

Wavefront Sensing

June 22, 2014

1 Introduction

In the study of a control algorithm for a high resolution imaging system, a knowledge of the sensor is required. Although the control algorithm is interested in the wavefront aberrations, the sensor measures only the intensity and information about the wavefront must be derived from there. The sensor model developed in this project expresses the intensity distribution through the point spread function (PSF). This document will discuss the relationship of the PSF and (1) the Fraunhofer approximation, (2) the effect of using a discrete Fourier transform and (3) the imaging system parameters - including focal length, lens diameter and wavelength. Further, the results of selecting a non-circular pupil function and the addition of noise to the model is discussed. Once the model for a single lens has been developed, it is extended to a lenslet array to model a Shack-Hartmann sensor.

2 Fraunhofer Approximation

The Fraunhofer approximation is a more stringent approximation of the Fresnel approximation which is valid when

$$z \gg \frac{2\pi(\xi^2 + \eta^2)_{\max}}{2\lambda}, \quad (1)$$

where λ is the wavelength, z is the distance between the object and the lens and (ξ, η) are the 2-dimensional positions on the aperture plane. It is used to compute wave propagation in the far field, whereas in the near field, the more general Fresnel approximation is used to compute wave propagation. Another condition for the Fraunhofer approximation, which as well applies to the Fresnel approximation, is that the incident waves must be within the paraxial regime, meaning that the direction of propagation must be small towards the optical axis of the system. In other words, this is simply a restriction to small incident angles. The less stringent Fresnel Approximation condition to Equation (1) is

$$z > \frac{2D^2}{\lambda}, \quad (2)$$

where D is the dimension of the aperture. Although the distance z is required to be large, the Fraunhofer diffraction patterns can still be observed at distances much closer than implied by the above conditions.

The observed field strength $U(x, y)$ can be calculated by the Fraunhofer approximation as

$$U_f(x, y) = \frac{e^{j k z} e^{j \frac{k}{2z}(x^2 + y^2)}}{j \lambda z} \iint_{-\infty}^{\infty} U_i(\xi, \eta) e^{-j 2\pi(f_X \xi + f_Y \eta)} d\xi d\eta \Big|_{f_X = \frac{x}{\lambda z}, f_Y = \frac{y}{\lambda z}}, \quad (3)$$

with $k = 2\pi/\lambda$ the wavenumber and where $U_i(\xi, \eta)$ is the field distribution at the aperture. From this equation it can be seen that the Fraunhofer approximation for the field distribution at the focal plane ($U_f(x, y)$) is simply the Fourier Transform (FT) of the distribution $U(\xi, \eta)$ times a multiplicative factor, evaluated at frequencies f_X and f_Y .

2.1 Diffraction Patterns at the Focal Plane of a Single Lens

Here we assume a diffraction-limited system with the input directly placed against the lens, to model the lens system that is used to measure the intensity in the focal plane. For our purposes it is convenient to consider the input at the lens and exclude any propagation prior to the lens in order to model the sensor system itself. The input $U_i(\xi, \eta)$ then contains the disturbance induced by the atmosphere and is illuminated by the solar system which has a flat wavefront. The field distribution behind the lens is then calculated by

$$U'_i(\xi, \eta) = U_i(\xi, \eta) P(\xi, \eta) e^{-j \frac{k}{2f} (\xi^2 + \eta^2)}, \quad (4)$$

where $P(\xi, \eta)$ is the pupil function of the lens system and is defined by

$$P(\xi, \eta) = \begin{cases} 1 & \text{inside the aperture} \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Moreover the exponential function in Equation (4), is the phase change introduced by the lens with focal length f . Using the Fresnel diffraction formula the field distribution in the back focal plane of the lens can be found as

$$U_f(x, y) = \frac{e^{j k f} e^{j \frac{k}{2f}(x^2 + y^2)}}{j \lambda f} \iint_{-\infty}^{\infty} U_i(\xi, \eta) P(\xi, \eta) e^{-j 2\pi(f_X \xi + f_Y \eta)} d\xi d\eta \Big|_{f_X = \frac{x}{\lambda f}, f_Y = \frac{y}{\lambda f}}. \quad (6)$$

Hence the complex amplitude field distribution in the back focal plane of the lens is simply the Fraunhofer diffraction pattern seen in Equation (3), but with the propagation distance f (between the lens and the image) instead of z . Thus the Fraunhofer diffraction patterns can be observed although not satisfying the

before discussed conditions (1) and (2). The real interest is the intensity across the focal plane and is computed by

$$I_f(x, y) = |U_f(x, y)|^2. \quad (7)$$

Also the model of the lens system as derived here is studied under the condition that the illumination is monochromatic.

