
1 Basic Imaging

Starlight falls on every square meter of Earth. Yet it must fall on a telescope, the telescope must focus the light into an image, and a detector must respond to the image before an astronomer can use starlight to learn about the heavens. This chapter is about the detection of starlight, and this book is about the step that follows detection—image processing—converting data from the detector into a measurement or an image that is informative and even beautiful.

1.1 Light

Light is variously described as *rays* that travel in straight lines, as *waves* of electromagnetic energy, and as *particles* called photons. These descriptions are each valid in the sense that light exhibits ray-like, wave-like, and particle-like properties. None of these descriptions is complete by itself, yet each is a valid description of some aspects of the behavior of light.

The ray description treats light as a purely geometric phenomenon—it says nothing about the nature of light. The Greeks knew that light moves in straight lines even though they didn't understand its nature. The ray description is useful for talking about optical systems. The wave description elegantly unravels the phenomenon of diffraction, and the particle description offers insight into the nature of CCD operation. The three seemingly different descriptions only serve to illuminate different aspects of the behavior of light.

Light is electromagnetic radiation. Moving charges generate self-sustaining electric and magnetic fields that propagate away from the source as radio, infrared, visible light, ultraviolet, x-rays, and gamma-rays. Only the wavelength and frequency of the waves distinguish the light we see from forms of electromagnetic radiation we do not see. The relationship between the wavelength and frequency for electromagnetic radiation is:

$$\lambda \nu = c \qquad \text{(Equ. 1.1)}$$

where λ is the wavelength of the light, ν is the frequency, and c is the speed of light (2.99×10^8 meters per second). The wavelength characteristic of yellow-green light (to which the human eye is most sensitive) is 550 nanometers, corresponding to a frequency of 545 Terahertz (5.45×10^{14} Hz).

Chapter 1: Basic Imaging

Although light carries energy, nothing in classical physics prepared turn-of-the-century scientists for the discovery that the energy is quantized, behaving exactly like a stream of particles each carrying a specific energy. The energy carried by a single photon of wavelength λ is:

$$E_{\text{photon}} = \frac{hc}{\lambda} \text{ [electron-volts]} \quad \text{(Equ. 1.2)}$$

where c is the speed of light and h is Planck's constant; their product, hc , is 1240 electron-volt nanometers. Each photon (i.e., each "particle") of yellow-green light carries an energy of 2.25 electron volts. Long-wavelength photons have lower energy per photon; short-wavelength photons carry more energy. Ultra-high energy gamma-ray photons pack as much punch as a cruising mosquito, but 2.25 electron-volts is barely more than the energy required to kick free one electron in a crystal of silicon.

When we discuss how light travels through space, how it bounces off mirrors, how it bends as it passes through transparent materials, and how it forms images, we will use the ray and wave descriptions of light. When we examine the interaction of light with detectors, where the quantized nature of light becomes evident, we speak of light as photons.

1.2 Image Formation

Light alone is not enough. We need to know what direction the light that reaches us comes from—for that we have eyes. Eyes form images, sorting the flood of photons by direction. Imagine an amoeba under a clear sky, bathed in a flux of billions of photons. Although the amoeba can tell the total amount of energy falling on it, it cannot determine with any accuracy where the photons came from.

Images are the marriage of intensity with direction. Images are patterns of light intensity in which the amount of light on any point corresponds to the direction of origin of the light. Once light has been organized into an image, it can be detected by a retina, a silver-halide emulsion, or electronic camera.

1.2.1 Pinhole Imaging

The most basic method for finding the direction of light is shadowing. Even the amoeba can tell that the side of its body facing a strong light source is warmer (or more chemically stimulated) than the side facing away, and turn toward (or away from) the light. X-ray and gamma-ray astronomers once used detectors that were only slightly more sophisticated than the amoeba's simple shadowing.

However, the flood of light can be converted into an image with a simple pinhole camera. Any closed box or room (i.e., a *camera* in Latin) with a small aperture (*apertura* = "opening" in Latin) serves as an image-forming device. Rays of light from outside objects enter the aperture and continue in straight lines to the opposite side. On the image surface, light from each source thus has a well-defined location: the angular positions of the light sources have been mapped onto

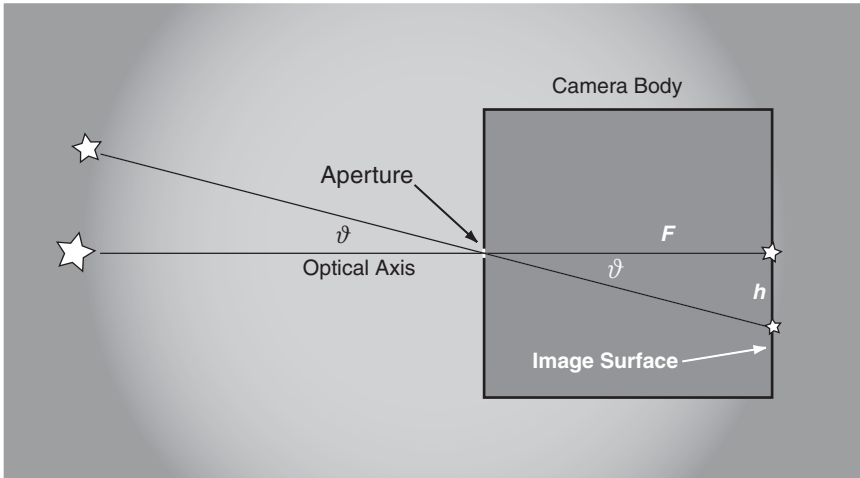


Figure 1.1 At the back of a pinhole camera, light sources at different angular positions in front of the camera are mapped onto locations in the image. The intensity of the light at each point in the image depends on the flux entering the camera from the light sources outside it.

(x, y) locations on a surface. The organized pattern of light intensities is an *image*.

The feature that distinguishes an image from illumination is the spatial organization of intensities in the image. The amoeba senses the total illumination without knowing where the light comes from, but in a *camera* the light is sorted into location by its direction of origin. As beings with eyes, we are so used to knowing the angular positions of light sources that it tends to be difficult for us to imagine light without knowing what direction it came from.

Let us now examine the properties of the images formed by a pinhole camera. A pinhole camera consists of a light-tight box. At the center of one end is the aperture and at the other end is a surface on which light falls. (Important to remember: pinhole cameras can be any shape and size, and the receiving surface need not be flat. The camera still works, but the math is more complicated.) Light from many sources falls on the front of the camera, and a tiny fraction of the light enters the aperture. Light that enters crosses the interior and falls on the receiving surface.

Inside a camera, we define a line between the aperture and the point that lies directly “under” the aperture as the *optical axis*. This point is the point on the receiving surface that is closest to the aperture—a distance called the *focal length*—and light reaching this point arrives perpendicular to the receiving surface. If we point the camera at a source of light—that is, align the optical axis of the camera with the direction of the source—then the image of the source lies on the receiving surface and on the optical axis.

Now consider the location of a source at an angular distance ϑ (theta) from the optical axis. The angle ϑ is measured at the aperture of the camera, and the

Chapter 1: Basic Imaging

image of this source forms some height, h , from the optical axis:

$$h = F \tan \vartheta \quad (\text{Equ. 1.3})$$

where F is the focal length. The greater the angular distance a source lies from the optical axis, the greater the linear distance its image lies from the optical axis.

Pinhole cameras can easily cover off-axis angles of 45° to 60° , so that the whole image spans a 90° to 120° angle. The rectilinear mapping of source position to image location has some interesting consequences. An array of sources that lie on a straight line in front of a camera will lie on a straight line on the image. In addition, the images of sources that lie on a great circle on the celestial sphere will lie along a straight line on the image. In terrestrial imaging, the sides of buildings are rendered as straight lines, and in astronomical imaging, horizons, equators, and longitude circles are rendered as straight lines. Images in the pinhole camera are said to be both rectilinear and free of distortion.

As imaging devices, pinhole cameras collect too few photons to be practical. If the aperture is enlarged to make the image brighter, light from different sources overlaps, and the image is blurred. If the aperture is too small, diffraction caused by the wave properties of light degrades the image. For practical imaging, light must be collected over a large area and focused into an image.

1.2.2 Lens Cameras

The aperture in a pinhole camera is simply a hole: it does nothing to the light that passes through it. Suppose that we place a carefully shaped disk of glass—a lens—in the aperture. Because light passes through glass more slowly than it passes through air, a wavefront encountering glass slows and changes its direction of propagation—that is, the ray of light bends when it passes from air to glass or from glass to air. By shaping the lens so that each ray is bent in direct proportion to its distance from the center of the glass, parallel rays of light from a source will converge until they cross. The rays cross at a point called the *focus*. After reaching the focus, the rays will continue on diverging paths unless they are intercepted at the focus by a viewing screen, a piece of photographic film, or an electronic detector, such as a CCD chip.

Rays that pass through the center of the lens are called principal rays. Principal rays pass through a lens and exit parallel to their original paths. Since the lens is shaped so that all rays from a given source cross the principal ray at the focus at the receiving surface, all of the rays from a source meet at the focus. A camera equipped with a lens collects all of the light from a star that falls on the lens and concentrates it into a single bright point.

Practical lenses almost always contain multiple lens elements made of different types of glass. As light passes through the multiple surfaces and glass types, the wavefront undergoes subtle manipulation. A great deal of art and effort goes into designing optical systems that bring light to an accurate focus.

However complex its internal design, a lens acts like an equivalent simple lens. The distance between the location of the equivalent simple lens and the focal

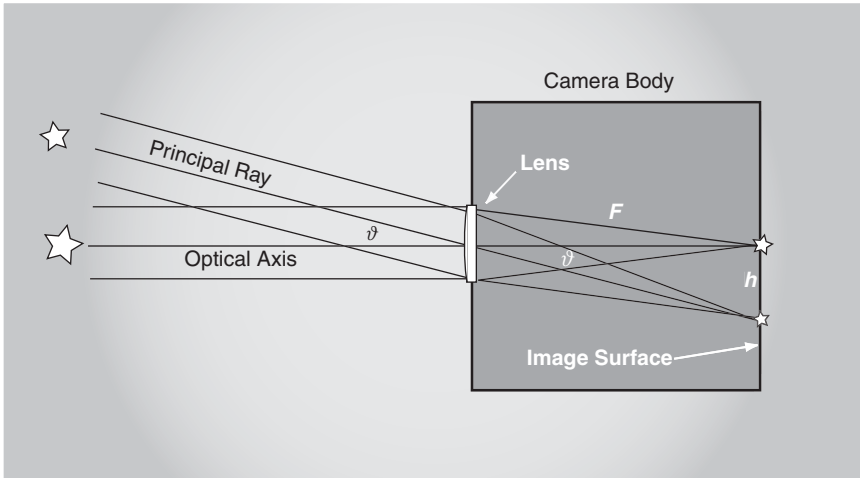


Figure 1.2 The lens of a camera directs rays from a large area to a focal point, resulting in a bright, sharp image. The image of each object falls in exactly the same place that it would in a pinhole camera because the lens directs all rays to converge on the principal rays, which remain parallel to their original paths.

plane is the *focal length* of the lens. Optical designers work hard to maintain a strict linear relationship between the focal length, F , the tangent of the angular distance from the optical axis, $\tan \vartheta$, and the ray height, h , from the optical axis:

$$h = F \tan \vartheta . \quad (\text{Equ. 1.4})$$

This is the same rectilinear relationship found in the pinhole camera. The lens directs the light from the sources in front of the camera to a tiny point of focus. Ideally, a camera equipped with a lens forms an image that is exactly the same as that formed by a pinhole camera, except that because the aperture of the lens is larger than that of a pinhole, the image is brighter.

