Local Minima Analysis of Phase Diverse Phase Retrieval Using Maximum Likelihood

David R. Gerwe, Michael M. Johnson
Boeing Directed Energy Systems, Los Angeles, CA
Brandoch Calef
Boeing Laser Technical Services, Kihei, HI

Abstract

Phase diversity techniques infer wavefront phase aberrations of an imaging system directly from the image data by processing of a set of two or more images with deliberately induced additional aberrations (usually defocus). This is a non-linear inverse problem that is typically solved by iteratively determining the aberration that maximizes an objective function or by iteratively applying a set of constraints. The robustness of these techniques can be severely degraded by local minima that start to appear and increase in density with increasing wavefront aberration strength. Local minima severely limit the dynamic range of wavefront aberration levels for which diversity techniques can serve as a reliable wavefront sensing method using the processing approaches currently in practice. This paper explores mechanisms behind the local minima and characterizes their density and spacing as dependent on SNR and wavefront aberration level. The performance of a global optimization method is compared to a single gradient search for robust WF estimation accuracy.

1. Introduction

The phase diversity concept, introduced by Gonsalves[1], measures two images of the same object through a common wavefront aberration that differ only in introduction of a deliberate change in the imaging system's focus setting. Using a variety of algorithmic methods, an estimate of the pristine object and the wavefront aberration are inferred. The key enabling ingredient is the "diversity" of measurements with controlled changes to the optical system that induce corresponding changes in the optical transfer function (OTF) and that are interrelated to the common unknown component of the wavefront aberration. The extra measurements provide constraints that improve the ability to determine a unique wavefront solution. Several variations on this concept have been proposed including using multiple images with a sequence of defocus levels, diversity aberrations other than defocus, sequence of diversity measurements with time-evolving aberrations and common object (referred to as phase diverse speckle), diversity in the spatial distribution of the aperture transmission function, and differences in wavelength for each image (wavelength diversity)[2-4]. When the object is known and need not be estimated, for example a known point source target such as an isolated star, PD reduces to phase diverse phase retrieval (PDPR).

Phase diversity methods offer many attractive advantages for wavefront sensing. They can work with both point and extended objects and with broadband illumination. The hardware design is simple, requiring only addition of a second defocused imaging channel in addition to the primary mission sensor. Calibration requires just models or measurements of the pupil transmission function, spectral response, magnification, and defocus level, plus standard correction for focal plane array non-uniformity and bad pixels. Because of these unique features, phase diversity has received strong attention for many applications most notably compensation for atmospheric turbulence for ground-to-space imaging applications, and for measurement and correction of the misalignments and deformations of the mirror petals of segmented aperture telescopes[2-28].

A current major deficit of the general class of PD techniques is their limited dynamic range. The most prominent methods for processing the data to infer the wavefront are based on iterative search techniques that seek the wavefront that maximizes agreement with the measured data in terms of a mathematical model of the optical system and the measurement noise properties in terms of a cost function[5,6,12,25]. Local minima in the cost function form traps leading the iterative techniques to converge to the wrong answer[18,23]. The subject of this paper is to better

understand the nature of the local minima in terms of their relation to the image formation physics, and their quantity and distribution in the wavefront parameter space as related to SNR and true wavefront aberration level. The paper is organized as follows. Section 2 presents the approach and describes methods of simulation and analysis. Section 3 discusses the local minima phenomenology and cost function topology. The final section summarizes the results of the paper and discusses areas for future research.

2. Analysis Approach

This research focuses on the restricted case of phase diversity phase retrieval (PDPR) in which the object is known to be a point source. This eliminates additional local minima problems associated with unknown object parameters so as to focus solely on local minima associated with the relation of the wavefront (WF) to the in-focus and de-focus channel PSFs. We can expect the phenomenology and trends to be similar for the more general case of PD. Characterization of the local minima problem was performed by application of PDPR processing algorithm to numerically simulated phase diversity image data generated over a range of strengths of the true aberration, SNR levels, and random realizations of the true WF structure and of sensor noise. The nominal imaging system model used is as follows:

- Monochromatic light
- 80x80 pixel FPA
- The focal plane array (FPA) pixels are spaced to provide Nyquist sampling of the PSF associated with 1m diameter aperture:
  - $D = 1m$ = aperture diameter
  - $IFOV = 1 / (2 * D)$
  - $Q = (\lambda/D) / IFOV = 2$
  - Fill factor = 100%
- Telescope clear aperture comprised of 6 hexagonal segments
  - Nominal model assumes segments are perfectly optically aligned (e.g. no contribution to wavefront aberration)
- In-focus + single de-focus diversity channel
  - 1 λ rms defocus
  - Boresight of all channels perfectly aligned
- Noise consisted of mix of Photon shot noise (Poisson) plus Gaussian read noise of 10 e- rms
- Exposure time scaled to achieve desired PSNR (as described below)

SNR is defined as peak-SNR (PSNR) equal to the signal level for the brightest pixel in the image divided by the noise standard deviation (shot plus read noise). Note that two WFs with same rms aberration level may produce PSFs with different peak levels, causing differences in PSNR, though the total signal (integrated over image) is the same. For each individual wavefront draw, the exposure time and thus integrated signal was scaled to achieve the desired PSNR. Random WF aberrations were generated as follows:

