# The Restoration of Astronomical Images

**By R. C. PUETTER**[1]

[1]Center for Astrophysics and Space Sciences,
University of California, San Diego
9500 Gilman Drive
La Jolla, CA, 92093-0111, USA
rpuetter@ucsd.edu

This lecture presents practical applications of pixon-based Bayesian estimation to the restoration of astronomical images. It discusses the effects of noise, effects of finite sampling on resolution, and special problems associated with spatially correlated noise introduced by mosaicing. Comparisons to other methods demonstrate the significant improvements afforded by pixon-based methods and illustrate the science that such performance improvements allow.

## 1. Introduction

Image restoration is the improvement of the quality of image-like data. This includes noise filtering and resolution enhancement. Image restoration is to be distingushed from image reconstruction in which the underlying image is estimated from data which is not in image-like form, but which is complexly related to the underlying image. An example of image reconstruction is radio interferometry in which one deduces the underlying spatial structure of sources from a sparsely sampled interferogram. A common example of image restoration is Point Spread Function (PSF) deconvolution, and we shall concentrate on this application in the present lecture.

Before presenting examples of pixon-based restorations, however, we discuss several important issues which limit our ability to perform perfect PSF deconvolutions. These include the presence of random noise, the finite sampling represented by digital detector arrays, and spatial noise correlation caused by the co-addition of multiple image frames.

## 2. The Effects of Noise on Image Restoration

In the absence of noise, the PSF image deconvolution problem is completely (or nearly completely) invertable. Non-uniqueness of the solution only results because of pathologies in the PSF. If viewed in Fourier space, for example, one can see that there is a straightforward solution to the deconvolution problem, i.e. that given by equation (2.7) of Lecture 1. In the absence of noise, this solution is well defined and completely unambigous so long as the Fourier transform of the PSF has no zeroes. In this case, each Fourier component of the underlying image can be recovered, and since the Fourier basis is complete, the solution is known exactly. If the Fourier transform of the PSF has only a few zeroes, then the problem is not too much worse. Then nearly all of the Fourier components of the underlying image are known and there is only small ambiguity in the solution. In the presence of noise, however, things change dramatically. Then even if the Fourier transform of the PSF has no zeroes, one is unable to accurately determing the Fourier components of the underlying solution over large regions of Fourier space. This occurs whenever the product of the Fourier transforms of the PSF and data is small relative to the noise.

As discussed in Lecture 1, directly using the Fourier deconvolution formula is a bad idea for performing image reconstruction. While the above considerations are valid for any deconvolution algorithm, the Fourier deconvolution method suffers greatly from noise propagation problems. Calculation of the Fourier components of the data and PSF is an inherently noisy process, and results in an estimate of the Fourier components which is much noiser than what actually generated the measured data and PSF. In addition, in order to estimate the underlying image, this noise propagation is further enhanced by taking the Fourier inverse of the noisy quotient $F(data)/F(PSF)$.

As can be seen, we can place some fundamental limits on the information content of a PSF blurred image, i.e. we know at what level we begin to lose information concerning certain Fourier components in the data. However this is not very useful for several reasons. First, knowing that certain Fourier components of the data and PSF are noisy, does not give an intuitive handle on the nature of the limitations to one's knowledge of localized features in the image, since such features are comprised of a wide range of Fourier components. Second, we have already noted that taking the Fourier transform of the data and PSF is a noisy process. Hence a good image restoration algorithm should be able to recover an image with Fourier components significantly smaller than the noise level in the computed Fourier components. This leaves us with an unsatisfactory feeling, and we are left searching for a more practical method for assessing the limitations placed by noise on a given image restoration.

A direct solution to this problem is offered by multiresolution pixon-based image reconstruction. Since we are normally interested in the properties of localized features in an image, multiresolution pixon-based methods are a natural choice. The language itself is phrased in terms of structures of various scales, and the pixon map informs us of the level of believable structure based on a statistical analysis of the noise. In more general terms, we can use Bayesian estimation to answer any question we like about the solution. For example, if we wanted to know what the relative likelihood was of the image having properties $\{x_i\}$ relative to it having properties $\{y_i\}$, we would simply calculate $p(\{x_i\}|D)/p(\{y_j\}|D)$. The properties under discussion could be the width of a given structure, its peak value, its shape, or any combination of a variety of properties.

What, then, has Bayesian estimation in general, and pixon methods in particular, taught us? Multiresolution pixon methods have taught us that at any given place in the image, the resolution limit is given roughly by the smallest statistically justifiable scale. Hence such structure must be more than a 1-$\sigma$ event. A good rule of thumb is that one can see the presence of a lump above a smooth background only if it contains roughly 3-$\sigma$ of signal above the background. Figure 1 illustrates the nature of the detectability of structure in the presence of noise. Displayed is a particular noisy data set along with the underlying true structure. On the figure is a feature which is labeled as detectable structure. This feature is detectable because relative to the underlying background, the signal contained in the bump is statistically significant, i.e. greater than 3-$\sigma$. Hence if one were to use multiresolution pixons to assess the local smoothness scale at this position, one would find Bayesian evidence for structure on the size scale of this bump.