The above text is based on (**Goodman,2005**) and is also advised for further reading.

3 Discrete Fourier Transform

The computation of the FT will be done numerically. This can be done in a highly efficient algorithm called the Fast Fourier Transform (FFT). Using the Discrete Fourier Transform (DFT) it approximates the continuous FT of a function $g(x)$, where the FT of a two dimensional function is defined as

$$\hat{g}(f_X, f_Y) = \iint_{-\infty}^{\infty} g(x, y) e^{-j2\pi(f_X x + f_Y y)} dx dy. \quad (8)$$

As a first approximation, the integral is restricted to the finite interval $-L_x/2 \leq x \leq L_y/2$ and $-L_y/2 \leq y \leq L_y/2$, giving us, for $L = L_x = L_y$,

$$\hat{g}(f_X, f_Y) \approx \iint_{-L/2}^{L/2} g(x, y) e^{-j2\pi(f_X x + f_Y y)} dx dy. \quad (9)$$

Where the size L_x and L_y must be large enough to cover the essential non-zero parts of the function $g(x, y)$. Second, the integral is approximated by replacing it by a finite sum. Thus, we get

$$\hat{g}(f_X, f_Y) \approx \Delta y \sum_{m=N_y/2}^{N_y/2} e^{-j2\pi m f_Y \Delta y} \Delta x \sum_{n=-N_x/2}^{N_x/2} g(n \Delta x, m \Delta y) e^{-j2\pi n f_X \Delta x}, \quad (10)$$

where the x, y -coordinates are sampled in the previously described intervals with N_x, N_y samples, when $N = N_x = N_y$, spaced at $\Delta x = \Delta y = L/N$ at positions $x_n = n \Delta x$ and $y_m = m \Delta y$ with $n = m = -N/2, \dots, N/2 - 1$. A connection to the DFT can be made by sampling the spatial frequencies with N values in the interval $-\Delta x/2 \leq f_X \leq \Delta x/2$ and $-\Delta y/2 \leq f_Y \leq \Delta y/2$, the samples being spaced by $\Delta f_X = \Delta f_Y = 1/L$ at frequencies $f_k = k \Delta f_X$ and $f_l = l \Delta f_Y$, with $k = l = -N/2, \dots, N/2 - 1$. This produces the exact DFT described by

$$\hat{g}(k \Delta f_X, l \Delta f_Y) \approx \Delta y \sum_{m=N_y/2}^{N_y/2} e^{-j2\pi l m / N_y} \Delta x \sum_{n=-N_x/2}^{N_x/2} g(n \Delta x, m \Delta y) e^{-j2\pi k n / N_x}. \quad (11)$$

In short the procedure is:

1. sample the function $g(x, y)$ in a grid at values (x_n, y_m) and store the values in a matrix of size (N_x, N_y) ,
2. do a DFT on this matrix and multiply it with the sample spacings Δx and Δy ,
3. plot the resulting matrix of size (N_x, N_y) at points (f_k, f_l) in the spatial frequency space.

The numerical calculation of the diffraction pattern is now easily done by applying the FFT to $U_i(\xi, \eta)P(\xi, \eta)$, adding the multiplicative factor as seen in Equation (6), and setting $(f_k = \frac{x_k}{\lambda z}, f_l = \frac{y_l}{\lambda z})$. The intensity distribution is then found by taking the modulus squared.

3.1 Analytic and Numerical Solutions

In figure 1 the difference is seen between the numerical solved diffraction pattern and the analytically solved diffraction pattern. Clearly the two are almost identical except for the outer rings where the flaws of numerical computation are seen through the not-disk like shape in the airy function. Because only the first few orders of the airy disk are important, this is a good approximation to be used for calculating the centroids.

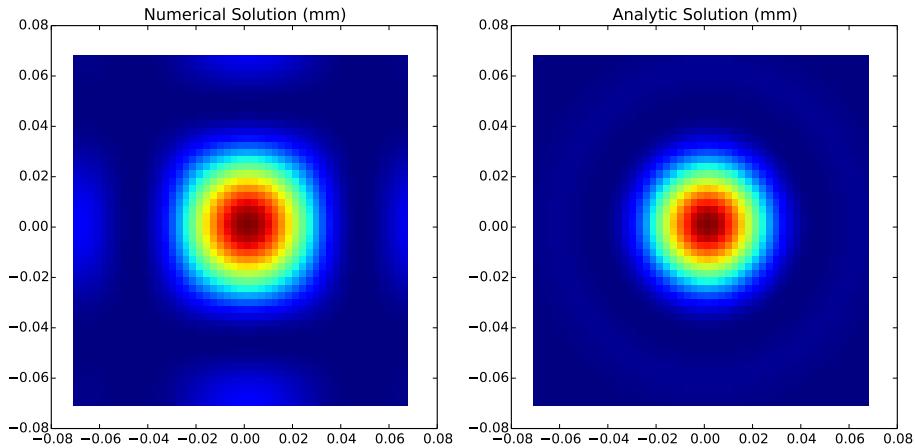


Figure 1: Numerical solution and analytical solution for the intensity distribution

3.2 Image Size and Sampling

The selection of the size of the imaging plane and sampling plane have a large effect on the numerical solution when found through the DFT. The more highly sampled, the more similar the numeric solution will be to the analytical solution, however the computational resources put a limit on the number of samples.

