Just How Good Are Flats?

John Menke
May 2005

Abstract

Using "flats" as a means of correcting various errors in CCD images is well known. What is not so well known is how well they work. This paper describes an investigation of different types of flats and how well they work. The techniques used here are described in sufficient detail that others can use them to evaluate their own systems.

Discussion

Many of us have trouble figuring out just what flats are measuring, and what they do, and why they work or don't work. This is NOT a talk about T-shirts vs. sky flats (though I will touch on that). It is a talk about some of the more subtle aspects of understanding just what flats are measuring, and when they may be misleading.

All of us have done flats of some sort at some time. My own first experience was when I tried imaging with a x3 Reducer on a 6 in. f/12 AP refractor. The vignetting was obvious, with a loss of some 30% of intensity toward the edges of the field (and this was with what is now considered a small field ST7). I tried doing a flat field, and lo and behold, the vignetting disappeared, leaving only a little extra noise in its wake. I was having little problem with dust in my camera, so vignetting was the only major problem. I tried different flat techniques (sky, Tshirt (or sheet over the scope), dome flat with a Styrofoam panel) and all seemed to give the same result. Of course, I kept reading the religious wars over flat techniques on the SBIG list and elsewhere, but it didn't seem to apply to me. But I also kept having this nagging thought—why does a dome flat from six feet away give the same result as a sky flat from five miles? And is it really the same?

My ignorance continued with no challenge until I started doing photometry, first of stars, then of asteroids. Some parts of the photometry were easy, some very confusing. But when I tried doing flats, I saw little effect (I was back to using an f/6 reducer on the C11, ST7E, with very little vignetting), so I often skipped them. But then my V-filter got dirty. Of course, I didn't know it at the time, as I was doing remote observing and didn't look at the equipment for weeks or months at a time. But image quality started deteriorating, and funny things started happening. For example, I would flip the GEM mount, and the star and references would show a changed relative brightness offset of perhaps .05 mag as they landed on different parts of the field.

Well, I thought, that just means that the effective sensitivity across the FOV is not constant, so I'll just do a flat to get rid of the effect. Well, that helped—sometimes. But other times the flat seemed to make things worse. So I began trying to think about flats, and what they measure, and what they don't. I must tell you that while I have made some headway in both thinking and experimenting, I do not have all the answers nor do I fully understand what is going on. But I will give you a hint of what is to come: I was baffled.

One of the first questions I tried to understand was what problems a flat can help. The purpose of the flat is to compensate for sensitivity variations across the FOV. You can have a sensitivity variation in the CCD itself (though I've never seen this), or it can be introduced by any variations in the light path from source to chip that are different from one part of the chip to another. There are two general effects we will discuss (1) dust doughnuts, (2) gradients or gradual sensitivity variations across the field.

Let's take the easiest case of something that will cause a variation of sensitivity at different parts of the chip—a speck of dust on the chip. But even that is not so simple: the speck is really not on the chip, it is probably on the cover slip that is on the chip. Furthermore, the speck is not opaque, so has some variable transmission. Let's look at a typical geometry for an ST7 on an F/6 scope as shown in Fig. 1.

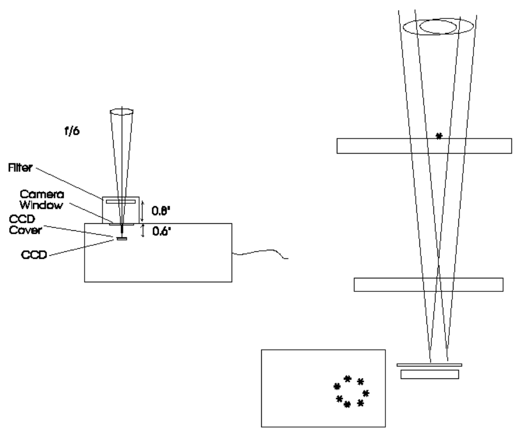

Figure 1 CCD Image Geometry

We talk about a "light cone", but really, every pixel is receiving its own light cone. But we can see that the speck will cast a partial shadow on the CCD as it intercepts each light cone, and the pattern of the shadows on the CCD will be a doughnut. The radius of the donut depends on how far away the speck is from the CCD, and the convergence (i.e., f number) of the light cone(s). You can measure the radius in pixels, multiply by pixel size and f number, and compute how far from the chip the dust is. The % of darkening or obscuration of the dust will depend primarily on the size of the speck, and on the f/number and so on. In general, the cone of light for one pixel is much, much larger than the dust speck where it crosses the speck so that even an opaque speck will cause only a partial blockage (attenuation) of the light.

