STATISTICAL CHALLENGES IN MODERN ASTRONOMY IV
ASP Conference Series, Vol. 371, 2007
G. Jogesh Babu and Eric D. Feigelson

The Characterization, Subtraction, and Addition of Astronomical Images

Robert Lupton

Princeton University Observatory, Princeton, NJ 08544, U.S.A.

Abstract. Astronomical images are characterized by a spatially-varying Point Spread Function ("blur"; PSF) which is usually not known ab initio. Knowledge of the PSF is crucial for the correct analysis of images (e.g. separating stars from galaxies; estimating the deconvolved shapes of objects), and we have ways of modelling the PSF, but are not currently solving the problem especially well — the fields may be (very) crowded; the data may be undersampled; and the spatial variation of the PSF may be considerable (in the case of X-ray data, varying in width by an order of magnitude). A related issue is that of finding what has changed between a pair of images taken under different conditions. Rather than solve for the PSF of both images, we generally solve for the convolution kernel that transforms one into the other, but the question of how to represent this spatial structure of this kernel is similar to the (unsolved) problem of the previous paragraph. Finally, we are often faced with the question of how to handle a set of several to many images of the same part of the sky taken under different conditions, and which are, in general, marginally sampled. A variety of ad-hoc solutions exist to recover the properties of the objects in the field, but I do not believe that we currently solve the problem optimally.

1. Introduction

Astronomical images are characterized by a spatially-varying Point Spread Function ("blur"; PSF) which is usually not known ab initio. Knowledge of the PSF is crucial for the correct analysis of images (e.g. separating stars from galaxies; estimating the deconvolved shapes of objects). To a statistical audience this is presumably obvious, as we can write down all of observational astronomy in terms of the model

$$I = \sum_{i \in \text{objects}} O_i \otimes \phi + \epsilon_i$$

where the O_i are our objects, ϕ is the PSF, and ϵ is the noise. If our objects are simple (e.g. if they're all stars, so $O_i = A_i \delta(x - x_i)$) then all that we have to do is estimate $3N$ numbers.

2. Astronomical PSFs

The PSF has at least three contributions: the detector; the telescope; and the atmosphere (this last term may be safely neglected for space-based instruments).

There are two main contributions to the PSF from the detector; charge diffusion within the photo-active material (usually Silicon), and variable sensitivity wihin the finite pixels. These contributions may be taken to not be a function of position, and are

small; critically sampling a PSF of rms size α requires a pixel of size $\sim \alpha$, so the pixel convolution increases the rms size of the PSF by $\sqrt{1 + 2/12} - 1 \sim 8\%$.

For optical telescopes, the optics are usually designed to play at most a small rôle in the overall error budget, whereas high-energy telescopes (e.g. the X-ray satellite Chandra) have optical systems which deliver severely degraded images far from the optic axis. Such drastic behaviour is generally avoided in the optical by the use of lenses, but image quality still typically degrades towards the edge of the field. In reality the goal of permitting the the optics to do no harm is not always achieved, as a telescope must be perfectly constructed, collimated, and focussed to achieve its design specifications. Almost all optical systems[1] also display 'ghosting', where internal reflections produce additional images of a star significantly displaced from the primary object (these ghosts are out of focus, so their profiles are very different from the sharply concentrated image that you are probably imagining; these ghost images are typically many arcseconds across and have the form of annuli with sharp edges and significant internal structure. Fortunately ghosts usually only contain of order 0.5% of the total power in the object. Scattering, either off dust on the mirrors or elsewhere in the system, also plays an important part in the outer parts of the PSF.

The atmospheric contribution comes both from turbulence at heights of tens of kilometers, and from a boundary layer within several hundred meters of the ground (including the telescope's dome); the two contributions at a good site are probably comparable. The high-atmosphere contribution is well described by pure Kolmogorov turbulence (Szeto and Butts 1998), with a coherence length (r_0) for the wavefront of 10 – 50cm in the optical; this leads to changes in the detailed structure of the PSF over scales of arcminutes. The low-atmosphere term has much larger spatial scales (as the angle subtended by turbulent cells is larger). Note that r_0 is much smaller than the size of a typical telescope, so the diffraction limit of the telescope is much smaller than the delivered PSF.

