Hartmann Wavefront Analyzer Tutorial

Spiricon®
Inc.
Laser Beam Diagnostics

Spiricon, Inc.
2600 North Main St.
Logan, Utah, USA   84341

Phone: **435-753-3729**
Fax: **435-753-5231**
E-mail: **sales@spiricon.com**

Visit our web site for current product information at:
        http://www.spiricon.com

Contents

Introduction

The Hartmann sensor was invented a century ago to perform optical metrology.[1] Subsequently these sensors have been adapted to a wide variety of applications including adaptive optics, ophthalmology, and laser wavefront characterization. In this document we will explain the operation of the Spiricon Hartmann Wavefront Analyzer (HWA), analyze the device limitations, and compare it to other similar technologies.

Basic Wavefront Sensor Operation

The Hartmann wavefront sensor consists of an array of apertures mounted a distance from a charge-coupled device (CCD). Figure 1 shows a schematic of the operation of the wavefront sensor. Light travels to the sensor along the z-axis. Each of the apertures acts like an optical lever, displacing the diffracted spot proportional to the average phase tilt over the aperture. The wavefront sensor measures the tilt over each aperture by comparing the measured positions of the diffracted spots to the positions of the diffracted spots for a reference input beam. The tilt measurements are then converted into a replica of the wavefront by performing a form of integration called wavefront reconstruction.
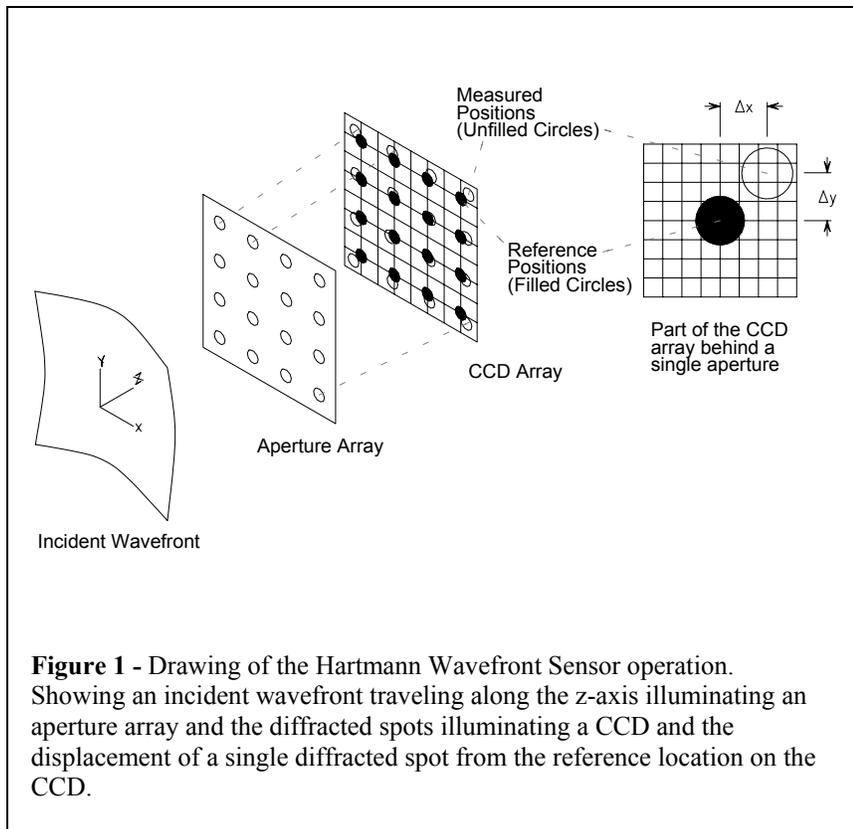

Figure 1 - Drawing of the Hartmann Wavefront Sensor operation. Showing an incident wavefront traveling along the z-axis illuminating an aperture array and the diffracted spots illuminating a CCD and the displacement of a single diffracted spot from the reference location on the CCD.

Diffracted Spot Location

The first step in measuring a wavefront with the Hartmann Wavefront Analyzer is to determine the location of the diffracted spots. We will use the coordinate system shown in Figure 1. The z-axis is the optical axis normal to the planes of the CCD and aperture arrays. The x and y axes are located in the plane of the CCD array. The sensor is calibrated by recording an image of a uniform plane wave. The diffracted spot locations are then determined from this image by calculating their centroids or first moments along both the x-axis and the y-axis. The centroid in the x direction, $\overline{x}$, of an arbitrary intensity pattern is given as,

$$\overline{x} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I(x,y) \cdot x \cdot dx \cdot dy}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I(x,y) \cdot dx \cdot dy}$$

Equation 1

where $I(x, y)$ is the intensity as a function of x and y, x is the location at which the intensity is measured along the x-axis, and y is the location at which the intensity is measured along the y-axis. This equation needs to be modified to apply it to Hartmann sensing. The CCD behind the Hartmann wavefront sensor is comprised of a set of discrete pixels that we describe by the variable i along the x-axis and j along the y-axis. This allows us to replace the integral with a summation. Furthermore, we have to change the limits of the integration to apply this formula to measuring an array of diffracted spots. We assume that the intensity pattern at the focus of a given aperture can be effectively bounded by the edges of the aperture. We typically use a rectangular grid of apertures when building Hartmann wavefront sensors to roughly match to the rectangular grid of photo-detectors in the CCD. To find the locations of all the diffracted spots, we first divide the image into an array of small rectangular sections of pixels that are centered on the apertures with sides as long as the aperture spacing. We call these rectangular sections of pixels integration areas because we limit the centroid integration for each aperture to these areas. Using index coordinates, we call the pixel corresponding to the lower-left corner of the integration area (i_{min}, j_{min}) and the pixel corresponding to the upper-right corner of the integration area (i_{max}, j_{max}). Therefore the general centroid equation can be rewritten as,