We can actually visualize this process by putting a cover on the scope that is pierced by small (1/8") holes. For example, a single hole will create a very narrow cone of light that will project the shadow of the speck onto the chip (this forms, in effect, the light cone from an f/900 scope). More holes, even spaced, will show how the pattern builds into the familiar doughnut. Some of these test results are shown in Fig. 2. On the right is an image through one hole, on the left is through 18 holes. Geometry shows that these dust doughnuts come from the CCD window..

In most astronomy optical systems, what you will probably find is that the smaller dust doughnuts you see (the intense, little guys with a diameter of 5% or less of the FOV) are inside the camera on the cover slip. Dust on the camera window will show much larger doughnuts (e.g. diameters 10-30% of the FOV). Dust on the filters will form even larger, but fainter, doughnuts. Clearly, once you leave the region close to the chip, dust has a rapidly decreasing effect in terms of obscuration: most of the light just goes around it to get to each pixel

But the question still remains: how do we correct for the attenuation, pixel by pixel? If we can measure that attenuation pixel by pixel, then we can correct the measured intensity at that pixel, i.e., do a flat calibration.

So how do we do that? In theory, we could put a parallel beam into the scope that creates

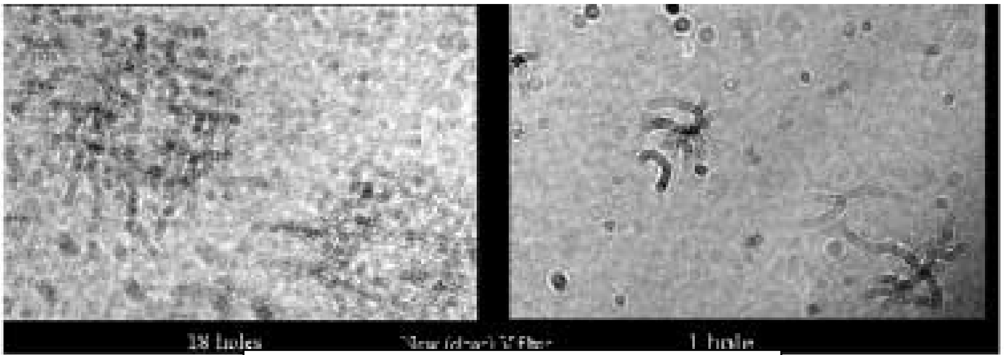

Figure 2 Doughnut Manufacture

a light cone the size of a single pixel. We could measure the sensitivity of that pixel, then move on to the next, and map out the whole system. This is probably the ideal way to measure a “flat”: we’ll come back to why that is so. But that is slow and difficult. So, instead of doing one pixel at a time, we do them all at once (so they each calibrate one another: We flood the chip with uniform light (at least, that is what we TRY to do), and measure the relative sensitivity of all the pixels at once.

