Chapter 7

Image Processing

There are three stages in the treatment of CCD data, which we call *acquisition*, *reduction*, and *analysis*. The techniques for data acquisition are presented in Chapter 12 and should be thoroughly studied before raw data are aquired. Data reduction can be further divided into image processing (properly calibrating your detector) and photometric reductions. We treat image processing now, and photometric reductions in a following chapter, rather than after the chapter on observing techniques, because intelligent data acquisition requires a knowledge of the types of observations necessary for the reduction process. The reduction of data from DN into magnitudes tied to the standard system can be a complicated process, but one that is required by many research projects. Careful reading of this material and the examples in the appendices will enable you to reduce any instrumental observations and place them on the standard system. Most of the third stage, data analysis, is left up to the individual. Analysis involves the calculation of such quantities as periods, orbital elements, and in general all calculations beyond the determination of magnitudes and colors. The analysis depends greatly on the purpose of the investigation and should be obtained from other sources.

1. A Data-Reduction Overview

You have some raw CCD frames taken of a program field. What are the steps necessary to complete the reduction? There are many different ways that data reduction can proceed. A general outline that fits most situations follows:

- subtract any dark current or bias signal from each raw image.
- flatfield the frames.
- extract instrumental magnitudes for all objects of interest.
- calculate the instrumental magnitude and colors. For differential photometry, calculate the magnitude differences between the variable and comparison star(s).
- determine the extinction coefficients and apply the extinction correction. This step is often unnecessary for differential photometry. If you intend to leave your differential photometry on the instrumental system, skip the next two steps.

- Use the standard stars to determine the zero-point constant and, if necessary, the transformation coefficients.
- Transform your instrumental measurements to the standard system.
- Estimate the quality of the night by comparing the transformed standard-star magnitudes and colors with their accepted values. For differential photometry, check the reproducibility of the comparison star measurements after correcting them for extinction.
- Perform any ancillary calculations such as time conversions that are necessary to make your observations useful and publishable.

Steps 1,2,3,5,6 and 7 and illustrated by a worked example in Appendix H. An example of step 4 is found in Appendix G. Step 9 is covered in detail in Chapter 5. In what follows, we review the concepts and difficulties of some of these steps and present a worked example of the data reduction associated with differential photometry. The one exception is flatfielding, which because of its importance, we have moved to its own chapter. However, several of the calibrations and tests described in the current chapter will make use of some sort of flatfielding system, so you may have to read the flatfielding chapter as a digression when encountering one of these steps to clarify the operation.

2. Bias Frames

Bias frames are those CCD images taken with the shutter closed, and an exposure time as close to zero length as possible. When you read out such an image, it will not have zero DN at each pixel. When the camera manufacturer designs a CCD system, they insert a small electronic offset (or bias) on the output of the amplifier. This bias level ensures that the input to the ADC is always positive.

In the best of worlds, the bias would be an easily subtracted constant for every pixel. In reality, there are a number of noise sources related to the bias level. For proper calibration, you not only want to subtract the proper mean level, but also reduce all noise sources to only a small contribution of the total data frame error budget.

The main bias level is usually set by a resistive voltage-divider analog that provides the offset. There is a small amount of noise associated with this signal, dependent on its Thompson noise. This noise source is Gaussian in nature, not Poissonian.

There may be incoherent noise external to the CCD system, such as that caused by motor spikes, 60Hz power supplies, fluorescent lights, computer monitors, etc. that couple into the readout electronics. In general, these noise sources are not repeatable, and the only methods of removing them are to eliminate the source of the noise. You can mitigate the effect in a master bias frame by taking many individual bias frames that are combined into the master. However, you should be aware that such noise sources will also appear in your data frames and thereby increase the inherent noise in those frames.

Just as for data frames, there is also readout noise in each bias frame. This is the major noise source for bias frames. To reduce this noise source, you need to combine many bias frames in the creation of a master bias. Cosmic rays will occur since there is a finite readout time; these events can be removed with minmax or median filtering from the stack of biases used to create a master bias.

In addition to these standard random noise sources, there are several coherent fixed-pattern (that is, unchanging) noise sources, such as:

- preamp glow. The on-chip amplifier of a CCD will act like an LED and actually glow when it is energized. This means that a small amount of increased signal will appear at the beginning of each row, usually with an exponential tail away from the amplifier. If the amplifier is left on during an exposure, the corner of the CCD nearest the amplifier will show a noticable increased signal level. Early camera systems that used the TI TC-211 CCD had obvious amplifier glow. You can lessen the impact of amplifier glow by ensuring the amplifier is turned off whenever it is not needed, and by changing the various voltage levels to some degree. Many CCD manufacturers move the amplifier as far away from the imaging area as possible through the use of extra pixels in the serial register.