$$\bar{x} = \frac{\sum_{i=i_{min}}^{i_{max}} \sum_{j=j_{min}}^{j_{max}} I(i, j) \cdot i}{\sum_{i=i_{min}}^{i_{max}} \sum_{j=j_{min}}^{j_{max}} I(i, j)} \cdot s \qquad \textbf{Equation 2}$$

where $I(i, j)$ is the intensity measured by the pixel in the ith row and j^{th} column and s is the spacing of pixels along the x or y axes. The same equation can be used to determine the position of the diffracted spot along the y-axis. The intensity used in the centroid calculation, $I(i, j)$, is actually a threshold limited intensity value. If the intensity is above the threshold the intensity value minus the threshold is used in the calculation. If the intensity is below the threshold, the intensity at this pixel is set to zero. Threshold limiting the intensity helps to minimize the effects of noise due to the camera and due to light diffracted out of the primary diffracted order.

Wavefront Slopes

The gradient of the wavefront can be calculated based on the motion of the centroid of a diffracted spot. To perform any measurement, the Hartmann Wavefront Analyzer is illuminated with a beam whose wavefront, $\Phi(x, y)$ is being measured and the diffracted spot positions are determined from the CCD image using the centroid algorithm. For this discussion we limit ourselves to considering the x-axis, but the y-axis can be treated in the same fashion. The change in the diffracted spot position along the x-axis, Δx, is found by subtracting the measured diffracted spot positions from the reference spot positions. The measured two-dimensional wavefront is designated by the variable: $\Phi(x, y)$.
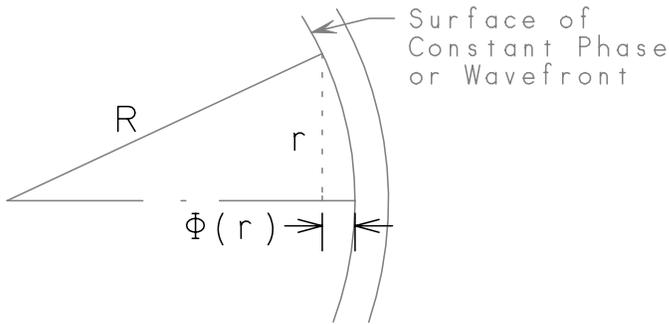
Figure 2 Showing two successive wavefronts with a radius of curvature, R, and the phase as a function of radius, r, from the optical axis.

From the geometry of Figure 2 we can write:

$$R^2 = r^2 + (R - \phi(r))^2 \qquad\qquad \textbf{Equation 3}$$

Solving for the wavefront, ϕ, and then expanding we get:

$$\phi(r) = R - \sqrt{R^2 - r^2}$$

$$\phi(r) \approx R - R \cdot \left(1 - \frac{r^2}{2R^2} + ... \right) \qquad\qquad \textbf{Equation 4}$$

$$\phi(r) \approx \frac{r^2}{2R}$$

where we have assumed that $R \gg r$. Next we look at the slope of the wavefront at r.
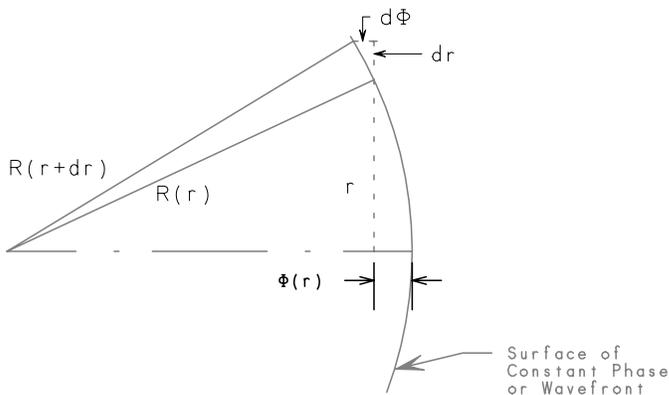

Figure 3 Showing the slope of the wavefront

The slope of the wavefront can be given by $\dfrac{d\phi(r)}{dr} \approx \dfrac{r}{R}$ **Equation 5**

In the Hartmann wavefront sensor this slope is determined by measuring the spatial displacement of a diffraction pattern from that due to a plane wave incident on the same aperture array.
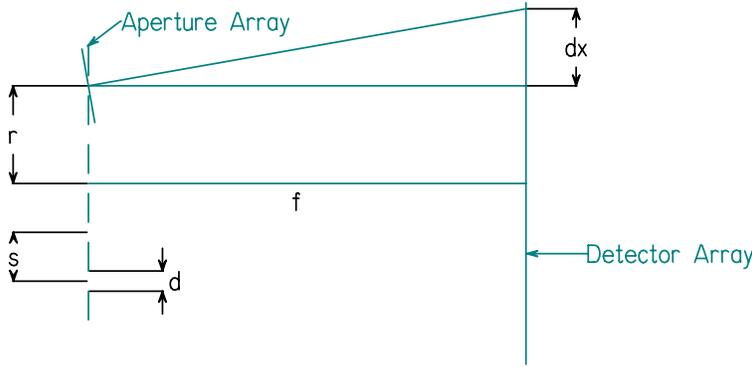
Figure 4 Schematic of the diffraction spot displacement on the detector due to the slope of the wavefront at the aperture array.

