Reduction and Analysis Techniques for Infrared Imaging Data

Mark McCaughrean

NRC Resident Research Associate
NASA Goddard Space Flight Center, Code 685, Greenbelt, MD 20771

1 Introduction

Infrared detector arrays are becoming increasingly available to the astronomy community, with a number of array cameras already in use at national observatories, and others under development at many institutions. As the detector technology and imaging instruments grow more sophisticated, more attention is focussed on the business of turning raw data into scientifically significant information.

Turning pictures into papers, or equivalently, astronomy into astrophysics, both accurately and efficiently, is the subject of this paper. I shall discuss some of the factors that can be considered at each of three major stages; acquisition, reduction, and analysis, concentrating in particular on several of the questions most relevant to the techniques currently applied to near infrared imaging.

2 Data Acquisition

Effective reduction and analysis begin with effective data acquisition. As infrared cameras have become available, observing techniques have been developed that combine features of conventional infrared photometry and raster scanning, optical CCD imaging, as well as new procedures to account for unique problems presented by imaging with small detector arrays at infrared wavelengths.

2.1 Source

Due to the bright sky background, broad band near infrared images become background limited quickly, on the order of a minute. Only very shallow survey projects will require less time on a given field. For deeper experiments, the telescope can be moved a few arcseconds after each background limited exposure is read out, before starting the next—this makes it easier to replace bad pixels, and reduce noise due to fixed, time-independent noise sources such as flat fielding errors. If the source fields are mainly blank, and sky images are being used as flat fields, this method also makes it possible create flat fields from the source data itself. This eliminates the need for separate sky measurements, and effectively increases the on-source integration time by a factor of two (see section 3.2).

2.2 Sky

Classical AC-coupled infrared photometry involves spending half the time measuring the sky; how important is it to measure the sky when using an imaging camera? Sky images have two main uses—measuring the sky brightness as in conventional photometry, and to provide a flat field. The infrared sky background can be roughly divided into three regimes:

- $\lambda < 2.5\,\mu$m : dominated by emission from the hydroxyl molecule OH*—highly structured, both spatially and spectrally, with peak-to-trough intensity variations of up to 50%, on time scales of less than an hour [1].

- $2.5\,\mu$m $< \lambda < 10\,\mu$m : thermal emission from telescope/warm optics dominates—thermally stable, so background is relatively stable.

- $\lambda > 10\,\mu m$: thermal emission from telescope/warm optics and atmosphere roughly equal—fairly rapid variations in atmospheric temperature and opacity make for less stable background.

Sky intensity monitoring

The near infrared sky background can change by up to 50% on a time sca $\div$ of about half an hour—therefore the sky brightness should be measured every 10–30 minutes. This may or may not require separate sky images:

- Blank field experiments—if any given source image is more or less empty, there is enough sky information in the source data alone: the median value in the image provides an accurate estimate of the sky brightness at the time it was obtained. Small time scale sky variations are almost irrelevant; all pixels in the array integrate simultaneously, and the total sky background contribution should be the same for each.

- Filled field experiments—if the source fills the field of view, or you are making a mosaic of a large region filled with low level nebulosity, some measurement of the sky away from the region is required. Current array detectors are small—it is usually possible to find a blank patch of sky near to most sources. In some places (e.g. the Galactic Centre), it may be harder to find a blank patch of sky—as array detectors get bigger, it will be hard to find blank sky anywhere. As long as a patch of sky can be found that fills less than one third of the array pixels with sources, the median value will provide a good estimate of the sky brightness. Time interpolate between adjacent sky measurements to calculate the brightness at the time when a given source frame was taken. However, sensitivity to any short time scale variability will remain function of the time elapsed between sky measurements. If short time scale sky variations dominate the experimental errors, a modified form of chopping can be used. The alternating sequence of source and sky images is co-added in separate sections of the instrument computer. Reading out and storing both co-added images once every few minutes, the effects of short time scale variations are reduced—both images contain a similar cumulative record of background variations at speeds up to the chopping frequency.

Taking skies for flat fielding

Using sky images for flat fielding is an effective way of obtaining very low residual noise. In blank field experiments, where objects are rare, each image is mostly filled with sky, and median stacking techniques will filter out the sources to provide a high accuracy flat field. No separate sky measurements are required. For filled field experiments, separate skies will be required. If no completely blank field can be found, images taken at slightly offset positions can be median stacked to remove sources.

In principle, the actual intensity of the sky background should not be important, and sky images taken many hours before or after a source image could be used to flat field it. In practice, changes in the near infrared sky brightness can seriously affect the quality of flat fielding—the changes in brightness are accompanied by changes in the OH* emission line ratios. Integrated across a broad band filter, the effective 'colour' of the sky background is changed on time scales of an hour or so. If separate sky images are being used to flat field source images at broad band near infrared wavelengths, they are best obtained at least once every half hour. This variability is of even greater concern in higher spectral resolution.

2.3 Additional considerations

As well as imaging the source and the sky, several other factors should be considered at the telescope:

- Instrumental offsets—measurements of system offsets may be required, including dark current and bias. The frequency of measurement will depend on the stability of the detector system. If a dark current image is used in the reduction procedure, the on-chip integration time should be the same as used for source and sky measurements, but a larger number of dark frames should be co-added, to minimise its contribution to the noise. Dark current frames can often be taken in twilight.

- Flat fields—to remove pixel to pixel variations in quantum efficiency and gain, a flat field image is needed. Typical flat fields include images of the sky close in time and space to the source frame; images of the sky at high and low airmass; images of the Moon; images of the inside of the dome; images of the twilight sky. The signal to noise in the flat field images should be significantly greater than in the source images, so that

Figure 1: A mosaic scheme

the flat fielding process does not degrade the signal to noise in the reduced data. This is easy for flat fields which can be obtained during daylight or twilight hours, but the time spent on those that must be obtained 'on line', i.e. during observing hours, must be carefully examined. The flat fielding process is discussed in section 3.2.

- Standards—to flux calibrate the data, images of standard sources will be required. The frequency of measurement will depend on the required accuracy. A range of standards is required to measure atmospheric extinction as a function of airmass, and photometric zero-points and colour terms. Standards can also be used to check system linearity. Standards are discussed further in section 4.2.

- Linearity calibrations—if the detector is known to exhibit non-linear behaviour, calibration data may be required to measure and remove it. Usually done for only once for a given detector—however, non-linearity can be quite sensitive to operating conditions such as gate voltage. For the highest precision work, linearity calibration data can be obtained in twilight each night. Non-linearity is discussed in more detail in section 3.1.

- Mosaic scheme—if a large scale view of a region is required, a mosaic scheme is needed. For example, figure 1 shows a fully overlapping grid of image positions, in which every point on the sky is imaged twice. This scheme has many advantages, as discussed in section 3.3.

If a mosaic is made through two or more filters, it may be worth imaging each mosaic position through each filter before moving to the next position: relative differences in seeing, airmass, and sky background are minimised. In the case of bad weather, some fraction of the mosaic will be complete in all filters, rather than all of the mosaic in just one. If the results of the experiment rely on multi-colour photometry, this is the preferred outcome.

3 Data Reduction

There are many subtleties to the process of reducing infrared imaging data, and more will undoubtedly be uncovered as the data are subjected to increasing scrutiny. Following are some key points.

3.1 Linearisation

The so-called Direct Read-Out (DRO) type multiplexer is common amongst the current generation of devices. The DRO uses a single source-follower for each detector, and the whole array is addressed sequentially via X and Y shift registers. This is simpler to fabricate and quieter to operate than a surface channel CCD, but exhibits a a drawback, namely non-linearity.

The DRO works as follows: the unit cell is reverse biased, to charge ip the cell capacitance. Dark current and photo-electrons formed in the detector substrate discharge the cell—measuring the voltage on the cell at the beginning and the end of an integration period, the total voltage discharge can be determined. The instantaneous rate of voltage discharge is determined by the current flow and the cell capacitance—but, a large fraction of the cell capacitance is junction capacitance, which changes as a function of the bias across the diode. Thus, as the cell discharges, the voltage discharge achieved by a unit of input current changes. This makes it difficult to relate the total voltage discharge to the total current integrated. Also, the amount of dark current generated depends on the changing bias across the diode junction, adding to the problem.

A simple analogy

The explicit semiconductor physics behind the operation of a DRO-type detector is covered in detail elsewhere [2, 3,4,1]. In practice however, we are less concerned about *why* the DRO is non-linear, than *how* the non-linearity appears in practice—a simple analogy may be drawn as follows. Figure 2 shows an optical CCD operating as a bucket, with current represented by water. Exposed to light, photo-current is collected in the bucket, as is dark current. At the end of an integration period, we measure the depth to which the bucket is filled, i.e. the voltage change. Because the bucket has straight sides, the depth is directly proportional to the amount of current collected. To determine the dark current contribution, a 'lid' is placed over the bucket. The rate of dark current flow is the same as it was with the lid off—therefore it is easy to calculate the dark current contribution by an appropriate linear scaling.

