Use of Zernike polynomials for efficient estimation of orthonormal aberration coefficients over variable noncircular pupils

Hanshin Lee
McDonald Observatory, University of Texas at Austin, 1 University Station C1402, Austin, Texas 78712, USA (lee@astro.as.utexas.edu)

Received February 1, 2010; revised May 24, 2010; accepted May 26, 2010; posted June 9, 2010 (Doc. ID 123632); published June 22, 2010

An efficient way of estimating orthonormal aberration coefficients on variable noncircular pupils is proposed. The method is based on the fact that all necessary pieces of information for constructing orthonormal polynomials (via the Gram–Schmidt process) can be numerically obtained during a routine least-squares fit of Zernike polynomials to wavefront data. This allows the method to use the usual Zernike polynomial fitting with an additional procedure that swiftly estimates the desired orthonormal aberration coefficients without having to use the functional forms of orthonormal polynomials. It is also shown that the method naturally accounts for the pixeltation effect of pupil geometries, intrinsic to recording wavefront data on imaging sensors (e.g., CCDs), making the coefficient estimate optimal over a given pixeltated pupil geometry. With these features, the method can be ideal for real-time wavefront analysis over dynamically changing pupils, such as in the Hobby–Eberly Telescope (HET), which is otherwise inefficient with analytic methods used in past studies. © 2010 Optical Society of America

OCIS codes: 010.1080, 080.1010, 120.3180, 220.4840.

Aberration coefficients of optical wavefront are affected by the physical state of systems where the wavefront has propagated (e.g., the Earth atmosphere or a telescope). Thus, useful insights into the observed behaviors of the systems can be acquired by estimating the coefficients from wavefront measurements. In doing so, the estimated coefficients are desired to be unique, and this is enabled by fitting a set of orthonormal polynomials to the measurement [1]. Zernike polynomials, which constitute one particular such set over a unit disk and are used in this Letter, have been applied [2–4]. However, there are increasingly many optical systems whose pupils are noncircular [5–7]. In addition, some systems exhibit variability in pupil shape as a function of field or pointing angles [8,9]. On such pupils, Zernike polynomials lose orthogonality, and it is desirable to use the coefficients of new orthonormal aberration polynomials.

Studies show that orthonormal polynomials can be analytically constructed, via the Gram–Schmidt (GS) process, for noncircular pupils in simple shape without shape variability, i.e., annulus, hexagon, ellipse, and rectangle [10–13]. The studies overlook issues in applying the analytic method to wavefront data measured over more complex pixeltated (due to pixel-based imaging sensors) pupil shapes with dynamic variability. The analytic method can be extended to general pupils, but the analytic calculations can be challenging over more complicated pixeltated pupil geometries, such as that of the Hobby–Eberly Telescope (HET). Any dynamic change in pupil shape complicates the execution of the method even more, as the analytic forms of orthonormal polynomials have to be obtained whenever the change happens. This can also slow the orthonormal aberration coefficient estimation in real-time wavefront analysis, thereby diminishing the advantage of having orthonormal coefficients over variable noncircular pupils.

We propose an efficient way of estimating orthonormal aberration coefficients that can resolve the problems mentioned previously. The estimation uses the usual numerical fit of Zernike polynomials to wavefront data recorded in a pixeltated format, but essentially without having to compute the functional forms of the orthonormal polynomials (which can be complex and time-consuming), thereby keeping simplicity and speed of the routine least-squares Zernike fit. The orthonormal coefficients obtained in this way, naturally account for the pupil pixeltation effect and are, thus, optimal over a given pixeltated pupil geometry. The details of the method are described as follows.

The Zernike polynomials \(Z_i\) can be identified by radial and angular orders \(|n|, |m|\) [1]. The polynomial index scheme used here follows: \(Z_{|n|,|m|}\) precedes \(Z_{|n|',|m|'}\) if \(|n| < |n'|\) if \(|n| = |n'|\) and \(|m| < |m'\|, Z_i\) precedes \(Z_j\) if \(|n| = |n'|, |m| = |m'|\), and \(|m| < 0\). \(Z_i\) precedes \(Z_j\). Suppose a wavefront function \(W(x, y)\) on a noncircular pupil \(E\) that is a subset of a unit disk \(S\). Assume that \(W\) can be described in terms of \(M\) Zernike polynomials,\n
\[ W = \alpha_1Z_1 + \alpha_2Z_2 + \ldots + \alpha_MZ_M. \]

The estimation of \(\alpha_i\) starts with the following (\(A\) the area of \(E\)):

\[ \sum_{j=1}^{M} \alpha_j F_{ij} = \sum_{j=1}^{M} \alpha_j \int_{E} \frac{Z_j Z_i}{A} \, dA = \int_{E} \frac{Z_i W}{A} \, dA. \]

\(Z_i\) are not orthonormal over \(E\), and thus a \(M \times M\) matrix \(F\), whose \((i,j)\) element is \(F_{ij}\), is neither unitary nor diagonal. The estimate of \(\alpha_i\) can still be given by

\[ \alpha_i = \sum_{j=1}^{M} G_{ij} \int_{E} \frac{Z_i W}{A} \, dA, \]

with \(G_{ij}\) being the \((i,j)\) element of \(G = F^{-1}\). \(\alpha_i\) is neither \(\alpha_i\) nor the desired orthonormal aberration coefficient (\(\beta_i\)) of \(W\) on \(E\), but closely related to \(\beta_i\) as follows.

