# Errata to Third Printing[1] of *The Handbook of Astronomical Image Processing, Second Edition*

1. **Section 2.1**, beginning after the sixth paragraph, has been revised to clarify the distinction between the case where we *know* the mean signal, and the case where we determine a mean signal from samples:

When counting events like photons, an astonishingly simple rule describes the expected variation:

- Given an mean signal of $\mu$ photons, an individual sample will contain $\mu \pm \sqrt{\mu}$ photons.

But what does the notation "$\mu \pm \sqrt{\mu}$" mean?

The "$\mu$" is the *mean value*. It is the average rate over a long time interval, the mean value of a large number of samples. The "$\sqrt{\mu}$" term is the *standard deviation*, a statistical measure of the departure of a typical sample from the mean value. Finally, the "plus or minus" symbol, "±", means that an individual sample may be either larger or smaller than the mean value.

The statement "$\mu \pm \sqrt{\mu}$" assumes that the distribution of sample values about the mean follows a bell-shaped curve called the *normal distribution*. In the normal distribution:

- 68.3% of samples will lie between $\mu - \sqrt{\mu}$ and $\mu + \sqrt{\mu}$,
- 95.4% of samples will lie between $\mu - 2\sqrt{\mu}$ and $\mu + 2\sqrt{\mu}$, and
- 99.7% of samples will lie between $\mu - 3\sqrt{\mu}$ and $\mu + 3\sqrt{\mu}$.

In other words, roughly two-thirds of the time, the measured sample value will differ less than one standard deviation from the mean value; and the other one-third of the time, the measured value will differ by more than one standard deviation from the mean value. However, departures greater than three times the

1.The printing of the book can be determined by looking at the book's copyright page.

standard deviation are rare, and departures greater than five standard deviations are *extremely* rare.

So—getting back to those ten pictures of the galaxy—what do you see?

Of course, you see the galaxy. In the pixels that comprise its image, you see pixel values scattered about the mean value that defines the galaxy. In the next section, we will dig deeper into the relationships between samples, signals, photons counts, and noise.

## 2.1 Signals and Noise

When we spoke about the mean value of the signal, $\mu$, we spoke as if we knew its value. In fact, we *do not know the value* of $\mu$. We can only *estimate* the value of $\mu$ by collecting samples of the signal. If we have taken one image, we have only one *sample* of the signal, which we will call $x_1$. If we take another image, we can call that sample $x_2$, and a third, $x_3$. If we take $n$ samples, the $n^{\text{th}}$ sample we take is called $x_n$.

To find the mean value of the signal, $\bar{x}$, we add the samples together and divide by the number of samples:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}. \qquad \textbf{(Equ. 2.1)}$$

The symbol $\bar{x}$ is read "x-bar," and the bar means "the average value of." It is important to remember that $\bar{x}$ does not equal $\mu$. As we sum more and more samples, our estimate of the mean signal, $\bar{x}$, will approach $\mu$, but a bit of uncertainty remains. How large *is* that uncertainty?

We can rewrite Equ. 2.1 for the mean value of a set of samples using summation notation:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \qquad \textbf{(Equ. 2.2)}$$

Although we don't know the *exact* value of $\bar{x}$, if we assume that $\bar{x} \approx \mu$, then assuming that $\mu \pm \sqrt{\mu}$, we expect the photon count of each sample be uncertain by $\sqrt{\bar{x}}$. However, because we have collected $n$ samples, we can lump together the photon counts from many samples (assuming, of course, that the signal does not change while we were collecting samples).

Given an accumulation of $n$ samples with a known mean signal of $\mu$, the total photon count in our "supersample" is $n\mu$, so we expect the photon count uncertainty of the supersample to be $n\mu \pm \sqrt{n\mu}$. Assuming again that $\bar{x} \approx \mu$, if we collect $n$ samples of a signal and compute that its mean is $\bar{x}$, then the uncertainty of the mean signal value $\bar{x}$ is $\sqrt{n\bar{x}}$.

It is important to distinguish between the case where we *know* the mean signal to be $\mu$, and the case where we measure a mean signal $\bar{x}$. In the first case, we know the signal and compute its uncertainty; in the second, we work backwards, averaging a set of samples to *estimate* the signal value and its uncertainty. The circumstances of CCD imaging are those of the second case: we take samples and estimate their mean.

This has a profound impact on the collection of astronomical data. If you take two “identical” images of the same object and compare them, you will find that they are *not* identical. They differ because each pixel captures independent samples of the mean photon rate, and each and every pixel in the two images has to follow the $\mu \pm \sqrt{\mu}$ photon-counting rule. To reduce the uncertainty (a.k.a., “noise”) in our images, we must use longer exposures or sum multiple image to accumulate a larger signal with a smaller uncertainty (i.e., less noise).

