Wavefront Sensing and Control Architecture for the Spherical Primary Optical Telescope (SPOT)

Bruce H. Dean, Jeff S. Smith
Optics Branch / Code 551

Jason G. Budinoff
Electromechanical Systems Branch / Code 544

Lee Feinberg
Instrument System Technology Division / Code 550

NASA Goddard Space Flight Center
Greenbelt MD 20771

Abstract

Testbed results are presented demonstrating high-speed image-based wavefront sensing and control for a spherical primary optical telescope (SPOT). The testbed incorporates a phase retrieval camera coupled to a 3-Mirror Vertex testbed (3MV) at the NASA Goddard Space Flight Center. Actuator calibration based on the Hough transform is discussed as well as several supercomputing architectures for image-based wavefront sensing. Timing results are also presented based on various algorithm implementations using a cluster of 64 TigerSharc TS101 DSP’s (digital-signal processors).

1. Introduction

Future light-weighted and segmented primary mirror systems require active optical control to maintain mirror positioning and figure to within nanometer tolerances. An image-based wavefront sensor offers a simple solution that differs from conventional wavefront sensing approaches (e.g., Shack-Hartmann or shearing interferometry) in that complicated optical hardware is replaced by a computational approach where the science camera itself serves as the wavefront sensor. Although the image-based approach is simpler to implement in hardware, and thus preferred for space-optics deployment due to reduced risk of optical system failure, the image-based approach trades hardware for a software solution and is therefore computationally intensive. As a result, near real-time operation of the wavefront sensor requires substantial floating-point performance in addition to high data transfer and communication rates.

In this paper, we give an overview of the technology created to address these problems. The 3-Mirror Vertex testbed (3MV) is an electro-optical system built at the NASA Goddard Space Flight Center, for prototyping and testing wavefront sensing and control algorithms. The hardware and wavefront sensing and control architecture is discussed in Section 2, which includes a Firewire phase retrieval camera, 1 a segmented spherical primary mirror under actuator control, as well as motorized translation stages for positioning the phase retrieval camera. An overview of the wavefront sensing and control algorithms is discussed in Section 3. In Section 4, we describe an implementation of the Hough transform which automates the characterization of actuator influence functions in a turbulent environment. The testbed also utilizes a cluster of 64 digital signal processors (DSP’s) for accelerated processing of the diversity defocus images collected by the phase retrieval camera. The DSP hardware, wavefront sensing and control architectures, and performance under a variety of data collection scenarios are presented in Section 5.

2. 3MV Testbed

The 3-Mirror Vertex (3MV) testbed originated as a breadboard, “SPOTLight,” for the SPOT (spherical primary optical telescope) project. 2 The SPOT architecture represents a new space telescope architecture based on a segmented spherical primary mirror. This type of mirror design, when combined with a high-speed wavefront sensing and control system, allows for continuous phasing of a segmented primary mirror, enabling cost-effective 3-meter class missions (e.g., Midex and Discovery) as well as 30-meter class space telescope designs. 3 As a result of using a replicated

1 Bruce.Dean@nasa.gov
spherical primary, these designs can be manufactured at a substantial savings in both time and cost compared to their corresponding aspheric primary mirror counterparts. A significant optical design challenge is compensating the spherical aberration induced by the spherical primary mirror. However, several potential solutions to this problem exist including multiple mirror corrector systems and refractive designs. Design approaches based on a spherical primary have been discussed by Mertz. A related 25-meter telescope design has been proposed by Moretto. Design studies for robotic assembly and testing of the 30-meter class telescopes in space are also being developed and actively pursued by NASA.

An important component of these designs is a continuously updated wavefront sensing and control system which facilitates missions in low-earth-orbit as well as missions that do not require expensive stable structures and thermal control systems. These features are enabled by a novel wavefront sensing optical design coupled to a compact high-speed portable supercomputing system. The computing system is based on a cluster of 64 DSP's (digital signal processors) operating in parallel under a shared memory architecture. The system is portable and callvable from a laptop computer using a single Ethernet port, reaching a sustained floating-point performance of approximately 100 Giga-Flops/sec. This floating-point performance translates to the algorithms being accelerated by a factor of ×1100 or 3 orders of magnitude beyond what was previously attainable from a state-of-the-art desktop computing system. Additional details of the DSP system and architecture are discussed in Section 5. A related computational approach based on the FPGA (field programmable gate array) is also being developed in parallel as a pathway toward a radiation-hard implementation of the algorithms for closed-loop on-board wavefront sensing and control.