The amazing thing is that this works so well: it is truly a quick and dirty technique. But with a little thought you can see immediately several problems with this approach. The major concern is that the flat is taken by flooding the field and these are VERY different conditions from how you actually take an image. Differences include

- Total amount of light hitting the CCD is different (usually much, much larger) and in a different pattern which may cause variations in response across the CCD (e.g., either from chip non-uniformity, or from electronic non-linearities as the electronics copes with the high readout values)
- With light flooding the field, the light reaching the speck affected pixel may be augmented by light scattered within the camera, or elsewhere in the optical system. This will affect the flat calibration accuracy.
- The scattered light reaching the camera may depend on the range of angles of light entering the telescope, the baffling of the telescope, and any field stops in the optical system. The range of angles of light rays entering the scope may be very different between Tshirts, dome flats, and sky flats. For example, the mix of light ray angles generated by imaging –i.e., in-focus- a 2 foot white circle at 500 feet away is not the same vs. a way out-of-focus T-shirt on front of the scope or a sky flat.
- The flat light may be a different color from the image light

Then there are always statistics: the flats you take always have statistical errors, so there will always be at least some adverse effect on the noise of the flat calibrated images.

So, how come flats work at all? Well, for imaging, the demands are not too high. If you can get rid of doughnuts and obvious vignetting, you will likely be satisfied with the resulting image. After all, if there is a decrease in brightness in one part of the image, who is to know that it is not in the object? Also, the intensity range across most images is in the 100:1 range, so a few percent of smooth variation will not be apparent.

The situation, of course, can be very different for photometry. An error in the flat between the target and the reference star(s) pixel locations may cause no problem if the target and reference stays on the same pixels all night (or if there is a uniform gradient in the direction of their motion)—e.g., the target and reference will always be X.XXX mag different. However, if the gradient is not uniform, or if the target and reference stars move to different pixels (imperfect tracking, a GEM flip which may even interchange target and reference, movement of an asteroid) then an error will have been introduced by the flat calibration. You may detect that an error is present, but it may be very difficult to correct for it.

So, what to do? How can you know how good your flats really are?

The easy answer is to take your photometry data and treat it with and without flat calibration—and hope the answers are the same, thus probably showing that you are ok. You watch carefully for signs of error—again, both with and without flats. You work with a system that is as clean as you can (to reduce doughnuts), and that has good baffling to reduce stray light.

The hard answer is that you can do experiments similar to what I will describe. Over the past several years, I have done the following sets of tests to try to understand what is happening:

- Compare Tshirts (or sheet on scope), dome flats, and sky flats
- Perform field tests in which I presented a well defined cylinder of light (first from a 50 gal drum, then a much fancier octagonal light source—which did no better) into the scope, i.e., control the angular spread of incoming light. I can compare this at very close distance (e.g.,, 20 ft) with long distance (120yd). Making a 30 inch diameter flat source has been an interesting project, but one that so far has not reached the 1% flatness I was seeking.
- Perform basement tests using sheets and lights in various configurations (distances up to 50 ft) both with the camera alone, and with the camera on a variety of telescopes
- Tests using true star fields to evaluate the uniformity of the CCD response, and to compare to flats.

Virtually all outdoor tests were performed with the ST7E, CFW8, on a C11 operated at f/6. I obtained and analyzed the images using MaximDL, which provides easy to use graphing and analysis tools.

In any test, there is always the question of whether a given artifact (e.g., a gradient across the field) arises from the flat source (or its lighting), the scope, or the camera. To try to identify the source of an asymmetric artifact, one can rotate each of the components in the system and try to track the change in the artifact. Keeping straight what you are doing is a real challenge (and even harder, is reconstructing what you did six months later). You may have the camera, filter, scope, scope hole pattern, flat, and flat lighting to rotate. Is that six test alternatives, or 2^6=64, or only 32. Well, it depends. But regardless, it is a pain because rotating some of these items is easier said than done. Also, errors that are symmetric about the optical axis cannot be teased out just using rotations.

As I noted, I had compared dome flats, Tshirt flats, and sky flats. They produced similar results, but they had disturbing variations under different conditions which I could not identify. For example, sometimes there were 4% gradients across the FOV, other times is was only 2%. While none of these are terrible for .05mag photometry, they did show that I did not understand my system.

So I stripped the system to the essentials. Working in my basement, I set up the experiment shown in Fig. 3. I aimed the camera through a 3” D by 36” tube fitted with field stops as shown. This was aimed at a white sheet five feet away, which was illuminated by a light as shown. I could put a diffuser (piece of paper) at A (on the camera nosepiece) or at B (end of the tube). (I also did other experiments, but will only report on this set).