## 3. Finite, Discrete Sampling and Image Restoration

A typical assumption made by many workers in the field of image reconstruction is that the finest resolution obtainable with a digital image is given by 1 pixel (some would say 2 pixels). This, however, is incorrect. There have been a number of studies that have demonstrated that digital images contain information at sub-pixel scales (see, for example, Weir and Djorgovsky 1990). The ability to estimate sub-pixel scale structure is illustrated in Figure 2. In this figure we present 3 sample situations. In the first case (illustrated by the underlying source and data set in the lefthand column), the

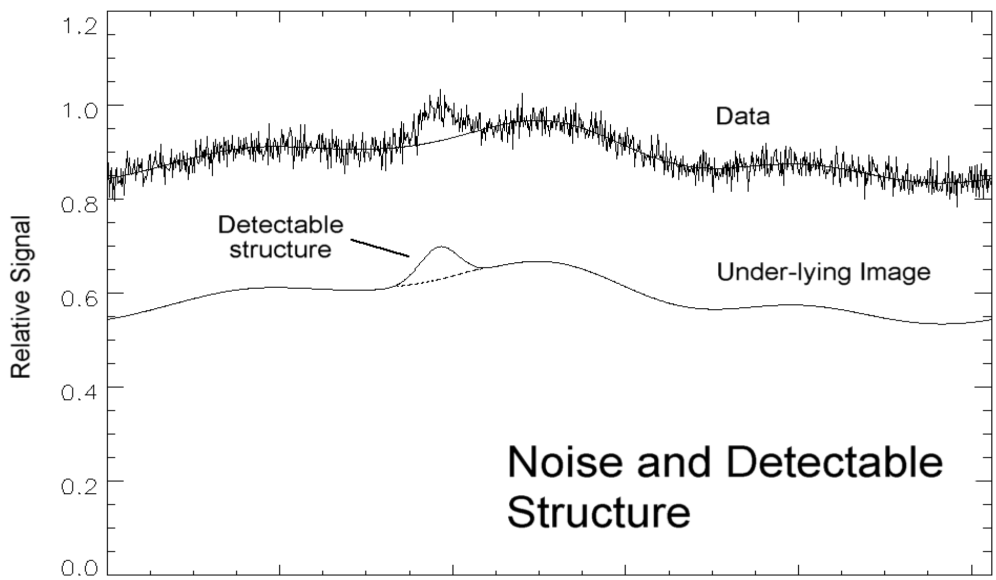


FIGURE 1. Ability to detect structure in the presence of noise. Pixon-based methods detect the structural scale present in an image by looking for Bayesian evidence of this scale.

true source uniformly fills a single pixel. When convolved with the PSF, the resulting data sampled with a resolution of 1 pixel is shown in the lower lefthand panel. If the underlying source had smaller spatial extent, e.g. filled a $1/3 \times 1/3$ pixel corner of a pixel, yet retained the same total flux, the situation would be as illustrated in the central column of Figure 2. In this case the resulting data looks significantly different, and is easily distingushable from the uniform source case. Thus the data set produced by the $1/3 \times 1/3$ pixel source contains information about sub-pixel sized structure. In some sitututations, however, sources with clear sub-pixel scale structure produce data which are ambigous. Such a case is ilustrated in the righthand column of Figure 2. Here the underlying source has sub-pixel sized structure, but the resulting data is identical to that produced by the uniform source.

While it appears that some sources are unable to "express" any evidence of sub-pixel scale structure, the real problem has to due with the particular placement of the sampling function. For example, for the cases presented in Figure 2, the ambiguity between the situations given in the lefthand and righthand columns can be removed by changing the sampling centers. Simply put, each of these the source distributions behaves differently as a top-hat sampling function is dragged over them; each case has a distinguishable response function, $R_i(\vec{x}) = \int dV_{\vec{y}} H(|\vec{y} - \vec{x}|) S_i(\vec{y})$, where $H$ is the top-hat sampling function of 1 pixel width and $S_i$ are the different source functions. Ambiguity is only present if the values for $R_i$ are given at only a few values of $\vec{y}$, or if the sampling function is pathological—deduction of sub-pixel scale structure is mathematically identical to the deconvolution problem which is usually invertable. This is quite valuable information. It tells us how to improve our ability to deduce sub-pixel scale structure. We can remove sub-pixel scale structure ambiguity by dithering our sampling function over the image. Indeed, if the signal-to-noise is high enough, the sampling dense enough, and the sampling function is not pathological, this approach should result in infinite resolution.

While we can now see that digitally sampled data contains sub-pixel scale information, and we have identified a manner in which to extract this information (i.e. dithering the

