures. In contrast, the human eye responds only to the rate at which photons reach the retina.

The type of detector needed depends on the wavelength region of interest. At optical wavelengths from 0.3 $\mu$m to $\sim$1 $\mu$m, light is energetic enough liberate electrons from their surfaces through the photoelectric effect. Detectors such as the charged-coupled device or CCD make use of this effect (Fig. 1). At longer wavelengths, in the infrared, detectors work on either of two processes. Thermal detectors capitalize on the rise in temperature of a material that occurs with the absorption of a photon. Photon or quantum detectors are based on the response of semi-conductors to the absorption of an IR photon. Detectors at other wavelengths will be discussed later. The sensitivity of detectors is often expressed as the quantum efficiency, which is the ratio of photons detected to those incident on the detector. Evidently, an efficiency of 100% is a perfect detection.

Here I’ll give a short overview of CCD, with essentially no details of the engineering. This write-up instead conveys the basic ideas behind the functioning of a CCD, and includes information needed for standard optical photometry and spectroscopy of astronomical objects. The figure at the front page of these note shows the Hyper-SuprimeCam, which sits on the 8.3 m telescope on Mauna Kea. It has 116 CCD chips, which covers a field of view of 1.5 degrees, with 870 million pixels. The HSC weighs 3 tons and is 9 feet high.

2. The CCD

Charge Coupled Devices (CCD) have a quantum efficiency that peaks at 90% and extends, for quantum efficiencies of over 10%, from $\sim$0.3 to 1.05 $\mu$m (Fig. 2). The range can be extended with, for example, various coatings and configurations (e.g. thin back-illuminated CCDs). This sensitivity is extraordinary, when compared to, for example, the few percent efficiency typical of photographic emulsion.

A CCD consists of basically two different kinds of silicon substrates. The top n-type silicon substrate contains phosphorus. This impurity creates fixed positive charges throughout this thin $\sim$2 $\mu$m layer. The bottom layer, p-type silicon substrate, is thicker ($\sim$260 $\mu$m) and doped with boron, which creates fixed negative fixed charges.

Before a CCD is exposed to radiation, the n-type substrate is purged of electrons. Once the negative free charges are flushed out of the n-type silicon, the positive holes, repelled by the overlying positive layer move to the bottom of the p-type substrate (the undepleted region). Taken from Bradt 2004.
lack of free charges prevents conduction. In the p-depleted layer an electric field in the downward direction (motion of a positive test charge) occurs as a result of the high positive charge in the n-type region, which has a greater concentration of fixed charges than does the p-type region. This sets the stage for capturing the photo-electrons, which are forced upward into the n-substrate, whose positive charge traps them.

3. Detection of photons

Photons are measured through their production of photoelectrons, such that for standard CCDs, one electron is produced for one detected photon. Silicon has a band gap energy of 1.14 eV. In a CCD, the silicon absorbs photons of energies of 1.1 eV to ~1.4 eV, thereby bumping up valence electrons into the conduction band. Such photons have wavelengths of 0.3-1.1 µm, since the photon’s energy is \( E(eV) = \frac{hc}{\lambda} = 1.2407/\lambda(\mu m) \).

The wavelength sensitivity, or the quantum efficiency of CCDs can be almost deduced from the path length of light as a function of wavelength in silicon (Fig. 3). The QE depends further on the width of the CCD. Photons greater than 0.35 µm and less than 0.8 µm are are absorbed even by the backside illuminated CCDs, which are only ~15 µm thick, as well as by the heftier front-side illuminated CCDs of ~300 µm width (Fig. 2). At wavelengths shorter than ~0.35 µm, however, 70% of the light is reflected off. An anti-reflection coating can partly remedy this and extend the sensitivity to 0.3 µm. At wavelengths longer than ~1 µm, depending on the CCD’s width, silicon is basically transparent.

4. Storage of photons

The electrons released into the conduction band do not recombine with the silicon, which they would do in a fraction of a millisecond, because they are quickly forced into the n-type layer as a result of the electric field. There they are trapped as a result of the potential well. At the end of the observation, the number of stored electrons depends on the integrated intensity of light. The amount of charge that a pixel can hold after normal integration is called its well capacity. One must be sure to limit integration times so as not to try to detect more photons than the detector has room for, i.e. so as not to exceed the well capacity of the detector. Such a goof-up will yield observations with flat peak fluxes.

Pixels are created by imposing a horizontal differ-
ential voltage across the n-type substrate. Each pixel generally has 3 electrodes of different positive potentials, so that the electrons gather at the spots within the array (and within the pixel) where the potential is highest. In Fig. 1, electrode “1” induces the highest potential, generally $\sim +5\,\text{V}$, and the photons gather at this place. CCDs consist of an array of $\sim 15 \,\mu\text{m}$ wide pixels, numbering up to 4096 x 4096 ($2^{12} \times 2^{12}$) elements.

