Optical surface reconstruction technique through combination of zonal and modal fitting

Julián Espinosa
David Mas
Jorge Pérez
Carlos Illueca

Universidad de Alicante
Departament de Optica Campus de San Vicente del Raspeig
P.O. Box 99
Alicante, 03690 Spain

Abstract. Videokeratometers and Scheimpflug cameras permit accurate estimation of corneal surfaces. From height data it is possible to adjust analytical surfaces that will be later used for aberration calculation. Zernike polynomials are often used as adjusting polynomials, but they have shown to be not precise when describing highly irregular surfaces. We propose a combined zonal and modal method that allows an accurate reconstruction of corneal surfaces from height data, diminishing the influence of smooth areas over irregular zones and vice versa. The surface fitting error is decreased in the considered cases, mainly in the central region, which is more important optically. Therefore, the method can be established as an accurate resampling technique.

Keywords: Zernike polynomials; wavefront error; surface reconstruction; corneal irregularities.

Paper 09447RR received Oct. 5, 2009; revised manuscript received Feb. 25, 2010; accepted for publication Feb. 26, 2010; published online Apr. 16, 2010.

1 Introduction

In the last few years, a large number of accurate devices for measuring corneal morphology have appeared. Among them, those based on Placido rings and Scheimpflug cameras are the most widely used. These devices can provide height data of the corneal surfaces that can be fitted to analytical surfaces through polynomial expressions (modal approach). These can be later used for aberration calculation and ray trace solving.

Zernike polynomials are often used as adjusting polynomials, but they have shown to be not precise when describing highly irregular corneal surfaces, since they do not fit properly to irregularities. In 2003, Smolek and Klyce concluded that Zernike polynomials did not fully characterize the surface features that affect vision. Later, Klyce, Karon, and Smolek found that using Zernike polynomials in normal eyes was acceptable, but for eyes after corneal surgery or eyes with corneal pathology such as keratoconus, the Zernike method fails to capture all clinically significant higher-order aberrations. Furthermore, it is very difficult to assess a priori how many terms are necessary to achieve an acceptable accuracy in the Zernike reconstruction for any given corneal shape.

Common measuring devices are based in imaging techniques. An image reflected from a pattern—linear or annular—is projected onto a charge-coupled device (CCD) and is digitally processed. From this information, a height map is obtained. Surfaces with radial structure do not fit well with squared detectors or matrices, since density of samples in the azimuthal coordinate depends on the radial coordinate, due to the fact that the number of different angular samples in the center of the matrix is lower than in the periphery. Unfortunately, acquiring devices are based on squared CCD, and little can be done to change this fact. Reconstruction methods may consider such limitations and obtain the best possible results. Traditional reconstruction of the optical surface with Zernike polynomials does not pay attention to the particular sampling distribution of the analyzed surfaces, and thus it does not provide optimum results in the central area around the apex, which is the most important zone under the optical point of view. One of the main limitations of the polynomial representation is that each term extends its influence over the entire pupil.

The goodness of fit of a surface model strongly depends on the number of samples and its distribution. Let us consider a spherical-like surface given by height samples in Cartesian coordinates. The number of points that contribute to the polynomial fit is proportional to the area of the considered surface. Thus the ratio between the number of samples included in an external annular region defined between radii \( R_1 \) and \( R_2 \) and the central part of aperture radius \( R_1 \) is quadratic, i.e., \( (R_2/R_1)^2-1 \). The effect is not very important for a spherical surface, but if the surface is affected by any irregularity in the periphery, it determines the coefficients in the central optical area, although the real effect of the irregularity on the retinal image is limited. Therefore, the fitting error is distributed over the entire surface, not just at those zones where the deformation is located.

B-spline polynomials are especially well suited for fitting complex-shaped surfaces. They have the advantages of being locally defined and having great flexibility, which allows control of their smoothness and polynomial degree. In Ref. 25, the authors propose an alternative technique for reconstructing the corneal shape from elevation data using sets of radial basis functions (RBFs) in a modal approach.
Due to the fast decay of RBFs, the modal approach exhibits features of the zonal approach, eventually capturing small deformations of the surface, which are missed by the polynomial fitting. As Ref. 25 states, we think that there is no unique or best approach to corneal surfaces, and a combination of techniques can be a good strategy.