When moving to a sampled system, two things must be specified: (1) the size of the support structure (the total width over which the system will be sampled) and (2) the sample frequency or spacing. The general sampling scenario is shown in Figure 2.

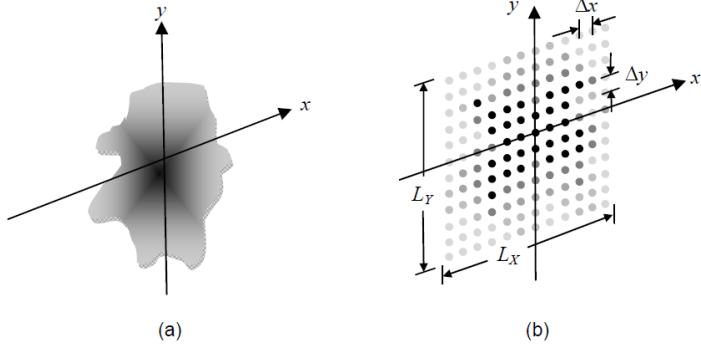


Figure 2: General two dimensional function: (a) analytic and (b) sampled [1].

The side lengths (L_x and L_y in Figure 2) should be large enough to contain the entire significant portion of the original function, without taking up more memory than required. Further the sample spacing (Δx and Δy) should be small enough to measure all of the features of the original function, however not so small to require the number of samples to be unreasonable.

The limit on the maximum size of the sample spacing for a bandlimited signal is defined by the Shannon-Nyquist sampling theorem [1], given in Equation 12 below.

$$\Delta x < \frac{1}{2B_x}, \Delta y < \frac{1}{2B_y} \quad (12)$$

If the Shannon-Nyquist criterion is not met, the signal is said to be undersampled. This leads to aliasing, where elements of the signal with high frequencies are mis-interpreted as low-frequency elements [1]. Figure 3 shows a signal (solid black line) that has been undersampled (grey 'x' markers) resulting in the appearance of a low frequency signal (dashed grey line).

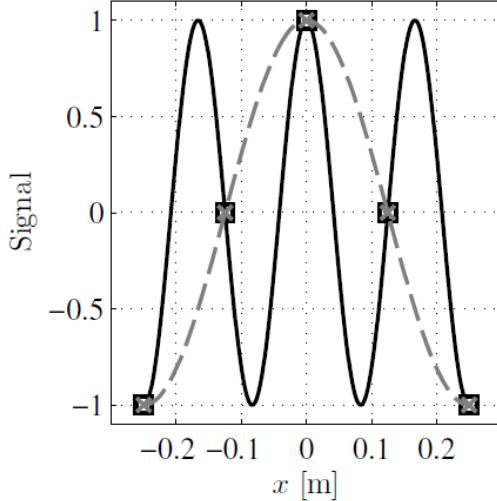


Figure 3: Aliasing of a high frequency signal through undersampling [2].

If the sampling rate is known, it is also possible to determine the maximum frequency that can be accurately represented by the samples. This is called the Nyquist frequency, and is given by Equation 13.

$$f_{NX} = \frac{1}{2\Delta x}, f_{NY} = \frac{1}{2\Delta y} \quad (13)$$

Since real signals are not bandlimited, effective bandwidths can be calculated such that the significant frequencies are preserved and the amount of ripple generated by the cut-off of frequencies is tolerable. According to [1], the effective bandwidth is found when the spectral power is 98% of the total power in the spectrum. This can be calculated using the method in [1] using Parseval's theorem. For example, the effective bandwidth for a circular function is given in Equation 14[1].

$$\text{forcirc} \left(\frac{\sqrt{x^2 + y^2}}{w} \right), B_{eff} = \frac{5}{w} \quad (14)$$

Combining Equation 14 with Equation 12 indicates that 10 samples across the radius of the circle is sufficient to encompass 98% of the spectral power.

In addition to aliasing, discretization results in "rippling and smearing in the spatial-frequency domain" and "virtual periodic replication in the spatial domain" [2]. The stages of the discrete fourier transform are shown graphically in Figure 4 from [2].