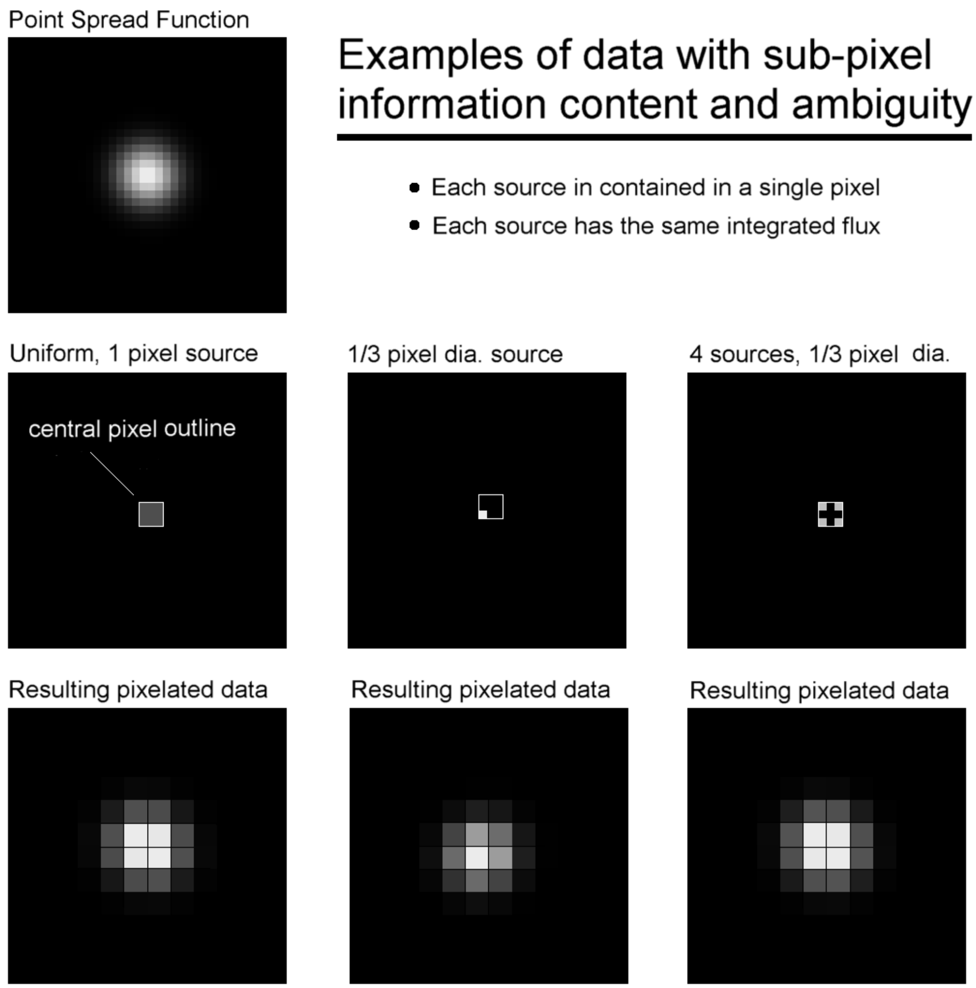


FIGURE 2. Illustration that pixelized data contains sub-pixel information. Upper left panel: PSF (Gaussian, FWHM = 2 pixels). Bottom two rows: sources and resulting data sampled at the pixel scale. Lefthand column: source uniformly fills a pixel with unit integrated flux. Center column: source fills the lower-left 1/9th of a pixel and has unit integrated flux. Right-hand column: source fills the 4, 1/9th pixel corners of a pixel and has unit integrated flux.

image sampling), in practical situations the spatial resolution typically is limited to being on the order of the pixel scale or perhap a factor of a few smaller. This is because noise counteracts our attempts to remove ambiguity. It makes each of the dithered images look similar, except in the case of the very brightest sources. Because high signal-to-noise ratios are required to resolve fine scale structure, it is usually only practical to measure a few dithered images. However, application of Bayesian methods makes the most use of this information. Bayesian methods allow us to select the most probable solution consistent with all of the information at our disposal—in this case the dithered images. So unlike direct inversions methods, which typically suffer terribly when information is incomplete, Bayesian methods work exceedingly well. Pixon-based methods, being Bayesian in their formulation, share these properties. In addition, since the srtuctural scale is part of the estimation process in pixon-based reconstruction, this method reports

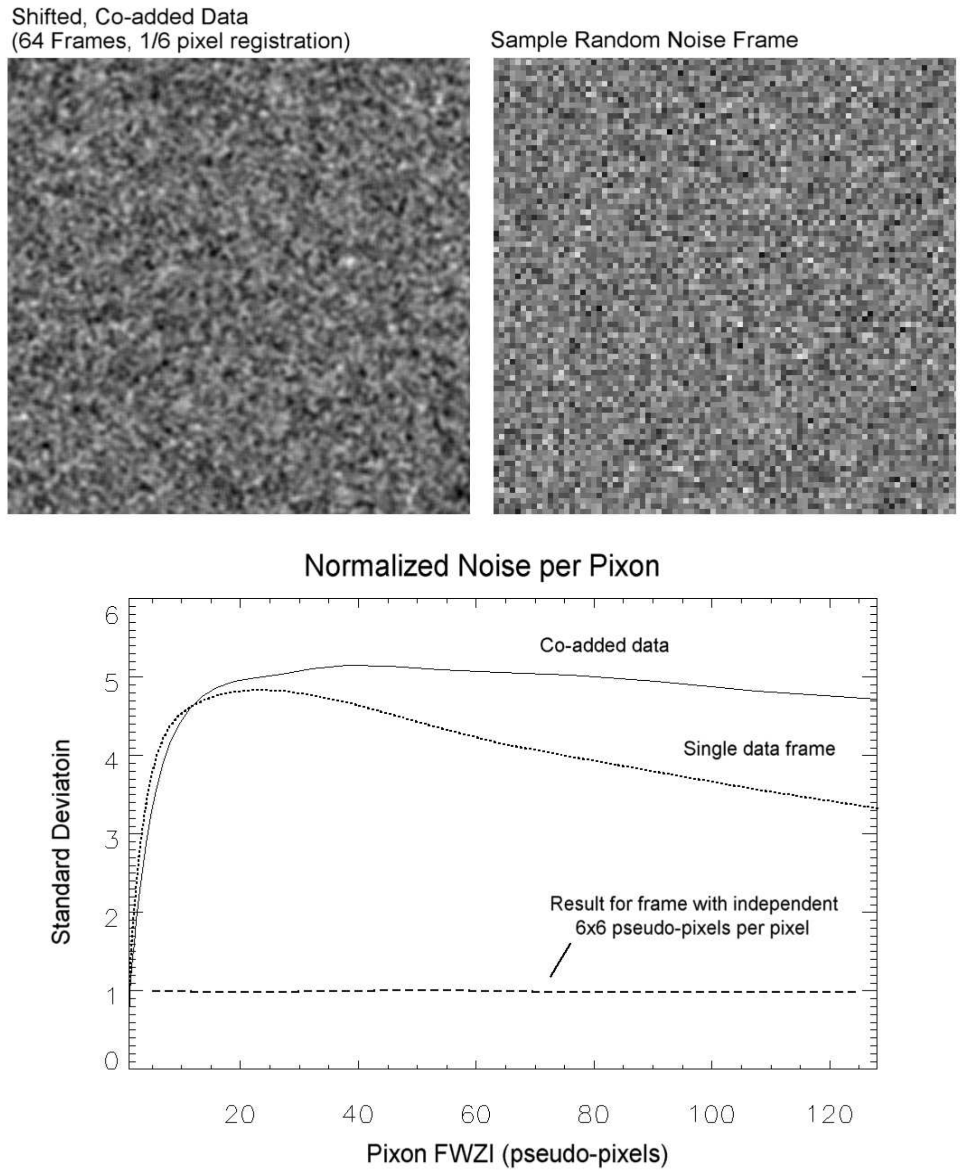


FIGURE 3. Illustration of the effects of sub-pixel registration on the spatial correlation of noise. Top left panel: Noise generated by randomly registering 64 independent noise frames to 1/6 of a pixel. Top right panel: A sample of one of the 64 noise frames used to make up the top left panel noise frame. Bottom plot: Noise per pixon as a function of pixon FWZI.

