Journal of the European Optical Society - Rapid Publications **2**, 07012 (2007) www.jeos.org

Computing Zernike polynomials of arbitrary degree using the discrete Fourier transform

Augustus J. E. M. Janssen Philips Research Europe, HTC 36
a.j.e.m.janssen@philips.com

Peter Dirksen Philips Research Europe, HTC 4
peter.dirksen@philips.com

The conventional representation of Zernike polynomials $R_n^m(\rho)$ gives unacceptable numerical results for large values of the degree n. We present an algorithm for the computation of Zernike polynomials of arbitrary degree n. The algorithm has the form of a discrete Fourier (cosine) transform which comes with advantages over other methods in terms of computation time, accuracy and ease of implementation. As an application we consider the effect of NA-scaling on the lower-order aberrations of an optical system in the presence of a very high order aberration. [DOI: 10.2971/jeos.2007.07012]

Keywords: Zernike polynomial, aberration, high-order, discrete Fourier transform, NA-scaling

1 INTRODUCTION

The (radial part of the) Zernike polynomials $R_n^m(\rho)$ are widely used in the representation of the aberrations of optical systems and in the computation of the diffraction integral defining the point-spread function of these systems [1]-[5].When we are dealing with smooth exit pupil functions, it is, in general, sufficient to consider the R_n^m for modest values of the degree n and azimuthal order m, say, $n+m \leq 12$. For such pupil functions, the conventional polynomial representation [1]

$$R_n^m(\rho) = \\ = \sum_{s=0}^{(n-m)/2} \frac{(n-s)!\,(-1)^s}{\left(\frac{n-m}{2}-s\right)!\left(\frac{n+m}{2}-s\right)!\,s!}\,\rho^{n-2s}\ ,\ 0 \leq \rho \leq 1\ , \tag{1}$$

can be used to calculate the Zernike polynomials. Some low order Zernike polynomials are shown in Table 1.

Degree n	m	$R_n^m(\rho)$
0	0	1
1	1	ρ
2	0	$2\rho^2-1$
2	2	ρ^2
4	0	$6\rho^4-6\rho^2+1$
3	1	$3\rho^3-2\rho$
3	3	ρ^3
	⋮	

TABLE 1 Low order Zernike polynomials

In the case that the exit pupil function contains discontinuities, or is roughly behaved in a more general sense, it is necessary to consider Zernike polynomials of much higher degree and order. For instance, when the pupil function has a central obstruction (0 and 1 on two concentric sets $0 \leq \rho < a$ and $a \leq \rho \leq 1$), the coefficient of $R_n^0(\rho)$ in the Zernike expansion of the pupil function can be shown to decay only like $n^{-1/2}$. Then Eq. (1) becomes cumbersome because of the high-order factorials that are required. Also, for $m=0$, it can be shown from the Stirling's formula that the largest coefficient of ρ^{n-2s} occurring in the series in Eq. (1) behaves like $(1+\sqrt{2})^n$. Accordingly, when computing with d digits, Eq. (1) produces errors of the order of unity or larger from $n = d/log(1+\sqrt{2})$ onwards. Hence, for the commonly used 15 digits precision, one has serious problems from $n=40$ onwards, as shown in Figure 1, top, right. An alternative to compute Zernike polynomials is to use recursions for them such as those found in Ref. [6]. These recursion schemes are, however, computationally more expensive and less direct than a formula like Eq. (1) and their accuracy due to error propagation is also an issue.

In this paper, we present a new computation scheme in which one can allow degrees as large as 10^5 without problems. This new algorithm is of the discrete-cosine transform (DCT) type, and is direct and transparent. Furthermore, the computation can be done using the FFT-algorithm which comes with the following advantages [7, 8]:

- Very favorable and well-established accuracy
- Simultaneous computation of all Zernike polynomials of the same degree n in as few as $O(n\,logn)$ operations

As an application we consider the effect of NA-scaling on the lower-order aberrations of an optical system in the presence of a very high order aberration. For this we use a recently found

Received March 16, 2007; published April 12, 2007 ISSN 1990-2573

formula [9], entirely in terms of Zernike polynomials, for the Zernike coefficients of scaled pupils.