- preflash structure. On some CCDs, there are sufficient charge traps that you must preface any exposure with a short uniform illumination exposure to fill the traps. This is often done in spectrographs, for example, since the spectra is usually weak and you cannot count on the sky background filling the pixels far enough to override the trapping process. At the same time, it is very difficult to make the preflash exposure uniform. If you take your bias frame in the same manner as an object frame, and include the preflash, then you can subtract this structure easily. Note, however, that any added preflash signal also adds noise to the bias frame, so limit the use and amount of preflash.

- RC exponential tails at the beginning of each line. Sometimes the signal chain electronics is AC-coupled and has RC constants to remove as much high-frequency noise as possible. Unfortunately, this means when the electronics are hit with a large signal, they take time to decay to their nominal level. Since the first column in a CCD often has a large signal due to bleeding from the adjacent non-illuminated silicon, you might see such a tail on each row in your CCD. Another common problem that results in a similar feature is inadequate power supply regulation, so that a CCD vertical transfer causes a slight shift in the voltages that takes a finite amount of time from which to recover. A typical bias frame that shows such an RC time constant is shown in Figure 7-1. Figure 7-2 is a plot of a typical row of this CCD, showing the exponential rolloff at the beginning of the line.

- cosmetic defects. Sometimes hot pixels will cause a partially blocked column; these pixels may appear during the readout phase because they are a partial short to one of the gates, and therefore show up in every frame including bias frames at about the same level. They can often be subtracted out during the bias subtraction phase. Figure 7-3 shows a bias frame with a blocked column.

Coherent noise sources such as those mentioned above can be removed from the science frame by subtraction of a mean bias.

Examination of bias frames can be very beneficial in understanding how your CCD system works and what impact extraneous noise sources have on the readout process. For example, we often multiplex operations at professional observatories, reading out a CCD while moving to a new field. Since the readout of a 2kx4k CCD might take a minute or two, this multiplexed operation can improve the observing efficiency. At the same time, if the use of the telescope slew motors or the dome rotation cause voltage spikes that get coupled into the CCD electronics, you may increase the average noise or even cause the loss of a segment of the image. It is easier to see low-amplitude noise sources in the low-level bias frames, both in actual standard deviation calculations and by visual inspection (the eye is a marvel at seeing correlated noise).

Note that if you do not have any noticable structure in your bias frames, then subtracting them from your data frames will just add noise. Often we just subtract a fit to the bias overscan columns instead.

Note also that a bias frame with zero second exposure still shows dark current since it takes a finite length of time to read the CCD. Other systems will not permit a zero-second exposure. For these two cases, you can obtain a true zero-exposure bias frame by taking two dark exposures of differing length. Say, for example, you take a dark exposure of one second duration and a dark exposure of two second duration. Then each frame contains the electronic bias plus some dark current. If you create a pseudo two-second dark by adding the one-second dark to itself, then that pseudo dark has two seconds of dark current like the true two-second dark, but has twice the electronic bias. So if you subtract the true two-second dark from the pseudo two-second dark, what remains is just the electronic bias.

In mathematical terms, we can say that the signal seen in a bias frame can be represented by the following equation:

$$bias_{x,y} = \frac{1}{g}[b_{x,y} + fp_{x,y} + t_{x,y}d_{x,y} + \sigma_b + \sigma_f p + \sigma_{ext} + \sigma_{rn} + \sigma_d] \tag{7.1}$$

where bias is the final signal; g is the gain of the system in electrons per DN; b is the signal from the bias offset; fp is the fixed-pattern signal, t is the readout time, d is the dark current, rn is the readnoise, and the various sigmas represent the noise associated with each of these parameters.

Note that, for most CCD systems, the reset process is quite fast so that all pixels are cleared of charge almost simultaneously. That is why there is more dark current on those pixels that are read last during the readout process, and therefore there is more inherent noise in the pixels furthest from the readout point. You can make the dark current virtually identical on each pixel if you modify the reset process so that it takes the same amount of time as for a standard read of the CCD (for example, do an initial read and discard of the detector, then open the shutter), and indeed something similar to this is what we do for infrared detector arrays. The disadvantage is that you double the readout time since the reset time takes just as long as the final read.

3. Bias Overscan

Most professional CCD software systems include more columns in the readout than are physically present on the CCD. After the charge from a row has been shifted into the serial register, and you read out each pixel of that register in turn, you can continue to 'read' the serial register. These extra pixels have no signal, but do have the electronic bias. Therefore, if your bias frame has no fixed-pattern structure and your dark current is negligible, then you can use this overscan as a bias level for your specific frame. Since the overscan can have a slow variation with time, the common practice is to average the overscan columns for each row to derive a vector of bias values. You can then fit this vector with either a constant or a low-order polynomial, and then subtract the fit from each pixel.

The bias overscan only corrects for the electronic bias. If you have dark current, for example, the overscan will not correct the accumulation of charge with time. In general, you will see a gradient in your bias frames since the last rows to be read will have had longer time to accumulate dark current.

