Representation of wavefronts in free-form transmission pupils with Complex Zernike Polynomials

Rafael Navarro\textsuperscript{a,*}, Ricardo Rivera\textsuperscript{a}, Justiniano Aporta\textsuperscript{b}

\textsuperscript{a}ICMA, Universidad de Zaragoza and Consejo Superior de Investigaciones Científicas, Zaragoza, Spain
\textsuperscript{b}Departamento de Física Aplicada, Universidad de Zaragoza, Zaragoza, Spain

Received April 7, 2011; accepted April 15, 2011

Abstract

\textbf{Purpose}: To propose and evaluate Complex Zernike polynomials (CZPs) to represent general wavefronts with non uniform intensity (amplitude) in free-form transmission pupils.

\textbf{Methods}: They consist of three stages: (1) theoretical formulation; (2) numerical implementation; and (3) two studies of the fidelity of the reconstruction obtained as a function of the number of Zernike modes used (36 or 91). In the first study, we generated complex wavefronts merging wave aberration data from a group of 11 eyes, with a generic Gaussian model of the Siles-Crawford effective pupil transmission. In the second study we simulated the wavefront passing through different pupil stop shapes (annular, semicircular, elliptical and triangular).

\textbf{Results}: The reconstructions of the wave aberration (phase of the generalized pupil function) were always good. The reconstruction RMS error was of the order of $10^{-4}$ wave lengths, no matter the number of modes used. However, the reconstruction of the amplitude (effective transmission) was highly dependent of the number of modes used. In particular, a high number of modes is necessary to reconstruct sharp edges, due to their high frequency content.

\textbf{Conclusions}: CZPs provide a complete orthogonal basis able to represent generalized pupil functions (or complex wavefronts). This provides a unified general framework in contrast to the previous variety of \textit{ad hoc} solutions. Our results suggest that complex wavefronts require a higher number of CZP, but they seem especially well-suited for inhomogeneous beams, pupil apodization, etc.
Introduction

The Zernike polynomial (ZP) expansion is widely used in optics because ZPs form a complete orthogonal basis on a circle of unit radius. Since many optical systems have a circular pupil, ZP expansion can be used to describe any real function at the pupil plane, such as the phase of a wavefront or the wave aberration. They are on the basis of many applications from optical design and testing, wavefront sensing, adaptive optics, wavefront shaping, corneal topography, etc.

In all these applications the main assumption is that the pupil (wavefront) or surface (topography) has a circular shape. However, in the human eye, the pupil may not be exactly circular, for example for peripheral visual angles, and its effective transmission is not constant but approximately Gaussian due to the wave guiding optical and its effective transmission is not constant but exactly circular, for example for peripheral visual angles, shape. However, in the human eye, the pupil may not be circular, but its shape (eccentricity) changes with visual field and its orientation changes with meridian. The problem or representing the change of low and high order aberrations or representing the change of low and high order aberrations across the 2-dimensional visual field in a compact and homogeneous way still lacks a proper solution. Several solutions were proposed in literature for particular cases. For instance, Zernike annular polynomials were introduced to deal with annular stops. Affine (linear) transformations applied to circular pupils (and Zernike polynomials) permit to compute the effects of rotations, translations or two-dimensional scaling to pass from circular to elliptical geometries and vice versa. It is also possible to orthogonalize Zernike polynomials for general aperture shapes.

A different but related issue is the case of inhomogeneous transmission pupils, or inhomogeneous illumination beams, or a combination of both. In the human eye, the SCE means that the effective pupil transmission of the eye is approximately Gaussian. Modern light sources such as lasers, LEDs, or new optical elements such as axicons proposed to compensate presbyopia produce with inhomogeneous amplitude wavefronts: Gaussian, Bessel or associated beams. Nowadays apodized multifocal intraocular lenses (with inhomogeneous pupil transmission)
may offer improved performance over standard IOLs after cataract surgery.\textsuperscript{17}

