Wave-front sensing

Lisa A. Poyneer
Lawrence Livermore National Lab

CfAO Summer School
UC Santa Cruz
Monday, August 11, 2003

This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract No.W-7405-Eng-48.
Wave-front sensing is the first key step in Adaptive Optics

- You can’t correct well if you can’t measure well
- Goals for WFS:
  - high accuracy (low or no bias)
  - low noise
  - ease, low cost of implementation and use
  - robust to changes in source
Course outline

- Basic AO and WFS concepts
- Optical WFS techniques
  - Curvature, Shack-Hartmann, Pyramid
- Advanced topics:
  - aliasing
  - in-depth analysis of the centroid algorithm
  - using scenes with Shack-Hartmann sensor
  - future problems: laser GS
What are we trying to measure?

- We are trying to measure the phase aberration (usually) along the optical path of a reference.
- Commonly measure a derivative of the phase at discrete locations across the pupil.

Other optical options:
- measure magnitude of zernike modes
- measure phase directly
Curvature Sensing

- Curvature sensing uses two defocused images in the focal plane
- Estimate the curvature with normalized difference of intensities

\[
\frac{I_2 - I_1}{I_2 + I_1} \approx z \nabla^2 \Phi
\]
More details of curvature sensing

- Amount of defocus to use depends on the phase aberration
  - can determine this for atmosphere
  - can change in closed-loop if designed correctly
- Acquisition is done in practice with time-multiplexing of a single detector to both image planes
- Requires 2 pixels per subaperture
- Usually used for low-order correction on dim stars
Sensor measures the average phase gradient across the subaperture.
This is Fourier optics - the spot shifts based on linear phase term.
Shack-Hartmann forms an array of spots

- Note illumination changes near edge
- Actual sampling of spots/spot size is variable and system dependent

WFS image from Abdul Awwal
More details for Shack-Hartmann

- Need to have a reference location that the spots shift away from (may not be null)
- Spots diffraction-limited by subaperture size or the seeing
- Requires at least 4 pixels per subperture
Pyramid sensing

- Pyramid sensor divides the field in the *focal plane* into 4 quadrants
- Reimage into pupil plane and compare intensities of same pixel in all subpupils

Figure from Richard Clare
Consider a focus phase term, with linear slope.

X- and y-slopes. Note the lack of linearity.
Modulation of pyramid leads to wider linear range

• Move the pyramid (or the spot) around in an optimal pattern

Modulation pattern of field intensity

X- and y-slopes.
Wider linear range
• Computation:
  • All require small amount of computation on pixels
  • Slope sensors w/ DM require phase reconstruction
  • Curvature sensor w/ bimorph applies curvatures directly (in theory)

• Signal-to-noise
  • Which is better is quite contentious.... and unresolved
  • Must consider read noise of detector
Mini-break!
Aliasing limits AO performance by fooling the system

High-frequency errors incorrectly measured as low-frequency errors

Aliasing of a cosine

- Actual cosine
- Aliased cosine

Spatial dimension vs. amplitude graph

0 2 4 6 8 10 12
-1 -0.5 0 0.5 1
Spatially-filtered WFS implemented with a field stop

• Spatial filter is a field stop of width $\lambda/d$ in focal plane
• Designed to reject phase past $1/2d$ in spatial frequency
Spatially-filtered WFS suppresses high-spatial-frequency phase

- Cosine phase aberration scatters light to a specific location
- Field stop will reject this light

For discussion on PSF formation, see Perrin et al “Structure of high Strehl...“ ApJ in press
Correction of a segmented primary

- Correct phase discontinuities between segment edges
- Model based on detailed info of Keck primary

Segmented primary mirror phase error - 110 nm rms

personal communication with M. Troy and C. Carrano
Spatial filter cleans up segment errors

Regular AO residual
23 nm

SFWFS AO residual
19 nm

This is a small difference in rms error but...

Poynneer
Correction of segments, ExAO noise-free simulation, D/d = 62
Filter produces a region out to 1 arcsec that has 50 times less light than without the filter.