Using the bias overscan in lieu of a master bias frame can reduce the noise added to a data frame from the bias subtraction process. It can also be used to measure the true bias level of even a master bias frame, and therefore leave just the added features of fixed pattern noise and dark current. Figure 7-4 shows a typical bias overscan vector for a liquid nitrogen cooled CCD, where there is no measurable dark current during a readout.

4. Dark Frames

Most professional CCD systems are either cooled to liquid Nitrogen temperatures, or are run in MPP mode with ThermoElectric Coolers (TEC). In either case, the dark current is so small as to be negligible in all but the most extreme cases.

This is not the case for amateur systems. Not only are they typically cooled to only -25C, but they are often run at maximum cooling without regulation, so that the CCD gets colder as the night gets colder. Even for regulated CCDs, the amount of regulation depends on where the temperature sensor resides relative to the CCD itself. This means the dark current varies throughout the night. Since dark current is an additive effect, it must be removed before flatfielding.

The obvious technique is to take a dark frame immediately after your data frame, and make the dark exposure length exactly equal to your data exposure. Assuming the CCD temperature is slowly changing, the dark frame will be close enough in time to be representative. Of course, the signal/noise in the dark frame will be low, and it is better to take several dark frames and average them to reduce the noise. If you do this for every data frame, you will be spending most of your valuable telescope time taking dark frames!

If you temperature-control your TEC, then you need only take a master dark frame sometime during the night and use it for all data frames. You should take 10-20 dark frames and average them to improve the signal/noise. We recommend taking the dark frames during the course of the night, rather than before or after your observing program, because some cameras have light leaks. By taking your dark frames during darkness, you lessen the effect of such light leaks.

Note, however, that even if the temperature of the CCD remains constant, you will either need to take a set of darks at each exposure time you use through the night, or else take a long dark frame and scale it for each exposure time. To use this technique, you will need to subtract a bias frame from the master dark frame in order to leave pure dark current in the master dark. Again, the object is to add as little noise as possible into your data frames, so the master bias and the master dark should have as high signal/noise as possible. This technique will work if the dark current is constant during the night and you know the exposure times fairly accurately. Don't assume that your vendor has temperature-controlled your CCD accurately; run tests yourself. Some systems have a temperature sensor as tightly coupled to the CCD itself as possible; others have the sensor just about any convenient place. In addition, the quality of temperature control can range from 0.1C regulation to several degrees of overshoot.

If you are not temperature controlled, but do have a temperature readout, then you can take a series of dark frames at different temperatures, and interpolate between them for the temperature of the data frame. You should be aware in this case that dark current does not scale linearly with temperature; it is a power law, doubling about every 6C of temperature change, as shown in Figure 7-5. Even this doubling factor changes depending on the starting temperature. In addition, there are a family of hot pixels that have nonlinear response to temperature and may not subtract cleanly. Figure 7-6 shows a histogram of the dark current values for a Loral CCD, showing both the multipeaked structure of an MPP chip as well as a high-valued tail of hot pixels.

If you are not temperature controlled, and have no temperature readout, then if you have either a frame-store CCD or a CCD with masked columns specifically intended for dark current measurement, you can determine how much dark current is present and use an appropriate dark. For example, the

Loral 442A CCD has a picture frame of covered columns and rows around the outside of the imaging area. Since these pixels are in every way identical to the imaging pixels, they accumulate dark current while the exposure takes place. You can use these pixels to estimate the average dark current over the imaging area, or to scale a master dark to the appropriate level. The latter is the preferred method in order to handle hot pixels as best as possible.

Otherwise, take at least one dark frame after each science frame. To improve the signal/noise, you can use the median value from a single dark frame to select a proper master dark for subtraction.

5. Defringing

While backside illuminated CCDs can have much higher quantum efficiency and less internal pixel structure than front illuminated CCDs, they do have one major fault: interference fringing due to night-sky emission lines. This is also called *etaloning*.

Silicon starts becoming transparent the further you go towards the red. In fact, you can use silicon in the near-infrared as a high-index lens material. As you enter the 800nm region of the electromagnetic spectrum, silicon becomes semitransparent and the QE decreases. The photons are not immediately absorbed; instead, they pass through the epitaxial layer and encounter both a layer of silicon dioxide (refractive index 1.5) as well as the polysilicon gates. This is a sizable discontinuity compared to silicon (refractive index 4), and acts to produce a large reflection. On the other boundary for thinned CCDs is an AR coating that is usually optimized for the blue; this AR coating is not as effective in the red and also acts like a mirror. The photons can transverse the thickness of the CCD several times, not only increasing the effective path length in the silicon but also setting up a standing wave pattern. Some of the photons are lost at both reflective surfaces and by absorption in the silicon, but at long wavelengths sufficient amplitude survives to cause significant constructive or destructive interference.

This condition – a transparent optical element with two highly reflective surfaces – forms a resonant optical cavity called an etalon. The equation governing an etalon is

$$I = \frac{I_{max}}{1 + (2Q/\pi)^2 sin^2(\frac{2\pi d}{\lambda})} \tag{7.2}$$

where Q is the finesse of the cavity, λ is the wavelength, and d is the distance between the two surfaces. When the distance between the two surfaces is an integer multiple of the wavelength, you will have constructive interference.