This brief overview suggests a wide potential field of application of a proper description of general complex wavefronts, with free-amplitude and free-phase distributions. The purpose of this work is to study the ability of CZPs to represent these general complex wavefronts in human eyes. To this aim we study two different cases: First, the Siles-Crawford apodization as an example of inhomogeneous pupil transmission (or inhomogeneous amplitude). This is especially relevant as this is an intrinsic property of the optical system of the eye, and also because methods such as laser ray tracing can measure both the amplitude (from the relative intensity of spots) and phase (from the centroid of spots) of the wavefront.\textsuperscript{18} Second, we study different types of pupil apertures. As we said above, there exist effective ad hoc solutions even for general pupil shapes.\textsuperscript{12} Also for inhomogeneous transmission the standard approach is to use two separate real functions for $T$ and $W$. In this context, there are two main potential benefits of using CZPs. On the one hand, they provide a unified and generalized solution for a wide variety of pupil shapes and transmissions, since the CZPs form a complete orthogonal basis able to represent any complex wavefront in monochromatic light. The only constraint imposed to the wavefront is that it has to be fully contained within a “reference” circle. On the other hand, as we discuss in Section 4, the generalized Nijboer-Zernike (N-Z) approach permits one to use the same set of coefficients to describe both the wavefront and the amplitude spread function (image quality), by simply changing the basis functions.\textsuperscript{19,20} This may be especially relevant in visual optics applications. In fact, Braat and co-workers\textsuperscript{21} introduced CZPs to compute PSFs using the generalized N-Z theory.

**Complex zernike polynomials**

From now on we will consider a monochromatic wavefront at the pupil plane described as a generalized pupil function (amplitude and phase) of spatial polar coordinates:

$$R(p, \theta) = T(p, \theta) e^{j \delta(p, \theta)} \quad (1)$$

This complex function is defined within a circle of unit radius, which means that the radial coordinate $p = r/R$ is normalized by the pupil radius $R$ of a reference circle which contains the wavefront. $T$ represents the wavefront amplitude, or effective pupil transmission and $k \delta/\delta$ is the phase, where $W(p, \theta)$ is the wave aberration and $k$ is the wave number.

**Basic formulation**

Let us start with a brief review of the formulation of Zernike polynomials. The expression for the real polynomials is (ANS Z80.28 standard) within the circle of unit radius is:

$$Z_m^n(p, \theta) = \begin{cases} N_n^m R_m^n(p) \cos m\theta \text{ for } m \geq 0 \\ -N_n^m R_m^n(p) \sin m\theta \text{ for } m < 0 \end{cases} \quad (2)$$

where $\rho, \theta$ are polar coordinates, and the radial part is given by:

$$R_m^n(p) = \frac{(-1)^s (n-s)!}{s! [0.5 (n+|m|) - s!] [0.5 (n-|m|) - s!]} \rho^{n-2s} \quad (3)$$

A normalization factor is included to guarantee orthonormality:

$$N_n^m = \sqrt{\frac{2 (N+1)}{1 + \delta_{nm}}} \quad (4)$$

The complex version can be obtained by considering couples of polynomials with angular frequencies $+m$ and $-m$, corresponding to the real and imaginary parts respectively. After the required normalization by a factor $\sqrt{2}$ it is straightforward to arrive to the expression for the complex ZPs:

$$C_m^n(p, \theta) = \frac{1}{\sqrt{2}} N_n^m R_m^n(p) e^{j m \delta} \quad (5)$$

This means that we construct a couple of a complex $C$ and its conjugated $C^*$ (or couple of $+m, -m$) from its respective couple of $+m, -m$ real $Z$ polynomials:

$$C_m^m = \frac{1}{\sqrt{2}} \left( Z_m^m - i Z_m^m \right) \quad \text{and} \quad C_m^{-m} = \frac{1}{\sqrt{2}} \left( Z_m^m + i Z_m^m \right) \quad (6)$$

To recover the real polynomials we only need to take the real and imaginary parts:

$$Z_m^m = \text{Re}(C_m^m) = \text{Re}(C_m^{-m}) \quad \text{and} \quad Z_m^{-m} = -\text{Im}(C_m^m) = \text{Im}(C_m^{-m}) \quad (7)$$

Figure 1 shows some examples of the amplitude and phase of the CZPs. Note that the amplitude $\frac{1}{\sqrt{2}} N_n^m \abs(R_m^n(p))$ only depends on radius, whereas the phase term $\text{sign}(R_m^n(p)) e^{j m \delta}$ is a function of both coordinates $(p, \theta)$ (the phase is a pure angular frequency only for $m = n$ that is when $R$ is always positive, such as for $C_0^0$).

