Closed loop, DM diversity-based, wavefront correction algorithm for high contrast imaging systems

Amir Give'on[1*], **Ruslan Belikov**[2], **Stuart Shaklan**[1], **Jeremy Kasdin**[2]

[1] *Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA*
[2] *Princeton University, Princeton, NJ*

[*] *amir.giveon@jpl.nasa.gov*

Abstract: High contrast imaging from space relies on coronagraphs to limit diffraction and a wavefront control systems to compensate for imperfections in both the telescope optics and the coronagraph. The extreme contrast required (up to 10^{-10} for terrestrial planets) puts severe requirements on the wavefront control system, as the achievable contrast is limited by the quality of the wavefront. This paper presents a general closed loop correction algorithm for high contrast imaging coronagraphs by minimizing the energy in a predefined region in the image where terrestrial planets could be found. The estimation part of the algorithm reconstructs the complex field in the image plane using phase diversity caused by the deformable mirror. This method has been shown to achieve faster and better correction than classical speckle nulling.

OCIS codes: (010.1080) Adaptive Optics.

References and links

1. Traub, W. A., ed. *Proceedings of Coronagraph Workshop 2006*. JPL Publication 07-02.
2. Tyson R.K. *Introduction to Adaptive Optics*. SPIE Press, 2000.
3. Poyneer L.A. and Macintosh B. "Spatially filtered wave-front sensor for high-order adaptive optics," J. Opt. Soc. Am. A **21**(5), 810-819 (2004).
4. Give'on A., Kasdin N. J., Vanderbei R. J., and Avitzour Y. "On representing and correcting wavefront errors in high-contrast imaging systems," J. Opt. Soc. Am. A **23** (2006).
5. Trauger J. T. , Burrows C. , Gordon B. et al, "Coronagraph contrast demonstration with the high-contrast imagaing testbed," Proc. SPIE **5487**, 1330-1336 (2004).
6. Belikov R., Give'on A., Trauger J.T., Carr M., Kasdin N.J., Vanderbei R.J., Shi F., Balasubramanian K., Kuhnert A., "Toward 10^{10} contrast for terrestrial exoplanet detection: demonstration of wavefront correction in a shaped-pupil coronagraph," Proc. SPIE **6265**, pp. 626-518 (2006).
7. Borde P. J. and Traub W. A., "High-contrast imaging from space: Speckle nulling in a low aberration regime," Astrophys. J. **638** (2006).
8. Malbet F., Yu J.W. and Shao M, "High dynamic range imaging using a deformable mirror for space coronography", PASP, **107**, pp. 386 (1995).
9. Kasdin N. J., Vanderbei R. J., Spergel D. N., and Littman M. G. "Extrasolar Planet Finding via Optimal Apodized-Pupil and Shaped-Pupil Coronagraphs" Astrophys. J. **582**(2), 1147-1161, (2003).
10. Kuchner M. J., Crepp J., Ge J. "Eighth-Order Image Masks for Terrestrial Planet Finding," Astrophys. J. **628**(1), 466-473 (2005).
11. Kasdin N.J., Belikov R., Beall J., Vanderbei R.J., and Littman M.G., Carr M., and Give'on A., "Shaped pupil coronagraphs for planet finding: optimization, manufacturing, and experimental results," Proc. SPIE **5905**, 128-136 (2005).

#84970 - $15.00 USD Received 9 Jul 2007; revised 6 Sep 2007; accepted 9 Sep 2007; published 13 Sep 2007

1. Introduction

The problem of direct detection of exoplanets requires high-contrast imaging of a dim point source (the planet) appearing adjacent to a much brighter source (the star). Many techniques have arisen, using various forms of coronagraphy, to achieve the very high contrast needed (up to 10^{-10}), to image extrasolar planets from space [1]. In this note we describe an algorithm using a single DM for estimating and correcting static wavefront error in a coronagraphic imaging system. Our approach works across the full spectrum of high-performance coronagraphs.

The performance of conventional adaptive optics (AO) typically has been limited by the accuracy of the wavefront phase estimation (known as phase reconstruction) and the ability of the deformable mirror (DM) to achieve the arbitrary shapes required [2]. As improvements have been made in these areas, the contrasts achieved by conventional phase conjugation have still been limited [3]. In [4], we demonstrated that this is due to intermodulation products of the high frequency content of the DM, call frequency folding. Frequency folding, not to be confused with aliasing due to wavefront sampling, occurs even with perfect estimation of the wavefront and a DM limited only by the highest spatial frequency it can achieve.