2 DCT FORMULA FOR ZERNIKE POLYNOMIALS

In Appendix A we show that

$$R_n^m(\rho) = \frac{1}{N}\sum_{k=0}^{N-1} U_n\left(\rho\cos\frac{2\pi k}{N}\right)\cos\frac{2\pi mk}{N}\,, \qquad 0\le\rho\le 1\,, \tag{2}$$

where N is any integer $> n+m$. In Eq. (2) we have integer $n, m \geq 0$ with $n-m$ even and ≥ 0 (as usual), and U_n is the Chebyshev polynomial of the second kind and of degree n. No matter how large n is, the evaluation of $U_n(x)$ is no problem since we have

$$U_n(x) = \frac{\sin(n+1)\,v}{\sin v}\,, \qquad x = \cos v. \tag{3}$$

Eq. (2) is a consequence of the formula (A.10) that represents $R_n^m(\rho)$ as an integral of a trigonometric polynomial of degree $n+m$ over a periodicity interval. Such an integral can be computed error-free as a series when the number of equidistant sample points N exceeds the degree $n+m$. It also follows from this that the right-hand side of Eq. (2) (and that of Eq. (A.10)) vanish when $m > n$ or when n and m have different parity (with again $N > n+m$).

Eq. (2) gives $R_n^m(\rho)$ as the m^{th} component of the DCT of the sequence $(U_n(\rho\cos 2\pi k/N))_{k=0,1,\dots,N-1}$, hence we get all $R_n^m(\rho)$, with $m \geq 0$ and $m = n, n-2, \dots$, using $O(N\log N)$ operations. Since $m \leq n$, it is sufficient to take N any integer $> 2n$.

In Figure 1, top, we show $R_n^m(\rho)$ as a function of ρ, $0\le\rho\le 1$, computed according to Eq. (1) and Eq. (2), using 16 decimal places, for $m = 17$, $n = 39$ and for $m = 0$, $n = 50$. We see that Eq. (1) gives unacceptable results for the case $m = 0$, $n = 50$ from $\rho = 0.8$ onwards.

In Figure 1, bottom, we show $R_n^m(\rho)$, computed according to Eq. (2), with $m = 0$ and $n = 10000$ and ρ very close to 1. We see that $R_n^m(\rho = 1) = 1$ which is in agreement with the theory [1].

3 AN APPLICATION: HIGH-ORDER ABERRATIONS AND SCALING

In lithographic imaging systems, the numerical aperture (NA) is varied intentionally below its maximum value so as to optimize the performance for the particular object to be imaged. In Ref. [9] the effect of NA-scaling on the Zernike coefficients describing the optical system has been concisely expressed in terms of Zernike polynomials. Thus we consider a pupil function

$$P(\rho,\vartheta) = \exp\{i\,\Phi(\rho,\vartheta)\}\,, \qquad 0\le\rho\le 1\,,\ 0\le\vartheta\le 2\pi\,, \tag{4}$$

in polar coordinates with real phase Φ, and we assume that Φ is expanded as a Zernike series according to

$$\Phi(\rho,\vartheta) = \sum_{n,m} \alpha_n^m\, R_n^m(\rho)\cos m\vartheta\,. \tag{5}$$

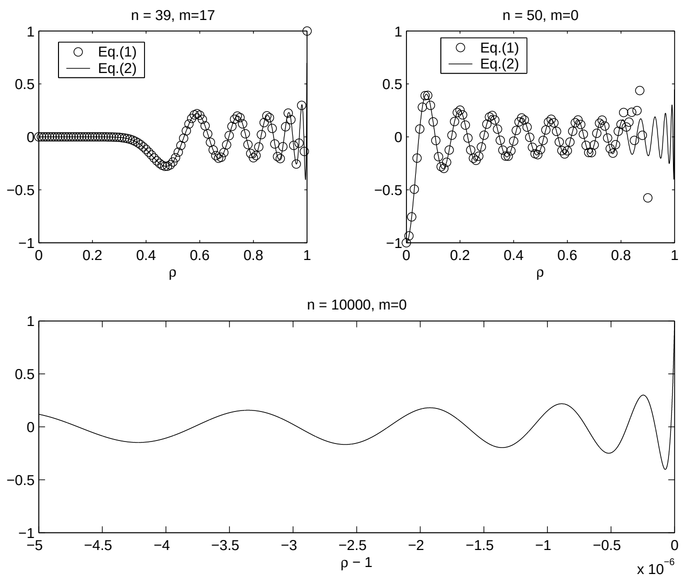

FIG. 1 **Top**: $R_n^m(\rho)$ as a function of ρ, $0\le\rho\le 1$, computed according to Eq. (1) and Eq. (2), using 16 digits, for $m = 17$, $n = 39$ and for $m = 0$, $n = 50$. **Bottom**: $R_n^m(\rho)$, computed according to Eq. (2), with $m = 0$ and $n = 10000$ and ρ very close to 1

