Optimized centroid computing in a Shack-Hartmann sensor

Sandrine Thomas

Cerro Tololo Inter-American Observatories, Casilla 603, La Serena, Chile

ABSTRACT

The wavefront-sensor is one of the most important components of any adaptive optics (AO) system. The simplicity of the Shack-Hartmann sensor has made it a popular choice for such systems. Its accuracy, which largely determines its performance depends on having a good and robust centroid algorithm. Despite a large number of studies, the general recipe for selecting the best centroiding algorithm and best pixel size in a Shack-Hartmann wavefront sensor is still lacking. We combine analytical theory with numerical simulations to compare various flavors of centroiding algorithms (thresholding, windowing, correlation, quad-cell) under different conditions of photon flux, read-out noise, and sampling. It is shown that the choice of the best method depends on those parameters. At very low signal to noise ratio, the performance of the quad-cell is close to optimum.

Keywords: Shack-Hartmann, Centroid algorithms, adaptive optics

1. INTRODUCTION

A Shack-Hartmann Wavefront Sensors (SHWFS) samples the incident wavefront by means of a lenslet array which produces an array of spots on a detector. The wavefront is then analyzed by measuring, in real time, the displacement of the centroids of those spots. In this paper we compare different methods of centroid measurements and select the best method for the case of low photon flux, typical of astronomical Adaptive Optics (AO) systems.

Optimization of SHWFS centroid computation has been given considerable attention in the literature. Different types of algorithms have been developed to improve the basic centroid calculation, e.g. Mean-Square-Error estimator,[4] Maximum A Posteriori estimator[5] or Gram-Charlier matched filter.[6] Arines and Ares[7] focused themselves on the thresholding method. On other hand, many parameters are involved in the estimation of the error of the centroid calculation as explained below, and therefore there are many ways to approach this problem. Irwan & Lane,[8] for example, just considered the size of the CCD and the related truncation problem. They concentrated on photon noise only, neglecting readout noise or other parameters. In our paper, the chosen CCD size in pixels is large enough to permit us to neglect the truncation and we focus on other causes of error. Every method is subjected to the Cramér-Rao (CR) bound – a lower boundary for the error of the centroid determination of a spot on a CCD. Winick[3] studied analytically this powerful tool. No method can give better results than this limit. Therefore, we computed and used it in our study as a lower boundary to compare with our simulations. Moreover, it also indicates whether the best method has been found or not.

Yet, most previous studies remained theoretical or compared performance to simple centroiding. There is still no clear recipe for designing an optimized SHWFS, and more importantly, there is no clear recipe at the limit of detection. Many modern astronomical AO systems use quad-cell centroiding.[11] Despite its associated non-linearity and slope problems, it appears to be a good method at very low light level.

Centroid measurements are corrupted by the coarse sampling of the CCD, photon noise from the guide star, readout noise (RON) of the CCD, and speckle noise introduced by the atmosphere. The first three factors are illustrated in Fig. 1. In case of strong RON and weak signal, the spot is completely lost in the detector noise, at least occasionally. This presents a problem to common centroid algorithms like thresholding that rely on the brightest pixel(s) to determine approximate center of the spot. When the spot is not *detected*, the centroid calculation with such methods is completely wrong. Thus, the lowest useful signal is determined by spot detection rather than by the centroiding noise *per se*.

Further author information: (Send correspondence to S.T.)
S.T.: E-mail: sthomas@ctio.noao.edu, Telephone: 56 51 205 325

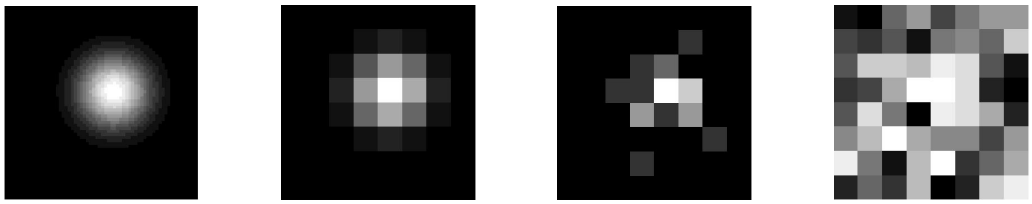

Figure 1. From the left to the right: Raw, well sampled spot image for one subaperture in a Shack-Hartmann wavefront sensor. **Sampling problem:** On the next frame is the same image obtained on the CCD, sampled with only 8x8 pixels. **Photon noise:** When the number of photons is very low, the detected spot is noisy. **Readout noise:** this noise has a normal distribution. The number of photons in the spot $N_{ph} = 30$ and readout noise $N_r = 2$. (This represents one of the worse shots for this set of parameter to emphasize the problem).

