

An adaptive algorithm for the cornea modeling from keratometric data

Andrei Martínez-Finkelshtein*

*Department of Statistics and Applied Mathematics, University of Almería, Spain, and
Instituto Carlos I de Física Teórica y Computacional, Granada University, Spain*

Darío Ramos-López

Department of Statistics and Applied Mathematics, University of Almería, Spain

Gracia M. Castro-Luna

Corporación VISSUM Almería, Spain

Jorge L. Alió

Corporación VISSUM Alicante, Spain

Abstract

In this paper we describe an adaptive and multi-scale algorithm for the parsimonious fit of the corneal surface data that allows to adapt the number of functions used in the reconstruction to the conditions of each cornea. The method implements also a dynamical selection of the parameters and the management of noise. It can be used for the real-time reconstruction of both altimetric data and corneal power maps from the data collected by keratoscopes, such as the Placido rings based topographers, decisive for an early detection of corneal diseases such as keratoconus.

Numerical experiments show that the algorithm exhibits a steady exponential error decay, independently of the level of aberration of the cornea. The complexity of each anisotropic gaussian basis functions in the functional representation is the same, but their parameters vary to fit the current scale. This scale is determined only by the residual errors and not by the number of the iteration. Finally, the position and clustering of their centers, as well as the size of the shape parameters, provides an additional spatial information about the regions of higher irregularity. These results are compared with the standard approximation procedures based on the Zernike polynomials expansions.

Keywords: Zernike polynomials; surface reconstruction; surface modeling; corneal irregularities; gaussian functions; radial basis functions; multi-scale methods

*Corresponding author. Send correspondence to Department of Statistics and Applied Mathematics, University of Almería, 04120 Almería, Spain. Phone +34-950015217, fax +34-950015167.

Email address: andrei@ual.es (Andrei Martínez-Finkelshtein)

1. Introduction

There is an increasing need of a reliable and precise modeling of corneal surfaces, motivated both by technological and clinical applications.

Given the significance of the shape of the front surface of the cornea to the refraction of the eye [1] and the ability to correct refractive errors by laser ablation of the front surface of the cornea, a detailed wavefront error analysis of individual corneal topography data is crucial for the clinicians as a basis for a customized treatment. It has been recognized that the corneal front surface generally provides the bulk of the ocular aberrations in the postsurgical or pathologic eye [2].

Corneal modeling can be used also as a tool for screening corneal diseases. Keratoconus, for example, distorts the corneal shape and results in a significant vision loss. Keratoconic patients should be screened for refractive surgeries because such treatments may worsen the corneal shape and lead to corneal transplantation. Hence, corneal modeling plays an essential role in diagnosing and managing keratoconus to assess suitability of a subject for the treatment and prevent improper refractive surgeries [3]. Also, the great role of the reliable visualization tools in clinical practice should not be underestimated.

Modern techniques of design and fit of soft contact lenses can take into account features of the patient's eye, adapting the back surface of a lens to match the specific elevations of the cornea. These methods require again a detailed corneal topographic analysis of the anterior face of the cornea.

With the introduction of the high-speed videokeratoscopy [4, 5] in the study of the dynamics of corneal surface topography [6] and tear film stability [7], the storage needs have become significant, motivating another important application of corneal surface modeling: data compression [8].

The vast majority of modern corneal topographers collects the data (either elevation, curvature, mire displacement or others) in a finite and discrete set of points. Typically, these points present a quasi-structure; for instance, those devices based on the Placido rings technology provide elevations and curvatures at the discretized images of the mires, whose deformation is proportional to the complexity of the surface. In any case, the data is contaminated by the error, which stems from several sources: the natural device noise, measurement and digitalization errors, algorithm errors (like those converting the displacement in elevation), rounding errors and others. Hence, we face the problem of the parsimonious fit of the actual surface data contaminated by noise, with a minimum number of coefficients or parameters, for its clinical and technological applications.

The solutions to this problem are usually classified into the so-called zonal and modal methods. In the former group, the domain where the data are collected is subdivided in more elementary subdomains (e.g., triangles), and the surface is approximated in each subdomain with a relative independence from the other regions. The standard tool here are splines, in particular, the numerically stable B-splines [9, 10, 11], widely used in the Computer-Aided Geometric Design.

In the modal methods of reconstruction the approximation is found as a (typically, linear) combination of functions from the given set or dictionary, defined by a number of parameters. In this sense, crucial decisions to make are the set of functions to use, the value of their parameters, and the number of functions needed to recover the relevant information without fitting (at least, in a large scale) the inevitable error (the so-called model selection problem).

Among the advantages of the zonal methods is the flexibility and the accuracy of fit, but they lack the simplicity of the modal approach, which renders functional expressions valid across the whole domain, suitable for further calculations (such as ray tracing and others). Zonal techniques are also substantially more computer-intensive and encode the final shape in a larger amount of data.

A standard functional basis for the modal reconstruction, used commonly in ophthalmology to express ocular wavefront error are the Zernike polynomials [12]. The coefficients of their expansions have interpretation in terms of the basic aberrations such as defocus, astigmatism or coma, along with higher order aberrations such as trefoil and spherical aberrations. As a fitting routine, Zernike polynomials are not limited to analysis of wavefront error surfaces, but can be applied to other ocular surfaces as well, including the anterior corneal surface [13, 14]. It has been suggested that Zernike analysis may be applicable in the development of corneal topography diagnostic tools, using the Zernike coefficients as inputs into corneal classification of neural networks [15, 16], replacing or supplementing the currently used corneal indices included with many topography devices.