Suppose that instead of examining the same pixel in a sequence of images, we examine one image in an area of sky that has no stars. Although you might expect that all pixels making up that area of sky would have the same value, they do not. Just as a sequence of samples of one pixel varies, each pixel in a set of pixels that are side-by-side is an independent sample of sky brightness, and it will follow the same rule that a series of samples at the one location does. By averaging the sky signal from a many pixels, we can characterize the brightness of the sky background with less noise.

It may not be immediately obvious from the mathematical symbols that *the more photons in the signal, the better the signal quality*. Why should this be so? Consider the following example: we have two known signals, one consisting of 25 photons and the other consisting of 100 photons. In the 25-photon signal, the expected variation is $\sqrt{25}$, or 5 photons; while in the 100-photon signal, the expected variation is $\sqrt{100}$ or 10 photons. Since the variation in the 100-photon signal is twice that of the 25-photon signal, does that mean the 100-photon signal is worse? In one sense it is—it has twice the standard deviation. However, the percentage variation in the 25-photon signal is $5/25$, or 20%; while in the 100-photon signal, the percentage variation is $10/100$, or 10%. Even though it has twice the standard deviation, as a ratio of the signal strength, the 100-photon signal has only half the standard deviation as the 25-photon signal.

We quantify the signal quality using the *signal-to-noise ratio*, or *SNR*. This ratio is the mean signal divided by the noise. For a known signal in which uncertainty in the photon count is the primary source of noise, the signal-to-noise ratio for each sample is:

$$\mathrm{SNR} = \frac{\mu}{\sqrt{\mu}} = \sqrt{\mu}\,. \qquad \textbf{(Equ. 2.3)}$$

The SNR of the 25-photon signal above is 5, and the SNR of the 100-photon signal is 10. The greater the signal-to-noise ratio, the better the image quality. Note, however, that within a given image, the signal is not the same at all places. The sky background will have a lower signal level than the bright center of a galaxy, so it is meaningless to assign an SNR to an entire image because signal-to-noise ratio is meaningful for only one signal level. Nevertheless, astronomers sometimes do quote an SNR for an entire image, and when they do, it refers to the SNR at the signal level of the sky background.

2. We have revised **Sections 2.4.3 through 2.4.5** for greater clarity, and present them with more fully worked-out examples:

### 2.4.3 Signal and Noise in a Dark-Subtracted Image

We are finally ready to explore the signal and noise found in dark-subtracted images. Dark subtraction removes both the bias and the dark current pattern from the raw image, but adds noise from the dark frame to the noise already present in the raw frame. If this somehow seems a bit "unfair," remember that to remove the dark current pattern, we've been forced to count more electrons, which, by the inexorable statistics of event counting, leads to greater uncertainty.

The photon signal in the dark-subtracted image is:

$$S_{\text{image}} = S_{\text{raw}} - S_{\text{dark}} \qquad \text{[ADUs]} \qquad \textbf{(Equ. 2.18)}$$

$$= \left(\frac{x_{\text{pe}}}{g} + \frac{x_{\text{dark}}}{g} + b\right) - \left(\frac{x_{\text{dark}}}{g} + b\right)$$

$$= 364 - 124$$

$$= 240.$$

The noise in the dark-subtracted image is the sum of the noise contributions from the raw image and from the dark frame:

$$\sigma_{\text{image}} = \sqrt{\sigma_{\text{raw}}^2 + \sigma_{\text{dark}}^2} \text{ [ADUs] r.m.s.} \qquad \textbf{(Equ. 2.19)}$$

Recall Equ. 2.13 as it applies for noise in a raw image and noise in a dark frame:

$$\sigma_{\text{raw}} = \frac{1}{g}\sqrt{\sigma_{\text{pe}}^2 + \sigma_{\text{dark}}^2 + \sigma_{\text{ron}}^2} \text{ and } \sigma_{\text{dark}} = \frac{1}{g}\sqrt{\sigma_{\text{dark}}^2 + \sigma_{\text{ron}}^2} \text{ ADUs r.m.s.}$$

Expanding this to show all of the noise components, we see that:

$$\sigma_{\text{image}} = \sqrt{\left(\frac{1}{g}\sqrt{\sigma_{\text{pe}}^2 + \sigma_{\text{dark}}^2 + \sigma_{\text{ron}}^2}\right)^2 + \left(\frac{1}{g}\sqrt{\sigma_{\text{dark}}^2 + \sigma_{\text{ron}}^2}\right)^2} \text{ [ADUs] r.m.s.} \qquad \textbf{(Equ. 2.20)}$$

Evaluating this equation gives us the noise in the dark-subtracted image:

$$\sigma_{\text{image}} = \sqrt{\left(\frac{1}{2.5}\sqrt{\sqrt{600}^2 + \sqrt{60}^2 + 8^2}\right)^2 + \left(\frac{1}{2.5}\sqrt{\sqrt{60}^2 + 8^2}\right)^2} \text{ ADUs r.m.s.}$$

$$\sigma_{image} = \sqrt{\left(\frac{1}{2.5}\sqrt{600 + 60 + 64}\right)^2 + \left(\frac{1}{2.5}\sqrt{60 + 64}\right)^2} \text{ ADUs r.m.s.}$$

$$\sigma_{image} = \sqrt{10.7629^2 + 4.4542^2} = 11.65 \text{ ADUs r.m.s.}$$

It is clear that dark current noise and readout noise each contribute twice to the total noise in a dark-subtracted image.

