CCD and CMOS Sensors

John Kielkopf

September 30, 2021

Abstract

Image data acquired with a charge coupled device camera on an astronomical telescope contain quantitative information on spatial, spectral, and temporal properties of the source that emitted the light. Between that source and the final data, the photons traverse an interstellar medium, and then for ground-based telescopes, also the Earth's atmosphere. They pass through an optical system including intentional selective filtering to the sensor that produces photoelectrons in nearly linear proportion to the incident light. For optical astronomy from the ultraviolet through the infrared, the most common detectors are built with the charge coupled device (CCD) or complementary metal oxide (CMOS) technologies that we review here..

1 Introduction

Historically, we think of an image formed by visible light as what we would see directly with our eyes if they were sensitive enough. Now, with electronics and computers we can capture only a few photons and turn information about them into an image thay conveys far more than our eyes could sense on their own. This is a guide to the acquisition and analysis of optical astronomical images. It is intended primarily to be useful for work with visible and near-infrared light and array detectors such as the charge coupled device (a CCD) and new complementary metal oxide (CMOS) camera chips, but it may also be useful for other work with data acquired in many ways and then stored in an image format. The concern we are addressing is how to detect the light and to take the first steps of processing the data into a measurement of the light's flux and the direction from which it came.

2 Radiometry of astronomical sources

If you suppose that light from a distant star reaches Earth's atmosphere, then we would describe the energy arriving per unit time and area as an irradiance E in units of (W/m^2).

Sometimes this is referred to as "flux", in the sense of a flow of energy through an imaginary surface. Also, astronomers may say "intensity" generically for terms that measure energy flow, though in the science of radiometry this is a misnomer. For more on this topic look at the Glossary in Section 9.

The energy we detect in photons originates with its production in the interior of the star. It has propagated out to the star's surface where it is emitted from the point of last scattering as light over a wide range of frequencies. For typical stars such as the Sun, most of the energy leaving into space is in the form of photons of visible light, but there is something at all frequencies from radio through gamma radiation. The photons travel through interstellar dust, molecules, and atoms that absorb and scatter some of it, and then arrive at Earth. The energy passes through Earth's atmosphere and is collected by the area of a telescope, where it is focused onto a detector. Ultimately, the signal measured is proportional to E with allowance for atmospheric transmission, telescope area and efficiency, and the conversion of photons to an electrical signal s. We will deal progressively with these factors that affect the measurement of the light from stars, with the understanding that they are usually interpreted on the astronomical magnitude scale. For a signal s measured relative to a reference s_0, magnitudes are given by

$$s/s_0 = 10^{-2/5(m-m_0)} \tag{1}$$

The scale is such that a difference of 5 magnitudes is a relative signal of a factor of 100×. The larger the magnitude of a star, the fainter it is. Inversely, magnitudes are logarithmic and approximate the non-linear response of the eye

$$m - m_0 = -2.5 \log_{10}(s/s_0) \tag{2}$$

Typically stellar magnitudes are referenced to a system of standard stars, represented here by s_0, and they are measured in specific spectral bands. That means that the magnitudes and fluxes depend on which wavelengths of light are included, and on the system by which the reference magnitudes are established. We see in Eq. 2 that the numerical value is larger for fainter stars, and it follows that the "0" point should be a bright and presumable non-variable star accessible from many observatories. The zero of the magnitude scale is set very nearly by the northern star Vega in Lyra, one of the three prominent naked-eye stars that make the Summer Triangle. Magnitudes set in this way are referred to as *relative* magnitudes because they are properties as observed, but will be in various filter bands and are usually considered to be corrected for the absorption of Earth's atmosphere and for instrumental effects. However, stars are actually at different distances if we want to find their *absolute* properties, we need to take into account how far they are from us and the effects of interstellar material.

Distances on an astronomical scale are conveniently measured in light years, the distance that light travels in one Earth year. It is a very fundamental unit, since it is given by the speed of light (a defined constant), the second which is defined by the property of a cesium atom, and the orbital period of Earth. A light year is 9.4605284×10^{15} m. Measurements of the distances to the nearest stars are made by parallax, the angular shift of a nearby star relative to very distant stars or galaxies resulting from Earth's motion around the Sun. The semimajor axis of Earth's orbit is 149 597 871 (km), and for simplicity is referred to as an astronomical unit (AU). When the Earth is displaced 1 AU through space, a star a distance d away will shift by the small angle

$$\theta = 2\tan(1/(2d)) \tag{3}$$
$$\theta \approx 1/d \tag{4}$$

if d is measured in AU and θ in radians. If this angle is 1 arcsecond, that is 1/3600 degree, then d is 206 265 AU. We say that a star that has a *parallax* of 1 arcsecond is 1 parsec from us, where a parsec is 206 265 AU, or 3.26 ly. The nearest neighboring star to our Sun Proxima Centauri, is about 4 ly or 1.2 pc away.

We determine the absolute magnitude of stars as the magnitude they would have if they were at 10 parsecs, 32.6 ly. Thus, at any other distance, the signal s we would detect would be given by the inverse square law from the signal s_0 at 10 pc by

$$s/s_0 = (10/d)^2 \tag{5}$$

Combining this with Eq. 1 gives

$$10^{-2/5(m-m_0)} = (10/d)^2 \tag{6}$$
$$m_0 - m = 5\log(d) - 5 \tag{7}$$

Conventionally the magnitude at a distance of 10 pc is represented by M, and we have the useful relationship between apparent magnitude m, absolute magnitude M, and the distance of the star d that

$$m - M = 5\log(d) - 5 \tag{8}$$

The greater the distance, the fainter the star and the larger its magnitude.

The light from a distant star arrives at Earth, passes through the atmosphere, and is captured by a telescope. The optics of the telescope focus the light onto a sensor, where it is detected through its electromagnetic interaction with electrons in the sensor material. The sensor can be in the retina of your eye, a photographic emulsion, or most commonly today a piece of silicon in an electronic camera. Usually, but not always, one photon produces one electron through the photoelectric effect, and that electron is measured or counted by analog-to-digital conversion devices and digital processing. The result is a number that is proportional to the number of photons that arrived on the sensor during the time it is exposed to light. Each photon has an energy

$$E = h\nu \tag{9}$$

and its arrival carries an energy flux that could be measured in W/m^2-second. The magnitude of a star m is related to this flux by Eq. 1. Thus one additional step for absolute calibration

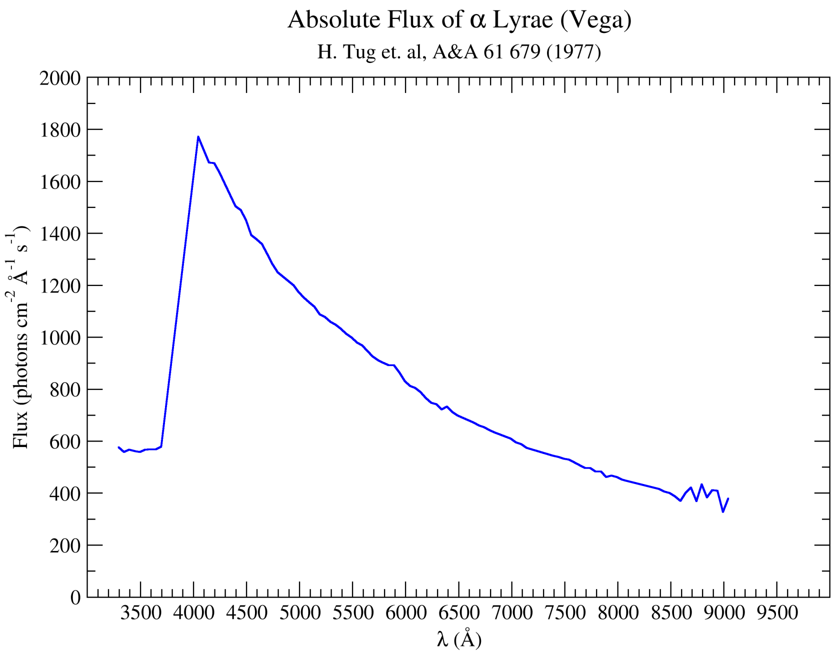


Figure 1: The reference flux from Vega [1]. Roughly there are 1000 photons arriving per square centimeter per second per Ångstrom in the visible. Note the very steep decline below the hydrogen Balmer series limit, and the approximately thermal distribution for a 10 000 K blackbody as wavelengths increase.

is required, and it is very difficult: determine the flux from Vega above Earth's atmosphere in units of either photons/s at every frequency of interest, or equivalently W/s-m^2. Since the star's emission is a continuum, not discrete narrow spectral lines, the reference flux is usually specified either per unit wavelength interval, or per Hz bandwidth, at each wavelength or frequency.