The average gradient of the wavefront over the aperture diameter along the x-axis is calculated using the equation,

$$\frac{d\phi(x,y)}{dx} = \frac{\Delta x}{f} \qquad\qquad \textbf{Equation 6}$$

where f is the distance between the aperture array and the detector array.

Wavefront Reconstruction

Once the local wavefront slopes have been determined, the wavefront can be reconstructed by performing a type of integration on the gradient measurements. The two primary types of wavefront reconstruction algorithms are zonal and modal. The zonal wavefront reconstruction is a type of numerical integration. The modal wavefront reconstruction fits the data to a set of orthogonal surface polynomials. The Hartmann Wavefront Analyzer can do two types of zonal reconstruction and one modal reconstructor. Each reconstructor will be explained in this section. Finally, a simple theory predicting the sensitivity of Hartmann wavefront sensors is presented.

Linear Integration

In this simplest of wavefront reconstructors, the computer begins at one edge of the wavefront sensor slope data and defines the wavefront height at each integration area as zero. The height of the wavefront in the next adjacent integration area along the scan direction is calculated as the previous wavefront height plus the slope of the previous integration area times the aperture separation. Mathematically, this is given for the x direction by,

$$\phi^x_{n,m} = \phi^x_{n-1,m} + \frac{\partial \phi_{n-1,m}}{\partial x} \cdot s' \qquad\qquad \textbf{Equation 7}$$

where ϕ is the wavefront surface and s' is the aperture separation. This is also performed in the y direction using the same formula.

After the linear integration is performed along both axes, the wavefront height is found by summing both individual scans, or,

$$\phi_{n,m} = \phi_{n,m}^{x} + \phi_{n,m}^{y}$$

Equation 8

While this reconstructor is very fast, its results are very often noisy.

Southwell Reconstructor

William Southwell introduced a simple zonal wavefront reconstructor that is very popular because of its high speed and high accuracy. [2] The idea behind the Southwell reconstructor is to minimize the error between the reconstructed wavefront and the measured wavefront slopes. To achieve this goal, a scan is made through the entire grid of wavefront slope data and the wavefront height is calculated at each point as average height predicted by each of the neighbors as follows,

$$\phi_{n,m} = \sum_{j=-1}^{1}\sum_{i=-1}^{1} I_{n+i,m+j} \cdot \left[\phi_{n+i,m+j} + \left(\frac{\dfrac{\partial\phi}{\partial x_{n,m}} + \dfrac{\partial\phi}{\partial x_{n+i,m+j}}}{2} \right) \cdot \Delta \right]$$

Equation 9

where $\phi_{n,m}$ is the wavefront at grid location (n,m), Δ is the separation between the (n,m) integration area and the $(n+i,m+j)$ integration area, and $I_{n+i,m+j}$ is the intensity measured in the $(n+i,m+j)$ integration area. The intensity weighting in the above equation is able to be toggled on and off by the user. It is present to give a measure of the validity of a wavefront slope measurement because generally integration areas with high intensity have a more accurate slope measurement. This algorithm takes many iterations through the measured wavefront sensor data to achieve an accurate result. The HWA code allows the user to adjust the number of iterations through this reconstructor. The absolute maximum number of iterations should be equal to the number of integration areas, although it will always converge much faster for real data.

Modal Zernike Reconstructor

There has been much work done on fitting Hartmann wavefront sensor data to Zernike polynomials because the Zernike polynomials are the standard set used for analyzing optical elements and the Hartmann sensor was originally designed for optical metrology. The Zernike polynomials are a set of polynomials used to describe a surface. These polynomials are orthogonal over a unit circle. Orthogonality of Zernike polynomials means that the inner product of any two Zernike polynomials is zero when the inner product is performed over the unit circle. The inner product is the integral of the product of one polynomial times the complex conjugate of the other over the specified limits of orthogonality. Some of the important low-order Zernike polynomials are listed in Table 1.

The derivatives of the Zernike polynomials are not orthogonal, so the derivatives may not be fit directly. Several authors including Galveriedes[3] and Acosta[4] have introduced a sets of orthogonal polynomials whose inner product coefficients give the Zernike coefficients directly.

Since the Zernike polynomials are only orthogonal over a unit circle, the user must specify the Zernike radius in the Zernike reconstructor section of the Setup dialog box under the methods tab. The Zernike radius can automatically be set to double the second moment with a check-box in the setup form. The Zernike circle location can automatically be set to the first moment of the measured intensity pattern on the HWA.