Figure 3 shows the equivalent bucket model for a DRO detector. Things are more complex—the the bucket is immersed in water, and it has sloping sides and a hole in its bottom. Again, dark and photo-current fill the bucket, the former flowing in through the hole in the bottom of the bucket, and latter coming in from above. However, the final depth to which the bucket is filled (i.e. the voltage change) is not linearly proportional to the amount of current collected, due to the sloping sides. As the bucket is filled, it takes more and more current to increase the depth of the water by a unit amount. This non-linearity is equivalent to that caused by the change in the junction capacitance as the diode discharges.

The non-linearity in dark current generation is seen when the bucket is covered. The dark current flows in through the hole in the bottom of the bucket—it flows fastest when the bucket is empty; as the bucket fills, the dark current flow decreases, stopping completely when the level of the water in the bucket reaches the level of the water surrounding it. As the rate of dark current flow depends on the depth to which the bucket is filled, it is not easy to relate the amount of current collected during a dark integration to the amount collected when photo-current is also filling up the bucket.

This simple analogy can also explain another useful characteristic of the DRO circuit. When the optical CCD bucket fills up, it overflows, and current spills over into adjacent buckets, the process known as blooming or column bleeding. However, with the DRO, there is no blooming. When the bucket fills up, and passes the saturation point (where the level of the water is the same inside and outside the bucket), the photo-current tries to fill the bucket further. The dark current, however, now flows *out* through the hole in the bottom of the bucket, effectively stopping the bucket from overflowing.

How bad is the non-linearity?

The device that has drawn most attention to the problem of DRO non-linearity is the SBRC 62 × 58 pixel InSb+CRC228 DRO array. Figure 4 shows the measured non-linearity of an SBRC InSb+DRO device: the array
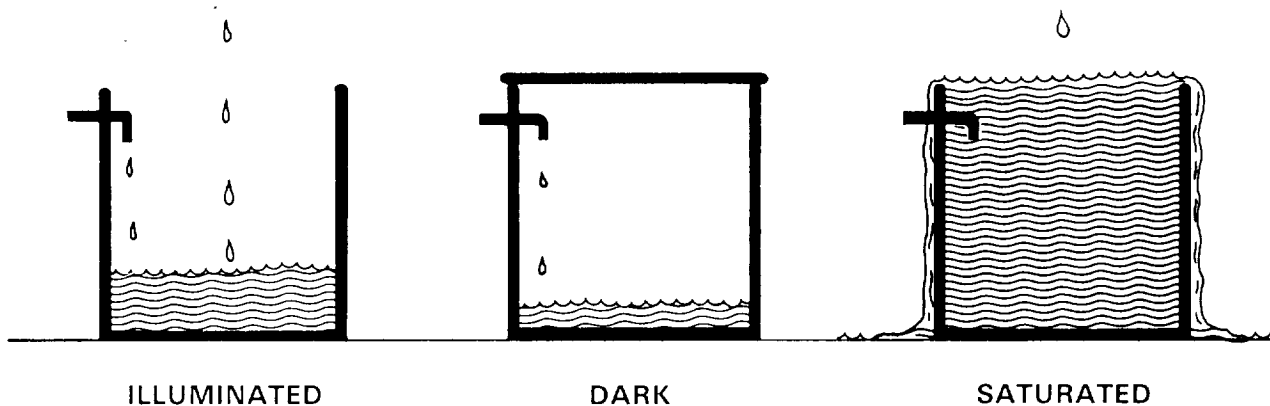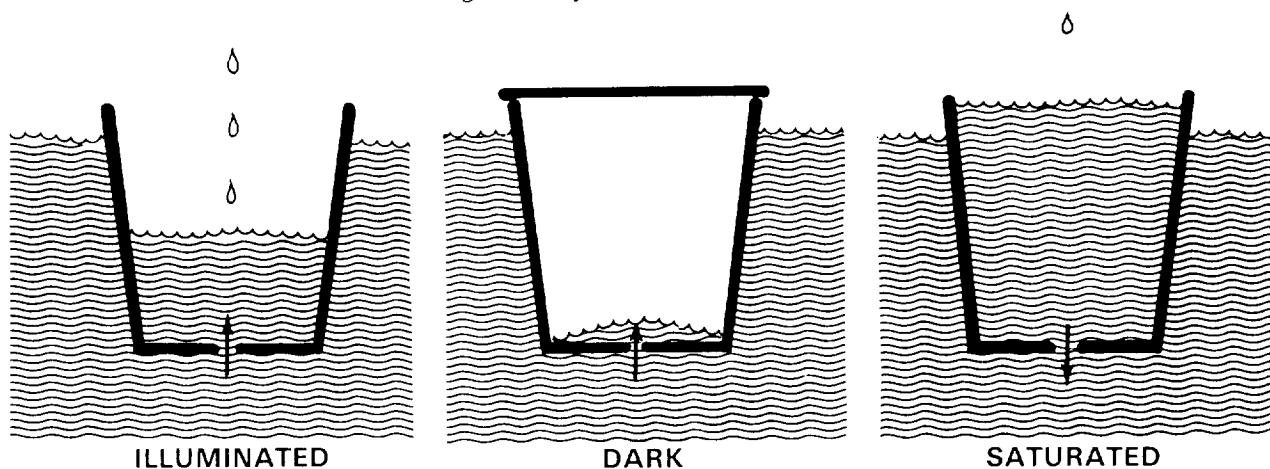
Figure 2: Optical CCD as a bucket

Figure 3: DRO as a bucket

was illuminated with a constant thermal background, by placing a piece of cardboard over the camera window and imaging through the K' (2.2 μm) filter. The mean output signal over a clean section of the detector is shown as a function of on-chip integration time. The non-linearity of the device is seen to be small and well behaved—a quadratic provides an accurate fit to the measured data. In fact, the measured non-linearity is a little less than predicted for the SBRC array [1]. This is probably due to parasitic capacitances in the actual device, unaccounted for in the modelling. These are independent of detector bias, and thus reduce the non-linearity.

The practical effect of the non-linearity depends on the dynamic range in the data, rather than how deeply the buckets are filled. The further apart in relative brightness two sources are *in the same image*, the larger the error in their measured relative brightnesses. In the worst possible case, with very little sky background, a very bright pixel close to saturation, and a very faint pixel close to zero, the ratio of the measured pixel brightnesses can be wrong by up to 10%, i.e. the bright pixel will appear about 10% fainter than it should when compared to a linear extrapolation of the faint pixel. Conversely, for two sources both close to a bright sky background, the non-linearity will be slight. An extensive discussion of the quantitative effect of DRO non-linearity under a range of typical observing conditions can be found elsewhere [1].

Removing the non-linearity

Removing the non-linearity of a DRO type device is not very difficult, and has been addressed in detail elsewhere, by Alan Hoffman of SBRC [4], and myself [1]—the non-linearity can easily be reduced to less than 1% across the whole dynamic range of the detector. There are some points worth noting about practical implementations of linearisation procedures:

Detector : FPA061 (low doped)
Filter : K (2.2 μm)
Imaging mode : 1.2"/pixel
Approx. photon flux : 3.5 x 10^5 γ/sec/pixel
Initial bias : − 250 millivolts
Initial dark current : 100 e⁻/sec/pixel
Gain : 24 e⁻/DN
Polynomial degree : 2
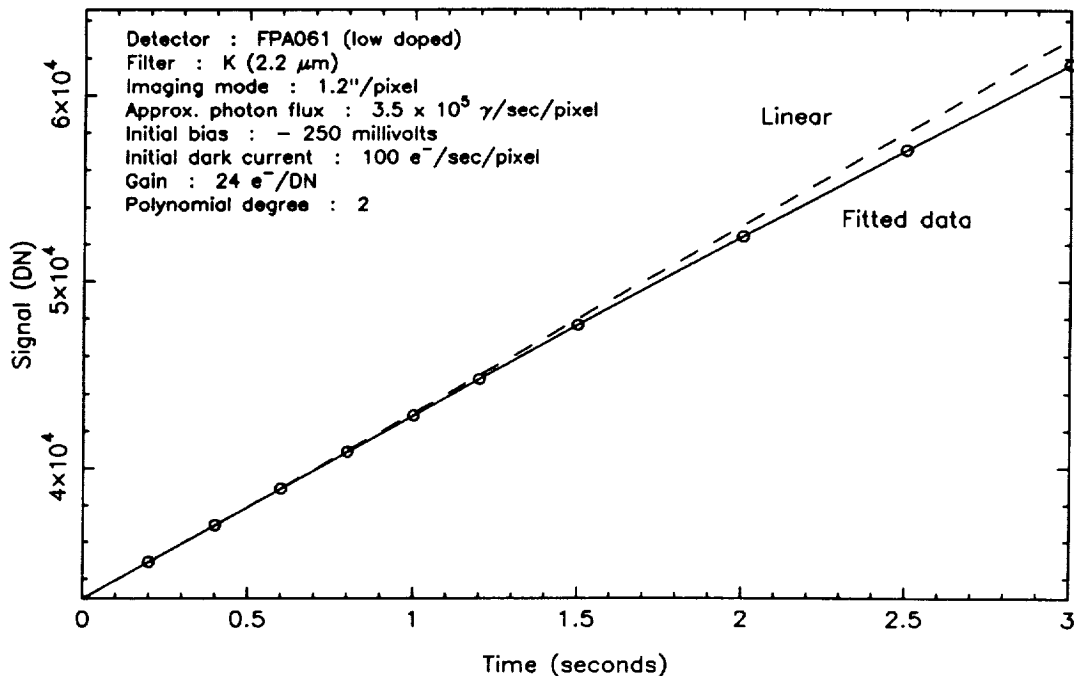
Linear

