School Lecture E2: Object Detection and Measurement

Robert Lupton

2011-06-02

How should I detect [faint] objects?

Let's restrict our attention to stars.
Let's restrict our attention to *isolated* stars.
We'll first answer a simpler question:

How should I measure a star's brightness?

Aperture fluxes

One solution would be to add up all the intensity within some region centered on a star.

Image

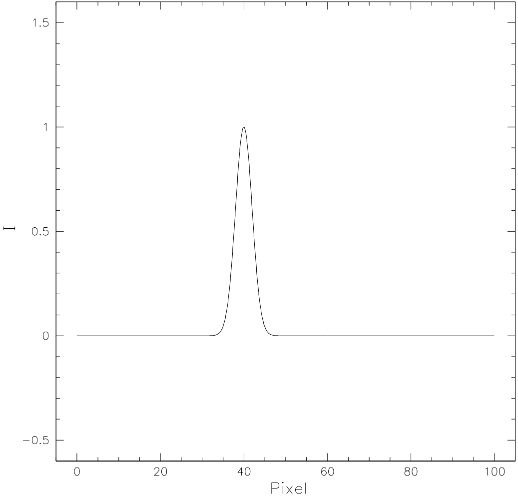
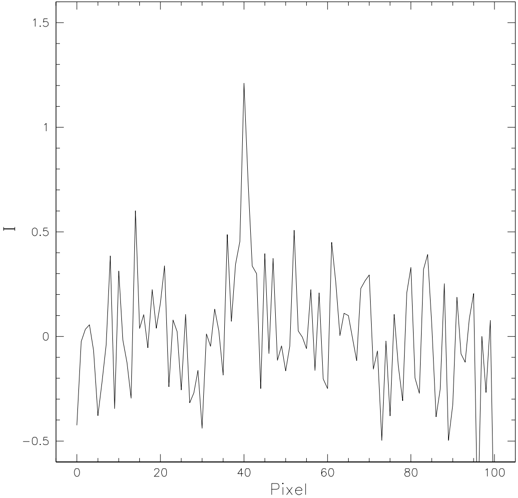
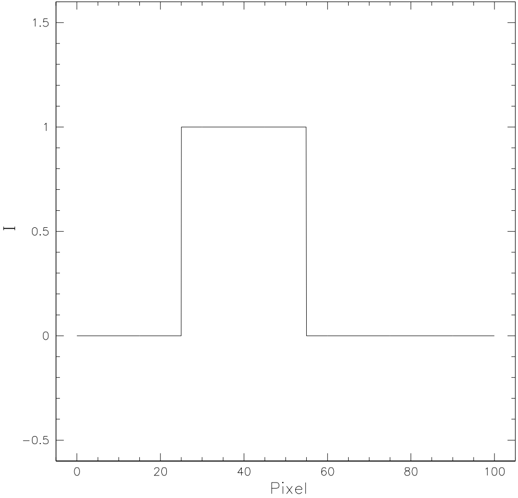


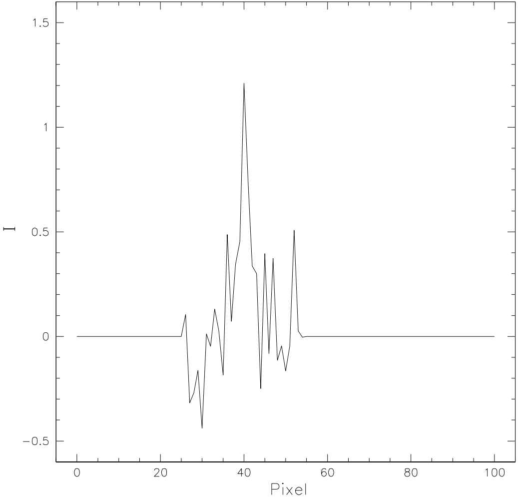
Image with noise



My Aperture



Aperture Flux with noise



This is obviously a noisy measurement

Isolated Stars

We have a PSF ϕ, so a part of the sky containing a single star with flux F_0 at x_0 may be modelled as

$$S + F_0\,\delta(x - x_0)$$

where S is the background, taken to be constant. Including the PSF and noise, N, we have:

$$I(x) = S + F_0\,\delta(x - x_0) \otimes \phi + N$$

I.e.

$$I(x) = S + F_0\,\phi(x - x_0) + N$$

Noise

In the optical, to a very good approximation, N is almost entirely shot noise due to the finite number of photons detected, *i.e.* it is a Poisson process. This may be approximated by a Gaussian with mean and variance equal to the signal, I.