5. Reading out CCDs

The electrodes or gates, generally 3, have voltage potentials that can be regulated. To read out the CCD, the potentials are sequenced simultaneously, so as to move the charges, generally along a column, which is connected in parallel, towards an “output register”, which is simply an unexposed row of pixels at the end of the exposed chip (Fig. 3). This shifting of electrons to the output registrar is achieved by altering the voltages in the 3 gates, simultaneously in all pixels. Effectively the voltage of electrode “1” is dropped to $\sim 2.5\,\text{V}$, while ”2” is raised to a higher voltage. In the next step “1” is dropped to zero, and 2 raised to 5V so that the electrons migrate to electrode “2”. In the next two steps, the electrons migrate to electrode “3”. In the final 2 steps the electrons move to electrode “1” in the adjacent pixel, which has been vacated by the same process, conducted simultaneously. Therefore, in 6 voltage adjustment steps, the electrons migrate to the adjacent pixel.

Once an entire row is shifted into this output register row, the charges on the “output register” row pixels are shifted, similarly using variable voltages, towards the output electronics. Here the voltage (analog) is amplified and converted to an output digital number.

Once the entire row is read out, the gate voltages are altered so that the charges, again, all shift towards the output register row, and the next row is readout. This sequence of shifting and reading out the rows is continued until all the rows are read out.

The CCD is then flushed of electrons by reading out the entire array a couple more times. The next observation can then begin.

5.1. Analog to digital conversion

The current is amplified with a field effect transistor (FET) and converted from an analog voltage to a digital number using something called, not surprisingly, an analog-to-digital converter (ADC). The data is now stored in digital ADU units, such that the number of electrons needed to produce 1 ADU is called the gain of the CCD. A typical gain is 1-10 electrons/ADU.

Now the ADU is digitized as an integer, and this attribute affects the precision of your data. Suppose that you are working with a CCD with a gain of 100 electrons/ADU, and your CCD detects 144 photons, which are stored as 144 electrons. Then after the data is converted to digital ADU units, you will have a result of 1 ADU. In effect, the gain affects the precision of your final result, which here is 100 photons or equivalently electrons. Why not always use a very low gain? That’s because the gain is an integer specified by a number of digital bits. For example a $2^{15} = 32,768$ bit ADC can only have ADU values from 0 to 32,767. To make sure that these span the full range
of electron counts possible to a CCD, you might build your electronics so that the gain equals the capacity of the full well depth divided by the largest number of possible ADU values. For example if the well capacity of your pixels is 150,000 electrons, then the gain could be set to 150,000 e\(^{-}\) / 32,767 ADU, that is \(\sim 4.5 \text{ e}^{-}/\text{ADU}\) (Fig 4). Of course, if you have very few digital bits available, e.g. as on a spacecraft, then you might end up with high gain and low precision, if you follow that protocol. You could alternatively set the upper limit at the point with the detector becomes nonlinear.

5.2. The integration time of an observation

Note that there are only so many photons that you can detect with a given CCD before reading them out. In fact, three factors control the upper limit: 1) the well capacity, 2) the maximum possible number in ADU units, and 3) the electron capacity above which the detector becomes non-linear. That is, one doesn’t want to stuff more electrons in the detector than can fit, nor record more than can be counted or than can be related linearly to the number of photons. The third limitation is the most dangerous, because there is no sign of a problem, and so one could gleefully misinterpret the data and come up with a false scientific result.

5.3. Readout Noise

Although faster readout times are possible, the readout time is generally kept to 50 \(\mu\text{s/pixel}\), so as to not heat the amplifier (with the higher current) and thus keep read noise to a minimum, which is usually around 10 electrons/pixel/readout. Read noise occurs in part during the analog-to-digital conversion, and also partly results from spurious electrons that are introduced during the entire process. One way to limit the read out noise is to bin the pixels before readout. Then if four pixels are binned, the signal is affected by only one read out noise, rather than four times the value. The output register pixels normally hold 5 to 10 times the normal pixels so that binning is possible even for high signals.

5.4. Images or Spectroscopy

Both photometric images and slit spectroscopy can be achieved with a CCD. The latter are produced by simply placing the detector at one of the telescope’s focal planes, and selecting a wavelength range of interest with the use of filters. To produce a spectrum, additional optical elements are added after the focal plane, often along with a slit, e.g. in the y-axis direction, such that light is dispersed along the x-axis. In this case, the CCD records spatial information, which runs in the y-axis along the CCD, and spectral information, which runs along the x-axis.

6. Data Reduction of CCD Images

A raw CCD image can look like a bit of a mess. The initially sad appearance of the data has a range of origins, for example from 1) the presence of bad pixels, 2) spuriously excited pixels by cosmic rays, 3) thermally excited electrons, 4) dust on the filters, and 5) a generally lack of uniformity in the sensitivity of the array. These problems are treated as follows.

6.1. Hot Pixels

CCD’s have a number of hot or bad pixels that register erroneous signals. In addition, cosmic rays or charged particles can trigger a high signal in one or two pixels. In both cases, the signal does not obey the point spread function (PSF) and thus can be easily identified. One way to deal with these anomalies is use some software and replace the pixels with a local average. A better way is to record a number of frames, if the readout noise is not a major factor, and take the median of the images.