Fitted data

Signal (DN)

Time (seconds)

Figure 4: Measured non-linearity of the SBRC InSb+CRC228 DRO

- A numerical approach—make no *a priori* assumptions about the physics of the device, and measure the non-linearity in a controlled experiment involving constant brightness sources across a wide range of parameter space.

The dark current is commonly assumed to be very small, that its non-linear behaviour can be ignored, and that subtracting a dark current image will remove it completely. This simplifies matters considerably. Only the shape of the curve as shown in figure 4 needs to be determined, i.e. the deviation from direct proportionality between the integration time and the output signal. Each pixel is linearised by applying a polynomial correction as a function of well depth.

If the dark current is not negligible (e.g. in low background spectroscopic experiments), the independent dark current non-linearity should be accounted for. This can be done by taking a series of blanked-off exposures with increasing integration time, as well as the corresponding set of illuminated exposures. The two curves are combined to remove the effects of both non-linearities [4,1].

- An analytical approach—assume that the device physics can be modelled, and that the non-linear behaviour of the device can be predicted from a few key measurements of the parameters of a given device.

There is often considerable uncertainty in parameters such as the doping levels and the fixed capacitances. Also, simplifying assumptions about the nature and temperature dependence of the dark current, and whether the p-n junction has a step or graded profile, may be required. These assumptions can lower the accuracy of the resulting linearisation algorithm.

- A "don't" approach—assume the non-linearity is negligible and ignore it; this technique involves the least work. Frequently, sources being measured are faint and close to the sky background, and therefore non-linearity *is* negligible. However, standard stars measured to photometrically calibrate the images might be bright enough that non-linearity in *their* measurement becomes a factor.

- One curve or many?—a single calibration curve is often used to linearise an entire array. Deriving a curve for each individual pixel is seen to be a more accurate technique, resulting in slightly improved signal to noise after flat fielding.

3.2 Flat fielding

Of all the reduction steps applied to imaging data, the flat fielding process remains the most arcane. Optical CCD users have been tackling this question for more than a decade, with no 'standard' methodology yet agreed on. The steps taken to flat field data can depend heavily on the goals of the experiment—one project might require obtaining the lowest possible noise across limited blank patches of sky in a search for extremely faint sources; another might be concerned with photometric repeatability across the whole detector for accurate crowded field point source photometry. The techniques used to achieve these disparate goals might not be the same.

Some possible flat fielding techniques

We can use simple algebraic representation of the factors determining how source photons arriving above the atmosphere are turned into raw numbers in our computer, to examine a few of the flat fielding techniques currently 'in vogue' in the infrared imaging community. First the definition of some variables, noting that the majority, except (hopefully) the on-chip integration time, will differ from pixel to pixel:

S — the extra-atmospheric source flux
σ — the sky flux at the position of the source
σ_1 — the sky flux at a low airmass (e.g. the zenith)
σ_2 — the sky flux at a high airmass (e.g. sec z = 2)
T — the total warm optical throughput (including sky and telescope transmission)
τ — the telescope and warm optics background (independent of airmass)
η — the total cold optical throughput (including transmission losses and detector QE)
I_d — the detector dark current
t — the on-chip integration time
G — the electronic gain
O — the electronic offset

Next generate some 'images' by combining the variables appropriately. Note, we are considering a linear system, or at least one in which the data can be made linear, following one of the prescriptions described in the previous section:

$$\text{Source frame} = ((((S+\sigma)T+\tau)\eta+I_d)tG)+O$$
$$\text{Sky frame} = (((\sigma T+\tau)\eta+I_d)tG)+O$$
$$\text{Low airmass frame} = (((\sigma_1 T+\tau)\eta+I_d)tG)+O$$
$$\text{High airmass frame} = (((\sigma_2 T+\tau)\eta+I_d)tG)+O$$
$$\text{Dark current frame} = I_d tG+O$$

Next consider some of the ways to retrieve the source flux S from some combination of these images. Implicit in the algebra is the fact that we are manipulating the images on an individual pixel basis.

- Optical reduction: A simple version of the technique used for optical CCD data. First, the additive dark current and electronic components are subtracted from the source and flat field frame. Then the source frame is divided by the flat field frame, in this case, an image of blank sky. The critical assumption made here is that the flat field (i.e. the sky) is indeed flat across the field of view (FOV) of the detector:

$$\frac{(\text{Source} - \text{Dark})}{(\text{Sky} - \text{Dark})} = \frac{(S+\sigma)T+\tau}{\sigma T+\tau} \tag{1}$$

At $\lambda < 2\,\mu\text{m}$, $\tau \sim 0$, i.e. there is no flux from the telescope and other warm optics ($\sim 300\,\text{K}$)—the result reduces to :

$$\frac{(\text{Source} - \text{Dark})}{(\text{Sky} - \text{Dark})} = \frac{S+\sigma}{\sigma} \tag{2}$$

Usually the flat fielded image is normalised. Multiplying by the mean of a clean area of the (Sky − Dark) image, i.e. $\overline{\sigma T \eta t G}$, images flat fielded by different sky frames can be compared. Noting that σ and $\overline{\sigma}$ cancel

207

because we have assumed the sky to be flat across the FOV, we obtain:

$$\frac{(\text{Source} - \text{Dark})}{(\text{Sky} - \text{Dark})} \times \overline{(\text{Sky} - \text{Dark})} = S\overline{T\eta tG} + \overline{\sigma T\eta tG} \tag{3}$$

Finally, the sky background is removed. We assumed the sky was flat across the detector FOV, thus we can subtract a mean value, determined either from the source frame itself (if it is not crowded and does not contain low level nebulosity), or from the blank sky frame, if it was taken nearby in space and time. In either case, the mean sky value is algebraically equivalent to $\overline{\sigma T\eta tG}$

$$\frac{(\text{Source} - \text{Dark})}{(\text{Sky} - \text{Dark})} \times \overline{(\text{Sky} - \text{Dark})} - \overline{(\text{Sky} - \text{Dark})} = S\overline{T\eta tG} \tag{4}$$

For each pixel we have the source flux falling on it multiplied by the mean system throughput and the integration time. The integration time is trivially removed by division, while the system throughput is 'removed' by comparing a given source measurement with that for a standard source, whose absolute brightness is known.

At $\lambda > 2\,\mu$m however, the telescope and warm optical emission τ is *not* zero. Returning to equation 1, the flat field component of the equation becomes $(\sigma T + \tau)$. If τ is flat across the FOV, then the algebra can be retraced to arrive back at equation 4. But, if the telescope and warm optical emission is *not* flat across the FOV, the process is invalidated. As these components are not in the same focal plane as the source, effects such as vignetting can lead to changes in thermal illumination across the FOV.

• Thermal reduction: Used at longer wavelengths, where the telescope and warm optical emission τ is the dominant component of the background, and where τ may well not be flat across the detector FOV. The atmospheric part of the thermal background is a function of airmass, and the telescope/warm optics part τ should not be. Therefore, a flat field is made by subtracting one sky frame taken at low airmass from another sky frame taken at high airmass, removing τ from the problem. The resulting flat field should be flat, if the *a priori* assumption that the sky is flat holds true. The source frame has its sky, τ, and other system offsets removed by subtracting a nearby sky frame; the result is divided by the flat field frame:

$$\frac{(\text{Source} - \text{Sky})}{(\text{High Airmass} - \text{Low Airmass})} = \frac{S}{\sigma_2 - \sigma_1} \tag{5}$$

This is normalised, multiplying by the mean of the flat field, $\overline{(\sigma_2 - \sigma_1)T\eta tG}$. This leads back to equation 4, again with the proviso that $(\sigma_2 - \sigma_1)$ for any given pixel is equal to $\overline{(\sigma_2 - \sigma_1)}$, i.e. that the sky is flat.