© 2010 Optical Society of America
It is widely known that new orthonormal polynomials, say \( U_i \), can be determined by the GS orthonormalization of a known set of orthonormal polynomials \( Z_j \) in our case\(^{14} \). To utilize the GS orthonormalization, we allow \( U_1 = Z_j \):

\[
V_j = Z_j - \sum_{k=1}^{l-1} C_{jk} U_k, \quad U_j = V_j \sqrt{A/\iint_E V_j V_j dA}. \tag{4}
\]

with \( C_{jk} = A^{-1} \iint_E Z_j U_k dA \) where \( j = 1, 2, \ldots, M \). Using Eq. (4), the following algebraic relation is obtained:

\[
\iint_E V_j V_j dA = F_{ij} + \sum_{m=1}^{l-1} \left\{ \sum_{n=1}^{l-1} C_{im} \delta_{mn} C_{nj} \right\} - \sum_{n=1}^{l-1} C_{in} C_{nj} = \sum_{m=1}^{l-1} C_{im} C_{mj}, \tag{5}
\]

with \( \delta_{mn} \) being the Kronecker delta. Given the orthogonality of \( V_j \), the left-hand side of Eq. (5) equals \( C_{ij}^2 \) for \( i = j \) or vanishes for \( i \neq j \). This reduces Eq. (5) to the following algebraic expressions of the GS orthonormalization coefficient \( C_{ij} \):

\[
C_{ij} = \frac{F_{ij} - \sum_{k=1}^{l-1} C_{ik}^2 C_{jk}}{C_{ij} - \sum_{k=1}^{l-1} C_{ik} C_{jk}}, \tag{6}
\]

\( C_{ij} \) is essentially the \((i,j)\) element of a \( M \times M \) lower triangle matrix \( C \), which is the Cholesky factor of \( F \), as noted in a different way elsewhere\(^ {12} \). From Eq. (4), \( Z_j = \sum_{i=1}^j C_{ij} U_j \) by which \( W \) can be given as

\[
W = \sum_{i=1}^M \alpha_{i,e} Z_i = \sum_{i=1}^M \sum_{j=1}^i \alpha_{i,e} C_{ij} U_j = \sum_{j=1}^M \beta_j U_j, \tag{7}
\]

leading to \( \beta_j = \sum_{i=1}^M \alpha_{i,e} C_{ij} \).

Note also \( U_i = \sum_{j=1}^i D_{ij} Z_j \), where \( D_{ij} \) is the \((i,j)\) element of \( D = C^{-1} \). The above orthonormalization process clearly indicates that knowing \( F_{ij} \) is central to computing \( C_{ij} \) and constructing analytic forms of \( U_i \), via \( D_{ij} \), which are fitted to \( W \) to determine \( \beta_i \). However, analytically computing \( F_{ij} \) over a complex pupil can be out of the question. Although computer algebra software or numerical quadrature could be helpful, it is still required to extract (or approximate) the pupil geometry from a measured pupil illumination or to divide the geometry into a number of integrable simpler subdomains, which would make the whole analytic approach increasingly complicated and vastly inefficient when large sets of orthonormal polynomials need to be obtained for variable noncircular pupils. In addition, it would be much harder to incorporate the pixelation effect, intrinsic to digital wavefront data, into analytic derivations of \( F_{ij} \) and \( U_j \) and ultimately into estimating \( \beta_i \).

However, realizing that the wavefront is recorded on \( N \) discrete pixels on \( E \) (a \( N \times 1 \) vector), Eq. (1) is given as a least-squares Zernike fit in a matrix form:

\[
\begin{aligned}
\begin{bmatrix}
\tilde{W}
\end{bmatrix} &= Z \tilde{\alpha} \\
\tilde{\alpha} &= (Z^T Z)^{-1} Z^T \tilde{W} = \hat{\alpha},
\end{aligned}
\tag{8}
\]

where \( \tilde{\alpha} \) is the column vectors of \( \alpha \), and \( Z \) is a \( N \times M \) matrix of \( M \) Zernike polynomials evaluated on the pixels \( (E_{\text{pix}}) \). \( \hat{\alpha}_{i,e} \) (the \((i,e)\)th element of \( \hat{\alpha} \)) is the estimate of \( \alpha_i 

new polynomials. Also, $\hat{D}_{ij}$ are close to $D_{ij}$ except for a few with small contributions. Because of analytically deriving $D$, method X took ~72 s longer than method Y on the same computer. Thus, the efficiency gain by using method Y in dealing with variable noncircular pupils can be immense.