One such measurement still widely regarded as definitive was made by Tüg, White and Lockwood in 1977. [1, 2] They compared the flux of Vega (α Lyrae) and of 109 Virginis to a standard blackbody source and established a reference flux for the visible and near-infrared by accounting for absorption of the light passing through the atmosphere. The result is shown in Fig. 1 below.

3 Astronomical data management

Before we continue on to look at the sensors and the actual data, it will help to understand how the data are managed and what information is available to process. Each photon arriving at the focal plane of a telescope carries information:

- Time of detection
- Integrated over time, the number of detections and the rate of arrival
- Frequency, wavelength, or energy of the photon
- Photon polarization
- Direction from which the photon came

In optical astronomy it is not possible as yet to simultaneously determine all of these factors. For instance, we can "time-tag" a photon but not if many of them are coming in rapidly. We can integrate over a time and find how many came in that time, and in the process lose the information about the arrival of each one. We can determine the photon energy through filters and spectroscopy, but not in the same measurement determine the direction from which it came. Ultimately, the information would be a multi-dimensional data set which can be sliced and viewed in many ways.

The most common slice is to take the arriving photons and sort them by direction onto a two-dimensional surface. This is, in effect, the role of the telescope or camera optics, to send the light to a sensor surface so that each point on the surface corresponds to a different direction in space. Divide that surface into cells of specific usually uniform area, and the photons that arrive in each cell accumulate for a defined time interval. Measure the number of photons in each of these pixels by recording a signal that is related to the number of photons in a data array. Let's call the array $S(i, j)$ where i and j are integers starting with 1 that locate a point in the image on a two-dimensional grid. An image that has $2^{12} = 4096$ pixels horizontally and vertically will have $2^{24} = 16\,777\,216$ or 16 mega-pixels of data. At each element of the array, that is at each pixel, an integer number is stored that we relate back to the actual number of photons that arrived while the sensor was exposed to light. Thus the data are on an integer grid for which we know the number of rows, the number of columns. Since phyiscally each sensor element has a spatial width and height (usually the same), and any photon arriving on the surface in that space may be detected and attributed to that pixel, it is useful to think of the coordinates on the sensor as a continuous floating point measuring position in fractions of a pixel, while the data are associated with the discrete integer positions of the centers of each one. By convention, we start the z the position measurement at the outside corner of one pixel such that the center of that first pixel is at (1., 1.). We most often display this image with that reference corner in the upper left. It's actual orientation on the scene or sky is arbitrary.

Most optical image sensors produce noise as well as real signal, and have a limited "depth" of response before they are no longer sensitive to additional photons while being exposed to light. This well-depth may be as small as 10 000 photons for the sensors used in cell phone cameras, to as large as 10 000 000 photons for some cameras designed from high dynamic range. A typical visible light camera used for astronomy has a full well capacity in each pixel of between $2^{16} = 65\,536$ and $2^{18} = 262\,144$ which we describe as 16- to 18-bits. However, the arrival of photons is a Gaussian random process, and in any given exposure time the probability of a number N of photons arriving has a Gaussian distribution about N with an uncertainty $\sqrt{N}$. The signal-to-noise ratio is $N/\sqrt{N} = \sqrt{N}$. Thus the noise in an 18-bit signal is $2^9 = 512$ and there is nothing gained by measure the least significant 9 bits of the data. Consequently, the hardware which is designed to measure and store images almost never exceeds 16-bits because this covers the useful signal for the full well, and at most misses data that would be significant only for very weak sources just above the actual electronic noise. The reason for mentioning this here is that the useful range of integer data is typically 16-bits, while 32-bits may be needed in exceptional cases. Generally raw data are stored as integers to save storage space, while processed data require floating point since at some stage physical units are introduced rather than the original digitization units of the device.

This sets the stage for managing the data. We need a file structure that accommodates the data itself as an array of values that may be floating point or integers requiring 16 or more bits to represent, and we need additional information that specifies at a minimum the dimensions of the array, the exposure time, the time and date at which it was recorded, and the instrumentation that produced it. In astronomy the preferred mechanism is a Flexible Image Transport System or FITS file format. While the details may be left to computer programmers and are subject to a defined protocol, this system allow images with floating point or integer values in any array dimension, and has a header component that informs the user of the size of the array and its data type, with specfic entries for other optional information describing the data and its processing. Lossless compression is often applied to the entire file if storage space is limited, though this also consumes processing time with the data are accessed. Here is an example of the header of a FITS file that originated with a camera on a telescope and was processed for precision photometry:

```
SIMPLE  =  True
BITPIX  =  -32
NAXIS  =  2
NAXIS1  =  4096
NAXIS2  =  4096
EXTEND  =  True
COMMENT  =    FITS (Flexible Image Transport System) format is defined in 'Astronomy
COMMENT  =    and Astrophysics', volume 376, page 359; bibcode: 2001A&A...376..359H
EXPTIME  =  100.0
DATE-OBS  =  2016-08-14T12:07:54.434
IMAGETYP  =  Light Frame
TARGET  =  KS36C048025
INSTRUME  =
CCD-TEMP  =  -10.06
FILTER  =   r_(530-700)
TELESCOP  =  CDK700
```

```
DATE  =  2016-08-14T12:09:53
JD_SOBS  =  2457615.00549
JD_UTC  =  2457615.00607
HJD_UTC  =  2457615.00802
BJD_TDB  =  2457615.00878
ALT_OBJ  =  45.71574672
AZ_OBJ  =  218.786028266
HA_OBJ  =  3.272129028
ZD_OBJ  =  44.28425328
AIRMASS  =  1.3951692652
RAOBJ2K  =  16.53116
DECOBJ2K  =  -54.5866472222
RA_OBJ  =  16.5535875184
DEC_OBJ  =  -54.6218490177
SITELAT  =  -27.7978611111
SITELONG  =  151.855416667
HISTORY  =  Previous Filename = ks36c048025_00100.fits
HISTORY  =  Bias corrected with bias.fits
HISTORY  =  Non-linear corrected with coefficients:
HISTORY  =  a0 = 0.0
HISTORY  =  a1 = 1.0
HISTORY  =  a2 = 8.0E-7
HISTORY  =  a3 = 0.0
HISTORY  =  Dark corrected with (deBiased) dark_100s.fits
HISTORY  =  and exposure time scaling factor = 1.0
```

The first lines describe the data type, and `BITPIX = -32` means it is floating point with 32 bits. There are two axes as expected, and each has 4096 values, that is, it is a 16 megapixel image. The exposure time was 100 seconds at UTC 2016-08-14T12:07:54.434 that may mark the beginning or the middle of the exposure depending on how the software that wrote the image was designed. Other details describe the filter that was used, the telescope, the object, and a history of processing. Additional header information may add the celestial coordinate system as well, though in this case that was not done.