One can write the total centroid measuring error σ_{err} as:

$$\sigma_{err}^2 = \sigma_{ph}^2 + \sigma_r^2, \tag{1}$$

where σ_{ph}^2 is the photon noise and σ_r^2 is the readout noise.
In the case of pure photon noise and a Gaussian spot, $\sigma_{err} = \sigma/\sqrt{N_{ph}}$, where σ is the rms size of the spot and N_{ph} is the number of photons. As shown by many authors,[3,8] this indeed proves to be the lower boundary of the centroid error. Moreover, this boundary is reached by simple centroid, which is thus the optimum method for no readout noise. However, for a diffraction spot other methods are superior.[4]

In the presence of CCD readout noise and/or background, the detailed expression of σ_r^2 depends on the method of calculation of the centroid and have been derived by Rousset[9] and Sandler[10] for example. However, if the simple center of gravity formula is used, the number of pixels p used in the calculation should be reduced to a minimum. For quad-cell, $p = 4$.

In this paper, we study different popular variants of centroid algorithms such as quad-cell, thresholding and windowing. Recently, correlation was suggested as a powerful technique of centroid calculation.[1] We determine optimum parameters for each of these methods and compare their performance. We do not pretend to be exhaustive and to give the best method. We emphasize here the detectability problem, often left aside.

2. METHOD AND RESULTS

In this section we present the results of our simulation of different centroiding methods in a SHWFS.

2.1. Simulation method

Here we model the spot by a 2D Gaussian curve, deferring the study of atmospheric speckle noise to the future. We consider only one 8x8 subaperture and its corresponding region on the CCD. The full width at half maximum (FWHM) of the Gaussian spot is N_{samp} CCD pixels. We calculate the Gaussian on a fine 256x256 grid and integrate it within each detector pixel, to account for the sampling. The pixel signals are then replaced by the Poisson random variable, with the average of N_{ph} photons per spot (e.g. per subaperture). A zero-mean normal noise with rms N_r is added to simulate the RON. For each set of parameters, our Monte-Carlo simulation includes 500 random realizations. At each iteration, the true center is placed randomly within the CCD pixel and then compared to the "measured" center.

Since the detectability limit is our main concern, we focused on low level of signal relative to the noise. Some additional strategies are therefore needed to assure that the centroid calculation is meaningful. Indeed, at low signal-to-noise ratio (SNR), the brightest pixel occasionally happens to be outside the spot. In a real AO system the spot centroids do not change from one frame to the next by very large amounts, permitting one to catch such cases and to reject them. We simulate this additional criterion by requiring the maximum of the image to fall within a radius of (1.5 N_{samp})x(1.5 N_{samp}) CCD pixels from the true centroid and the value of the maximum to be at least twice the RON. Finally, the error variance of a single frame must be at least smaller than three

times the rms size of the spot to be significant. If those conditions are not fulfilled, the measure is rejected. The number of rejected cases gives us information on the detectability limit: when more than half of images are rejected, we consider that the centroid measurements has failed. Otherwise, the rms centroid error σ_{err} is computed on the retained images.

The error σ_{err} depends mainly on the *number of photons*, the *readout noise* of the CCD, the *number of CCD pixels* per subaperture and the pixel size which is related to the *sampling of the spot*. There are additional parameters for each specific method such as the threshold, T for the thresholding and the size of the window, R_w for the windowing. Table 1 summarizes the parameters with the range of values explored here. For the RON of $N_r = 0, 3, 5$ electrons we consider only low number of photons because we focus on the centroiding limits at low SNR. Moreover, with the values presented in Table 1, we cover a representative range of SNR. The strategy is to find the best set of parameters that minimizes the centroiding error σ_{err}.

We compare σ_{err} with the photon noise limit $\sigma/\sqrt{N_{ph}}$ as well as the CR lower bound. We will consider both limits for completeness and coherence. We also compare σ_{err} with the Gaussian rms width σ. When $\sigma_{err}/\sigma > 1$, the spot "effective" SNR is less than 1 and the spot detection is problematic. Thus, $\sigma_{err} = \sigma$ is the reasonable upper limit for the centroid error.

Table 1. Parameters of the SHWFS centroiding

Parameter	Range	Units
Number of photons per spot N_{ph}	[10;150]	photons
Readout noise N_r	[0;5]	electrons
Spot FWHM N_{samp}	[0.2;3]	CCD pixels
Sub-aperture size N	8x8	CCD pixels
Threshold T	[0.1; 0.2]	fraction of the maximum
Window radius R_w	R_w = [1;3]σ	CCD pixels

Throughout the paper, we give the normalized variance σ_{err}/σ as a result of the simulation.