The Kolmogorov-dominated PSF's Fourier transform is given by Fried (1966):

$$\exp\left(-6.8839\left(2\pi k\lambda/r_0\right)\right)^{5/3}/2\Big) \tag{1}$$

where

$$\lambda/r_0 \equiv \mathrm{FWHM}/0.976 \tag{2}$$

(The 6.8839 comes from the definition of the Fried parameter r_0; the 0.976 comes from a numerical inverse Fourier transform; FWHM is the full width at half maximum [$2\sqrt{2\ln 2}\alpha$ for a Gaussian]). This is not especially convenient, but fortunately this form of the PSF is well described by the sum of two Moffat functions (Racine 1996). A sum of two Gaussians, $G(\alpha) + 0.1G(2\alpha)$, is also a convenient and reasonably accurate representation of Eq. 1.

Note that this form of the PSF has extensive wings, and in reality there are power-law components that extend further still; Fig. 2. shows the radial profile of an SDSS (Abazajian et al. 2003) PSF extending out to 30" and with a dynamic range of over 10^5.

2.1. Chromatic Effects

In reality, all of these contributions is a function of the colour of the object.

[1] the exception is off-axis purely-reflecting systems such as those used for CMB studies

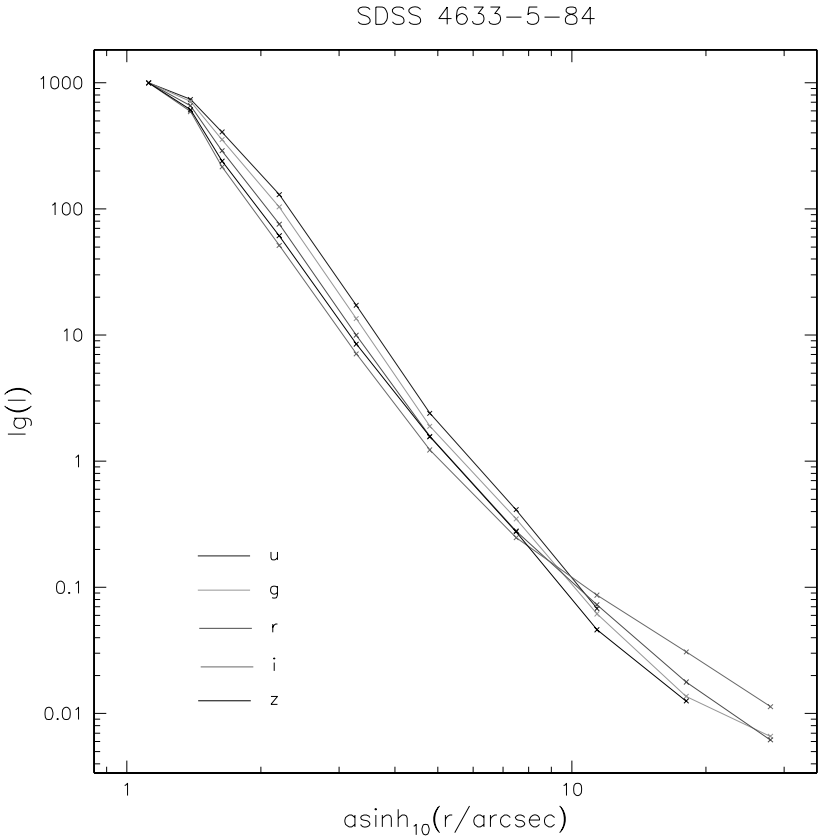

Figure 1. The measured PSF for SDSS frame 4633-5-84 in the u, g, r, i, and z bands. It will be noted that the bluer bands are generally more extended, although the outer part of the i-band profile shows the effect of scattering within the detector, and the z-band data includes a significant contribution from optical aberrations

The charge-diffusion component of the detector PSF depends on where in the device the photon is converted, and some detectors (e.g. the thinned SITe CCDs used by SDSS) produce halos around objects due to scattering within the detector packaging; the amplitude is a strong function of photon energy, especially in the red.

For purely reflecting systems the telescope optics are achromatic, but we almost always employ refractive correctors which lead to PSFs which are a function of wavelength, λ. If the system is diffraction limited, which usually requires that the telescope be above the atmosphere, the size of the PSF scales directly with λ.