Because images can be large spatially, and require two or more bytes per pixel, astronomical image file sizes are often much bigger than those encountered from consumer cameras. For example this particular image takes 67 megabytes to store the header and the $4096 \times 4096 \times 4$ bytes of the image. A single night of observing may capture 200 or more images, that is 10's of gigabytes per night for a small telescope. The data rate for the Large Synoptic Survey telescope's 3.2 gigapixel camera will be 20 terabytes per night, to be processed and distributed in real time. [3] The management of such large data sets is science itself. One of the issues is that in processing images the software that is used is critical, and it may change with time. Some file systems that do not use the FITS standard are designed to track not only the data, but details of the processing software, making a larger data structure that

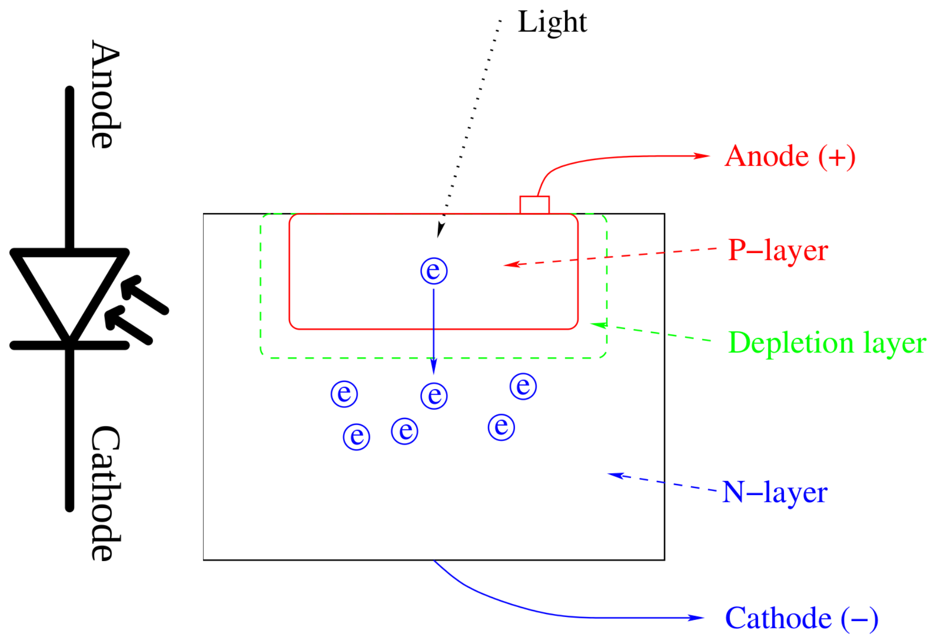


Figure 2: Photodiode electronic symbol (left) and physics (right).

includes everything. Obviously this becomes an even bigger management problem which is currently being studied and developed in preparation for the so-called era of "Big Data".

4 Silicon Photodiodes

An optical electronic image sensor converts the incident photon to one or more electrons which is then subsequently measured and digitally stored. The key element in modern image sensors is a photodiode based on silicon, illustrated in Fig. 2.

Light is incident on a P-type semiconductor layer on silicon that is produced by diffusing boron to a depth of the order of 1 μm (10^{-6} meter) into the substrate. Below that is an N-type region and between them is a depletion layer boundary. Photons of sufficient energy promote an electron over the band gap from the valence band into the conduction band. The potential difference creates a field that accelerates the electrons and their corresponding holes in opposite directions across the depletion layer. The result is that positive charge builds up in the P-layer valence band and negative charge in the N-layer conduction band. This charge may be stored in the intrinsic capacitance of the device, stored in an external capacitor, or delivered as a current to an external circuit. The charge is proportional to the number of photons incident and the quantum-efficiency of the detection is determined by the wavelength of the light and its penetration into the regions that are photosensitive. [4]

5 Photodiode arrays

While a single photodiode measures the light at a single point in an image, an array of photodiodes captures the entire scene. The early efforts at building such devices were derived from technology invented to serve as computer memory, and were spatially linear with electronic switches to connect each diode to a common bus. The electrons saved in the capacitance of diode would be moved to the bus and then measured as a voltage, one diode at a time. For spatial imaging, clearly the technical difficulty is how to make a connection each of the sensor elements, since an array with 1024 pixels in both directions would have $1024^2 = 1\,048\,576$ pixels presumably covering the surface. More typically for astronomy the arrays will be even larger, with 16-megapixels being typical over an area 50 mm on the diagonal. Two methods are in use today enabling communication from each sensor element to the host processing system. In one, the charges are passed bucket-brigade style off the sensor or "chip", and in another, transistor switches make the connection. The bucket-brigade transfer imager is called a "charge coupled device" or a CCD. The switched transfer imager, which is much more complex because of the large numbers of transistors it requires, uses a multilayered structure of "complimentary metal oxide on silicon" and is called a CMOS imager. Until recently the best CCD sensors have had lower noise, larger charge storage capacity, better quantum efficiency, and more spatial uniformity than CMOS sensors. From their first appearance as devices with a 100,000 elements built one with care in the lab, to today's highly engineered commercial scientific CCDs, they were more suitable for astronomy and replace photographic film within a few years. Comet Shoemaker-Levy 9 that struck Jupiter and left dark blotches in its atmosphere was discovered by Carolyn Shoemaker with confirmation by her husband Eugene and David Levy using carefully processed film and a large telescope. Today, CCD sensors on small backyard telescopes detect as many or more photons because of their high quantum efficiency, and make the amateur's telescope as powerful as professional telescopes of only a few decades ago.

Another revolution, perhaps less impactful but just as technically significant, is the development of CMOS sensors that are lower in intrinsic noise, higher in quantum efficiency, and read out much more rapidly than CCDs. Sometimes termed "scientific CMOS" by manufacturers of cameras for biology and astronomy, the best of these sensors are also incorporated into commodity products – your cell phone and professional digital cameras. With these devices, the spatial scale of the pixels is usually finer than CCDs, and tiny chips may have many millions of pixels on a spacing of under 5 microns. Typically CCDs have larger pixels, providing greater capacitance and capable of storing more electrons. Thus the choice of sensor may depend on the astronomical application, though post-detection processing can bin pixels and effectively add up the electrons from adjacent pixels to reduce the noise while virtually making the set of pixels a single superpixel.

Both CMOS and CCD devices are structured as a uniform array of photodiodes in rows and columns on a single piece or substrate of silicon. The diodes cannot cover the entire surface since they have to be insulated from one another, and there has to be a provision for connections to them. Especially in CMOS sensors, the fraction of the surface covered, the "fill-factor", affects the efficiency of the device as well as its ability to render fine detail. In a common design, the sensitive regions are illuminated through transparent gate electrodes, and incorporate microlenses to collect light from the entire surface and direct it onto the photodiode pixels. Another solution is to use a very thin substrate and illuminate it from the back such that the light goes directly to the sensor element without the intervening connections. Backside illuminated image sensors were difficult to produce in the first imaging arrays and are still uncommon and expensive in CCD technology, but Sony and other manufacturers are producing them in large quantities for their CMOS sensors, including ones found in the latest cell phones, digital single lens reflex cameras, and low light surveillance cameras.

5.1 CCDs

After an exposure, charge is passed across the device to a column, one pixel at a time out of each row. The charge in the buffer column is passed to a floating diffusion sense node, and then to a linear source follower amplifier. The floating diffusion node has a potential level that varies depending on the charge present, and it is this potential that is then amplified. The amplifier's output is converted into a voltage and is current amplified in order to drive electronics off the chip. Once transferred, amplified, it is measured one pixel at a time by an analog-to-digital converter. In scientific CCD's the converter does a 16-bit digitization of the signal and a computer program stores a number from 0 to $2^{16} - 1 = 65535$ that is proportional to the captured charge. Since one electron is produced by each photon, at most 65535 photons can be counted in each exposure. However, because the noise in reading the device is of the order of 10 electrons, and because the quantum or "shot" noise is $\sqrt{N}$ where N is the number of photons, it is not necessary to have a 1:1 correspondence of photons to digital units. Typically, this gain is of the order of 1.5, such that 65535 represents about $65535 \times 1.5 = 98302$ photons. This allows conservative use of digital storage, but does not lose information because of the noise inherent in the signal, and a typical well depth of 100 000 electrons for astronomical devices. Consumer devices, or high speed sensors with small pixels, may operate effectively with 14-, 12- or even 8- bit digitization.