Despite the best efforts of their designers, lenses exhibit aberrations, or departures from perfect imaging. Aberration means that the rays from a point source (such as a star) fail to focus at a common focal point. In spherical aberration, for example, rays at different distances from the principal ray converge ahead of or behind the focus point. Coma and astigmatism are aberrations that affect images away from the optical axis. Lenses suffer from chromatic aberration, in which rays of different wavelength fail to meet. Ordinary camera lenses perform well up to 20° to 30° from the optical axis, but at larger angles the aberrations degrade image quality.

Lenses focus much more light into an image than a pinhole. The ratio between the diameter of the bundle of rays entering the lens and the focal length is called the *focal ratio*. If a lens has a diameter of 100 millimeters and a focal length of 500 millimeters, its focal ratio is $f/5$.

The smaller the focal ratio, the greater the concentration of light onto the fo-

Chapter 1: Basic Imaging

cal surface. Typical pinhole cameras have focal ratios of $f/500$ (i.e., the aperture is $1/500$ of the focal length), but ordinary camera lenses have focal ratios as low as $f/2$. Because the light-gathering area of the $f/2$ lens is 62,500 times greater than the area of the small pinhole, the image formed by the lens is 62,500 times brighter than a pinhole image.

1.2.3 Telescopes

Telescopes are image-forming optical systems designed and optimized for astronomy. Because astronomical sources are faint, astronomers want a telescope with a large aperture; and because large lenses are difficult to manufacture, telescopes larger than about 300 millimeters aperture usually employ mirrors rather than lenses to gather and focus light.

As all amateur astronomers know, telescopes come in three varieties: refracting, reflecting, and catadioptric (combinations of reflecting and refracting). Quality refractors employ several different types of glass to correct chromatic aberration and bring all wavelengths to nearly the same focus. Reflectors employ mirrors which lend themselves to large apertures; the reflectors most common in astronomy are the Newtonian and the Cassegrainian configurations. Catadioptrics use mirrors to gather and focus the light, and lens elements to correct residual aberrations.

To a good approximation, telescopes obey the rectilinear relationship:

$$h = F \tan \vartheta , \quad (\text{Equ. 1.5})$$

and their images are therefore rectilinear. Because telescopes cover much smaller off-axis angles than camera lenses, seldom more than 1 or 2 degrees, this formula can be simplified. Over small angles $\vartheta \cong \tan \vartheta$, where ϑ is measured in radians; so for quick calculations with small angles, the formula becomes:

$$h \cong \frac{F \vartheta^\circ}{57.3} , \quad (\text{Equ. 1.6})$$

where ϑ° is the off-axis angle in degrees, F is the focal length of the telescope, and h is the off-axis distance in the focal plane.

In astronomy, high-quality images are essential. Telescopes are designed and manufactured to form diffraction-limited images; that is, images as nearly perfect as the laws of physics allow. To form a sharp image, light waves from a distant source must meet at the focus of a telescope in phase; that is, with the peaks and troughs of light waves from different parts of the aperture lined up. For this to happen, the total distance along every optical path must be the same to within a fraction of a wavelength of light, or about 100 nanometers (4 micro-inches). Although fraction-of-a-wave accuracy is a demanding criterion, do-it-yourself amateur telescope makers routinely attain this accuracy in their homebuilt telescopes.

Even in an optically perfect telescope, however, the light from a star cannot focus to a perfect point. As the waves converge toward focus, they arrive in almost perfect phase in a region surrounding the geometric point of focus; so instead of

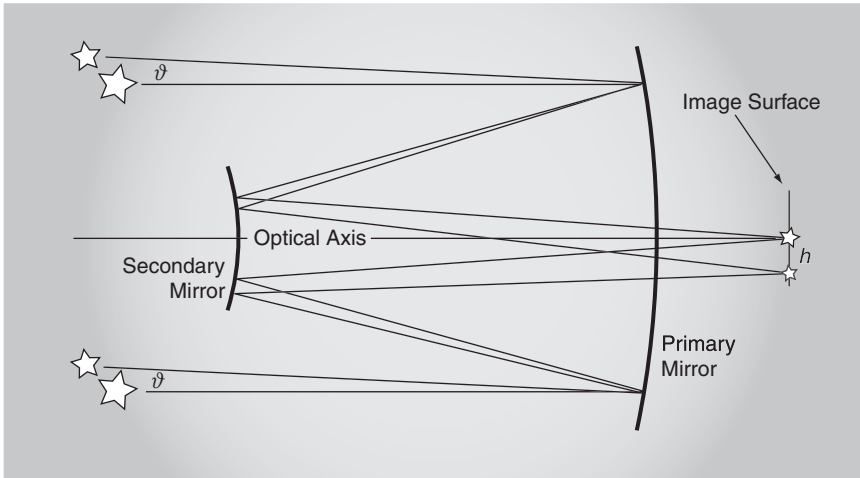


Figure 1.3 In a telescope—schematically shown here as a Cassegrainian—parallel rays from a source come together at a focus. Although principal rays would be blocked by the secondary mirror in a real telescope, their paths nonetheless define the focal length of the optical systems.

being concentrated in an infinitesimal point, light appears in a small spot called the Airy disk, or diffraction disk. Due to the wave nature of light, some of the light is deposited outside the Airy disk in a faint pattern of rings. In a perfect unobstructed optical system with a circular aperture, the Airy disk contains 84% of the incident energy, and the diffraction rings contain the remaining 16%. If the maximum path-length errors in a telescope can be made smaller than approximately 20% of the wavelength of the light, diffraction is the limiting factor in the quality of images, rendering improvement of the optical system beyond a fraction of a wavelength less and less practical.

The distribution of light at focus, whether a perfect diffraction figure or a degraded blur resulting from optical aberrations in the telescope, is called the point-spread function of the telescope. The point-spread function describes how the telescope forms an image of a perfect mathematical point. Because every plane wavefront that enters the telescope is transformed into the point-spread function at the focus, the image consists of countless overlapping copies of the point-spread functions. This is true for images of galaxies, nebulae, and planets as well as stars.

The angular diameter of the Airy disk sets an important constraint on what can be seen or imaged with a telescope, and its linear diameter at the focal plane is a key parameter in detecting images. The angular diameter of the Airy disk to the first dark diffraction ring, ϑ_{Airy} , is:

$$\vartheta_{\text{Airy}} = 2.44 \frac{\lambda}{A} \text{ [radians]}, \quad \text{(Equ. 1.7)}$$

where λ is the wavelength of the light, and A is the aperture in the same units. For

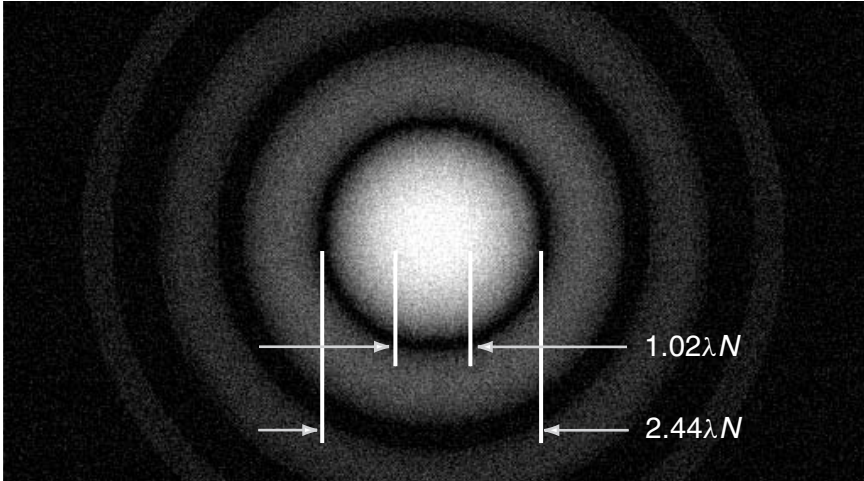


Figure 1.4 In a high-quality unobstructed telescope, 84% of the light from a star lies inside the first dark diffraction ring, and the remaining 16% of the light goes into the surrounding diffraction rings. However, a small, bright inner core of the diffraction disk contains nearly half of the star's light.

example, the angular diameter of the Airy disk at the focus of a telescope with an aperture of 200 millimeters (8 inches) at a wavelength of 656 nanometers is:

$$\vartheta_{\text{Airy}} = 2.44 \times \frac{656 \times 10^{-9}}{200 \times 10^{-3}} = 8 \times 10^{-6} \text{ [radians]}, \quad \text{(Equ. 1.8)}$$

which corresponds to 1.6 seconds of arc. (Note that all measurements were converted to the same unit—meters. To convert from radians to degrees, multiply by 57.3; to convert from radians to minutes of arc, multiply by 3438; to convert to seconds of arc, multiply by 206,265.) The linear diameter of the Airy disk, d_{Airy} , out to the first dark diffraction ring, is:

$$d_{\text{Airy}} = 2.44 \lambda \frac{F}{A} = 2.44 \lambda N, \quad \text{(Equ. 1.9)}$$

where A is the aperture and F the focal length of the telescope, and N is its focal ratio, F/A . The first term in the formula above is simply the angular diameter times the focal length of telescope.

However, nearly half of the image-forming light is concentrated in the small, bright central core of the diffraction disk, a much smaller region defined by the diameter at which the light has fallen to half its central intensity. This region is the full-width at half-maximum (FWHM) of the diffraction disk, ϑ_{FWHM} . The angular diameter of the small, bright core of the point-spread function is:

$$\vartheta_{\text{FWHM}} = 1.02 \frac{\lambda}{A} \text{ [radians]}, \quad \text{(Equ. 1.10)}$$

and its linear diameter is:

$$d_{\text{FWHM}} = 1.02\lambda\frac{F}{A} = 1.02\lambda N, \quad (\text{Equ. 1.11})$$

where d_{FWHM} is the FWHM of a perfect star image.