- WFs were represented by Zernike modes 2-36 (e.g. through 8th order with zero piston)
- Given a desired rms aberation level:
  - WF coefficients were randomly generated such that covariances matched Kolmogorov statistics[27]
  - Coefficients are scaled to produce specific desired wavefront aberration strength
  - This procedure is repeated until a coefficient vector is drawn such that all coefficients are less than 2.5 times their variance under Kolmogorov statistics at the specified rms aberration strength (further constraining the space of possible wavefronts)

Figure 1 illustrates the clear aperture, an example wavefront, and the diffraction limited PSF. Figure 2 shows examples of the in-focus and de-focus images for different WF aberration and SNR levels.
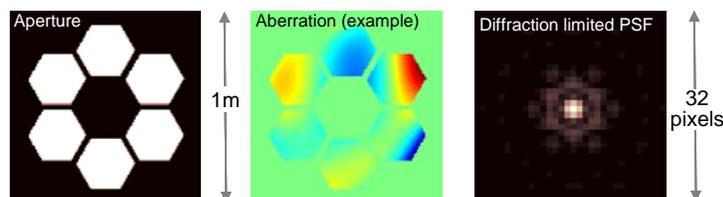

Figure 1. (left) Clear aperture function, (center) example random WF draw using Zernike modal decomposition with coefficient covariances matched to Kolmogorov turbulence statistics, and

(right) the diffraction limited PSF (zero WF aberration) over a 32×32 sub-region of the FPA region.
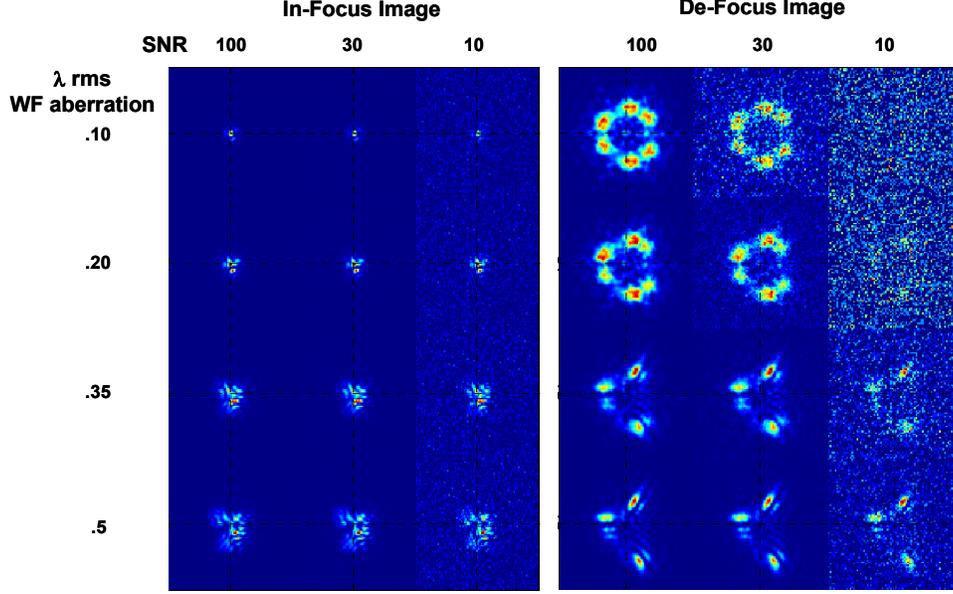

Figure 2. Example in-focus and de-focus images at different SNR and WF aberration levels (all use scaled version of same WF aberration)

A few variations from these nominal settings were explored:
- Defocus of .5 λ rms (deviation from nominal defocus = 1.0 λ rms)
- A 4-channel diversity sensor with defocus levels of = { -0.5, 0.0, 0.5, 1.0 } λ rms
- Panchromatic imaging
  - Modeled as sum of 5 monochromatic images
  - Wavelengths span .4-.6 μm ($D\lambda/\lambda_o = .77$)
- Wavefront aberration consists of piston/tip/tilt for each of 6 primary mirror segments (18 modes total)

The PDPR algorithm is implemented as described in C. Luna [26] and D. R. Gerwe [28], with the modification that the object is considered to be a point source of known magnitude centered on the FPA for both channels and is not estimated. The L-BFGS-B quasi-Newton algorithm is used to conduct the gradient-based descent[29]. The following Log-Likelihood cost function was used that approximately matches the probability distribution function (PDF) for a combination of photon shot noise with Poisson statistics and additive read noise with Gaussian statistics [30-32].

$$LL(\bar{\alpha}) = \sum_{c,\bar{x}} \left\{ \begin{array}{l} + \left[d_{c,\bar{x}} + \sigma_{RN}^2 \right] \log \left(\left[g_{c,\bar{x}}(\bar{\alpha}) + \sigma_{RN}^2 \right] \right) \\ - \left[d_{c,\bar{x}} + \sigma_{RN}^2 \right] \log \left(\left[d_{c,\bar{x}} + \sigma_{RN}^2 \right] \right) \\ + d_{c,\bar{x}}(\bar{\alpha}) - g_{c,\bar{x}}(\bar{\alpha}) \end{array} \right\}$$

(1)

c = diversity sensor channel index
$\bar{x}$ = image pixel 2D index
$\bar{\alpha}$ = WF coeffient vector
$d_{c,\bar{x}}$ = value of channel c measured image at pixel $\bar{x}$
$g_{c,\bar{x}}(\bar{\alpha})$ = estimate of $d_{c,\bar{x}}$ at WF $\bar{\alpha}$
σ_{RN}^2 = read noise variance