The accepted solution is to perform measurements on the final science camera in conjunction with, for example, the speckle nulling algorithm [5, 6]. While proven effective, classical speckle nulling requires many iterations, potentially imposing unreasonable stability requirements on the system. The algorithm we propose here is substantially more efficient. The correction part of the complete algorithm is based on a method developed by Bordè et al. [7] for correcting the field in a classical Lyot coronagraph which in turn has its origin in the linear solution of the 'dark hole algorithm' developed by Malbet et al. [8] . We have generalized these approaches by incorporating the full, nonlinear expression for the aberrations, including scattered light and the nominal diffraction of the system, and by formulating the correction in a manner applicable to a wide variety of high-contrast imaging systems.

2. Energy Minimization: a closed-loop correction algorithm for high contrast imaging systems

In this section we briefly describe the details of the correction algorithm which we call Energy Minimization. These are divided into two pieces: a correction stage where the DM actuators are set based on the measured field and an estimation, or reconstruction, stage where the real and imaginary parts of the complex electric field at the DM is computed from multiple measurements in the image plane. It is important to note that while we present a complete closed-loop algorithm, these two steps are functionally independent and could be implemented with other companion approaches.

2.1. The correction stage

Let the response of an optical system, operating in monochromatic light, be modeled by the linear transformation C between the electric field at the deformable mirror (DM) plane, E_0, and the electric field at the science camera plane, E_f, where the measurements will take place,

$$E_f = C\{E_0\} \tag{1}$$

We are particularly interested in using shaped pupil coronagraphs for high-contrast [9]. There, $C\{E_0\} = \mathscr{F}\{SE_0\}$, where $\mathscr{F}$ represents the Fourier transform and S represents the shaped pupil function (the DM is assumed to be at a conjugate plane to the shaped pupil). Similarly, for the band-limited Lyot coronagraph [10], $C\{E_0\} = (\mathscr{F}\{SE_0\}M) \otimes \mathscr{F}\{L\}$, where S is the entrance pupil of the coronagraph, M is the image plane mask, L is the Lyot pupil and $\otimes$ represents the convolution (the DM is assumed to be at the same plane as the entrance pupil). Since all real

#84970 - $15.00 USD Received 9 Jul 2007; revised 6 Sep 2007; accepted 9 Sep 2007; published 13 Sep 2007

optical systems have aberrations induced by errors in the optics, the input field in Eq. 1 can be modeled,

$$E_0 = Ae^{\alpha+i\beta}e^{i\psi}, \tag{2}$$

where A is the un-aberrated, ideal, electric field, α and β represent the amplitude and phase aberrations, respectively, and ψ represents the phase difference caused by the deformation of the DM surface.

Using Eqs. 1 and 2, the electric field in the science camera field for a system with phase and amplitude aberrations and a DM is thus given by,

$$\begin{aligned} E_f &= C\left\{Ae^{\alpha+i\beta}e^{i\psi}\right\} \approx C\left\{A\left(1+\Phi\right)\left(1+i\psi\right)\right\} \\ &= C\left\{A\left(1+\Phi+i\psi+i\psi\Phi\right)\right\} \approx C\left\{A\left(1+\Phi+i\psi\right)\right\} \\ &= C\left\{Ae^{\alpha+i\beta}\right\} + iC\left\{A\psi\right\} \end{aligned} \tag{3}$$

where we used the first order Taylor series for the effect of the DM, $\Phi = e^{\alpha+i\beta} - 1$ was defined for expansion purposes and the cross terms between ψ and Φ is assumed to be negligible.

The total energy in the dark zone, $\mathscr{E}$, is given by the inner product of the electric field, $\mathscr{E} = \left\langle E_f, E_f \right\rangle$, or,

$$\mathscr{E} = \left\langle C\left\{Ae^{\alpha+i\beta}\right\} + iC\left\{A\psi\right\}, C\left\{Ae^{\alpha+i\beta}\right\} + iC\left\{A\psi\right\} \right\rangle \tag{4}$$

where $\langle f,g\rangle = \int\int f^* g\, d\xi\, d\eta$ and the asterisks represents a complex conjugate.