The atmosphere leads to colour-dependent PSFs in at least two ways. Standard Kolmogorov theory (Fried 1966) implies that $r_0 \sim \lambda^{6/5}$, so the PSF scales as $\lambda^{-1/5}$. Usually more important, is the fact that the atmosphere behaves like a prism, deflecting blue light more than red. When observing at moderate distances from the zenith, the amplitude of this effect is about 2" from the red to the blue (and smaller in practice by a factor of a, as we observe through filters with bandpasses of order $100nm$)

2.2. Sampling and Dithering

A telescope of diameter D band-limits the image of the sky at D/λ, and the atmosphere severly attenuates beyond r_0/λ; in general $r_0 \ll D$ and we design the instrumentation to sample a signal limited at the lower frequency r_0/λ but (as Eq. 1 shows) the signal has some power at higher (angular) frequencies. Unfortunately, detectors (especially the HgCdTe arrays used in the infra-red) are expensive, and some data is taken which is not fully sampled. In many cases, multiple exposures with the telescope's pointing varied a little are taken to recover the fully sampled image, but it is not clear that the techniques used (e.g. Fruchter and Hook 2002) are optimal.

3. Estimating the PSF

Astronomers would be thrilled to find a site where the finite size of stars had any impact on their appearance (the angular diameter of α-Centauri is only about 7 milliarcseconds (mas), and the largest angular diameter of any star other than the sun is only about 60 mas); accordingly we may take the image of any star in an image to represent the PSF. Unfortunately, we aren't usually blessed with a superfluity of candidate stars; we require reasonably bright,[2] isolated stars, so statistical efficiency is important.

I have been blithely referring to "Estimating the PSF", but what is it that I wish to estimate? We'd be happy with a pixel-based represention — basically an image — but the outer parts of the PSF (even without considering ghosting) extend to several tens of arcseconds (cf. Fig. 2.), and this would involve estimating the value of at least 20000 pixels. Three approaches suggest themselves: employing a purely analytic approximation; using an exact (i.e. pixel based) representation of the inner parts of the PSF and an analytic approximation to the outer parts; and a hybrid scheme whereby we use an analytic form, and also carry a residual image (Stetson et al. 1990). Because the PSF has spatial structure, we have to also ask how to represent this in any of these approaches.

---

[2] But not too bright; the dynamic range of a CCD is only about 50000, assuming a full well of 65000 and setting the gain so that we have 3 bits to sample the noise, i.e. $\sigma \sim 1$

From the discussion in Sec. 2. of the various contributions to the PSF you might think that it would make sense to model the PSF as a convolution of the constant-in-time component (the detector and telescope), and the constant-in-space component (the atmosphere). To the best of my knowledge this has not been attempted, although Jarvis and Jain (2004) have modelled the PSF in an ensemble of images by studying its variation at a given point in the focal plane as a function of exposure number.

3.1. The SDSS PSF Algorithm

For the sake of argument, and because I was responsible for it, let us consider the SDSS algorithm (Lupton et al. 2001). In this case, the data was taken in Time Delay Integrate (TDI) mode, which has the unfortunate effect of making *temporal* variations in the PSF appear as (1-dimensional) *spatial* structure; this means that the spatial variations of the PSF are complex, and might encourage you to try the approach of the previous paragraph (as it would reduce the dimensionality of the required spatial model), but the complexity of the modelling pursuaded me not to try, but rather to model it heuristically using a Karhunen-Loève (KL) transform (Hotelling 1933; Karhunen 1947; Loève 1948).

KL Expansion of the PSF We can use a set of objects identified as stars to form a KL basis, retaining the first n terms of the expansion:

$$P_{(i)}(u,v) = \sum_{r=1}^{r=n} a^r_{(i)} B_r(u,v) \tag{3}$$

where $P_{(i)}$ is the i^{th} PSF star, the B_r are the KL basis functions, and u, v are pixel coordinates relative to the origin of the basis functions. In determining the B_r, the $P_{(i)}$ are normalised to have equal peak value.

Once we know the B_r we can write

$$a^r_{(i)} \approx \sum_{l=m=0}^{l+m\leq N} b^r_{lm} x^l_{(i)} y^m_{(i)} \tag{4}$$

where x, y are the coordinates of the centre of the i^{th} star, N determines the highest power of x or y to include in the expansion, and the b^r_{lm} are determined by minimising

$$\sum_i \left(P_{(i)}(u,v) - \sum_{r=1}^{r=n} a^r_{(i)} B_r(u,v) \right)^2 ; \tag{5}$$

note that all stars are given equal weight as we are interested in determining the spatial variation of the PSF, and do not want to tailor our fit to the chance positions of bright stars. An alternative way to achieve this would have been to weight each star by $\sigma^2 + \Upsilon^2$ where σ is a measure of the uncertainty in the star's flux, and Υ is a floor designed to prevent bright stars from dominating the fit; in the limit $\Upsilon \to \infty$ we would recover the equal-weights scheme that we in fact adopted.