Correction of segments, ExAO noise-free simulation, D/d = 62
Broadband light reduces the area of improvement in PSF

- Short $\lambda$ light sees stop as too big
- Long $\lambda$ light sees stop as too small
- Extra error around $1/2d$ due to aliasing and missing info

Broad-band WFS simulation used 700-900 nm light with field stop set for 800 nm
The x-slope is based on the difference between linear combinations of the pixels on the right and left sides.

Most CfAO systems use the Shack-Hartmann because it is inexpensive and easy to implement.

\[ \hat{x} = \frac{r - l}{t} \]
For centroid on $N \times N$ pixels, performance can be analyzed

- Weight pixels by distance for center
- $N \times N$ has center at value $N/2 - 1/2$

\[
\begin{align*}
    r &= \sum_{i=N/2}^{N-1} \sum_{j=0}^{N-1} (i - N/2 + 1/2)s[i, j] \\
    l &= \sum_{i=0}^{N/2-1} \sum_{j=0}^{N-1} (-i + N/2 - 1/2)s[i, j] \\
    t &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} s[i, j]
\end{align*}
\]
Dealing with photon and read noise

- Each pixel in the spot has an normalized expected number of counts (poisson) scaled by exposure time + white read noise

\[ s[i, j] = f \times p[i, j] + n[i, j] \]

- Mean and variance are

\[ m_s[i, j] = f \lambda[i, j] \quad \sigma_s^2[i, j] = f^2 \lambda[i, j] + \sigma_n^2 \]
Use analysis to determine algorithm behavior

- You should fully characterize your algorithm
  - how does noise propagate, both photon and read?
  - what happens if the spot size/shape changes?
  - do you need background subtraction?
Break up into two independent terms: photon and read noise

- Error due to photons is inversely related to SNR
  \[
  \text{Var}(\hat{x})_p = \frac{\sigma_r^2 + \sigma_l^2}{f m_t^2}
  \]
  RMS error is inverse power law in SNR

- Error variance due to read noise depends on the square of the total number of pixels
  \[
  \text{Var}(\hat{x})_n = \sigma_n^2 \left( \frac{N^2(N^2 - 1)}{12m_t^2} \right)
  \]
Background light reduces gain of centroid estimate

- With no background, the expected slope is
  \[ E[\hat{x}] = \frac{m_r - m_l}{m_t} \]

- With poisson background (mean \( b \)) added in, expected slope is now
  \[ E[\hat{x}] = \frac{m_r - m_l}{m_t + N^2b} \]

Incorrect background subtraction, even by a few counts, can lead to significant losses in gain and increased error.
The quad-cell method suffers from gain changes with spot size

- Quad-cell is 2x2 centroid
- As spot size changes, so does gain
- Lost light off edge of pixel also biases estimate

Est is 0.28
Est is 0.14

Figure from Marcos van Dam
Using 4x4 pixels ameliorates the gain problem

- 4x4 centroid has a more regular response with spot size than quad-cell.
Mini-break!
What are other algorithmic options?

- If you have very accurate knowledge of the spot profile, you can do a maximum likelihood estimate.

- Can correlate the spot against a reference:
  - much lower noise propagation
  - expected estimate is insensitive to background levels
  - requires more computation
Correlation has lower read noise propagation

• Error variance due to noise does not depend of number of pixels

\[
\text{Var}(\hat{x})_n \propto \frac{\sigma_n^2}{8(m_0 - m_1)^2}
\]

• Data from Vision AO system show this clearly in comparison to centroiding

<table>
<thead>
<tr>
<th></th>
<th>Noise prop</th>
<th>(\sigma_x)</th>
<th>(\sigma_y)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Correlation</strong></td>
<td>1.58 \times 10^{-5}</td>
<td>3.8 \times 10^{-3}</td>
<td>4.23 \times 10^{-3}</td>
</tr>
<tr>
<td><strong>Centroider</strong></td>
<td>212 \times 10^{-5}</td>
<td>19.6 \times 10^{-3}</td>
<td>20.9 \times 10^{-3}</td>
</tr>
</tbody>
</table>
Correlation works with spot deformities

- Just like background, spot features throw off the centroider
- Correlation (with a spot reference) is insensitive to these features

Centroider (w/ thresholding)

Correlation

WFS data provided by Kai LaFortune
Summary: algorithm options

- What do you know about the point source?
  - how well do you know the spot profile?
  - do the spots deform or change size?
- What are your performance constraints?
  - are you computationally limited?
  - how much noise can you tolerate?