#	Form of the Polynomial	(n,l)	Name
TABLE 1: ZERNIKE TERMS USED IN THE HWA APPLICATION			
0	1	(0,0)	Piston, Not Used
1	$\rho \cos(\theta)$	(1,1)	Tilt x, (about y axis)
2	$\rho \sin(\theta)$	(1,-1)	Tilt y, (about x axis)
3	$\rho^2 \cos(2\theta)$	(2,2)	Astigmatism x, (0°)
4	$2\rho^2 - 1$	(2,0)	Power or Focus
5	$\rho^2 \sin(2\theta)$	(2,-2)	Astigmatism y, (45°)
6	$\rho^3 \cos(3\theta)$	(3,3)	Trefoil x
7	$(3\rho^3 - 2\rho)\cos(\theta)$	(3,1)	Coma x
8	$(3\rho^3 - 2\rho)\sin(\theta)$	(3,-1)	Coma y
9	$\rho^3 \sin(3\theta)$	(3,-3)	Trefoil y
10	$\rho^4 \cos(4\theta)$	(4,4)	Tetrafoil x
11	$(4\rho^4 - 3\rho^2)\cos(2\theta)$	(4,2)	Secondary Astigmatism x
12	$6\rho^4 - 6\rho^2 + 1$	(4,0)	Primary Spherical
13	$(4\rho^4 - 3\rho^2)\sin(2\theta)$	(4,-2)	Secondary Astigmatism y
14	$\rho^4 \sin(4\theta)$	(4,-4)	Tetrafoil y
15	$\rho^5 \cos(5\theta)$	(5,5)	Pentafoil x
16	$(5\rho^5 - 4\rho^3)\cos(3\theta)$	(5,3)	Secondary Trefoil x
17	$(10\rho^5 - 12\rho^3 + 3\rho)\cos(\theta)$	(5,1)	Secondary Coma x
18	$(10\rho^5 - 12\rho^3 + 3\rho)\sin(\theta)$	(5,-1)	Secondary Coma y
19	$(5\rho^5 - 4\rho^3)\sin(3\theta)$	(5,-3)	Secondary Trefoil y
20	$\rho^5 \sin(5\theta)$	(5,-5)	Pentafoil y
21	$\rho^6 \cos(6\theta)$	(6,6)	
22	$(6\rho^6 - 5\rho^4)\cos(4\theta)$	(6,4)	Secondary Tetrafoil x
23	$(15\rho^6 - 20\rho^4 + 6\rho^2)\cos(2\theta)$	(6,2)	Tertiary Astigmatism x
24	$20\rho^6 - 30\rho^4 + 12\rho^2 - 1$	(6,0)	Secondary Spherical
25	$(15\rho^6 - 20\rho^4 + 6\rho^2)\sin(2\theta)$	(6,-2)	Tertiary Astigmatism y
26	$(6\rho^6 - 5\rho^4)\sin(4\theta)$	(6,-4)	Secondary Tetrafoil y
27	$\rho^6 \sin(6\theta)$	(6,-6)	

TABLE 1: ZERNIKE TERMS USED IN THE HWA APPLICATION			
28	$\rho^7 \cos(7\theta)$	(7,7)	
29	$\left(7\rho^7 - 6\rho^5\right)\cos(5\theta)$	(7,5)	
30	$\left(21\rho^7 - 30\rho^5 + 10\rho^3\right)\cos(3\theta)$	(7,3)	Tertiary Trefoil x
31	$\left(35\rho^7 - 60\rho^5 + 30\rho^3 - 4\rho\right)\cos(\theta)$	(7,1)	Tertiary Coma x
32	$\left(35\rho^7 - 60\rho^5 + 30\rho^3 - 4\rho\right)\sin(\theta)$	(7,-1)	Tertiary Coma y
33	$\left(21\rho^7 - 30\rho^5 + 10\rho^3\right)\sin(3\theta)$	(7,-3)	Tertiary Trefoil y
34	$\left(7\rho^7 - 6\rho^5\right)\sin(5\theta)$	(7,-5)	
35	$\rho^7 \sin(7\theta)$	(7,-7)	

Wavefront Sensitivity

The sensitivity of the wavefront sensor depends on the sensor's ability to determine the position of the diffracted spots on the CCD array. The error in determining the diffracted spot location includes a set of factors such as CCD detector noise, coherent optical crosstalk between diffracted spots from different apertures, and CCD digitization error. The RMS error in the diffracted spot location for the wavefront sensor was measured by evaluating the spot position of a stable wavefront repeatedly with the sensor and measuring the RMS error in the spot position location. It is possible to have errors in the measurement of the diffracted spot locations that are larger due to a variety of factors including coherent optical crosstalk between adjacent detector array elements.

The RMS wavefront error is measured by illuminating the sensor with a beam of uniform-intensity with a planar wavefront. The sensor evaluates the wavefront many times and determines the RMS difference between the individual measured wavefronts and an average wavefront evaluated at each of the apertures. The RMS wavefront error, ϕ_{RMS}, associated with a Hartmann wavefront sensor is related to the quality of the reconstruction algorithm, the diffracted length of the apertures, the diameter of the apertures, and the RMS spot location error. The ϕ_{RMS} can be expressed as,

$$\phi_{RMS} \approx F_{reconstructor} \cdot \frac{d}{f} \cdot \Delta x_{RMS}$$
Equation 10

where $F_{reconstructor}$ is a factor determined empirically by the type of reconstructor used, d is the diameter of a aperture in the aperture array, and f is the aperture diffracted length. The reconstruction factor is usually between 1 and 10 for most reconstructors.

When the HWA application reports the RMS wavefront error, it is simply a measure of the average deviation of the reconstructed wavefront from the plane wave case:

Equation 11 Reported WF rms

$$\phi_{RMS} = \sqrt{\sum_n \sum_m \phi_{n,m}^2 \cdot \frac{1}{N}}$$

Differential Wavefront Measurements

The Hartmann wavefront sensor can be used to make differential wavefront measurements in addition to making measurements using an optical flat as a reference. The advantage of this type of measurement is that the aberrations in the optical system can be calibrated out and only a change in the optical system will be recorded. This is especially useful for adaptive optics and thermal lensing measurements.