For a CCD, the reason for fringing is monochromatic night sky lines, primarily OH airglow lines. A spectrum of the night sky is shown in Figure 7-7. The prominent lines blueward of 700nm are primarily forbidden Oxygen [OI] at 557.7nm and 630.0nm. These do not normally contribute to fringing since the silicon absorbs these wavelengths well (though we have seen fringing at 630nm due to aurora!). However, in the I-bandpass, there are numerous OH lines that can cause fringing since the silicon has become semitransparent. For CCDs, the airglow light is monochromatic, but the eptitaxial layer is not uniform in thickness. This thickness variation does not have to be large; since silicon has an index of refraction of around 4, the effective path length through silicon is 4 times that through air and therefore the thickness variations are also multiplied by 4. The thickness variations give rise to interference fringes as the thickness passes through integral wavelength dimensions. A typical fringe pattern is shown in Figure 7-8.

You can also run across fringing if you use a red narrow-band filter, since this has the effect of producing monochromatic illumination of the CCD. In this case, all objects in the field are affected by the fringing, and the fringing can be removed with the usual flatfielding techniques.

Fringing can be decreased by increasing the epitaxial layer thickness, which decreases the blue response but has the advantage of increasing the red response since you have a longer pathlength of silicon to absorb the photons (and thereby reducing the number of photons that survive to be reflected and cause interference). You can also use a red-optimized AR coating, removing one of the reflective surfaces.

The fringing is an additive feature (just an addition/subtraction to the night sky, depending on the amount of airglow emission), and remains relatively spatially constant with a given CCD and filter. It can change intensity as the amount of airglow changes. It can also change spatially as the relative intensity of the airglow lines change.

You cannot remove fringing with a flatfield frame. If you are using twilight sky or dome/lightbox flats, the light source is polychromatic and therefore will not create fringes. Since fringing impacts the ability to do photometry near the sky background limit in the I-bandpass, you need some technique to remove it.

The obvious technique for creating a master fringe frame is to form a master sky flat in addition to any other flatfielding method you might use (these are all mentioned in the next chapter). Since the fringes will appear in the sky flat and not in, say, a dome flat, dividing the sky flat by the dome flat will remove the normal flatfield variations but leave a master fringe frame, where the fringe amplitude is shown as a percentage change from unity. You can form a master sky flat from a single field if you take multiple exposures of the field and move the telescope sufficiently so that you can median-filter out all stars. This works as long as your field does not contain extended objects like galaxies. This on-target sky frame is discussed by Melnick et al. (1999) and is shown to be effective in reducing scattered-light gradients as well as fringing. The main concern is to *dither* your images by a significant fraction of the field width so that stars median-filter out cleanly. If you have bright stars in your frames that leave residuals using a median filter, you can also try a minmax reject and remove a larger fraction of the high pixel values, keeping the lower values to average together.

Tyson (1986) discussed a technique called adaptive modal filtering, which computes the absolute difference between the mean and the median of values associated with each CCD pixel within a set of images, and rejects deviant values until the mean and median are the same within a specified error or until a maximum number of values have been rejected. Then each pixel is median-filtered using the surviving values from all images. This can be done easily in IRAF, for example, since that system allows you to prescale images by the mean/median/mode before combining with median techniques.

Ostrov (1997) suggested a filtering technique that can be applied to single frames. The concept is that fringing has a low frequency spatial scale, so you should be able to use multiple pixels to follow the fringe pattern. In Ostrov's case, he used a single-pixel-wide rotating column whose length was short enough to follow the winding pattern, yet long enough to improve the signal/noise over a single pixel. If you take the median of all pixels at a particular orientation and derive the dispersion at that orientation, then rotate the window and perform the operation at all possible angles, the minimum of the dispersion will be when the window aligns with the fringe. If you store a value proportional to the inclination for that pixel in an inclination map, you will end up with a two=dimensional map of the same size as your original image that indicates the orientation at each pixel of the fringes. Then invert the process: use the inclination map to orient the window in your original image, calculate the median value, and form a fringe frame using those median values.

Once you form a master fringe image, the techique to remove fringing from any given frame is to first flatfield the frame in the normal manner. Then find the median or modal value of the sky background. Scale the master fringe image to this value, and subtract it from the frame. Note that the subtraction means another noise source is being added to the final processed image; you should make the highest quality master fringe image that you can. Since fringing has low spatial frequency, you may be able to smooth the master fringe image slightly to improve the signal/noise.

6. Image Processing Software

We do not wish to recommend any particular image processing system for your CCD camera. There are many public domain systems, such as IRAF, PCVISTA and MIDAS, that do entirely adequate processing; there are many more commercial products that work equally well. You may be restricted by the computer and operating system; you may have requirements such as image sharpening, stacking, or psf-fitting photometry that limit your choices.