However, potential limitations in this approach have been reported in the literature [12, 17]. There is a growing concern that the Zernike fitting method itself may be inaccurate in abnormal conditions. Furthermore, it is very difficult to assess a priori how many terms are necessary to achieve acceptable accuracy in the Zernike reconstruction of any given corneal shape [18]. It is known [17] that limiting Zernike analysis to only a few orders may cause incorrect assessment of the severity of the more advanced stages of keratoconus [1]. This information is particularly needed in the discriminant analysis of the disease markers, or when selecting the numerical inputs for neural network-based diagnostic software such as corneal classification and grading utilities for condition severity.

In this sense, several alternatives to the modal least-square fit with Zernike polynomials have been recently suggested. Some of them intent to combine the modal and zonal approaches in order to preserve the best of both worlds [19] or to use non-linear methods [8]. The idea of the possibility of getting the accuracy and flexibility of the zonal methods within the framework of the linear modal approach by means of localized radial basis functions has been expressed in [20], but without developing an actual implementation or procedure. In this paper we describe an adaptive and multi-scale working algorithm for the parsimonious fit of the surface data, based on residual iteration with knot insertion, that allows to adapt the number of functions used in the reconstruction to the conditions of each cornea.

The residual iteration is well-known in many branches of mathematics; it is related for instance with the iterative refinement methods of solution of systems of linear equations. In the context of purely radial basis functions an adaptive greedy approximation algorithm using interpolation has been proposed in [21]. The adaptive increase of the number of basis functions as a technique used to improve the quality of a given initial approximation is also standard, see for instance [22, Ch. 21] for a general discussion and references.

The method proposed here allows to build iteratively an approximation function as a linear combination of anisotropic gaussian basis functions, implementing also a dynamical selection of the parameters and the management of noise. It can be used to reconstruct altimetric data, corneal power maps, and others. Although it has been tuned up for the Placido ring based keratoscopes, with the data nodes located in almost concentric rings, the technique is actually applicable to any scattered data approximation. Part of this approach was announced in [23].

2. The fitting procedure

2.1. The general setting

The input data is a 3D cloud (x_k, y_k, z_k) , $k = 1, \dots, N$, corresponding to either elevation or corneal power z_k by a corneal topographer at the node \mathbf{P}_k of the anterior corneal surface with Cartesian coordinates (x_k, y_k) . We will discuss the case when z_k corresponds to elevation. Taking into account

the global shape of the cornea, a standard procedure is to “flatten” the data by fitting it with the best-fit sphere [24] of the form:

$$S(x, y) = z_0 + \sqrt{R^2 - (x - x_0)^2 - (y - y_0)^2}$$

where R and (x_0, y_0, z_0) are its radius and the Cartesian coordinates of its center, respectively. Although the common practice is to fit with the standard linear least squares, better results are obtained with a weighted least square fit, using $(1 + \|P_k\|)^{-1}$ as the weight, in accordance with the typical error distribution [25].

As a result of the previous step, the residual errors $\varepsilon_k^{(1)} = z_k - S(x_k, y_k)$ contain both the relevant information at different scales and noise. Our aim is to fit these residuals by a function $E(x, y)$ in such a way that an analytic expression for the corneal height is given by

$$\text{Cornea}(x, y) = S(x, y) + E(x, y), \quad (1)$$

In this way, S accounts for the global shape of the cornea, while E captures the small irregularities in the surface. Function E is given by a linear combination of n functions from a given dictionary,

$$E(x, y) = E_n(x, y) = \sum_{j=1}^n c_j h_j(x, y). \quad (2)$$

In an ideal setting, n depends on the actual data, and should be large enough to allow all relevant information from the elevation measured modeled by E , but not too large to prevent from overparametrizing the problem and fitting pure noise. In order to circumvent the difficulties of the Zernike polynomials mentioned above we use as basis functions the gaussians of the form

$$h(x, y) = \exp(-\|\mathbf{P} - \mathbf{Q}\|_A^2), \quad \mathbf{P} = (x, y)^T,$$

where the superscript T denotes the matrix transpose, $\mathbf{Q} = (Q_x, Q_y)^T$ is a certain point on the plane (“center”), and A is a positive-definite matrix in $\mathbb{R}^{2 \times 2}$. For such a matrix the A -norm of a point (column vector) \mathbf{P} in \mathbb{R}^2 is defined as

$$\|\mathbf{P}\|_A = \sqrt{\mathbf{P}^T \mathbf{A} \mathbf{P}} = \sqrt{\alpha_x x^2 + \alpha_y y^2 + 2\alpha_{xy} xy},$$

for $A = \begin{pmatrix} \alpha_x & \alpha_{xy} \\ \alpha_{xy} & \alpha_y \end{pmatrix}$ with $\alpha_x > 0$ and $\alpha_x \alpha_y > \alpha_{xy}^2$.

In general, these are anisotropic radial basis functions, that boil down to standard radial basis gaussian functions (RBGF) when both eigenvalues of A coincide (in other words, when A is a positive multiple of the identity matrix I_2).

One of the advantages of these functions is that they are quasi compactly supported: for $|x| > 1.7308$, $e^{-x^2} < 0.05$, that is, achieves less than 5% of its maximum height.

Hence, we seek the expression of the form

$$\text{Cornea}(x, y) = S(x, y) + \sum_{j=1}^n c_j \exp(-\|\mathbf{P} - \mathbf{Q}^{(j)}\|_{A_j}^2), \quad \mathbf{P} = (x, y)^T.$$

Clearly, a fitting routine should allow for an adequate selection of all parameters, namely

- centers $\mathbf{Q}^{(j)}$;
- shape matrices (or shape parameters) A_j ;
- scaling factors c_j ;
- number of terms n in the functional representation.