Scaling to a pupil with relative size $\varepsilon = \mathrm{NA}/\mathrm{NA}_{\max} \leq 1$ requires computation of the Zernike coefficients $\alpha_n^m(\varepsilon)$ of the scaled phase function $\Phi(\varepsilon\rho,\vartheta)$. For $m = 0, 1, \dots$ the $\alpha_n^m(\varepsilon)$ are given in terms of the α_n^m as

$$\alpha_n^m(\varepsilon) = \sum_{n'} \alpha_{n'}^m\,[R_{n'}^n(\varepsilon) - R_{n'}^{n+2}(\varepsilon)]\,, \qquad n = m, m+2, \dots\,, \tag{6}$$

where the summation is over $n' = n, n+2, \dots$ ($R_n^{n+2} \equiv 0$). In case of a non-smooth phase function Φ, one should expect significant values of $\alpha_{n'}^m$ for very high degrees n'. Also, scaling is normally done using values of ε close to its maximum 1, where Eq. (1) produces the largest numerical error. Thus, formula (6) is not practicable in these cases when Eq. (1) is used to evaluate $R_{n'}^n(\varepsilon) - R_{n'}^{n+2}(\varepsilon)$, but becomes so when Eq. (2) is used instead.

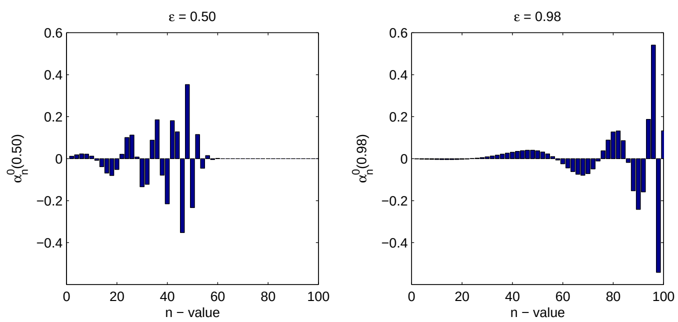

FIG. 2 The disturbance $\alpha_n^0(\varepsilon)$ of the aberration of order $n = 0, 2, \cdots, 100$, due to the presence of an aberration of amplitude 1 and of the order $n' = 100$ when the system is scaled to relative size $\varepsilon = 0.50$ (left) and $\varepsilon = 0.98$ (right).

As an example, we consider the effect of a *single* high order aberration term $\alpha_{n'}^m$ on the totality of α_n^m with $n = m, m+2, \dots, n'$ while scaling to relative size ε. We take $\alpha_n^m = 0$ for $n \neq n'$ and $\alpha_{n'}^m = 1$, and get from Eq. (6)

$$\alpha_n^m(\varepsilon) = [R_{n'}^n(\varepsilon) - R_{n'}^{n+2}(\varepsilon)],\quad n = m, m+2, \dots, n', \tag{7}$$

while $\alpha_n^m(\varepsilon) = 0$ when $n > n'$. The numbers $R_{n'}^n(\varepsilon) - R_{n'}^{n+2}(\varepsilon)$ required in Eq. (7), with n' fixed and $n = m, m+2, ..., n'$, can be computed simultaneously using $O(n' \log n')$ operations by employing Eq. (2) in its DCT-mode. Figure 2 shows the result for $m = 0$ and $n' = 100$, $n = 0, 2, \cdots, 100$, $\alpha_n^0(\varepsilon)$ with $\varepsilon = 0.50$ and $\varepsilon = 0.98$.

A PROOF OF THE MAIN RESULT

We write for integer $n, m \geq 0$ with $n - m$ even and ≥ 0

$$z_n^m(\nu, \mu) = Z_n^m(\rho, \vartheta) = R_n^m(\rho) \cos m\vartheta \,, \qquad \text{(A.1)}$$

in which the Cartesian coordinates ν, μ and polar coordinates ρ, ϑ are related according to $\nu = \rho \cos \vartheta$, $\mu = \rho \sin \vartheta$ and $0 \leq \rho \leq 1$, $0 \leq \vartheta \leq 2\pi$. Furthermore, we let

$$f_n^m(\nu) = \frac{1}{2(1-\nu^2)^{1/2}} \int\limits_{-\sqrt{1-\nu^2}}^{\sqrt{1-\nu^2}} z_n^m(\nu, \mu)\, d\mu \,, \qquad -1 \leq \nu \leq 1 \,. \qquad \text{(A.2)}$$

According to the formula for the Radon transform of Z_n^m we have, see Ref. [10], Eq. (8.13.17),

$$f_n^m(\nu) = \frac{1}{n+1} U_n(\nu) \,, \qquad -1 \leq \nu \leq 1 \,, \qquad \text{(A.3)}$$

with the Chebyshev polynomial U_n given in Eq. (3). We consider next the Zernike expansion of $f_n^m(\nu)$,

$$f_n^m(\nu) = \sum_{n',m'} \beta_{n,n'}^{m,m'}\, z_{n'}^{m'}(\nu, \mu) \qquad \text{(A.4)}$$

in which the β's are given, due to the orthogonality[1] of the z's, by

$$\beta_{n,n'}^{m,m'} = \frac{(n'+1)\varepsilon_{m'}}{\pi} \iint\limits_{\nu^2+\mu^2 \leq 1} f_n^m(\nu)\, z_{n'}^{m'}(\nu, \mu)\, d\nu\, d\mu \,. \qquad \text{(A.5)}$$