Any optimal correction algorithm requires some model for the DM surface height, ψ as a function of the actuator commands. Here, we assume an *influence function model*,

$$\psi = \sum_{k=1}^{N_{dm}} \sum_{l=1}^{N_{dm}} a_{k,l} f_{k,l}, \tag{5}$$

where the DM consists of an array of $N_{dm} \times N_{dm}$ actuators, $a_{k,l}$ is the kl^{th} coefficient (actuator command), and $f_{k,l}$ is the DM's influence function, centered at the location of the kl^{th} actuator. The influence function is defined as the surface profile of the DM when one actuator is commanded (note that this model assumes no coupling between actuators and that superposition holds).

We choose as our correction criteria to minimize the energy in a region of the image plane. We then check to verify that this minimization created adequate contrast. The condition for minimum energy is that for all k and l, $\frac{\partial \mathscr{E}}{\partial a_{k,l}} = 0$, or,

$$\sum_{m=1}^{N_{dm}} \sum_{n=1}^{N_{dm}} a_{m,n} \Re\left\{\left\langle C\left\{Af_{k,l}\right\}, C\left\{Af_{m,n}\right\}\right\rangle\right\} = -\Im\left\{\left\langle C\left\{Af_{k,l}\right\}, C\left\{Ae^{\alpha+i\beta}\right\}\right\rangle\right\}, \tag{6}$$

where $\Re$ and $\Im$ represent the real and imaginary parts, respectively, and we used the relations $C\left\{A\psi\right\} = \sum_{k=1}^{N_{dm}} \sum_{l=1}^{N_{dm}} a_{k,l} C\left\{Af_{k,l}\right\}$, $\frac{\partial C\{A\psi\}}{\partial a_{k,l}} = C\left\{Af_{k,l}\right\}$ and the linearity of C. The criteria for minimum energy in Eq. 6 can be written in matrix form as

$$\begin{bmatrix} G_{1,1} & \cdots & G_{1,N_{dm}^2} \\ \vdots & \ddots & \vdots \\ G_{N_{dm}^2,1} & \cdots & G_{N_{dm}^2,N_{dm}^2} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_{N_{dm}^2} \end{bmatrix} = \begin{bmatrix} H_1 \\ \vdots \\ H_{N_{dm}^2} \end{bmatrix}, \tag{7}$$

#84970 - $15.00 USD Received 9 Jul 2007; revised 6 Sep 2007; accepted 9 Sep 2007; published 13 Sep 2007

where $G_{r,q} = \Re\left\{\left\langle C\{Af_r\}, C\{Af_q\}\right\rangle\right\}$ and $H_r = -\Im\left\{\left\langle C\{Af_r\}, C\left\{Ae^{\alpha+i\beta}\right\}\right\rangle\right\}$ and $r,q = 1,2,\cdots,N_{dm}^2$. Therefore, in order to find the coefficients for the DM configuration that minimizes the energy in the dark zone, we need to estimate $C\left\{Ae^{\alpha+i\beta}\right\}$ which is essentially the complex valued electric field in the science camera plane of an uncorrected, aberrated system.

2.2. The reconstruction stage

For each DM configuration ψ_k, following similar steps as the ones leading to equation 3, the electric field in the final science camera plane is given by

$$E_k \approx C\left\{Ae^{\alpha+i\beta}\right\} + C\{A\Psi_k\} \tag{8}$$

where $\Psi_k = e^{i\psi_k} - 1$ and we assume the cross term $C\{A\Phi\Psi_k\}$ is negligible.

The intensity of light in the final science camera plane for any DM configuration, ψ_k, is therefore approximately given by,

$$I_k = \left|C\left\{Ae^{\alpha+i\beta}\right\} + C\{A\Psi_k\}\right|^2. \tag{9}$$

Note that this is an image-plane measurement, in contrast to conventional adaptive optics, where the electric field at the pupil is measured directly using a wavefront sensor such as a Shack-Hartmann (SH). While the SH sensor allows direct field estimation, it requires the use of additional optics not in the science path. This introduces uncorrectable non-common path error larger than the desired contrast. Our image-plane measurements with the science camera eliminate this problem. However, because we are measuring the intensity of the field, one image plane measurement is inadequate for estimating the complex electric field. We therefore take multiple images at different DM settings to estimate the field. Let the first such image, I_0, be taken with $\psi = 0$. The intensity of light in the image plane is then

$$I_0 = \left|C\left\{Ae^{\alpha+i\beta}\right\}\right|^2. \tag{10}$$

The remaining images are taken with different DM configurations, ψ_k. Combining equations 9 and 10, for each image taken gives,

$$I_k - I_0 - |C\{A\Psi_k\}|^2 = \left(C\left\{Ae^{\alpha+i\beta}\right\}\right)^* C\{A\Psi_k\} + C\left\{Ae^{\alpha+i\beta}\right\}(C\{A\Psi_k\})^*. \tag{11}$$