Schematically we can picture an element of a CCD pixel as shown in Fig. 3. The pixel wells are controlled by gates that are set by external electronics, clocked to shift the charge from one pixel to another. Charge movement in unwanted directions is controlled by modifying the silicon to stop its migration, so that rows are isolated from one another, while columns may have charge moved across and ultimately off the chip. Fig. 4 illustrates the process for an On-Semi CCD originally designed by Kodak in which the readout arrangement on the array is shown in Fig. 5. Fig. 6 summarizes the specifications for one of the largest CCD's manufactured in commercial quantities. Note the parameters of pixel size ($9\,\mu m$), quantum efficiency at the wavelength of peak sensitivity (60%), and readout rate (10 MHz). In practice, to precisely measure the electrons in a deep CCD well, readout rates are much slower and 500 KHz is common in the best astronomical CCD cameras. At that rate, an image with 16 million pixels is retrieved from the sensor in $16 \times 10^6/500 \times 10^3 = 32$ seconds

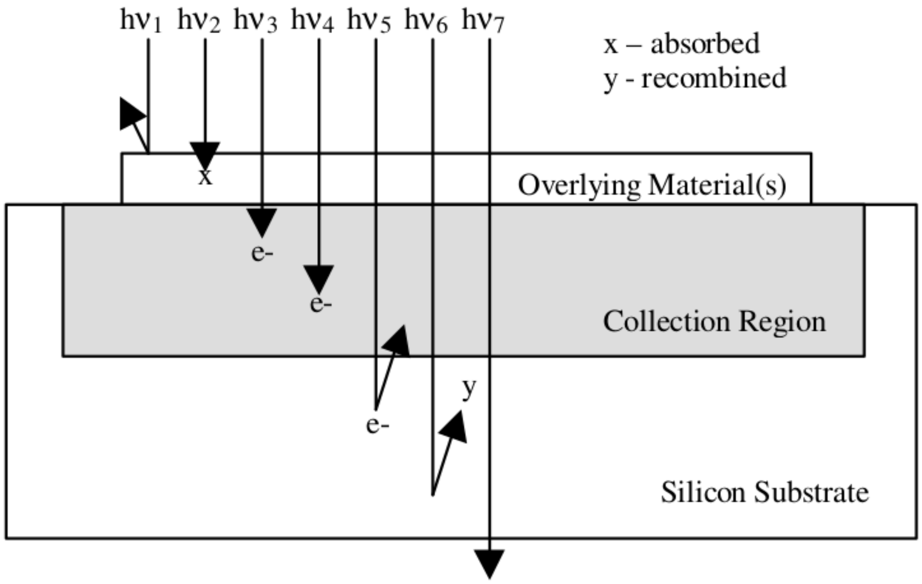


Figure 3: Photons interact with the photodiode to produce electrons that are collected on the device. [5]

if read serially through one amplifier.

This specific chip is the largest of a family that included others in different pixel sizes. For many years the KAF family was the most widely used in astronomical cameras for small telesscopes. It was originally designed by the Eastman Kodak Company and intended to be its gateway to electronic imaging when photographic film was in steep decline. Kodak went bankrupt, and the team that made these sensors became a Truesense, a new company that survived for several years and was purchased by OnSemi, a global semiconductor manufacturer. The entire line of CCD sensors of this type was discontinued by OnSemi in 2020 because of the rise of CMOS.

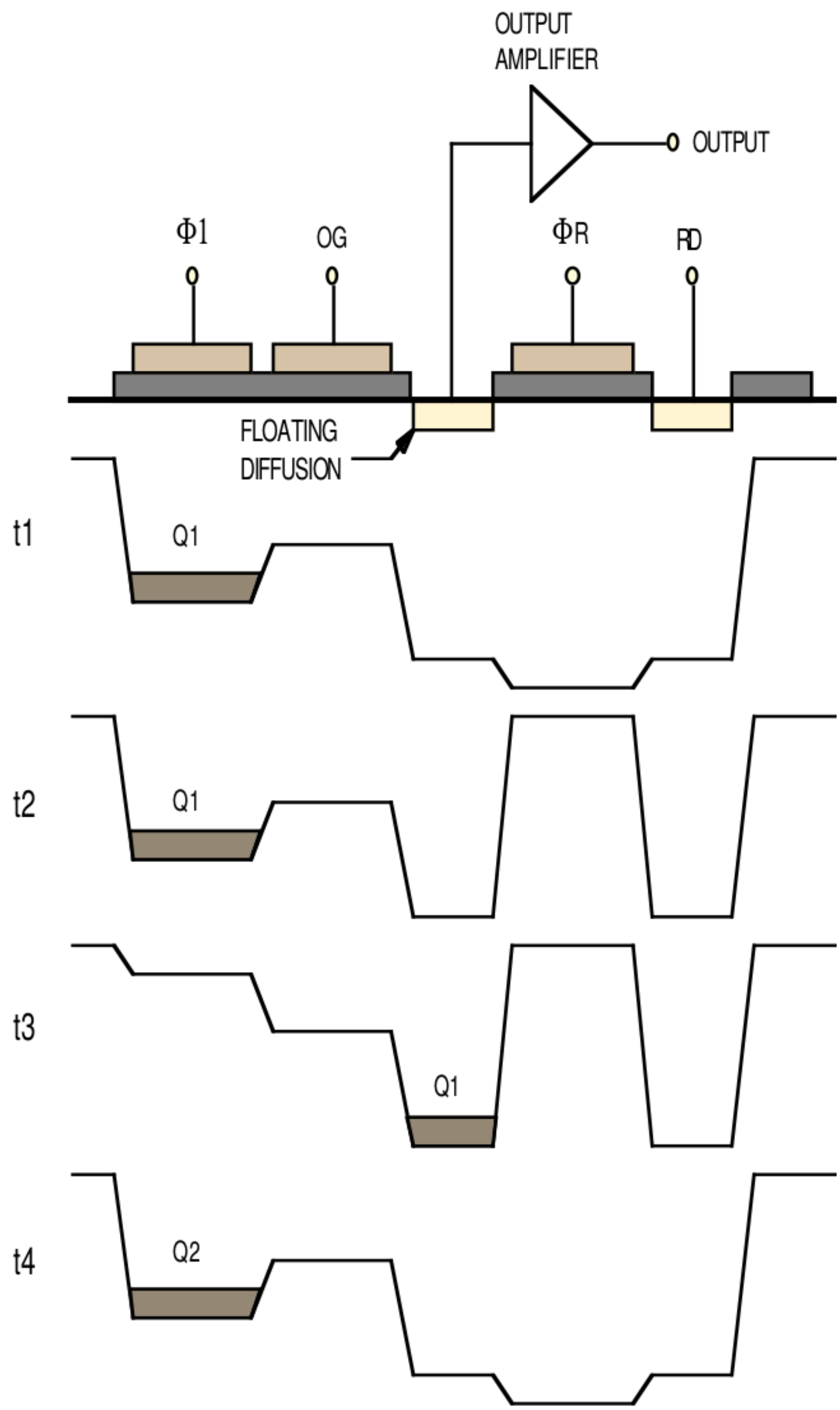


Figure 4: Electrons are transferred from the photodiode collector to an amplifier. [5]

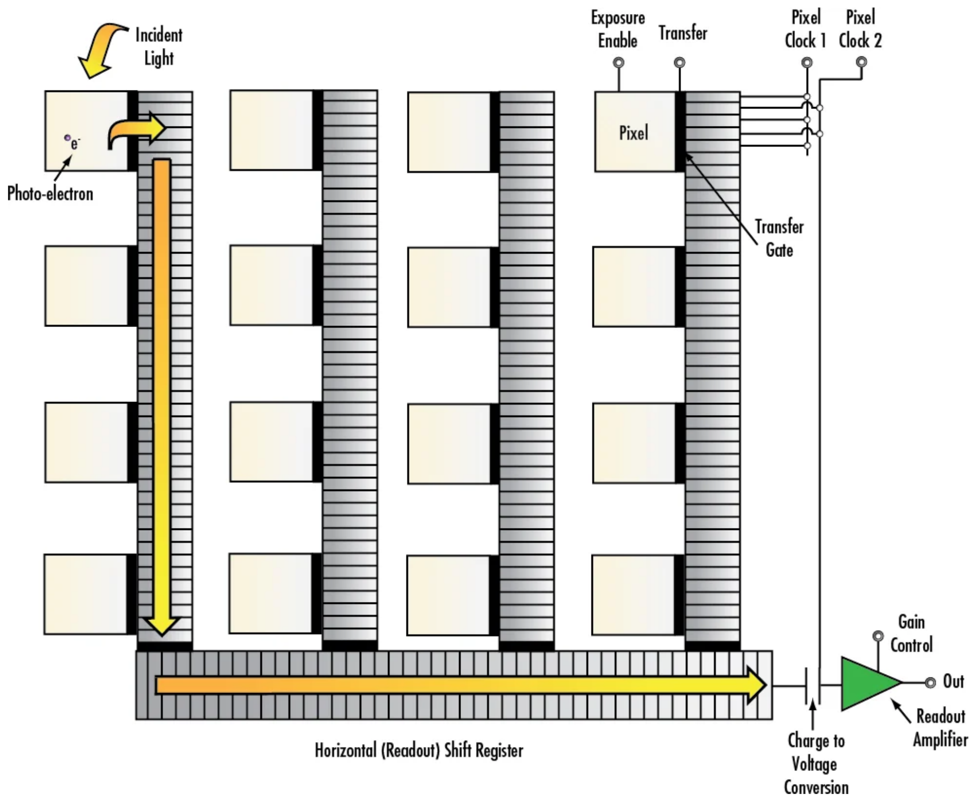
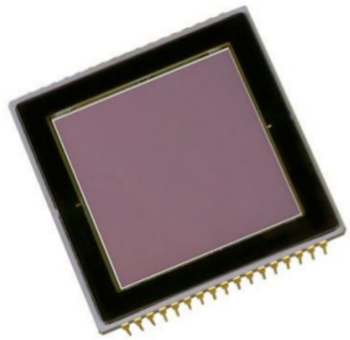