We propose an iterative algorithm of reconstruction, such that in each step we fit partially the residual error by one anisotropic RBGF (A-RBGF), and compute the new residuals, which will become the input for the next iteration (residual iteration with knot insertion). To preserve the maximum possible degrees of freedom, the centers, the shape parameters and the scaling factors will be chosen dynamically depending on the residual data in each step.

2.2. Description of the iterative algorithm

Let E_{j-1} be already computed (we take $E_0 \equiv 0$). The input data for the j 's iteration ($j = 1, 2, \dots$) is the cloud $(x_k, y_k, \varepsilon_k^{(j)})$ of nodes $\mathbf{P}_k = (x_k, y_k)^T$ and the corresponding residuals $\varepsilon_k^{(j)}$, $k = 1, 2, \dots, N$; recall that $\varepsilon_k^{(1)} = z_k - S(x_k, y_k)$ are the residual errors after the weighted best sphere fit. We perform the following steps:

STEP 1: selection of the center $\mathbf{Q}^{(j)}$.

The problem of the selection of a center of a radial basis function has been discussed in [26]. There the center is chosen among the data nodes \mathbf{P}_k using the criterion of maximum cross-correlation. Another criterion uses the power function (see e.g. [21], [22], [27]). Both methods, although computationally demanding, can be implemented to perform Step 1. However, in our practice we found the much simpler criterion of maximal residual (strategy for so-called greedy approximation) to be as satisfactory, at a minimum cost; it correlates also with the geometry of the A-RBGF. Hence, we choose

$$\mathbf{Q}^{(j)} = (x_{k_0}, y_{k_0})^T \quad \text{where} \quad k_0 = \arg \max_k |\varepsilon_k^{(j)}|$$

and denote

$$m^{(j)} := \varepsilon_{k_0}^{(j)}.$$

STEP 2: dynamical filtering.

As it was mentioned before, the altimetric data obtained from measuring devices such as a keratographer are contaminated by noise. Although there are some indications about statistical distribution of these errors, the information is still very limited. In order to cope with this problem we need to filter out those data that clearly correspond to the measurement error and thus spoil the quality of the reconstruction. We can do it in advance, before starting the fitting procedure, like in [28]. But we have chosen a simpler alternative, that yields satisfactory results: once the center $\mathbf{Q}^{(j)}$ has been selected, we check the number, ℓ_k , of nodes \mathbf{P}_k lying in the largest disk, centered at $\mathbf{Q}^{(j)}$ and containing only nodes with the residues of the same sign as $m^{(j)}$. If $\ell_k < 20$, we consider $\mathbf{Q}^{(j)}$ an outlier and exclude it from consideration at this iteration. This can be done by simply setting $\varepsilon_{k_0}^{(j)} = 0$, after which we return to Step 1. Otherwise, we proceed to the next step.

Obviously, this step can be ignored if we know that the error is negligible.

STEP 3: selection of the shape parameters.

We determine first the influence nodes $\mathcal{P}_j(s)$, defined as the maximal set of at most s nodes \mathbf{P}_k closest to $\mathbf{Q}^{(j)}$ with residues of the same sign than $m^{(j)}$. Observe that $\mathcal{P}_j(s_1) \subset \mathcal{P}_j(s_2)$ if $s_1 \leq s_2$. It

is convenient to parallelize the subsequent computations for several values of s : we have performed experiments using the vector of values $s = [s_{\min}, 100, 150, 200, 300]$, where $s_{\min} = \min(\ell_k, 50)$, with ℓ_k defined in Step 2.

The interpolating conditions

$$\varepsilon_{k_0}^{(j)} h_j(x_k, y_k) = \varepsilon_k^{(j)}, \quad k \in \mathcal{P}_j(s),$$

are equivalent to the overdetermined linear system

$$\begin{aligned} \alpha_x (x_k - x_{k_0})^2 + 2\alpha_{xy} (x_k - x_{k_0})(y_k - y_{k_0}) \\ + \alpha_y (y_k - y_{k_0})^2 = \log \left(\frac{\varepsilon_{k_0}^{(j)}}{\varepsilon_k^{(j)}} \right), \quad k \in \mathcal{P}_j(s) \end{aligned} \quad (3)$$

in the 3 unknown entries of the shape matrix

$$A_j = \begin{pmatrix} \alpha_x & \alpha_{xy} \\ \alpha_{xy} & \alpha_y \end{pmatrix}.$$

We solve this system in the sense of weighted linear least squares (WLS), where the k -th equation is multiplied by the weight $\eta_k := (1 + \|\mathbf{P}_k - \mathbf{Q}^{(j)}\|^2)^{-1}$ in order to account for the bigger influence of the neighboring nodes on A_j . This solution is obtained by standard methods, using either the QR factorization of the collocation matrix corresponding to (3) or its singular value decomposition, see e. g. [29].

Observe that due to the selection of the active center $\mathbf{Q}^{(j)}$,

$$t_k := \log \left(\frac{\varepsilon_{k_0}^{(j)}}{\varepsilon_k^{(j)}} \right) \geq 0, \quad k \in \mathcal{P}_j(s).$$

However, this condition does not guarantee that the solution A_j of (3) in the sense of the WLS described above will be positive definite. This can typically fail in the periphery of the convex hull of the nodes, where the lack of data in some direction might yield non-positive definite A_j . Although the corresponding function h_j might fit the data correctly locally, it is not valid globally due to its exponential increase in the direction of the eigenvector of a negative eigenvalue of A_j .