Background Estimation

I believe that background estimation is an unsolved problem. One approach is to use the median of the image as an estimator.

Question: what is $\mathrm{median} - \mathrm{mean}$ for a Poisson distribution in the limit of large mean?

Answer: $\frac{1}{6}$.

For today let's assume that S is known, and accordingly set it to 0.

Likelihoods

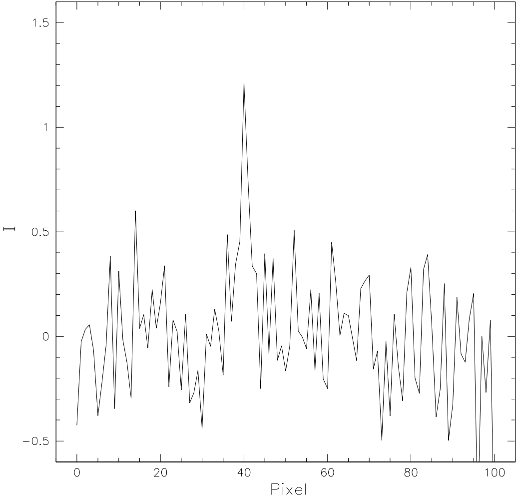
We can write the (log-)likelihood of our star as

$$\ln \mathcal{L}(I|F_0, x_0) \propto -\sum_i \frac{(I_i - F_0\,\phi_i)^2}{\sigma_i^2}$$

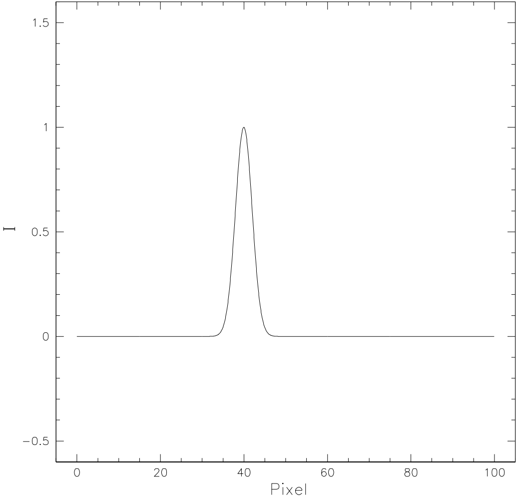
where $\sigma^2 = S + I$.
Differentiation with respect to F_0 we find that this is maximised at