Assuming that the telescope has an $f/10$ optical system, the FWHM of a perfect star image is:

$$d_{\text{FWHM}} = 1.02 \times 656 \times 10^{-9} \times 10 = 6.7 \times 10^{-6} \text{ meters}, \quad (\text{Equ. 1.12})$$

or 6.7 micrometers. The diameter of the FWHM of the diffraction disk is a realistic measure of the smallest detail contained in an astronomical image.

Note: The scientific unit for 10^{-6} meters is the *micrometer*. However, many engineers and technicians use a less formal unit, the *micron*. In this book, we use both the formal and the informal units as they are normally used by engineers and scientists.

1.3 Detectors

Detector is a technical term that refers to a device that generates a signal in response to a phenomenon such as light. Amateur astronomers generally use one of three types of detector: the retina in their eyes, a photographic emulsion, or an electronic sensor. In this section, we examine how each works.

1.3.1 The Human Retina

The human eye combines an image-forming lens and a versatile detector into one compact unit. The detector in the eye is called the retina; it is an array of cells that detect light by the breakdown of the chemical rhodopsin. The breakdown products of rhodopsin trigger nerve responses that send an encoded signal to the brain through the optic nerve, where the signals are interpreted as a visual scene.

The lens in the eye has a focal length of about 16 millimeters, and a pupil that varies in aperture from 7 to 1.5 millimeters (yielding focal ratios from $f/2.3$ to $f/11$) to adjust the amount of light admitted to the eye. The retina is located on the curved interior of the eye, opposite the lens. Because the retina is fully integrated into the structure of the eye, it cannot be attached directly to a telescope, but must be used in conjunction with an *afocal* lens system consisting of an eyepiece and the objective lens of the eye itself.

Nevertheless, for comparison with photographic and CCD detectors, we shall examine the properties of the retina as a detector. Located on the curved interior of the eye, the retina, if it were flattened, would be approximately 40 millimeters in diameter. The retina contains some 100,000,000 light-sensing cells of two types: rods and cones. Rod cells cover the entire retina, and work well at low light levels. Cone cells are clustered near the optical axis of the lens, and operate best at high light levels. Individual cone cells are optimized to detect light at dif-

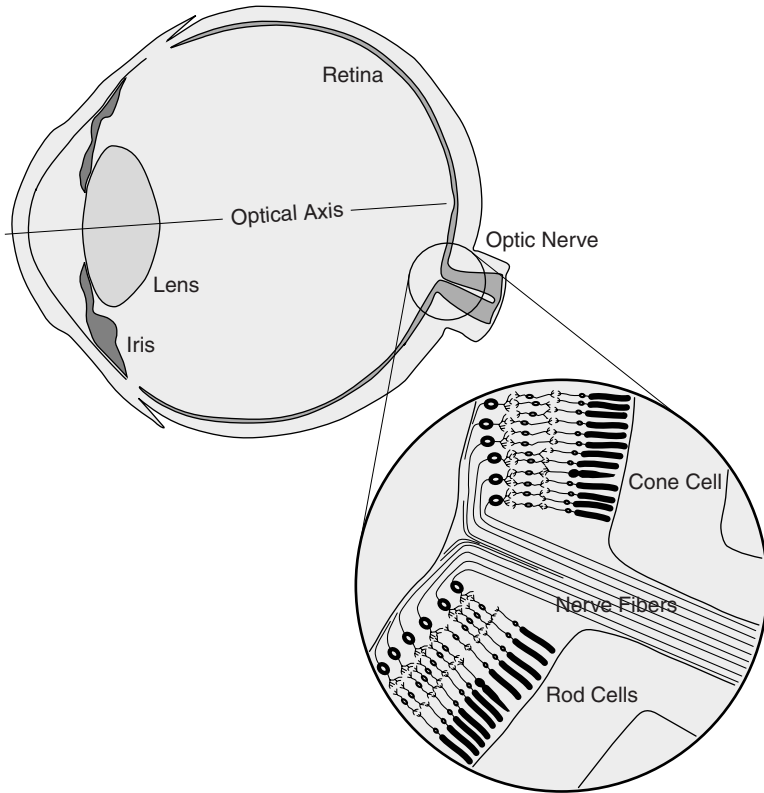


Figure 1.5 The eye is a high-performance imaging system containing a lens, detector array, and an on-board image processing system in the neural network overlying the rods and cones. The eye is so sensitive that under optimal conditions, an observer can see a flash of light consisting of ten photons.

ferent wavelengths to provide color vision.

Near the optical axis, the retina consists of cone cells—each about 2 microns in diameter—packed to a density of about 1,000 per square millimeter. The cones match the diffraction limit for the lens of the eye, providing an angular resolution of about 80 seconds of arc. Away from the optical axis, resolution is much lower. The light-sensitive parts of the rods and cones are nearly 100% efficient at absorbing photons, but because they are located on the back side of the retina, approximately half the incident light is lost passing through the overlying neural network. This layer of nerve cells mediates against noise by sensing light only when several adjacent cones or rods are simultaneously triggered. Because of these losses, the overall quantum efficiency of the retina is about 15% at the peak rod-cell sensitivity at 505 nm wavelength. The effective integration time of the retina is 100 to 200 milliseconds (consistent with its function in a moving animal), and the generation

of the signal is continuous.

In addition to filtering noise, the network of nerve cells preprocesses the signals generated by the light-sensing cells so as to detect edges, lines, and small differences in color. Thus, the signals that travel to the brain are not raw brightness data, but partially processed information on the shape, size, and color of objects in the visual field.

1.3.2 Photographic Emulsions

From 1880 to 1970, photographic emulsions were the primary detectors that professional astronomers used to record images. A photographic emulsion consists of silver halide crystals dispersed in a gelatin matrix, coated on a glass plate backing or a flexible plastic film backing. Although silver bromide is the primary silver salt, traces of chloride and iodide may be present. In the manufacture of photographic materials, the size and shape of the crystals is carefully controlled; the most sensitive crystals are approximately one micron across and flattened. In addition, the spectral sensitivity of the finished emulsion depends on dyes that absorb photons and transfer their energy to the silver bromide crystals.

Photographic emulsions differ from the retina and the CCD because exposure to light produces irreversible changes in the detector; in other words, you can use a film or plate once and once only. When a photon of energy greater than about 2 electron-volts enters a silver bromide crystal, it creates a defect in the crystal structure. Such defects can migrate through the crystal, and may spontaneously “heal” after several seconds or minutes. When several such defects have been created in a crystal, however, they coalesce into a stable defect called a development center. A stable defect involving only a few dozen atoms is enough to render the entire crystal—containing billions of silver atoms—chemically unstable.

When the emulsion is subsequently placed in a solution of photographic developer, the defect triggers the chemical reduction of the entire crystal to metallic silver. The capture of three to four photons can thus precipitate the formation of a grain of silver containing 10 billion silver atoms—a remarkable feat of chemical amplification. After development, the emulsion is washed in a fixer (a bath that dissolves undeveloped crystals of silver halide), then thoroughly rinsed and dried.

The developed photographic image thus consists of a clear backing coated with a gelatin layer containing microscopic grains of metallic silver where light struck the emulsion. Because the image appears dark or opaque where light fell, it is called a *negative*. By passing light through the negative to another sheet of photographic emulsion coated on white paper, you can make a negative of the original negative, which is called a positive print.

Photographic images have interesting properties. The process is inherently nonlinear with respect to its exposure to light, because once a grain crosses the development threshold, it is developable; and further exposure cannot make it more developable. Because of this, the number of developed grains is less than proportional to the number of photons. Furthermore, since a developed emulsion is three-dimensional, developed grains shadow other developed grains, resulting in a fur-

Chapter 1: Basic Imaging

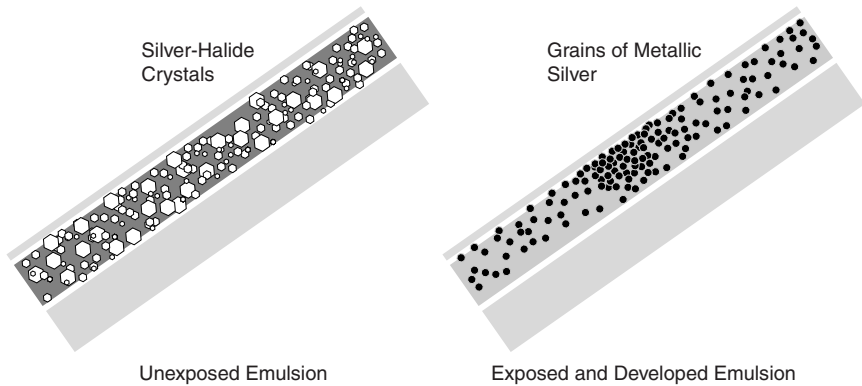


Figure 1.6 Photographic detectors are made by coating a thin sheet of plastic or glass with gelatin containing microscopic crystals of silver bromide, iodide, and chloride, protected by a top layer of plain gelatin. After exposure to light and chemical development, the image consists of tiny grains of metallic silver.

ther undercount of the original number of photons.

Photographic emulsions are also nonlinear with respect to exposure time. Unless three or four defects are created in a sufficiently short interval to create a development center, individual defects may disappear before coalescing; and the grain will not be developable. Photographic materials, therefore, record a smaller fraction of the incident photons in long exposures to low levels of light than they record when the same number of photons arrive in a short time. At the light levels found in astronomical images, faint parts of the image are recorded less efficiently than are bright parts of the same image.

Finally, because the silver halide grains in the emulsion were scattered randomly, the number of developed grains in small regions that have received the same exposure varies randomly around an average value. The random distribution of developed grains contributes to the “grainy” appearance of photographs.