Using multiple DM configurations allows us to solve for $C\left\{Ae^{\alpha+i\beta}\right\}$ via the following matrix equation,

$$\begin{bmatrix} I_1 - I_0 - |C\{A\Psi_1\}|^2 \\ \vdots \\ I_k - I_0 - |C\{A\Psi_k\}|^2 \end{bmatrix} = 2 \begin{bmatrix} \Re\{C\{A\Psi_1\}\} & \Im\{C\{A\Psi_1\}\} \\ \vdots & \vdots \\ \Re\{C\{A\Psi_k\}\} & \Im\{C\{A\Psi_k\}\} \end{bmatrix} \begin{bmatrix} \Re\left\{C\left\{Ae^{\alpha+i\beta}\right\}\right\} \\ \Im\left\{C\left\{Ae^{\alpha+i\beta}\right\}\right\} \end{bmatrix} \tag{12}$$

implying that

$$\begin{bmatrix} \Re\left\{C\left\{Ae^{\alpha+i\beta}\right\}\right\} \\ \Im\left\{C\left\{Ae^{\alpha+i\beta}\right\}\right\} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \Re\{C\{A\Psi_1\}\} & \Im\{C\{A\Psi_1\}\} \\ \vdots & \vdots \\ \Re\{C\{A\Psi_k\}\} & \Im\{C\{A\Psi_k\}\} \end{bmatrix}^\dagger \begin{bmatrix} I_1 - I_0 - |C\{A\Psi_1\}|^2 \\ \vdots \\ I_k - I_0 - |C\{A\Psi_k\}|^2 \end{bmatrix} \tag{13}$$

#84970 - $15.00 USD Received 9 Jul 2007; revised 6 Sep 2007; accepted 9 Sep 2007; published 13 Sep 2007

where † represents the pseudo inverse (typically solved by a least squares-based method such as the Singular Value Decomposition). The condition for the existence of a solution at a given point is that there exists m and n, such that

$$\Re\{C\{A\Psi_m\}\}\Im\{C\{A\Psi_n\}\} - \Re\{C\{A\Psi_m\}\}\Im\{C\{A\Psi_n\}\} \neq 0 \tag{14}$$

The physical meaning of this equation is that there must be at least 2 DM settings that create 2 different electrical fields for every CCD pixel in the image plane that mix with the original field. Only then will there be enough diversity to reconstruct the field at each CCD pixel.

3. Experimental results

In this section, we briefly demonstrate the method described in the previous section on a real shaped pupil coronagraph at the Princeton Terrestrial Planet Finder Laboratory [11], and compare the performance to classical speckle nulling.

Our lab is designed to take an image of an artificial star with a mockup of a coronagraphic telescope and measure the resulting contrast. We used a 632nm He-Ne laser as our artificial star, and a 32 x 32 Boston Micromachines deformable mirror for correction. Our optics had $\lambda/20$ surface figure and were 6" in diameter (only a small portion of which was used).

We iteratively applied 2 types of correction: Classical Speckle Nulling and Energy Minimization. (See [6] for a brief description of our implementation of Classical Speckle Nulling.) For the case of Energy Minimization, we used two DM configurations ($k = 1,2$), so that the total number of images per iteration is 3. Each DM configuration was a sum of cosine ripples, designed to create a grid of overlapping PSFs in the image plane that completely covers the dark zone. The two DM configurations were the same except they were in quadrature phase (in which case they satisfy Eq. 14). Figure 1 shows the experimental results.

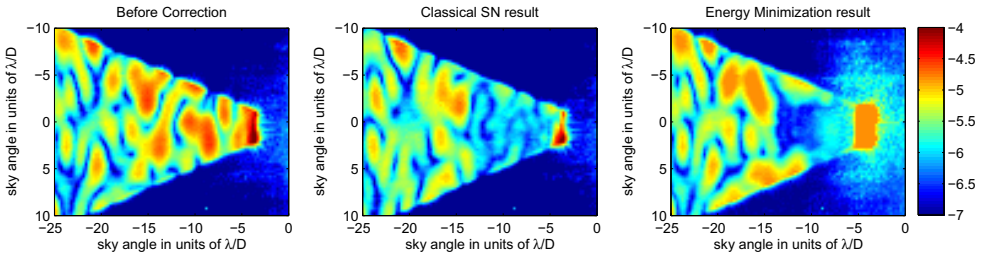

Fig. 1. CCD images. *Left*: Before correction. Average contrast across $(5-14\lambda/D)$ is about 10^{-5}. *Center*: After Classical Speckle Nulling correction bottoms out. Contrast across $(6-14\ \lambda/D)$ is about 10^{-6}. *Right*: After Energy Minimization bottoms out. Contrast across $(6-14\ \lambda/D)$ is about 6×10^{-7}.