This technique is 'more correct' when the warm optical thermal background is a factor. However, the use of four images rather than three will result in slightly higher noise. Also, there may be large enough differences in the emission at low and high airmass (e.g. colour) that the whole technique might fail. Finally, if the thermal emission background structure *is* a function of airmass (e.g. due to flexure), the technique is further invalidated.

• Raw reduction : This technique is fast and easy—a sky frame is subtracted from the source frame:

$$(\text{Source} - \text{Sky}) = S\overline{T\eta tG} \tag{6}$$

The technique is incorrect, as it does not divide out any multiplicative pixel to pixel differences. However, it is a useful way of obtaining a quick look at the source data. In the limit of a blank field with no source flux, one sky frame is being subtracted from another; all the multiplicative factors and additive factors cancel out. All that is left is noise (shot, read-out, etc.)—thus it is an effective way of looking for minor perturbations near the background, i.e. very faint sources. It should not be used for photometric work, where the source flux is not close to zero.

This list of techniques by no means covers the whole range of possible approaches. For example, a modified version of the 'thermal reduction' can be used at non-thermal wavelengths: a raw sky frame is subtracted from the raw source frame, effectively removing the additive components, such as system offset, dark current, and sky background. The result is divided by a flat field constructed from images of (say) the Moon or the dome, in order to correct for colour dependent effects in the detector quantum efficiency.

208

Some practical considerations

There are additional subtleties in the flat fielding process. For example:

- The flat field 'colour'—not only will the quantum efficiency be variable from pixel to pixel across the detector, it is also likely to be wavelength dependent. The two effects will probably be correlated if quantum efficiency is related to thickness of the detector material, i.e. the spatial structure of the flat field may be wavelength dependent.

  Therefore, an accurate flat field should not only be flat, but should also be the same 'colour' as the source: a very red source should be divided by a very red flat field, while conversely, a blue source should be divided by a blue flat field. 'Coloured' flat fields are created in the optical by imaging the inside of the dome, illuminated by tungsten lamps covered with a range of colour modifying filters. Also, the twilight sky is observed at different times relative to sunset and sunrise, as the colour of the sky changes [5]. Similarly, 'coloured' flat fields can be achieved in the near infrared: an out of focus image of a lunar mare provides a flat field with approximately solar colours [6], and illuminating the dome with fairly dim tungsten lamps ($T\sim$ 2000–3000 K) will provide an analogue for late type stars. It is more difficult to create a flat field with an equivalent colour of $J-K\sim5\overset{m}{.}0$, or one that matches the colour of the sky seen through narrow band filters, or a Fabry-Pérot étalon: flat fielding these very narrow band images can be tricky. Note that the nighttime sky background from 1–2.5 μm is virtually all line emission, with an effective integrated broad band 'colour' unlike any black-body or dust reddened source [1]. There is a certain circularity in choosing the colour of the flat field source to match the colour of your program objects—there will always be some colour related errors in flat fielding process.

- Stacking sky flats—if a number of sky images are obtained during the night for use as flat fields, they can be stacked to create a master flat, with higher signal to noise. At broad band near infrared wavelengths (J,H,K), making a master flat out of sky images taken up to an hour either side of the source frame can yield the expected reduction in residual noise after flat fielding. However, adding in data from outside this time frame, the residual noise starts to increase again. This is not unexpected—the the OH* emission component of the sky background can change by 50% on time scales of a half hour, and in addition, changes in sky background will occur as the telescope changes in airmass. Intensity changes are usually accompanied by colour changes, reducing the effectiveness of the stacked flat field. Long term drifts in instrumental parameters such as electronic gain, dark current, and detector temperature, can also make it undesirable to stack a whole night's sky flats. Little is gained by combining more than $\sim$ 5 individual skies to create a master flat—beyond that, only about another 10% reduction in residual noise will be obtained.

How accurately can we flat field?

The accuracy that can be achieved in the flat fielding process depends largely on the definition of "accuracy", and to illustrate this we shall examine two extreme cases:

- Residual background noise—in the case of deep searches for very faint sources, the aim is to reduce the pixel to pixel noise over relatively limited regions of the array, in order to spot small enhancements above the noise. Gross structures across the field are not important if they have a much larger scale size than the faint objects being sought. The limiting case is blank field, which is just an image of the sky background. The best flat field in this case is an image of another piece of blank sky, taken nearby in both space and time. In the limit, source frames can be used to flat field themselves. After each image is read out, the array is moved on the sky by an amount somewhat larger than the objects being searched for. Provided real sources fill no more than about one third of each image, a master flat field can formed by taking the median for each pixel through the images stacked directly on top of one another. Each source frame is divided by the master flat, registered up on common objects, and co-added to provide a master image. This approach has been used to locate very faint galaxy populations, both with optical CCDs [7], and near infrared arrays (McLean, private communication). A residual noise of less than 0.01% of the sky background can be achieved.

- Photometric uniformity—program sources and standard stars may fall on different parts of the detector array, and accurate measurement of their relative brightnesses will depend on the large scale uniformity, including vignetting, image ghosting, and colour sensitivity differences across the array. Intra-pixel variations

may also be important. Taking a large number of images of a standard star, each time moving the star to a slightly new position on the array, the total flux can be measured at each new position, and the degree of photometric uniformity across the array can be derived. Using a number of flat fielding techniques involving coloured dome flats and images of the twilight sky, uniformities on the order of 3% have been measured using this technique (Forrest, private communication).

Flat fielding techniques used to achieve the very lowest small scale pixel to pixel noise might not result in the best photometric uniformity, and conversely, the technique that best corrects for photometric non-uniformities might result in increased pixel to pixel noise in the background. Different techniques may be applicable depending on the experimental goal.

3.3 Mosaic making

Making large scale mosaics out of many smaller images is an effective way of showing the large scale structures in a region at seeing limited spatial resolution, previously almost impossible in the infrared. Although mosaic making should be reduced in importance with the advent of much larger format arrays, the use of $\sim 64 \times 64$ pixel arrays is likely to continue for a few years yet. Also with no equivalent to the combination of Schmidt telescope and photographic plate, the infrared astronomer will still want to make mosaics even when 256×256 pixel arrays are common.

Positional offsets

Few telescopes can offset with an accuracy equal to one seeing limited pixel ($\sim 0.5''$), and therefore, for accurate mosaicing, there should be overlap between adjacent images to allow registration on common objects. Point sources are most effective—centroids can be defined to better than 1/10th of a pixel if fully sampled. In the absence of point sources, extended regions of nebulosity can be used, but will result in lower accuracy registration.

The calculation of positional offsets is trivial: common sources can be manually identified on an image display, their centroids found, and offsets derived. Alternatively, more sophisticated automatic cross correlation procedures can be used, searching for the best fit offset around a box centred on the nominal offset position.

Sky background and transparency changes

These effects can be difficult to remove; one is additive, the other multiplicative. One possible approach is as follows: once the positional offset between a pair of images has been calculated, two intensities for any given point on the sky in the overlap region can be obtained, one from each image. By plotting pairs of values as x and y, a straight line should result (see figure 5). The intercept on the y-axis determines the difference in additive sky background, and the slope determines the relative change in the multiplicative transparency.

To produce a well defined straight line, a large dynamic range is needed, i.e. pairs of points measuring both faint and bright sources in the overlap region. If there *are* bright sources in the overlap region, they will usually be stars. Unless images are accurately registered to significantly less than one FWHM, stars will not exactly overlap, and the plotted value pairs will not be well correlated. Registering fully sampled images to less than one FWHM requires shifting to a fraction of a pixel—the interpolation required will propagate bad pixels. In many situations, only faint sources and nebulosity close to the sky background will be imaged, resulting in a much less correlated plot—see figure 6. Making an accurate linear fit in this situation is more difficult.

A less direct approach can be taken. Regular measurements of a nearby blank field for flat fielding purposes will also monitor changes in the sky background. Time interpolating, the background component of each source image is calculated and subtracted. Regular measurement of a nearby isolated standard star monitors transparency changes, removed by multiplication. Applying these externally calculated corrections, mosaic images usually are reasonably well matched—any small residual offset is assumed (perhaps erroneously) to be additive and is calculated from the mean difference between overlapping pixel pairs.