During the decades when emulsion photography was king in astronomy, astronomers found numerous ways to enhance the performance of photographic emulsions. Bathing plates in a dilute solution of ammonium hydroxide a day or two before a night’s observing made the plate more sensitive. By exposing plates to faint light just before exposure in the telescope, astronomers found they could create one or two defects in each crystal, reducing the number of photons necessary to cross the development threshold. Baking plates and films for several hours before exposure likewise created defects, enhancing the performance of some emulsions. However, when researchers discovered that the primary reason that single-photon defects decayed was that water and oxygen were present in the emulsion, astronomers began baking films and plates in vacuums or gas mixtures containing hydrogen. This process—hydrogen hypersensitizing—drove out the

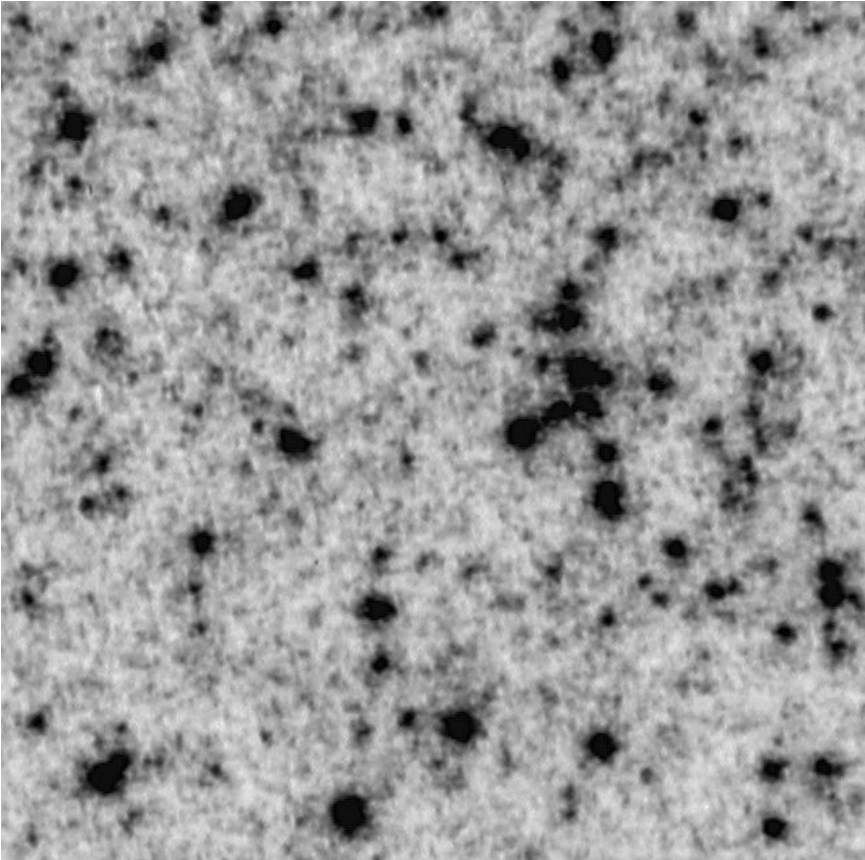


Figure 1.7 This highly enlarged section of an astronomical photograph shows the random distribution of silver grains in a developed image. Random variation in the number of developed grains in a given area causes the characteristic “grainy” appearance of photographic images.

water and reduced the oxygen, and produced a twenty-fold improvement in the fine-grain emulsion of Kodak 2415 Technical Pan. For serious amateur astrophotographers, hydrogen-hypered Tech Pan is the ultimate astronomical film.

Modern black-and-white films contain multiple layers of emulsions with differing crystal size to compensate for the limited dynamic range of simple emulsion coatings. Modern color films consist of multiple layers of emulsions with differing color sensitivity to separate and record the full gamut of color, and sophisticated chemistry to replace developed silver grains with clouds of colored dyes.

Overall, photographic emulsions serve as remarkably good image detectors. With hydrogen hypering, modern emulsions like Technical Pan perform well in exposures of several hours. Even without special treatment, modern black-and-white and color films are still efficient in exposures ranging from 5 to 20 minutes.

Although their overall quantum efficiency ranges from around 0.5% to about

Chapter 1: Basic Imaging

4%, with spectral sensitivity peaking in the blue and green end of the spectrum, photographic detectors are readily available in large sizes. Standard 35-mm film has a detector area 24 mm by 36 mm, standard 120-format roll film produces images 60 mm by 70 mm, and standard 4 × 5-inch sheet films have active areas of 100 mm by 125 mm. Although individual grains are only a few microns across, the smallest effective area capable of producing a good signal-to-noise ratio ranges from 5 to 20 microns. Assessed in the terms used for electronic sensors, a single frame of fine-grain 35-mm film offers the resolution of a 10 megapixel electronic detector.

Unfortunately, because photographic detectors are inherently nonlinear, and because a piece of film can be used only once, high-precision measurements of light intensity are not possible. With typical quantum efficiencies of 1%, photographic exposure times must be 20 to 60 times longer than comparable exposures with electronic sensors. For capturing physically large images, however, photographic emulsions are competitive with electronic sensors, especially in applications where the detector must be simple, compact, rugged, and inexpensive.

1.3.3 Electronic Detectors

Since their first astronomical use in 1976, electronic sensors have steadily gained ground as the detectors of choice in astronomy. In amateur astronomy, *charge-coupled devices* (CCDs) have been the dominant type. CCDs can detect light over a broad range of wavelengths, and they offer both high quantum efficiency and low noise. Furthermore, dark current and nonuniformities in sensitivity can be subtracted or divided out, thereby minimizing these shortcomings.

Challenging the CCD is another class of electronic sensor, the CMOS device. CMOS stands for *complementary metal-oxide semiconductor*, referring to the method of making them. While CMOS offers lower manufacturing costs, the resulting sensors are less sensitive and noisier than CCDs—a situation that will almost certainly change as the technology improves.

Today, electronic sensors incorporating CCDs and CMOS devices are used widely in amateur astronomy. Electronic cameras designed for astronomical work almost exclusively use CCDs, as do high-end digital cameras; but consumer-grade digital cameras and webcams increasingly rely on CMOS devices.

How CCDs Work. CCDs consist of an array of identical metal-oxide semiconductor capacitors formed on a silicon substrate. Each element in the array is called a photo-detector junction or photosite. The charge in each photosite is isolated from the others by a voltage applied through conductive channels on the surface of the silicon. At the beginning of an exposure, the capacitors are charged positively and then disconnected. As photons enter the silicon crystal lattice and are absorbed, they raise electrons from a low-energy valence-band state to a high-energy conduction-band state, partially discharging the capacitors. The degree of discharge of each capacitor is proportional to the number of photons that strike each element of the array during the exposure. At the end of the exposure, the electrons remaining in the photosites are sequentially shifted (or “clocked”) to a

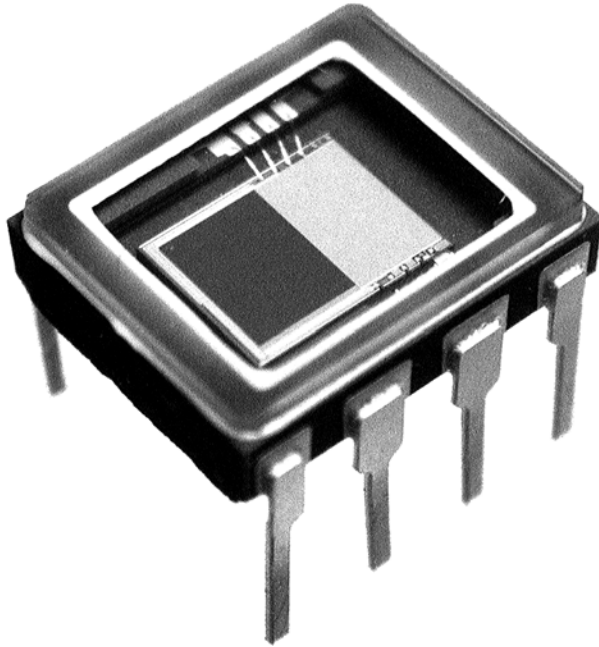


Figure 1.8 Electronic detectors have revolutionized astronomy. They are compact, highly sensitive to light, and—unlike a photographic emulsion—an electronic sensor can be used over and over again. Shown here is a Texas Instruments TC237, a 640x480-pixel CCD in a package just half-an-inch long.

charge-sensing node, amplified; and the signals are passed to external circuitry to be digitized and stored.

The detector characteristics of CCDs are the direct results of their construction. In silicon, the energy gap between the valence band and the conduction band is 1.1 electron volts, so that only photons with energies of more than 1.1 electron volts can boost an electron into the conduction band and be detected. This energy limit corresponds to a wavelength of 1100 nanometers, in the near infrared part of the spectrum. At shorter wavelengths, however, silicon becomes progressively more reflective so that the photons never enter the crystalline lattice, and hence cannot be absorbed. Silicon CCDs, therefore, reach peak quantum efficiencies of 40% to 90% between 500 and 950 nanometers wavelength.

In CCDs, the number of electrons boosted into the conduction band is directly proportional to the incident flux of photons. CCDs are highly linear as long as the total charge that collects in a photosite is too small to leak past the charge barriers that separate each photosite from its neighbors. In practice, most CCDs are linear as long as the photosites retain at least half their original charge.

Photosites on a CCD are arranged in columns and lines. The size of a photosite ranges from 6 to 25 microns depending on the CCD. The lower limit is set by manufacturing difficulties, and the practical problem is that small photosites have

Chapter 1: Basic Imaging

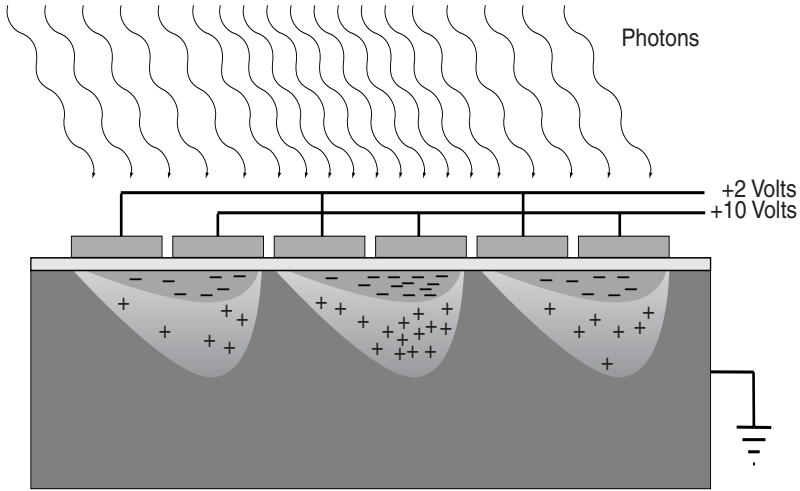


Figure 1.9 CCDs consist of adjacent charge wells, each collecting photo-electrons during integration. To shift the electrons down to a charge detection circuit on the CCD chip, the charge on the electrodes is cycled so that accumulated photo-electrons move from one charge well to the next.

small collection areas and therefore intercept few photons, and they can hold only a limited amount of charge. At the upper end, there is no need for photosites that are too large to capture all of the information present in images formed by camera lenses and telescopes.