**Representation of real and complex functions**

The classical expansion of a real function, such as the wave aberration $W$ in terms of ZPs is

$$W = \sum_{m,n} a_{m,n} Z_m^n + \sum_{m,n} b_{m,n} C_m^n \quad (8)$$

where $a_{m,n}$ are real and $b_{m,n}$ are complex coefficients respectively. Note that in the complex expansion the polynomials are conjugated. It is immediate to show that for real functions, the complex coefficients can be computed from the real ones:
Zernike polynomials. The three examples correspond to \(C_6\), \(C_3\) and \(C_1\), respectively.

\[
b_{n,m}^{(i)} = \frac{a_{n,m}^{(i)} + ia_{n,m}^{(i)*}}{\sqrt{2}} \quad \text{and} \quad b_{n,m}^{(r)} = \frac{a_{n,m}^{(r)} - ia_{n,m}^{(r)*}}{\sqrt{2}}
\]

And conversely,

\[
a_{n,m}^{(i)} = \frac{b_{n,m}^{(i)} + b_{n,m}^{(r)}}{\sqrt{2}} = \sqrt{2} \Re(b_{n,m}^{(i)}) \quad \text{and} \quad a_{n,m}^{(r)} = \frac{b_{n,m}^{(r)} - b_{n,m}^{(r)}}{\sqrt{2}} = \sqrt{2} \Im(b_{n,m}^{(r)})
\]

As a consequence of these expressions, for real functions \(b_{n,m}^{(i)} = b_{n,m}^{(r)}\) the coefficients with \(-m\) and \(+m\) are conjugated, which means that they are not independent (but redundant). This is a general property of this type of expansions, such as complex Fourier series, etc. Therefore, it is totally equivalent to use real or complex versions of ZPs to represent wave aberrations (or real functions in general), and Eqs. 9 and 10 permit to pass from the real to the complex basis and conversely. However, these relations do not hold for complex functions in general.

For complex wavefronts, we simply combine Eqs. 1, 5 and 8 to obtain:

\[
R(\rho, \theta) = \sum_{n,m} b_{n,m}^{(i)} C_{n}^{(i)} = \sum_{n,m} b_{n,m}^{(r)} C_{n}^{(r)} = \frac{1}{\sqrt{2}} N_{1}(\rho) N_{2}(\theta) e^{-i\theta}
\]

where \(b_{n,m}^{(i)}\) are the complex coefficients of the expansion. Note the negative sign in \(e^{-i\theta}\) to take into account the complex conjugation \(C_{n}^{(i)}\). The real coefficients \(a\) are not defined in this case. Nevertheless, when \(W\) is given as an expansion of real ZPs it is straightforward to express the relationship between the real and complex coefficients:

\[
\sum_{n,m} b_{n,m}^{(r)} C_{n}^{(r)} = T(\rho, \theta) e^{im\theta} \quad \text{or} \quad \sum_{n,m} a_{n,m}^{(i)} C_{n}^{(i)} = \frac{1}{\kappa} \ln(\sum_{n,m} b_{n,m}^{(r)} C_{n}^{(r)} / | \sum_{n,m} b_{n,m}^{(r)} C_{n}^{(r)} |) \quad (12b)
\]

For practical implementation we will assume a limited number of coefficients and sampling points in the wavefront, so that these equations can be expressed in vector-matrix notation (see next Section).

![Amplitude (left) and phase (right) of complex Zernike polynomials. The three examples correspond to \(C_6\), \(C_3\) and \(C_1\), respectively.](image)

**Figure 1** Amplitude (left) and phase (right) of complex Zernike polynomials. The three examples correspond to \(C_6\), \(C_3\) and \(C_1\), respectively.

### Implementation and results

#### Numerical methods

In the numerical implementation, we work with discrete (sampled) wavefronts so that the continuous expressions translate into a matrix-vector formulation. We applied a square sampling grid, and took 3720 points within a circle (34 samples along its radius.) The samples of the complex wavefront were arranged as the 3720 components of a column vector \(p\). In the series expansion of Eq. 11, we considered two cases with maximum order \(n = 7\), that is 36 polynomials or modes, and \(n = 12\), which means 91 modes. In each case we constructed a complex matrix \(C\), of 3720 \times 36 and 3720 \times 91 respectively. Then, Eq. 11 becomes \(p = \mathbf{C}b\) where \(b\) is another column vector formed by either 36 or 91 complex coefficients. To compute the coefficients \(b\) of the expansion we applied a standard least square fit to this strongly oversampled set. This is equivalent to apply the pseudoinverse of \(\mathbf{C}\) to the data:

\[
b = (\mathbf{C}^{\dagger})^{-1}\mathbf{C}^{\dagger}p \quad (13)
\]