In all cases, however, you should have the following minimum functions if you want to do quality photometry of the processed images.

- Image arithmetic, especially subtraction (bias frames), addition (coadding images), and division (flatfielding).
- median filtering. Highly useful for creating master flats and bias frames.
- image display with good contrast adjustments.
- cursor readback.
- ability to read your CCD images
- ability to read and write FITS images

Performing all processing steps in floating-point arithmetic is extremely useful, as is the reading and writing of floating-point images.

7. Data Archival

Archiving your data is important. The number of times that we have gone back to our data to reprocess or retrieve different information is almost embarrassing. How careful you are in the archiving process is up to you, but we include here some recommendations of why, how, and what to save. Listed below are several files that you might consider part of your archive:

- Logs of observations
- The nightly raw images
- The processed (flatfield/darksubtracted) images
- Any extracted starlists
- Fully transformed starlists

- Notes regarding nightly processing

You might find other files useful to archive, depending on your individual approach to research. For example, some observers will save input files to processing programs that might include transformation coefficients or object finding thesholds. For other observers, these values might be included in their starlists. Still other observers will save the source code or executable programs actually used for processing their data on a nightly basis. These programs are typically similar in size to a single CCD frame and so do not take up appreciable space on the archival media.

You might consider building a database of your images. Such a database might include the date/time/filter/exposure time, etc. for each image; perhaps central coordinates or RA,DEC of the corners of the image; and of course what disk or tape the files were archived upon. Such databases are a convenient method to search for a specific frame and be able to restore it, or to see if you might have recorded a precursor to some transient phenomenum.

7.1. media

Only a decade or so ago the main archival media was magnetic tape. Technology has advanced dramatically on the archival media front. The primary drivers have been home archive of large hard drives, and the business market where large databases need to be archived for several years. In general, do not assume any technology will be present for more than 5-10 years, and do not buy the current state-of-the art. The most encouraging aspect of data archival is that the media capacities and cost are continually improving and are keeping up with the CCD size. Listed here are some of your current options.

Magnetic tape. Yes, 9-track reel-to-reel tapes still exist. Controllers were developed for the ISA bus soon after the PC-AT was on the business market. We would never recommend using magtape for current archival storage as the technology has been far surpassed. Magtape suffered from low density (up to about 200MB/tape was the maximum), media damage during spooling (we have lost many nights of data when a tape misfed and crumpled), and imprinting between layers due to the magnetic field. On rarer occasion, we have seen tapes destroyed due to heat and closeness of a magnetic field. Magtape needs to be stored in a vertical fashion in a cool, low humidity environment. Every couple of years you should respool the tape to prevent printthrough. However, the early CCD results from the national observatories were all recorded on magtape and if you want to look at these old files, you need some method of reading the tapes. If you have a stack of old tapes, we highly, highly recommend transfering them to some newer media.

Floppy disk. This technology underwent rapid development from the first 8-inch single-sided low density format to the current 3.5-inch hardshell media. About a decade ago, vendors standardized on the 1.44MB 3.5-inch disks, and that appears to be the final development of the technology. There have been a few attempts to increase the density, most notably a 2.88MB model, but they have not been widely accepted by the community. The standard 1.44MB floppy is durable and a nice size to archive. However, the capacity is very small in CCD standards, and while the access is random, the access speed is quite slow. We use floppies to store source code and starlists, where the files are small and rapidly evolving. The floppy should be stored in a library case in a vertical position, again in a low humidity cool environment. The floppy can be damaged by heat and destroyed by force. The larger concern is that each floppy has to be formatted with track and sector information, and that formatting may or may not be compatible with different drives (alignment) or readable in the future. Use floppies for temporary storage of small files.

Magneto optical and floptical disks. These media are the same length and width as a 5.25-inch floppy disk, but thicker. They can store 128MB-1.2BG of raw data. They are somewhat more stable than magnetic media since they are recorded magnetically but read optically. Both drives and media tend to be expensive.

The primary type of magneto optical disk is called a WORM (Write Once Read Many). This medium was used for early archival at the Space Telescope Science Institute since the capacity was large for a randomly accessible archive, ideal for access to image files such as the Digital Sky Survey. The 5.25" variety had about 800MB storage capacity, 400MB per side. Some WORM drives used even larger platters (12 inches) and could store several GB per disk. The disk is written by heatng up a magnetic layer with a laser until it reaches its Curie point so that it can be magnetically written. When cooled, the particles retain their orientation for indefinite time periods, and a lower-power laser is used to read the polarized reflected light.

A more modern development are the MO drives with 3.5-inch media. These are available for PCs in 230MB capacity. While convenient since they are re-writable (though the writing speed is slower than the read speed), they have not widely available and our thought is that they are ok as local short-term archive but should not be used for permanent storage.