The signal-to-noise ratio, $SNR_{image}$, given a sky background level of 240 ADUs to be:

$$SNR_{image} = \frac{S_{image}}{\sigma_{image}} = \frac{240}{11.65} = 20.6 . \qquad \textbf{(Equ. 2.21)}$$

Although an image with this signal-to-noise ratio would look somewhat grainy, faint deep-sky objects would be readily visible in it. Note that this is a *sky-limited image* because the overwhelmingly dominant source of noise comes from the photon statistics of the sky background, whose 600 photoelectrons contribute 9.8 ADUs worth of noise. For comparison, each of the two readout noise contributions of 8 electrons root-mean-square each add 3.2 ADUs of noise, and each of the two 60 electrons of dark current each add another 3.1 ADUs of noise. Added quadratically, these contributions total to 11.65 ADUs.

Under dark skies, however, dark current noise or readout noise can easily become the dominant noise source. At rural dark-sky sites, you might find a sky contribution of 40 electrons in a 60-second exposure. Photon statistics would then account for a mere 2.5 ADUs of noise, but the readout noise and dark current noise would still be 3.2 and 3.1 ADUs, respectively. Under these conditions, your images would be *detector limited* or *photon limited*.

How good is this image? Against a noise level of 11 ADUs, a single pixel 11 ADUs brighter than the mean level of the sky background—240 ADUs—would not stand out enough to be visible; but a star image consisting of a dozen pixels, or a faint galaxy image built from a few hundred pixels would appear brighter than the surrounding sky. A large, bright nebula would be clear and unmistakable.

In this example, subtracting the dark frame raised the noise level from 10.76 to 11.65 ADUs r.m.s.—and in general, subtracting a dark frame increases the random noise level. However, dark subtraction also removes hot pixels, dark-noise patterning, and the bias signal. As a result, dark subtracted images are cleaner and look better than raw images.

However, it is not necessary to accept even this modest loss. By making several dark frames and averaging them together, you can reduce the $\sigma_{dark}$ contribution to a negligible level. By averaging ten dark frames—see Equation 2.7—you could reduce $\sigma_{dark}$ to 1.40 ADUs. Subtracting this averaged dark frame (with 1.40 ADUs of noise) from the raw image (with 10.76 ADUs of noise) yields a $\sigma_{image}$ of 10.85 ADUs. If it's done properly, dark-frame subtraction causes a negligibly small loss in image quality—and it removes the hot-pixel noise pattern and bias.

If you do your dark frames right, you'll preserve every photon your camera captures!

### 2.4.4 The Effect of the Sky Background on Signal and Noise

In the preceding section, we have considered various noise sources and their impact on the signal-to-noise ratio. We hinted that the sky background has a significant impact on a camera's ability to detect extended astronomical objects such as galaxies and nebulae. In this section, we examine the role of the sky background on imaging performance.

In Section 2.4.1 we defined the count of photons detected, $x$. In astronomical observation, however, photons come from two sources: an object of interest, and parasitic illumination from the night sky:

$$x = x_{\text{object}} + x_{\text{sky}} \text{ [electrons].} \qquad \textbf{(Equ. 2.22)}$$

We now define a ratio between object photons and sky photons as:

$$\text{OSR} = \frac{x_{\text{object}}}{x_{\text{sky}}}, \qquad \textbf{(Equ. 2.23)}$$

in which OSR stands for *O*bject-to-*S*ky *R*atio. Sky brightness varies greatly from rural to suburban to urban sites, and also varies with the phase of the Moon. Furthermore, the range of object brightness is enormous. Suffice it to say that for very bright objects (the Moon and planets), the OSR is 1000 or more, and for exceedingly faint objects, it may be as low as 0.001.

Because the OSR varies over such a wide range, you may wish to determine the OSR of objects in your own images specific to your skies and local conditions. From a calibrated image, measure the following:

- $S_{\text{object}}$, the pixel value of the object in ADUs; and
- $S_{\text{sky}}$, the pixel value of the sky, in ADUs.

Remember that when you measure the pixel value of the object, it is the sum of the object and the sky. Because of this, the OSR must be calculated as follows:

$$\text{OSR} = \frac{S_{\text{object}} - S_{\text{sky}}}{S_{\text{sky}}}. \qquad \textbf{(Equ. 2.24)}$$

To illustrate the impact of sky brightness, we'll now consider the signal-to-noise ratio for three cases: a dark rural sky, a fairly good suburban sky, and a bright urban sky. Our celestial object is a galaxy that yields a photon flux of 250 electrons per pixel in 60 seconds. For this exercise, we'll use a camera with the gain, dark current, and readout noise from the preceding sections.