With a diffuser at B, we have a sort of f/36 light source: the largest angles of the light with respect to the axis are as if it were f/36. With such simple geometry, I should be able to understand what is going on.

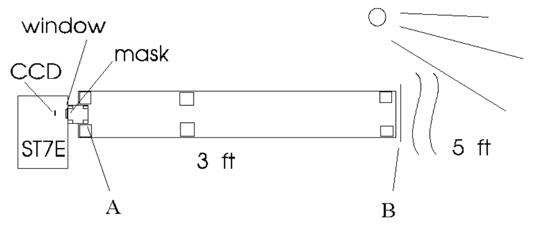

Figure 3 Diffuse Light Source Test Setup

Fig 4A (upper left) shows a terrible vertical gradient of more than 5%, with Fig. 4B showing a substantial gradient across the middle of about 1.5%. This seemed terrible! What is going on? Remember, this is an f/36 like system—the "shadow" in Fig 4A is NOT the familiar case of the guide chip mirror

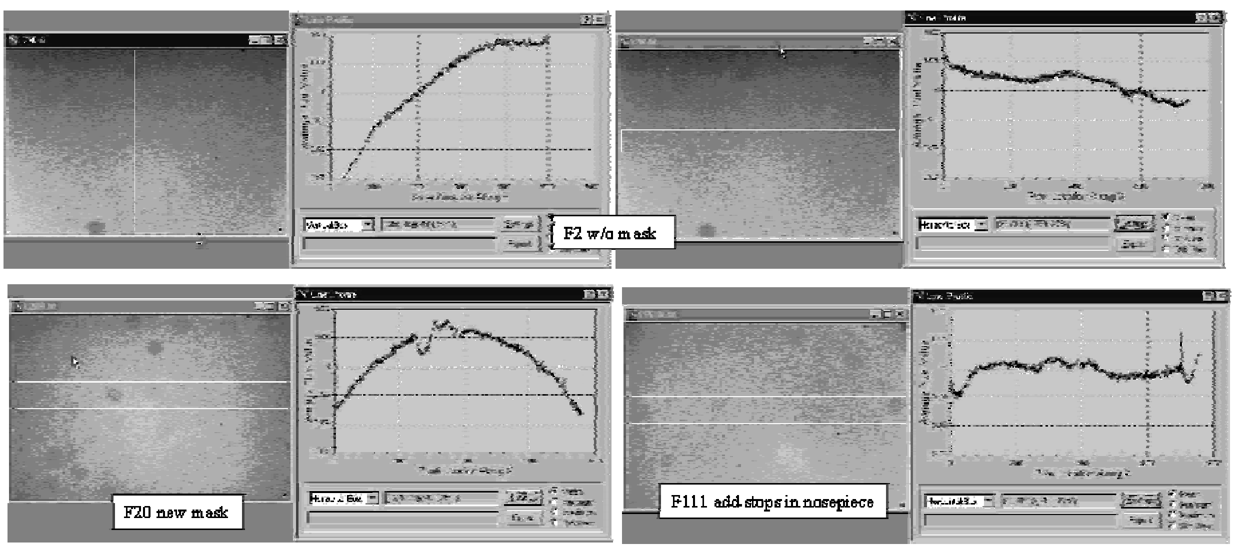

Figure 4 Diffuse Light Source Tests

Figure 4 shows some of the results. The top pair 4A&B are for the unmodified camera. In this and following figures, I have normalized each image to its total brightness so that we can easily compare sensitivities. Note that a brightness change from 1.01 to 1.02 is a 1% (.01 mag) difference. Also remember that positive Y is downward. The Fnn designation is my own nomenclature for my images (not f/number).

shadowing the image which can produce a very narrow dark band at the top of an image taken with a fast optical system. I had also seen similar behavior as this in my flats taken with the C11 at f/6 and even f/10.