Here the predicted image $g_{c,x}(\alpha)$ (which is simply just the PSF for the restricted PDPR case considered here) is the square of the Fourier transform of the aperture coherent transmission function.

$$PSF(\bar{x}) = \left| \int d\bar{u}\, T(\bar{u}) \exp\left(\frac{i2\pi}{\lambda} OPD(\bar{u}) \right) \exp\left(i2\pi \bar{u} \square \bar{x} \right) \right|^2$$

$OPD(\bar{u})$ – optical path distance wavefront aberration

$$= \sum_j \alpha_j Z_j(\bar{u})$$

(2)

$\bar{x}$ – image plane 2D coordinate
$\bar{u}$ – pupil plane 2D coordinate
$\bar{\alpha}$ – Zernike coefficients vector
$Z_j(\bar{u})$ – Zernike mode j
$T(\bar{u})$ – telescope amplitude transmission function

Iterations were terminated when the gradients or changes in cost value diminished to within a factor of 10 of numerical precision double-precision floating point math. This typically took 20-80 iterations with the number of iterations tending to be larger for stronger WF aberrations and fewer for weaker WF aberrations.

Only Zernike modes 2-21 (e.g. through 6th order) were estimated with the higher order coefficients fixed at zero. This provides realism to the analysis that a modal decomposition cannot cover the full space of possible physical wavefronts presented to the system and thus will suffer some inaccuracies as a result. Noll [33] describes the residual wavefront error for Kolmorogov turbulence statistics after compensation for the 1st J Zernike modes, giving specific values for J=1-21 and the approximation of $0.2944 J^{-sqrt(3)/2}(D/r_o)^{5/3}$ rad^2 for J>10. For simplicity consider D= r_o, then for piston compensation only the residual is 1.0299 rad^2, for compensation of modes 1-21 the residual is 0.0208 rad^2, and for compensation of modesl 1-36 the residual is .0132 rad^2. Modes 1-36 then represent 1.0299 - .0132 = 1.0167 rad^2 of aberrations while modes 1-21 cover 1.0299 - .0208 = 1.0091 rad^2. The fraction missed by estimating only the 1st 21 modes is thus (1.0167 - 1.0091) = .0076 rad^2. Thus if the 1st 21 modes are estimated perfectly the residual error is sqrt(.0076)=.0872 rad rms compared to a total rms aberration of sqrt(1.0167)= 1.0083 rad rms. This means that if PD perfectly estimates the 1st 21 modes, the rms residual error associated with the unestimated modes 22-36, is on average about 10% of the full aberration level.

For each of the following combination of PSNR and WF aberration level
- PSNR   = {10, 25, 50, 100, 100}
- Wavefront aberration levels   = {. 05, .01, .15, .20, .25, .35, .50} λ rms

four random true WF aberrations were randomly generated and corresponding noisy PDPR image pairs were generated. PDPR was repeatedly applied to each pair with the initial WF estimate seeded to the following set of values:
- WF coefficient vector initialized to true value
- WF coefficient vector initialized to zero
- WF coefficient vector initialized to 250 randomly drawn guesses generated as follows:
    o WFs coefficient vectors were randomly generated such that coefficient covariances matched the Kolmogorov statistics corresponding to 2.5× the rms aberration level of the true WF
    o Individual draws were repeated until one was found in which all elements of the coefficient vector were less than 2.5× their Kolmogorov variances.
    o These two restrictions served to focus the interogated region of the parameter space on values that are statistically likely (though it is possible a deep and wide local minima outside this constrained space could draw the search algorithm to it).

The rational for the method of generating random starting guesses is that one may generally have *a priori* expectations of the strength of aberration likely to be encountered. Assuming that the rms aberration level is known, it seems reasonable to limit the search space to a region in which captures covers the majority of WF possibilities that can statistically be expected to occur (the factor of 2.5 times used here being somewhat arbitrary).

3.   Local Minima Phenomenology and Cost Function Topology

As shown in Figure 3, local minima in the negative log-likelihood (-LL) function are associated with incorrect WFs that produce a similar speckle pattern to the true PSF. When the relative difference between the predicted and measured PSF is small the –LL function approximately corresponds to the weighted mean-square-sum of the difference of the measured and predicted images, spatially weighted by the local pixel noise level. Thus a good match can occur when the major PSF speckle lobes of the measured and predicted image align to each other well or when there is good alignment between regions of generally high energy (e.g. a complex composite of many closely positioned minor speckle lobes). As illustrated in Figure 4, adjusting the wavefront coefficients to morph the associated PSF speckle pattern toward the true PSF requires first reducing the match to the true speckle pattern before it can improve again. The phenomenology involved is nicely illustrated in Figure 5, by considering a telescope aperture consisting of two mirror segments and a WF aberration corresponding to an independent WF tip/tilt error on each of the individual segments. A general qualitative trend that was observed in visually comparing the local minima wavefronts for many different true wavefront over a range of true aberration strengths and SNR levels is that the estimated WF matches the coarse global structure of the true WF quite well with the discrepancy tending to be strongly locally concentrated. Examples of local concentration of the discrepancy are highlighted with dashed red circles in Figure 3. This trend bodes well for use of phase diversity in closed loop systems. If multiple phase diverse measurements associated with the same aberration can be made, it may be possible to exploit this