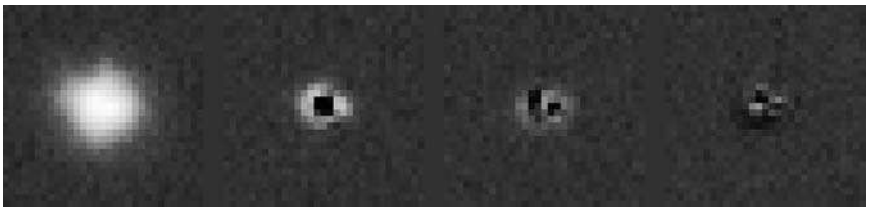

Figure 2. The KL basis images for frame 756-z6-700, using a histogram-equalised stretch.

The Outer Parts of the PSF The KL model of the PSF extends to only about 7", but the outer parts contain a significant part of the total flux, and need to be modelled. One difficulty is that stars with sufficient signal-to-noise in the outer parts have cores that saturate the detectors. In practice we use partial radial profiles from stars of a range of magnitudes to construct a single profile in each filter, such as the one shown in Fig. 2.. We can use our KL model to estimate the central parts of the PSF at any position in the image, and then use this composite profile to model the outer parts.

Application to SDSS data For each CCD, in each band, there are typically 15-25 stars in a frame that we can use to determine the PSF. We impose an upper limit (usually 100) on the number of stars used in order to save memory and processing time; when imposing this limit, we keep the brightest stars. We usually take $n = 3$ and $N = 2$ (i.e. the PSF spatial variation is quadratic).

The 4 KL basis images from an early SDSS run are shown in Fig. 2 and the resulting reconstructed PSFs at 35 points within the frame are shown in Fig. 3; this was early data, and the images are clearly astigmatic, with the astigmatism varying as a function of position.

Additionally, we force the KL basis images with $n > 0$ to have 0 mean in their outermost pixels, those more than approximately 7" from the center of the star (7" is approximately the radius of the region used to determine the KL PSF) This doesn't force the PSF to be zero in its outermost parts, but it does mean that the PSF is taken to be constant at 7".

This modification to the basis functions means that they are no longer strictly orthogonal, but it makes the PSF determination significantly more robust to errors in the background level.

Gauging the Success of the KL Expansion In regions of low star density (especially a problem in the u band), or when there happen to be no stars in part of the field, or when the seeing is changing too fast, this scheme to model the PSF's spatial structure can run into trouble; these troubles tend to manifest themselves towards the edge of the field, where we are forced to extrapolate. We accordingly monitor its performance, and modify the algorithm as appropriate. After determining the spatial expansion of Eq. 5 we estimate the PSF at about 120 positions around the edge of the field, and also a few points in the interior. At each of these positions we calculate both an aperture