2.2. Centroiding improved by thresholding

In this method, the selection of the significant pixels is made considering their intensity relative to the maximum. The pixel with the maximum intensity I_{max} is first determined, then the threshold C is set at $T \cdot I_{max}$, but no less than $3N_r$. For low SNR, using the $3N_r$ lower bound for the threshold removes most of the noisy pixels without removing the signal. $3N_r$ is optimum; $2N_r$ still admits noise while $4N_r$ cuts a part of the signal. Improvements of the results' reliability using the readout noise as a reference for the threshold, occur mainly when the SNR is low – either low number of photons or large N_{samp}. Indeed, in that case, the maximum value is relatively small. Using a simple percentage of the maximum will therefore increase the number of pixels above threshold resulting from noise compared to the number of pixels belonging to the spot, leading to false estimations.
C is then subtracted from the detected image over the whole subaperture. The pixels with negative values after subtraction are set to zero. Finally, the simple center of gravity is computed for this thresholded image.

Fig. 2 shows the best value for the parameter T for $N_r = 3$ as well as the best sampling. The best N_{samp} is 1.2 whatever the readout noise and the number of photons. $T = 0.2$ gives slightly the best performances, only noticeable for high SNR. Indeed, as T increases, the curves corresponding to the highest number of photons have their best sampling shifted toward the higher values – here, for $T = 0.3$, the best sampling is 1.4 and we found that for $T = 0.5$, it becomes almost 2.

The sharp raise for $N_{samp} < 1$ is due to the sampling. The light is concentrated on a small number of pixels increasing the signal to noise ratio, but also reducing the number of pixels for the center of gravity estimation. Again, for low SNR, the threshold will be defined by $3N_r$ rather than by $T \cdot I_{max}$ to avoid as much as possible the contribution of the noise in the centroid calculation and to reach the detectability limit possible for this method.

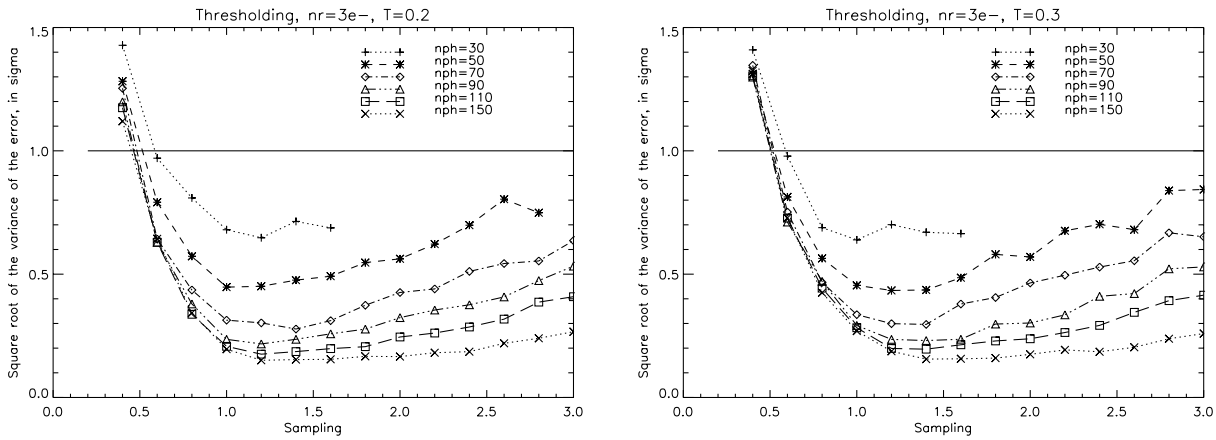

Figure 2. Relative centroiding error σ_{err}/σ as a function of the spot sampling N_{samp} for the thresholding. Test of the best threshold: $T = 0.2$ on the left and $T = 0.3$ on the right. The readout noise is $N_r = 3$. Similar results are obtained with $N_r = 5$.

2.3. Centroid improved by Windowing

This method consists in using a circular window defined by $W(r) = 1$ for $r < R_w$ and $W(r) = 0$ otherwise and centered on the pixel with the maximum intensity in order to remove the noisy pixels at the edge. In the program, the parameter R_w is the radius of the window in CCD pixels. Fig. 3 shows example of windows.

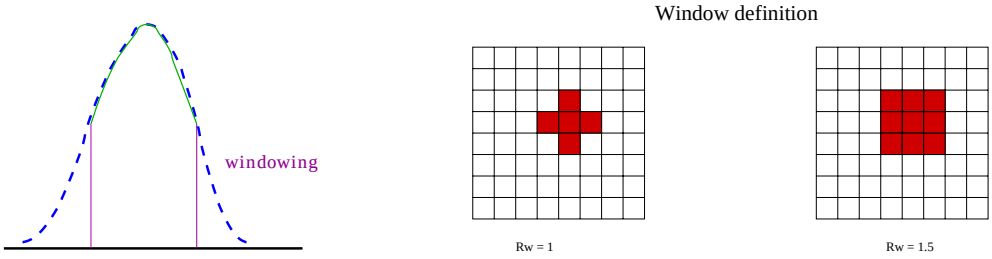

Figure 3. Windowing method. All pixels outside a circular window of diameter $2R_w$ are set to zero.