In order to overcome this problem we examine the solution A_j of (3): if it is not positive definite, we interpret that there is a lack of data in a neighborhood of $\mathbf{Q}^{(j)}$ and force h_j to be an isomorphic (a bona fide) radial basis function: $A_j = \alpha I_2$. In this way, (3) is reduced to

$$\alpha \left\| \mathbf{P}_k - \mathbf{Q}^{(j)} \right\|^2 = t_k, \quad k \in \mathcal{P}_j(s), \quad (4)$$

whose solution in the sense of the WLS is

$$\alpha = \sum_{k \in \mathcal{P}_j(s)} \theta_k t_k,$$

with

$$\theta_k = \frac{\eta_k^2 \left\| \mathbf{P}_k - \mathbf{Q}^{(j)} \right\|^2}{\sum_{t \in \mathcal{P}_j(s)} \eta_t^2 \left\| \mathbf{P}_t - \mathbf{Q}^{(j)} \right\|^4}$$

and $\eta_k = \left(1 + \|\mathbf{P}_k - \mathbf{Q}^{(j)}\|^2\right)^{-1}$. Observe that in this case α is positive by construction, and we define

$$h_j(x, y) = \exp\left(-\alpha \|\mathbf{P} - \mathbf{Q}^{(j)}\|^2\right), \quad \mathbf{P} = (x, y)^T.$$

STEP 4: selection of the scaling factor.

We can calculate the coefficient c_j from

$$c_j h_j(x_k, y_k) = \varepsilon_k^{(j)}, \quad k \in \mathcal{P}_j(s),$$

using the WLS with the same weights η_k :

$$c_j = \sum_{k \in \mathcal{P}_j(s)} \gamma_k \varepsilon_k^{(j)},$$

with

$$\gamma_k = \frac{\eta_k^2 h_j(x_k, y_k)}{\sum_{t \in \mathcal{P}_j(s)} \eta_t^2 h_j(x_t, y_t)^2}.$$

It should be noted however that numerical experiments show that in many cases the much simpler interpolation condition $c_k = m^{(j)}$ yields comparable results.

STEP 4: computation of the new residuals.

With the values of c_j and A_j just computed we update

$$\varepsilon_k^{(j+1)} = \varepsilon_k^{(j)} - c_j h_j(x_k, y_k).$$

As it was mentioned before, all the computations have been performed in parallel for different values of s (typically, from 3 to 5 values between 50 and 300), and hence, different nested sets of influence nodes $\mathcal{P}_j(s)$. We now keep the value of s (and the corresponding values of c_j and A_j) that yields the smallest norm of the residue vector ($\varepsilon_k^{(j+1)}$), and discard the other values. As a result, we find the new approximation, $E_j = E_{j-1} + c_j h_j$.

As a final step, we check the stopping criterium that will be discussed below. If this is not satisfied, we increment the iteration counter j in 1 and return to Step 1.

2.3. Stopping criteria

In theory, the algorithm run indefinitely yields an interpolating function, and in consequence, a zero residue vector. In the real life situation of the data contaminated by errors, a very important problem is that of the model order selection: we want to capture all the relevant information without over-parametrizing the model and without fitting the noise. Many solutions to this problem are described in the literature. For instance, the choice of the number of Zernike polynomials for the modal reconstruction of the altimetric data has been discussed in [30, 31].

The statistical methods of selection of the appropriate number n in (2) usually make assumptions about the noise. However, a priori information about the measurement error bounds or measurement error distribution is limited. According to [25], the errors cannot be assumed i.i.d. random variables, although the assumption that they are normally distributed (with the variance proportional to the square of the distance of the node to the center) is apparently reasonable. They are also computationally intensive, [31, 32].

Less demanding methods use information theory criteria, such as the Akaike Information Criterion (AIC), or the Efficient Detection Criterion (EDC) [33], which studies the evolution of

$$EDC_j(p) = N \log(MSE_j) + j(N \log N)^p, \quad 0 < p < 1,$$

with

$$MSE_j = \frac{\rho^2}{N} \sum_{k=1}^N \left(\varepsilon_k^{(j)} \right)^2, \quad \rho = \max_k \|P_k\|. \quad (5)$$

The value of p is usually tuned up experimentally.

However, we can gain information analyzing directly the behavior of the normalized errors MSE_j defined in (5). Typically, these errors start decaying with an exponential rate and average order greater than 1. After a number of iterations we observe a stabilization in this rate of decay that becomes linear; this typically happens when values of MSE_j are between 10^{-3} and $10^{-4} \mu\text{m}^2$. Based on this experience we have used successfully the following stopping criterion: we allow the algorithm run for up to 50 iterations (this takes less than 2 seconds to complete) and calculate the weighted slopes

$$\delta MSE_j = \frac{\log(MSE_{j+1}/MSE_j)}{j}.$$

The sequence δMSE_j , although oscillatory, is negative and tends to zero, so we find the last iteration $1 \leq J_1 \leq 50$ when $\delta MSE_j \geq -0.02$. If $MSE_{J_1} < 10^{-3} \mu\text{m}^2$, we fix $J_1 + 1$ as the stopping iteration. Otherwise we seek for the last iteration $1 \leq J_2 \leq 50$ when $\delta MSE_j \geq -0.01$, and stop the algorithm at the $(J_2 + 1)$ -th iteration.