**Dark Rural Sky.** In this setting, we assume a signal of 40 electrons per pixel in 60 seconds, so for our celestial object, the OSR is 6.3; that is, the object is considerably brighter than the sky background. In a dark-subtracted image, a sky pixel measures 16 ± 6.8 ADUs and an object pixel measures 116 ± 9.3 ADUs. The object is 100 ADUs brighter than the surrounding sky, and it will stand out clearly against the noise in the dark rural sky background.

**Fairly Good Suburban Sky.** For our suburban sky, we take a signal of 600 electrons per pixel in 60 seconds. Our celestial object is now less bright than the sky background; its OSR for this sky is 0.42. In a dark-subtracted image, a sky pixel measures 240 ± 11.6 ADUs and an object pixel measures 340 ± 13.3 ADUs. The object is still 100 ADUs brighter than the surrounding sky, but both the sky and the object noise are greater than they were under the rural sky.

**Bright Urban Sky.** Urban skies are detrimental to good imaging because they are so bright. For the urban sky, we assume a signal of 4000 electrons per pixel in 60 seconds. The OSR of the object has fallen to 0.06. In a dark-subtracted image, a sky pixel measures 1600 ± 26.07 ADUs and an object pixel measures 1700 ± 26.83 ADUs. Our object is still 100 ADUs brighter than the surrounding sky, but the noise levels from the bright sky have become a significant fraction of the object's pixel value, and the image looks quite noisy. Nevertheless, even though the urban sky is 100 times brighter than the rural sky, the celestial object remains visible.

In the section that follows, we investigate the improvement obtained when you average multiple images. One seemingly paradoxical result is that it is necessary to accumulate more total exposure time under a bright sky than that required to achieve the same result under a dark sky.

### 2.4.5 Signal and Noise in Multiple Averaged Images

In preceding sections, you observed that by averaging several dark frames you could avoid adding noise. In this section, you will see that by shooting an ample number of dark frames and by averaging multiple raw images, you can cut noise and build image quality.

Begin by reviewing the signal and noise information you have on hand:

- You know $S_{\text{raw}}$, the raw-image signal in ADUs,
- you know $\sigma_{\text{raw}}$, the raw-image sky noise in ADUs,
- you know $S_{\text{dark}}$, the dark-frame signal in ADUs, and
- you know $\sigma_{\text{dark}}$, the dark-frame noise in ADUs.

Suppose that you shoot a number, $N_{\text{raw}}$, of raw images plus some other number, $N_{\text{dark}}$, of dark frames. What signal and noise levels do you expect? You average the raw images, you average the dark frames, and then you subtract them.

Here is the general equation for the expected signal:

$$S_{\text{combined}} = \frac{N_{\text{raw}} S_{\text{raw}}}{N_{\text{raw}}} - \frac{N_{\text{dark}} S_{\text{dark}}}{N_{\text{dark}}} = S_{\text{raw}} - S_{\text{dark}} \,. \qquad \textbf{(Equ. 2.25)}$$

When you average your raw images and dark frames, the resulting *signal level does not depend on the number of raw images and dark frames*.

Here is the general equation for the noise:

$$\sigma_{combined} = \sqrt{\frac{\sigma_{raw}^2}{N_{raw}} + \frac{\sigma_{dark}^2}{N_{dark}}}. \quad \textbf{(Equ. 2.26)}$$

The fact that the $N_{raw}$ and $N_{dark}$ are inside the square-root operation means that more raw images and more dark frames improve the image quality only as the square root—so getting a factor of two improvement requires four times more images. This difficulty notwithstanding, you're happy to shoot lots of images if it means that you'll have outstanding results.

Here is the signal and noise information for the suburban-sky example:

- $S_{raw}$ is 364.0 ADUs,
- $\sigma_{raw}$ is 10.76 ADUs,
- $S_{dark}$ is 124.0 ADUs, and
- $\sigma_{dark}$ is 4.45 ADUs.

The signal will always be:

$$S_{combined} = S_{raw} - S_{dark} = 240 \text{ [ADUs]}. \quad \textbf{(Equ. 2.27)}$$

The noise depends on the number of raw images and dark frames:

$$\sigma_{combined} = \sqrt{\frac{(10.76)^2}{N_{raw}} + \frac{(4.45)^2}{N_{dark}}} \text{ [ADUs]}. \quad \textbf{(Equ. 2.28)}$$

Suppose that you make 10 raw images and 10 dark frames—what signal-to-noise ratio can you expect? Evaluating the equation gives $\sigma_{combined} = 3.68$ ADUs, producing sky-level signal-to-noise ratio of 65—suggesting that can expect a combined image of excellent quality!