R_w is either proportional to the size of the spot – $R_w = P_w * N_{samp}$ – or set to 1 if the resulting R_w is smaller than one pixel. Figure 4 shows the simulation results for $P_w = [1; 1.5; 2]$. The best value of the radius parameter is $P_w = 1.5$ for median to high SNR and $P_w = 1$ for very low SNR. Considering the error bars for $N_{ph} = 30$ (cf. Fig 9), we keep $P_w = 1.5$ as a better window in the following. The best sampling being $N_{samp} = 1.2$, the resulting window is a 3x3 box. This result does not depend on the signal to noise ratio.

A model of error for windowing can be derived. As presented in Section 1, the total error variance of the centroid calculation using a CCD is $\sigma_{err}^2 = \sigma^2/N_{ph} + \sigma_r^2$. For the windowing method, σ_r^2, computed from the center of gravity formula, is expressed by:

$$\sigma_r^2 = \frac{\kappa N_r^2}{N_{ph}^2}, \quad \kappa = \sum_{i,j} x_{i,j}^2, \tag{2}$$

where $x_{i,j}$ is x-coordinate of the pixel $\{i, j\}$ in the window. In our simulation, the best window is found to be a 3x3 box ($R_W = 1.5$), leading to $\kappa = 6$. Therefore, the model for the windowing is:

$$\frac{\sigma_{err}}{\sigma} = \frac{\sqrt{N_{ph} + 6N_r^2/\sigma^2}}{N_{ph}}. \tag{3}$$

