DSFP: Image Processing

Robert Lupton

Princeton University

2016-08-02

Image Processing

Astronomical image processing is an enormous topic, and one that we'll return to in future Schools. For example, I worry about:

Image Processing

Astronomical image processing is an enormous topic, and one that we'll return to in future Schools. For example, I worry about:

- Correcting for the effects of CCD electronics on the data

Image Processing

Astronomical image processing is an enormous topic, and one that we'll return to in future Schools. For example, I worry about:

- Correcting for the effects of CCD electronics on the data
- Estimating the contamination due to moonlight and optical ghosts

Image Processing

Astronomical image processing is an enormous topic, and one that we'll return to in future Schools. For example, I worry about:

- Correcting for the effects of CCD electronics on the data
- Estimating the contamination due to moonlight and optical ghosts
- Finding the PSF given an image

Image Processing

Astronomical image processing is an enormous topic, and one that we'll return to in future Schools. For example, I worry about:

- Correcting for the effects of CCD electronics on the data
- Estimating the contamination due to moonlight and optical ghosts
- Finding the PSF given an image
- Looking for what has changed between pairs of images

Image Processing

Astronomical image processing is an enormous topic, and one that we'll return to in future Schools. For example, I worry about:

- Correcting for the effects of CCD electronics on the data
- Estimating the contamination due to moonlight and optical ghosts
- Finding the PSF given an image
- Looking for what has changed between pairs of images
- Measuring the fluxes and colours of faint galaxies

Image Processing

Astronomical image processing is an enormous topic, and one that we'll return to in future Schools. For example, I worry about:

- Correcting for the effects of CCD electronics on the data
- Estimating the contamination due to moonlight and optical ghosts
- Finding the PSF given an image
- Looking for what has changed between pairs of images
- Measuring the fluxes and colours of faint galaxies
- Disentangling sets of overlapping images

Image Processing

Astronomical image processing is an enormous topic, and one that we'll return to in future Schools. For example, I worry about:

- Correcting for the effects of CCD electronics on the data
- Estimating the contamination due to moonlight and optical ghosts
- Finding the PSF given an image
- Looking for what has changed between pairs of images
- Measuring the fluxes and colours of faint galaxies
- Disentangling sets of overlapping images
- Subtracting the OH lines from faint spectra
- …

Image Processing

Astronomical image processing is an enormous topic, and one that we'll return to in future Schools. For example, I worry about:

- Correcting for the effects of CCD electronics on the data
- Estimating the contamination due to moonlight and optical ghosts
- Finding the PSF given an image
- Looking for what has changed between pairs of images
- Measuring the fluxes and colours of faint galaxies
- Disentangling sets of overlapping images
- Subtracting the OH lines from faint spectra
- …

I'm only going to talk about an easier problem today.

How should I detect objects?

How should I detect objects?

Let's restrict our attention to stars.

How should I detect objects?

Let's restrict our attention to *isolated* stars.

How should I detect objects?

Let's restrict our attention to *isolated* stars.

We'll first answer a simpler question:

How should I detect objects?