It is hard to write infallible automated software to perform these tasks alone. Once the first order corrections have been automatically applied, the resulting mosaic should be manually inspected for joins, and the necessary corrections applied to remove them. The human eye is very sensitive to edges—in practice, intensity offsets as small as 0.5σ can be spotted when mosaicing fields with gaussian noise distributions, standard deviation σ.
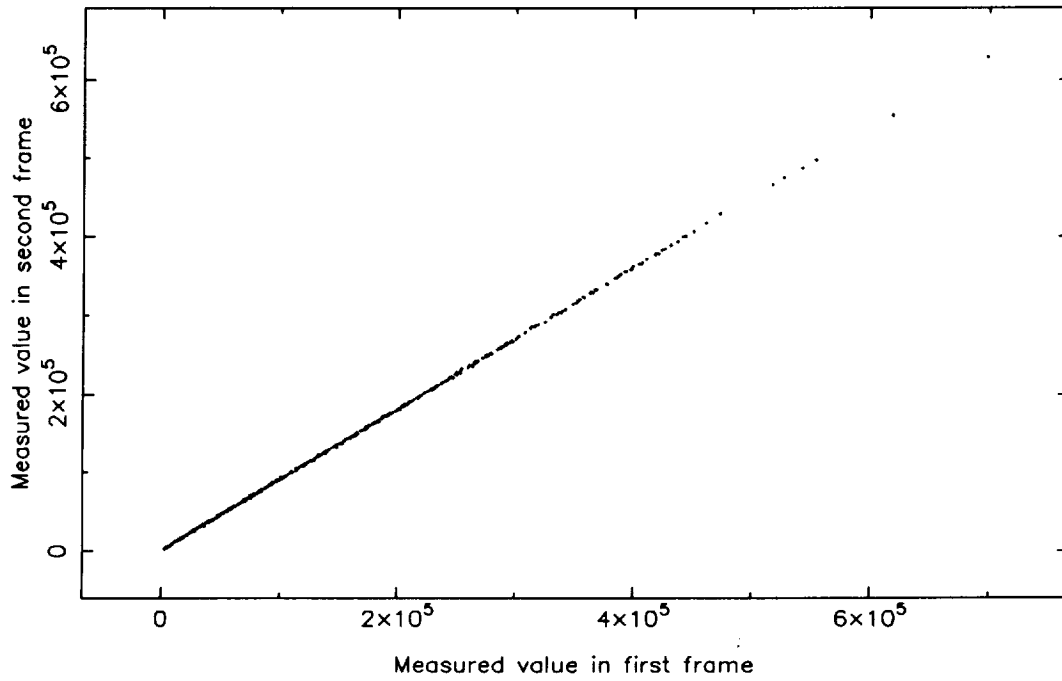
210

Sky transparency/offset calculation

Figure 5: Background and transparency matching—the ideal case

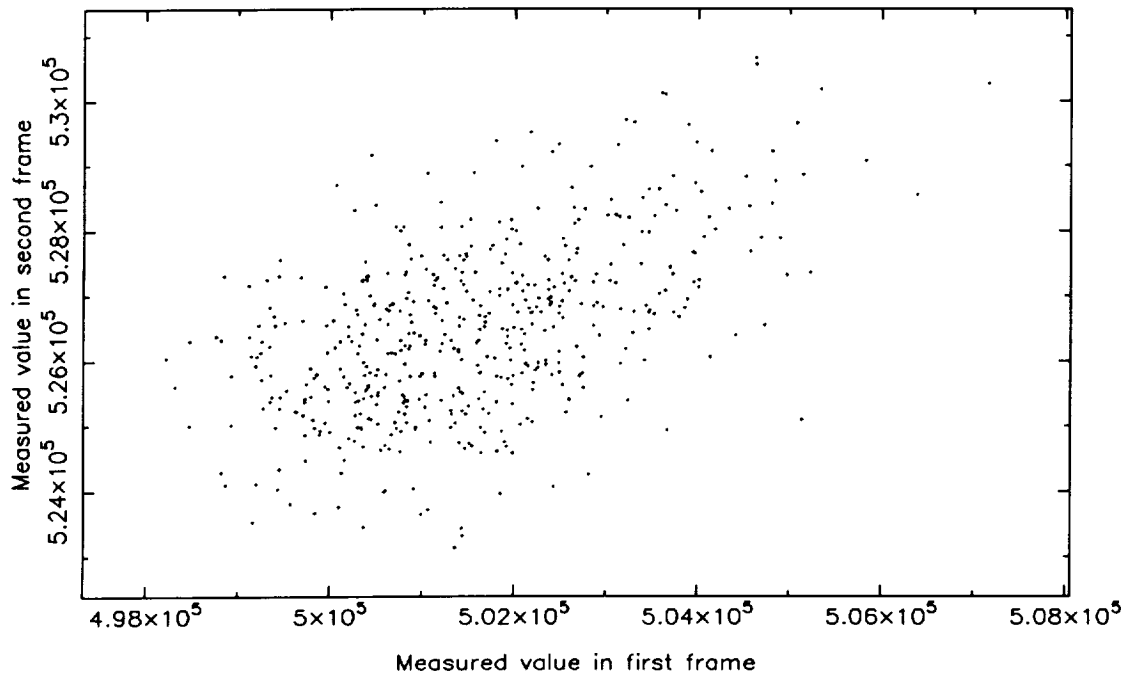Sky transparency/offset calculation

Figure 6: Background and transparency matching—a more typical case

211

Combining pixel values

To calculate the intensity of each point on the sky, the mosaic can be considered as a series of photographs laid down on top of each other, the value for each point just being the value of the corresponding pixel in the last image laid down. However, this wastes the information in any pixels 'covered up' by images laid on top of them.

A better method is to take the mean in the z-direction, i.e. for any given point on the sky, take the mean of the values in the corresponding pixels in all the images that cover that point. If two images both cover a given region, a $\sqrt{2}$ signal to noise gain would be expected; three overlapping images should yield a gain of $\sqrt{3}$, and so on.

If there is incomplete overlap, taking the mean and increasing signal to noise in the areas that *are* covered more than once may result in a patchy looking mosaic, some areas having higher signal to noise than others—a factor of $\sqrt{2}$ difference is readily discernable. A fully overlapping mosaic scheme as shown in figure 1 will avoid this problem.

The mean can be skewed by bad pixels, of which there are two kinds: fixed ones, which are always bad; and sporadic ones, which can crop up anywhere at any time. The former might be 'leaky', 'dead', or might not flat field properly: they can be catalogued, and a 'bad mask' or 'bad pixel list' used to flag them, and exclude them when calculating the mean. A sporadic bad pixel might be due to a cosmic ray hit or electronic glitch, and are less easily dealt with. If they are infrequent, they can be ignored, included in the mean, detected in the final mosaic, and then suitably replaced. This works well if the image scale fully samples the seeing—single pixel events are easily identified as bad. In undersampled modes where stars may appear as single pixels, this approach can be dangerous.

Taking the median rather than the mean of the pixel values in the z-direction will remove sporadic bad pixels. However, for a gaussian distribution, the mean is a better estimator of the true value than the median, and for small numbers of overlapping frames (< 10), the mean will result in a greater improvement in signal to noise.

Another form of 'bad data' can occur, more insidious than the isolated single pixel event. Infrared cameras have relatively complex optics, and there is a possibility of ghost images and other diffuse features being generated, perhaps near to a bright source, or perhaps far from it, depending on the optical scattering paths. Although these spurious 'sources' are often easily recognised by their peculiar shape, sometimes they are not. A fully overlapping mosaic scheme proves invaluable in detecting and replacing them—any any source that only appears in one frame of an overlapping pair can be considered spurious. Flagging the pixels containing the spurious 'source' as bad, good data from the overlapping image can be used to 'patch' the final mosaic. Mosaic schemes using only a small overlap make it more difficult to detect and remove these undesirable features.

An example

The best scheme will depend on the data being combined. Compromise approaches may be appropriate, with bad masks used to identify fixed bad pixels, a median technique to identify sporadic ones, and flagging to identify spurious 'sources'; after all types are rejected, the mean of the remaining pixels can be used to calculate the final value.

As an example, take a mosaic of 145 2.2 μm images of the Trapezium cluster in the Orion Nebula, obtained using the mosaic scheme shown in figure 1. Further details are given in section 3.4. Figure 7 shows the result of taking the mean at every pixel, with no regard to bad pixels. Even the fairly clean SBRC InSb+DRO array used for the experiment results in quite a mess when mosaiced. However, figure 8 shows the result of throwing a switch in the mosaic making software, telling it to account for known bad pixels. Bad data in one image was replaced with good data from another overlapping one; no replacement by artificial data has taken place. The effect is dramatic, with only a few bad pixels remaining, some scattered, and some near the edge of the mosaic, where each point on the sky was imaged once only—the effectiveness of a fully overlapping mosaic scheme is well demonstrated.