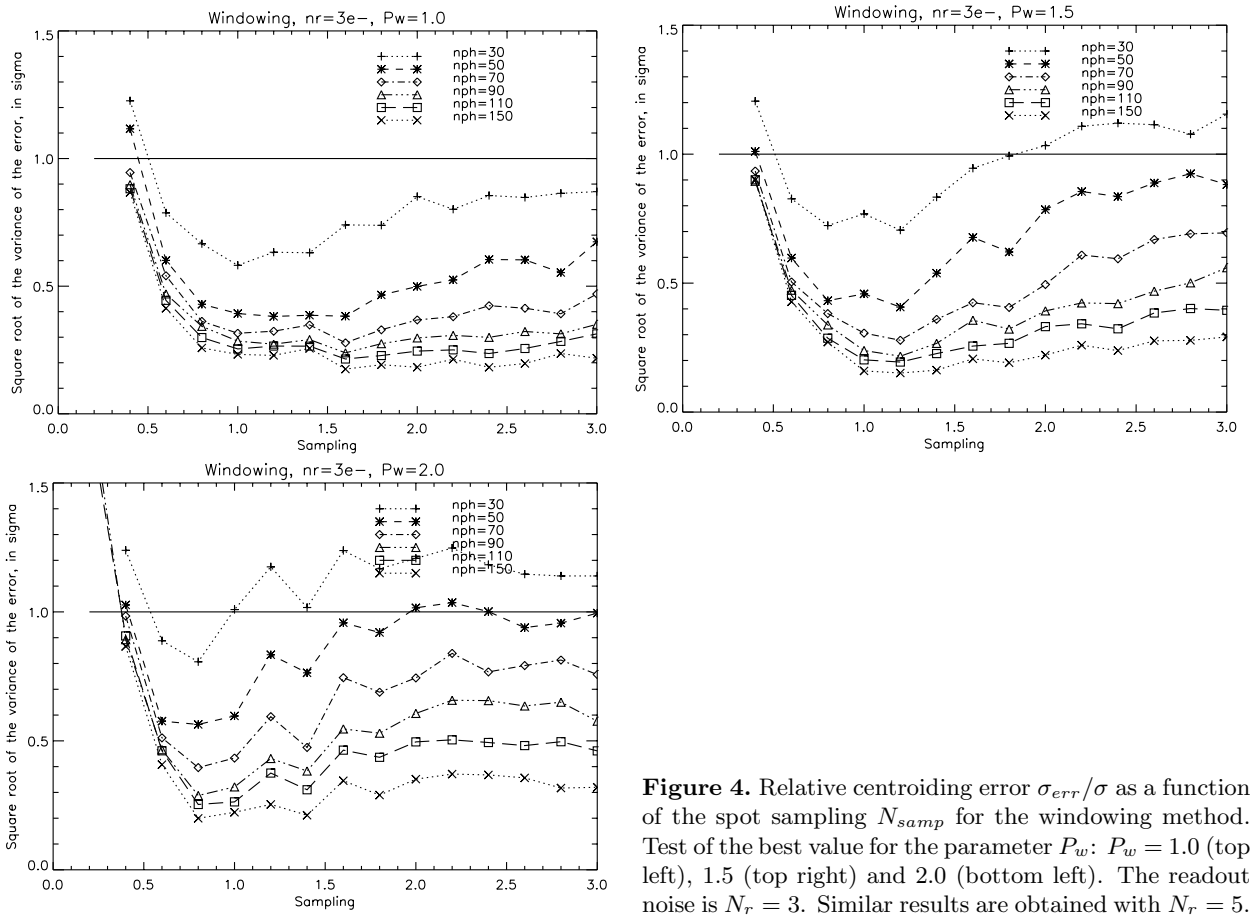

Figure 4. Relative centroiding error σ_{err}/σ as a function of the spot sampling N_{samp} for the windowing method. Test of the best value for the parameter P_w: $P_w = 1.0$ (top left), 1.5 (top right) and 2.0 (bottom left). The readout noise is $N_r = 3$. Similar results are obtained with $N_r = 5$.

We will use this model in Section 3, to evaluate the results of our simulations.

2.4. Correlation

This method computes the correlation between the image obtained on the CCD and a reference spot (template).[1] The template is a 2D Gaussian function with rms size σ. Note that we compute the template by evaluating the Gaussian at the center of each CCD pixel, and thus do not take into account the averaging within pixels. The correlation is calculated using Fourier transform, then the position of the maximum of the resulting image in "x" and "y" direction is determined, either by a simple centroid calculation (Tarik Noel[2]) or by fitting a parabola to 3 points around the maximum (Lisa Poyneer[1]). We used the second method in our study. Since the correlation is not based on the centroid calculation, it appears to be very good at suppressing the noise from pixels outside the spot, as shown in Fig. 5 for $N_{ph} = 50$, $N_r = 3$ and $N_{samp} = 1$. Moreover, correlation is known to be the best method of signal detection ("matched filtering").

In Fig. 6 the centroid measured as a function the true centroid is plotted. In a perfect world, the relation would be linear. With increasing noise, the scatter of points becomes wider, until complete confusion is reached. As explained in Section 2, a measurement is not taken into account if the error is higher than 3σ and if the number of rejected cases is more than half the number of iterations. The rejected measures are not plotted on the final curves. The detectability limit in this case is about 15-20 photons for a readout noise of 3 electrons. Indeed, the left graph of Fig. 6 shows the expected linear dependence of the measured centroid vs the true centroid, still for $N_{ph} = 15$ in spite of some scatter.

Our results are quite consistent with those of Lisa Poyneer.[1] In figure 7 our results for the correlation

Figure 5. Illustration of the correlation method. On the left is a raw image obtained on the CCD and on the right is its correlation with the template. The parameters are: $N_{ph} = 50$, $Nr = 3$, and $N_{samp} = 1$.

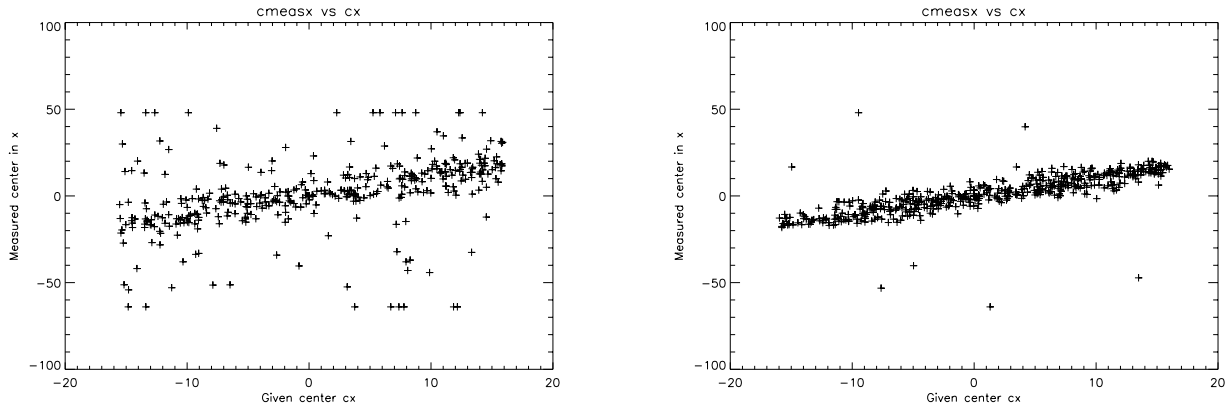

Figure 6. Example of detection limit in the case of the correlation method: measured centroid vs true centroid. c_x is in "fine" pixels. The readout noise is 3e- and the number of photons is 15 on the left and 30 on the right. Note that the condition on the error being less than 3σ is not applied here to highlight the linear trend in spite of the scatter. For $N_{ph} < 15$, the scatter prevails and no linear trend is visible.

method are shown along with L. Poyneer's results for $N_{ph} = 100$ and $N_r = 3$ – left graph – and $N_r = 5$ – right graph. Lisa uses a seeing-limited point source instead of a Gaussian source.