3. Experimental results

In this section, we present a comparison of the numerical results obtained with our method applied both to simulated and real cornea surfaces. All the procedures were implemented in Matlab 7 and run on standard platforms (Windows PC and Mac with average configuration). The altimetric and curvature data from in-vivo corneas used for experiments described below were collected by the CSO topography system (CSO, Firenze, Italy), which in ideal conditions digitizes up to 24 rings with 256 equally distributed points on each mire.

Since the procedure is meant for real-time reconstruction of the corneal data, any extremely computer-intensive methods should be discarded. However, in our Matlab implementation the execution time was always below 2 seconds, so this becomes less and less of a concern with the progress of the software optimization and computing power.

We demonstrate first the power of the proposed methodology using some elementary surface models, starting with the simplest example: a surface given by a linear combination of three exponentials,

$$\text{Cornea}(x, y) = \sum_{j=1}^3 c_j \exp\left(-\|\mathbf{P} - \mathbf{Q}_j\|_{A_j}^2\right), \quad \mathbf{P} = (x, y)^T,$$

with $\mathbf{Q}_1 = (0.3, -0.4)^T$, $\mathbf{Q}_2 = (0.7, 0.1)^T$, $\mathbf{Q}_3 = (-0.4, 0.3)^T$, $c_1 = 0.02$, $c_2 = -0.015$, $c_3 = 0.02$, and

$$A_1 = \begin{pmatrix} 10 & -7 \\ -7 & 20 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 15 & 3.5 \\ 3.5 & 10 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 20 & -6 \\ -6 & 12 \end{pmatrix}.$$

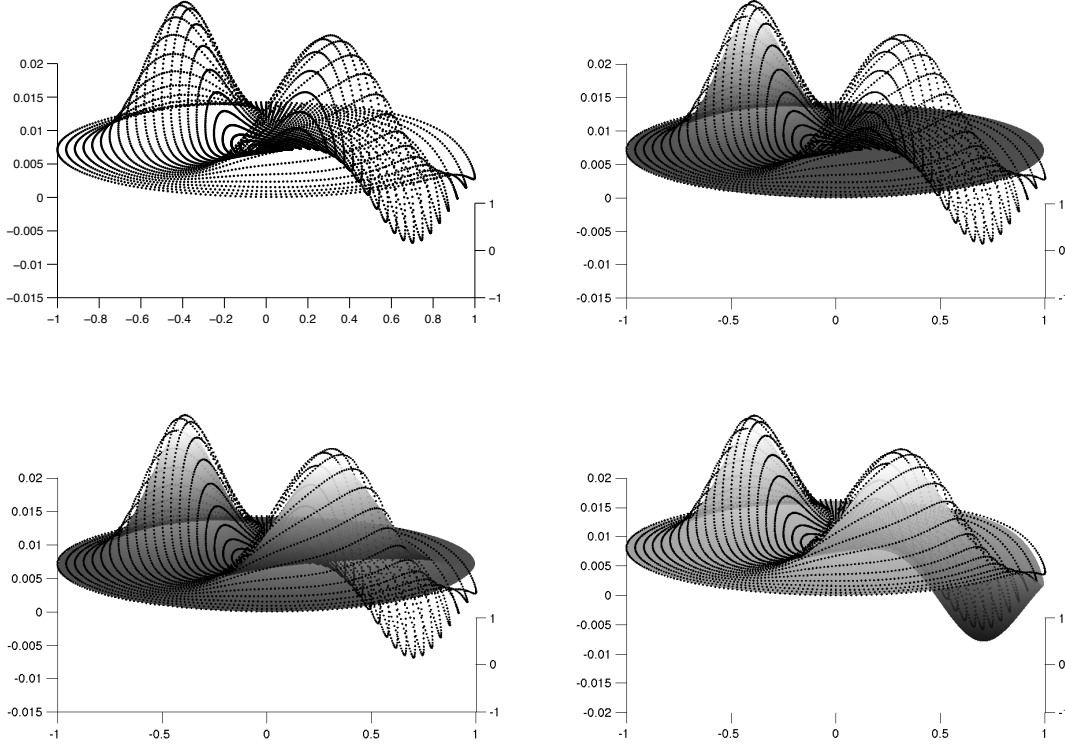


Figure 1: Example of the performance of the algorithm in the simplest case of a surface comprised of three gaussians (upper left): the first iteration (upper right) matches the highest peak, while the next two capture the rest of the features of the surface.

The surface is obviously exaggerated with respect to the typical residue of the standard cornea, but this is made with illustrative purpose. The result of the first three iterations of the algorithm are shown in Figure 1.

The measurement white noise (zero-mean and Gaussian noise process) of variance of 10% of the maximum height was added to the surface. In most optical applications this would correspond to a very high level of the measurement noise. It should be noted that the knowledge of the distribution of the measurement noise is not necessary for our algorithm. As a result, the three centers of the Gaussians were correctly estimated (see Figure 2).

The global error, measured by MSE_j , tends to decrease monotonically with the exponential rate. A typical behavior for the reconstruction of the altimetric data for both synthetic and real corneas can be observed in Figure 3. In particular, in all cases we observe the feature mentioned above of the clear transition from an over-exponential to linear rate of decay, that is used as the stopping criteria. For comparison, we have reconstructed the same data with the Zernike polynomials using the linear least squares, which is the standard procedure in the clinical practice, implemented in most topographers. The MSE_j for the Zernike reconstruction is plotted in the same Cartesian coordinate