The floptical disk is similar in size to a 3.5" floppy, but holds about 120MB/disk. The drive itself can be used to read and write standard 1.55MB floppies. However, as for the MO drives, the floptical is not an adopted standard and so should be considered for local, short-term storage only.

Compact Disks. These are optical disks in a 4.6-inch format. Three types are common: the CD-ROM is produced from a master and is used to mass-copy a single disk. This technology is used for making audio CDs and in some cases astronomical, widely distributed, catalogs, but is not normally used for data archival. The CD-R (CD Recordable) technology is similar to the WORM in that you can write once and then read multiple times. Finally, the CD-RW is a multiple read and write disk technology. All CDs have about 650MB capacity per disk. The CD-ROM and CD-R media use a common recording standard (ISO-9660), which describes how the data is written to disk and also the limitations on directory and file naming (8 character names plus 3 character extensions, for example). There have been multiple extensions (called *Books*) to the basic ISO-9660 standard to allow larger file names, multi-session recording, etc. Most of these extensions are common enough that you can feel reasonably confident that later drives and operating systems will still be able to read the older media. We tend to be extremely conservative and just record using the original ISO-9660 format. CD-RW media have a directory structure much like your typical hard drive, and so are more prone to operating system imcompatibilities for long-term archival. We recommend using CD-RW for temporary storage of image data, and CD-R for long-term storage. We use CD-R media exclusively in our own archives due to its cheapness (disks are in the $0.20 range now, drives are well under $100) and standardization. According to industry, CD-R media should last for 100 years; CD-RW media for a few decades. CD-RW media can be rewritten hundreds of times. They should be stored in plastic jewel cases for safety, and placed on shelf in a vertical fashion. Several software packages are available to both write CDs and to create labels. While you can place a glued label on the CD itself, we prefer to just use a felt-tip pen to label the disks and not have a label that can eventually work loose or add weight and thickness to the disk. Specialized felt-tip pens are available, but we've found that Sharpie permanent markers work about as well as anything and are very inexpensive. Be careful and don't use ball-point pens to mark on CDs. The top surface can be scratched, which affects the laser read beam and can ruin your disk. The plastic disk itself is quite durable (at least one user tried to purposely break a disk, and it took a sledgehammer). The random access is slower than for a hard drive, but once positioned, an image can be read off of a CD quite fast in streaming mode. The National Institute for Standards and Technology (NIST) has a nice document on the care of CDs and DVDs; look at http://www.itl.nist.gov/div895/carefordisc/CDandDVDCareandHandlingGuide.pdf

DVD media. The Digital Versatile Disk has been under development for a decade now. Recently, several incompatible recording methods have been proposed, and a DVD-RAM drive made available for PCs. The capacity is quite high; 2.6GB per side or 5.2GB per disk. Other standards, such as DVD+R, give 4.7GB on one side. In the future, DVD-RAM will probably replace CD technology for data storage, but we feel that the current marketplace is still evolving and the media cost is still high.

Removable hard drives. Several vendors are making pluggable hard drives. While these offer the utmost in access speed, the drives are quite expensive and much more fragile than optical media. They are useful in a different market than CCD image archival and are not recommended except perhaps for local image transfer.

The most well-known removable hard drive is the Zip disk, pioneered by Iomega. It currently comes in two sizes (100MB and 250MB); the physical media is 3.5inches and about twice as thick as a floppy. The read/write time is between that of a floppy and a hard drive. The media tend to be much more expensive than CD-R. This technology is useful for short-term local archive.

Another removable method is solid-state disk. Here battery-protected static memory is used to store data. Memory costs have been decreasing rapidly, and such solid-state disks can have high capacity at low cost. Digital camera storage media is probably the largest driver for this technology, with current capacities in the 128MB range. These solid-state drives have the fastest possible access times and are a good choice for remote computer systems, where the lack of moving parts improves reliability. Their capacity is still too small and cost too high for CCD image archival. The batteries do eventually fail, so these devices are also not good long-term archival media. A variation on the solid-state disk theme is the thumb drive, a USB-pluggable static memory device that can be obtained in sizes up to about 1GB at the time of this writing. These are great for transferring files or carrying presentations from one computer to another, but still have too small capacity for permanent storage of images.

As disk capacity increases, the chances of a head failure means far more data will be lost when it does occur. The simplest protection method is called *disk mirroring* where you make an exact copy of one disk onto another. Another alternative is RAID (Redundant Arrays of Inexpensive Disk), which safeguards against a hard drive failure by making the data redundant (much like ECC memory). With such RAID systems, on-line storage of a Terabyte or more is possible, though still not as secure as writing onto removable media.

Cassette tape. These tapes are similar to the audio cassette tape. In fact, the earliest PCs used standard audio tapes for program storage. The QIC (Quarter Inch Cartridge) was based on this format but with far better recording technology. Current storage capacity is up to 2GB/tape; however, more common capacities are in the few hundred MB/tape range. Any tape system has serial access; you must fast-forward or rewind to the proper record before reading or writing. As such, tapes are recommended for raw image storage where you may want to go back and restore an image on infrequent basis. Strictly speaking, cartridge tape refers to systems with one reel, and cassette tape refers to two-reel systems. The Travan tape uses an 8mm tape with longer length to give higher capacity (up to about 1.6GB/tape uncompressed).