CCD imagers range in size from 1.3 mm to approximately 70 mm across the diagonal, containing arrays of 102×102 to 8192×8192 photosites—and even bigger CCDs are being designed all the time. Unfortunately for amateur astronomers, the price of CCDs rises exponentially with the physical size of the chip. Typical CCDs in amateur astronomy measure 4 to 24 mm across the diagonal, contain between 32,000 and 12,580,000 photosites, and cost between \$60 and \$10,000+. Although many CCDs can be read out 60 times per second, astronomical CCD cameras are seldom read faster than 50,000 photosites per second to allow precise measurement and digitization of the signal. Readout times vary from a fraction of a second to about 60 seconds to read a complete image.

The electronics industry makes CCDs in a variety of configurations, from robust camcorder CCDs designed for readout at 60 fields per second, to digital camera CCDs designed to produce sharp megapixel images, to science-grade sensors optimized for high quantum efficiency and low noise. The configurations encountered most often in amateur astronomy are the progressive scan, frame transfer, and interline transfer. In addition, CCDs can be either front-illuminated (as most CCDs are) or back illuminated (as are high-performance scientific CCDs).

Progressive-scan CCDs contain a rectangular array of photosites with a spe-

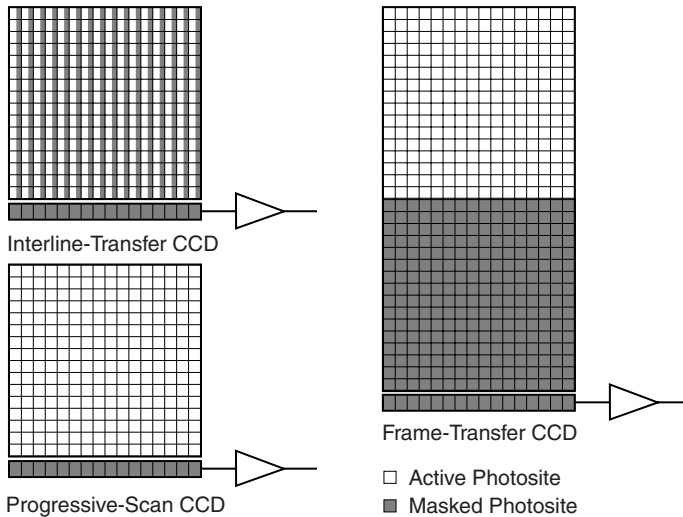


Figure 1.10 CCDs used in astronomy are almost always one of three types: interline-transfer, frame-transfer, and progressive-scan, shown schematically above. In all three types of CCDs, charge is shifted line by line from the array to a serial register, and then shifted to a charge-sensitive amplifier.

cial row of high-capacity photosites called the serial register at the bottom line in the array. After exposure, the electrons in all of the columns in the array are “clocked” down one line, and the bottom line enters the serial register. The serial register is then clocked one element at a time into the charge detection node. When the serial register has been emptied, all columns are clocked again to refill the serial register. The sequence is repeated until all charges on the array have been read and digitized. Progressive-scan CCD cameras must have a shutter to prevent light from generating new photoelectrons during the readout phase.

Frame-transfer CCDs have the same basic architecture, except that the lower half of the array is covered by an opaque mask. After the exposure, all of the columns are clocked rapidly to move the accumulated charge in the top half of the array to its bottom half. The bottom of the array is then clocked and read out just like a progressive-scan CCD. Frame-transfer CCDs do away with the need for a mechanical shutter, since charge in the upper half of the array can be transferred to the lower half in about 1 millisecond.

Interline-transfer CCDs are designed for even faster shuttering than frame-transfer ones. In this configuration, alternate columns on the CCD are masked, and the electrodes overlying the CCD are arranged so that charge in the uncovered columns can be clocked into the covered columns in a microsecond. The covered columns are then read out slowly. The primary difficulty with interline transfer is that half the detection area of the CCD is covered with masked columns, but a new generation of interline CCDs is being made with tiny integral lenses that redirect

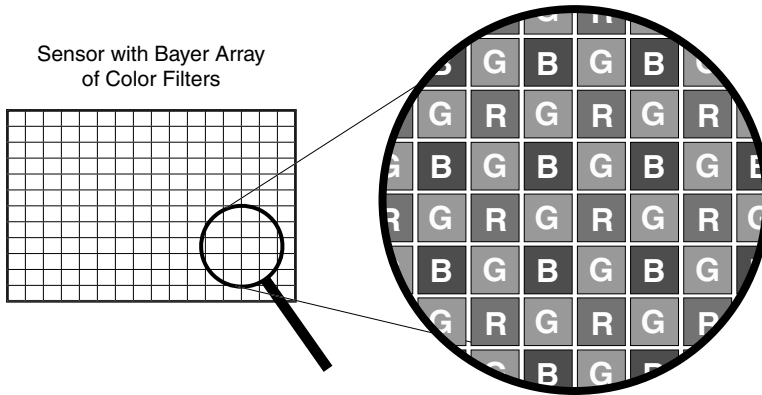


Figure 1.12 For color images, electronic sensors can be made with a checkerboard array of red, green, and blue filters. Each filter covers just one photosite, so the resulting raw image has a checkerboard pattern of pixels made with different filters. The final color image must be reconstructed in software.

light from the masked columns to the active sensing columns.

The vast bulk of CCDs are front-illuminated models, meaning that the photosites are formed on a “thick” silicon wafer, and the electrically conductive gates necessary for charging and clocking charge are laid on top. This construction means that to reach the light-sensitive bulk silicon, photons must penetrate the gate structures, which are sometimes nearly opaque to short-wavelength (blue) light. Because of this, front-illuminated CCDs seldom exceed a peak quantum efficiency of 60% at the long-wavelength (red) end of the spectrum.

Back-illuminated CCDs are made the same way as front-illuminated ones, but the silicon wafer is etched to a thickness of around 10 microns, and the silicon mounted so that photons enter the array from the back side. Thinned back-illuminated CCDs are expensive, but their quantum efficiencies are high across the spectrum from 500 to 950 nm, with peak values approaching 90%.

One of the greatest advantages that CCDs enjoy is that the same detector is used again and again, and the output is highly repeatable. This means that the odd-ball characteristics of a particular CCD can be calibrated out. An image from a CCD contains the following components:

- a bias voltage that is constant,
- systematic variations in the bias voltage,
- random variations in the bias voltage,
- a temperature-dependent dark current,
- systematic variation in the dark current,
- random variations in the dark current,
- a random variation (readout noise) in the output amplifier, and
- a signal generated by photons falling on the CCD.

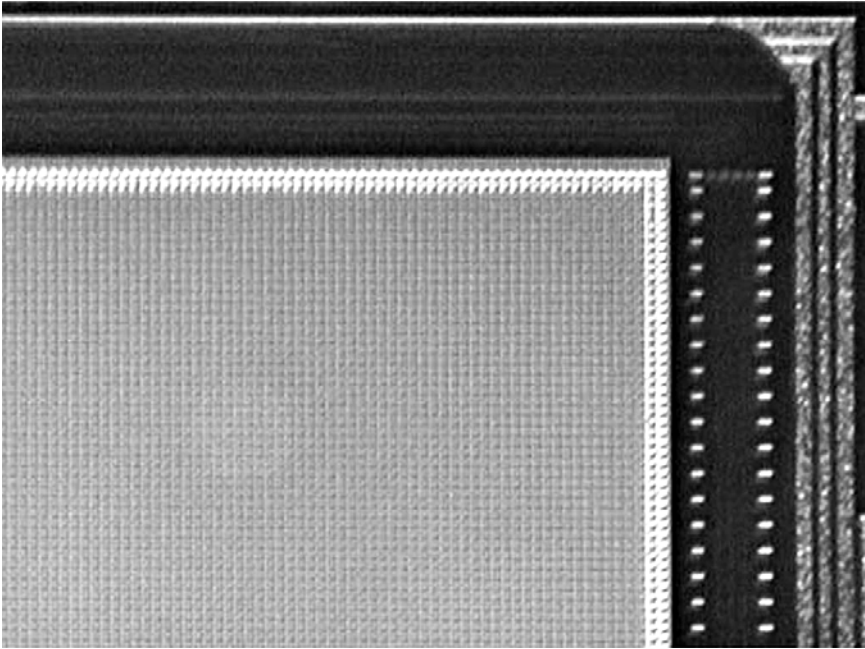


Figure 1.11 In this highly enlarged view of one corner of a CMOS sensor, you can see the regular array of individual pixels. In a CCD, accumulated charge from the whole chip is transferred to a single amplifier; in a CMOS device, each pixel has its own individually addressable amplifier.

In addition, photosites vary from one to the next in quantum efficiency, so the sensitivity of the array is not constant. However, because you can use the same CCD over and over, all of the constant and systematic effects can be removed. After calibration, a CCD image faithfully reproduces the amount of light that fell on each photosite in the array.

CMOS Devices. Like the CCD, the CMOS device consists of a large array of photosites on a silicon substrate. Unlike the CCD, however, the photosites in a CMOS device are individually addressable; that is, by activating a grid of conductors, any photosite can be read at any time. Instead of reading out the entire image, a small group of photosites can be read; or the CMOS sensor can be read out one line at a time while the remaining lines continue to accumulate signal.

Offsetting these advantages, however, each photosite must have an amplifier to sense charge and transistor switches to return the output signal. The auxiliary electronics take space that on a CCD would be collecting light, and hard-to-control variations in amplifier characteristics make CMOS devices less uniform than CCDs. Nevertheless, CMOS devices have proven their worth in webcams and a growing number of digital cameras.