But—are 10 raw images and 10 dark frames the best use of precious telescope time? Intuitively, the answer is "no" because the noise contribution from the dark frame is so much smaller than that of the raw images. If you're willing to accept equal noise contributions from dark frames and raw images, then if you shoot some number of raw images, $N_{raw}$, you should shoot $N_{dark}$ dark frames:

$$N_{dark} = \left(\frac{\sigma_{dark}^2}{\sigma_{raw}^2}\right) N_{raw}. \quad \textbf{(Equ. 2.29)}$$

In the suburban skies example, if you shoot one dark frame for every six raw images, the noise contribution from the dark frames will be the same as the noise from the raw frames. With 17 raw images and 3 dark frames, Equ. 2.28 gives a total noise of $\sigma_{combined} = 3.66$ ADUs, for a signal-to-noise ratio of 65—essentially the same signal-to-noise ratio as shooting 10 raw images and 10 dark frames.

Which would be better: 10 raws and 10 darks or 17 raws and 3 darks? That depends on whether you must reduce the number of raw images to provide time to make dark frames. If you have a CCD camera with accurate temperature regulation, your best option would be to shoot 17 raw images and as many dark frames as you can manage during twilight—and preferably as many more than three as possible.

But, suppose that you have a non-temperature-regulated CCD camera and there is only enough time for you to shoot some total number of images (that is, the total of raw images plus dark frames), then the optimum ratio of dark frames to raw images is:

$$\frac{N_{\text{dark}}}{N_{\text{raw}}} = \frac{\sigma_{\text{dark}}}{\sigma_{\text{raw}}}. \qquad \textbf{(Equ. 2.30)}$$

For example, if the total number of images you can shoot is 20, then:

$$\frac{N_{\text{dark}}}{N_{\text{raw}}} = \frac{4.45}{10.76} = \frac{1}{2.4}, \qquad \textbf{(Equ. 2.31)}$$

or 1 dark frame for each 2.4 raw images. Given a total of 20 images, you should shoot 14 raw images and 6 dark frames, and you would achieve a signal-to-noise ratio of 70 for your nebula—a small but significant improvement over the signal-to-noise ratios for 65 from the 10 plus 10 or 17 plus 3 scenarios.

It should come as no surprise, however, that a darker sky will give you better results. For the dark-sky example, for which $S_{\text{combined}} = 16$ ADUs and $\sigma_{\text{raw}} = 2.5$ ADUs, your images are noise limited. You must make enough dark frames to subtract the dark current as accurately as possible; for dark skies, Equation 2.28 calls for 3.1 dark frames for each raw image. If you shoot 10 raw images and 10 dark frames, the result is $\sigma_{\text{combined}} = 1.98$ ADUs—considerably less noise than you could attain under suburban skies. Paradoxically, the signal-to-noise ratio has fallen to 8.1 at the sky-brightness level, but that's only because the sky is so very dark. To make a direct comparison, compute the signal-to-noise ratio for a nebula with a pixel value of 240 ADUs, and you will see that the signal-to-noise ratio rises to an outstanding 121.

If you take Equation 2.25 seriously for the dark rural sky and decide to shoot 10 raw images and 32 dark frames, you get $\sigma_{\text{combined}} = 1.11$ ADUs and a signal-to-noise ratio for the sky background of 14.4; but for a 240-ADU nebula, the signal-to-noise ratio is 216. Dark skies make a huge difference in image quality, especially if you shoot enough dark frames to make photon statistics the dominant source of noise.

3. **Equation 2.34** should read:

$$\sigma_{\text{flat}} = \left(\sqrt{\frac{(71.6)^2}{10} + \frac{(3.44^2)}{10}}\right) = 22.67\ [\text{ADUs}] \qquad \textbf{(Equ. 2.34)}$$

The resulting signal-to-noise ratio is 565.

4. In **Table 4.3**, the Full Well heading should read:

Full Well (thousands)

5. In **Section 6.2.1.3**, replace the final paragraph with:

There are three basic methods of combining multiple bias frames: averaging, taking the median, and computing the $k\sigma$ clipped mean. If you operate your CCD in an electrically noisy environment, individual bias frames sometimes show large noise spikes. Averaging bias frames makes these abnormal events part of the master bias frame. The median of the bias frames excludes the abnormal values, but does not reduce random noise as effectively. The clipped mean excludes only the deviant pixel values, and is nearly as effective as mean at reducing noise. If you operate in an electrically noisy environment or your CCD is prone to cosmic rays; use the $k\sigma$ clipped mean.

6. In Section 6.2.2.7, replace the second and third paragraph with:

If you average multiple dark frames to create a master dark frame, a single cosmic ray hit can mar all images calibrated with it. Rather than averaging, it is good practice to compute the $k\sigma$ clipped mean. The average value is raised by the strong signal from the cosmic ray trail, but the clipped mean rejects extreme pixel values such as those arising from a cosmic ray hit.