Light distributions propagated from the cornea can be numerically evaluated by means of a Fresnel propagation algorithm. Numerical calculation of convergent Fresnel patterns through fast Fourier transform usually requires a large number of samples ($4 \times 10^6$) to fulfill the Nyquist condition around the focus. In Ref. 26 the authors present a simple method that permits subsampling the Fresnel pattern while maintaining the Nyquist condition, and thus preventing the appearance of aliasing effects in the calculation. Later, in Ref. 29 the authors proposed modifying the initial wavefront and relaxing the Nyquist condition, thus giving a more efficient numerical algorithm. Corneal height data sampling provided by corneal topographers is still lower than the one imposed by the Nyquist condition. Hence, we need a method that allows an accurate resample of corneal height data.

In this work, we propose a zonal Zernike fitting (combination of zonal and modal approaches) of corneal height data. It consists of obtaining Zernike coefficients over local areas. A local mask is applied on the surface, and polynomial fit is done for this part of the surface. The coefficients so obtained have no physical meaning due to the fact that Zernike polynomials are usually defined over a unit disk. However, we used the method as a reconstruction tool in a way similar to Lundström, Unso, and Gustafsson. The mask is then displaced over the surface and the process is repeated on each local area. The surface reconstruction takes into account the fittings in all zones, diminishing the influence of smooth areas over irregular ones and vice versa. The method limits the influence of the peripheral irregularities over the central corneal area calculation, thus giving accurate reconstruction of the central optical zone. This fact will be of special interest in the evaluation of wavefront aberrations of irregular corneas (e.g., keratoconus, pellucid marginal corneal degeneration, corneal etc.). Thus, we join the advantage of a simpler analytical form of the modal representation together with the benefit of local definition of the zonal one. We have compared zonal Zernike fitting with the modal approach, evaluating both techniques over two different surfaces: a theoretical irregular surface and height data of a real keratoconic cornea.

2 Methods

Zernike polynomials are often used as an expansion of corneal height data and to analyze of optical wavefronts, being low order terms directly related to classical Seidel aberrations. As a complete modal set, any surface can be approximated by a linear combination of circular polynomials as follows:

\[ W(\rho, \theta) = \sum_{j=0}^{p-1} c_j Z_j(\rho, \theta) = G(\rho, \theta); \quad j = \frac{1}{2}[n(n+2)+m], \]

where \( n \) and \( m \) are the radial order and the azimuthal frequency, respectively, \( p \) is the number of terms in the expansion, \( c_j \) are the Zernike coefficients associated with their Zernike polynomial \( Z_j(\rho, \theta) \), \( \rho \) is the normalized distance from the origin, and \( \theta \) is the angle.

The modal fitting principle of the Zernike representation is shown in Eq. (1). The data \( W(\rho, \theta) \) are approximated by a polynomial function extended over the whole domain (Fig. 1). Our basic equation is \( W = Zc \), which is a linear transformation on \( c \), where \( c \) is the expansions coefficients, \( W \) is a discrete set of elevation data, and \( Z \) is a matrix of discrete Zernike polynomials. The problem is choosing the coefficients \( c \) to minimize, in some sense, the difference between the observed \( W \) discrete elevations and the prescribed output of the system \( Zc \). The function to be minimized is

\[ \phi(c) = c^T c + t^T (W - Zc), \]

where \( t \) is the vector of Lagrange multipliers and the super index \( T \) means matrix transposition. The minimization yields

\[ c = Z^+ W, \]

where the super index+ stands for matrix pseudoinverse, developed by Moore and later by Penrose. To calculate the pseudoinverse, we used a method called singular value decomposition (SVD), implemented in MATLAB based in linear algebra package (LAPACK) routines. This method provides a solution to Eq. (3) regardless of being a determined, undetermined, or overdetermined system.

Zonal Zernike fitting is developed in the same way as a modal one, but the equation \( W = Zc \) is not solved over the whole domain. The surface is divided into zones that overlap each other (see Video 1). It takes advantage of increasing the influence of the central zone in the adjustment, thus providing a better description of the more interesting optical area. On the contrary, in the modal fitting, periphery data have a heavier influence than central ones due to the fact that they are more numerous.