Higher capacity cassette tapes use helical scan recording technology. Two such tape systems are the

Exabyte (proprietary format up to about 5GB/tape, though recent models have touted 25GB/tape uncompressed) using the VHS 8mm tape and the DAT (Digital Audio Tape), using a 4mm tape (about 4GB/tape). Data compression is often used to increase the capacity. All magnetic tape suffers from the same inherent problems of imprinting (from one layer to the next) as well as the fragility of the media itself (you can mangle a tape when reading or rewinding, and tape melts easily). The tape media itself comes in two grade qualities, one for video/audio recording and one for binary data. Use the binary tapes for archiving CCD images. If you are spooling files serially onto a tape for long-term archive, data rates can be quite high; several MB/sec can be easily stored.

Digital Linear Tape. The Digital Linear Tape (DLT) is a outgrowth of the CD market, using a similar technology to record on a plastic tape. The storage capacity is high, 40GB or more, especially if compression techniques are used, and the tapes can be written multiple times. However, again the access is serial (average seek time is about one minute) and so is recommended for storage of large quantities of data that are not going to be accessed frequently. We have DLT drives here primarily for hard disk backups. The cost per tape is high and the long-term operating system incompatibilities limit the use for archival purposes. Shelf life is quoted at about 10 years.

Jukeboxes are available for just about every technology to offer access to large quantities of data stored on multiple media. However, such access is quite slow since the media have to be loaded mechanically. If you need a jukebox, we suggest that you investigate the next generation of storage media that offer sufficiently increased capacity that a loader is not necessary.

7.2. archiving techniques

Once an exposure is completed, the image is read into memory. For most multitasking operating systems, the next exposure can then be started if the highest repetitive rate is needed. The raw image is written to disk, and optionally processed and written to disk a second time (we always keep the original, raw image). The image is then displayed, and optionally pipeline-processed to produce a real-time starlist.

The method used at NOFS is to archive the raw frames onto magtape (Exabyte) since we do not expect to ever read these frames again, but need them for security. Normally, the hard drive raw frames are processed later and then written to CD. Every so often, the starlists are written to CD. The Exabyte and the CD are kept in different places for safety. You should keep track of technology changes; every so often, back your data up to the next generation so that it can always be read. Note that it is not just hardware; the software and operating system continually change and may not support a particular media in the future. Also, the way files are written can change. For example, many camera vendors provide software that prefers to write images in a native format. That format may not be readable in the future if the company folds or evolves.

7.3. data compression

Archival of small CCD frames is not difficult; many nights will typically fit on a CD-R. However, when you start investigating high-speed all-night time-series photometry, or start using a large format CCD, even a CD-R fills up quickly. Observers soon start thinking about compressing their data so that fewer media are used (read: lower cost and smaller library space).

There are two types of data compression that might be used: *lossless* and *lossy*. Lossless compression means that the data can be restored to the exact same numbers as before the compression took place. Lossy compression means that some information has been lost. We will discuss each of these methods below. The reader should be aware that almost every possible compression scheme has been patented (often several times!), so home use of the algorithms is generally accepted, but commercial use requires the payment of royalties. There are many books written on data compression. Some common paper references for astronomical data compression include Sabbey et al. (1998), Johnson (1999), White and Percival (1994),

Lossless compression is only possible if the data is not truely random (random data cannot be compressed). For example, on astronomical images, only a few pixels approach the saturation limit; most are close to the sky background level. Under such circumstances, algorithms have been developed that preserve all of the information contained within the image. The common general-purpose compression algorithm, as used in PKZIP, uucompress and other system-level compression programs, is Lempel-Ziv-Welch (LZW), which works well with highly redundant data. LZW is a substitutional compressor that creates a dictionary of possible data values and then substitutes an index in the file when a given data value occurs. For astronomical images, LZW yields compression ratios of 1.4-2.0 and so can have a major impact on your archival size.

More image-specific compression algorithms have been developed. Richard White (1994) discussed an algorithm for compression of the Digital Sky Survey called *hcompress*. This algorithm is a wavelet transform called the H-transform (a two dimensional implementation of the Haar transform). For lossy applications, the H-transform is followed by a quantization to discard noise and a quadtree coding of the quantized coefficients. Rice (1991) described a predictive algorithm, where previous pixels forecast upcoming pixel values, and you encode the difference between the prediction and the actual value. SDSS and QUEST both use the postage-stamp technique, where you save the actual pixels surrounding objects and discard other sky pixels, highly effective in sparse regions.