Figure 5: A CCD readout moves data in serial along a column pixels, bucket-brigade style, and then out along a non-imaging row called a shift register at the edge of the sensor. [6]



Parameter	Typical Value
Architecture	Full Frame CCD; Square Pixels
Total Number of Pixels	4145 (H) x 4128 (V) = 17.1 Mp
Number of Effective Pixels	4127 (H) x 4128 (V) = 17.0 Mp
Number of Active Pixels	4096 (H) x 4096 (V) = 16.8 Mp
Pixel Size	9 µm (H) x 9 µm (V)
Active Image Size	36.8 mm (H) x 36.8 mm (V) 52.1 mm Diagonal
Aspect Ratio	1:1
Horizontal Outputs	1
Saturation Signal	100 ke^-
Output Sensitivity	22 $\mu V/e^-$
Quantum Efficiency (550nm)	60%
Responsivity (550 nm)	28.7 $V/\mu J/cm^2$
Read Noise (f = 4 MHz)	9 e^-
Dark Signal (T = 25 °C)	3 e/pix/sec
Dark Current Doubling Temperature	6.3 °C
Linear Dynamic Range (f = 4 MHz, T = 25 °C)	80 dB
Blooming Protection (4 ms exposure time)	> 100 X saturation exposure
Maximum Data Rate	10 MHz
Package	CERDIP (Sidebrazed, CuW)
Cover Glass	AR coated, 2 sides and Taped Clear

All parameters above are specified at T = 25 °C, unless noted otherwise

Figure 6: Specificiations for the Kodak/Truesense/OnSemi KAF-16803 used with many small telescopes is similar to other sensors of the same family originally developed by Eastman Kodak. [7]

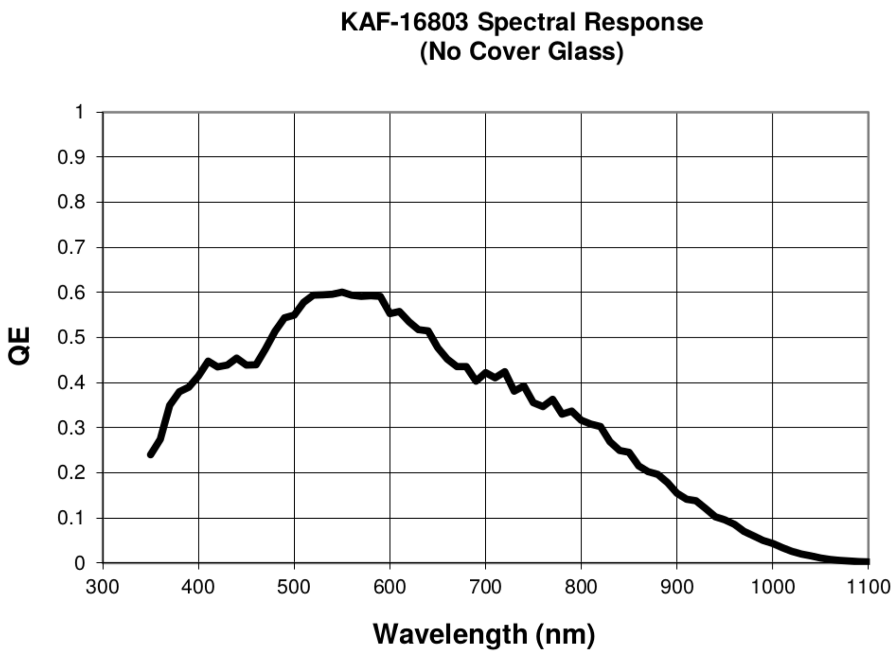


Figure 7: Quantum efficiency of the KAF-16803 CCD. It responds from 350 nm in the ultraviolet to 850 nm in the infrared with an efficiency above 25%. The near infrared response extends to 1 μm at lower efficiency. [7]

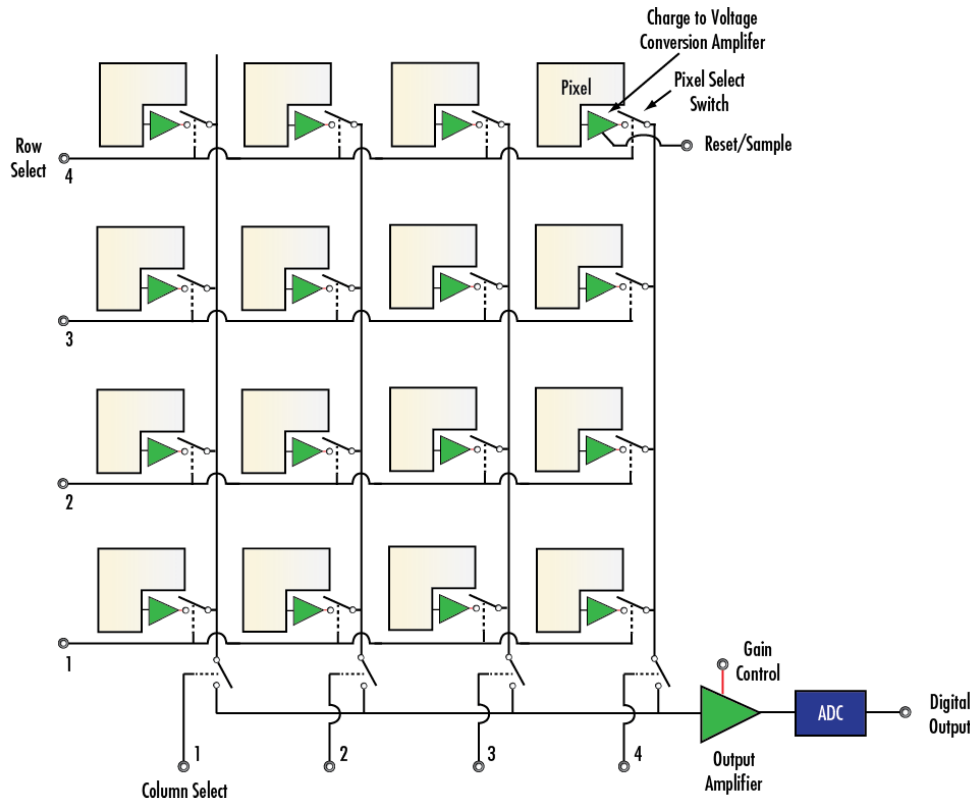


Figure 8: A CMOS readout moves data by switching each pixel to a bus. [6]

6 Scientific CMOS sensors

The architecture of a CMOS image sensor with backside illumination is illustrated in Fig. 8. It may provide multiple output amplifiers and with many paths off the device and fast switching, it is usually faster to read a full image than a CCD. With this design, CMOS sensors can do ultra high speed video even reaching thousands of per second, while similar CCD sensors are reading at many seconds per frame. Although the transistor switching introduces noise, recent designs minimize readout noise to less than statistical noise in the signal.