• ***Tip:*** ***AIP4Win*** *supports both dark-frame averaging and determining the dark-frame* $k\sigma$ *clipped mean. In any case, it is a good idea to inspect the individual frames for cosmic ray hits before combining them, and not include frames with obvious cosmic ray hits from a master dark frame.*

7. In **Section 6.2.3.3**, replace the last paragraph with:

Because the brightness of the sky changes, you cannot simply average sky flats. Instead, you must subtract an average dark frame from each image, measure the average brightness of a small region near the center of each frame, and then multiply the images so that all have the same average pixel value near the center. This process is called "normalizing" the images. Once the flats have been normalized, you can take the $k\sigma$ clipped mean pixel value (to eliminate the star images) from the set of dark frames as you create the master flat.

8. **Equations 7.5 and 7.6** should read:

$$R \geq |x - x_0| \qquad \textbf{(Equ. 7.5)}$$

$$R \geq |y - y_0| \,. \qquad \textbf{(Equ. 7.5)}$$

$$R \geq |x - x_0| > r \quad \text{(Equ. 7.6)}$$

$$R \geq |y - y_0| > r\,. \quad \text{(Equ. 7.6)}$$

9. In **Section 7.4.1.6**, the last paragraph should read:

In a set of normally distributed random numbers, 68% of them should lie between $\bar{P} - \sigma$ and $\bar{P} + \sigma$. The standard deviation is the characteristic width of a Poisson or Gaussian distribution; thus in a random sample, the standard deviation is a direct measure of what we loosely call “noise.”

10. In **Section 7.4.1.**9, the first paragraph should read:

The “mean of the median half” is a hybrid statistic designed to use information from many pixels while excluding widely disbursed values. The mean of the median half is the mean of the middle half of a set of sorted pixel values. In the sequence: 1, 2, **7, 8, 8, 8, 11**, 19, 99, the mean of the median half is 8.2. The bottom quartile and top quartile were rejected in forming this value.

11. Add **Section 7.4.1.10**:

### 7.4.1.10 The $k\sigma$ Clipped Mean

The $k\sigma$ clipped mean is another hybrid statistic intended to include information from as many pixels as possible—while excluding the most extreme values. To find the clipped mean, the algorithm computes the mean and standard deviation, $\sigma$, of a set of pixel values. Pixel values that departs by more than $k$ standard deviations from the mean are excluded, then a new mean and standard deviation are computed. Pixel values that depart by more than $\pm k\sigma$ are excluded again, and the process is repeated until all pixel values lie within $k\sigma$ of the mean.

In cases where pixel values are expected to vary main from photoelectron statistics and normally distributed noise sources, the $k\sigma$ clipped mean is extremely effective in removing deviant values with only minor loss to the signal-to-noise ratio. These cases include cosmic rays, airplane trails, satellite trails, guiding errors that affect a small number of images in a set of images. The value of $k$ is usually chosen as 3.0, which, in a normally distributed set of pixel values will exclude only about 1% of the values.

12. **Equation 7.10** should read:

$$r \leq \sqrt{(x_0 - x_i)^2 + ((\text{aspect ratio})(y_0 - y_i))^2} \quad \text{(Equ. 7.10)}$$

13. **Equation 7.13** should read:

$$d = \sqrt{(x_1 - x_2)^2 + ((\text{aspect ratio})(y_1 - y_2))^2} \quad \textbf{(Equ. 7.13)}$$

14. **Equation 9.2** should read:

$$Y = -\frac{\sin\delta_0 \cos\delta \cos(\alpha - \alpha_0) - \cos\delta_0 \sin\delta}{\cos\delta_0 \cos\delta \cos(\alpha - \alpha_0) + \sin\delta_0 \sin\delta} \quad \textbf{(Equ. 9.2)}$$

15. In **Section 9.2.2**, the paragraph following **Equ. 9.7** should read:

In real life, however, $(\alpha_0, \delta_0)$ is inevitably somewhat off the center of the image, and the $x$ and $y$ axes of the detector will be rotated through some angle (see Figure 9.2). However, these complications do not matter. When we apply the mathematical transform from standard coordinates to plate coordinates, the rotation, plate tilt, and arbitrary center are automatically compensated.

16. In Section 9.2, in the paragraph immediately above Equ. 9.9, "linear" should read "affine."

17. In Section 9.2.3, the fourth paragraph should read:

To find the minimum value of the sum of the squares of the residuals, the equations of condition are differentiated with respect to the unknowns, producing three so-called normal equations in which *a, b,* and *c*, and *d, e,* and *f* are the unknowns, and their coefficients are terms in *x* and *y*. Since the normal equations are linear, their solution is straightforward.