After much experiment, thought and false steps, I postulated that it might be due to scattered light entering the chip. I installed a simple mask, cut large so as not to cut off f/6 light rays. The result is shown in Figure

4C—a major change! The vertical gradient virtually goes away, and the horizontal gradient is now much more regular.

Clearly, baffling is important, and this is proof that at least in my ST7 there can be a major problem with scattered light. What is it doing? It is entering the camera, bouncing around, and changing the pattern of light on the chip itself. That is, the scattered light contributes to the flat. And what is wrong with that? What is wrong is that it causes the flat calibration process to produce errors in the corrected images! This is because the stars we usually measure are superimposed on a very weak background of similarly scattered light which causes no problem because we are doing aperture photometry. But the flat calibration introduces a calibration variation that varies across the FOV, and which does not in fact represent actual, varying sensitivity. Which is what I had suspected.

As I looked at Fig. 4C, I realized I now did not understand the bell curve: was it the mask cutting off the edges of the light cones? But no—the mask should not have been doing that in this f/36 system. Again, after more thought, I installed small stops in the camera nosepiece which gave Fig 4D. The bell went away, and the curve is flat to within 0.5%. Clearly, some of the f/36 rays were reflecting off the interior of the nosepiece or part of the camera into the central area of the FOV.

The basement tests (which included test with several telescopes, and use of a diffuser at A) helped me understand flat processes, and showed me that I was having a real problem with scattered light which the mask helped. Obviously, in a well baffled, high f-number system, these effects will be small. But how could I actually prove how good my system in the observatory really is?

I then did star tests. Remember, I said it would be ideal to run tests on each pixel, one at a time? Well, I didn't do that, but I could do a subset of those tests using stars.

One way to do this would be to take an image of a star field (using the C11 at f/6, mask on camera), then translate and/or rotate the scope/camera and evaluate the measurements to determine the spatial pattern of response across the image. Doing this thoroughly for even 10-20 stars would be a lot of work, so I decided to try just two stars. I picked two stars of similar brightness about 5 a-min apart on an E-W line (there aren't very many!). My FOV is about 12 a-m wide. I could thus center StarA image, then center StarB. When centered, A & B in the two images are at the same spot on the chip, so have the same calibration which we use as the reference. We can thus flick the scope back and forth and build a sequence of images. We can measure the relative brightness of the star pair Left, then Right, to see whether there is a consistent difference. While this won't detect a constant gradient, it will identify a symmetric error or non constant gradients. I then rotated the scope and camera 180deg, and repeated the test. I then rotated the camera on the already rotated scope, and retested.

Even using multiple images, the scatter in the data limited the accuracy of the measure of gradients to about 0.5%. However, the data did show that there was less than about this much (0.5%) variation across the image. Great! Although the measurements could not rule out that there are still some issues of scattered light, there is clearly no systematic problem of calibration, at least in the horizontal band studied.

I then did a final series of flat tests including a Tee shirt flat through my Red filter of the daytime sky, a night-time sky flat (using local light pollution), and sky flats with the tube/camera rotated, then only the camera rotated back.

When I plotted the results, I found that the curves are not as nice (smooth) as those from the basement tests although they were generally consistent. The deviations showed I still have some scattered light entering the chip, but far less than when I started.

Because the star measurements showed less than 0.5% gradient errors, if I used typical flats with 0.5-1% variations, I would apparently be introducing errors! However, it is also clear that all these effects contribute less than .01mag error across the chip (now that I have the mask on), so I can ignore them. And, of course, using the flats would remove doughnuts that might be more than 0.5% errors. I may or may not use a flat, but at least I DO know what some of the limits on the system are! And I think I understand doughnuts.

Flat Summary

At least for my camera, scattered light can cause systematic errors in flats, and introduce systematic errors into your data if you use those flats in calibration. Use of wide angle flat sources (sky flats, Tee shirt flats, perhaps dome flats) in fast systems are likely to be at greatest risk of this effect. Using star measurements, you can verify at least to a limited degree the actual performance of your system. In faster systems, you may not notice the dust doughnuts on the flats taken through filters, so you should inspect and clean the filters regularly.