phenomena to resolve ambiguities. It was also noticed that in most cases the local minima produce a good match to the PSF in one channel but not the other. In the two leftmost panels of Figure 3, the estimated WF associated with a local minima match the in-focus channel well but not the defocus. In the 100 or so local minima cases visually investigated, this was almost always the situation. However, there were a few rare instances in which the match was striking in the de-focus channel and weaker for the in-focus channel as illustrated in the rightmost panel of Figure 3.
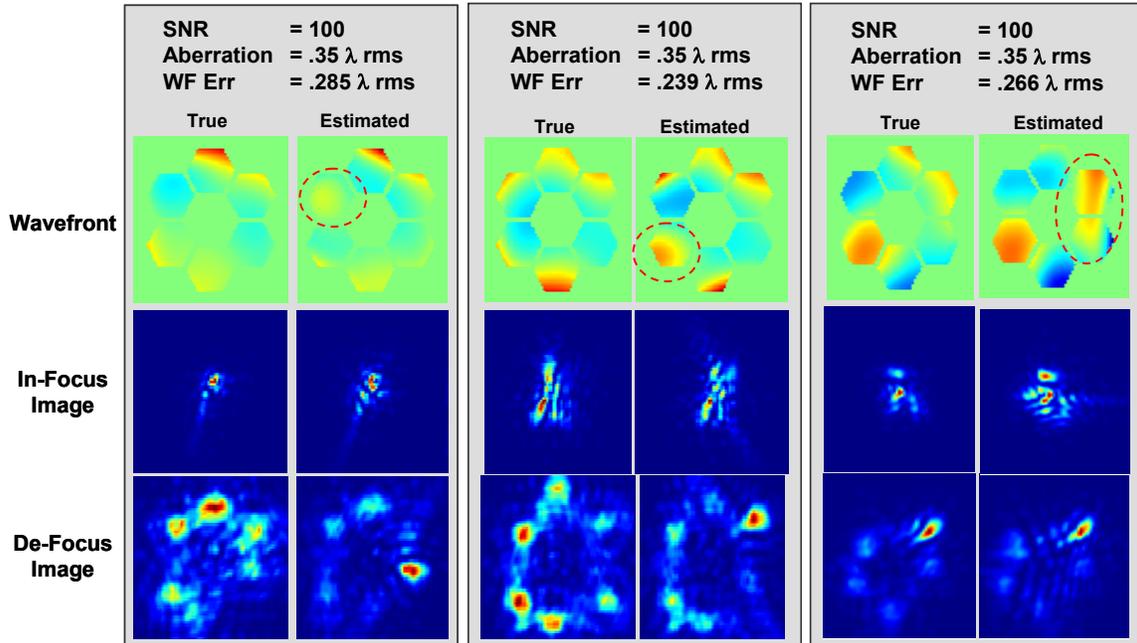
Figure 3. Local minima occur for incorrect WFs that produce a PSF speckle pattern that matches well to the true PSF. A general trend that was observed was for the WF corresponding to the local minima to match the general global structure of the true WF well with the differences tending to be locally concentrated (as highlighted by dashed red circles in the figure).
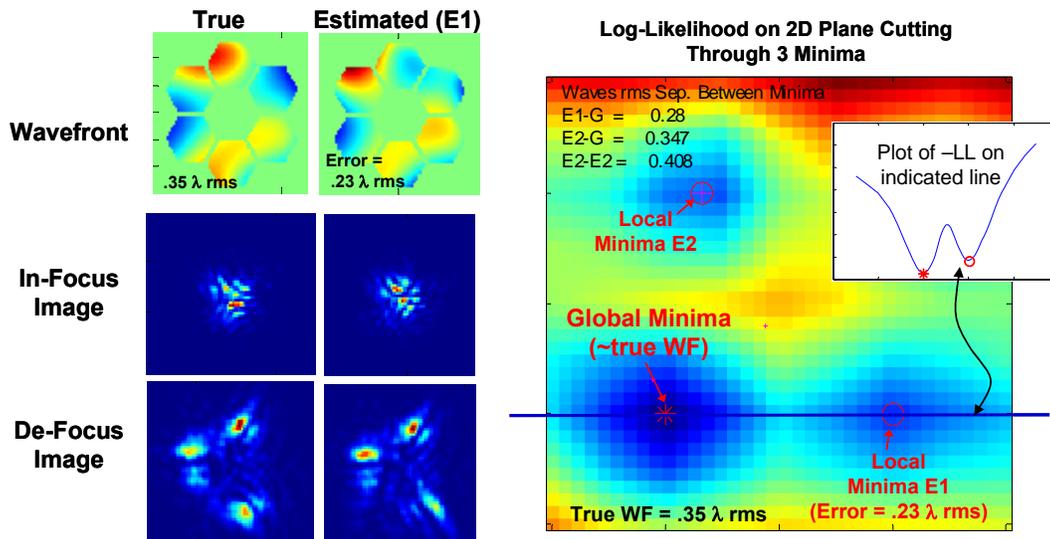