18. **Equation 11.2** should read:

$$\sin\iota - \sin\delta = mn\lambda \quad \textbf{(Equ. 11.2)}$$

19. In **Figure 11.6**, the two labels that read HII should read HI.

20. In **Section 12.2**, **Equations 12.3**, **12.4**, and **12.5** should read:

$$x' = x\cos\vartheta - y\sin\vartheta \quad \textbf{(Equ. 12.3)}$$

$$y' = x\sin\vartheta + y\cos\vartheta \quad \textbf{(Equ. 12.3)}$$

$$x' = x_0 + (x - x_0)\cos\vartheta - (y - y_0)\sin\vartheta \qquad \textbf{(Equ. 12.4)}$$

$$y' = y_0 + (x - x_0)\sin\vartheta + (y - y_0)\cos\vartheta\,. \qquad \textbf{(Equ. 12.4)}$$

$$x = x_0 + (x' - x_0)\cos\vartheta + (y' - y_0)\sin\vartheta \qquad \textbf{(Equ. 12.5)}$$

$$y = y_0 - (x' - x_0)\sin\vartheta + (y' - y_0)\cos\vartheta\,. \qquad \textbf{(Equ. 12.5)}$$

21. In Section 12.5, the terms “flipping” and “flopping” are hard to remember. The preferred term for “flipping” is “a top-to-bottom reflection” or “reflection in *y*,” and the preferred term for “flopping” is “a left-to-right reflection” or “reflection in *x*.”

22. In **Section 13.5**, the paragraph beginning, “In image processing...” should read as follows for greater clarity:

In image processing software, transfer functions can be precomputed and the results stored in an array called a *look-up table*. The reason for this is simple: speed. If you are going to determine a new brightness for every pixel in an image that is 16 bits deep by 1024 x 1024 pixels on a side, you would need to compute the transfer function 1,048,576 times. If you construct a look-up table for each of the 65,536 possible pixel values, you benefit because you can consult the look-up table 1,048,576 times because each of the lookups is nearly instantaneous.

23. In **Section 14.1**, the line of pseudocode describing the convolution algorithm should read:

```
new(x) = new(x) + k(i) * old(x + i)
```

24. In **Section 16.4.1**, the last paragraph should read:

$\Delta x$ and $\Delta y$ describe the translation of the slave image relative to the master, in units of pixels. Positive values of $\Delta x$ mean the reference point has moved to the left; negative values mean the reference point in the slave image is right of that in the master image. For registration, the slave image should be translated by $\Delta x$ and $\Delta y$. Section 12.1 describes the mechanics of image translation.

25. In **Section 16.4.2**, the paragraph following **Equ. 16.14** should read:

The final step is to translate, rotate, and scale the slave image to match the master, for which the translations are $\Delta x$ and $\Delta y$, the rotation is $\Delta\vartheta$, and the scaling factor is $1/s$. Algorithms for translation, rotation, and scaling are covered in Section 12.4.

26. **Equation 17.15** should read:

$$N(x, y) = \log(1 + |S(u, v)|^2) \quad \textbf{(Equ. 17.15)}$$

27. The beginning of **Section 18.1** is revised for greater clarity:

## 18.1 The Wavelet Transform

An efficient way to compute the wavelet transform is to use a function like the Mexican Hat transform to isolate each wavelet scale (i.e., each band of spatial frequencies) from the original image in turn. For astronomy, the *à trous* algorithm is particularly effective, and it has the added advantage of being computationally fast and efficient. The name of the algorithm—*à trous*—means "with gaps," since, as we shall see, at the higher wavelet scales, the wavelet convolution kernel has big holes in it.

The *à trous* wavelet transform builds wavelet scales by iteratively splitting the low-frequency and high-frequency components of the image. It begins by applying a blurring filter to the original image, removing the high-frequency components. The blurred image is subtracted from the original image to obtain the high-frequency component, that is, scale 1 wavelet. (Section 14.2.3.1 on low-pass filter kernels and Section 14.2.3.2 on high-pass filter kernels illustrate how a smoothing kernel can be used to extract high-frequency detail.)

As the algorithm iterates, it applies progressively more dilute smoothing filters—filters "with gaps"—to the residual images to create scales 2, 3, and higher. This filter is diluted by applying its outer elements to pixels at progressively greater distances. At each new scale, the radius of the high-pass filter doubles: scale 1 corresponds to a radius of 1 pixel, scale 2 to a radius of 2, scale 3 to a radius of 4, scale 4 to a radius of 8, and so on to radii of 16, 32, 64, 128, 256. Successive wavelet scales contain progressively lower-frequency structures, as shown by the example in Figure 18.1. By scale 8, the residual contains very little structure. However, when the residual and all the wavelet scales are added together, they recreate the original image exactly.

Because the mean of each scale removed from the original is zero, the total pixel value of the residual remains the same as the original image. This has the interesting consequence that—without making the image darker or lighter—we can multiply, divide, or threshold wavelet scales to enhance, control, or remove image features based on their spatial frequency content.