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

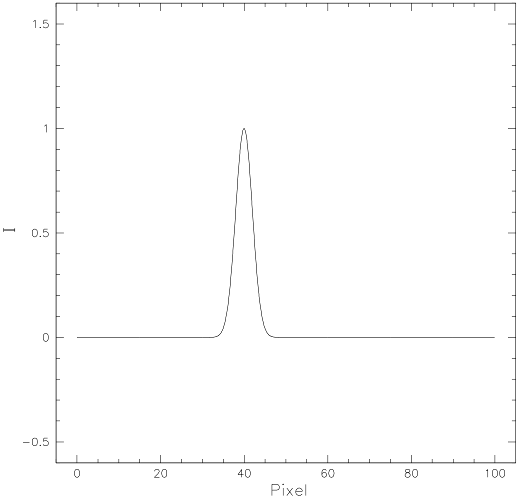
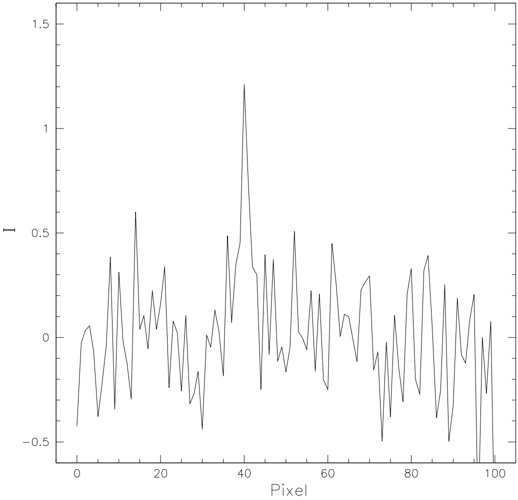
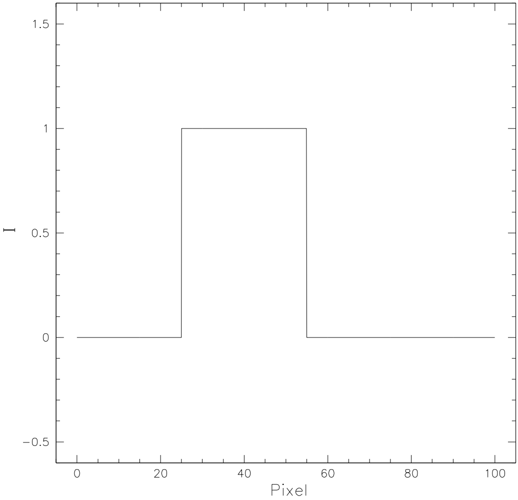


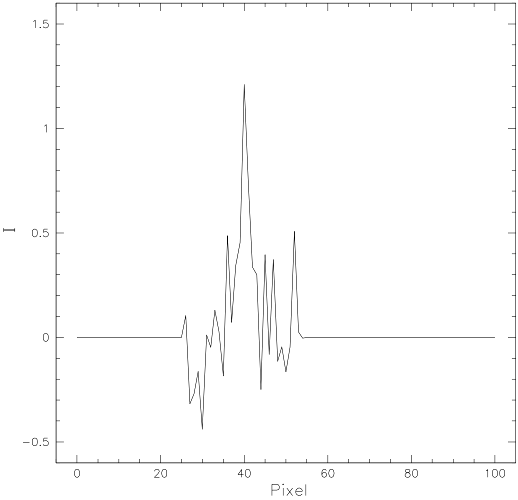
Image with noise



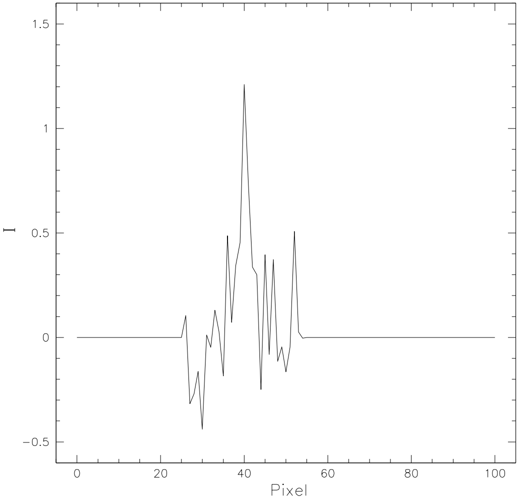
My Aperture



Aperture Flux with noise



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.

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:

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

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:

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

I.e.

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

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.

Background Estimation

I believe that background estimation is an unsolved problem.

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

Maximum Likelihood

As Maria explained this morning, there is a general way of taking a model of our data I and deriving asymptotically optimal estimates of the model parameters θ by the magic of Bayes' theorem.

Maximum Likelihood

As Maria explained this morning, there is a general way of taking a model of our data I and deriving asymptotically optimal estimates of the model parameters θ by the magic of Bayes' theorem.

$$P(\theta|I) = \frac{P(I|\theta)P(\theta)}{P(I)}$$

Maximum Likelihood

As Maria explained this morning, there is a general way of taking a model of our data I and deriving asymptotically optimal estimates of the model parameters θ by the magic of Bayes' theorem.

$$P(\theta|I) = \frac{P(I|\theta)P(\theta)}{P(I)}$$

i.e. calculating the *Likelihood* $P(I|\theta)$ (the probability of the data given the model) I get a handle on $P(\theta|I)$ (the probability of the model given the data).
Our model is that

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

where n is Gaussian with standard deviation σ. The only unknown parameter θ is F_0.