The systems engineering philosophy adopted for 3MV has been an "end-end" approach rather than based on the procurement of expensive high-tolerance components. I.e., a primary goal in building the 3MV is to realize a fully functional wavefront sensing and control prototype within a very limited budget. As such, all software and hardware interfaces for each of the 6 sub-systems illustrated in Figure 1 have been completed and realized in the laboratory at the Goddard Space Flight Center Instrument Development Labs (IDL). The 3MV optical correction and control is based on a radius of curvature test configuration as shown in Figure 1. The mirror control shown in Figure 1 based on a 6MV (6-mirror-vertex) design which is currently being mounted to the testbed (looking forward to Figure 6).

The design consists of six main components: (1) a point source is created at the primary mirror ROC, the beam is split after reflection by the primary and then imaged upon return with a Firewire CCD camera. The images are then sent, (2), directly to a laptop control computer for pre-processing, and then to the DSP unit (3) where the wavefront sensing algorithms process the data into an exit pupil phase estimate. The pupil phase estimate (4) is then sent back to the laptop control computer via Ethernet which is then converted to (5) actuator commands that are sent to control actuators (6) attached to the mirror.

To facilitate the overall system development of 3MV within a reasonable time frame and cost, the mirror tolerances have been considerably relaxed in the area of radius of curvature matching. The mirrors used for 3MV were purchased as off-the-shelf circular mirrors each of 8" diameter from Edmund Scientific with a radius of curvature error specified at 2%. Phase retrieval was performed on each mirror separately to validate the λ/4 peak-valley wavefront error (λ =
633 nm) prior to integration into the testbed (see Figure 2). Zonal errors dominated by concentric trough regions dominate the surface figure of all 3 mirrors. In practice, we found the smallest radius of curvature delta between segments to be within 0.5% as measured using our phase retrieval camera. The largest radius of curvature delta was approximately 1%. The mirrors were then cut into hexagonal segments using the water-jet process as shown in Figure 3.

The entire wavefront sensing and control sequence (Figure 1) including image capture, image processing, upload to DSP, algorithm processing, download from DSP, and subsequent control processing have been accomplished end-end at a 1 Hz rate using a single 512×512 pixel image. Additional performance enhancements are expected with an updated fast-frame camera and second generation DSP hardware based on the TigerSharc TS201 (double clock-speed compared to the TS101). Additional discussion on the DSP architecture and results is given in Section 5. A schematic of the Goddard 64 processor DSP layout is shown in Figure 4.

Figure 2. Phase retrieval was performed on each mirror segment.

Figure 3. Water-Jetting Mirrors into Hexagonal Segments for 3MV Design.

Figure 4. 64 processor DSP layout.
3. Wavefront sensing and control Algorithms

The data collected from the 3MV testbed has been processed using the Goddard Space Flight Center Hybrid Diversity Algorithm (HDA). Details on the HDA algorithm are reported elsewhere in these proceedings. A GUI front-end to the algorithms has been developed to facilitate data collection and Fourier calibration of the images in the lab (see Figure 5). The estimation theory discussion for the approach has been presented in using the Cramer-Rao bound. In addition, optimal diversity defocus values that yield maximum contrast in the spectral power, and therefore improve estimation performance, were presented and identified as a Talbot imaging phenomena in. Algorithm performance on the DSP will be discussed in Section 5. At a later date, the 6MV testbed will be available following the mounting of the ½ meter 6-segment hexagonal mirrors (see Figure 6).