A new reference file must be established before the differential measurement can be used. To perform a differential measurement, the HWA is setup in an optical system and an image is recorded by the CCD array. The image must be processed to identify the new integration areas and the new reference spot positions. The algorithms in the HWA program to identify the focal spots and their integration areas are discussed in this section.
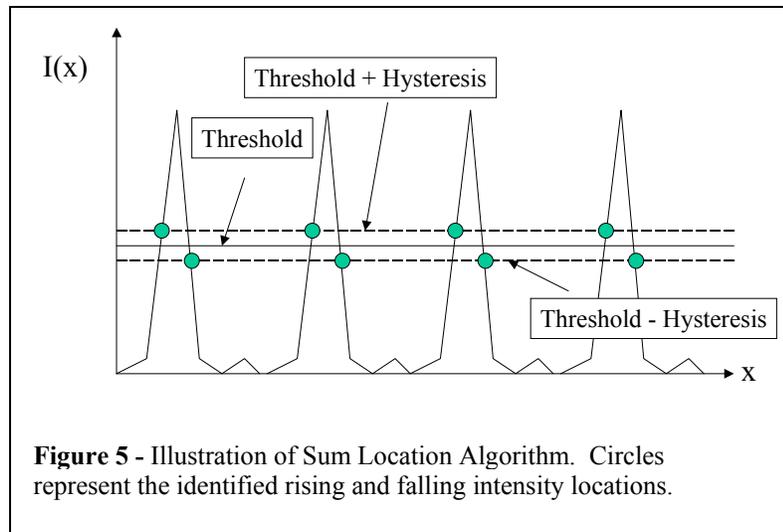
Peak Location

This algorithm scans through the image and identifies the approximate location of focal spots by looking for local intensity peaks. An intensity peak is a pixel where the neighboring 8 pixels have equal or less intensity than the central pixel. To avoid intensity noise being mistaken as a legitimate diffraction peak, the peak must be at least above a certain intensity level specified as a percent of the full intensity range by the user. Furthermore, the separation between peaks must be at least a distance equal to the user-specified peak separation factor times the aperture separation to avoid noisy intensity profiles being mistaken as intensity peaks.

Blob Location

This algorithm uses the classic image analysis technique called blob identification. The intensity profile is simplified into two states (on or off) by applying a user-specified count threshold to the image. Then blobs are identified as groups of neighboring pixels above the threshold. Each separate blob is assigned a number. To avoid noise being counted as a blob, each blob must have a minimum number of pixels that is specified by the user. Each blob with the minimum number of pixels is identified as a diffracted spot from an aperture. The integration area for each diffracted spot is defined as a square with each side as long as the aperture separation.

Sum Location

Sum location begins by summing the image along the two axes so that a one-dimensional intensity pattern is found for each axis. If the apertures are properly aligned to the CCD, the resulting sum is a series of peaks corresponding to the diffraction spots rows and columns. To work with the same intensity counts, we divide the sum by the number of pixels along the orthogonal axis. We begin by scanning the along the sum until the intensity value exceeds a count value equal to the sum of the

Figure 5 - Illustration of Sum Location Algorithm. Circles represent the identified rising and falling intensity locations.

threshold and the hysteresis, both specified by the user. The location of this point is noted and the scan continues until the intensity level drops below the threshold minus the hysteresis. At this point, both sides of the diffracted spot are identified and the location of the diffracted spot is assumed to be exactly between the two sides of the peak. The analysis is repeated for both axes and the approximate location of the rows

and columns of the diffracted spots has been determined. Then an integration area is arbitrarily defined at the location of each intersection of the identified row and column with a size equal to the aperture separation. If the sum of the intensity points exceeds 100 counts in the integration area, the integration area is recorded. Otherwise it is neglected due to lack of enough information to create a meaningful spot position.

Integration Area Post-Analysis

After determining the location of the integration areas, several routines are used to fine tune their position and determine the regular grid of the spots. The centroid is performed on all the integration areas and the position of each integration area is then re-centered on its centroid. Then the grid is established relative to the minimum x and y coordinate of the list of integration areas. This coordinate forms the origin $(0,0)$ of the grid, as shown in Figure 6. Each integration area is then placed on the grid by subtracting the minimum x and y position and dividing each by the aperture separation in pixels. The grid in the code is then an array of pointers to the different integration areas in the integration area linked list. Finally, the centroid of all the integration areas is determined and the set of integration areas and their centroid is called a reference.

Figure 6 - Illustration of the $(0,0)$ grid location.
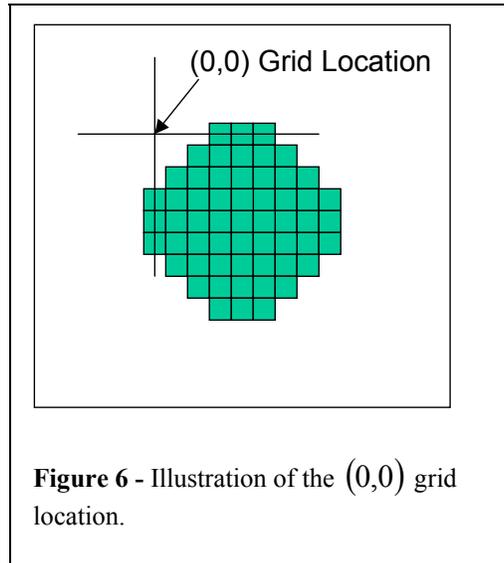
Limitations to Differential Wavefront Analysis

Creation of a reference using the HWA is limited in some respects. First, all of the subapertures must be illuminated and the illumination across the entire detector must be uniform to within 10%. Second, the wavefront slopes at any given subaperture must not exceed the dynamic range of the instrument. By looking at the live image and displaying the integration areas, you can see when the diffracted spots begin to cross into other integration areas, when this occurs it is not possible to make a usable reference. Because of these limitations your test setup must include some attempt to obtain a relatively flat wavefront to use as your reference and a similar arrangement for your test surfaces. See the installation and setup guide for ideas on such setups.

Wavefront Sensor Data Post-Processing

Once the wavefront sensor has made the measurement of the intensity and phase of the incident optical beam, many different forms of data processing can be performed on the measurement. The HWA code has the ability to calculate the first and second moments, wavefront slope statistics, radius of curvature determination and Hermite-Gaussian modal decomposition. Each of these types of analysis will be described in this section.