From a practical point of view, once an electronic sensor has passed an image to the computer, the image data from CCDs and CMOS devices look very

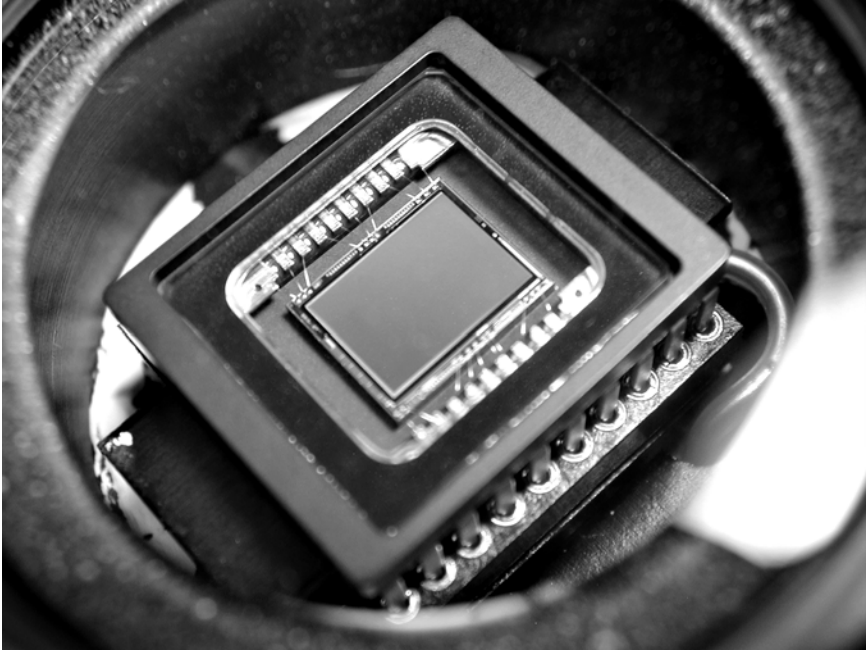


Figure 1.13 In CCD cameras designed for astronomical imaging, the CCD is isolated in a hermetically sealed chamber behind a clear glass window. This allows the detector to be cooled, thereby reducing its dark current, while avoiding problems from water condensing and freezing on the detector.

nearly the same. To get the best possible results from either type of detector, it is necessary to make calibration frames and apply them to raw images.

Color Imaging with Electronic Sensors. Although they are intrinsically sensitive to light over a wide range of wavelengths, CCDs and CMOS devices are monochrome sensors; that is, they record total incident flux of photons with no color information. To obtain color, observers must use one of three methods:

1. Make separate exposures through color filters—usually red, green, and blue. Each of the filtered images records the photon flux in one band of wavelengths, or one color channel. Observers often back up a tricolor set of images with an unfiltered luminance image that records all three color channels. To construct a color image, the three separated color channels and luminance must be merged into a color image.
2. Make a single exposure using a CCD or CMOS device with an integral color filter matrix, usually called a Bayer array. The filter array is a tiny checkerboard of red, green, and blue filters, each large enough to cover just one photosite. Thus, one exposure records information for all three color channels, at the expense of reduced image sharpness. To construct a color image, the image data must be resampled to provide every pixel with all three color channels.

3. Make a single exposure using a special CMOS device that has three sensing layers. The top layer responds to blue, the middle layer to green, and the bottom layer to red. At the time of this writing, multi-layer sensors are a new technology and not yet available for use in astronomy.

Making separate filtered exposures is by far the most flexible technique because it allows the observer to select a set of filters suited to the imaging task at hand. A sensor with a Bayer array is, however, the easiest way to make color images because “generic” red, green, and blue filters are built into the detector.

1.3.4 Linearity, Saturation, and Blooming

Astronomers prize CCDs for their linear response to light, that is, that the output signal is directly proportional to the number of photons that fell on each photosite during an exposure. When the number of photoelectrons accumulating during an exposure reaches the holding capacity of the charge well, the photosite is said to be saturated. Ideally, when a CCD reached saturation, it would cease responding to further photons. Unfortunately, that does not happen.

Bright stars in the CCD field of view continue to generate photoelectrons, and eventually the excess electrons overflow the charge wells and spill into adjacent photosites. On the finished image, a bright star displays streaks called *blooming trails* extending from the star image. Science-grade CCDs are linear, and as a consequence, most of them suffer from blooming trails from bright stars.

In camcorders and digital cameras, blooming trails are unacceptable. To combat blooming, CCD manufacturers incorporate electronic *drains* called *anti-blooming gates* to absorb overflow electrons during heavy exposure. With an active anti-blooming gate, it is possible to make a multi-minute exposure of a brilliant star without blooming trails. Unfortunately, the anti-blooming gate drains away electrons even before a photosite reaches saturation. Instead of trending steadily upward as the exposure increases, the number of photoelectrons rolls off—giving the CCD a nonlinear response to light.

CMOS devices tend to be free of blooming trails because the photosites are relatively independent of one another, but the small and necessarily simple amplifiers in CMOS photosites are themselves nonlinear.

1.4 Sensor Geometry

Images formed by a pinhole camera, lens camera, or telescope consist of differing amounts of light organized by their angular positions relative to the optical axis of the camera. At any given location, the intensity of the image corresponds to the amount of light coming from some particular direction. Detectors break the image into thousands or millions of discrete areas, each represented by a pixel in the resulting digital image.

Pixel means picture element. Pixels are to digital images what photosites are to CCDs. A pixel may correspond to a photosite on the CCD, or if the chip is

Chapter 1: Basic Imaging

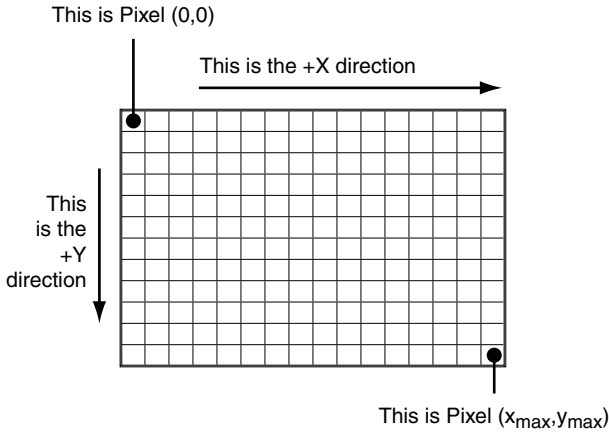


Figure 1.14 Each pixel in a digital image has an “address” corresponding its place in the image and the location of a photosite on the detector. The first pixel to be read from the detector is pixel (0,0) and the last is pixel (x_{\max}, y_{\max}) . The image shown here is 16 pixels wide by 12 pixels high.

binned, a pixel may correspond to two or more photosites on the CCD. Each pixel in an image has three key properties:

- its x -axis or column address,
- its y -axis or row address, and
- its pixel value.

The x -axis location is the pixel’s location in the coordinate that is clocked more rapidly from the CCD. By custom, the x -axis is displayed horizontally on computer screens. It may also be called the i -axis, or the *sample* axis.

The y -axis location is the pixel’s location in the coordinate that is clocked more slowly from the CCD. By custom, this axis runs vertically on computer screens. It is sometimes called the j -axis, or the *line* axis.

The numerical value of a pixel (its *pixel value*) is encoded in bytes stored on the computer’s hard disk. Pixel value is a property of pixels, just as mass is a property of matter. Pixel value can be expressed in different units of measurement, just as mass can be measured in grams or atomic mass units. In raw images, where the pixel has obtained its numerical value directly from the output of the CCD camera’s analog-to-digital converter, the units of pixel value are ADUs (analog-to-digital units) or DN (data number).

ADUs can later be converted into pixel values expressed in units of electrons, $\text{ergs}/\text{cm}^2/\text{sec.}$, or any other physical unit of measurement you want the pixels in your images to have.

In color images, the pixel value is usually expressed as numerical values of the three additive color primaries: red, green, and blue. These *RGB triads* may be encoded as 8-bit integers (256 gray levels), as 12-bit integers (4096 gray levels),

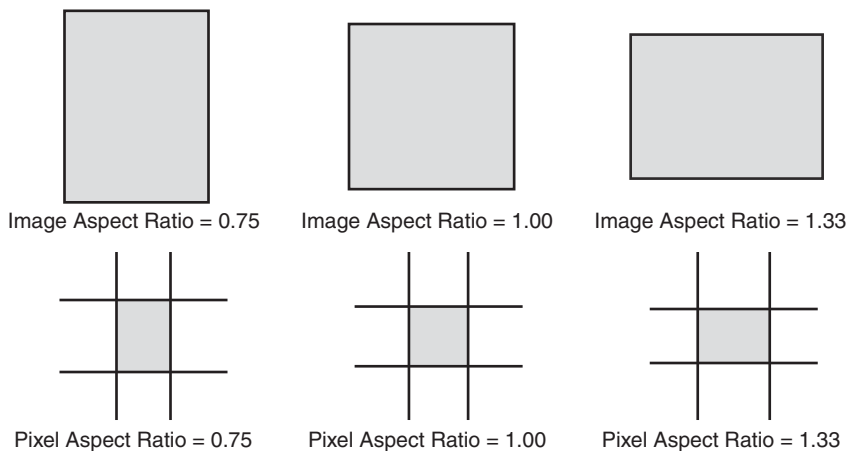


Figure 1.15 Aspect ratio is defined as width divided by height. For images, the width of the entire image is divided by its entire height. For pixels, the width of a single pixel is divided by its height. The wider the image or pixel is, relative to its height, the greater is its aspect ratio.

or 16-bit integers (65,536 gray levels), depending on the camera and the intended use for the image.

1.4.1 Aspect Ratio

To reconstruct the image sampled by the sensor, it is necessary to know how the mass of bytes is organized. This topic is treated fully in Chapter 3. This information is usually conveyed from the electronic camera to the computer in a file header. When a computer program opens an image, it begins by reading the header. The image can then be reconstructed on a computer monitor in the pattern of columns and rows matching that of the photosites. If this is done correctly, the computer screen will be a fairly faithful reproduction of the pattern of light that originally fell on the detector.

Aspect ratio defines an important characteristic of detectors: their shape. Electronic sensors have two types of aspect ratio: the *image aspect ratio* that defines the size of the entire image array on the sensor, and the *pixel aspect ratio*, which comes from the shape of the photosites.

The image aspect ratio, α_{image} , is the ratio of the image width to the image height:

$$\alpha_{\text{image}} = \frac{\text{image width}}{\text{image height}}. \quad (\text{Equ. 1.13})$$

Note that you must use the *physical* width and height of the image array, not the number of pixels, because the width and height of the pixel may differ. The same definition applies to photographic formats: the standard 35-mm film frame is 36 mm wide by 24 mm high, giving an image aspect ratio of 1.5.

Chapter 1: Basic Imaging

The pixel aspect ratio, α_{pixel} , is the ratio of the pixel width to the pixel height:

$$\alpha_{\text{pixel}} = \frac{\text{pixel width}}{\text{pixel height}}. \quad (\text{Equ. 1.14})$$

Note that you must use the physical width and height of the pixel, usually available in the manufacturer's specification sheet. On Kodak KAF400 CCDs, for example, two 9-micron-wide photosites are often binned to form a pixel 18 microns wide, and two 9-micron-high photosites are binned to form an 18-micron pixel, so the pixel aspect ratio is 1.00. On the Texas Instruments TC245, however, two 8.5-micron-wide photosites are binned to form a pixel 17 microns wide, but the 19.75-micron-high photosites are left unbinned, yielding a pixel aspect ratio, for the binned composite pixel, of 0.8608.