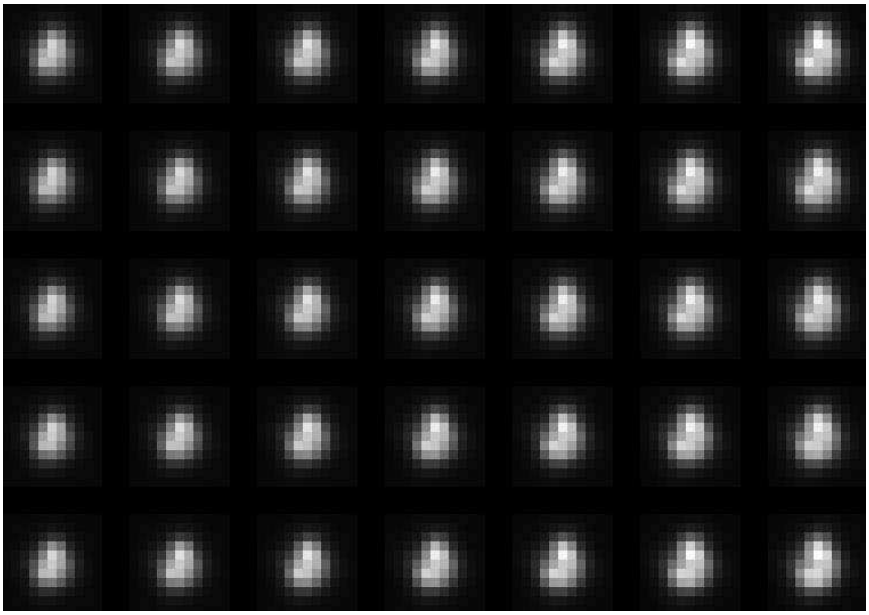

Figure 3. The estimated PSF for 35 positions in frame 756-z6-700, using a linear stretch. This is early SDSS data, and for one of the CCDs with worst image quality, and the astigmatism is clearly seen. The first component of Fig. 2 includes this astigmatism, although it is not obvious with the stretch used for that figure.

and a PSF flux and thus an aperture correction.[3] The PSF model used in estimating the flux is based on a *local* Gaussian fit. We then reject this PSF model if the maximum or minimum aperture correction seen differs from unity by more than 30%. It is not obvious that this is a good measure of failure, but it is simple, is related to a quantity that interests us deeply, and has served us well.[4]

Once we have what appears to be an acceptable PSF model based on examining the form of the reconstructed PSFs, we proceed to ask how well the model reproduces the stars that were used to construct it. We do this by comparing the PSF photometry based on the modeled KL PSFs to the aperture photometry for the same (bright) stars.

The width of the distribution of these differences is typically 1–2% and indicates an upper limit on the accuracy of the PSF photometry. Without accounting for the spatial variation of the PSF across the image, the photometric errors would be as high as 15%.

Problems with this approach One obvious deficiency of this algorithm is that it requires that we are able to identify isolated stars; we may hope that most of the signature of neighbours is restricted to high KL components (and thus discarded), but this is clearly only an approximation — after all, the first component is simply the mean of all the input objects.

Another problem is that we assume that all input objects are centered within a pixel; more precisely, we assume that we are able to shift them without error. For well-sampled data this is probably acceptable (we use a cosbell-tapered sinc shift), but once the seeing becomes too good this is an obvious difficulty.

Both of these issues could be avoided by estimating the properties of n fully-sampled "KL" images from the data, rather than by constructing the basis directly. I don't know of any cases where this has been applied to PSF reconstruction, but a similar approach is used in the SDSS `idlspec2d` spectroscopic data reduction (Schelgel and Burles 2007). Such an approach could use robust estimators and outlier rejection to minimise the effects of contamination.

Another widely used approach (Stetson et al. 1990; Schechter et al. 1993), is to fit a number of objects simultaneously using analytic PSF models, and this could be combined with the "KL" based-approach of the previous paragraph. Spatial variation of the PSF can be included in the model fit.

The models used for spatial variability are generally polynomials. In the case of SDSS data this is probably not a very good basis (due to the coupling of time- and spatial- variability due to TDI). For long exposures on large telescopes polynomials appear to work reasonably well, although it has been shown that, unsurprisingly, rational functions perform rather better (Van Waerbeke et al. 2005). However, once we come to short exposures better models may well be justified; I shall return to this in the discussion of Image Subtraction in Sec. 4.

---

[3] A *PSF flux* is an object's flux based on fitting a model; an aperture flux involves merely integrating the data over some region. The former is less noisy, but also less robust.

[4] We are asking questions about the form of the KL PSF, so there is no variance estimate available which would enable us to ask questions about goodness-of-fit

3.2. Using an Ensemble of Exposures

In many applications there are simply not enough stars to map the PSF's spatial variations with sufficient accuracy. One can imagine visiting dense star fields in order to measure the PSF in glorious detail (Hoekstra 2004), but unfortunately conditions (e.g. telescope focus) will in general have changed by the time that you are ready to take your science images. An interesting approach is to study a large number of exposures taken with a given telescope — including rich star fields — and ask about the statistical properties of this ensemble (Jarvis and Jain 2004). They use a 4^{th} order polynomial fit to the spatial structure, and use 8–10 components to analyse 651 exposures with 10^5 stars per exposure — a significant computational load as they perform QR decomposition on a $10^5 \times 2000$ matrix.