Figure 4. (left) True WF and corresponding in- and de-focus images. (center) Local minima WF gradient searches can be trapped in. (right) color-scale plot of –LL on a 2D plane that cuts through the global minima and 2 local minima with lowest –LL. Adjusting the WF coefficients from their values at E1 to morph the associated PSF speckle pattern toward the true PSF requires reducing the match to the true speckle pattern before it can improve again.
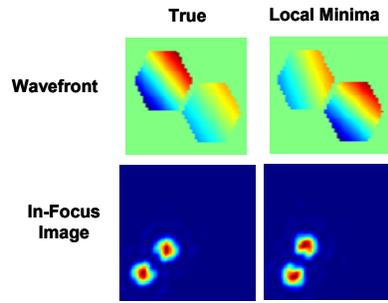
Figure 5. The phenomenology involved in how two WFs can produce similar PSFs is nicely illustrated by considering a telescope aperture consisting of two mirror segments and a WF aberration corresponding to an independent WF tip/tilt error on each of the individual segments. The tip/tilt error for a segment causes the reflected light to be directed at an angle from nominal and corresponding to a translational shift in the image speckle (coherent PSF) it produces. If the tip/tilts of each mirror differ enough, the two projected spots will not overlap and coherent interference will be negligible. Swapping the tip/tilt tilt errors of the two segments, causes the position of their speckles to also swap producing a PSF pattern that is almost identical to the original. Now consider a WF over a full aperture region decomposed into a composite of many small segments each with its own independent piston/tip/tilt aberration. This is similar to the representation of atmospheric wavefront aberrations as composed of many r_o size patches over which the piston/tip/tilt is relatively statistically independent and which each produces a PSF speckle with translational position related to the tip/tilt. Swapping the tip/tilt values between various any two patches produces the same speckle projection but with switched relation between the speckles and source regions. It is immediately evident that at large D/r_o there will be many WFs that produce similar speckle patterns.

One particular local minima frequently occurs as a result of the fact that rotating a WF by 180° and flipping its sign produces the identical PSF as the original WF. This can be easily shown using the PSF relation in Eq. (2). For brevity this particular WF transformation will be referred to as the -Rot180 WF hereafter. The defocus channel sees the common WF aberration plus an additional defocus, thus the image corresponding to the –Rot180 WF will not be the same as the true measured image (as illustrated in Figure 6). However, since the LL function is the sum of the individual LL for each imaging channel, a good match in one channel can still often cause a local-minima. Examination of the Monte-Carlo simulation results showed that the –Rot180 issue was found to correspond to a local minima in the majority of cases at all WF aberration strengths and SNR levels considered. Rank ordering the minima by their –LL values, in most cases the –Rot180 WF was 2nd after the global minima even at SNR ≥ 100.
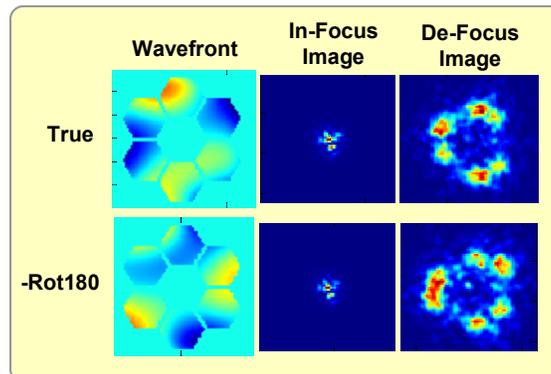

Figure 6. Rotating the wavefront aberration 180° and flipping its sign (-Rot180) produces the identical PSF. However, the defocus channel sees the –Rot180 wavefront plus defocus which produces a different PSF than the original wavefront plus defocus. Despite the mis-match in the defocus channel, the perfect match in the in-focus channel is still usually enough for the –Rot180 WF to result in a local minima.

As the SNR decreases such that the noise variance rises above the strength of weaker speckles, the match of these speckles becomes inconsequential and the density of local minima that occur tends to increase. This phenomena is illustrated in Figure 7.
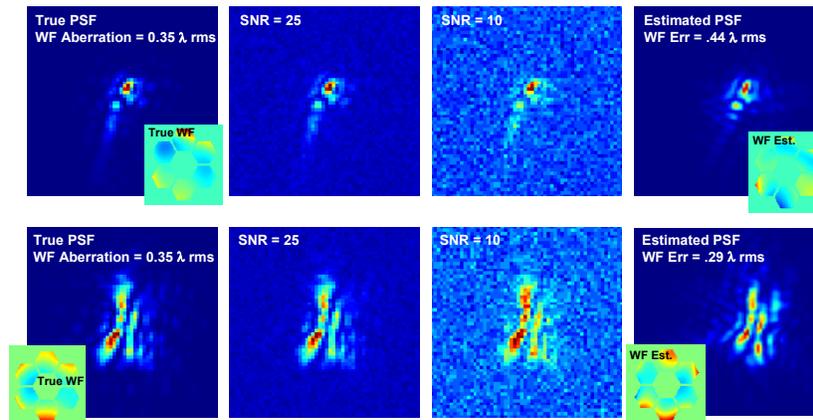

Figure 7. At low SNR the match between the weaker speckles of the true and estimated PSF becomes inconsequential relative to noise level tending to produce a higher density of local minima and generally degrading PDPR WF estimation reliability and accuracy.