$$F_0 = \sum_i \frac{I_i\phi_i/\sigma_i^2}{\phi_i^2/\sigma_i^2}$$

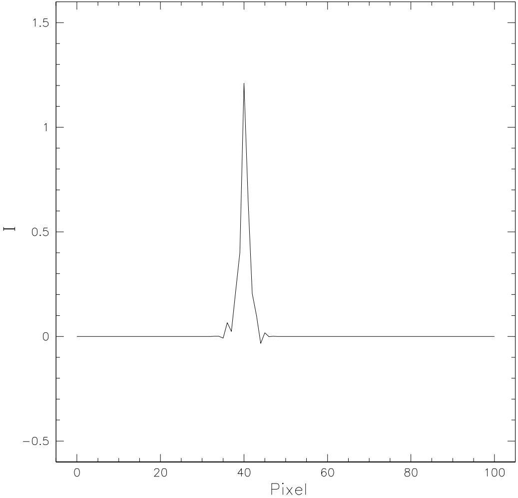
Image with noise



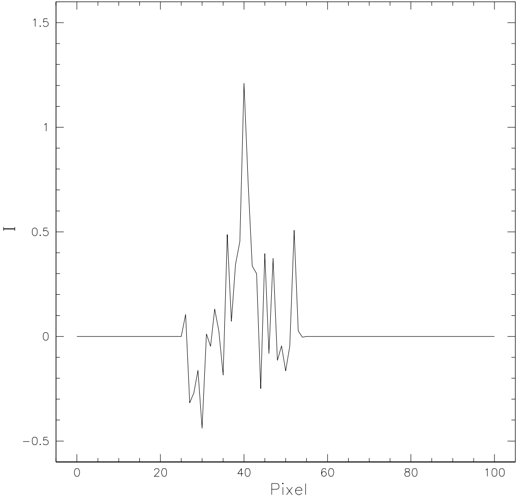
PSF



PSF flux



Aperture flux



Source Detection

I'm primarily interested in faint sources, so the noise is dominated by S which is the same in all pixels. We then have

$$\ln \mathcal{L}(I|F_0, x_0) \propto -\sum_i \left(I_i - F_0\,\phi_i(x_0)\right)^2$$

$$\ln \mathcal{L}(I|F_0, x_0) \propto -\sum_i I_i^2 + 2F_0 \sum_i I_i\phi_i(x_0) - F_0^2 \sum_i \phi_i^2(x_0)$$

The only term that depends on x_0 is $\sum_i I_i\phi_i(x_0)$, a convolution (actually correlation) with ϕ.
The maximum likelihood estimate of the position of our object is thus given by the maximum of the initial data, convolved with the PSF.

Do I get to use an FFT?

FFTs cost $N/2\log_2 N$ multiplies; for an $M \times M$ image that's

$$M^2\left(\frac{1}{2} + \log_2 M\right)$$

multiplies to convolve I with ϕ.
If ϕ is represented as an $n \times n$ image, the direct cost would only be $n^2 M^2$
Or $2nM^2$ if the filter is separable.
And don't forget about cache efficiency...

Measuring fluxes using the Psf

For faint sources (so all pixels have the same variance, σ^2), the flux is given by

$$F_0 = \sum_i \frac{I_i \phi_i}{\phi_i^2}$$

So each photon is weighted by the PSF's profile, *i.e.* the probability that it belongs to the source.
In this limit, the noise in the measurement is

$$\frac{\left(\sum_i \phi_i\right)^2}{\sum_i \phi_i^2}\,\sigma^2 \equiv n_{\mathrm{eff}}\,\sigma^2$$

If the PSF is Gaussian $N(0, \alpha^2)$, $n_{\mathrm{eff}} = 4\pi\alpha^2$

Aperture Fluxes

We can now see what went wrong with our aperture measurement; we assumed that the object's profile was a top-hat and paid the (noise) penalty.

You will sometimes meet astronomers who think that an aperture flux is somehow more "natural" than fitting a model; you now know why they are wrong.

For bright objects things are different; now the noise is dominated by photon noise in the source, and a (large) aperture has higher signal to noise. We can understand this probabilistically too; as the background is negligible, all photons should be assumed to come from the source and lovingly counted.

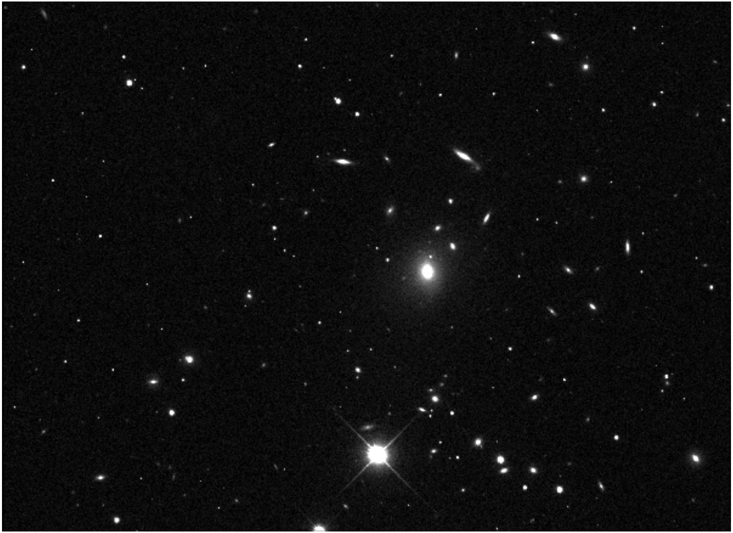
What about Galaxies?