First and Second Moments

The first and second moment analysis can be performed on the overall intensity profile measured by the HWA. The intensity of each integration area is the sum of the intensity at each of the pixels above the centroid threshold. The first moment, or centroid, is described above for a single integration area. The same formula applies to the entire intensity profile. By summing the intensity over an integration area we find an intensity value for the corresponding sub-aperture. By applying our first moment formula to these intensity values at the locations of the sub-apertures, we find a first moment for the entire beam profile.

$$\overline{x}' = \frac{\sum\limits_{n=0}^{n_{\max}} \sum\limits_{m=0}^{m_{\max}} I(n,m) \cdot n}{\sum\limits_{n=0}^{n_{\max}} \sum\limits_{m=0}^{m_{\max}} I(n,m)} \cdot s'$$

Where n and m are the indices designating sub-apertures.

The position indicated is now the centroid of the whole beam intensity profile since the indices refer to sub-aperture locations, the spacing is the separation of the sub-apertures and the intensity is the sum of intensity over the integration area corresponding to the indexed sub-aperture. The x and y values indicated are reported as the beam profile centroid location, measured from the center of the sensor.

The second moment calculation is an indication of the beam profile radius. For the beam half-width (radius in the case of a circularly symmetric beam) in one direction: $w_x = 2\sigma'_x$ where,

$$\sigma'^2_x = \frac{\sum\limits_{n=0}^{n_{\max}} \sum\limits_{m=0}^{m_{\max}} I(n,m) \cdot \left(n - \frac{\overline{x}'}{s'}\right)^2}{\sum\limits_{n=0}^{n_{\max}} \sum\limits_{m=0}^{m_{\max}} I(n,m)} \cdot s'^2$$

This is the second moment of the beam intensity profile. The HWA application reports the second moment based beam widths $d_{\sigma_x} = 4\sigma'_x$.

Wavefront Slope Statistics

Statistics on the wavefront slope are sometimes helpful in determining the performance of the wavefront sensor. The HWA application can report the average wavefront slope and the RMS wavefront slope. The RMS wavefront slope is directly proportional to the RMS centroid accuracy when measuring a flat wavefront, and is thus a good measure of the wavefront sensor noise performance.

Radius of Curvature Determination

The HWA code can scan through the wavefront slope data and fit it to the equation of a line in the two axes. The linear fit coefficient is the reciprocal of the wavefront radius of curvature. The constant term is the average slope in that axis. This fit is useful for determining a Gaussian beam's curvature or for calibrating the separation between the Hartmann array and the imager.

Hermite-Gauss Modal Decomposition

The Hermite-Gauss modes are an orthogonal set of stable resonant modes.[5] It is often useful for laser manufacturers to know about the modal content of the laser during its fabrication since most laser builders want only the lowest order mode to be amplified. The HWA code takes the measured intensity and phase data and performs the overlap integral of the data with an ideal set of Hermite-Gaussian modes with a waist specified by the user in the Zernike reconstructor section. The waist radius can automatically be set to double the second moment with a check-box in the setup form.

Practical Limitations of Hartmann Sensing

Hartmann wavefront sensors can be used for a wide variety of applications, but like any sensor technology, there are some limitations to their applicability. In this section, we will discuss some of the limitations and their origins. Specifically, we will discuss limitations in spatial resolution and wavefront height measurement range. We will note that there is active research addressing all of the sensor limitations, but we will limit our discussion to the standard usage of the Hartmann sensor.

Spatial Resolution Limitations

The spatial resolution of any Hartmann sensor is limited by the fact that each aperture measures the average tilt over that aperture. The Nyquist criterion indicates that we can only measure a spatial frequency equal to half the inverse of the aperture spacing. Practically, at least 6 samples should be taken per spatial frequency period to get a good measurement of a sinusoidal wavefront.

With this in mind, the wavefront sensor is not good at measuring high spatial frequency aberrations. One way to increase the spatial resolution of a Hartmann sensor is to use a lens system to magnify the incident beam. Unfortunately, the height of the wavefront remains constant as the beam is magnified. This means that the wavefront tilt decreases by the same magnification factor as the lateral dimension increase. Thus, spatial resolution is obtained at the cost of wavefront tilt sensitivity. This magnification effect makes the Hartmann wavefront sensor more sensitive for measuring large optics, but less sensitive for measuring small optics.

Wavefront Height Measurement Error

The minimum wavefront height is limited by the noise in measuring the position of each of the diffracted spots on the CCD. The error in measuring the diffracted spot position is attributable to a variety of sources including temporal and fixed-pattern CCD noise, digitization noise, coherent aperture diffraction cross talk, and manufacturing defects. Since there are many other factors that contribute to the spot position determination error, this noise has to be measured for each new Hartmann sensor system. Despite the complexity of the noise analysis, modeling shows that camera dynamic range and the number of pixels per diffracted spot directly affect the noise of the diffracted spot position measurement. The error in the spot position measurement decreases as the square root of the number of pixels per diffracted spot and linearly with the true intensity dynamic range.

The RMS wavefront slope measurement can be calculated by dividing the RMS spot position determination error by the separation between the CCD and the Hartmann array. An estimate of the RMS wavefront error then can be made by multiplying the RMS wavefront slope by the separation between apertures and by a scaling factor of order unity corresponding to the accuracy of the wavefront reconstructor. The wavefront reconstructor factor for a Zernike reconstructor is about three. For a zonal reconstructor like the Southwell one, it is typically five.

As an example, assume an RMS focal spot position measurement error of 0.2 microns, an aperture separation of 280 microns, and a CCD to Hartmann array separation of 8mm. The RMS wavefront tilt is about 25 micro radians. For a reconstructor factor of five, the RMS wavefront measurement error is 35nm.