In the second example, the variance of the wavefront on pupil B [Fig. 1(b)] is given in three ways: the true variance of the data $(\sigma^2)$, the sum of squares of the first 15 of $\hat{\beta}_1(\sigma^2(\beta))$, and the sum of squares of the first 15 of $\hat{\alpha}_i(\sigma^2(\alpha))$. Note the dynamic change in the pupil B shape as a function of the pupil position parameter. These variances are plotted in Fig. 2. As expected, $\sigma^2(\beta)$ closely approximates $\sigma^2$, while $\sigma^2(\alpha)$ differs from the true value. As the pupil shape becomes closer to the unit circle (i.e., position $\sim 0.7$), $\sigma^2(\alpha)$ asymptotes the other two curves, but quickly diverts away as the pupil shape starts departing from the unit disk again. The last 21 of $\beta_i$ are also estimated independently, and the sum of squares of these $(\Delta \sigma^2(\beta))$ matches the true difference $\Delta \sigma^2 = \sigma^2(\beta) - \sigma^2$.

As demonstrated, the proposed method can be used to efficiently estimate orthonormal aberration coefficients of wavefront data over noncircular pixelated pupils with shape variability. It uses the routine least-squares fit of Zernike polynomials to wavefront measurements, but without having to know the functional forms of the orthonormal polynomials. This effectively eliminates the complex and time-consuming part from the estimation, thereby keeping the simplicity and speed of the usual least-squares fit, while ensuring the optimality of the resultant orthonormal aberration coefficients over a given pixelated pupil geometry. These features not only are ideal for real-time wavefront analysis over dynamically varying pixelated pupils but are also useful in extending the method to estimating orthonormal slope aberration coefficients, where mean wavefront slope is measured over a further complicated geometry formed by variable noncircular pupils and a coarse slope sampling grid [16].

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\textbf{i} & \textbf{True $\beta_i$} & \textbf{Method X} & \textbf{Method Y} & \textbf{|X - Y|} \\
\hline
1 & -0.0076 & -0.0076 & -0.0072 & 0.493 \times 10^{-3} \\
2 & -0.1133 & -0.1134 & -0.1134 & 0.051 \times 10^{-3} \\
3 & -0.0233 & -0.0234 & -0.0235 & 0.055 \times 10^{-3} \\
4 & 0.2741 & 0.2740 & 0.2741 & 0.103 \times 10^{-3} \\
5 & -0.0501 & -0.0502 & -0.0502 & 0.012 \times 10^{-3} \\
6 & 0.1706 & 0.1706 & 0.1706 & 0.023 \times 10^{-3} \\
7 & -0.1108 & -0.1110 & -0.1107 & 0.247 \times 10^{-3} \\
8 & -0.0409 & -0.0411 & -0.0411 & 0.225 \times 10^{-3} \\
9 & -0.1064 & -0.1062 & -0.1062 & 0.070 \times 10^{-3} \\
10 & -0.0369 & -0.0365 & -0.0365 & 0.001 \times 10^{-3} \\
11 & -0.0648 & -0.0648 & -0.0645 & 0.310 \times 10^{-3} \\
12 & -0.0057 & -0.0058 & -0.0058 & 0.063 \times 10^{-3} \\
13 & -0.1070 & -0.1071 & -0.1072 & 0.109 \times 10^{-3} \\
14 & -0.0750 & -0.0748 & -0.0749 & 0.111 \times 10^{-3} \\
15 & 0.0464 & 0.0466 & 0.0466 & 0.018 \times 10^{-3} \\
\hline
\end{tabular}
\caption{Table 1. $\hat{\beta}_i$ by Method X and Y on Pupil A}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
\textbf{i} & \textbf{D}_{ij} & \textbf{D}_{ij} & \textbf{D}_{ij} & \textbf{D}_{ij} & \textbf{D}_{ij} \\
\hline
01,01 & 1.000 & 1.000 & 08,08 & 1.214 & 1.214 \\
02,02 & 1.025 & 1.025 & 09,02 & -0.006 & -0.003 \\
03,03 & 1.022 & 1.022 & 09,07 & -0.006 & -0.003 \\
04,01 & 0.109 & 0.108 & 09,09 & 1.068 & 1.068 \\
04,04 & 1.355 & 1.355 & 10,03 & -0.008 & -0.011 \\
05,05 & 1.166 & 1.166 & 10,08 & -0.007 & -0.009 \\
06,01 & -0.004 & -0.004 & 10,10 & 1.385 & 1.385 \\
06,04 & -0.009 & -0.008 & 11,04 & 0.827 & 0.827 \\
06,06 & 1.105 & 1.104 & 11,04 & 0.098 & 0.096 \\
07,02 & 0.495 & 0.493 & 11,06 & -0.007 & -0.006 \\
07,07 & 1.221 & 1.221 & 11,11 & 1.632 & 1.634 \\
08,03 & 0.484 & 0.484 & --- & --- & --- \\
\hline
\end{tabular}
\caption{Table 2. D_{ij} by Analytic Integration and D_{ij} by Proposed Method on Pupil A}
\end{table}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Fig_2.png}
\caption{(Color online) Variance estimates by different methods.}
\end{figure}

References
16. H. Lee, M. Hart, and G. J. Hill are preparing a manuscript to be called "Optimal estimation of wavefront slope aberrations on variable non-circular pupils."