The parameters used to describe the PSF are not specified; for weak lensing studies it may be sufficient to measure a few weighted moments, but one can certainly use a richer representation — for example the components of a KL basis of all the stellar images, or an expansion of the PSF into some favoured orthogonal basis.

4. Image Subtraction

In many situations, only the time variable part of an image is of interest; the classic case is searching for gravitational micro-lensing in the direction of the Galactic bulge. In this case, we have two or more images of the same part of the sky, taken under different conditions; in particular the PSF will be different in the two exposures.

The classic solution to the problem of searching for variability is to measure the brightness of each source in both images and compare the resulting catalogues with each other (Alcock et al. 1996; Udalski et al. 1993).

An obvious alternative is to subtract the two images directly, allowing for the difference between the seeing in the two images. It might seem that we've reduced our difficulties to a previously solved problem; after all, given two images $I_1 \equiv S \otimes \phi_1$ and $I_2 \equiv S \otimes \phi_2$, we can write the Fourier-transform of the (seeing-matched) difference as $I_1(k) - I_2(k) \times (\phi_1(k)/\phi_2(k))$ (Tomaney and Crotts 1996) and we now know how to estimate ϕ. Unfortunately, it's difficult to measure the outer parts of the PSFs well enough to carry out this Fourier division, and the problems only increase as the density of stars increases.

To avoid this difficulty, Alard and Lupton (Alard and Lupton 1998) realised that what really mattered was how well the subtraction worked, and that the the residuals left by subtracting objects that *hadn't* varied should be as small as possible; that is, we should find the kernel K such that $R \equiv ||I_i - K \otimes I_2||$ be minimised. They further proposed that K be written as a sum of a set of basis functions: $K(u,v) \equiv \sum_r a_r B_r(u,v)$. The task of subtracting two images is thus released to the problem of finding a set of a_r that minimise R; if we use an L_2 norm, this is a simple least-squares problem that may be solved by inverting a matrix (possibly using SVD).

Alard (Alard 2000) has generalised this approach to handle spatially-varying PSFs, using a similar approach to that of Eq. 4, and writing

$$K(u,v;x,y) = \sum_{r=1}^{r=n} \sum_{l=m=0}^{l+m \le N} b_{lm}^r x_{(i)}^l y_{(i)}^m B_r(u,v) \tag{6}$$

This technique has been widely applied (Zebrun et al. 2001; Bond et al. 2001; Becker et al. 2004), but the questions of the optimal choice of basis functions, the best way of representing K's complex spatial variation, and the choice of the 'reference image' I_2 are still open.

Many colleagues have investigated the use of basis functions of the form

$$B_r(u, v) = \delta(u - u_r)\delta(v - v_r); \tag{7}$$

i.e. solve directly for K's pixel values and their spatial variation; they have found that the resulting kernel is unacceptably noisy. While it would be possible to regularise (e.g. with a total variation cost function?), another approach comes to mind, although I do not believe that it has been tried:

> Use Eq. 7 to estimate the kernel K and its spatial structure, without any regularisation, then reconstruct K at a dense set of points within the field. Perform a KL decomposition of these estimates of K, keep the significant terms, and re-expand their spatial structure.

This is not a subtle procedure (and indeed it's likely that one could achieve a similar result without actually re-expanding the initial estimate), but it has the advantage of being simple to implement and separating the kernel estimation problem from that of estimating the spatial variation of a known function — indeed, the second part of the procedure is exactly the one that we used to estimate the spatial structure of the SDSS PSF, so when the assembled Statistical Wisdom solves the PSF estimation problem, I'll be able to apply an identical algorithm to kernel estimation for image subtraction.

Let us return to the choice of the reference image, I_2. In general we have a set of N images and we want to find the variability in the entire ensemble. One approach would be to choose the image with the best seeing as I_2 (so as to avoid forcing K to do any deconvolution), but this has a number of problems: It increases the noise in the difference image by $\sqrt{2}$; it leads to noise correlations; it requires that we mask out a region the size of the convolution kernel K around the edges of the field and any defects (e.g. cosmic rays; bad columns; saturated stars) in I_2; and finally if I_2 is not well sampled we have all the difficulties of resampling an aliased image.

The obvious remedy to all of these woes is to generate a fully sampled, defect free image from the ensemble of N images; its noise will be down by a factor of $\sqrt{N}$ from an individual frame, and will make a negligible contribution to the noise in the difference image. Unfortunately, I don't know any really good ways to generate such an image; it is the subject of the next section, on Image Addition.