The Sony IMX455 is a back-illuminated CMOS sensor intended for surveillance use applications. It is equivalent to a 35 mm camera's full frame measuring 43 mm on a diagonal with 9568×6380 active 3.76 μm pixels. A built-in 16-bit analog-to-digital converter (ADC) and 8 lanes of output allow video frame rates. With a mask of color filters it, and its even more advanced variants, have many applications. Without the mask, as a monochrome imager where each pixel is unfiltered, it has been adopted for astronomy as well. The quantum efficiency in Fig. 9 is higher than competitive CCDs.

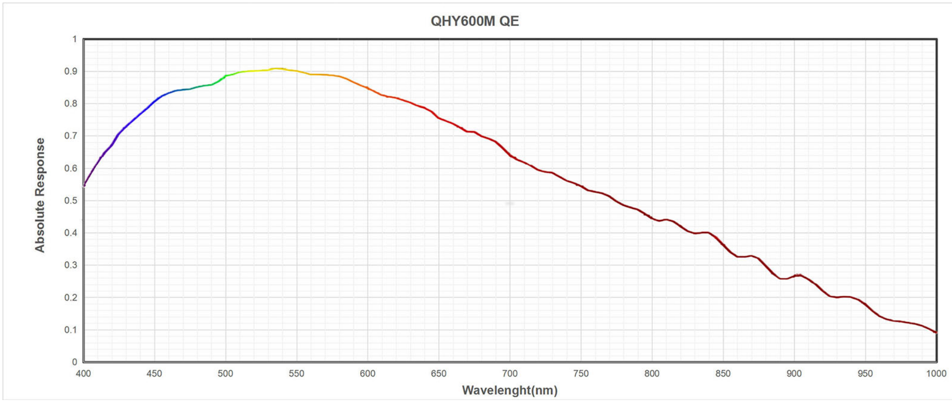


Figure 9: Sony IMX455 back-illuminated CMOS image sensor has a peak quantum efficiency of 90%. It covers 350 to 900 nm with more than 25% efficiency that remains high into the blue. Like most silicon CMOS and CCD sensors its sensitivity declines in the near infrared and vanishes above 1 μm. [8]

7 Commercial cameras

For NASA space missions, defense applications, and the largest of major observatories imaging cameras are developed in-house, usually in collaboration with sensor fabricators. For everyone else, there are companies that specialize in scientific imagers. The following links to leading manufacturers cover those focused on amateur and university astronomy, and those with broad (and expensive) products that are at the cutting edge of technology. Typically, small CMOS or CCD cameras cost from a few hundred to $30,000, while the very large cameras with specialty sensors for high speed or low noise may cost from $50,000 to more than $100,000 each. This might be compared to the cost of the best of commercial cameras for photography, which typically are in the $5,000 to $10,000 range. The differences are in the technology – scientific cameras are designed for low noise, not necessarily maximum pixel count; they have larger pixels; and they must be cooled so that long exposures will not accumulate thermally generated electrons. Some of these links provide educational pages about the technology and diverse applications.

Diffraction Limited /SBIG

ASI/ZWO

QHY

Fingerlakes Instruments

Oxford Instruments Andor

Teledyne Princeton Instruments

8 Basic processing of image data

8.1 Gain and non-linearity

We assume that the raw image data have been pre-processed to remove the bias and dark signals such that a numerical 0 is in fact no light, apart from residual *read noise*. The remaining signal for most devices will increase linearly with the number of photons that fall on each pixel during an exposure. The most common defect that may appear in short exposures is the effect of the shutter, which may have more open time at the center of the image than at the edge, or have a flower-like pattern as its leaves open and close with exposure time variations across the device. Before proceeding it is good to be aware of these effects, though only rarely necessary to compensate for them. For example, the exposure time variation due to shutter transit time is always of the order of a few milliseconds. For an exposure of 10 seconds, this would be 1 part in 10^3, a milli-magnitude in astronomical terms, and well within the noise due to photon statistics for a single exposure saturating at less than 10^6 photons per pixel. The exceptions would be for shorter exposures (say 1 second), slow shutters, or co-added frames of very many short exposures.

Thus we are left with a digital value that monotonically increases with the exposure time and the total number of photons to arrive during that time. Non-linearity arises in either the transfer of the charge out of the storage site, or in its amplification. Typically the charge transfer efficiency is so high that, excepting loss when the well is full or saturated, the remaining effect is a small non-linear coefficient in the relationship of the digital signal to the total photon count. If full well capacity is 10^5 photoelectrons, a typical value, then the statistical noise in one pixel is $1/\sqrt{10^5} = 3.2 \times 10^{-3}$, equivalent to 3 milli-magnitudes. A non-linear effect of this order might be significant for a single pixel, but is surely significant when many pixels are added such that the total count exceeds 10^6 coming from some pixels exposed to near saturation, and others with much lower signals. Thus whether non-linearity needs to be in included in processing at this stage depends on the expected outcome, and the precision required.

The non-linear transformations are applied to the pre-processed data with bias and dark removed. Ideally, this removal is also done with non-linear corrections in the bias and dark frames, but that is not really necessary as can be seen from the size of those corrections compared to signal noise in the final data set. One would have a set of coefficients that represent the response of the detector to p photons, where $p = \mathcal{F} \cdot t$ for a "flux" $\mathcal{F}$ at the sensor and an exposure time t. We do not usually consider these terms separately, though they do have their own effects. In sensors such as self-gain devices like avalanche photodiodes, high fluence may mean missed photons (i.e. non-linearity) that would not happen at lower fluence. In those cases long exposures of faint sources may yield a different total than what would seem to be the equivalent short exposure of a bright source. More commonly, all we are concerned with is the amplifier non-linearity once the photoelectrons have been transferred across the device to the output node. We might represent the measured signal s this way

$$s = (b_1 p^1 + b_2 p^2 + b_3 p^3 + ...) \tag{10}$$

Since the signal is now a voltage and not a digital count, there is an overall gain that may not be known relating the count p to the digital value s. Here we have incorporated the gain into each term, while in practice we may factor it out to become a global parameter. Typically the gain in front-illuminated Truesense (Kodak) detectors supplied in Apogee cameras is of the order of 1. The non-linear coefficients b_n are then all relative to unity when gain is factored out, and typically will be $b_n \approx 10^{-3} p_{\mathrm{fw}}^{-n}$ where p_{fw} is a full well charge of the order of $10^5\, e^-$.

Given these small b_n, non-linear corrections are usually necessary only when comparing signals of considerably different signal strengths where one of them is close to the full well capacity. This may occur in precision photometry to find small variations in a signal, as is done in measuring exoplanet transit events, stellar pulsation, and other sources of periodicity, when the point-spread-function is oversampled or the image is highly defocused. The response at each pixel is then taken to be have a linear factor that depends on the pixel, a global non-linear response due to the charge transfer process and amplifier, and a systemic gain that puts the signal on an absolute scale. Thus after flat fielding we would apply a transformation

$$p = gs(1 + a_1 s^1 + a_2 s^2 + ...) \tag{11}$$

where g is the inverse gain (photons/adu), and the polynomial in parenthesis compensates for non-linearity. The compensation depends on the signal at that pixel but with non-linear coefficients that are everywhere the same. The gain is the calibration in the small signal limit. In the following discussion of setting the absolute calibration we will consider only this limit since it applies in first order to large signals as well, and may be approximately corrected by non-linear coefficients in the transformation from signal value to photon count.

8.2 How do we measure the non-linear coefficients?

For those precision applications where it is needed, one obvious issue is how do we determine coefficients so that from a non-linear signal s we can infer a correct p? In practice it is not possible to control the actual flux from a source such that for a constant exposure time various total photon counts are delivered to the detector. Nor are astronomical standards available with enough dynamic range and precision, not to mention freedom from color and dependence on Earth's atmospheric extinction. One solution is to find them by fitting measured signals for a series of exposures of a stable uniform source at exposure times long enough not to have a shutter effect, and short enough to be practical for source stability. [9] With longer exposure, more photons are delivered to the pixel, and more photoelectrons are produced to go through the charge transfer and amplification. The precision measurement of exposure time replaces the precision measurement of light source flux.