Fig. 1 Area of the height data where modal fitting is obtained.
Let us call $W(u,v)$ a surface described by an $N \times N$ matrix, with $u$ and $v$ being discrete indexes from 1 to $N$. A local area of $M \times M$ pixels, being $M < N$, is selected in this matrix where we solve Eq. [1]. Zernike decomposition for this region is obtained, thus having a local description of the surface. The calculation area is displaced along the entire surface, and a zonal Zernike polynomial fit is obtained each time. Reconstruction on one point is done by evaluating all Zernike decompositions, describing the local areas that overlap on this point and taking the mean value of the obtained height. Calculation windows for selecting local regions are implemented by constructing auxiliary matrices of size $N \times N$, whose elements are equal to one in the region of interest and zero elsewhere. These matrices are described by:

$$O_{a,b}(u,v) = \begin{cases} 1: & u \in [a,a+(M-1)] \land v \in [b,b+(M-1)] \land a \leq N-1, b \leq N-1, \land (a+b) \leq N-1 \land (a+b) \leq (M-1) \land (a+b) \leq N-1 \land (a+b) \leq (M-1) \\ 0: & \text{otherwise} \end{cases}$$

(4)

where $1 \leq a \leq N-(M-1)$, and $1 \leq b \leq N-(M-1)$. The Hadamard product of matrices $O_{a,b}(u,v)$ with the initial one $[W(u,v)]$ will select the local areas. Mathematically, auxiliary matrices that contain the local fittings are described by $W_{a,b}(u,v)$. Thus, Zernike decomposition over the local domain defined by $M$ can be written as:

$$W_{a,b}(u,v) = \sum_{j=0}^{p-1} c_{j}^{[a,b]}Z_{j}[\rho_{(a,b)}, \theta_{(a,b)}]; \quad u \in [a,a+(M-1)] \land v \in [b,b+(M-1)]$$

(5)

with $[\rho_{(a,b)}, \theta_{(a,b)}]$ being the polar coordinates of element in $(u,v)$ of the global matrix. From Eq. [5] it can be seen that the Zernike fitting is performed over a square mask, thus a complication arises, since Zernike polynomials are orthogonal on the unit disk. In a way similar to Lundström, Unsbo, and Gustafsson,\textsuperscript{3} who proposed a direct reconstruction method for elliptical pupils, we have solved the problem. The region of missing data between the mask and the whole pupil can be ignored, provided there are enough valid measured heights to perform a fit. Once the coefficients are determined, the surface can be reconstructed over the entire domain of the pupil, even though valid data used to compute the coefficients are available only in the limited area of the mask. There will thus be an extrapolated part of the surface outside the mask that has no physical relevance. When the reconstructed surface is evaluated, this part is removed.

As we said before, one point of the surface may belong to different local regions. Thus, it is necessary to take into account how many regions overlap on a single pixel. According to our previous description, the reconstructed surface can be written as

$$L(u,v) = \frac{\sum_{a=1}^{N-(M-1)} \sum_{b=1}^{N-(M-1)} W_{a,b}(u,v)}{\sum_{a=1}^{N-(M-1)} \sum_{b=1}^{N-(M-1)} O_{a,b}(u,v)}$$

(6)

We would like to point out that the surface can be interpolated at any $(u,v)$ point, now expressed like real numbers. The reader should also note that central area samples are included in more local matrices than those in the periphery, and thus calculation becomes more accurate. According to the process just described, notice that all elements integrated into the submatrix defined between elements $(M,M)$ and $(N-M+1,N-M+1)$ are evaluated $M^2$ times. The existence of such a central matrix, which contains $[N-2(M-1)]^2$ elements, requires $M \leq (N+1)/2$. Selection of an optimum value for $M$ has been done, balancing the size of the local kernel and the number of elements in the central zone, i.e.:

$$f_{N}(M) = M^2[N-2(M-1)]^2.$$

(7)

From the Eq. [7] derivative, we obtained a maximum that provide an optimum value for $M$. 

Video 1 Areas of the height data where zonal Zernike fitting is obtained. Note the overlap of masks (QuickTime, 823.5 KB). [URL: http://dx.doi.org/10.1117/1.3394260.1]
\[
M_{opt} = \text{round}\left(\frac{N + 2}{4}\right).
\] (8)

Traditionally, conical curves are used for approximating ophthalmic surfaces in the eye. The method just presented has been evaluated over two different surfaces. On the one hand there is a theoretical irregular surface composed of a sphere plus a centered double-peak surface with a single hole, known as Franke’s function [Eq. (9)], which has been extensively used by researchers to test new fitting schemes, as a highly complex waveform:

\[
K(x, y) = \frac{3}{4} \exp\left\{-\frac{\left((9x - 2)^2 + (9y - 2)^2\right)}{4}\right\}
+ \frac{3}{4} \exp\left\{-\frac{\left((9x + 1)^2 + (9y + 1)^2\right)}{49} + \frac{\left((9x - 1)^2 + (9y - 1)^2\right)}{10}\right\}
+ \frac{1}{2} \exp\left\{-\frac{\left((9x - 7)^2 + (9y - 7)^2\right)}{4}\right\}
+ \frac{1}{5} \exp\left\{-\frac{\left((9x - 4)^2 + (9y - 7)^2\right)}{9}\right\}.
\] (9)