(through the pixon map) the resolution achieved at each point in the image. Hence one is always aware of the level of spatial structure ambiguity since this is directly estimated.

## 4. Spatially Correlated Noise

A finally problem, often associated with digital imaging, is the introduction of spatially correlated noise by mosaicing a number of independent data frames. Of course, the best method of performing an image restoration is to use each of the data sets directly, and not to use an intermediate data product such as a mosaiced data frame. If this prescription is followed, then the noise properties of the data are simple and clear. Each data frame has independent (and hopefully random) noise. This can be dealt with easily. However, often dealing with a large number of data frames in a restoration is computationally too expensive. This is the major reason for coadding or mosaicing all of one's images into a single data frame. One must clearly understand, however, that if this is done, then the noise properties of the resulting mosaic are very complex.

To illustrate the non-Gaussian nature of mosaiced or frame-coadded data, we present the results of Figure 3. In the example illustrated here, we have taken 64 independent, random Gaussian noise frames, shifted them randomly on a grid with 1/6th pixel centers (call this the pseudo-pixel grid) and coadded them. The maximum displacement of one data frame relative to the other was 108 pseudo-pixels (108 pseudo-pixels = 6 pseudo-pixels/pixel $\times$ 18 pixels), and the 64 displacements were randomly and uniformly distributed over these 108 positions in both the $x$ and $y$ directions. The result is presented in the upper lefthand image of Figure 3. A sample of one of the 64 Gaussian noise frames used to make up the mosaiced image is presented in the upper righthand image. The full pixel grid size is apparent in this image.

As can be seen from the mosaiced image, the noise is far from Gaussian, i.e. a random Gaussian noise field at the pseudo-pixel scale. Despite the randomness of coadding a large number of images, the noise is clearly spatially correlated. Furthermore, this spatial correlation does not go away as the number of coadded frames goes to infinity. It is a feature of coadded noise. The properties of this noise can be analyzed in a number of ways. One could, for example, look at the Fourier power spectrum of the noise. However, this is not what we are interested in here. We are interested in the effects that this noise has on our ability to detect localized structure. We are not interested in how this noise affects our ability to measure sine waves. Hence the appropriate analysis is to study the noise as a function of structural scale, i.e. a pixon-based analysis. This is what we have done in the bottom plot presented in Figure 3. In this diagram we plot the noise contained in a given radially symmetric, truncated parabolic pixon as a function of pixon full width at zero intensity (FWZI), i.e. we plot $\sigma_{N,pixon}/\sigma_{G,pixon}$ where

$$\sigma_{N,pixon} = \sqrt{\sum_i \frac{1}{N_{pseudo-pixels}} \left( \sum_j k(\vec{x_i}, \vec{y_j}, \delta(\vec{x_i})) N(\vec{y_j}) \right)^2} \quad , \tag{4.1}$$

$$k(\vec{x}, \vec{y}, \delta(\vec{x})) = 1 - \frac{|\vec{x} - \vec{y}|^2}{\delta(\vec{x})^2} \quad , \tag{4.2}$$

$$\sigma_{G,pixon} = \sigma \sqrt{\sum_j k(\vec{x_i}, \vec{y_j}, \delta(\vec{x_i}))} \quad , \tag{4.3}$$

where summations are over the pseudo-pixel grid, $N$ is the noise frame in question, $\delta$ is the pixon FWZI, $\sigma_{G,pixon}$ is the value of $\sigma_{N,pixon}$ for the case in which the noise is a random Gaussian noise field on the pseudo-pixel grid, and $\sigma$ is the standard deviation of the noise at the pixel scale..

Note that in the diagram at the bottom of Figure 3, we have presented results for both a random Gaussian field over the pseudo-pixel grid (dashed line—the value is expected to be 1 since the noise is normalized to this case), as well as for the shifted co-added data, and the single sample Gaussian noise frame at the pixel resolution. Note that both of these cases have reduced noise below the pixel scale (i.e. 6 pseudo-pixels). This property is directly apparent from visual inspection of the images. Again, this is valuable knowledge. Co-adding data affects the spatial structure spectrum of the noise. In calculating a pixon map, this noise spectrum must be acknowledged. There is less noise at small scales. Of course at sub-pixel scales there usually is also less signal—the signature of sub-pixel scale structure is relatively weak in noisy dithered, mosaiced images. Use of the appropriate noise level for each pixon scale, however, is necessary if one wishes to achieve an optimum restoration of the structure present in the data at each scale.

## 5. Examples of Pixon-Based Reconstruction

In this section we present 2 pixon-based image reconstructions. The first example compares pixon-based reconstructions to Maximum Entropy methods for a mock data set, i.e. a completely made up data set in which the underlying image, PSF, and noise frames are manufactured and hence known precisely. Mock data sets are useful in comparisons of image restoration/reconstruction methods since there can be no argument about the goal of the reconstruction. The true, underlying image is known perfectly *a priori*. Furthermore, the noise and all parameters of how the input data was made are completely specified. The second example presents an elliptical pixon-based reconstruction of Keck $2\mu m$ imaging data for the gravitational lens of FSC 10214+4724. This example illustrates the use of non-radially symmetry pixon bases. We shall go into some detail in the discussion of this example principally because most pixon reconstructions presented in previous work have used radially symmetric basis functions and because the science allowed by the reconstruction is quite interesting.