Galaxies have many more degrees of freedom than stars. The simplest plausible galaxy model is probably a Sérsic model:

$$I(r) = I_0 \exp(-(r/r_e)^{-1/n})$$

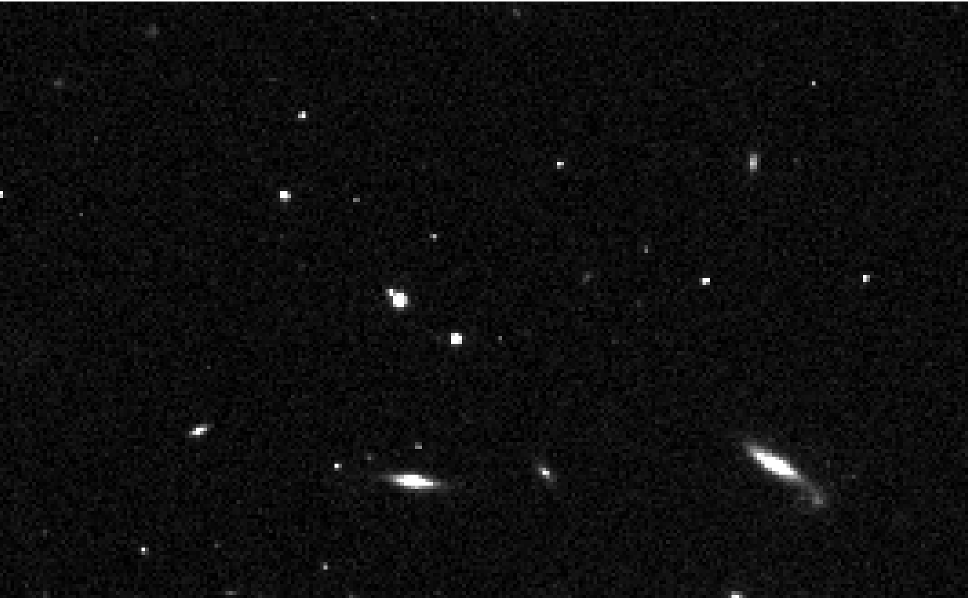
where r is the major axis of an elliptical isophote; that's 7 parameters: x_0, y_0, I_0, r_e, n, a/b, θ

An SDSS field



A reconstructed SDSS field

An SDSS field



A reconstructed SDSS field

Forward modelling

At least to SDSS depths (22.5 5σ), even simple models appear to capture most of the information present in high-Galactic latitude fields.

How should I process a set of images?

Given a set of images of the same part of the sky, how should I process them to obtain deeper data?

- How far does $\sqrt{N}$ take you?
- Should I add the images together?
- What's a good algorithm to add images?
- Is there an optimal algorithm?
- Do I need an optimal algorithm?

How should I coadd a set of images?

There are (at least) three ways to think about processing repeated images:

- Add the images together somehow
- Process each image separately and add the results
- Process all the images simultaneously

Adding images together

Among the problems are:

- Correlated noise
- Sampling
- Discontinuous PSFs
- No opportunity for non-linear analysis in the processing (e.g. 3σ clips).
- Average over moving/variable objects

  On the other-hand, it has the great advantage of being computationally relatively simple and cheap.

Processing each image separately

An easy alternative is to process each exposure separately, and add the resulting measurements.

- Only objects detected in at least one frame are measured
- There is no guarantee that the same objects will be detected in each exposure
- It seems unlikely that the errors in all measurements (e.g. galaxy effective radii) will scale as $\sqrt{N}$.

  There are ways around some of these problems; for example, we could *detect* on a coadded frame and then use this master catalogue to measure each of the input images.