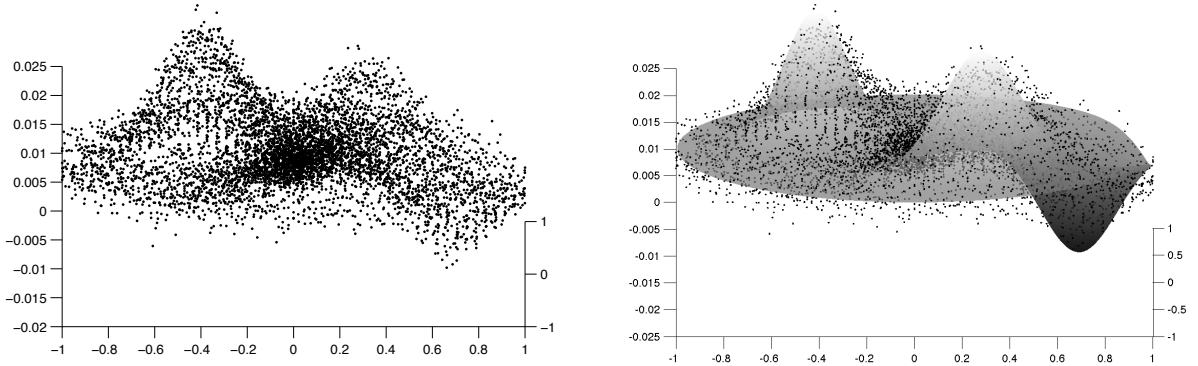


Figure 2: Performance of the algorithm with the surface analyzed in Figure 1 contaminated by white noise of 10% of the maximum height: the data (left) and the result of the first three iterations (right).

system, where j indicates the total number of Zernike polynomials employed. Recall that j varies from 1 to $(m + 1)(m + 2)/2$, where m is the maximal radial order used.

We observe two curious features of these plots. First, the Zernike polynomials easily capture the global shape of the reconstructed surface, which is expressed in a typical fast decay of the corresponding error. However, smaller details on the surface (such as areas of localized steepening) are much less suited for this tool. It explains the clear saturation observed in the Zernike error behavior after a few (typically, between 21 and 36 polynomials, corresponding to $5 \leq m \leq 7$). This is not the case of the reconstruction with A-RBGF, whose multi-scale and adaptive character allows to adjust the parameters adequately in each iteration.

On the other hand, numerical experiments show another interesting phenomenon related with the stopping criterion we use: in a majority of situations the stopping time corresponds to a j for which MSE_j is approximately equal for A-RBGF and Zernike polynomials.

However illuminating these graphs are, the global error is not the best way to compare both reconstruction approaches. Recall that the modal method with Zernike polynomials is suited precisely to achieve a maximum reduction of the MSE_j for each j , while the iterative algorithm presented here has a totally different goal: locating the most salient feature of the residual surface and incorporating it into the analytic expression (1).

This can be illustrated by fitting a synthetic “cornea” having a simulated scar, used already in [20]. Its contour plot is depicted in Figure 4, upper left. The upper right plot shows the contour plot of the surface reconstructed with the adaptive procedure described here using as few as 20 functions. The other two contour plots correspond to the same surface reconstructed with 36 (order ≤ 7) and 136 (order ≤ 15) Zernike polynomials.

The situation becomes even more apparent if we compare the residual errors along the mire 8 for both methods (Figure 5, left): while the Zernike polynomials work perfectly on smooth portions of the surface, they find difficulties in adapting to fast varying shapes, where the A-RBGF algorithm uses its multi-scale and adaptive character to fit the surface almost perfectly after a few iterations. It takes a really big time for the MSE errors of Zernike polynomials to progress, as illustrated in