Maximum Likelihood

As Maria explained this morning, there is a general way of taking a model of our data I and deriving asymptotically optimal estimates of the model parameters θ by the magic of Bayes' theorem.

$$P(\theta|I) = \frac{P(I|\theta)P(\theta)}{P(I)}$$

i.e. calculating the *Likelihood* $P(I|\theta)$ (the probability of the data given the model) I get a handle on $P(\theta|I)$ (the probability of the model given the data).
Our model is that

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

where n is Gaussian with standard deviation σ. The only unknown parameter θ is F_0.
Rearranging, we find that

$$n(x) = F_0\,\phi(x - x_0) - I(x)$$

and this is a Gaussian, $N(0, \sigma(x)^2)$; *i.e.*

$$P(n(x)) = \frac{1}{\sqrt{2\pi\sigma(x)^2}} e^{-n(x)^2/2\sigma(x)^2}$$

Likelihoods

If we have pixellated data it's convenient to write $n_i \equiv n(x_i)$, so we can write the total likelihood of our data as

$$\mathcal{L}(I|F_0) = \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} \mathrm{e}^{-n_i^2/2\sigma_i^2}$$

but it's easier to work with the logarithm:

$$\ln \mathcal{L}(I|F_0) = -\frac{1}{2}\left(\sum_i \ln 2\pi\sigma_i^2 + \sum_i \frac{n_i^2}{\sigma_i^2}\right)$$

Likelihoods

If we have pixellated data it's convenient to write $n_i \equiv n(x_i)$, so we can write the total likelihood of our data as

$$\mathcal{L}(I|F_0) = \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-n_i^2/2\sigma_i^2}$$

but it's easier to work with the logarithm:

$$\ln \mathcal{L}(I|F_0) = -\frac{1}{2}\left(\sum_i \ln 2\pi\sigma_i^2 + \sum_i \frac{n_i^2}{\sigma_i^2}\right)$$

The values of σ_i are known so the first term is irrelevant, and we can write

$$\ln \mathcal{L}(I|F_0) \sim -\sum_i \frac{n_i^2}{\sigma_i^2}$$

Likelihoods

If we have pixellated data it's convenient to write $n_i \equiv n(x_i)$, so we can write the total likelihood of our data as

$$\mathcal{L}(I|F_0) = \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-n_i^2/2\sigma_i^2}$$

but it's easier to work with the logarithm:

$$\ln \mathcal{L}(I|F_0) = -\frac{1}{2}\left(\sum_i \ln 2\pi\sigma_i^2 + \sum_i \frac{n_i^2}{\sigma_i^2}\right)$$

The values of σ_i are known so the first term is irrelevant, and we can write

$$\ln \mathcal{L}(I|F_0) \sim -\sum_i \frac{n_i^2}{\sigma_i^2}$$

I'm going to ignore the priors for now.

Likelihoods

If we have pixellated data it's convenient to write $n_i \equiv n(x_i)$, so we can write the total likelihood of our data as

$$\mathcal{L}(I|F_0) = \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-n_i^2/2\sigma_i^2}$$

but it's easier to work with the logarithm:

$$\ln \mathcal{L}(I|F_0) = -\frac{1}{2}\left(\sum_i \ln 2\pi\sigma_i^2 + \sum_i \frac{n_i^2}{\sigma_i^2}\right)$$

The values of σ_i are known so the first term is irrelevant, and we can write

$$\ln \mathcal{L}(I|F_0) \sim -\sum_i \frac{n_i^2}{\sigma_i^2}$$

I'm going to ignore the priors for now. Unfortunately this isn't always possible.

Likelihoods

Substituting our formula for n this becomes

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

where $\sigma^2 = S + I$.

Likelihoods

Substituting our formula for *n* this becomes

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

where $\sigma^2 = S + I$. The MLE results from minimising that sum of squares. We've recovered a simple χ^2 estimator

Likelihoods

Substituting our formula for n this becomes

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

where $\sigma^2 = S + I$. The MLE results from minimising that sum of squares. We've recovered a simple χ^2 estimator -- but now we know it's because we assumed that the noise was Gaussian and chose an asymptotically optimal approach.