Processing all the images simultaneously

It seems clear that this is the right thing to do, but it is expensive:

- All data vectors are longer by the number of exposures
- Reading all the data into memory may be difficult

Estimating a Picture of the Universe

N.b. I stole some of these ideas from Nick Kaiser.
If we decide to create a coadd, we can write down the ML estimate of the Universe U given an image, I, and a (known) PSF, ϕ:

$$I(x) = U(x) \otimes \phi(x) + \epsilon(x)$$
$$I(k) = U(k) \times \phi(k) + \epsilon(k)$$

Let us assume that all objects are fainter than the sky, so ϵ is an $N(0, \sigma^2)$ variate.

$$\ln \mathcal{L} \propto -\ln \sigma - \frac{1}{2} \frac{(U\phi - I)^2}{\sigma^2}$$

Estimating a Picture of the Universe

If we have multiple images, I_i, this becomes:

$$\ln \mathcal{L} \propto -\sum_i \ln \sigma_i - \frac{1}{2} \sum_i \frac{(U\phi_i - I_i)^2}{\sigma_i^2}$$

so, differentiating with respect to the Universe,

$$U(k) = \frac{\sum_i I_i \phi_i / \sigma_i^2}{\sum_i \phi_i^2 / \sigma_i^2} \equiv \frac{D(k)}{P(k)}$$

An Optimal Algorithm

$$U(k) = \frac{D(k)}{P(k)}$$

$$D(k) \equiv \sum_i I_i \phi_i / \sigma_i^2; \qquad P(k) \equiv \sum_i \phi_i^2 / \sigma_i^2$$

I.e.

$$U(x) = D(x) \otimes^{-1} P(x)$$

where

$$D(x) = \sum_i I_i \otimes \phi_i / \sigma_i^2; \qquad P(x) = \sum_i \phi_i \otimes \phi_i / \sigma_i^2$$

Note that the function $P(x)$ is discontinuous wherever the exact set of input exposures changes; in general this will lead to a very large number of very small chunks. One suggestion is to generate a separate coadd for each object.

Estimate the properties of the Universe

Another alternative is to fit directly to the input data. This is straightforward for e.g. PSF magnitudes. Harder problems include:

- Sky estimation
- Object detection
- Deblending
- Shape measurements

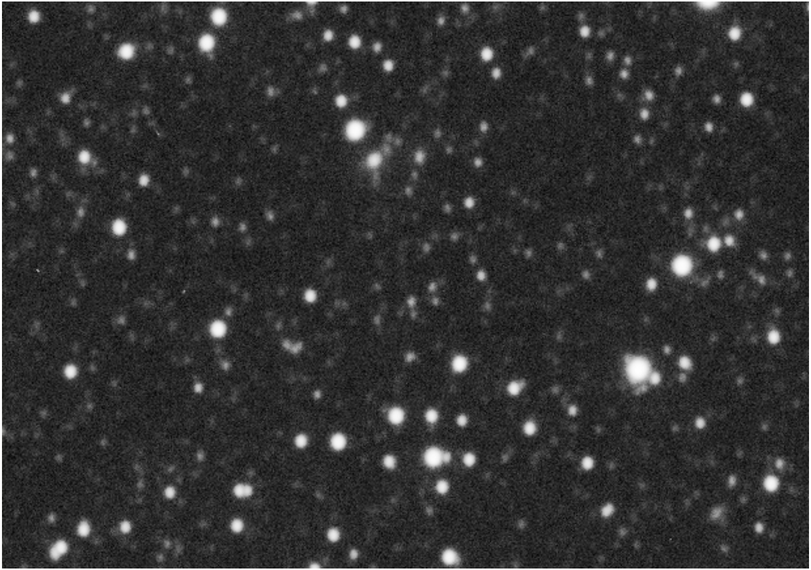
Some of these are hard (e.g. deblending); some are just expensive.

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.