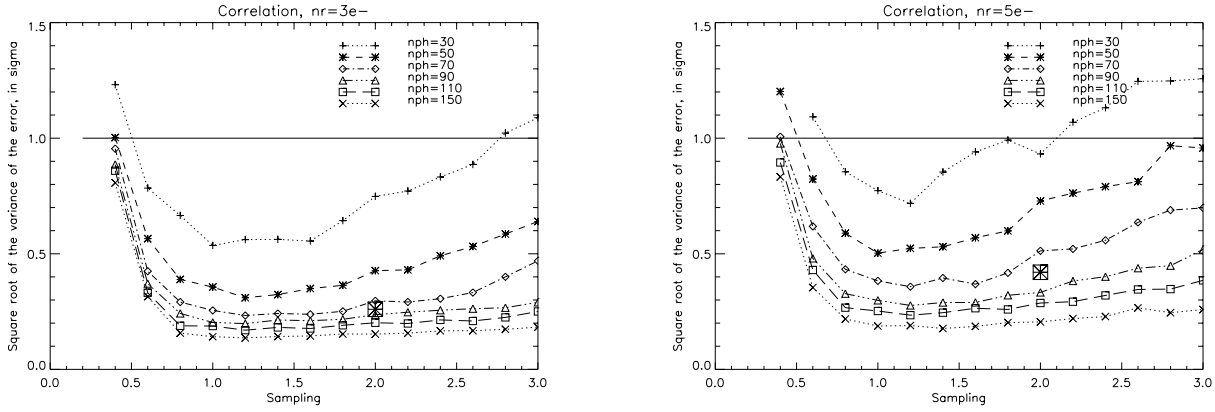

Figure 7. Relative centroiding error σ_{err}/σ as a function of the spot sampling N_{samp} for the correlation method. L. Poyneer's results in the case of 100 photons, 3 or 5 electrons are represented by a cross on both graphs: $\sigma_{err}/\sigma = 0.26$ for $3e^-$ and 0.42 for $5e^-$.

Moreover, she pointed out the poor reliability of measurements for $N_{ph} < 30$, even at $N_r = 0$. In our simulations, we find that reliable measurements are possible for $N_{ph} < 30$ even when $N_r = 3$. The differences of our results may be due to difference in simulated spot shapes (real spot after propagation in the atmosphere vs

Gaussian spot).
Note that the computation cost of the correlation method does not represent a serious limitation: nowadays there are computers powerful enough to run this type of algorithm at reasonable speed for AO systems. Indeed, correlation centroiding is routinely used in the solar AO.[12]

2.5. The quad cell

A quad-cell is the specific case of a centroid for a 2x2 array. It is widely used in astronomical AO systems, e.g. Altair,[11] because with $p = 4$ the weak signal from guide stars is better detected against RON and because with a small number of pixels the RON can be further reduced by a slow CCD readout. A quad-cell algorithm calculates the centroid in each direction from the differential intensity between one half of the detector and the other. Depending on the size and the shape of the spot, non-linearity and saturation will degrade the accuracy of the measurements even for high SNR. The main concern would then be the size of the spot, which has to be big enough to avoid the saturation and small enough to stay within the subaperture and give good accuracy.

As for the windowing method, a model can be derived to compare with the results of the simulation. The centroid in x direction is defined as:

$$C_x = A \times \frac{I_l - I_r}{I_l + I_r}, \tag{4}$$

where I_l and I_r are the intensity respectively on the left and the right halves of the detector and A is a constant. A simple derivation shows that $A = \sigma\sqrt{\pi/2}$ for a Gaussian spot. The total variance of the error is thus expressed by:

$$\frac{\sigma_{err}}{\sigma} = \sqrt{\frac{\pi}{2}} \left(\frac{\sqrt{N_{ph} + 4N_r^2}}{N_{ph}} \right). \tag{5}$$

Equation 5 is plotted in the next section along with the quad-cell simulation.

3. COMPARISON BETWEEN METHODS

After studying centroiding methods independently to determine the best set of parameters for each (Table 2), the methods are inter-compared.