### 5.1. *Mock Data*

We present in this section pixon-based and MEMSYS 5 restorations of a mock-data set. For the comparisons presented here, the ME algorithms chosen are those embodied in MEMSYS 5, the most current release of the MEMSYS algorithms. The MEMSYS code represent a powerful set of ME algorithms developed by Gull and Skilling (Gull and Skilling 1991). The MEMSYS algorithms probably represent the best commercial software package available for image reconstruction. The MEMSYS reconstruction was performed by Nick Weir of Caltech, a recognized MEMSYS expert, and were supplemented with his multi-channel correlation method which has been shown to enhance the quality of MEMSYS reconstructions (Weir 1991; Weir 1994). The true, noise-free, unblurred image presented in the top row is constructed from a broad, low-level elliptical Gaussian (i.e. a 2-dimensional Gaussian with different FWHMs in perpendicular directions), and 2 additional narrow, radially symmetric Gaussians. One of these narrow Gaussians is added as a peak on top of the low-level Gaussian. The other is subtracted to make a hole. To produce the input image, the true image was convolved with a Gaussian PSF of FWHM=6 pixels, then combined with a Gaussian noise realization. The resulting input image is displayed in the top row. The signal-to-noise ratio on the narrow Gaussian spike is roughly 30. The signal-to-noise on the peak of the low level Gaussian is about 20. The signal-to-noise at the bottom of the Gaussian "hole" is 12.

As can be seen, the FPB reconstruction is superior to the multi-channel MEMSYS result. The FPB reconstruction is free of the low-level spurious sources evident in the MEMSYS 5 reconstruction. These false sources are due to the presence of unconstrained degrees of freedom in the MEMSYS 5 reconstruction and are superimposed over the entire image, not just in the low signal to noise portions of the image. Furthermore, the FPB

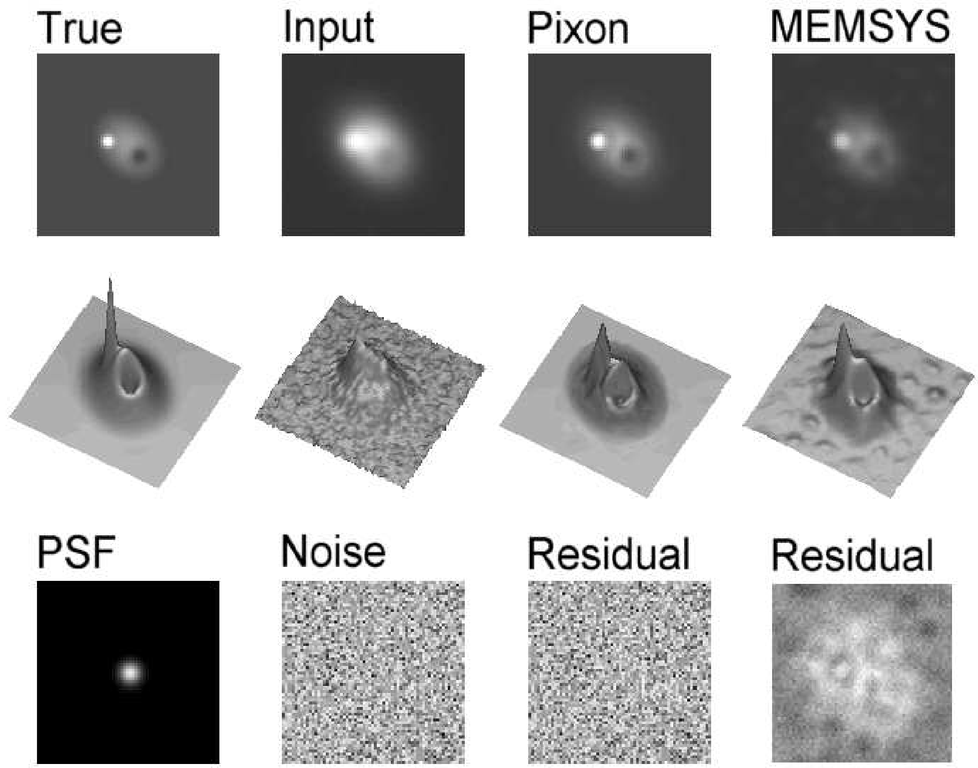


FIGURE 4. Comparison of pixon-based image restoration with Maximum Entropy restoration for a mock data set. Shown is the PSF (Gaussian with FWHM=6 pixels), the true, underlying image, and the Gaussian noise. The restored image and residuals for radially symmetric, truncated parabolic fuzzy pixons are shown. Also shown are the results and residuals for the multi-channel MEMSYS 5 method. Reproduced from Puetter 1994.

reconstruction's residuals show no spatially correlated structure, while the MEMSYS 5 reconstruction systematically under-estimates the signal, resulting in biased photometry.

### 5.2. *The Gravitational Lens of FSC 10214+4724*

We next discuss the application of pixon-based methods to image restoration of Keck K-band ($2.3\mu m$) imaging of the gravitational lens system FSC 10214+7424 (data presented in Graham and Liu 1995). This object was initially thought to be the most luminous object in the universe with a luminosity approaching $10^{15}$ solar luminosities. However, it is now recognized as a gravitational lens with a magnificantion of roughly 100 (Broadhurst and Lehar 1995; Close et al. 1995; Graham and Liu 1995). Hence the luminosity is about 100 times fainter than initially estimated.