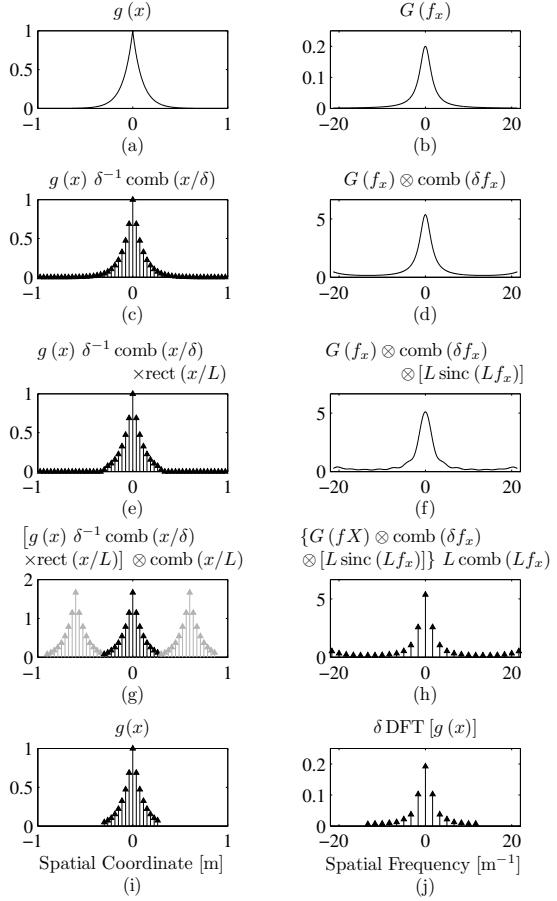


Figure 4: Steps of discrete fourier transform shown graphically [2].

Figure 4 shows pairs of images, where the left images are in the spatial domain ($g(x)$) and the right images are in the frequency domain ($G(f_x)$). The first two images show the original function and its continuous Fourier transform. The second two images account for sampling with a spacing δ . Image (d) shows the first artifact of the discretization process in the 'lifting' of the ends of the function. This is due to the periodic replication in the frequency domain. The third set of images limits the functions to a finite image size (L), and in image (f) the rippling and smearing distortions are clear. The fourth set of images show the sampling of the function in the frequency domain, with the result of virtual periodic replication in the spatial domain shown in image (g). Finally, image (i) shows the actual input function (with its limited spatial coordinate) and image (j) shows the sampled, band-limited DFT output. More information on the generation of this series of images and the distortions demonstrated can be found in [2].

The discretization effects can be mitigated by proper sampling. To reduce rippling, the support width (L_x and L_y) should be increased [2]. Similarly, to reduce aliasing, the sample spacing should be decreased (ie. the number of samples should be increased). However, this returns to the original problem of limited computational resources. The two requirements must be balanced based on the needs of the system.

4 Imaging System Parameters

The design of the sensor includes selection of three important parameters: the aperture diameter D , the focal length z and the wavelength λ . To see how changing these parameters will influence the image quality we can look at the point spread functions (PSF), but, first, it may be useful to consider the Rayleigh Criterion.

4.1 Rayleigh Criterion

To see the effect of the imaging system on the Fraunhofer diffraction patterns discussed up to this point, the system resolution can be determined. The Rayleigh Criterion assesses the quality of an optical system in terms of its resolution - how close two points can be while still being able to be identified separately - for the limiting case of a circular aperture. It is given in Equation 15.

$$\Delta\ell = 1.220 \frac{f\lambda}{D}. \quad (15)$$

In this equation, $\Delta\ell$ is the spatial resolution of the system, f is the focal length, λ is the wavelength of the light and D is the diameter of the aperture. The Rayleigh criterion is derived directly from the expression of the PSF. Rayleigh proposed that two sources could be distinguished from one another (ie. resolved) if their separation was sufficient that the position of their maximum intensities were at least $\Delta\ell$ apart [3]. This distance is the spacing between two peaks when the intersection of two PSFs is at half of the maximum intensity of each of the sources. This is illustrated in Figure 5 where as two sources move closer together they shift from being easy to identify separately to appearing as a single source.