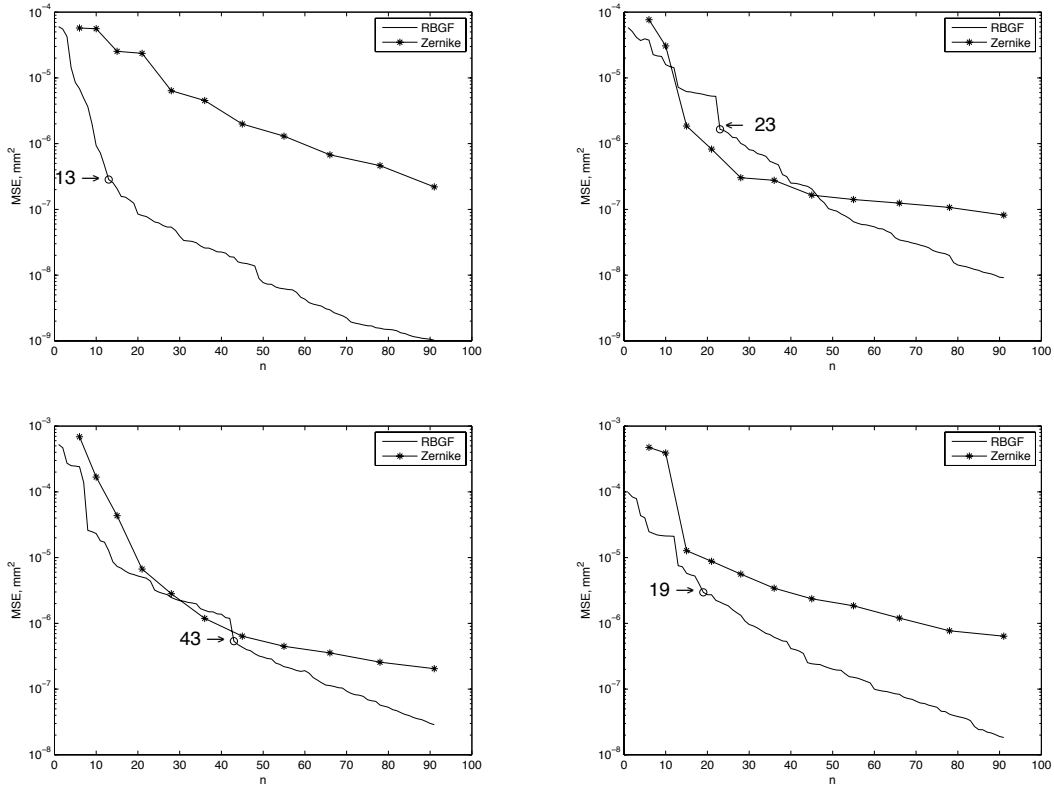


Figure 3: Comparison of the MSE_j for altimetric data reconstruction by Zernike polynomials and by the A-RBGF algorithm. Upper left: synthetic cornea made of a linear combination of 10 gaussian functions. Upper right: normal cornea. Lower row: keratoconic corneas. The arrows indicate the stopping time of the algorithm according to the criterion described in Subsection 2.3. The horizontal axis (n) indicates the number of functions used.

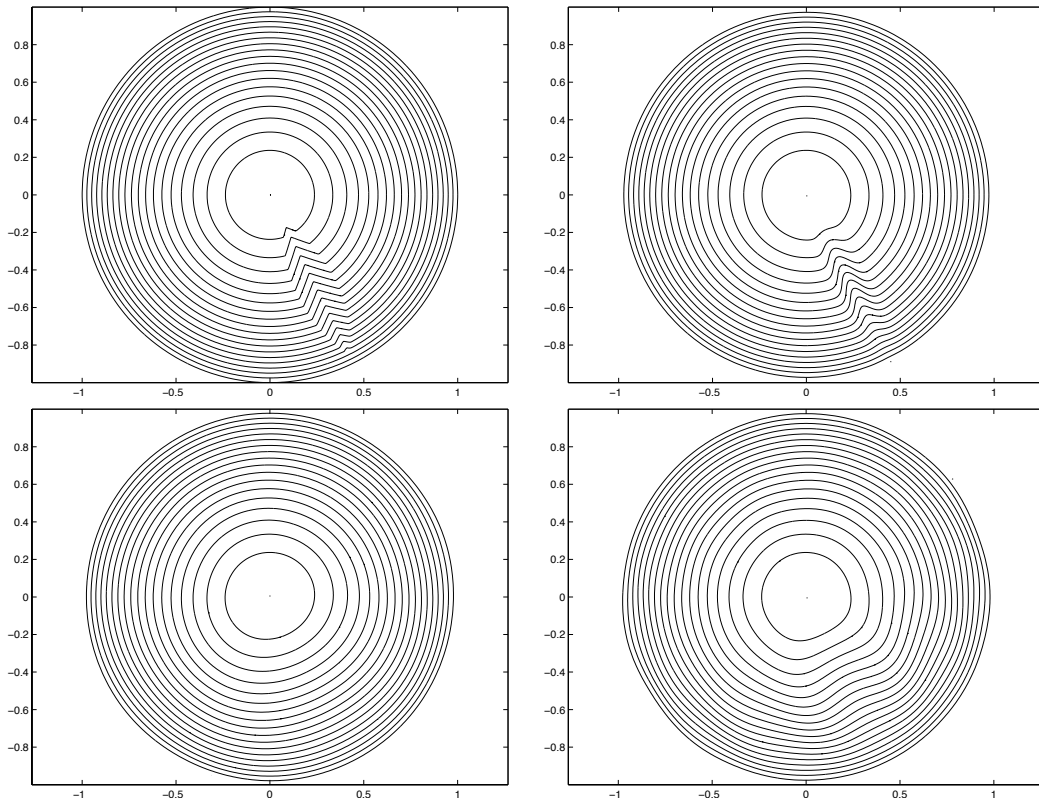


Figure 4: Comparison of the reconstruction of a synthetic “scar” (upper left) with 20 iterations of the adaptive algorithm (upper right), and 36 (lower left) and 136 (lower right) Zernike polynomials.

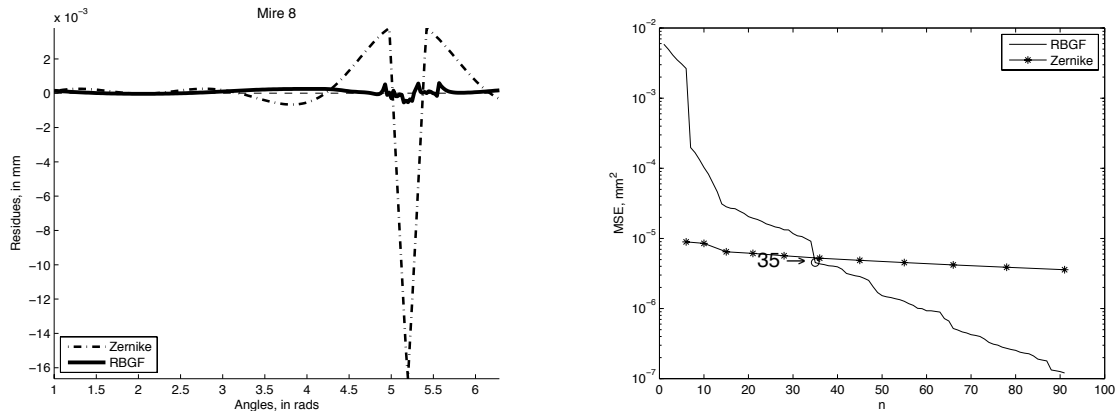


Figure 5: Left: residual errors along the mire 8 for the synthetic “scar” reconstructed with 20 iterations of the adaptive algorithm (bold line) and up to 36 Zernike polynomials (dotted line). Right: MSE_j for the “scar” reconstruction by Zernike polynomials and by the A-RBGF algorithm.