Lossy compression generally divides an image into three basic regions: a background with low spatial frequency, a high frequency component that contains important information, and noise. Most lossy algorithms make some assumption about the importance of the first two regions, and then threshold or filter to remove the noise component. Remember that noise for some applications is signal for others; lossy compression always comes with loss of information. If you intend to use a lossy algorithm, use one that has been developed for astronomical applications. Towards that end, we recommend the hcompress lossy algorithm at present. With compression ratios around 10x, this algorithm has been shown to not impact greatly either photometry or astrometry of reasonably well exposed objects. Most algorithms will start to systematically affect objects that are faint and near the sky background. Visually, it is very difficult to tell the difference between a raw image and a 10x hcompressed image, except that the compressed image seems less noisy.

Should you compress or not? Should you use lossless or lossy algorithms? These are questions that only the individual user can answer. Our personal opinion is that using any kind of compression scheme can cause future problems. Once the data are archived, you need to uncompress before being able to use the archive; an extra step. The compression program may or may not be around when you need to uncompress, and you need to know which program and which version were used to compress in the first place. For projects that produce enormous quantities of data, our usual recommendation is to process the images into starlists in near-real-time so that the original images have less importance. Then you can either discard the images or else use compression for archiving them, under the assumption that you will rarely need to go back to these images. Our experience is that after starlists are created, we further process about 10percent of the hard disk images (coadding to go deeper, inspecting images that yielded poor photometry, etc.) before deleting them. Less than one percent of archived images are inspected at a later date.

7.4. FITS

The current standard for image storage is the Flexible Image Transport System (FITS). This protocol was developed at NRAO for storage of radio and optical data, and is flexible enough that every satellite and ground-based data archive, regardless of passband or detector, has used this protocol for their storage. FITS is not elegant, and there has been extensive dialog on expansion of the system. Complete documentation can be found at http://www.cv.nrao.edu/fits/

In its basic form, the following rules are followed.

- storage The file is stored in 2880-byte records. The normal file extension is .fit, .fts or .fits.
- header An ASCII header record(s) preceeds the image section. Each header record is divided into 36 80-byte lines, without any line termination characters. Each line includes a keyword followed by any argument. Each keyword is a maximum of 8 characters. There are five defined keywords in the basic FITS format: SIMPLE, which indicates the file type (T = simple image, F = complex image or table); NAXIS, which indicates how many dimensions in the image (usually 2); NAXIS1, which is the size of the first axis (and NAXIS2 if the image is two-dimensional, etc.); BITPIX, which indicates the number of bits per pixel; and END, which indicates that this is the end of the header. If there are multiple header records, only the last record will contain the END keyword. All other keywords are optional.
- data The image data follows, with the highest dimension varying first and with each pixel stored in little-endian format where the most significant byte is stored first and the least significant byte is stored last. The last record is null-padded to 2880 bytes.

You can read the header if you are careful by setting your display to 24x80 mode and using something like 'more' or 'type' to page through the file. Stop before you reach the data section, however, as the data will appear as ASCII control characters and may influence the display. There are shareware programs to view and edit fits headers. Most recent astronomical image processing programs also enable viewing and editing of fits headers.

For display of fits images themselves, there are shareware and public domain programs; for instance, DS9 at

http://hea-www.harvard.edu/RD/ds9/

and FITSVIEW, available at

http://www.nrao.edu/software/fitsview

We do not recommend that you write your own routines for reading and writing fits files. If you need to embed such routines in your own data acquisition or image processing software, there is a nice library of fits routines available at http://legacy.gsfc.nasa.gov/docs/software/fitsio/fitsio.html which is callable by Fortran and C amongst other languages.

REFERENCES

. Howell, S.B., 1992. Astronomical CCD Observing and Reduction Techniques, ASP Conf. Ser. 23, p. 256.

. Melnick, J., Selman, F., and H. Quintana, 1999, PASP 111,1444.

Tyson, J. A. 1986. J. Opt. Sco. Am. A 3,2131.

Wu, et. al. 1972. AAS Photobulletin 1, 9.

Holst, G. C. 1998. CCD Arrays Cameras and Displays (Bellingham: SPIE).

Sabbey, C. N., Coppi, P. and A. Oemler, 1998, PASP 110, 1067.

Johnson, H. 1999, Technical Note 55, http://tass-survey.org.

White, H. L, Percival, J. W. 1994, SPIE 2199 (Advanced Technology Optical Telescopes V, ed. L. M. Stepp), 703.

Rice, R. E. 1991, JPL Publ. 91-3 (Pasadena: JPL). See also http://a188-1004.rit.edu/richmond/rice/rice.html

This plano tables was prepared with the AAS LaTeX macros v4.0.

Contents

7 Image Processing **1**

1 A Data-Reduction Overview . . . 1
2 Bias Frames . . . 2
3 Bias Overscan . . . 4
4 Dark Frames . . . 5
5 Defringing . . . 6
6 Image Processing Software . . . 8
7 Data Archival . . . 8
7.1 media . . . 9
7.2 archiving techniques . . . 12
7.3 data compression . . . 12
7.4 FITS . . . 13