1.4.2 Pixel Count

Given the pixel count and the physical width of the detector, you can find the pixel size by dividing the detector width by the pixel count:

$$d_{\text{pixel}} = \frac{d_{\text{CCD}}}{N}. \quad (\text{Equ. 1.15})$$

To find the size of the photosites in a webcam with a 4×3 mm detector listed as having 640×480 pixels, divide 4 mm by 640 pixels:

$$d_{\text{pixel}} = \frac{4}{640} = 0.00625 \text{ [millimeter]}, \quad (\text{Equ. 1.16})$$

or 6.25 microns. The same holds true for the pixel height, and in this case, the pixels are 6.25 microns square. Given the detector size and pixel size, you can find the pixel count. For the TC245 CCD, Texas Instruments gives an image width of 6.4 mm and an image height of 4.8 mm. With binned pixels 17 microns wide by 19.75 high, the pixel count is 377 pixels wide by 243 pixels high.

Most science-grade CCDs are made with pixels having the same width and height, and they often have pixel counts that are some power of two: 256×256 , 512×512 , 1024×1024 , or 2048×2048 . These devices have an image aspect ratio of 1.000 and a pixel aspect ratio of 1.000. (When the pixel aspect ratio is 1, the device is said to have *square pixels*.) These dimensions make geometric computations easier, and allow easy use of a spatial-filtering technique called the Fast Fourier Transform.

Older CCDs manufactured for use in video cameras, such as the TC241 and TC245, usually have an aspect ratio of 4:3, or 1.3333—an image aspect ratio derived from the original 24-mm wide by 18-mm high Kodak/Edison 35-mm movie film frame. When television appeared, it retained the 4:3 aspect ratio, so that CCDs made for use in television cameras also have that aspect ratio. The pixel dimensions are based on NTSC standard interlaced video framing.

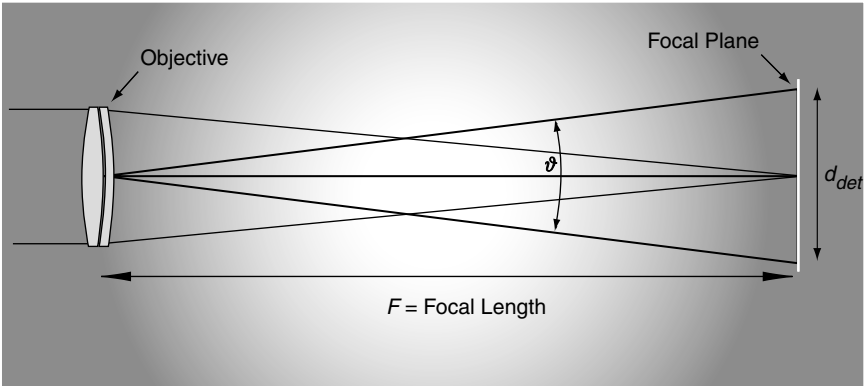


Figure 1.16 Whether it's a piece of photographic film or a CCD chip, a detector having a dimension d_{det} on a telescope of focal length F covers an angular field of view ϑ . To cover a larger field of view, either the detector must be made larger or the focal length of the telescope must be made shorter.

Many of the current generation of CCD and CMOS-based video cameras and webcams employ the 4:3 image aspect ratio with square pixels, so they are compatible with VGA and XGA computer displays. Pixel counts tend to fall on or close to the 4:3 ratio “magic numbers” of 640×480 , 1024×768 , 1600×1200 , and 2048×1536 . The Sony CCDs used in the Starlight XPress CCD cameras also have a 4:3 image aspect ratio, but some have pixel aspect ratios that depart slightly from square.

CCDs made for technical digital video and digital cameras, such as the Kodak KAF-series used in Santa Barbara Instruments CCD cameras, use an image aspect ratio derived from the 3:2 image aspect ratio of the standard 35-mm film frame, but with square pixels and power-of-two pixel counts (768×512 , 1536×1024 , and 3072×2048) in image height.

1.5 Image Capture

The most important characteristics of an astronomical image are its angular field of view and its angular pixel size. The field of view determines whether objects that you wish to image will fit inside a single image, and the resolution determines whether the photosites on the detector are the correct size to capture all of the desired image detail.

1.5.1 Angular Field of View of a Detector

The detector occupies an angle at the focus of a lens or telescope called the angular field of view. Recall that the image of an object is formed at a height h from the optical axis:

$$h = F \tan \vartheta, \quad (\text{Equ. 1.17})$$

Chapter 1: Basic Imaging

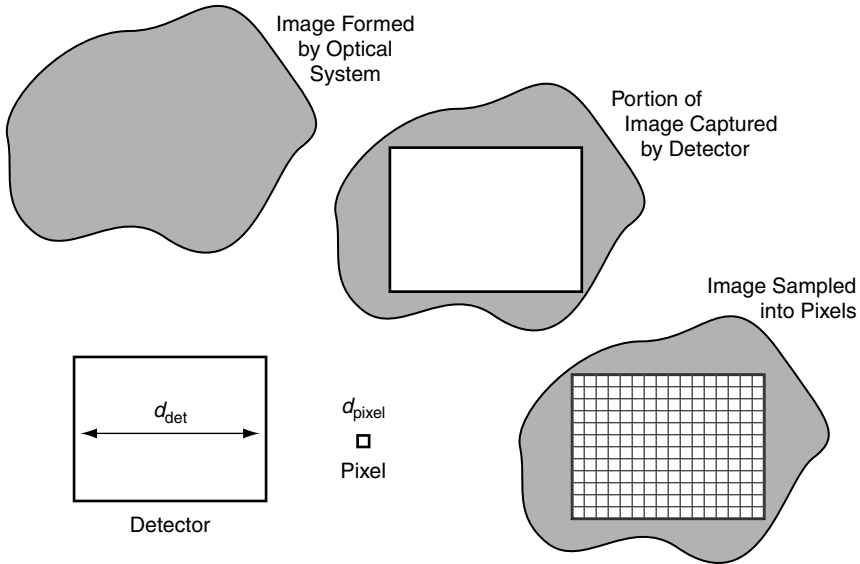


Figure 1.17 Electronic sensors capture the portion of the image that falls on them. If the image formed by a telescope is larger than the telescope image, then part of the image will be lost. During image capture, the photosite array on the sensor dissects the image into picture elements ("pixels").

where F is the focal length of the telescope and ϑ is the angular distance from the optical axis. Assuming that the center of the detector, with a dimension d_{det} , is placed on the optical axis of the telescope, then $d_{\text{det}} = 2h$, and a detector captures a field of view ϑ_{fov} :

$$\vartheta_{\text{fov}} = 2 \arctan\left(\frac{d_{\text{det}}}{2F}\right) \text{ [radians]} \quad \text{(Equ. 1.18)}$$

for a telescope with a focal length F . (Note: to convert radian measure to degrees, multiply by 57.3; to minutes of arc, by 3438; and to seconds of arc, by 206,265.)

As an example, consider a CCD detector that measures 6.4 mm wide by 4.8 mm high placed on a telescope with a focal length of 1,000 mm. What is the field of view of the detector on this telescope?

Perform the calculation for the 6.4-mm width as follows:

$$\vartheta_{\text{fov}} = 2 \arctan\left(\frac{6.4}{2 \times 1000}\right) \text{ [radians]}. \quad \text{(Equ. 1.19)}$$

$$\vartheta_{\text{fov}} = 2 \arctan(0.0032)$$

$$\vartheta_{\text{fov}} = 0.36669^\circ = 22'0''$$

For those unfamiliar with trigonometric functions, most scientific calculators will

perform the inverse tangent function and return the answer in degrees.

For the 4.8-mm height, perform the calculation as follows:

$$\vartheta_{\text{fov}} = 2\arctan\left(\frac{4.8}{2 \times 1000}\right) \quad [\text{radians}]. \quad (\text{Equ. 1.20})$$

$$\vartheta_{\text{fov}} = 2\arctan(0.0024)$$

$$\vartheta_{\text{fov}} = 0.27502^\circ = 16'30''$$

For the CCD detector and telescope in the example, the field of view is 22'00" wide by 16'30" high. You can compare these dimensions with a star atlas to see whether objects that you want to record will fit in a single image. Note that these equations work equally well for photographic film and electronic detectors.

1.5.2 Sampling the Image

Detectors do not reproduce images *in toto*, but rather, they *sample* the image. This means that the image is broken into discrete small chunks. With electronic sensors these are called *picture elements*, or *pixels*. The pixel structure dictates the smallest features that will be visible in a digital image.

Electronic sensors sample images in a very simple way: photons that fall on a photosite are lumped together as a single pixel value in the image. The CCD thus samples the image at the focus of the telescope in a regular grid, with each element in the grid represented by a single numerical value.

The number of photons captured by a photosite is proportional to the flux of photons times the collecting area of the photosite, and the number of electrons generated is the product of the quantum efficiency and the number of photons. Given that n is the average number of electrons generated, because the photons arrive at random, the number of electrons generated by a photosite during any particular integration is $n \pm \sqrt{n}$, and the signal-to-noise ratio is \sqrt{n} .

The ability of CCDs to store charge during an exposure ranges from 30,000 to 500,000 electrons in the charge well; so the expected random variations are 173 and 707 electrons, respectively. In addition, reading the charge from the CCD adds 6 to 50 electrons of random noise. When we apply the photometric bit-depth equation, we find that a normally exposed CCD image contains from 7 to 8 bits of useful information. However, to distinguish n gray levels containing that useful information requires a 12-bit to 16-bit analog-to-digital converter.

1.5.3 Angular Size of a Single Pixel

The equations for calculating the angular size of a photosite on the CCD and the corresponding pixel in the image are the same as those in the preceding section; although, of course, the pixel angles are much smaller.