A final point to ponder is how to handle data taken at a variety of airmasses; as discussed under PSFs, red and blue light images are slightly displaced from one another, and in general by different amounts than in the master image. A brute force approach would be to generate master templates at different airmasses. It seems probably that using the known colour distribution of the pixels in the image would be advantageous; but the obvious approaches (write the observed image as a linear combination of templates taken at different airmasses, or a template and a colour-template) do not appear to work, as we cannot write our data as a linear combination of the inputs.

5. Image Addition

There are some situations when we probably really should want to add images; the generation of reference images for image subtraction is one such case. In other cases, it may well be better to handle an ensemble of N images as an estimation problem with N realisations of the data.

In general each of these images has a different (and possibly spatially varying) background level, its own footprint on the sky, its own geometrical distortion, and its own PSF structure. Let us assume that the latter two are known.

Traditional Approaches; SWarp and Drizzle Historically, most astronomers have used empirical approaches to adding images together. For example, SWarp (Bertin et al. 2002) reads input images and, optionally, corresponding variance maps. It constructs and removes a background image from each input, resampled onto the desired output projection (using a user specified interpolation scheme, but typically a Lanczos$_3$ kernel), and then averages all of the resampled input images; the average is typically a mean, but other options (e.g. a median or minimum χ^2 estimator) are possible.

A procedure such as this works well for well sampled data if you don't care about the PSF (and especially its continuity), but there are a number of problems with this approach (as SWarp's author is well aware, although it is possible that most consumers are more naïve).

Firstly, subtracting background levels is a tricky business — as Fig. 2. showed stars have wings that extend to c. 30", but galaxies are far worse. A galaxy profile is typically of the form $\exp(-r^{1/n})$ with $1 \lesssim n \lesssim 4$, so a large fraction of the total flux is at large radii and low surface brightness. This means that estimating the background is something that should be done as part of the science analysis of the final high signal-to-noise image, not as part of the image combination step. Fortunately the solution is simple: rather than estimate the background, estimate the *differences* in the background between the images.

Secondly, the PSF in the final image will, in general, be discontinuous wherever the number of images contributing to a given output pixel changes — note that this may be due to defects in the input images (e.g. cosmic rays) as well as edge effects. This could be partially alleviated by applying a kernel such as those of the previous section to normalise the seeing before resampling, but this will have the effect of correlating the noise; it might be argued that the resampling has *already* correlated the noise, but this should be a minor effect if the scale hasn't changed, and the resampling kernel is close to sinc (the input noise is close to white and uncorrelated, at least for faint objects).

Thirdly, the input images may not be well sampled. This isn't a fundamental limitation — given enough images we can recover the fully sampled image, but it means that the interpolate-and-average approach will fail. One popular strategy, especially for HST data, is to use Drizzle (Fruchter and Hook 2002). The basic idea here is to map the observed pixels onto a fully-sampled grid, but to pretend that all the flux in concentrated into the central p% (typically $p \sim 50$) of the physical pixel. This preserves surface brightness and has good noise properties, but fails to preserve the PSF.

For the case that the PSF is known and constant for all exposures, and in the absence of noise, we may use standard iterative techniques for non-uniformly sampled data (Feichtinger et al 1995; Fruchter 2005) although it is unclear how well these will scale for very large, noisy, datasets. As far as I know, noone has generalised these techniques for the case that each input image, while band-limited, has its own PSF.

5.1. Estimation of the True Sky

An alternative approach that has not, so far as I know, been employed in anger was proposed by Kaiser (Kaiser 2004). If an point in an image is given by f_p where the pixel index p runs over all the pixels in a set of images and the p^{th} pixel is associated with a PSF g_p, in the limit of uncorrelated Gaussian noise n_p, we may write down the likelihood of the (Fourier transform) true sky $\boldsymbol{I}$, thought of as a vector, as

$$\mathcal{L}(\boldsymbol{I}) = \boldsymbol{I}\boldsymbol{\phi} - \frac{1}{2}\boldsymbol{I}A^{-1}\boldsymbol{I} + \text{constant} \tag{8}$$

where $\boldsymbol{\phi} \equiv \sum_p f_p g_p/\sigma_p^2$ and $A \equiv \sum_p g_p g_p/\sigma_p^2$. This is minimised at the point