Note that Eq. 12a means that vector \(p = t \cdot e^{i \varphi}\), where the dot product means element by element. Conversely \(a = \frac{1}{\kappa} (Z^{T}Z)^{-1}Z^{T}(\ln|\mathbf{C}b| - \ln|\mathbf{C}b|)\).

#### Stiles-Crawford apodization

The first numerical study consisted of representing the generalized pupil function (Eq. 1) of a group of 11 human eyes with CZPs with a variety of pupil sizes and RMS wavefront errors. The experimental wave aberration data \((W)\) were taken from a previous study, whereas a generic Gaussian model was used to describe the amplitude for all eyes:

\[
T(\rho, \theta) = 10^{-0.07\rho^{2}/2} \quad (14)
\]

where the argument is divided by 2 to pass from intensity to amplitude. (Note \(r\) is the physical radial coordinate is \(r = \rho R \ln m\) mm.) The quality of reconstructions with CZPs is high for most eyes. The RMS error for the wave aberration (phase) is of the order of \(10^{-4}\) wavelengths in all cases. The average is \(5 \times 10^{-4} \pm 2 \times 10^{-4}\) \(\lambda\) both for 36 and 91 modes.
Acknowledgments

We would like to thank the staff at the Massachusetts Eye and Ear Infirmary for their help and support during this study. We also acknowledge the contributions of Michael McGraw and William Collier for their technical assistance. This work was supported by grants from the National Institutes of Health (EY016453 and EY022783) and an unrestricted grant from Research to Prevent Blindness, Inc.

References

2. Lotmar, H. "Das Auge", 1911.
5. The role of the eye in the communication of visual information.

Figures

Figure 1: Schematic diagram of the human eye showing the location of the cornea, lens, and retina.

Figure 2: Photograph of a normal eye showing the cornea, iris, and pupil.

Figure 3: Photograph of a myopic eye showing the enlarged pupil and corneal curvature.

Figure 4: Photograph of a hyperopic eye showing compressed corneal curvature.

Figure 5: Photograph of an astigmatic eye showing a distorted pupil shape.

Figure 6: Photograph of a patient with spherical aberration showing a plano-convex lens.

Figure 7: Photograph of a patient with coma showing a field defect.

Figure 8: Photograph of a patient with astigmatism showing a elliptical pupil.

Figure 9: Photograph of a patient with irregular astigmatism showing a distorted pupil shape.

Figure 10: Photograph of a patient with irregular astigmatism showing a distorted pupil shape.

Figure 11: Photograph of a patient with irregular astigmatism showing a distorted pupil shape.

Figure 12: Photograph of a patient with irregular astigmatism showing a distorted pupil shape.

Figure 13: Photograph of a patient with irregular astigmatism showing a distorted pupil shape.

Figure 14: Photograph of a patient with irregular astigmatism showing a distorted pupil shape.

Figure 15: Photograph of a patient with irregular astigmatism showing a distorted pupil shape.
transmission is assumed to be one (maximum) within the aperture and zero outside. In particular we simulated annular (Fig. 4), semicircular (Fig. 5), elliptical or triangular (Fig. 6) pupil stops. The phase is pure coma in Figure 4 and pure spherical aberration in Figure 5. In Figure 6 only the amplitude is displayed as the phase reconstruction was similar to that of previous figures. We also simulated the case of Gaussian apodization (as in Section 3.2) for pure coma and pure spherical aberration.