3.4 Overlaying multiple images

A 'true-colour' composite image can be created by superimposing three images taken through different filters, representing one each as blue, green, and red, usually in increasing wavelength order. The technique has been extensively applied in the optical, to both photographs [8] and CCD images [9], and has already found popularity
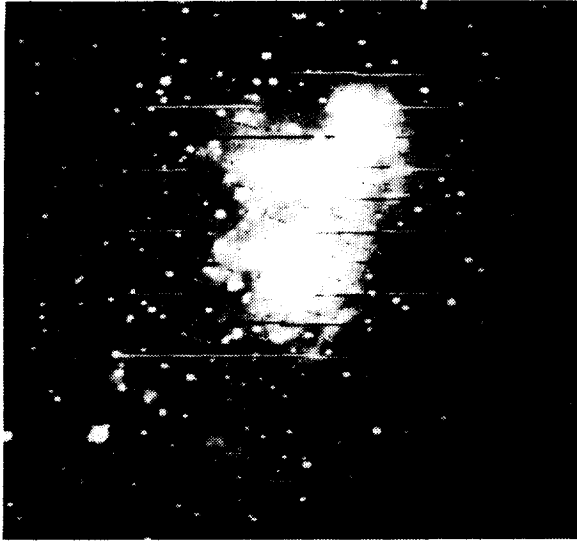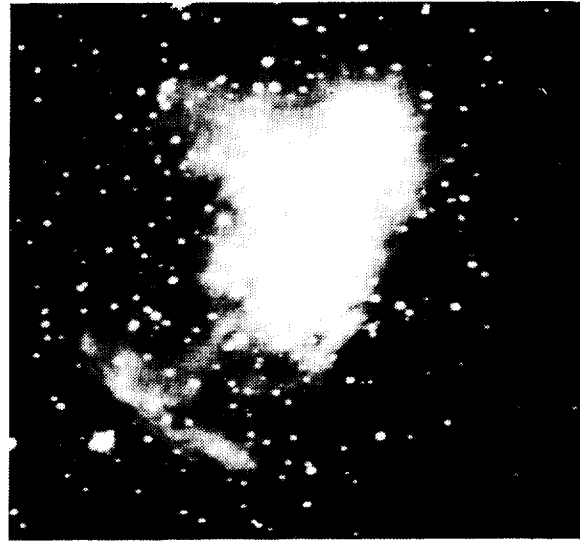
212

Figure 7: No bad pixel handling

Figure 8: Automatic bad pixel handling

in the infrared imaging community, as it can graphically reveal cool and reddened sources, and regions of increased extinction.

- Pixel scale—the pixel scale may be wavelength dependent if transmissive reimaging optics are used, and images taken at different wavelengths should be resampled on to a common pixel scale. With a number of bright point sources common to each image, the relative pixel scales can easily be determined—otherwise, a large scale mosaic can be used. At longer infrared wavelengths, few point sources are visible at all—the problem of determining absolute and/or relative pixel scales is harder.

The relative difference in pixel scales may only be ~ a few percent and is often overlooked, particularly if a image contains only one point source and some diffuse emission. Ignoring the pixel scale difference could lead to systematic errors when deriving the radial dependence of the nebular $J - K$ colour, for example—the images should be resampled to a common pixel scale first.

- Dynamic range compression—the dynamic range in infrared images can be very large, and effectively display-ing bright point sources and faint nebulosity in a single image sometimes requires non-linear compression. Logarithmic compression is often used, although square, cube, and higher roots of the data have been found empirically to give the necessary compression without reducing contrast quite as much. In some cases, tech-niques such as histogram equalisation may be useful—these account for the actual numerical distribution of bright and faint data points in a given image, redistributing them across the dynamic range more evenly.

- Intensity scaling—choosing the relative intensity scaling levels for the multiple images is as much an aesthetic issue as scientific. In the visible, the 'true-colour' technique is usually used to create a colour image of a source as it would be seen by the human eye if the retina was more sensitive to low flux levels, while maintaining the same general colour balance it has at high flux levels [8]. However, extrapolating the technique to the infrared, there is no point of reference for how an infrared sensitive eye might perceive 'colours'. For example, what should be forced to come out looking white in a composite? An object with equal flux in detected photo-electron units through each filter, or one with equal extra-atmospheric photon units, i.e. accounting for the total system throughput at each wavelength? Many regions imaged in the infrared have a large amount of dust extinction towards them—the dust could be effectively 'stripped away' by skewing the intensity scaling according to how much each filter is affected by extinction. The flux from stars is predominantly black-body continuum, and the 'colour' of a star in a composite can be physically meaningful. However, extended emission is often due to a complex combination of continuum and line,

213

reflected and intrinsic, molecular and atomic, and the amount of extended emission seen through each filter will have more to do with which lines happen lie within its bandpass, rather than some simpler diagnostic, e.g. the gas temperature.

An example

Each case must be treated separately. As an example of how a reasonable compromise may be reached, take the colour composite image of the Orion Nebula shown in figure 9.

The data were obtained by myself and Colin Aspin using the common-user $1-5\,\mu$m camera IRCAM on the 3.8 m United Kingdom Infrared Telescope, during the week 24–31 December 1987. IRCAM uses 62×58 pixel SBRC InSb+DRO array—additional details of this instrument are given elsewhere [10]. A 145 position mosaic scheme was used, as shown in figure 1: the image scale was ~ 0.62 arcseconds per pixel, and the total mosaic covers a 5×5 arcminute region centred on the Trapezium OB association. Each position was observed at J, H, and K before moving to the next. Nearby sky frames were taken after every third mosaic position, and standard stars were monitored frequently.

The data were reduced using the 'optical' technique described in section 3.2, with stacked sky images used as flat fields. Positional and intensity offsets were calculated from the image overlaps, and the sky and standard images. The full reduction procedure will be discussed elsewhere, as will the results of imaging photometry of the almost 500 members of the Trapezium cluster [11].

The three raw mosaics were overlaid on a 24-bit colour image display, one each in the red, green, and blue planes. They were registered at the mosaic centre (the brightest Trapezium star, θ^1 Ori C), and the relative pixel scales were calculated from positions of a number of stars around the edge of the mosaic. The three images were bilinearly interpolated to a common pixel scale.

Each image has a dynamic range of about 4000—displayed linearly to show the structure in the nebulosity, the region surrounding the Trapezium OB stars becomes completely saturated, and displayed to show detail in the Trapezium, the nebulosity is almost invisible. Taking the $\sqrt{\ }$ of each pixel intensity shows the Trapezium and nebular structure simultaneously, while retaining a reasonable amount of contrast in the fainter nebular structure.

To give a wide colour differentiation among the stars, the high intensity scaling levels were chosen such that a star came out white if its $J-H$ and $H-K$ colours were equal to the mean stellar colours of normal cluster members. These mean colours are roughly $1\overset{m}{.}1$ and $0\overset{m}{.}7$ respectively. The low intensity scaling levels were set to make the south-west corner of the composite come out black. While somewhat arbitrary, it has the effect of giving most of the nebulosity a bluish tinge, consistent with much of the emission in the region being hot thermal continuum from the ionised hydrogen gas, at an electron temperature around 10^4 K, and brightest through the J filter.

The chosen display levels give the desired effect—there is a reasonable spread in stellar colours; the hot OB stars appear blue-white; most of the optically visible stars, known to be mostly young late types with some reddening towards them appear white; and the 'infrared' stars (i.e. those not seen at optical wavelengths) appear various shades of yellow, orange and red, depending on stellar type, extinction, and evolutionary status. The HII region nebulosity appears bluish, while reflection nebulosity around the embedded BN-KL complex, and molecular hydrogen emission from beyond the bar and outside the HII region are seen as reddish.

4 Data Analysis

Having arrived at the 'final' image, the next stage is to determine the brightness of sources, point-like and diffuse, within the image. There are two distinct components to this process—the measurement of the 'instrumental' flux from the source, and the conversion of those units to a standard system.