Likelihoods

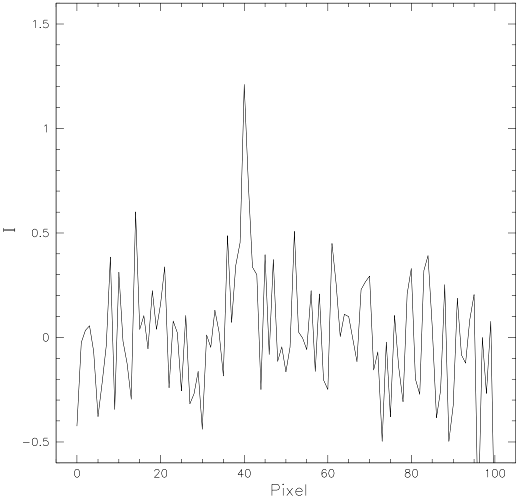
Substituting our formula for n this becomes

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

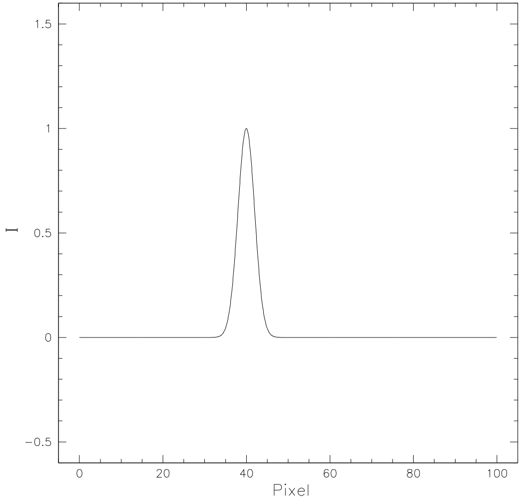
where $\sigma^2 = S + I$. The MLE results from minimising that sum of squares. We've recovered a simple χ^2 estimator -- but now we know it's because we assumed that the noise was Gaussian and chose an asymptotically optimal approach. Differentiation with respect to F_0 we find that this is maximised at

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

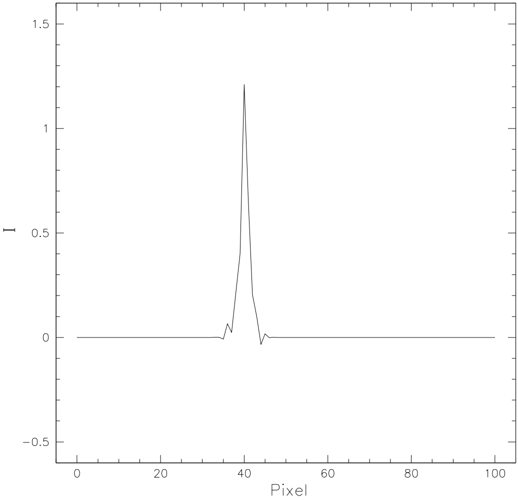
Image with noise



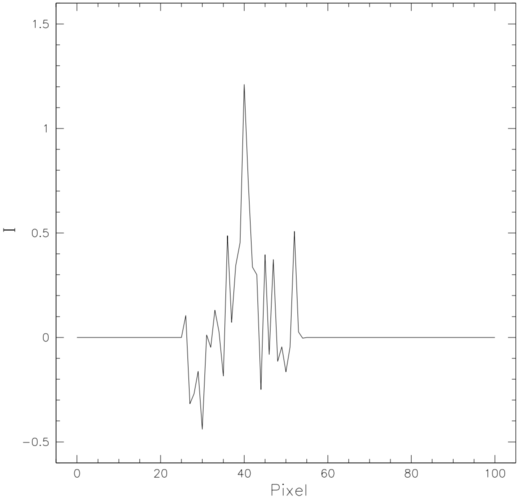
PSF



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

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

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.

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.

Measuring fluxes using the Psf

For faint sources (so all pixels have the same variance, $\sigma^2 \equiv S$), the flux is given by

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

So each photon is weighted by the PSF's profile, *i.e.* the probability that it belongs to the source.

Measuring fluxes using the Psf

For faint sources (so all pixels have the same variance, $\sigma^2 \equiv S$), the flux is given by

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

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.

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.

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.

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.
If you like algebra, you can take the expression

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

and substitute $\sigma_i^2 = F_0 \phi_i$ to find that

$$\hat{F_0} = \sum_i I_i$$

i.e. an aperture flux.