![Figure 5. SpotLight GUI Front-End to Algorithms.](image)

4. Hough Transform for Influence Function Characterization

Automated course alignment of the primary mirror requires that the actuator response to a unit impulse command (influence function) must be characterized and calibrated. For 3MV, these controls correspond to mirror piston, tip, and tilt. The process of calibration of the actuator impulse functions is critical, because it is these impulse functions that are used for correction. The approach presented here is both an efficient and robust estimator that is independent of the quantity and state of other segments.

In general, improved accuracy in the influence function characterization is achieved by increasing the number of measurements and decreasing the actuator step sized used in the calibration. Yet, in doing so, the complexity of the actuator characterization increases for a number of reasons: (a) large step size increases the probability that a segment

![Figure 6. 6-Mirror Vertex (6MV) Primary Mirror](image)
ID is lost by moving off the detector or on top of another image (it is difficult to predict a lost segment ID because it is the actuator itself that is being characterized); (b) if measurements are taken for one actuator at a time, with a small step size, then it becomes difficult to determine which segment is moving. The source of movement can either be from an actuator command, segments that move from jitter of the optical system, or from the apparent motion of the detected images due to lab or atmospheric turbulence (or all of the above). Furthermore, if there are \( N \) segments and \( k \) steps taken for each actuator, one must find the correct \( k \) positions among the \( k \times N \) possible positions. This results in a binomial factor of possible configurations:

\[
\text{combinations} = \binom{k \times N}{N},
\]

which is often impractical for all but the simplest cases.

To addresses these issues, the Hough transform\(^{10,11}\) is used. The Hough transform outlines an efficient method to select a subset of points that form a constrained function from the entire set. For this application, a straight line is used:

\[
y_i = ax_i + b,
\]

but higher order polynomials could also be used as the fitting function. The Hough transform will generate a discrete image of the parameter space, such that the rows and columns of the image correspond to ‘\(a\)’ and ‘\(b\)’, respectively. However, equation (2) has a slope that approaches infinity as the line approaches vertical. Thus, the polar representation of a line is used to avoid an infinite slope:

\[
\rho = y_i \sin \theta + x_i \cos \theta.
\]

From the parameter space image, the maximum value for \( \rho \) and \( \theta \) are used to determine the corresponding lines. The actuator basis function is determined by normalizing this line for the \( k \) steps and the step size.

The Hough transform was applied to course-align the segmented primary mirror. Two of the segments have 3 New-Focus piezo-electric actuators (6 total) for piston, tip, and tilt, adjustment and the third segment is fixed. As illustrated in the example of Figure 7, the images from each segment must eventually be co-aligned as shown in the right-most figure. To achieve the required sub-pixel level command accuracy in tip and tilt, each of the actuators must be fully characterized before moving from the constellation configuration on the left to the aligned state on the right.

The calibration steps are illustrated in Figure 8 where a single segment is characterized by sending 20 step commands to one actuator. The image is shown in Figure 8 (a), illustrating (for clarity) the superposition of images resulting from each actuator command. After collection of the data, the location of the image corresponding to each segment is needed. To obtain this information from the image data, a series of image-processing steps are performed: each image is first flat-fielded and then background subtracted. Next, the images are processed via thresholding, eroded, and then finally the correlated components are extracted. The weighted center of each region of interest is used as the segmented image location. The position data is used for the Hough transform to produce Figure 8 (b) and represents the entire parameter space of \( \rho \) and \( \theta \). Figure 8 (c) is a zoomed version of (a) in the area of interest.

The maximum value in the parameter space is the \( \hat{\rho} \) and \( \hat{\theta} \) that correspond to the response function for the given actuator. Finally, this function must then be normalized for the given number of steps and the step size.
Next we demonstrate the robust nature of the Hough transform, compared with a least squares estimator, in the presence of a non-stationary noise due to a neighboring segment undergoing motion as a result of lab turbulence and jitter. In this case the optical system is the same, yet one of the segments is positioned along the path of the segment being calibrated. Even though one of the segments is placed along the path, and could potentially interfere with the ongoing segment calibration, as shown in Figure 9 (b), the Hough transform is able to efficiently filter the extraneous data points. Determining the maximum of the parameter space is then no different for Figure 9 (c) than it is for the data of Figure 8 (b).