The bracketed repeat-exposure method [10, 9] takes single and multiple exposures of different durations: bias, $2\,s$, $2 \times 2\,s$,$2\,s$, $3 \times 2\,s$,$2\,s$, ..., with interspersed $2\,s$ exposures that monitor the source stability. Readout time is minimized by measuring only a small part of the CCD frame since the idea is to find a global fit to the non-linearity. By including many pixels, however, the total count is large and the statistical noise is minimized. The signals are fitted with a polynomial in which the correction coefficients add only a few percent to the total observed signal close to saturation, so it takes a total count of 10^6 or more to find the coefficients accurately. In Eq. 11 the values of s are then total signals over a region of the CCD including 100 or more pixels, and p is the expected signal as a multiple of the 2 s exposure. The ratio of p/s is fitted dependent on s so that the coefficients obtained are the ones needed to convert from measured s to actual photons p.

8.3 Exposure time and the relationship to flux standards

In astronomy we usually refer to the process of making a precise measurement of the light emitted by a source as "photometry", in deference to the historical use of the human eye. In other fields this would be called "radiometry" to reflect a process that has a physical, rather than physiological, basis. We have to measure the light emitted by point sources (stars at great distance) and extended sources (Earth's airglow and aurora, planetary surfaces, distant diffuse nebulae). The mathematics and terms used in these two cases differ, though the same factors determine the signal in both cases.

Brightness in conventional usage applies to an extended source, and is more appropriately termed *radiance* given in units of $W/m^2 - sr$, or $photons/m^2 - sr$, or (for astronomers) $photons/cm^2 - sr$ at a surface. We will use B to represent this, with the understanding that it may also be given per unit wavelength or frequency interval rather than integrated over the entire spectrum. In astronomy we often measure energy and area in erg/s and cm^2, using CGS rather than $joule/s$ or W and m^2 in *SI* units. The *Rayleigh* is a convenient standard CGS unit from airglow and auroral studies, $10^6/4\pi$ $photons/s - cm^2 - sr$, where a steradian is 4.25×10^{10} $arcsec^2$. Its conversion to power depends on the wavelength of the light, with a multiplication by $h\nu$ or hc/λ for each wavelength contributing to the flux. In some planetary science literature the term *intensity* is ambiguously used to refer to radiance, a point of confusion that is usually resolved in context.

For a distant extended source that is observed from Earth with an optical system of area A, we take the area of the source that contributes to the signal to be the area of the optical system. The solid angle in radiance or brightness is the solid angle of the sensor element seen by the source, that is the solid angle on the sky that the pixel covers. Let F be the focal length of the telescope, D its diameter, and d the spatial size of a square pixel on the detector. The solid angle of a CCD pixel is $\delta\Omega = d^2/F^2$. The area of the telescope is $A = \pi D^2/4$. For a source radiance B in *Rayleigh* designated B_R, we have P photons arriving at the sensor element in time t given by

$$P = B \cdot A \cdot \delta\Omega \cdot t \tag{12}$$

$$P = B_R \cdot (10^6/4\pi) \cdot (\pi D^2/4) \cdot (d^2/F^2) \cdot t \tag{13}$$

at *each* pixel. The right hand side simplifies to

$$P = 10^6 \cdot B_R \cdot (1/16f^2) \cdot d^2 \cdot t \tag{14}$$

where $f = F/D$ is the telescope f-ratio. Notice that the *telescope* area goes out of the equation, leaving a flux of photons at the pixel proportional to f^{-2}. Clearly a fast optical system will produce a much larger signal than a slow one for the same pixel size. Of course, a small aperture will also have a larger image scale (arcseconds/pixel) for the same pixel size at fixed focal ratio. We take as an example a 12 μm square pixel on a 530 mm focal length f5 telescope, and the North American Nebula (NAN or NGC 7000) with a radiance of 850 ± 50 *Rayleigh* Hα [11], we would see a flux at each pixel of

$$P/t = 850 \times 10^6 \cdot (1/16 \times 25) \cdot 1.44 \times 10^{-6} \quad (15)$$

$$P/t = 3.06 \text{ photons/s} \quad (16)$$

The signal received for a point source, by contrast, depends on the area of the telescope. Consider, for example, a star with an irradiance E in photons per unit area at the top of the Earth's atmosphere. This light, arriving at an optical system, is collected by a telescope of area $A = \pi D^2/4$ and focused onto one or more pixels. We are interested in the *total* photon count arriving at the detector in this case, rather than the count per pixel. With sufficient image quality and large enough detector pixel size, all of the light will enter only one element. More often, atmospheric blurring, defocusing (sometimes intentionally), and diffraction will spread the light over many pixels. In the subsequent analysis of the CCD image data, the photons from the one star are summed by adding the count in all the pixels and removing an estimate of the background from the sky or other stars. The detected signal will be

$$P = E \cdot A \cdot t \quad (17)$$

$$P = E \cdot (\pi D^2/4) \cdot t \quad (18)$$

The star Vega is a convenient example because it sets the zero point of the astronomical magnitude scale. The irradiance from Vega at the top of Earth's atmosphere is approximately 1000 $photons/s - cm^2$–Å throughout the visible spectrum. [1, 2]. Now the problem is to compute how much of the spectrum is actually allowed to get to the sensor, and of that, how much the sensor can respond to. If naively we assume a bandwidth of 1000 Å as is approximately typical for a standard astronomical broadband optical filter (e.g. B, V, R, or g, r, i), then with 10^6 $photons/s - cm^2$ at the same telescope as our NGC 7000 example, we have

$$P/t = 10^6 \cdot \pi 10.6^2/4 \quad (19)$$

$$P/t = 2.81 \times 10^7 \quad (20)$$

We see that Vega produces an enormous signal in one or a few pixels, while the NAN produces a much smaller signal per pixel, but because of its angular size may still have a large total integrated signal. To continue this example, since Vega is zeroth magnitude, a star of magnitude 15 would provide 10^{-6} of this signal or only 28 *photons/s* for this small telescope. Nevertheless, if the goal is a 1 percent measurement, then only 10^4 photons are required. At this rate, the exposure time would be approximately 6 minutes. A limiting magnitude for 10 percent accuracy (100 photons) in one hour (3600 seconds) would be 18.9 if the sky were dark enough not to contribute significant noise.

9 Glossary

For reference, some of the terms introduced in the text are collected here. Astronomical photometry has its own language and usage which may differ from the conventions of optical literature. Often the details are clear in the context. These are the more formal definitions that should be followed if possible.

Radiometry nomenclature

There is a prevalence in the physics and astronomy literature for inconsistent use of words such as "intensity" and "flux", which mean some measure of how much light arrives in a given time. In this work we try to use them with the presently accepted definitions given by Born and Wolf [12], the Oslo Optics Reference [13], and Wikipedia [14]. Many of the issues developed because of dual usage of terms in the description of a physical measurement of energy (radiometry), and a physiological response to light (photometry). Unfortunately, in astronomy, even the term *photometry* is used to describe the physical measurement of the energy from stars. The system of astronomical photometry, dependent as it is on choices of optical filters to mimic the eye's response, is historically based on visual measurements. The standard measure of a stellar magnitude, for example, is in the "V" for *visual* spectral band. In fact the V-band is actually mostly green light, and is the wavelength region to which the human eye is most sensitive.

Flux

Flux is the rate of change of energy, denoted by Φ or occasionally by F. It is given in SI units of watts (W), CGS units of (erg/s), or more simply ($photons/s$) for monochromatic light.

Irradiance

Irradiance is the flux per unit area on a surface. It is denoted by E and given in SI units of watts per square meter (W/m^2).

Intensity

The "trouble child" of radiometry, *intensity* is often misused to represent the detected power in a signal, but it is actually a radiometric measure of total flux per unit solid angle *from* a source. It is denoted by I and given in SI units of watts per steradian (W/sr).

In radiative transport theory as applied to stellar and planetary atmospheres, the term *specific radiative intensity* is used, often shortened simply to "intensity" and also represented by I. Although it is a radiometric quantity, it is not the same as "intensity" used in radiometry, but instead dimensionally equivalent to brightness (see below) per unit frequency. In SI units, specific intensity is watts per square meter per steradian per hertz ($W/sr - m^2 - Hz$).