Method	Sampling N_{samp} (CCD pixels)	Other parameter (T or P_w)
Thresholding	1.2	0.2
Windowing	1.2	1-1.5
Correlation	1.0	*
Quad Cell	*	*

Table 2. Summary of the best sampling as a function of the readout noise and the number of photons. The number of photons is $N_{ph} = [10, 150]$, the readout noise is $3, 5$ electrons and the sampling is $N_{samp} = [0.2; 3]$.

Figure 8 shows the results of the comparison for two values of readout noise as well as the upper and lower boundaries described in Section 2.
The windowing and thresholding methods, ameliorated by the different conditions enumerated in Section 2, give similar results. The signal is detectable down to 25-30 photons for $N_r = 3$ and 35-40 photons for $N_r = 5$.
Using the correlation method leads to better accuracy in the centroid calculation at low signal to noise ratio (SNR). Moreover, the detectability is also improved since it achieves useful results down to 15-20 photons for $N_r = 3$ and 25-30 photons for $N_r = 5$. However, it does not present any benefit at high SNR.
It is more robust at low SNR than thresholding and windowing. Finally, the quad-cell give results close to the optimum for low SNR and even reaches the Cramér-Rao bound for higher SNR – cf. Figures 8 and 9.

Yet, there might be an advantage to use correlation instead of other methods if dealing with non-symmetric and/or resolved sources.[12] Indeed, in order to calculate the centroid in this case, one needs to sample it more

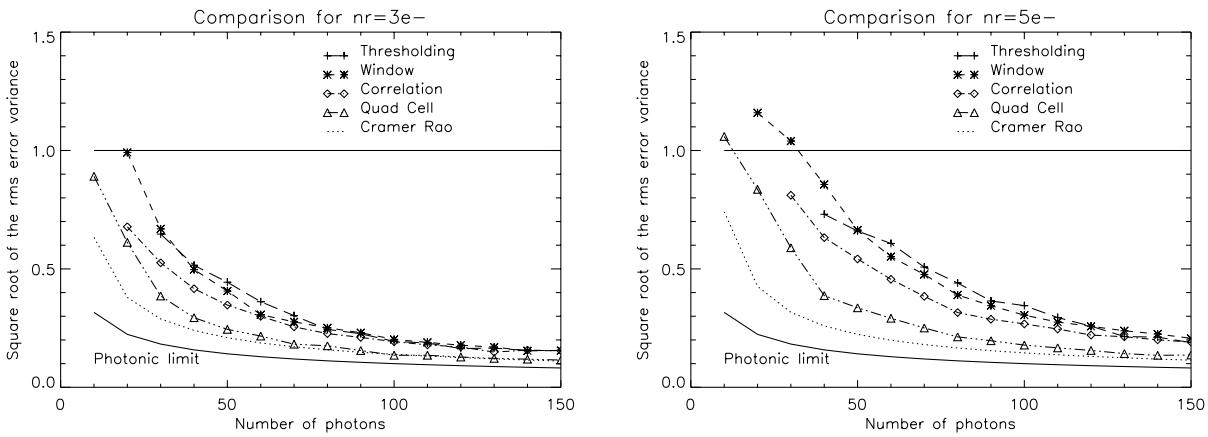

Figure 8. Comparison of the centroid methods using the best parameters for each method. The relative errors σ_{err}/σ is plotted as a function of N_{ph}. The RON is 3 on the left and 5 electrons on the right. The Cramér-Rao limit is represented by the dotted line and the photon noise by the full line.

finely than by a 2x2 array. Correlation would then be the best method.

The case of very high SNR is not shown here although we can predict some of the curves' behavior. For large number of photon, the Equation 5 becomes $\sigma_{err}/\sigma \approx 1.25\,(1/\sqrt{N_{ph}})$. This shows that the quad-cell error will always be higher than the photon noise limit by 25%, while the CR bound and the other methods will eventually reach it.

Figure 9 gives a comparison with the theory for windowing method and quad-cell. For the windowing, the theoretical and simulated curves match very well. For the quad-cell, the theoretical and simulated curves match for high SNR while a deviation occurs for low SNR. This partly comes from the big uncertainties obtained when the noise is high.