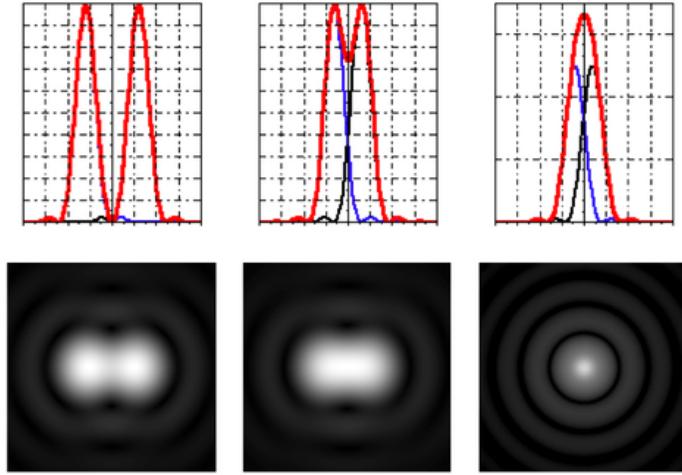


Figure 5: Rayleigh Criterion: (a) PSF for two easily resolved sources, (b) PSF for two just resolvable sources, (c) PSF for two unresolvable sources [4].

To reduce the resolution the system is limited to, a thin PSF is desirable. The equation for Rayleigh's criterion shows that this can be varied by selecting the aperture diameter (typically the lens diameter), the focal length, and the wavelength of the light.

4.2 Varying System Parameters

When the intensity distribution (PSF) is wide with a weak central peak, the energy is very spread out across the detector and the resolution of the system will suffer. In the last section on the lenslet array, it will be discussed that when a wavefront is incident on a lens at a certain angle, it leads to an image that is offset from the center of the detector by a certain amount. If we consider two lenses that are next to each other, and imagine that the wavefront incident on each lens is at an angle such that the two images approach one another on the detector plane, the spatial resolution will limit how steep the slope of the wavefronts can be, such that their images can still be resolved. This will put a practical limit on our sensor model and thus the three main parameters (diameter, focal length and wavelength) should be chosen with this in mind.

Increasing the diameter of the aperture, decreasing the focal length or decreasing the wavelength, will all decrease the diffraction of the system - and thus improve the system's resolution. The following PSFs in Figure 6 show the results of these changes while holding all other parameters constant. In these figures, an intensity distribution with a peak that has a small diameter and is sharp (tall), is preferred as it indicates better resolution than a peak that is short and has a large diameter.

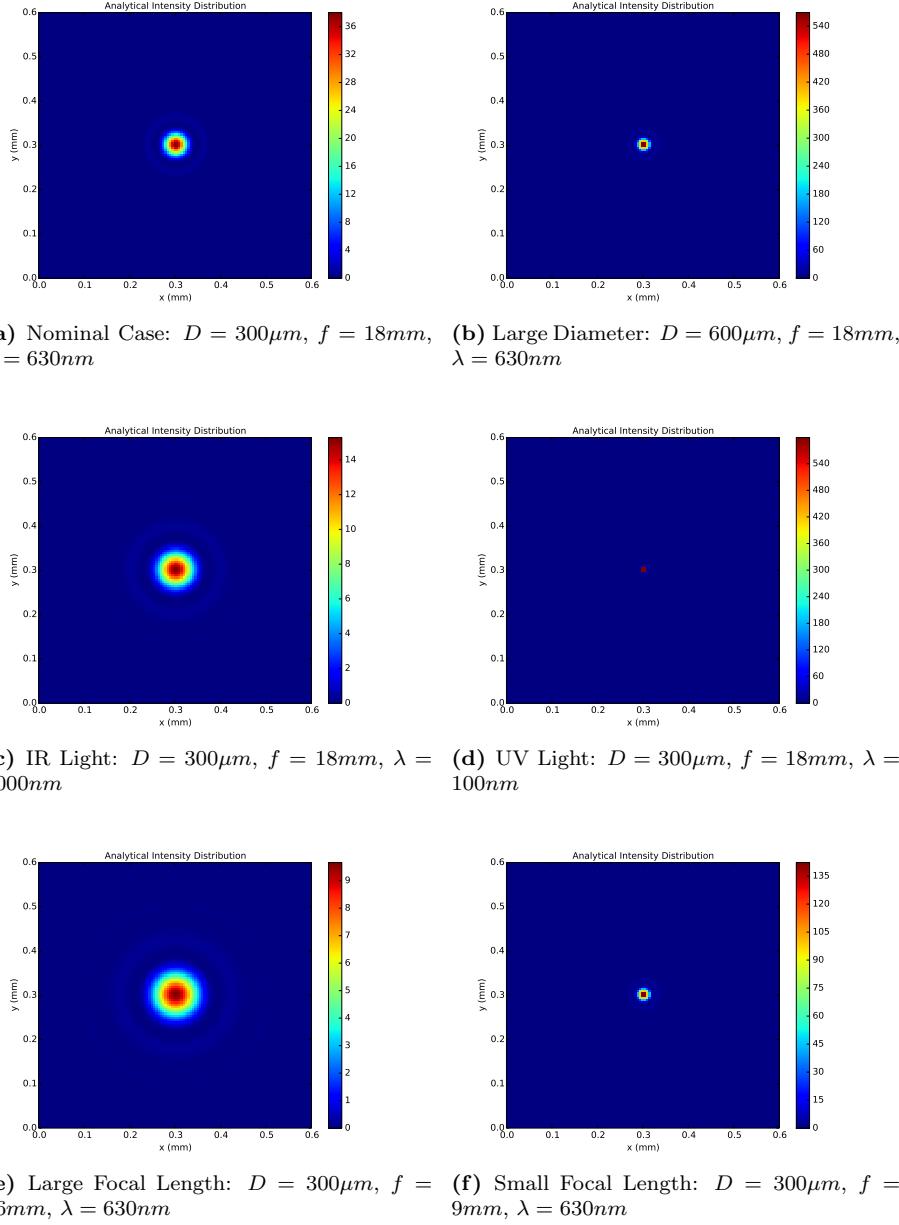
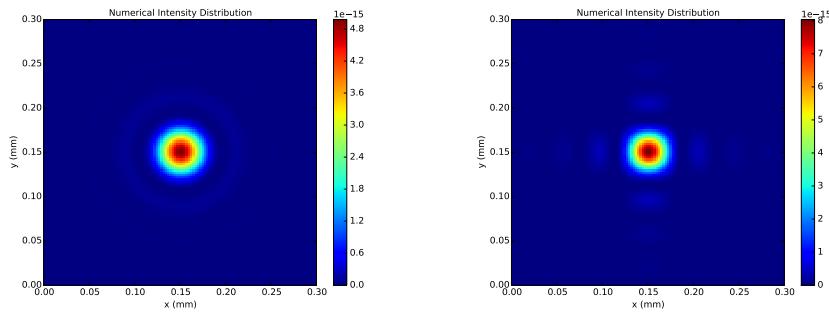