6.2. Bias Offset

When CCD images are converted from analogue to digital, an offset, or bias signal is purposefully introduced to prevent the signal from being negative at any point. The signal is kept positive so as to use all the bits (in the A/D converter) on the signal rather than the minus sign. The value of the bias varies with position across the CCD and with the read-out time. The bias can be measured with with overscan strips or the measurement of a bias frame. Overscan strips are added columns to the CCD, which are recorded and read-out with no signal, after the entire CCD image is read-out. They are averaged for each row, thereby giving a measure of the zero signal point, and subtracted from each image pixel in that row. Another, and better, way to eliminate the bias is to record bias frames, i.e. images taken at the beginning and end of the night, with no integration time and a closed shutter. These images thus represent the zero point in the
absence of photoelectrons. A master bias frame is produced by taking a median of all bias frames. This is subtracted from all images recorded during the night to normalize the data.

6.3. Dark Current

The dark current arises from the thermally excited electrons, particularly in IR observations, and changes with pixel and with time. The amount of dark current changes with pixel and with time, although it particularly depends on the operating temperature of the detector.

To regulate the dark current, CCDs are cooled to -100°C, with little (±0.1°C) variation, using liquid nitrogen. In this case, the dark current of the CCD may be low enough to ignore. It is best to evaluate this on an instrument to instrument basis. If the dark current is significant then the data can be dark-subtracted. Here at least 10 dark current frames are recorded in the beginning or the end of the night, with integration times that preferably match those of the observations, and with a closed shutter, of course. A master dark current frame is obtained by taking the median of the individuals and (if the integration time differs from that of the observations) dividing this by the integration time (to derive the dark current per second). This master dark current is then scaled to the exposure time of each observation and subtracted from the observation.

6.4. Flattening the Array Sensitivity

The pixels that constitute detector arrays, both for CCDs and IR quantum detectors, do not all have the same sensitivity due to manufacturing problems and the obstruction of light at the edge, vignetting, of the telescope field of view (Fig. 5). In addition there can be dust on the filters and dewar window; these add little donuts to the image (Fig. 6). It is therefore necessary to normalize the pixels. This is done by recording a uniformly illuminated light source, called a flat field, at each grating and filter setting of the observations. A light source can be shined on a screen in the dome and observed out of focus; these are called dome flats and are recorded at the beginning or end of the night, with a long enough integration time to quantify the sensitivity. Here the integration time is easy to determine as it will be based on the intensity of the lamp. However, the problem with this technique is that light enters the telescope differently than it does in the observations. Another way to record flat fields is to observe the twilight sky, however with sky flats you may need to play around with the exposure time. A master flat is then calculated by taking the median of all those that pertain to a grating and filter setting, to eliminate cosmic rays. One could say that the flats are combined or median stacked. It is then divided by the mean value of all the pixels so that it is normalized to unity. The data is divided by the flat that corresponds to the same grating and filter setting.

A problem with flat fielding a detector arises because the sensitivity of the detector depends on the wavelength. So ideally you would like to the same wavelength distribution over the bandpass in use. One could obtain night-time sky flats at different spatial offsets, and median these to remove the stars and planets to make a master flat. However this would take a lot of time.

An attribute that looks like fringes and are called such can occur in a CCD image, as a result of the interference of light being scattered within the CCD. These can be caused by strong emission lines of the terrestrial sky, particularly those due to OH that are powered by sunlight during the day, but which decay slowly, thereby making their appearance in the night sky. These emissions can be highly variable, even in a given night.
Fig. 6.— Flat field indicates the effects of dust.

6.5. Summary

In summary, here are a list of the frames that need to be taken to produce a fully reduced set of CCD observations.

A reduction of CCD data would be:

\[
\text{Final Image} = \frac{\text{Raw Data} - \text{Bias Frame}}{\text{Flat Field} - \text{Bias Frame}}
\]

if the CCD is cooled with liquid N\(_2\) and thus has insignificant dark current. If the CCD is not cooled, or if we are working with IR data and IR detectors, then the dark current must be subtracted from the raw data and the flat field.
### Table 1
Photometry CCD Measurements

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Integration</th>
<th>Filter/Grating</th>
<th>Shutter</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias Frames</td>
<td>0</td>
<td></td>
<td>Closed</td>
<td>10 (median)</td>
</tr>
<tr>
<td>Dark Frames</td>
<td>Match Obs</td>
<td></td>
<td>Closed</td>
<td>5 (median)</td>
</tr>
<tr>
<td>Flat Field</td>
<td>Set</td>
<td>Match Obs</td>
<td>Open</td>
<td>5 (median)</td>
</tr>
<tr>
<td>Object/Standards</td>
<td>Set(^1)</td>
<td>Set</td>
<td>Open</td>
<td>depends(^2)</td>
</tr>
</tbody>
</table>

\(^1\)Don’t exceed linear range, ADU maximum count level, & full well depth.

\(^2\)This depends, among other things, on the brightness of the source. For bright sources one might take a number of frames, and median them to get rid of the cosmic ray hits. For a dim source, one would like to integrate as long as possible, often set by the sky background, to limit the read noise.