### 18.1.1 The Wavelet Function

The Mexican Hat function is a continuous function equal to the difference between a positive Gaussian function and a negative Gaussian function of twice the width and half the value. The smallest equivalent one-dimensional convolution kernel is:

$$\begin{bmatrix} -0.25 & 0.50 & -0.25 \end{bmatrix} \cdot \qquad \textbf{(Equ. 18.1)}$$

The smallest equivalent two-dimensional wavelet kernel can be formed as the convolution of the row kernel with the column kernel:

$$\begin{bmatrix} -0.25 & 0.50 & -0.25 \end{bmatrix} \otimes \begin{bmatrix} -0.25 \\ 0.50 \\ -0.25 \end{bmatrix} \qquad \textbf{(Equ. 18.2)}$$

The blurring kernel complementary to the Mexican hat kernel is:

$$\begin{bmatrix} 0.25 & 0.50 & 0.25 \end{bmatrix} \cdot \qquad \textbf{(Equ. 18.3)}$$

and its two-dimensional counterpart blurring kernel is:

$$\begin{bmatrix} 0.25 & 0.50 & 0.25 \end{bmatrix} \otimes \begin{bmatrix} 0.25 \\ 0.50 \\ 0.25 \end{bmatrix} \qquad \textbf{(Equ. 18.4)}$$

Because this is a separable kernel, convolving the image to produce successive wavelet scales is fast and efficient.

### 18.1.2 Properties of the *À Trous* Wavelet Transform

Understanding the properties of the wavelet transform is the key to using it effectively. Here is the nomenclature:

- The *original image* to be transformed is $S_0$. It has dimensions $x_{\max}$, $y_{\max}$; a pixel in image $S_0$ at location (*x,y*) is $S_0(x, y)$.
- There are $J$ wavelet scales. The index $j$ denotes to the wavelet scale. Values of $j$ run from 1 to $J$.
- The wavelet transform consists of a three-dimensional array of *wavelet coefficients* with dimensions $x_{\max}$, $y_{\max}$, $J$. The wavelet coefficient at location (*x,y*) in the *j*-th scale of the wavelet transform as $w(x, y, j)$.
- *Residual images* are $S_1, \ldots, S_j, \ldots S_J$; a pixel in image $S_j$ at location (*x,y*) is $S_j(x, y)$.
- The wavelet kernel $\varphi$ contains elements $\varphi(r, s)$. For the $3 \times 3$ Mexican hat kernel, the indices take values of -1, 0, 1.

The *à trous* algorithm begins with the original image $S_0$ and $j = 1$. In the first iteration, The low-pass kernel is convolved with the image:

$$S_j(x, y) = \sum_{r, s} \varphi(r, s) S_{j-1}(x + 2^{j-1} r, y + 2^{j-1} s) . \qquad \textbf{(Equ. 18.5)}$$

The first iteration generates the residual image $S_1$ from $S_0$; the second iteration generates a new residual image, $S_2$, from $S_1$, and so on. The term $2^{j-1}$ equals 1 for $j = 1$, 2 for $j = 2$, 4 for $j = 3$, 8 for $j = 4$, and so on, doubling with each iteration. This causes the kernel, effectively with gaps or holes, to act over larger and larger regions of the image.

After the wavelet coefficients for each scale are computed by subtracting the blurred image from the previous residual image:

$$w(x, y, j) = S_{j-1}(x, y) - S_j(x, y) , \qquad \textbf{(Equ. 18.6)}$$

the algorithm loops back and computes the next $S_j$.

Each successive scale samples its scale at twice the size, half the spatial frequency, and half the resolution of the previous scale; the factor-of-two step from one scale to the next is called *dyadic*.

Values of $j$ typically run from 1 to 8, over a 1 to $2^8 = 256$ range of resolution. Wavelet coefficients in $w_j$ are both positive and negative, and have a mean value of zero.

After decomposing an image with the wavelet transform, the image can be reconstructed exactly using the inverse wavelet transform:

$$S_0 = \sum_{j=1}^{J} w_j + S_J , \qquad \textbf{(Equ. 18.7)}$$

where $J$ is the maximum scale in $w_j$, and $S_J$ is the last residual image after the extraction of scales $w_1$ through $w_J$. This equation simply states that the original image is the sum of the wavelet scales plus the last residual image. To reconstruct any pixel $(x,y)$ in $S_0$, it is only necessary to sum the corresponding wavelet coefficients plus the residual image.

28. **Equation 18.8** has units of electrons:

$$n = gS - gB \text{ [electrons]}. \qquad \textbf{(Equ. 18.8)}$$

29. Equation 20.1 should read:

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5} \cdot \frac{1}{e^{hc/\lambda kT} - 1} \qquad \textbf{(Equ. 20.1)}$$

30. In the interests of clarity, the paragraph following **Equ. 20.17** in **Section 20.3.2.2** should read:

where $Z_{\mathrm{G2V}}$ is the angular distance between the standard star and the zenith, and $k_R$, $k_G$, and $k_B$, are the extinction coefficients for each passband. In this formulation, the absorption of light by the atmosphere is entirely removed, so that objects appear as they would from outer space.

31. In **Section 20.3.2.2**, insert the following short paragraph after the listed extinction coefficients:

These extinction values are approximations derived from stellar photometry, as described in Section 10.2.