Figure 9 compares the accuracy of the Hough transform to a linear least squares estimator to the data set presented in Figure 9 (a). The normal equations do provide the minimal Euclidean norm of the residual, yet for this derivation it is assumed that all the data points are a set of the characteristic function, even the unintended extraneous data points. An important point to make is that because of the robust nature of the Hough transform approach; the calibration of multiple segments can be performed simultaneously, thus reducing the system commissioning time by the number of segments in the system.

5. High-Speed WFS &c Using a Cluster of DSP’s

Image-based wavefront sensing has a significantly higher computational requirement than analog optical wavefront sensing. Additionally, the computational requirements are proportional to the accuracy of the wavefront measurement. E.g., course alignment and course phasing use less complicated models to determine the necessary corrections; similarly, fine phasing requires more complex algorithms. Without sacrificing accuracy, it is optimal to minimize the time between image acquisition and correction. Additionally, to increase the resolution of the wavefront requires an increase in the data set size, and thus computational complexity.

The HDA requires numerous iterations from the image plane to the pupil plane via the Fourier transform pair, see Figure 11. Current desktop and server line general-purpose central processing units (CPU) have been optimized for
performing multiple tasks and use principles of data locality to increase performance, yet for performing large 2-D FFT’s, performance dramatically decreases. For example, a current general-purpose CPU can take several seconds for a double precision 2-D FFT of size 2048x2048. This is mainly due to the fact that the memory architecture is not optimized for such large data sets. To perform the numerous large 2-D FFT’s efficiently, several application specific architectures have been developed and are discussed further below. Many 1-D and 2-D FFT algorithms exist, yet for each element of the output, the algorithms require access to every element of the input. Thus, as one naive approach to parallelization, an image cannot be divided into sub-components that are processed completely independent. Typically, the 2-D FFT is computed as a series of 1-D FFT’s. Thus, a 1-D FFT is performed on each row; then, the 1-D FFT is performed on each column of the result.

The process of performing wavefront sensing on N Diversity-Defocused images is highly parallel for each image, and as such, past approaches to increase performance have used 1 to N general-purpose CPUs as a cluster. Consequently, this provides a maximum factor of N improvement, while having the negative effect of increasing power requirements, footprint, and cooling requirements by N. The primary solution is to provide N application specific, highly optimized computational cores. An earlier iterative transform algorithm implementation using the Analog Devices Hammerhead DSP was given.

The first solution for the computational core presented is a single Analog Devices TS101 Digital Signal Processor (DSP). The single DSP has many advantages over a single CPU: the instruction set architecture (ISA) is specific for digital signal processing, higher performance/watt, and a scratch-pad cache are a few of the features that separate the two. For our tests, the 1-D FFT of size 512 and 2048 on the TS101 at 250 MHz is 2.5 and 3.9 times faster, respectively, than a Pentium 4 at 2.4 GHz. Additionally, the scratch-pad cache allows one to optimize the communication requirements for the 2-D FFT sub-component processing. For example, if the 2-D FFT is performed as a series of 1-D FFT’s, the next row or column can be moved from the slower SDRAM to the faster, but smaller, scratch-pad cache SSRAM. Using this method, the single DSP is \( \times 7.6 \) and \( \times 10.2 \) times faster for the 512x512 and 2048x2048 case. This results in an overall improvement of \( N \times 10.2 \), where \( N \) is the number of diversity-defocus images. This is sufficient in certain applications such as small simulations or stable laboratory environments.

For other applications, more computational performance is needed, and thus multiple DSP’s must be used. The fundamental problem to be solved is scaling the system architecture to minimize bottlenecks. The first architecture considered is shown in Figure 12 and consists of a cluster of 4 DSP’s with a shared bus to memory, SDRAM. The communication requirements of the all-to-all transmission of the 2-D FFT sub-components showed diminishing returns on performance as the number of DSP’s were increased. The improvement from 1 DSP to 4 DSP’s was only by a factor of \( \times 2.1 \), with 4 DSP’s only increasing performance by 10% over 3 DSP’s. To expand beyond 4 DSP’s, various grid architectures have been explored.