- The choice of best algorithm depends on your system
Scene-based wave-front sensing uses the scene, not a point source

- Subimage of scene will shift just like a point source
- Need to field stop down

WFS camera image of one quadrant of subapertures

Camera image from C. Thompson and R. Sawvel
Estimate shift with a correlation-based algorithm

- Compute correlation to find minimum MSE between images

\[ \hat{x}_0 = \Delta x + \frac{0.5(C_{-1} - C_1)}{C_{-1} + C_1 - 2C_0} \]
Performance can be rigorously analyzed

- Probabilistic analysis of scene quality based on image content
- In the zero-shift (closed-loop case) the error variance is

\[ \sigma_x^2 = \frac{\sigma_1^2 - \sigma_{-1,1}^2}{8(m_0 - m_1)^2} \]

sharpness of autocorrelation
Scene performance depends on frequency content

- Performance varies along x- or y-axis
- Need to find best field of view for expected scene scales

\[ \sigma_x = 0.013 \]
\[ \sigma_y = 0.011 \]
\[ \sigma_x = 0.009 \]
\[ \sigma_y = 0.019 \]
\[ \sigma_x = 0.013 \]
\[ \sigma_y = 0.008 \]
\[ \sigma_y = 0.087 \]
Scaling laws - exposure time

- In the case of changing exposure time, the error st. dev. follows as inverse power law in SNR

\[ \sigma_x(f) = \frac{1}{\sqrt{f}} \left( \tilde{\sigma}_1^2 - \tilde{m}_1 - \tilde{\sigma}^2_{-1,1} \right)^{1/2} \frac{1}{2\sqrt{2}(\tilde{m}_0 - \tilde{m}_1)} \]

Same inverse power law in SNR as point sources
Scaling laws - angle of observation

- Generate radiometric models for background levels
- Performance falls off with too much background

\[
\sigma_x(b, f) = \left\{ Nb^2 + 2bf^2(\tilde{m}_0 - \tilde{m}_2 + \tilde{f}^{-1}) + f^3[\tilde{\sigma}^2_1 - (f - 1)f^{-1}\tilde{m}_1 - \tilde{\sigma}^2_{-1,1}] \right\}^{1/2} / 2\sqrt{2f^2(\tilde{m}_0 - \tilde{m}_1)}
\]
Future problems: Laser Guide Stars

• The LGS ‘spot’ is elongated for off-axis subapertures
• This elongation poses significant problems
  • how many pixels are necessary?
  • what is the best algorithm for slope estimation?
    • noise propagation concerns
    • bias and gain concerns

Sample 3x elongated spot
Papers to look up (not exhaustive)

- Shack-Hartmann WFS
  - van Dam & Lane, JOSA A 17, p1319-1324
  - Tyler & Fried, JOSA A 72, p804-808
  - Veran & Herriot, JOSA A 17, p1430-1439
  - dePater et al, Icarus 160, p359-374
- Solar AO
  - Rimmeele et al, SPIE 4007, p218-231
  - Rimmeele et al, SPIE 1542, p186-193
- Correlation/Scene-based WFS
  - Poyneer, Applied Optics in press
  - Poyneer, SPIE 5162 in press
- Spatial-filtering
  - Poyneer & Macintosh, submitted to JOSA A
  - Poyneer & Macintosh, SPIE 5169 in press
- Pyramids
  - Clare and Lane, SPIE 5169 in press
  - and many others...