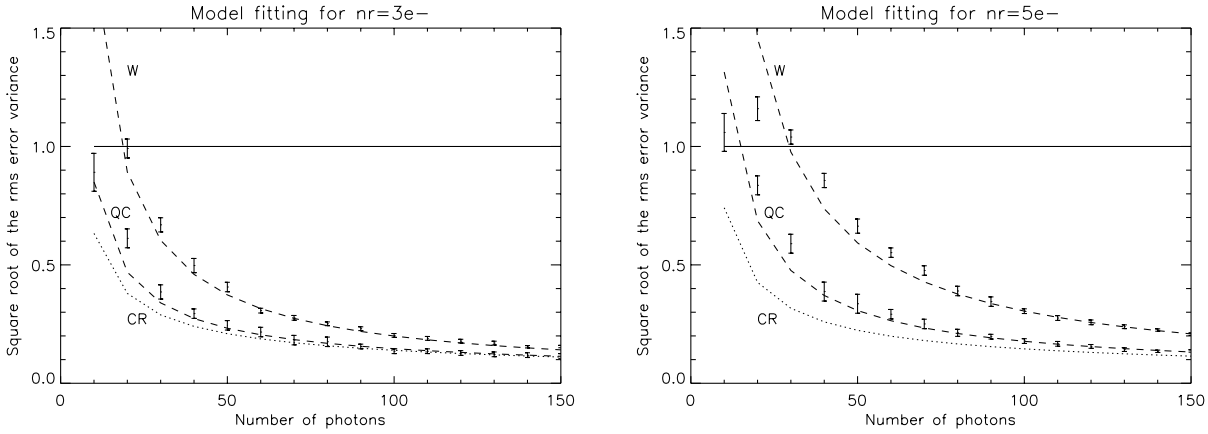

Figure 9. Results of simulation compared to the model for the windowing method and the quad-cell. The results are represented by dots with the error bars, the models in dashed lines. The dotted line represents the CR bound. The readout noise is 3 electrons on the left and 5 on the right. W=windowing, QC=quad-cell and CR=CR bound.

An important comment is that the quad-cell, in addition to its excellent accuracy for medium SNR, is also optimal in speed as far as computation is concerned.

4. CONCLUSION

Without any readout noise, the best method would be a simple centroid. As soon as the readout noise increases, the quad-cell method becomes the best one since it gives the lowest variance error. The correlation, as second "winner", still gives good results and good detectability. Then, the FWHM of the spot has to be about one CCD pixel. At very low SNR, though, the CR bound indicates that a better method of centroiding than the ones presented in this report might exist.
Finally, at high SNR, all method except the quad-cell are valid.

We proved in this paper that the limit of detectability of a Shack-Hartmann wavefront sensor can reach 10 photons per subaperture in presence of a $3e^-$ readout noise.

The next step is to check if these conclusions are still valid when we simulate spot images of a natural or laser guide star distorted by turbulent atmosphere (instead of Gaussian spot).

ACKNOWLEDGMENTS

I first want to aknowledge my supervisor Andrei Tokovinin for his patience and his long discussions with me on the subject. I also want to thank Rolando Cantarutti for his help in debugging some of my routines. Finally I thank Lisa Poyneer for her interesting discussion and Brooke Gregory for making useful comments on my manuscript.

REFERENCES

1. L.A. Poyneer, K. LaFortune, and A.A.S. Awwal, "Correlation wavefront sensing for Shack-Hartmann-based Adaptive Optics with a point source", Lawrence Livermore National Lab Document September 2003 (Livermore, CA 94551), 2003
2. N.Tarik, "Caractérisation spatiale et temporelle de la turbulence atmosphérique par analyse de front d'onde", Thesis ONERA, Paris, 1997
3. Kim A. Winick, "Cramer Rao lower bounds on the performance of Charge-Couple-Device optical position estimator", *JOSA*, **3**, pp. 1809-1815, 1986
4. M. A. van Dam and R. G. Lane, "Wave-front slope estimation", *JOSA*, **17**, pp. 1319-1324, 2000
5. S.A. Sallberg, B.M. Welsh, and M.C. Roggemann, "Maximum a posteriori estimation of wave-front slopes using a Shack-Hartmann wave-front sensor", *JOSA*, **14**, pp. 1347-1354, 1997
6. J.-M. Ruggiu, C.J. Solomon and G.Loos, "Gram-Charlier matched filter for Shack-Hartmann sensing at low loght levels", *Opt. Lett.*, **23**, pp. 235-237, 1998
7. J. Arines and J. Ares, "Minimum variance centroid thresholding", *Opt. Lett.*, **27**, pp. 497-499, 2002
8. R. Irwan and R. G. Lane, "Analysis of optimal centroid estimation applied to Shack-Hartmann sensing", *Appl. Opt.*, **38**, pp. 6737-6743, 1999
9. G. Rousset, "Wave-front sensors", in "Adaptive optics in Astronomy", ed. F. Roddier, Cambridge University press, p. 115, 1999
10. D. Sandler, "The design of laser beacon AO systems", in "Adaptive optics in Astronomy", ed. F. Roddier, Cambridge University press, p. 294, 1999
11. G. Herriot, S. Morris, A. Anthon/ et al., "Progress on Altair: the Gemini North adaptive optics system", *Proc. SPIE*, **4007**, pp. 115-125, 2000
12. Ch. U. Keller, C. Plymate, and S.M. Ammons, "Low-cost solar adaptive optics in the infrared", Proc. SPIE, **4853**, pp. 351-359, 2003