Wavefront Height Measurement Maximum

The maximum wavefront height a Hartmann sensor can measure is determined by the maximum throw of the position of the diffracted spots. Figure 7 illustrates the maximum motion of a diffracted spot within an integration area. We will define the maximum throw of the diffracted spots as when the diffracted spot edge move to the edge of its integration area, which is equal in size to the aperture spacing. Assuming a Hartmann array composed of an array of circular apertures with diameter d and a spacing s being illuminated by light with wavelength λ separated from the CCD by a distance f, the maximum wavefront slope can be written as,
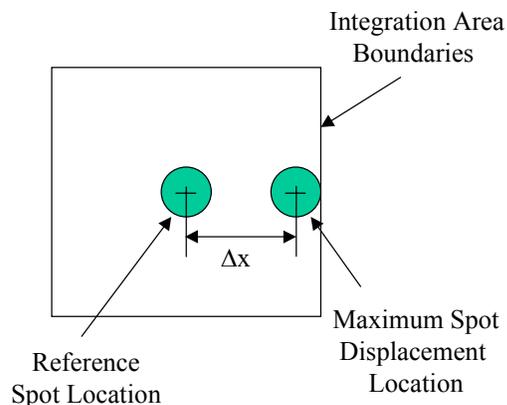

Figure 7 - Diagram illustrating maximum diffracted spot motion in a given integration area.

$$\frac{\partial\phi}{\partial x} = \frac{\left(\frac{s}{2} - 1.22 \cdot \frac{f\lambda}{d}\right)}{f}$$

Equation 12

For example, consider a Hartmann array with an aperture separation of 280 microns and an aperture diameter of 140 microns separated from a CCD by 8mm being illuminated by 633nm light, the maximum wavefront slope that the sensor can observe is 12 milliradians.

Although the maximum wavefront slope is straightforward to calculate, the maximum peak-to-valley wavefront is more difficult since it depends on the spatial frequency of the wavefront being measured. For example, consider a sinusoidal wavefront given by,

$$\phi(x) = A \cdot \sin\left(\frac{2\pi x}{\Lambda_{spatial}}\right)$$

Equation 13

where 2A is the wavefront peak-to-valley amplitude and $\Lambda_{spatial}$ is the spatial period of the aberration. The wavefront slope is then given by,

$$\frac{\partial\phi(x)}{\partial x} = A \cdot \left(\frac{2\pi}{\Lambda_{spatial}}\right) \cdot \cos\left(\frac{2\pi x}{\Lambda_{spatial}}\right)$$

Equation 14

Thus the peak-to-valley wavefront slope is inversely proportional to the spatial period of the aberration. High spatial frequency aberrations will therefore produce larger peak-to-valley wavefront slopes than low spatial frequency aberrations of the same amplitude.

If we limit our discussion to a beam with only spherical curvature, we can establish a maximum peak-to-valley wavefront distortion based on the maximum wavefront slope. If we assume that the wavefront is given by,

$$\phi(r) = \frac{r^2}{2R}$$

Equation 15

where r is the spatial radius of the beam and R is the wavefront radius of curvature. Then the wavefront slope is given by,

$$\frac{\partial\phi}{\partial r} = \frac{r}{R}$$

Equation 16

Setting this equal to the maximum slope equation, the minimum radius of curvature is given by,

$$R_{min} = \frac{r_{max}}{\frac{\partial\phi}{\partial r_{max}}}$$

Equation 17

where r_{max} is the maximum spatial radius of the sensor, which is usually half the diagonal sensor length. Then the maximum peak-to-valley spherical wavefront distortion the sensor can measure is given by,

$$\phi_{max} = \frac{r_{max}^2}{2R_{min}}$$

Equation 18

Using the above example information and assuming a CCD with an 8mm diagonal length, we calculate a minimum radius of curvature of 333mm and a maximum peak-to-valley spherical wavefront distortion of 24 microns.

Other Optical Metrology Technologies

There are many different devices for performing optical metrology besides Hartmann sensing like interferometers, Foucault knife-edge testing, and curvature sensing. In this section, we will briefly compare Hartmann sensing to these techniques and outline where the Hartmann sensor is superior.

Interferometry

Interferometry is especially good for measuring high spatial frequency aberrations and low amplitude aberrations. Unfortunately, air motion and mechanical vibrations make obtaining an image with an interferometer difficult, especially when testing large beams or large optics. Also, sophisticated software is necessary to extract meaningful and accurate information from interferograms. In contrast to interferometers, the Hartmann wavefront sensor can be made effectively vibration insensitive by averaging several frames together to remove the temporal noise. Furthermore, the interpretation of the wavefront sensor data is much more straightforward. Finally, commercial interferometers typically cost $100,000, where a high quality wavefront sensor is typically much less expensive.

Foucault Knife Edge Tests

Foucault knife-edge testing involves moving a knife-edge through the focus of a beam and observing the intensity pattern on a screen. Like interferometry, knife-edge testing allows high-spatial frequency aberrations to be observed, but it requires very accurate alignment of the knife-edge to the beam focus. Hartmann sensing requires only very simple tip-tilt alignment and the HWA software provides simple feedback to the user on the alignment of the sensor.

Curvature Sensing

Curvature sensing is a technique used typically in adaptive optics to measure the Laplacian of the wavefront by subtracting intensity profiles from an exact distance before and after the focus of a lens. For adaptive optics systems, the image of before and after the focus is usually switched mechanically making the systems quite noisy during operation. In contrast, the Hartmann wavefront sensor contains no moving parts making it more reliable. The largest disadvantage of curvature sensing is that it has not been made widely commercially available yet.