Figure 6: PSF for various values of aperture diameter, wavelength and focal length.

5 Various Pupil Functions

The selection of an arbitrary pupil shape and size requires the variation of only the Pupil Function given in Section 2 Equation 5. It can be specified discretely, however the discrete sampling should be at least as small as the wavefront sampling. The result for a square pupil and a round pupil are shown below in Figure 7.



(a) Circular pupil function of diameter D : $D = 300\mu\text{m}$, $f = 18\text{mm}$, $\lambda = 630\text{nm}$ (b) Square pupil function of width D : $D = 300\mu\text{m}$, $f = 18\text{mm}$, $\lambda = 630\text{nm}$

Figure 7: Intensity Distribution for different pupil shapes.

The intensity distribution for the square pupil shows a similarity to the diffraction pattern for a slit along the x and y axes. The central peak is a rounded-square with fringes out along the x, y axes and in the diagonals.

6 Sensor Noise

The wavefront sensor will experience several noise sources that limit the accuracy of the centroid algorithm and the control loop. Most notably are the discrete sampling of the intensity distribution, the incident photon noise and the detector readout noise. The latter two are discussed here.

6.1 Photon Noise

Photon noise, also called Shot noise, is due to the stochastic characteristics of light [5]. Photons arrive at the sensor as an average number of events per pixel, with a certain fluctuation in number of arriving pixels. This noise will vary with the exposure time and the strength of the source. It is modelled by a poisson distribution with the signal's variance equal to its intensity. If there are more than 1000 photon events, this can be modelled with a normal distribution instead, however numerically it is best to use the poisson distribution [6].

6.2 Readout Noise

Readout noise is attributed to the CCD or CMOS sensor that is used to convert the photons into an electrical signal. These devices are not perfect and the conversion and amplification process results in readout noise. It includes the dark current shot noise, dark offset fixed pattern noise, sense node reset noise, source follower noise, and ADC quantising noise. Although these can be found separately, they are typically modelled as the single readout noise as validation of the noise model for a sensor is done by the readout noise. More information on each of these separate noise sources can be found in [5].

Since this noise is from the electronics, it is present no matter the exposure time, signal intensity etc. The readout noise is modelled as white Gaussian, with a zero mean and normal distribution with standard deviation σ_r , and where the noise of each pixel is mutually uncorrelated [6].

Sensor specifications for readout noise are given as a number of electrons generated by the sensor due to the readout noise, however the standard deviation is given as the number of noise electrons relative to the mean signal brightness. Typically the number of electrons due to readout noise is between 0 and 5 electrons, and this is normalized with a mean signal brightness of 1000 electrons [7]. Thus the standard deviation for the readout noise, σ_r is nominally 0.005.