In order to make sure that our image description language would not be biased against faint Einstein rings, and was able to describe inherently thin structures, we decided it would be wise to use a set of elliptical pixon shape function as our language basis. Our elliptical pixon language is displayed in Figure 5 and contains 196 elliptical pixon shape functions with 6 different ellipticities, having major to minor axis ratios logarithmically spaced between 1 and 7.25. For each ellipse there are a number of major axis orientation angles. The ellipses of high ellipticity have a larger number of orientations.

Our elliptical pixon restoration of the Keck K-band image is shown in Figure 6. Also displayed in this figure is the Maximum Entropy reconstruction of Graham and Liu

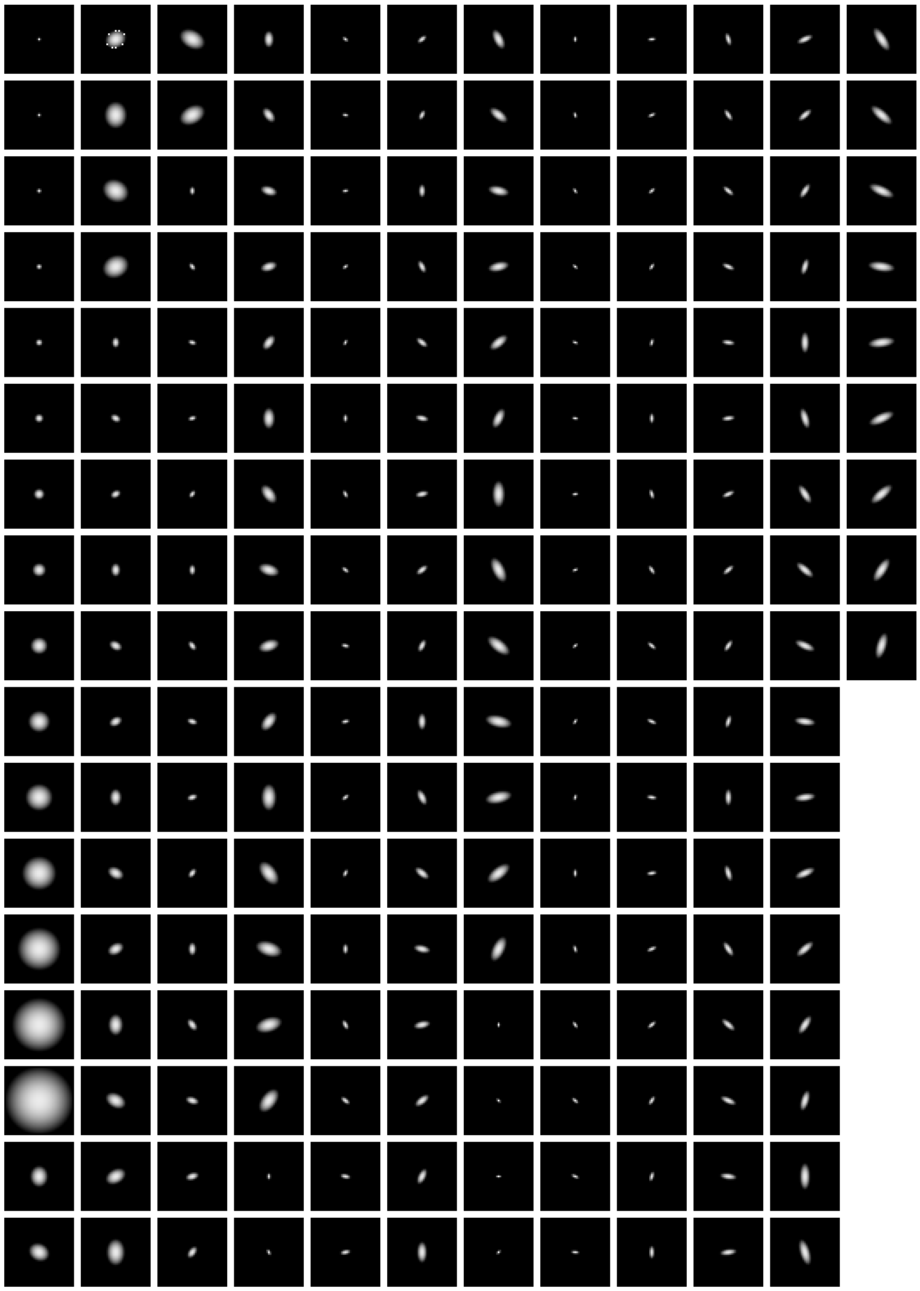

FIGURE 5. The elliptical language used to reconstruct the image of FSC 10214+7424.

(Graham and Liu 1995). As can be seen from this figure, our reconstruction reveals the entire Einstein ring. This low-level ring was not detectable by the Maximum Entropy methods employed by Graham and Liu because of the large number and strength of the spurious sources produced by that method. These sources are obvious in the Maximum Entropy result displayed in Figure 6. We have also plotted in Figure 7 our reconstruction overlaid with the local pixon shape functions. In this image one can easily see how the elliptical pixons follow the the image's spatial structure (information content). Smaller