A dark hole was generated in the region from 5 to 14 λ/D for the first 17 iterations of energy minimization and the first 50 iteration of speckle nulling, and then switched to 6 to 14 λ/D in the horizontal direction and -2 to 2 λ/D in the vertical. (We use sky angle as image plane coordinates, in units of $1\lambda/D$, which is the angular diffraction limit of a telescope of aperture D and wavelength λ). Figure 1 shows the images taken before correction, after Classical Speckle Nulling correction bottomed out, and after Energy Minimization bottomed out. The optical axis in all images is close to the right edge of the image and the images show the left dark zone out to about $25\lambda/D$, together with a portion of the image plane stop which blocks the core of the star. The colormap is logarithmic, with contrast levels shown in the colorbar to the right. It is evident that in both cases of correction, a dark hole appears with improved contrast.

#84970 - $15.00 USD Received 9 Jul 2007; revised 6 Sep 2007; accepted 9 Sep 2007; published 13 Sep 2007

Figure 2 shows the average contrast in the dark hole as a function of the total number of images taken for the two algorithms. The black and red curves correspond, respectively, to classical speckle nulling (which used 7 images per iteration) and energy minimization (which required 3 images per iteration). It is evident that, at least in this case, the energy minimization algorithm (a) improved contrast faster than classical speckle nulling (computational times were not significant) and (b) achieved a modestly deeper null. (Though it is not shown on the graph, speckle nulling asymptotes around 10^{-6} on the Princeton testbed.) It should be noted, however, that classical speckle nulling was observed to be much less sensitive to noise and errors in the model than energy minimization. The limiting factor in our classical speckle nulling experiment

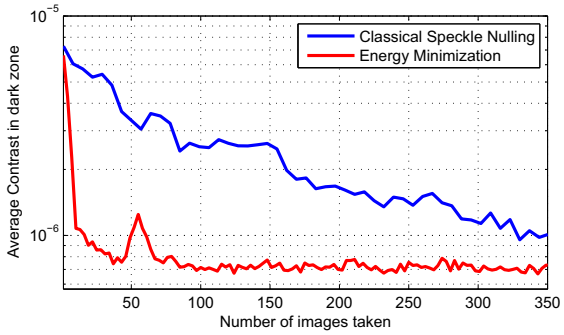

Fig. 2. A log-log plot of average contrast in the control region vs. number of images taken by the estimation and control algorithms.

is, based on simulations, believed to be the inefficiency of Classical Speckle Nulling to correct for spatially small speckles, which arise at around the 10^{-6} level. Energy minimization does not suffer from this limitation and therefore achieves a deeper null. However, it also hits a limit at roughly 6×10^{-7} in our lab. The limiting factor is the haze around $4\lambda/D$ that is apparent on the rightmost image in Figure 1. This residual halo behaves as incoherent light, i.e. adding in intensity rather than amplitude (it is not known yet what it actually is) and thus is not eliminated by the DM correction. One of the appealing features of the subtraction based estimation formula in Eq. 11 is that any incoherent light common to each measurement is eliminated, resulting in an estimate of the desired coherent portion only. The contrast of this residual coherent light estimate is roughly 10^{-7}, averaged across the control region. At this level, quantization of the DM voltage signal is believed to be the limiting factor. Simulations further show that with no quantization (and no other non-fundamental limiting factors such as incoherent light) the contrast reaches 10^{-10} with energy minimization.

4. Summary and conclusion

We have presented and demonstrated in monochromatic light a fast wavefront estimation and correction algorithm. In only a few iterations, the contrast improves by a factor of 10, and the contrast of the coherent light estimate in the image improves by a factor of 100. Limiting factors are being explored and are believed to be in hardware rather than the algorithm. Even though we have only considered monochromatic light, it is possible to generalize this algorithm for multi-spectral operation, though this is beyond the scope of the present paper.

Acknowledgments

This work was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

#84970 - $15.00 USD Received 9 Jul 2007; revised 6 Sep 2007; accepted 9 Sep 2007; published 13 Sep 2007