6.3 Measurement Noise

The combination of photon and readout noise are independent, but both result in errors for the centroid-ing algorithm. Of interest is the total noise at the output (Equation 16 from [8]).

$$\sigma_0^2 = g^2 \sigma_p^2 + g^2 \sigma_r^2 \quad (16)$$

Where σ_0^2 is the total noise output variance, σ_p^2 is the photon noise variance and σ_r^2 is the readout noise variance. The photon gain of the CCD detector, g , relates the number of electrons generated by a photon at the detector. Since it follows a Poisson distribution, the variance of the photon noise is equal to the mean number of photon events (intensity), p_i . Using the photon gain it can be expressed in terms of the number of output events since $p_i = gp_o$ [8]. This gives the relationship in Equation 17.

$$\sigma_0^2 = gp_o + g^2 \sigma_r^2 \quad (17)$$

This σ_0 is interesting for determining the error on the centroid location, however, numerically, it is preferable to calculate the photon and readout noises separately then add each of them to the intensity distribution [8]. In implementing the noise, the intensity distribution function should be normalized before adding the noise. This will remove the need to consider the detector specific characteristics such as the photon gain when adding noise to the output. Figure 8 shows the effect of measurement noise for a lenslet array. Brighter pixels (more red) have higher intensities. Since ideally the background would have

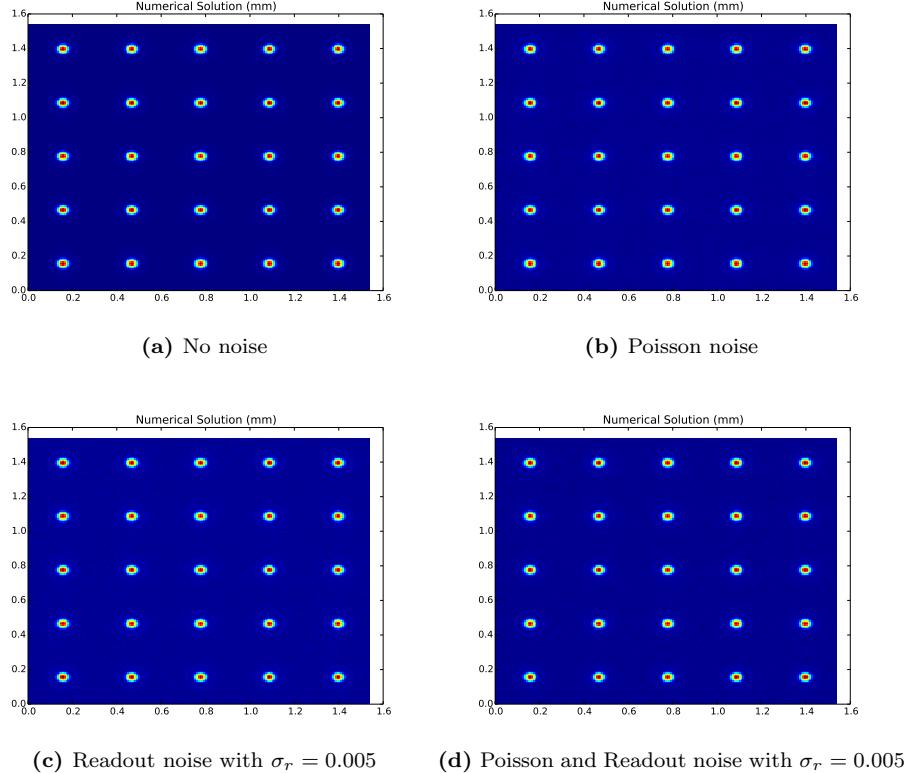


Figure 8: Effect of Measurement Noise

zero intensity (dark blue), a brighter background indicates more noise. It is clear that for the parameters modelled here, readout noise has a larger effect on the measurement noise than the poisson noise.

7 The Intensity Distribution of a Lenslet Array

Every intensity distribution of a lenslet is calculated separately by the procedure discussed above in Sections 2 and 3. If the incident rays enter on a to big angle, the FFT will be incorrect. To solve this, the average tilt of the incident phase on a lenslet is calculated. From this tilt a shift of the intensity distribution can be calculated. Putting this into a loop we calculate every intensity distribution of every lenslet with the correct coordinates. Then the distributions are interpolated on a grid that represent the Charge-Couple Device (CCD) giving the complete intensity distribution.