We continue the example above, noting that the photosites on the CCD are 17 microns wide by 19.75 microns high. Set up the equation as follows:

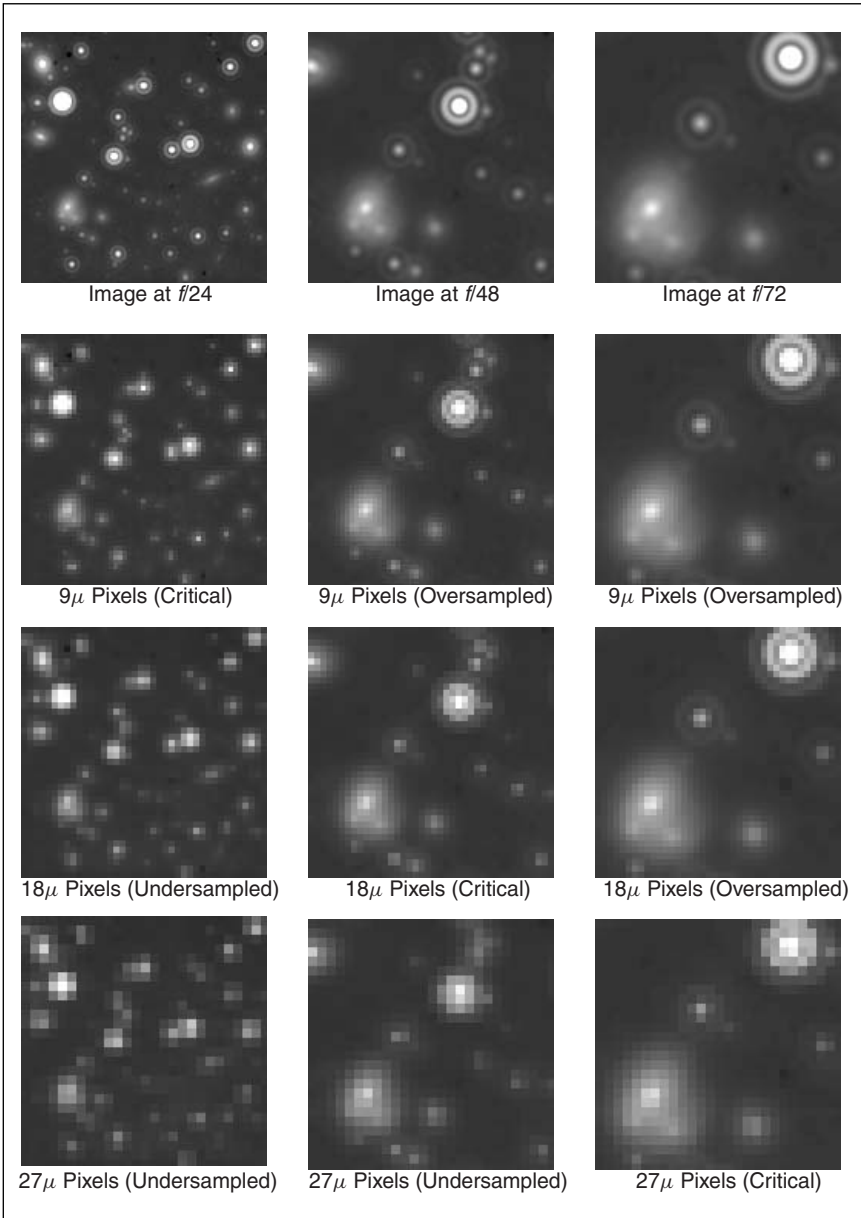


Figure 1.18 Matching pixel size and telescope focal ratio is essential for high-resolution imaging. The top row shows highly enlarged images at the focal plane of telescopes used at $f/24$, $f/48$, and $f/72$, with an effective wavelength of 730 nm. The succeeding rows show the same images as captured by detectors with 9-, 18-, and 27-micron pixels. At a focal ratio of $f/24$, 9-micron pixels satisfy the Nyquist sampling criterion, with two pixels across the Airy disk. With 18-micron pixels, the image is critically sampled at $f/48$, and with 27-micron pixels, critical sampling occurs at $f/72$. Images with finer sampling are called oversampled; those with coarser sampling are called undersampled.

$$\vartheta_{\text{fov}} = 2 \arctan\left(\frac{0.017}{2 \times 1000}\right) \quad [\text{radians}]. \quad (\text{Equ. 1.21})$$

$$\vartheta_{\text{fov}} = 0.000974^\circ = 3.5''$$

If you carry out the second calculation, you will discover that the pixels are 3.5 arc-seconds wide by 4.07 arc-seconds high.

1.5.4 Matching Pixels to the Point-Spread Function

The point-spread function of a telescope (ideally, the Airy disk) defines a characteristic dimension for the smallest details in a telescope image. To reproduce all of the detail present in the image, the sample size must be small enough to define the bright central core of the diffraction disk reliably. The Nyquist sampling theorem in communication theory states that in sampling a wave, the sampling frequency must be two times the highest frequency present in the original. Music recorded on CDs is therefore sampled at 44 kHz, a bit more than twice the highest frequency (20 kHz) most people can hear.

Applied to image sampling, the Nyquist theorem suggests that the size of a pixel must be half the diameter of the diffraction disk as defined by its full-width half-maximum dimension. Images sampled at this rate are called “critically sampled,” because the image has been broken into *just enough* pixels to capture all detail in the image.

Images sampled with pixels larger than half the full-width half-maximum of the diffraction disk are *undersampled*, because some of the fine structure in the telescope image will be lost. Undersampling is not necessarily a bad thing, since it may be a trade-off necessary to cover a large field of view.

Images sampled with more than two pixels across the core of the diffraction disk are *oversampled*. Oversampling with three to five pixels across the diameter of the full-width half-maximum diffraction disk insures that none of the information present in the continuous telescopic image is lost because the image is broken into discrete samples.

Atmospheric turbulence, telescope shake, poor guiding, and slightly out-of-focus images often enlarge the diffraction disk to many times the size of the Airy disk. A practical observer matches the pixel size not to the Airy disk, but to the size of the best seeing encountered at the observing site.

To match pixel size and diffraction disk, recall the formula for d_{FWHM} , the diameter of the core region of the diffraction disk:

$$d_{\text{FWHM}} = 1.02\lambda \frac{F}{A} = 1.02\lambda N, \quad (\text{Equ. 1.22})$$

where A is the aperture of the telescope, F is the telescope focal length, and N is its focal ratio. Because the real point-spread function d_{PSF} is always equal to or larger than the Airy disk of a perfect telescope, the following holds:

Chapter 1: Basic Imaging

$$d_{\text{PSF}} \geq d_{\text{FWHM}}. \quad (\text{Equ. 1.23})$$

Since the condition for critical sampling is that two pixels of dimension d_{pixel} must equal the linear dimension of the diffraction disk at the focus of the telescope,

$$2d_{\text{pixel}} = d_{\text{PSF}}. \quad (\text{Equ. 1.24})$$

This implies that the minimum focal length, F , required for critically sampling a telescopic image is:

$$F \leq \frac{Ad_{\text{pixel}}}{0.51\lambda}, \quad (\text{Equ. 1.25})$$

and that the focal ratio necessary for critically sampling a telescopic image is:

$$N \leq \frac{d_{\text{pixel}}}{0.51\lambda}. \quad (\text{Equ. 1.26})$$

To better understand the implications, consider an example: you have a telescope with an aperture of 200 mm, a focal length of 2,000 mm, a CCD camera with 9 micrometer photosites, and you want to take diffraction-limited images in red light at a wavelength of 630 nanometers. What is the optimum focal length?

Converting these units to meters and substituting into the above:

$$F \leq \frac{200 \times 10^{-3} \times 9 \times 10^{-6}}{0.51 \times 630 \times 10^{-9}}, \quad (\text{Equ. 1.27})$$

works out to $F \leq 5.6$ meters, or 5,600 millimeters. Under conditions of perfect seeing, using a 3x Barlow lens to raise the telescope's focal length of 2,000 mm to 6,000 mm would give you a slightly longer focal length than necessary for critical sampling. However, if a combination of seeing, drive errors, and telescope shake were to triple the effective diameter of the point-spread function, the images would be critically sampled at the $f/10$ focus of the telescope.

Here is another example: you want to take diffraction-limited images of Jupiter in green light (550 nanometers) with a camera that has a CCD with 12-micron pixels using the excellent optics of your 16-inch $f/6$ Newtonian. What is the best focal ratio to use?

Set up the solution by converting to meter units and substituting:

$$N \leq \frac{12(24) \times 10^{-6}}{0.51 \times 550 \times 10^{-9}}. \quad (\text{Equ. 1.28})$$

The result is $N \leq 43$ for critical sampling. To capture all the detail present in the image on a night of exquisite seeing, you will need to enlarge the image of the planet 7 times using eyepiece projection from the $f/6$ focus.

On nights of poor seeing, of course, you would use a lower focal ratio because the blurry image of a star would be much larger than a perfect diffraction disk. By dropping to $f/30$ or even $f/20$, you could use shorter integration times,

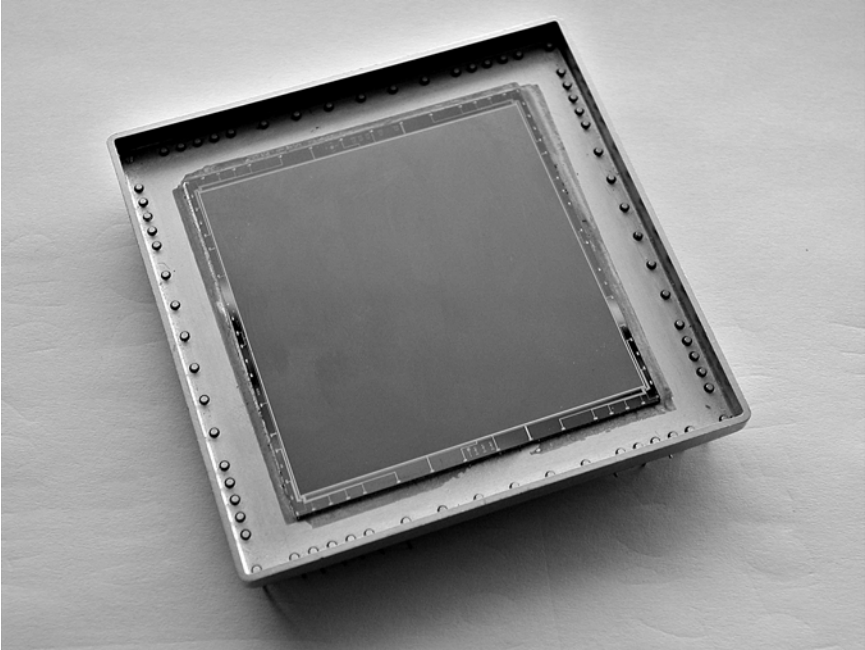


Figure 1.19 To match the focal length and seeing conditions at large modern telescopes, professional astronomers use CCDs that have large pixels and large pixel counts. This CCD is 50 millimeters square, and has 24 micron pixels. On an amateur-size telescope, it would grossly undersample the images.

thereby raising your chances of “freezing” moments of steady air. Today’s web-cam and video observers acquire hundreds, or even thousands, of images at a critical-sampling focal ratio; scan them to find moments of best seeing; and then combine those best moments into a diffraction-limited image.

Chapter 1: Basic Imaging