In some work, the symbol B is used instead, suggestive of brightness. Whereas radiance or brightness refers to the light emitted from an area, specific intensity is usually used in the context of energy flow, representing the balance of radiative energy through an area that may be emitting and absorbing, or may be free space. Radiance would be the preferred term but intensity is ingrained in the astrophysical literature. [15, 16, 17]

Radiance or Brightness

The radiance is the flux *emitted* per unit solid angle per unit area. In physics and astronomy literature it is often denoted by B and given in SI units of watts per steradian per square meter ($W/sr - m^2$). The use of B to represent radiance comes from an alternative term, *brightness*, for the same quantity. As noted above, in the context of radiative transport theory, radiance (usually represented by B) is referred to as "specific radiant intensity", or simply "intensity", and may be given per frequency unit as well, or in units of ($W/sr - m^2 - Hz$).

Usage of the term brightness instead of radiance is discouraged in illumination engineering literature because of the physiological association of the term. Instead, luminance, which would be given in units of candela per square meter (cd/m^2), is used in analyses of illumination for visual use. Both the physical radiometric radiance and the physiological photometric luminance are also sometimes called *brightness* in the optical science literature, and in those cases its meaning has to be found by context.

In some planetary science literature the term *intensity* is ambiguously used to refer to radiance, a point of confusion that is usually resolved in context.

Rayleigh

The *Rayleigh* is a CGS unit from airglow and auroral studies, $10^6/4\pi$ $photons/s - cm^2 - sr$, where a steradian is 4.25×10^{10} $arcsec^2$. Its conversion to power depends on the wavelength of the light, with a multiplication by $h\nu$ or hc/λ for each wavelength contributing to the flux.

Lambertian

Following Born and Wolf ([12] section 4.8.1), a source having a radiance independent of direction is called "Lambertian". Blackbody sources are Lambertian, and most diffusing self-luminous sources are approximately Lambertian.

Consider light emitted emitted from a surface element of area dA into a solid angle $d\Omega$ in the direction (θ, ϕ) at the point (x, y) on the surface. It is given by a radiance $B(x, y, \theta, \phi)$ in SI units of watts per steradian per square meter ($W/sr - m^2$). (Note: Born and Wolf call this brightness, but see our note about the term *radiance*.) The flux from a surface element would then be

$$dF = B(x, y, \theta, \phi) \cos(\theta) \, dA \, d\Omega \tag{21}$$

where θ is the angle in spherical coordinates from the local surface normal to the direction of the light emission. The intensity I (given in W/sr) is the flux per steradian from B, integrated over the scattering surface area

$$dI = \frac{dF}{d\Omega} = B(x, y, \theta, \phi)\cos(\theta)\, dA \tag{22}$$

$$I(\theta, \phi) = \int B(x, y, \theta, \phi)\cos(\theta)\, dA \tag{23}$$

The radiance of a surface depends on its nature (mirror smooth or sandpaper rough, for example), composition, and whether it is illuminated or self-luminous. The simple Lambertian assumption is that *the radiance B is independent of direction*, although not necessarily of position. It yields

$$dI = \frac{dF}{d\Omega} = B(x, y)\cos(\theta)\, dA \tag{24}$$

$$I(\theta) = \int B(x, y)\cos(\theta)\, dA \tag{25}$$

$$I(\theta) = I_0 \cos(\theta) \tag{26}$$

$$\tag{27}$$

where

$$I_0 = \int B(x, y)\, dA \tag{28}$$

For a Lambertian surface, the intensity (W/sr) is proportional to the cosine of the angle of emission to the surface normal.

BRDF

The **b**idirectional **r**eflectance **d**ististribution **f**unction (BRDF) is a radiometric factor describing the scattering (redistribution) of light incident on a surface. The "B" in bidirectional comes from its reciprocally equivalent dependence on the directions of incident and scattered light. It is considered a function not only of surface material and the position of a scattering point if the surface is not uniform, but also a function of direction, spectrum, and polarization of both fields. In most cases the scattered power is simply proportional to the incident power. Second order effects such as harmonic generation would not enter, but absorption and fluorescence will alter the spectrum. With that viewpoint we can consider the scattered power relative to the incident power as a function of these parameters.

Based on the ratio of scattered to incident radiant power, Nicodemus [18] identified the distribution function

$$f(\theta_i, \phi_i, \theta_s, \phi_s) = \frac{B_s}{B_i \cos(\theta_i)\, \Omega_i} \tag{29}$$

as a factor in units (sr^{-1}) to characterize the directional scattering property of a surface. The incident radiance B_i $(W/m^2 - sr)$ gives the power that flows through a solid angle Ω_i

onto a surface area with its normal making an angle θ_i to the incident ray. The same surface area scatters light with radiance B_s into solid angle Ω_s. With this definition, the ratio of scattered to incident powers is

$$\frac{P_s}{P_i} = f(\theta_i, \phi_i, \theta_s, \phi_s)\,\cos(\theta_s)\,\Omega_s \tag{30}$$

for an surface area so small that the BRDF and Ω_s factors do not vary across it significantly. Because of conservation of energy expressed in the Helmholtz reciprocity theorem ([18], [12] page 423)

$$f(\theta_i, \phi_i, \theta_s, \phi_s) = f(\theta_s, \phi_s, \theta_i, \phi_i) \tag{31}$$

and the incident and scattered directions in the BRDF are interchangeable. This means that if the sensor and the light source are exchanged, the sensor will measure the same signal.

References

[1] H. Tüg, N.M. White, and G. W. Lockwood. "Absolute energy distributions of α-Lyrae and 109 Virginis from 3295 Å to 9040 Å". In: *Astronomy and Astrophysics* 61 (1977), pp. 679–684.

[2] H. Tüg, N.M. White, and G. W. Lockwood. "Absolute energy distributions of α-Lyrae and 109 Virginis from 3295 Å to 9040 Å, erratum". In: *Astronomy and Astrophysics* 66 (1978), p. 469.

[3] LSST. *NEO FAQ*. 2016. URL: https://www.lsst.org/science/solar-system/neo_faq.

[4] Hamamatsu Photonics K.K. Editorial Committee. *Opto-Semiconductor Handbook*. Japan: Hamamatsu Photonics K.K. Solid State Division, 2014.

[5] *CCD Primer*. 1999.

[6] Edmund Optics. *Imaging Electronics 101*. 2021. URL: https://www.edmundoptics.com/knowledge-center/application-notes/imaging/understanding-camera-sensors-for-machine-vision-applications/ (visited on 09/29/2021).

[7] *KAF-16803 Image Sensor Device Performance Specifications*. 2012.

[8] QHYCCD. *QHY 600 M*. 2021. URL: https://www.qhyccd.com/qhy600m-c/ (visited on 09/30/2021).

[9] Ivan K. Baldry. "Time-Series Spectroscopy of Pulsating Stars". PhD thesis. University of Sydney, 1999.

[10] R. L. Gilliland, T. M. Brown, and H. Kjeldsen. "A search for solar-like oscillations in the stars of M67 with CCD ensemble photometry". In: *Astronomical Journal* 106 (1993), pp. 2441–2476.

[11] F. Scherb. “Hydrogen production rates from ground-based Fabry-Perot observations of Comet Kohoutek”. In: *Astrophysical Journal* 243 (1981), pp. 644–650.

[12] Max Born and Emil Wolf. *Principles of Optics*. 7th ed. Cambridge: Cambridge University Press, 1999. ISBN: 0-521-642221.

[13] Lambda Research. *Oslo Optics Reference*. Little, Massachusetts: Lambda Research Corporation, 2005.

[14] Wikipedia. *Radiance*. Jan. 2013. URL: http://en.wikipedia.org/wiki/Radiance.

[15] Wikipedia. *Specific radiative intensity*. May 2013. URL: http://en.wikipedia.org/wiki/Specific_radiative_intensity.

[16] Sara Seager. *Exoplanet Atmospheres: Physical Processes*. Princeton, New Jersey: Princeton University Press, 2010.

[17] Subrahmanyan Chandrasekhar. *Radiative Transfer*. New York: Dover, 1960.

[18] Fred E. Nicodemus. “Directional reflectance and emissivity of an opaque surface”. In: *Applied Optics* 4 (1965), pp. 767–773.