4.1 Measurement of source flux

The technique required to measure the flux from a 'source' in instrumental units such as ADUs or electrons, will depend on the type of source, its surroundings, and the desired accuracy. Sources might be point-like (e.g. stars), or extended (e.g. reflection nebulae, galaxies); they might be isolated, or confused, in a crowded cluster; there might be just sky beneath the source, or it might be embedded in structured nebulosity. Much of the following discussion will concentrate on point source measurement—this is the hardest to do accurately. In this context,
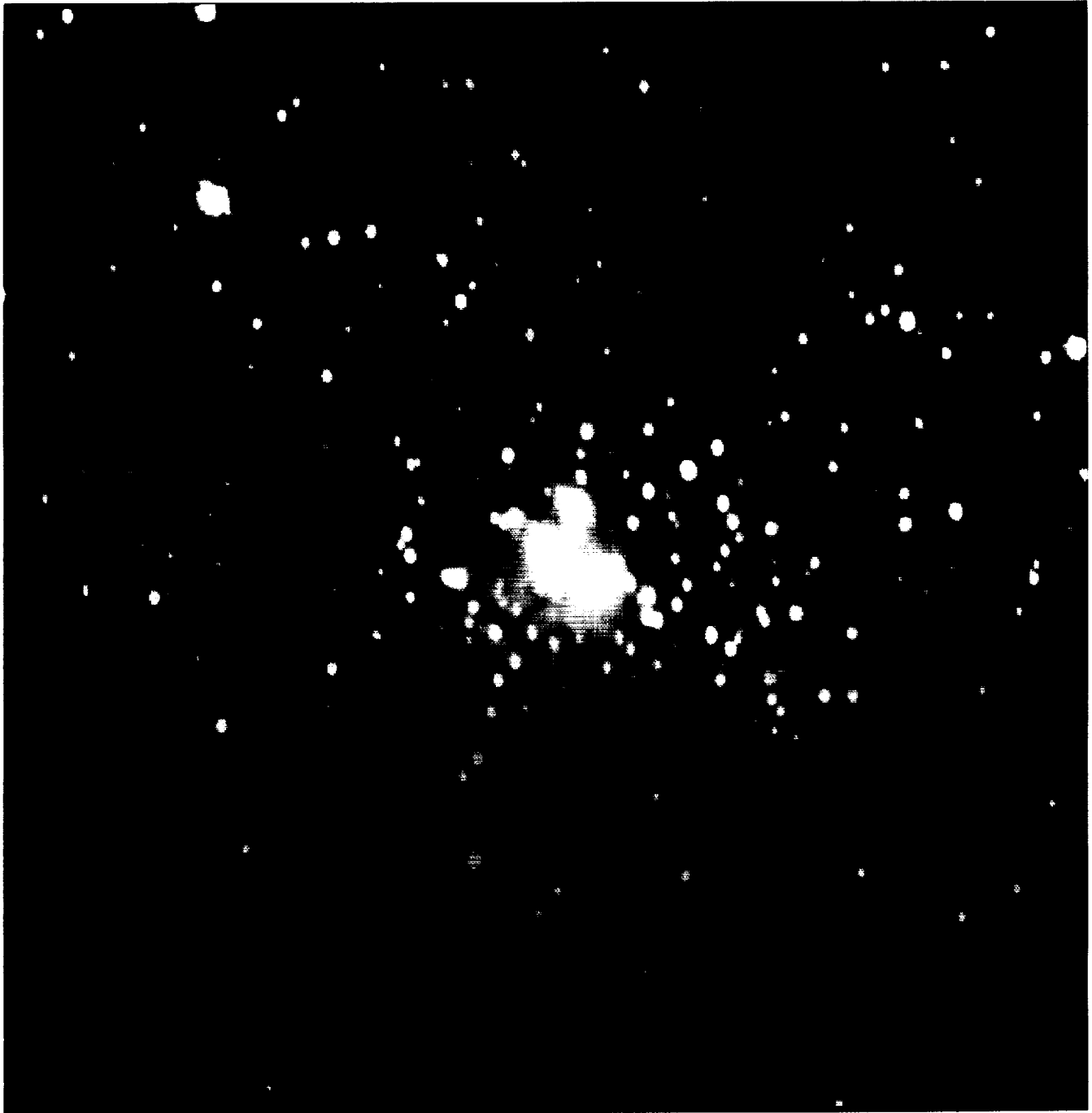
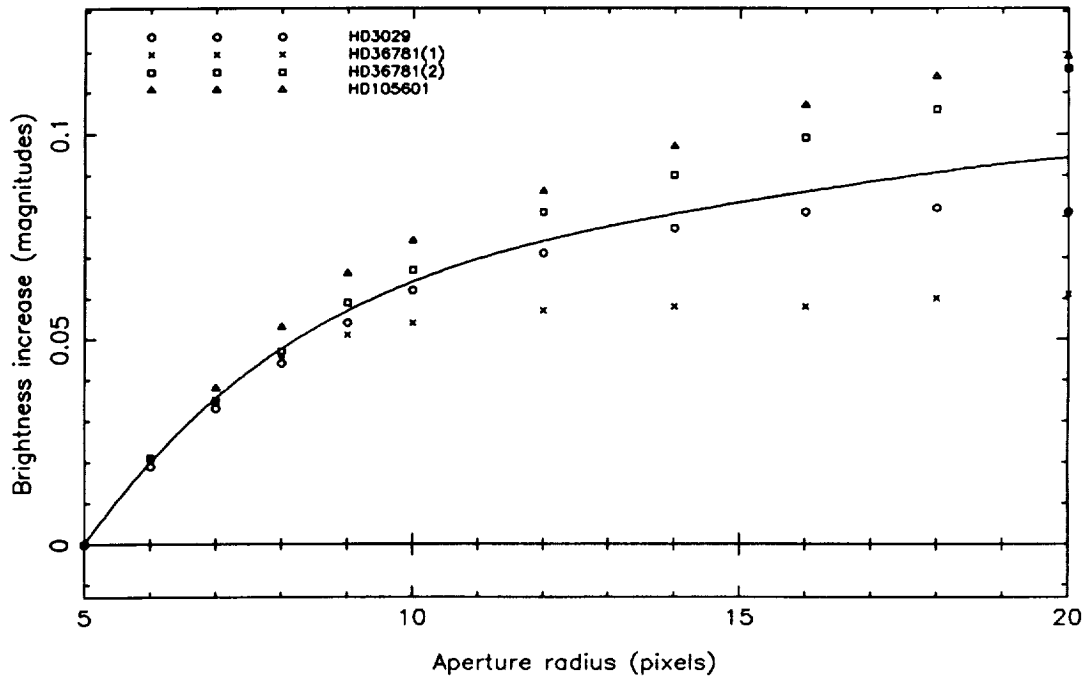Figure 9: The Orion Nebula in the Near Infrared

Figure 10: An example aperture curve of growth

'accurate measurement' implies errors $\sim$ a few percent, the best routinely achieved by conventional photometry, and the level at which many subtle systematic and random errors can crop up.

Software aperture

This technique simulates a single element photometer system, in which all the flux falling within a given aperture is added together. A software aperture can take a variety of shapes and sizes, and can be carefully chosen to avoid contamination by fainter nearby objects. Another software aperture placed elsewhere in the image is used to calculate the sky and/or nebular component—this necessarily assumes that these latter components are the same at both the on-source and off-source positions.

For extended emission, this technique can be very useful—the only real alternatives are contour maps or calibrated grey scale images. Software apertures are also often used to measure point sources, and there the technique must be more critically regarded.

Some finite sized aperture is used to measure the flux from a point source. However, for any reasonably small radius aperture ($\leq$ 5 arcseconds), some flux will remain unmeasured outside the aperture. The difference between the flux measured through the finite sized aperture and the 'total flux' from the point source as would be measured through an infinite diameter aperture is derived from the aperture 'curve of growth' (COG).

Unfortunately, changes in the seeing can badly affect the COG. Seeing profiles measured with solid state imaging devices can be well represented analytically by a Lorentzian, with a narrower core yet more extended wings than given by the more usually assumed gaussian profile—as the seeing gets worse, more flux is seen in the wings of the profile [12].

Figure 10 shows COGs derived from measurements of standard stars during one night of observing. The wavelength and camera focus were the same for all measurements, and none of the standards have faint companions. All the curves have been normalised to pass through zero at an aperture radius of 5 pixels—the diagram only shows the increase in flux beyond that point, not the amount contained in the core. The data marked by crosses and squares represent two measurements made of the same star some four hours apart in time. The cross data show an increase in measured magnitude out to an aperture radius of about 12 pixels ($\sim 7.5''$), beyond which

not much increase is seen. The square data however, show a continuing increase out beyond a radius of 20 pixels ($\sim 12.5''$). At this point, the difference between the two curves is about 0^m07, i.e. about 7% in flux terms. For reference, the conventional seeing parameter, the FWHM of the point source profile, was $\sim 1.45''$ for the cross data, and $\sim 1.55''$ for the square data, a small change.

Therefore, if a small aperture is used to measure program stars, the correction to an infinite diameter aperture could easily be several percent out, if calculated from a COG derived from a standard star measured under only slightly different seeing conditions.

The larger the arrays become, the lesser this problem. If all the program stars can be measured on one image, seeing changes will matter less—all stars will need the same aperture correction, and any error in the COG derived from a standard star measurement taken at another time will only result in a systematic shift in the program star magnitudes. With small arrays, only a limited number of sources can be measured at any one time, and seeing variations from image to image will create random errors, a worse situation.