The precise procedure for the calculation of the normalized intensity distribution of a lenslet array goes as followed:

1. A spatial grid (x_t, y_t) with the dimension $l_x \times l_y$ of the lenslet array is created to contain the intensity distribution in the focal plane. This grid has the same sample spacings Δx and Δy as the incident wavefront on the lenslet array.
2. A spatial support grid (ξ_{sup}, η_{sup}) with the dimension $L \times L$ is created, where L can be adjusted to get a high enough sampled diffraction spot per lenslet. The sample spacing is taken as Δx because it is a square grid.
3. A spatial grid (x_{sup}, y_{sup}) with the dimension $\frac{\lambda f}{\Delta x} \times \frac{\lambda f}{\Delta x}$ is created to contain the FFT needed for the calculation of the diffraction spots. The sample spacing can be obtained by $\Delta x_{sup} = \lambda f / L$.
4. Now for every lenslet the corresponding phase plate $\phi_n(\xi, \eta)$, for $n = 1, 2, \dots, N_{lens}$ where N_{lens} is the number of lenslets, is extracted from the incident phase $\phi(\xi, \eta)$ over the whole lenslet array.
5. To avoid problems with aliasing we want to capture the main dynamics of the diffraction pattern in the middle of the grid (x_{sup}, y_{sup}) . This can be done by removing the average phase tilt $\overline{\nabla \phi_n}(\xi, \eta)$ of the phase $\phi_n(\xi, \eta)$, thus the processed phase plate is

$$\hat{\phi}_n(\xi, \eta) = \phi_n(\xi, \eta) - \overline{\nabla \phi_n}(\xi, \eta). \quad (18)$$

6. Still the shift in the focal plane that is caused by the tilt of the phase has to be applied to the spatial grid (x_{sup}, y_{sup}) . To do this the spatial shifts σ_x and σ_y needs to be computed from the average phase tilt $\overline{\nabla \phi_i}(\xi, \eta)$. This can be done by

$$\sigma_x = \frac{\overline{\nabla \phi_n}(\xi, 0)}{k} f, \quad \text{and} \quad \sigma_y = \frac{\overline{\nabla \phi_n}(0, \eta)}{k} f. \quad (19)$$

This relation can be confirmed by the calculation of Equation (6) with oblique incident wavefronts, which applies as long as the incident angles are small.

7. Now calculate the complex amplitude of each lenslet in the focal plane. This is done by taking the FFT of $P(\xi, \eta)U_i(\xi, \eta)$, where in this case

$$U_i(\xi, \eta) = e^{j\hat{\phi}_n(\xi, \eta)} \quad (20)$$

and $P(\xi, \eta)$ is the pupil function. Then multiplying this by the multiplicative factor from Equation (6).

8. The complex amplitude $U_f(x_{sup}, y_{sup})$ is now known in its relative dimensions for every lenslet and the intensity distribution in the right absolute spatial dimensions per lenslet is now calculated by

$$I_f(x, y) = |U_f(x_{sup} - \sigma_x - l_{centre,x}, y_{sup} - \sigma_y - l_{centre,y})|^2, \quad (21)$$

where $l_{centre,x}$ and $l_{centre,y}$ is the center position of the lenslet on the x and y-axis respectively.

9. By interpolation the intensity distributions of every lenslet is collected into the spatial grid (x_t, y_t) and normalized as

$$\tilde{I}_t(x_t, y_t) = \frac{I_t(x_t, y_t)}{\|I_t(x_t, y_t)\|_\infty} \quad (22)$$

References

- [1] D. Voelz, *Computational Fourier Optics*. SPIE Press, 2011.
- [2] J.D. Schmidt, *Numerical Simulation of Optical Wave Propagation with examples in Matlab (SPIE Press Monograph Vol. PM199)*. SPIE Press, Aug. 2010.
- [3] M. Verhaegen, "Lecture notes on control for High Resolution Imaging," May 2012.
- [4] Pinhole Works. (2014, May 12). *Goerge Airy vs. Lord Rayleigh* [Online]. Available: <http://pinholeworks.com/wp/goerge-airy-vs-lord-rayleigh/>
- [5] M.V. Konnik and J.S. Welsh, "On numerical simulation of high-speed CCD/CMOS-based wavefront sensors for adaptive optics," *Proc. SPIE 8149*, September 2011.
- [6] C.S. Smith et al., "Iterative linear focal-plane wavefront correction," *J. Opt. Soc. Am. A*, vol. 30, no. 10, pp. 2002-2011, October 2013.
- [7] L. Gilles and B. Ellerbroek, "Shack-Hartmann wavefront sensing with elongated sodium laser beacons: centroiding versus matched filtering," *Applied Optics*, vol. 45, no. 25, pp. 6568-6576, September 2006.
- [8] G. Cao and X. Yu, "Accuracy analysis of a Hartmann-Shack wavefront sensor operated with a faint object," *Optical Engineering*, vol. 33(7), pp. 2331-2335, July 1994.