On the other hand, we have evaluated height data of a real keratoconic cornea measured with a Pentacam (Oculus Optikgeräte GmbH, Wetzlar, Germany). Evaluation of the goodness of reconstruction is done through the minus logarithm of the root mean square deviation (/RMSD):

\[
/RMSD = -\log\left(\frac{\sum_{u=1}^{N} \sum_{v=1}^{N} \left|F(u,v) - W(u,v)\right|^2}{N^2}\right)^{1/2},
\] (10)

where \(F(u,v)\) stands for the reconstructed surface points using modal, \(F(u,v) = G(u,v)\), or zonal, \(F(u,v) = L(u,v)\), Zernike fitting, and \(W(u,v)\) are the surface height data.

3 Results

Obtained /RMSDs for the height data of a real keratoconic cornea with a pupil diameter of 4 mm have been calculated and are shown in Fig. 2. We have applied modal and zonal Zernike fittings using the same masks as before. Again, it is shown that zonal Zernike fitting provides a better adjustment than the global one, as we saw in the theoretical surface case.

Contrary to what one may think, increasing the number of terms in the Zernike polynomial decompositions does not give a more realistic approximation in numerical image analysis. In practice, the SVD method can always provide a unique solution, although the matrix \(Z^*\) is ill-conditioned and numerically rank deficient. This rank deficiency may be diagnosed through the condition number. The problem arises when the condition number is too large, because the SVD provides a bad least-squares solution. This fact can be seen in the falls of the /RMSD plotted in Fig. 6. There is a number of polynomials from which the solution obtained is degraded.

Fig. 2 Reconstruction error between the Franke’s plus sphere surface and the reconstructed surfaces. /RMSD for modal and zonal Zernike fitting with a 21×21 and 11×11 px mask.

Fig. 3 Difference between the Franke’s plus sphere surface and the reconstructed surface obtained using the zonal Zernike fitting that provides the best adjustment for 11×11 px (43 polynomials).
Figures 7–9 present the height differences between real keratoconic corneal data obtained from the Pentacam and the reconstructed ones using the polynomial decomposition that provides the best adjustment for each method. All differences are set to the same scale to emphasize the influence of zonal Zernike fitting. Results are quite similar to those obtained for the simulation. Again, results show that the keratoconic corneal height description through Zernike polynomials over the whole surface produces fitting errors in the whole area, whereas the zonal Zernike analysis proposed here provides a lower surface fitting error, mainly in the central region (see Fig. 7), which is optically the most important.

4 Conclusion

In this work, we develop a zonal Zernike fitting of corneal height data. It consists of obtaining Zernike coefficients over local areas. A local displacing mask is applied on the surface, and polynomial fit is done for this part of the surface. The surface reconstruction takes into account the fittings in all the zones, thus diminishing the influence of smooth areas over irregular ones and vice versa. Zonal Zernike fitting has the same advantage of the simple analytical form of the modal representation, together with the benefit of local definition of the zonal one. If we compare for the same polynomial order (28 coefficients) our results with those presented in Ref. 24, they achieve a RMSD around $3 \times 10^{-3}$ mm for an irregular surface (modeling a keratoconic cornea) using radial basis functions, while our analysis on a similar case (see Fig. 6) shows that zonal Zernike fitting reaches between 2 and 4 orders of magnitude lower, depending on the size of the zonal mask we used. Furthermore, we compare our technique with that proposed in Ref. 22 on the same complex surface.
Franke’s function), and we find that the rms of the reconstructed surface is four times lower in our model. Regarding the computational cost of our method, it is at least 2 orders of magnitude higher than the traditional Zernike calculation over the entire surface.

Our goal is to propose a combination of techniques that could be a good strategy for fitting, reconstructing, and resampling corneal height data. Local fitting reduces the influence of regions far from the calculation point. The implementation of the method also shows that the central corneal surface part is better evaluated than the outer parts, since calculation is more intensive in this zone and is not affected by peripheral irregularities. Thus, the presented method allows for better reconstruction of optical surfaces, mainly of the optical central zone, than the traditional one based on polynomial fitting of the whole surface.

Acknowledgment
The authors acknowledge support of the Consellería de Educación project GV/2009/002 from the Generalitat Valenciana.

References

23. M. Ares and S. Royo, “Comparison of cubic B-spline and Zernike-