In a crowded field, the measurement aperture for any given star will occasionally also include flux from nearby stars. The error for any given star is random—some stars will have nearby companions, some will not. The overall effect is systematic however—flux from another source included in the aperture can only increase the calculated brightness, and averaged over a whole crowded field, the stellar frequency function will appear systematically bright.

Point spread function (PSF) fitting

The technique of choice for crowded field point source photometry is PSF fitting, the best known example of which is Peter Stetson's DAOPHOT package [13]. The technique assumes that all point sources within an image have the same PSF, and that bright and faint stars only differ by a linear intensity scaling of the same profile.

The PSF is defined from the mean profile of a few fairly bright, well isolated stars in the image. All other point-like sources in the image are identified, either automatically or by hand—aperture photometry provides a first approximation to the stellar brightnesses. A model of the point sources in the image is generated from a series of overlapping PSFs, and the brightness and position of each model star is varied until the resulting structure converges in its similarity to the original image. All but a few bright stars are subtracted from the image—aperture photometry performed on the latter provides a normalisation between the profile magnitudes and large diameter aperture photometry. At that point, comments applied to software aperture photometry become relevant again.

Well tried and tested on optical CCD data, the PSF fitting technique has been applied to infrared imaging data with limited success. The main pit falls are as follows:

- One of the underlying assumptions of PSF fitting photometry is that the PSF is the same for all stars being fitted. Due again to the limited size of the current detectors, only a small number of stars can be seen in any one image, and at least one of those stars must be used as a template PSF. Combined with the effects of bad pixels, this makes it likely that a large scale mosaic will be used. However, as the individual mosaic images are taken sequentially, the PSF will not be the same across the mosaic. Also, making a mosaic where multiple images of the same star are co-added in an overlap region, the resulting stellar profile will be further distorted. Even if every individual field had one isolated bright PSF template star and a small group of crowded program stars, problems with the final aperture corrections still apply, as standard stars will have been measured at a different time, with a different curve of growth.

- Another assumption made when PSF fitting is that the background underneath the stars is flat, or at the very worst, uniformly varying on a large scale. Due to the very nature of infrared astronomy, many regions have bright emission and reflection nebulosity associated with point sources—the nebulosity can seriously affect the point source PSF fitting. Modified procedures are needed to account for both the stellar PSF and a structured background beneath the star.

- The limited size of current arrays, and the desire to image very faint low surface brightness sources often leads to the use of images scales of $1-2''$/pixel, undersampling the seeing. For PSF fitting photometry the seeing must be fully sampled, and pixel sizes in the range $0.2-0.5''$ are optimum. In the background limit, which is almost always achievable by on-chip integration, the only thing that large pixel scales buy is field of view. With the current generation of detectors, field of view *is* at a premium, but the use of heavily undersampled pixel scales will probably decrease as infrared arrays grow in size.

Contour maps and grey scale images

For regions of extended emission, contour maps are the most effective way of showing the calibrated intensity in terms of flux per unit surface area. By comparing the total measured signal from a standard star with its known flux, it is easy to convert an image from instrumental units to flux units on a per pixel basis. However, contour maps can be difficult to 'read' if there is small scale structure (e.g. stars) in the extended emission, and a grey scale image can show these details more effectively, and still provide a reasonable indication of the surface brightness, with an accompanying calibrated intensity wedge.

4.2 Conversion to a standard system

To make meaningful astrophysical deductions about sources in the images (e.g. equivalent black-body temperature, stellar type, evolutionary status, dust extinction), fluxes measured in instrumental units must be converted to a 'standard system'.

As well as using standards as zero-points, colour terms should be considered. These terms can be quite large even for transformations between single element photometers: a highly reddened object with a $J - K$ of $\sim 5^m0$ on the AAO system would have a colour of $\sim 4^m5$ on the CTIO system, a significant difference [14]. Infrared imaging systems are quite different to their predecessors, the single element photometers—re-imaging mirrors may be used; new large diameter infrared filters are being used; the detector material may be anti-reflection coated, or made of a material such as HgCdTe, little used in single element photometers. Thus, the 'colour equations' for infrared imaging systems might be quite different to those determined for the single element detector systems.

Correspondingly, measurements of a number of standard sources, both bright and faint, and both blue and red will be required to accurately calibrate infrared imaging data. Bright sources can be found in the usual lists of photometric standards, such as those of Elias *et al.* [15]. These stars are typically $\sim 7^m0$ at K, and rapidly saturate array cameras on large telescopes. If instrument systematics such as non-linearity are taken seriously, standards much closer in brightness to the program objects are required, certainly as faint as $K = 10^m0$–12^m0, and perhaps even as faint $K = 15^m0$–17^m0. Some limited work has been done using faint white dwarfs, but it would be extremely useful if a consortium of faint standards could be developed, perhaps based on a homogenised system (e.g. [16]).

4.3 Measurement errors

There are a number of ways in which systematic and random errors can creep into the measurement of a point source before it is finally placed on the HR diagram, for example. Considering the number of steps involved in reducing and analysing infrared imaging data, some of which have been discussed in this paper, cumulative errors $\sim 10\%$ in the magnitude of a star, and errors $\sim 15\%$ in its position on an HR diagram seem likely at this time. In addition, it is likely that imaging photometry with the current generation of detector arrays will involve mosaic making, and errors due to changes in seeing and aperture curves of growth, and residual errors in additive and multiplicative intensity corrections, should also be considered.

As an object lesson in quite how hard imaging photometry can be, it is worth reading the recent paper on optical CCD photometry of the globular cluster M92 by Stetson and Harris [5], in which they painstakingly analyse the errors that arise in converting raw images into calibrated luminosity functions, quoting a final accuracy in their tie in to a standard system of ±1%. Bearing in mind that optical CCDs have been in serious astronomical use for over a decade, there are undoubtedly many lessons yet to be learned about obtaining accurate calibrated photometry with infrared imaging arrays.

5 Conclusions

In this paper, I have drawn on my own experiences in infrared imaging, mainly using the UKIRT 1–5 μm camera, IRCAM, and its 62 × 58 pixel SBRC InSb+DRO array. I have only covered some of the possible approaches to acquisition, reduction, and analysis, those most suited to the kinds of data I myself have obtained. Hopefully, I have made some points of a general enough nature to provoke deeper thought about the techniques currently being developed throughout the field.

The most important point to reinforce is that obtaining accurate data from infrared arrays is not easy, and that we must be prepared to expend considerable energy in extracting the undoubtedly exciting new information that will result from the introduction of imaging into infrared astronomy.

Acknowledgements

REFERENCES AND NOTES

[1] M. J. McCaughrean. PhD thesis, University of Edinburgh, 1988.

[2] J. L. Vampola. Internal Memorandum 2118-244, Santa Barbara Research Center, September 1982.

[3] M. H. Minshull and S. E. Botts. Internal Memorandum 2113-020, Santa Barbara Research Center, November 1984.

[4] A. W. Hoffman. In C. G. Wynn-Williams and E. E. Becklin, editors, *Infrared Astronomy with Arrays*, pages 29-35, Institute for Astronomy, University of Hawaii, 1987.

[5] P. B. Stetson and W. E. Harris. *Astron. J.*, 96:909-975, 1988.

[6] W. J. Forrest, A. Moneti, C. E. Woodward, J. L. Pipher, and A. W. Hoffman. *Pub. A. S. P.*, 97:183-198, 1985.

[7] J. A. Tyson. *J. Opt. Soc. Am. A*, 3:2131-2138, 1986.

[8] D. F. Malin and P. Murdin. *Colours of the Stars*. Cambridge University Press, 1984.

[9] R. E. Schild. *Sky and Telescope*, 75:144-147, February 1988.

[10] I. S. McLean, T. C. Chuter, M. J. McCaughrean, and J. T. Rayner. In D. L. Crawford, editor, *Instrumentation in Astronomy VI*, pages 430-437, Proc. SPIE vol. 627, 1986.

[11] M. J. McCaughrean, C. Aspin, II. Zinnecker, and I. S. McLean. *Ap. J.*, 1989. In preparation.

[12] F. Diego. *Pub. A. S. P.*, 97:1209-1214, 1985.

[13] P. B. Stetson. *Pub. A. S. P.*, 99:191-222, 1987.

[14] M. S. Bessell and J. M. Brett. *Pub. A. S. P.*, 100:1134-1151, 1988.

[15] J. II. Elias, J. A. Frogel, K. Matthews, and G. Neugebauer. *Astron. J.*, 87:1029-1034, 1982.

[16] J. Koornneef. *Astron. Astrophys. Suppl. Ser.*, 51:489-503, 1983.