The TS101 processor has 4 bi-directional link ports that allow the direct connection between two chips. Other processors offer similar communication ports. The 2-D FFT sub-components requirements, such as all-to-all communication, imply low-diameter architectures. Hence, the maximum of the shortest-paths between any two nodes is minimized. For example, referring to Figure 13, the node’s unique identifier and shortest path from 0 are listed. For graph (a), a cube, the shortest-path between node 0 and all other nodes is 3 or less, but for graph (b), a
1-möbius cube, the *shortest-path* between any two nodes is 2 or less.\textsuperscript{19, 20} For graph (b), the 2-D FFT is 17% faster than graph (a).

For the 16 to 32 DSP's per image, only the sub-clustered routing architecture was explored which is shown in Figure 14 (c). For processing, i.e. 1-D FFT's and other mathematical routines, the improvement between 16 and 32 DSP's is nearly linear. Thus, the computational routines provided a factor of \(\times 2\) improvement. For most parallel computing applications, communication is the bottleneck preventing linear scalability. Contrary to bus architectures, there was an improvement for increasing the number of nodes. This is due to the fact that increasing the number of DSP's also increases the total number of link port channels in the system, and hence, we are increasing the overall effective bandwidth. For the all-to-all communication, the 32 DSP's architecture of Figure 14 (c), was \(\times 1.2\) times faster than the 16 DSP's architecture, Figure 14 (b), for an overall improvement for the 2-D FFT of \(\times 1.7\). Table 1 summarizes the actual timing for several relevant configurations (it is of interest to note that the 4-image 512\times 512 case of Table 1 is the configuration baselined for the James-Webb-Space-Telescope wavefront sensing). The 16 and 32 DSP's per image use the architectures shown in Figure 14 (b) and (c) respectively. Furthermore, there are 4 of these clusters operating on each image. For the algorithm listed in Figure 11, there were 20 outer-loops and 5 inner-loops for a total of 100 iterations; increasing or decreasing the total number of iterations will scale the timing accordingly by that amount.
Table 1. Timing of the iterative-transform algorithm as shown in Figure 11; all times listed in seconds.

<table>
<thead>
<tr>
<th>Image Size</th>
<th>Pentium 4, 2.4 GHz</th>
<th>16 DSP per image</th>
<th>32 DSP per image</th>
</tr>
</thead>
<tbody>
<tr>
<td>64×64</td>
<td>20.49</td>
<td>0.0561</td>
<td>0.0332</td>
</tr>
<tr>
<td>512×512</td>
<td>219.2</td>
<td>0.491</td>
<td>0.288</td>
</tr>
<tr>
<td>2048×2048</td>
<td>8890</td>
<td>19.6</td>
<td>12.2</td>
</tr>
</tbody>
</table>

In addition to space optics wavefront sensing and control, additional application areas that can benefit from the application of digital signal processors include ground-based wavefront sensing, telescope image processing, laboratory optical processing, system design and tolerancing, Monte-Carlo simulations, and finite element modeling. High performance DSP’s do not provide the easiest path toward a radiation-hard implementation such as that required for autonomous space optics control. However, the FPGA (field programmable gate array) or ASIC (applications specific integrated circuit) can lead to a high performance solution in a radiation saturated environment. Currently, NASA Goddard Space Flight Center is exploring several solutions including reconfigurable computing to develop systems of FPGA’s to meet the requirements of high-speed space-based image-processing as well as wavefront sensing and control.

6. Acknowledgements

The authors thank Scott Acton and Bruce Jacobs for discussion and support of this work. We also thank Damon Bradley and Semion Kiznher for early collaboration and many helpful discussions. The machine shop and fabrication support of Greg Woytko and Dave Clark are also gratefully acknowledged.

7. References


J. Budinoff, D. Pfenning, L. Feinberg, J. Hagopian, B. H. Dean, J. Howard, A. Morell, A Long-Stoke Nanometric Actuator for the SPOT Segmented Telescope Demonstration


15. M. Frigo and S. G. Johnson. FFTW: An Adaptive Software Architecture for the FFT.