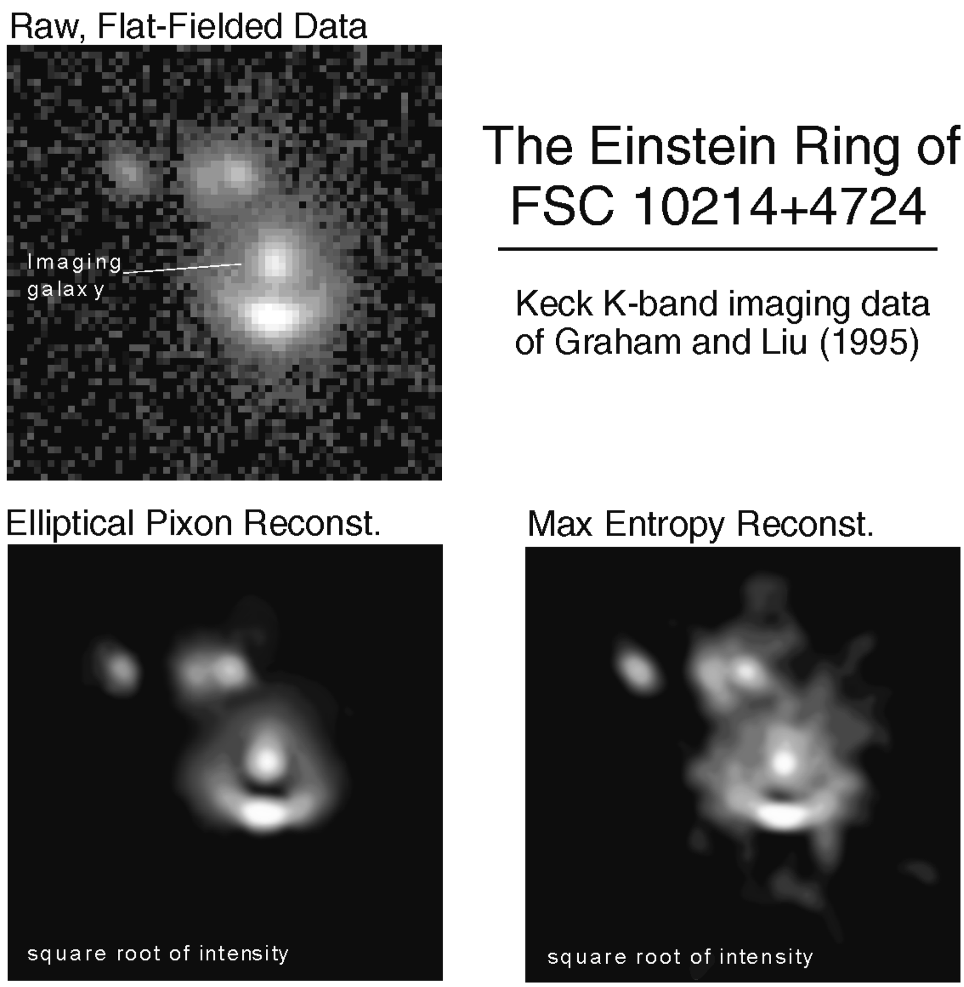


FIGURE 6. Comparison of image reconstruction techniques for the Einstein ring of FSC 10214+4724. The elliptical pixon reconstruction used 196 different elliptical pixon kernel functions. Relative to radially symmetric reconstructions, the elliptical pixon reconstruction used roughly 50% fewer DOFs in the image model. Reproduced from Puetter 1996.

pixons are used in locations of detailed structure. Larger pixons are used in regions in which the image is smooth, and the ellipses align with the structures present so that the largest pixon can be used to describe the image at each location. As argued in Lectures 2 and 3, relative to simpler languages, e.g. radially symmetric pixons, this richer elliptical pixon language should be able to describe the image using fewer degrees-of-freedom. We also have performed a reconstruction for this data with radially symmetric pixons and find that the elliptical pixon restoration uses half as many pixons.

The science that our better reconstruction allows for this data is quite interesting. First, Graham and Liu concluded on the basis of their inability to see an entire Einstein ring, that the size of the emitting source (which has a redshift of 2.3) is 0.25 arcsec or less. They point out that given the lens geometry, if a complete ring was seen, then the size of the emitting region would have to be at least 0.5 arcsec. Hence our first conclusion is that the emitting region of FSC 10214+7424 has significant emission out