To further understand the topological trends of the –LL function, color-scale plots of the –LL were examined on rectangular regions of planes cutting through the global minima and two lowest valued secondary minima. A few examples are shown in Figure 8. The structural characteristics had similar trends for all cases and suggested the following properties of the nature of the multi-dimensional –LL function (20 WF coefficient parameters):

- Topology about local-minima is a simple well defined dip and was not typically strongly elongated in any direction.
- Complex convoluted canyons did not occur
- On straight line connecting 2 minima the –LL typically varies smoothly up then back down without intermediate oscillations
- Minima are generally distinct from each other, separated by at least several 1/100's λ rms (and usually several 1/10's of a λ)
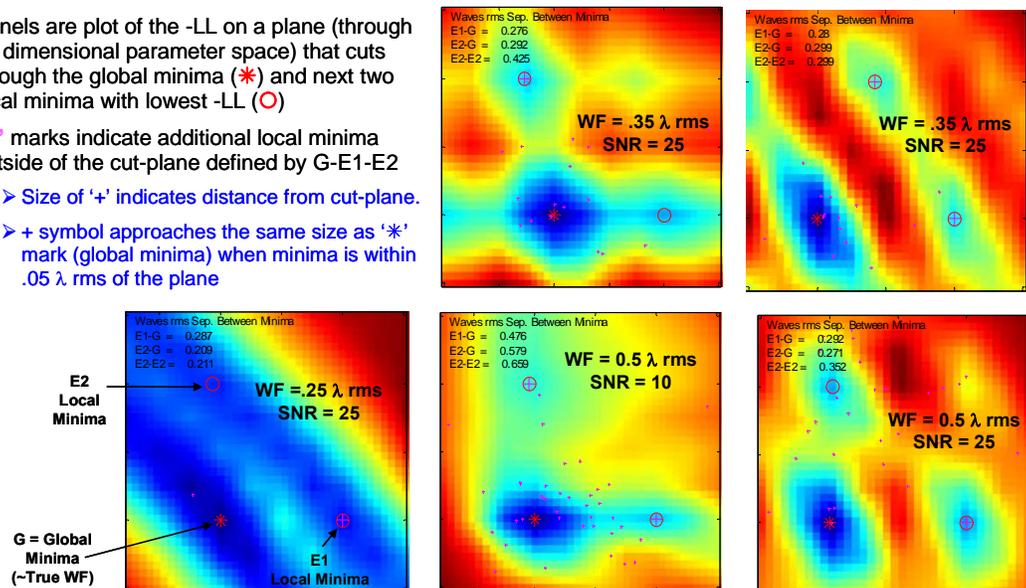- Minima were fairly uniformly distributed in the space without obvious clustering patterns

Figure 8. Examples of cuts through the 20-dimensional –LL function on planes passing through the global (G) and next two lowest valued local minima (E1 & E2).

To further examine the statistics of the distributions of the local minima, for each random PDPR image draw histograms two histograms were generated. The first was of the rms error for each local minima found for the particular PDPR image pair data. The second histogram was of the rms WF distance between all pairs of local minima found. This required identifying the set of unique local minima found for the 250 starting guesses tried for each case. Two gradient search results initiated at different starting guesses may converge to essentially the same local-minima though the exact values will differ because of numerical precision issues. WF estimate vectors for which all Zernike coefficients were within $1/1000$ λ rms of each other were attributed to the same local minima. The results were found to be insensitive to changes in this threshold by ± one order of magnitude. Typical examples are shown in Figure 9 and graphs of the mean number of local minima vs. wavefront aberration level and SNR are shown in Figure 10. The number of local minima increased strongly with increasing WF aberration level and to a lesser degree with decreasing SNR. This may be in part due to the fact that for larger true WF aberrations levels the domain from which starting WF guesses for the iterative search were drawn was increased proportionally. Several trends were observed.

- At WF aberration levels $\leq .1$ λ rms the majority of the time either no secondary minima occurred or a single secondary minima corresponding to the –Rot180 WF occurred. However there were certainly a non-trivial number of exceptions in which multiple local minima occurred and/or the local minima was not the –Rot180 WF.
- 1-2 minima in addition to –Rot180 ambiguity typically occur for aberration levels of .15 λ rms, even at SNR ≥ 100
- At a wavefront aberration level of .2 λ rms, ~10 minima typically occur within a .5 λ rms hyper-sphere.
- For all cases tested the global minima corresponded to the lowest wavefront error of all identified minima.

These statistics show that the local minima problem is an important challenge and will significantly reduce PDPR reliability if not addressed. It does show, however, that the number of minima is closely related to the expected range of WF aberration levels that can be expected to occur for a particular situation and that it appears likely that this corresponds to a polynomial-type growth law (currently under investigation), and that restrictions may be placed on the required density of interrogation needed for a global search.
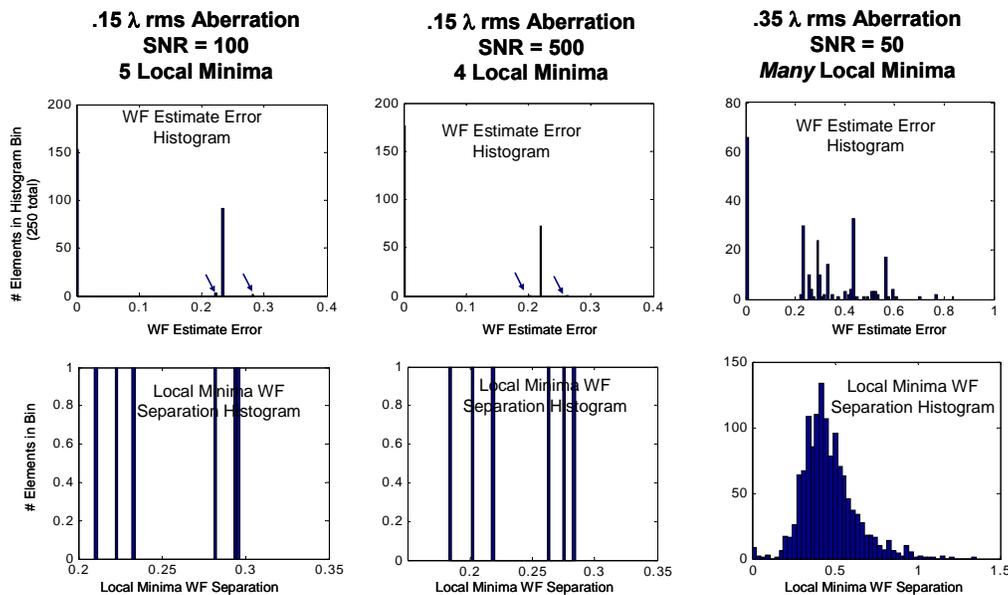