Exposure 1a

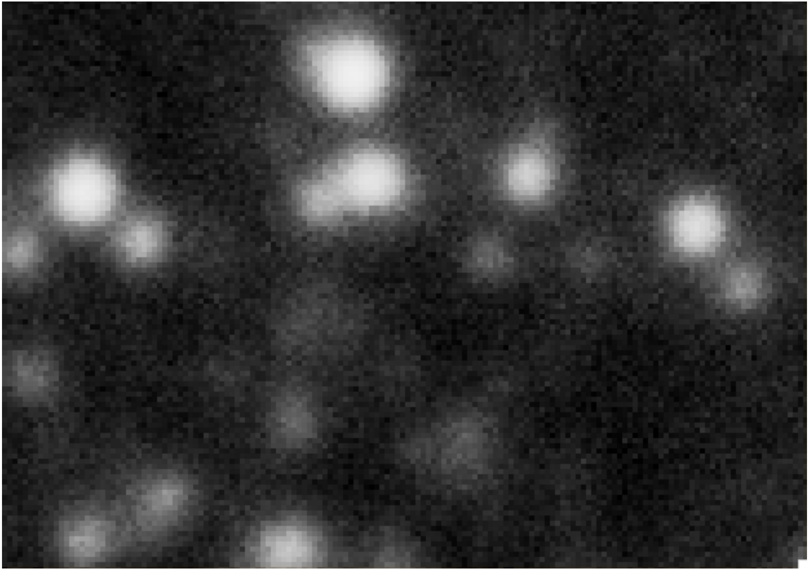
Exposure 1b



Catalogues

The classic solution to problem of searching for variability is to measure the brightness of each source in both images and compare the resulting catalogues.

Exposure 2a



Exposure 3b

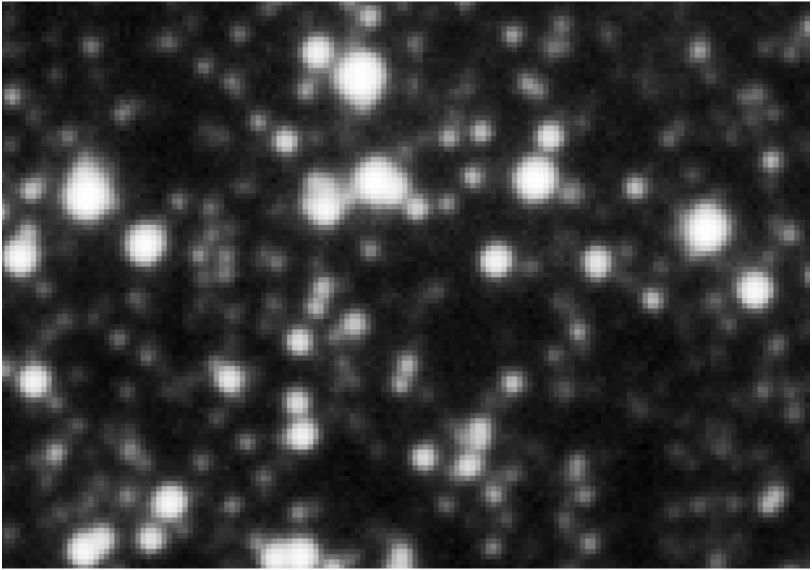


Image Subtraction

An obvious alternative is to subtract the two images directly, allowing for the difference between the seeing in the two images.

Given two images $I_a \equiv S \otimes \phi_a$ and $I_b \equiv S \otimes \phi_b$, we can write the Fourier-transform of the (seeing-matched) difference as $I_a(k) - I_b(k) \times (\phi_a(k)/\phi_b(k))$.

Unfortunately, it's difficult to measure the outer parts of the PSFs well enough to carry out this Fourier division.

What really matters is 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

$$R \equiv ||I_i - K \otimes I_b||$$

be minimised.