The accuracy obtained for the reconstructions, in terms of the RMS difference between reconstructed and original data (amplitude and phase respectively) are listed in Table 1 for case (2), coma. The results for the other two cases (spherical aberration and real ocular wave aberration) are not included since they are totally equivalent (close RMS values). The phase reconstruction was good in all cases, no matter the type of amplitude function or pupil shape. The RMS phase error is of the order of $6 \times 10^{-4}$ ($\lambda$ units) independently from the number of modes considered, as in the former study. On the contrary, the reconstruction of the amplitude is strongly dependent on the initial transmittance function $T$. For the circle (trivial case, first column in Table 1) the reconstruction is basically perfect, since the circle is fully represented by the piston term $c_0 = 1$. When $T$ is Gaussian, the reconstruction is good, but improves further by increasing the number of modes (about three orders of magnitude when passing from 36 to 91 modes). This result is also consistent with that obtained with real eyes. In the other cases (annular, semicircle, ellipse, triangle) the number of modes used here seems insufficient to accurately represent the sharp edges of the stop. In fact, we can observe ringing or wavy-like artifacts in the reconstructions, which tend to improve by increasing the number of modes. The amplitude RMS errors are now of the order of $10^{-1}$ even for 91 modes. This seems the main limitation of this method: the reconstruction of sharp edges in the amplitude function requires higher order modes (frequencies), whereas the method seems to work well when both modulus and phase are smooth functions.

Discussion and conclusions

In this article we propose the use of complex Zernike polynomials to represent complex wavefronts, or generalized pupil functions, with free-from amplitude and phase. The main advantage is that the CZPs basis provides a unified framework as opposed to a series of ad hoc solutions published in the literature for each type of aperture. Our numerical results are highly satisfactory whenever both
amplitude and phase are smooth functions within the reference circle. This could be especially useful to represent pupil apodization, complex filters, or inhomogeneous beams (Gaussian, Bessel, etc.). Nevertheless, our results suggest that the number of modes (or order of polynomials) needed to obtain a high fidelity representation increases with the degree of complexity. A totally different behavior was found for amplitude and phase. The reconstruction of phase was always good (regardless of the number of modes), whereas the amplitude showed a strong dependency on both the number of modes and the amount of high order aberrations. A plausible explanation for the good phase reconstructions could be that the initial values of phase were also given in a pupil apodization, complex filters, or inhomogeneous beams (Gaussian, Bessel, etc.). Nevertheless, our results suggest that the number of modes (or order of polynomials) needed to obtain a high fidelity representation increases with the degree of complexity. A totally different behavior was found for amplitude and phase. The reconstruction of phase was always good (regardless of the number of modes), whereas the amplitude showed a strong dependency on both the number of modes and the amount of high order aberrations. A plausible explanation for the good phase reconstructions could be that the initial values of phase were also given in a pupil apodization, complex filters, or inhomogeneous beams (Gaussian, Bessel, etc.).

Table 1  RMS reconstruction errors for the different pupil transmissions and for 1 λ of pure coma. The errors for other aberrations are totally equivalent

<table>
<thead>
<tr>
<th></th>
<th>Circle</th>
<th>Gaussian</th>
<th>Annular</th>
<th>Semicircle</th>
<th>Ellipse</th>
<th>Triangle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplitude</td>
<td>36 modes</td>
<td>1.9 × 10⁻⁴</td>
<td>2.6 × 10⁻³</td>
<td>1.5 × 10⁻¹</td>
<td>1.5 × 10⁻¹</td>
<td>1.7 × 10⁻¹</td>
</tr>
<tr>
<td></td>
<td>91 modes</td>
<td>2.5 × 10⁻⁵</td>
<td>3.1 × 10⁻⁶</td>
<td>1.0 × 10⁻¹</td>
<td>1.2 × 10⁻¹</td>
<td>1.3 × 10⁻¹</td>
</tr>
<tr>
<td>Phase</td>
<td>36 modes</td>
<td>5.7 × 10⁻⁴</td>
<td>5.7 × 10⁻⁴</td>
<td>5.8 × 10⁻⁴</td>
<td>5.7 × 10⁻⁴</td>
<td>6.2 × 10⁻⁴</td>
</tr>
<tr>
<td></td>
<td>91 modes</td>
<td>5.7 × 10⁻⁴</td>
<td>5.7 × 10⁻⁴</td>
<td>5.7 × 10⁻⁴</td>
<td>5.7 × 10⁻⁴</td>
<td>6.3 × 10⁻⁴</td>
</tr>
</tbody>
</table>
the optical design of advanced optical elements. The implementation and practical applications of Eq. 15 will be the subject of future work.

**Acknowledgements**

This work was supported by the Comisión Interministerial de Ciencia y Tecnología, Spain, under Grant FIS2008-00697.

**Conflict of interests**

Authors declare that they don’t have any conflict of interests.

**References**