Figure 9. (top) Example histograms of the WF estimate errors resulting from the 250 random search seeds for various SNR and WF aberration levels. Arrows highlight local minima that only a few random starts converged to and which are thus hard to see. (bottom) corresponding histograms of the separations between the resulting set of unique local minima found.
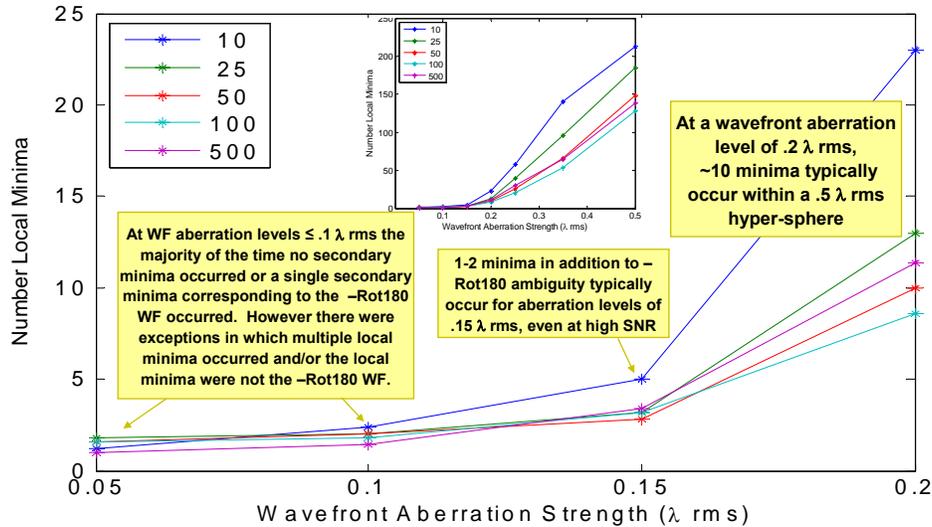
Figure 10. Mean number of local minima found as function of wavefront aberration strength.

4. Summary

The non-linear relation between a WF and its corresponding PSF allows two WFs to produce PSFs with speckle lobe distributions that match well enough to cause local minima in a cost function. This paper examined a Log-Likelihood function that approximately models a mix of Poisson shot noise and additive Gaussian read noise, but the results can be expected to hold for other cost function models. These local minima cause convergence of iterative PDPR approach to the correct global minima to be unreliable even at high SNR and for low WF aberration levels. The phenomenology of the problem is well illustrated by considering the fact that each patch of the WF forms a PSF speckle transversely positioned in proportion to the local WF tip/tilt and that swapping the tip/tilt values of two patches will direct speckles to the same location (just the source regions are switched) and thus produce similar PSF patterns. An additional systematic local minima that occurs is related to the fact that rotating a WF 180° and flipping its sign produces an identical PSF to the original WF. Examining the –LL function on 2D planes cutting through its higher dimensional space for many randomly drawn WF aberrations indicated that the structure of the –LL function was generally simple with a topology that was primarily related to the positions of the local minima and no other complex structures. The local minima were found to be generally fairly evenly distributed across the WF coefficient space without any obvious clustering patterns and the number of local minima within a space 2.5 times the true aberration level was found to grow rapidly with increasing aberration level. Local minima often occured even at low aberration levels and high SNR showing that unless addressed this presents a major issue to reliable use of phase diversity techniques for wavefront estimation. Though this paper focused on the restricted case of PDPR in which the object is known and not estimated, we can expect the issues and trends to be similar for the more general PD WF sensing case.