In Eq. (A.5) we have that m, m', n, n' all have the same parity and $n' \geq m'$, and $\varepsilon_{m'} = 1$ for $m' = 0$ and $\varepsilon_{m'} = 2$ for $m' = 1, 2, ...$ (Neumann's symbol). According to Eq. (A.3) we have

$$\beta_{n,n'}^{m,m'} = \frac{(n'+1)\,\varepsilon_{m'}}{(n+1)\,\pi} \int\limits_{-1}^{1} U_n(\nu) \left(\int\limits_{-\sqrt{1-\nu^2}}^{\sqrt{1-\nu^2}} z_{n'}^{m'}(\nu, \mu)\, d\mu \right) d\nu \,. \qquad \text{(A.6)}$$

Then using Eqs. (A.2), (A.3) with n', m' instead of n, m, we find

$$\beta_{n,n'}^{m,m'} = \frac{2\varepsilon_{m'}}{\pi(n+1)} \int\limits_{-1}^{1} U_n(\nu)\, U_{n'}(\nu)(1-\nu^2)^{1/2}\, d\nu = \frac{\varepsilon_{m'}}{n+1} \delta_{n,n'} \,, \qquad \text{(A.7)}$$

where δ denotes Kronecker's delta, and where we have used the orthogonality of the U's, see Ref. [11], 22.2.5 on p. 774.

We conclude from Eqs. (A.3), (A.4), (A.7) that

$$U_n(\nu) = \sum_{m'} \varepsilon_{m'}\, z_n^{m'}(\nu, \mu) \,, \qquad \text{(A.8)}$$

i.e., that

$$U_n(\rho \cos \vartheta) = \sum_{m'} \varepsilon_{m'}\, R_n^{m'}(\rho) \cos m'\vartheta \,. \qquad \text{(A.9)}$$

By orthogonality of the $\cos m'\vartheta$, $\vartheta \in [0, 2\pi]$, it follows that

$$R_n^m(\rho) = \frac{1}{2\pi} \int\limits_0^{2\pi} U_n(\rho \cos \vartheta) \cos m\vartheta\, d\vartheta \,. \qquad \text{(A.10)}$$

Finally, $U_n(\rho \cos \vartheta) \cos m\vartheta$ is a trigonometric polynomial of degree $n + m$. Therefore, the integral in Eq. (A.10) can be evaluated using the sample values of the integrand at the points $2\pi k/N$, $k = 0, 1, ..., N-1$ when $N > n + m$. This yields Eq. (2).

Note. Eq. (A.10) can also be used to get accurate stationary phase approximations to $R_n^m(\rho)$ when n gets large. Accordingly, $R_n^m(\rho)$ is vanishing small when $0 \leq \rho \leq \frac{m}{n} - \varepsilon$, and is oscillatory and of amplitude $O(n^{-1/2})$ when $\frac{m}{n} + \varepsilon \leq \rho \leq 1 - \varepsilon$ ($\varepsilon > 0$ fixed, $n \to \infty$). This can be used to explain some of the phenomena that one observes in Figure 2.

ACKNOWLEDGEMENTS

The authors thank the various referees of this paper for their valuable and constructive comments.

References

[1] M. Born and E. Wolf, *Principles of Optics* (7th rev. ed., Ch. 9 and App. VII, Cambridge University Press, Cambridge, MA, 2001).

[2] D. Malacara, *Optical Shop Testing* (2nd ed., Wiley, 1992).

[3] V. N. Mahajan, "Zernike circle polynomials and optical aberrations of systems with circular pupils" Appl. Optics. **33**, 8121–8124 (1994).

[4] R. J. Noll, "Zernike polynomials and atmospheric turbulence" J. Opt. Soc. Am. **66**, 207–211 (1976).

[5] See http://www.nijboerzernike.nl .

[6] A. Prata and W. V. T. Rusch, "Algorithm for computation of Zernike polynomials expansion coefficients" Appl. Optics **28**, 749–754 (1989).

[7] D. Calvetti, "A stochastic roundoff error analysis for the Fast Fourier Transform" Math. Comput. **56**, 755–774 (1991).

[8] See http://en.wikipedia.org/wiki/Fast˙Fourier˙transform .

[9] A. J. E. M. Janssen and P. Dirksen, "Concise formula for the Zernike coefficients of scaled pupils", J. Microlith. Microfab. Microsyst. **5**, 030501 (2006).

[10] S. R. Deans, *The Radon transform and some of its applications* (Wiley, New York, 1983, Sec. 7.6, Eq. 6.11).

[11] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970).