Constructing a linear system

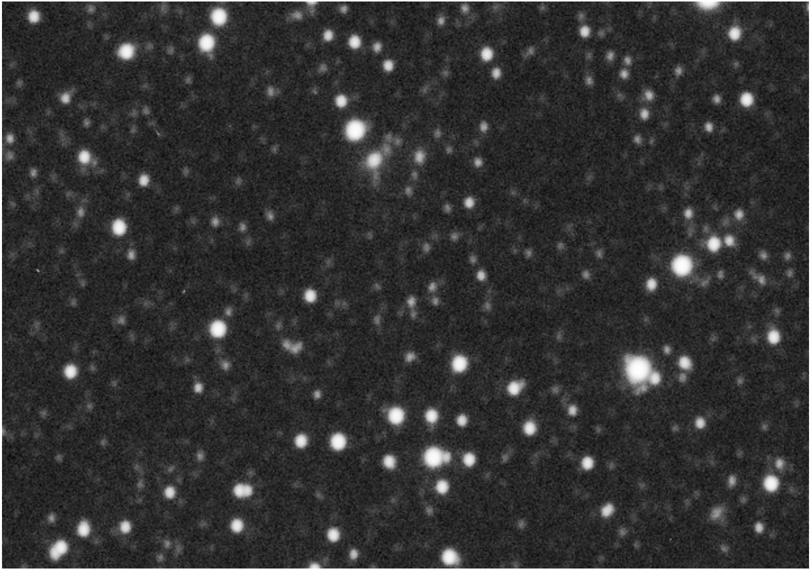
Let us write K as a sum of a set of basis functions:

$$K(u, v) \equiv \sum_r a_r B_r(u, v).$$

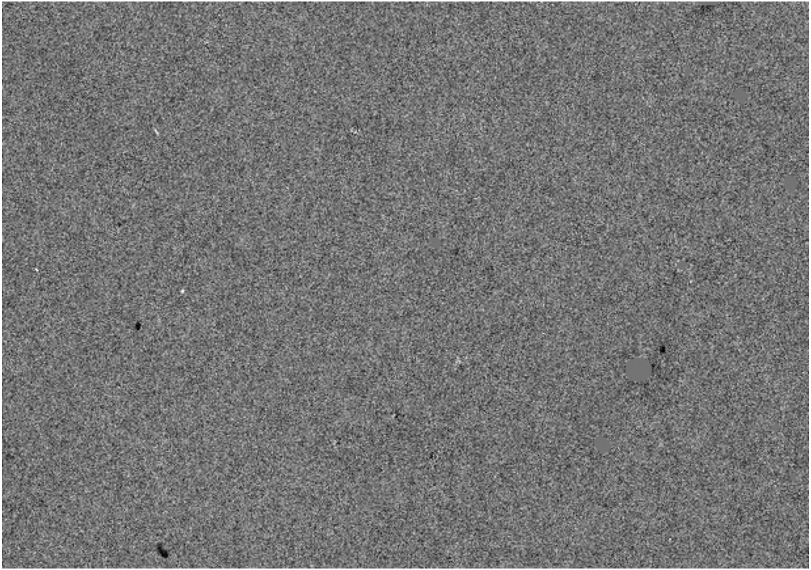
The task of subtracting two images is thus released to the problem of finding a set of a_r that minimise R; if we use an L_2 norm, this is a simple least-squares problem.
The most widely-used form for B_r is probably that originally proposed, Gaussians multiplied by polynomials. Various authors have considered using δ-function bases, but without conspicuous success.