5. Acknowledgements

6. References

1. Gonsalves, R. A., "Phase retrieval and diversity in adaptive optics," Opt Eng 21, pp 829-832, 1982.

2. B. J. Thelen, M.F. Reiley, R.G. Paxman, and D.W. Amlin, "Fine resolution multispectral imaging using measurement diversity," presented at the International Symposium on Spectral Sensing Research, Melbourne, Australia, Nov. 26 – Dec. 1, 1995.
3. Harry R. Ingleby and Donald R. McGaughey, "Experimental results of parallel multiframe blind deconvolution using wavelength diversity," Proc. SPIE 5578 8-14 (2004).
4. D. J. Lee, M. C. Roggemann, B. M. Welsh, "Cramer–Rao analysis of phase-diverse wave-front sensing," JOSAA 16(5) 1005-1015 (1999).
5. Paxman, R. G., Schulz, T. J., Feinup, J. R. "Joint estimation of object and aberrations by using phase diversity", J. Opt. Soc. Am. A 9, pp 1072- 1085, 1992.
6. Torgersen, T. C., Tyler, D. W. "Practical Considerations in Restoring Images from Phase-Diverse Speckle Data", Pubs. Of The Astron. Soc. Of The Pac. 114, pp 671-685, 2002.
7. M. R. Bolcar and J. R. Fienup, "Method of Phase Diversity in Multi-aperture Systems Utilizing Individual Sub-Aperture Control," Proc. SPIE **5896** (2005).
8. D. A. Carrara, B. J. Thelen, R. G. Paxman, "Aberration correction of segmented-aperture telescopes by using phase diversity," Proc. SPIE 4123 (2000).
9. J. J. Dolne, P. Menicucci, D. Miccolis, et al, "Real Time Phase Diversity Advanced Image Processing and Wavefront Sensing," Proc. SPIE (2007).
10. R. L. Kendrick, J.-N. Aubrun, R. Bell, et al, "Wide-field Fizeau imaging telescope: experimental results," Appl. Opt. **45**(16) 4235-4240 (2006).
11. R. L. Kendrick, D. S. Acton, and A. L. Duncan, "Phase-diversity wave-front sensor for imaging systems," Appl. Opt. 33(27) 6533-6546 (1994).
12. D. J. Lee, M. C. Roggemann, B. M. Welsh, and E. R. Crosby, "Evaluation of least-squares phase-diversity technique for space telescope wave-front sensing," Appl. Opt. 36(35) 9186-9197 (1997).
13. M. G. Löfdahl, R. L. Kendrick, A. Harwit, K. E. Mitchell, A. L. Duncan, "A Phase Diversity Experiment to Measure Piston Misalignment on the Segmented Primary Mirror of the Keck II Telescope," Proc. SPIE 3356 1190-1201 (1998).
14. M. G. Löfdahl, G. B. Scharmer, "A Predictor Approach to Closed-Loop Phase Diversity Wavefront Sensing" Proc. SPIE 4013 (2000).
15. R. G. Paxman, J. H. Seldin, M. G. Lofdahl, G. B. Scharmer, et al, "Evaluation of Phase-Diversity Techniques for Solar-Image Restoration," Astrophys. J. 466 1087-109 (1996).
16. R. G. Paxman, B. J. Thelen, R. J. Murphy, et al, "Phase-Diverse Adaptive Optics for Future Telescopes." Proc. SPIE 6711 (2007).
17. R. G. Paxman and J. R. Fienup, "Optical misalignment sensing and image reconstruction using phase diversity," JOSAA 5(6) (1988).
18. D. A. Carrara, B. J. Thelen, R. G. Paxman, "Aberration correction of segmented-aperture telescopes by using phase diversity," Proc. SPIE 4123 56-63 (2000).
19. D. Redding, S. Basinger, D. Cohen, "Wavefront Control for a Segmented Deployable Space Telescope," Proc. SPIE 4013 546-558 (2000).
20. R. L. Kendrick, R. Bell, A. L. Duncan, "Closed loop wave front correction using phase diversity," SPIE 3356 844-853 (1998).
21. M. G. Löfdahl, H. Eriksson, "Resolving Piston Ambiguities when Phasing a Segmented Mirror," SPIE 4013 774-782 (2000).
22. X. Rondeau, E. Thiébaut, M. Tallon, and R. Foy, "Phase retrieval from speckle images," JOSAA 24(10) 3354-3365 (2008).
23. B. J. Daniel, M. R. Bolcar, J.. R. Schott, and J. R. Fienup, "Phase Retrieval in Sparse Aperture Systems with Phase Diversity: A Trade Space Study," Proc. SPIE 6958 (2008).
24. T. C. Torgersen and D. W. Tyler, "Practical considerations in restoring images from phase-diverse speckle data," Astron. Soc. of the Pacific **114** 671-685 (2002).
25. B. J. Thelen, R. G. Paxman, J. H. Seldin, and D R. Rice, "Bayes estimation of dynamic and fixed aberrations, and object from phase-diverse speckle data," SPIE **2827** 32-43 (1996).
26. C. Luna, D. R. Gerwe, and B. Calef, "Phase diverse wavefront estimation and image restoration with a magnification change between imaging channels," Proc. SPIE **5896** (2005).
27. M. C. Roggemann, B. Welsh, Imaging Through Turbulence, CRC Press, 1996.
28. D. Gerwe, J Stone, C. Luna, B. Calef, "Comparison of Maximum-Likelihood Image and Wavefront Reconstruction using Conventional Image, Phase Diversity, and Lenslet Diversity Data," SPIE Proc. **6307** 213-225 (2006).

29. C. Zhu, R. H. Byrd, and J. Nocedal, "L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization," ACM Transactions on Mathematical Software 23, pp. 550-560, (1997).
30. D. L. Snyder, A. M. Hammoud, and R. L. White, "Image recovery from data acquired with a charge-coupled-device camera," JOSA A 10(5) 1014–1023 (1993).
31. D. L. Snyder, C. W. Helstrom, A. D. Lanterman, M. Faisal, and R. L. White, "Compensation for readout noise in CCD images," JOSA A 12(2), 272–283 (1995).
32. D. R. Gerwe, P. S. Idell, "Cramér-Rao Analysis of Orientation Estimation: Viewing Geometry Influences on the Information Conveyed by Target Features," JOSA A 20(5) 797-816 (2002).
33. R. Noll, "Zernike polynomials and atmospheric turbulence," JOSA 66(3) 207-211 (1976).