$$\bar{\boldsymbol{I}} = A^{-1}\boldsymbol{\phi} \tag{9}$$

As Kaiser points out, multiplying by A^{-1} is tantamount to deconvolving the orginal data, but nevertheless the data f_p enter into the estimator $\boldsymbol{I}$ only through the sum $\boldsymbol{\phi}$; in otherwords, $\boldsymbol{\phi}$ is a sufficient statistic for $\boldsymbol{I}$. In real space,

$$\boldsymbol{\phi}(\boldsymbol{r}) = \sum_i \left(g_i^\dagger \otimes f_i\right)_{\boldsymbol{r}} /\sigma_i^2 \tag{10}$$

which has effective PSF

$$\sum_i \left(g_i^\dagger \otimes g_i\right)_{\boldsymbol{r}} /\sigma_i^2 \tag{11}$$

I would like to thank many colleagues for discussions and enlightenment on the topics discussed here. In particular, Željko Ivezić and Jim Gunn for the SDSS related sections; Gary Bernstein for discussions on PSFs and inspiration on how one might not need to coadd images at all; Andy Fruchter for keeping me abreast of his thoughts on undersampled images; and Nick Kaiser for insight into handling multiple images, even if he does think in k-space. Finally, to Bohdan Paczyńki who, although he claims to understand nothing, has been a continuing source of encouragement as I've worried about these topics for far too many years.

References

Abazajian, K., et al., 2003, AJ 126, 2081.
Alard, C. and Lupton, R. H., 1998, ApJ 503 325.
Alard, C., 2000 AAS 144, 363.
Alcock, C. et al. 1996 ApJ, 461, 84.
Becker, A. C., Wittman, D. M., Boeshaar, P. C, et al. 2004 ApJ 611, 418.
Bertin, E., Mellier, Y., Radovich, M., et al., in *Astronomical Data Analysis Software and Systems XI*, edited by Bohlender, D. A., Durand, D., and Handley, T. H., ASP Conference Proceedings, 281, 228
Bond, I. A., Abe, F., Dodd, R. J., et al. 2001, MNRAS, 327, 868
Jarvis, M. and Jain, B, 2004 astro-ph/0412234
Feichtinger, H. G., Grchenig, K., and Strohmer, T, Numerische Mathematik, 1995, 69, 423.
Fruchter, A. S., and Hook, R. N., 2002, PASP 114, 144.
Fruchter, A. S., 2005 *The 2005 HST Calibration Workshop*, (Koekemoer, A. M., Goudfrooij, P., and Dressel L. L., eds.) 433.

Fried, D. L. 1966, J. Opt. Soc. Am., 56, 1372.
Hoekstra, H, 2004 MNRAS, 347, 1337.
Hotelling H., 1933, J. Educ. Psychol., 24, 417&498.
Kaiser, N, 2004 Pan-STARRS document PSDC-002-011-00.
Karhunen, K. 1947, Ann. Acad. Sci. Fennicae, Ser. A137 (Translated by I. Selin in "On Linear Methods in Probability Theory", T-131, 1960, The RAND corp., Santa Monica, Ca)
Loève, M., 1948, "Fonctions Alineare de Secondre Ordre." in P. Levy, *Processes Stochastique et Mouvement Brownien*, Hermann, Paris.
Lupton, R. H., Gunn, J.E., Ivezić, Ž., Knapp, G.R., Kent, S., & Yasuda, N. 2001, in *Astronomical Data Analysis Software and Systems X*, edited by F. R. Harnden Jr., F. A. Primini, and H. E. Payne, ASP Conference Proceedings, 238, 269
Racine, R., 1996 PASP, 108, 699
Schechter, P. L., Mateo, M., and Saha, A. 1993, PASP, 105 1342.
Schlegel, D.J., and Burles, S. In preparation.
Stetson, P. B., Davis, L. E., and Crabtree, D. R. 1990, ASPC 8 289.
R. K.-H. Szeto and R. R. Butts 1998, J. Opt. Soc. Am., 15, 6,
Tomaney, A. B., and Crotts, A. P. S. 1996 AJ 112, 2872
Udalski, A., Szymanski, M., Kaluzny, J., et al. 1993 AcA 43 289
Van Waerbeke, L., Mellier, Y., and Hoekstra, H. 2005 AA 429 75
Zebrun, K., Soszynski, I., Wozniak, P. R., et al. 2001 AcA 51, 317.