Exposure1a

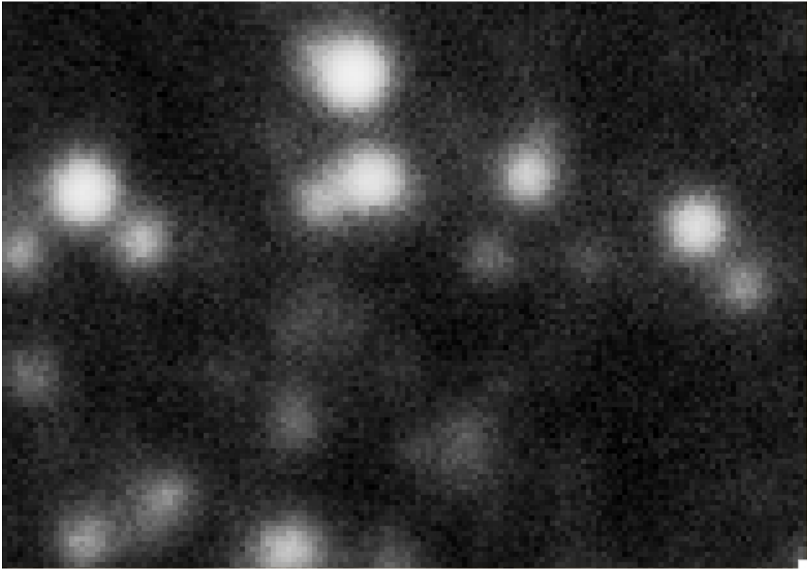
Exposure1b



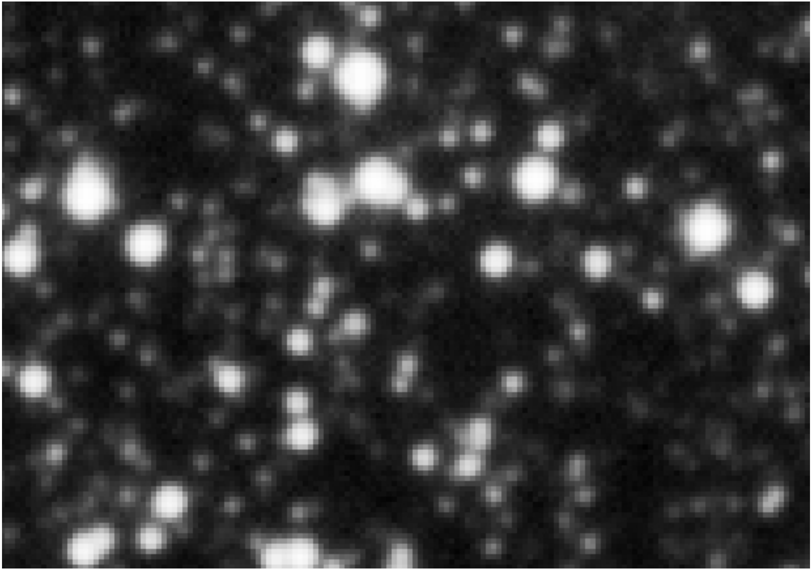
Difference Imaging 1



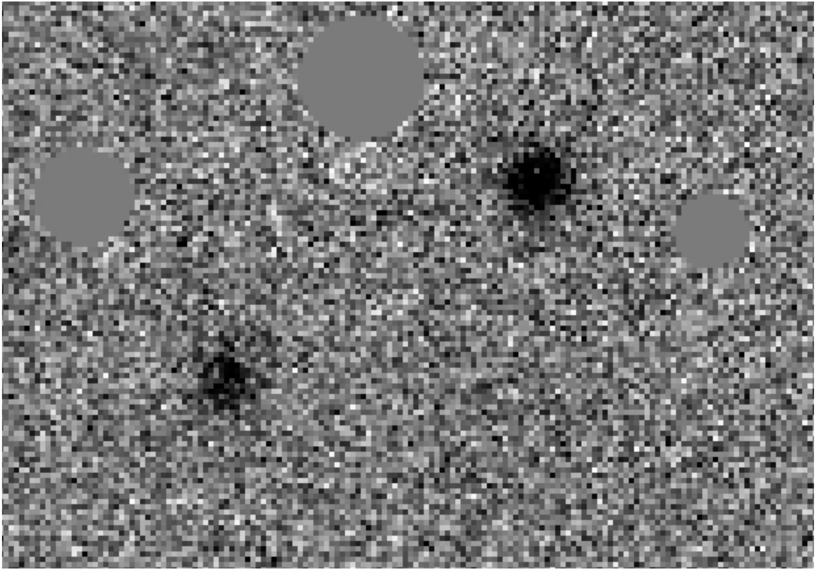
Exposure 2a



Exposure 2b



Difference Imaging 2



Spatially varying kernels

Spatially varying kernels can be handled by writing

$$K(u, v; x, y) = \sum_{r=1}^{r=n} \sum_{l=m=0}^{l+m \leq N} b^r_{lm} x^l_{(i)} y^m_{(i)} B_r(u, v)$$