Shack-Hartmann

What is the difference between the Shack-Hartmann & Hartmann wavefront sensors? When Hartmann invented his wavefront sensor, he was working in an optics laboratory with a bright light source, so he was able to use the diffraction from an array of apertures in a sheet of metal to measure the effect that a test optic had on the optical wavefront. In 1971 Roland Shack and Ben Platt modified the Hartmann senor by replacing the array of hard apertures that Hartmann had used with an array of discrete micro-lenses. There were two advantages to this architecture. The Shack-Hartmann wavefront sensor offered better photon efficiency. This was critical to the astronomical community that was at the time searching for a wavefront sensor for their adaptive optics systems, since they work in a photon-starved regime. Also, it was shown that the diffraction pattern at the focus of a lens is exactly the Fourier transform of the electric field pattern illuminating that lens. Thus, the position of the focal spot was only proportional to the average wavefront slope over the aperture and was independent of higher-order aberrations and intensity profile variations. Finally, the Shack-Hartman wavefront sensor examines the entire wavefront, not just small samples like the Hartmann sensor does.

How significant are these differences? Today, most of the users of Hartman-type wavefront sensors, either Shack-Hartmann or Hartman, are working in a photon rich environment, making the photon efficiency benefit negligible. To address the effect of intensity profile changes affecting the spot position, Spiricon

has performed rigorous Huygens-Fresnel propagation simulations, which have shown that when the light illuminating the aperture is allowed to diffract over a substantial distance, the effect of intensity profile variations is below the sensor's ability to sense it. Finally, although the Hartmann sensor is discarding a portion of the light illuminating the sensor, the accuracy of the sensor is not compromised because there is an equivalent spatial resolution limitation to both sensors. Thus, the only case where any difference will be observed between the two sensors is when the wavefront contains high spatial frequency components that cannot be measured by either sensor anyway. After much modeling and analysis, our conclusion is that unless you are working in a photon-starved regime like astronomy, the Hartmann wavefront sensor will perform equivalently to the Shack-Hartmann wavefront sensor.

Why choose a Hartmann over a Shack-Hartmann wavefront sensor?
There are several significant advantages of the classical Hartmann architecture over the Shack-Hartmann architecture. The cost of fabricating an array of lenses is significantly higher that that of fabricating an array of apertures. To make money on the device, the manufacturer has to pass that cost on to the consumer, making the Shack Hartmann wavefront sensor more expensive.

It is much easier to fabricate an array of apertures than an array of lenses. All the known techniques for fabricating lens arrays, including gray mask technology, reflowing photoresist, and binary optics, are very susceptible to manufacturing errors causing the lens arrays not to behave as the ideal Fourier transform element of the theory and making them susceptible to mapping intensity variations into wavefront noise. Furthermore, the manufacturing defects can also create scatter points, which create more false wavefront signals. Measuring the focal length of a lens array is difficult due to their small size and small numerical aperture. Often processing mistakes change the lens array focal length by as much as 10%. If the manufacturer does not know the focal length, they cannot place the focal plane on the CCD and are again susceptible to intensity variations. All known lens arrays are highly susceptible to chromatic aberrations due to the fact that all refractive materials suffer from dispersion. Thus the focal plane varies with respect to wavelength, making the device only truly independent of intensity variations at one wavelength. Finally, since the focal length of the commonly used lens arrays is less than 25mm, extraordinary care must be taken in mounting the lens array to place the focal plane exactly on the CCD. If the mounting is off any significant amount, the Shack-Hartmann wavefront sensor will again be susceptible to intensity variations creating wavefront noise.

Hartmann wavefront sensors are inherently slightly susceptible to intensity variations creating wavefront noise, but not through as many mechanisms as the Shack-Hartmann wavefront sensors. The manufacturing defects are minimal at best due to the use of mature silicon processing technology capable of writing 50nm features. The far-field diffraction pattern produced by illuminating an aperture is independent of wavelength, thus eliminating classical dispersive chromatic aberration problems. The manufacture of the Hartman wavefront sensor is also easier since precision mounting components are not required. Any imprecision in the mounting can be measured during the calibration and its effect can be removed from the sensor operation.

One final advantage of Hartmann sensing over all of the above methods is that it is almost entirely wavelength independent. The Hartmann sensor works as well with incoherent light as it does with coherent light. In fact some modeling indicates it works better with incoherent light due to reduced interference effects. This effect allows the user to easily switch between different illumination wavelengths, such as during non-linear optics experiments.

Conclusion

The Hartmann sensor is a tool for optical metrology that is widely applicable in any optics laboratory. It allows the user to make a wide variety of measurements that would be difficult with interferometry, Foucault testing, or curvature sensing. The Hartmann sensors have been used in optics shops for almost a century now and are being widely applied today in adaptive optics and laser metrology.

References

[1] J. Hartmann. "Bemerkungen uber den Bau und die Justirung von Spektrographen," Z. Instrumentenkd., **20**, 47 (1900).

[2] W. H. Southwell. "Wave-front estimation from wave-front slope measurements," J. Opt. Soc. Am. **70**, 998-1006 (1980).

[3] A. Gavrielides. "Vector polynomials orthogonal to the gradient of Zernike polynomials", Opt. Let. Vol **7**. No. 11, November 1982.

[4] E. Acosta, S. Bara, M. A. Rama, and S. Rios. "Determination of phase mode components in terms of local wave-front slopes: an analytical approach", Opt. Let. **20**, 1083-5.

[5] A. E. Siegman. *Lasers* (University Science, 1986).