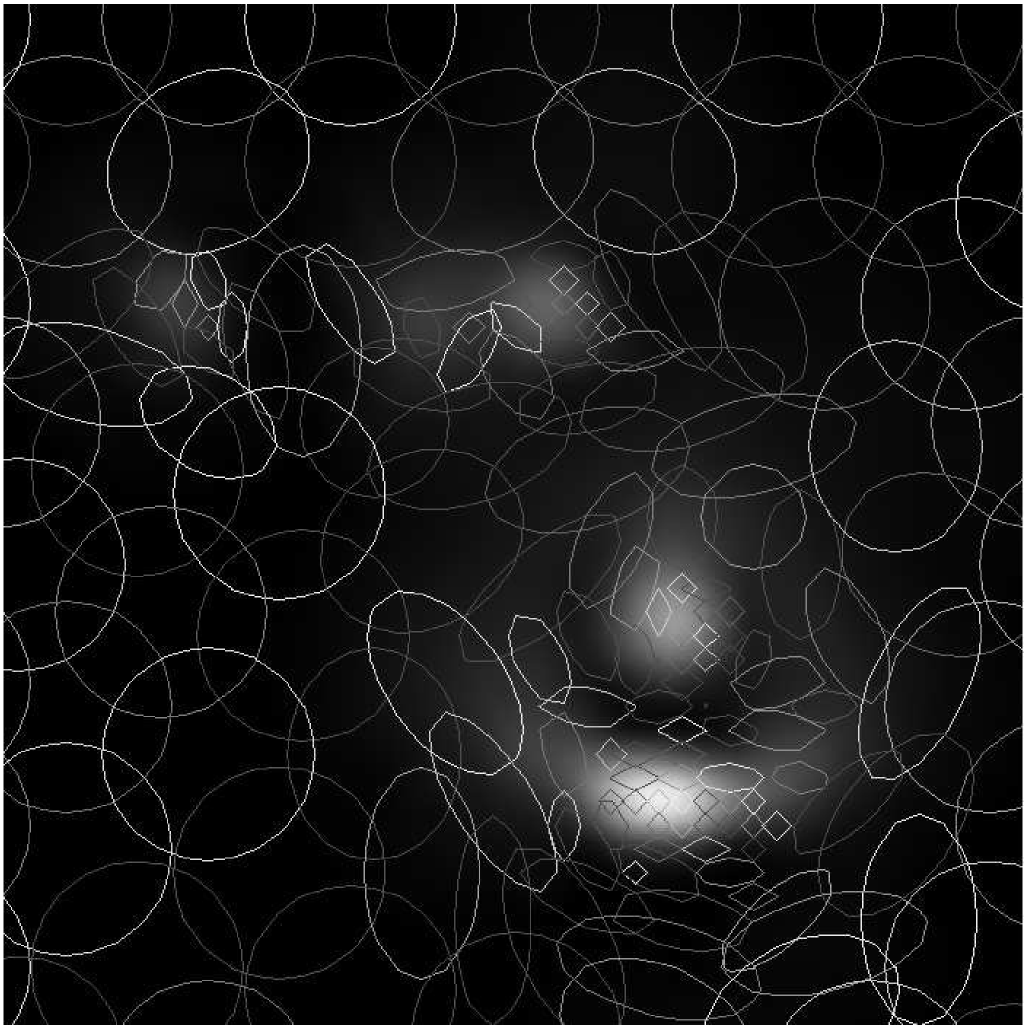

FIGURE 7. Reconstruction of the Einstein ring of FSC 10214+7424 with overlaid pixon kernel functions. Note that for a fuzzy pixon reconstruction a pixon kernel shape is assigned to each pseudo-pixel in the image. However plotting this many kernel functions would be confusing to the eye. Thus for the purposes of illustration, we have plotted a new pixon kernel function shape such that the area overlap is 25%. Pixon kernel contours are plotted at 50% of maximum.

to a size of 0.5 arcsec, roughly twice that deduced by Graham and Liu. Our second result is related to the nature of the gravitational lens. Recent HST images of the Einstein ring at far-UV rest-frame wavelengths ($0.25\mu m$) were obtained by Eisenhardt et al. (Eisenhardt et al. 1996). From the narrow nature of the arc, these authors conclude that the size of the UV source is of order 0.005 arcsec (40 pc), and that the best fit to the shape of the arc requires a contribution to the gravitational lensing potential from the next brightest elliptical galaxy near the main lensing galaxy. Our restored image directly confirms the gravitational contribution of this object to the lensing potential. The K-band Einstein ring is clearly warped in the direction of this source, directly confirming the speculations of Eisenhadt et al. Further discussion of these results will appear in a paper which is in preparation.

## 6. Conclusions

This lecture has discussed the effects of random noise on detectable structure in a restoration, the effects of discrete sampling, and the nature and effects of spatially correlated noise caused by data frame co-adding. We have outlined a rule of thumb to estimate the level at which random noise limits spatial resolution and discussed how image dithering can greatly improve our ability to detect sub-pixel scale structure. We have also described the nature of the spatial correlation resulting from data frame coadding and have indicated how these effects can be incorporated into pixon-based methods. Finally, we have presented 2 sample restorations. The first was of a mock data set and directly demonstrated the advantages of pixon-based methods over the best Maximum Entropy methods. The second was an elliptical pixon-based reconstruction of K-band imaging data of the gravitational lens system FSC 10214+7424. This example also confirmed the greater sensitivity afforded by pixon-based methods relative to Maximum Entropy methods. Furthermore we illustrated how these benefits translate directly into science.

## 7. Acknowledgements

The author would like to thank the many people who have contributed over the years toward making pixon-based methods a reality. In this regard I would like to especially thank Robert Piña, the co-inventor of the pixon. I would also like to thank James Graham for providing his imaging data for FSC 10214+7424 and Amos Yahil for valuable discussions during the development of the elliptical pixon-based code.

REFERENCES

BROADHURST, T.; LEHAR, J. 1995, A Gravitational Lens Solution for the IRAS Galaxy FSC 10214+4724, *Ap. J. (Letters)*, **450**, pp. L41-4.

CLOSE, L.M.; HALL, P.B.; LIU, C.T.; HEGE, E.K. 1995, Spectroscopic and Morphological Evidence that IRAS FSC 10214+4724 is a Gravitational Lens, *Ap. J. (Letters)*, **452**, pp. L9-12.

EISEN HADRT, P. R., ARMUS, L., HOGG, D. W., SOIFER, B. T., NEUGEBAUER, G., AND WERNER, M. W. 1996, HST Observations of the Luminous IRAS Source FSC 10214+7424: A Gravitaionally Lensed Infrared Quasar, *Ap. J.*, in press.

GRAHAM, J.R.; LIU, M.C. 1995, High-Resolution Infrared Imaging of FSC 10214+4724: Evidence for Gravitational Lensing, *Ap. J. (Letters)*, **449**, pp. L29-32.

GULL, S. F., AND SKILLING, J. 1991, *MemSys5 Quantified Maximum Entropy User's Manual.*

PUETTER, R.C. 1994, Pixons and Bayesian Image Reconstruction, *Proc. S.P.I.E.* **2302**, pp. 112-131.

PUETTER, R. C. 1996, Pixon-Based Multiresolution Image Reconstruction and Quantification of Image Information Content, *Proc. MaxEnt '95*, in press.

WEIR, N. 1991, Applications of Maximum Entropy Techniques to HST Data, in Proc. of the ESO/ST-ECF Data Analysis Workshop, April 1991, ed.s P. Grosbo and R. H. Warmels, p. 115. ESO

WEIR, N. 1994, A Maximum Entropy-Based Model for Reconstructing Distributions with Correlations at Multiple Scales, *J. Opt. Soc. Am.*, in press.

WEIR, N. AND DJORGOVSKY, S. 1990, MEM: New Techniques, Applications, and Photometry, in *The Restoration of HST Images and Spectra*, Proc. of a Workshop at the Space Telescope Science Institute, 20-21 August 1990, Baltimore, ML, ed.s R. L. White and R. J. Allen, pp. 31-38.