Figure 5, right.

Another indication of a consistent behavior of the iterative algorithm proposed here is the evolution of the parameters computed dynamically in each iteration. Although the eigenvalues of a shape matrix A_j tend to grow and can become eventually large (when fitting a small and steep area), their ratio (the spectral condition number of A_j) remains bounded with some exceptions; recall that the condition number 1 corresponds to an (isotropic) radial basis function. On the other hand, the scaling factors c_j steadily decrease, in concordance with a gradual reduction of the residual errors, see Figure 6.

Regarding the stopping time, the experiments performed with real and synthetic corneas show that the reasonable number of iterations lies between 20 and 40; there is no clear correlation between the number of iterations and the condition of the cornea, as Figure 3 shows. This is why we consider appropriate to perform 50 iterations (taking advantage of the speed of the algorithm) in order to decide the right value for n in (2).

However, more correlation exists with the location and grouping of the centers $Q^{(j)}$: for the normal corneas the centers typically cluster at the border of the area, where most of the oscillations occur, while for corneas affected by keratoconus we observe how some centers match the deformation already at the first iterations, see Figure 7.

The adaptive algorithm described here is suitable for a reliable reconstruction of any surface for the discrete set of data. In particular, we can also reconstruct a corneal power map or the wavefront, see an example in Figure 8. Taking into account the typical shape of such a surface, it is convenient here to skip the fit by a sphere or Zernike polynomials of a low order, making $S \equiv 0$ in (1).

Although the main goal of the algorithm proposed here is the reconstruction of the topography of the anterior surface of the cornea, it is convenient to assess the quality of the approximation by analyzing the resulting point spread function (PSF). Fig. 9 shows the effect of modeling the corneal surface data of a simulated highly irregular eye on the estimated PSF. The original PSF corresponding to a wavefront described by an expansion in 136 Zernike polynomials and a corneal diameter of 8 mm is shown together with a Zernike polynomial approximation of radial order ≤ 5 and ≤ 9 (21 and 55 functions, respectively), and the A-RBGF approximation with 21 functions. Clearly, the latter

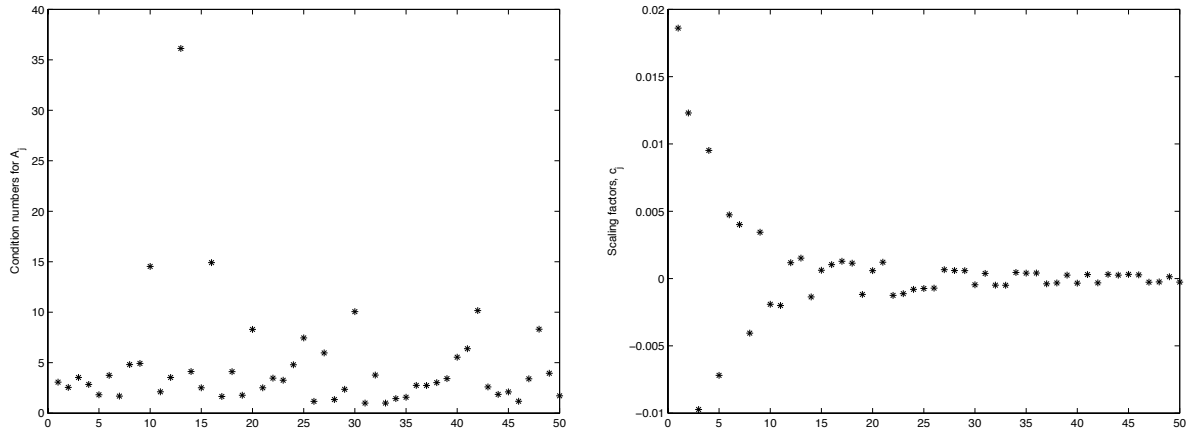


Figure 6: A typical evolution of the condition numbers of the shape parameters (left) and the scaling factors c_j (right), plotted against the number j of the function in (2).

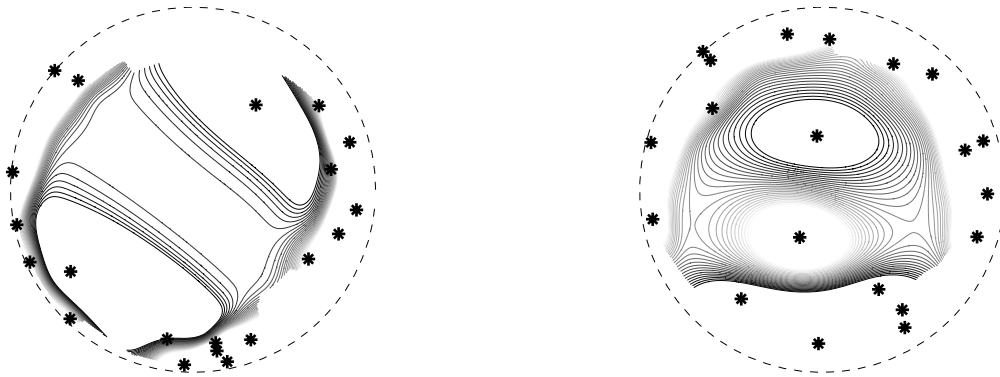


Figure 7: Center location for the reconstruction with 20 A-RBGF of a normal cornea (left) and a keratoconic one (right). The contours represent the level curves of the residues of the altimetric data with respect to the best fit sphere.

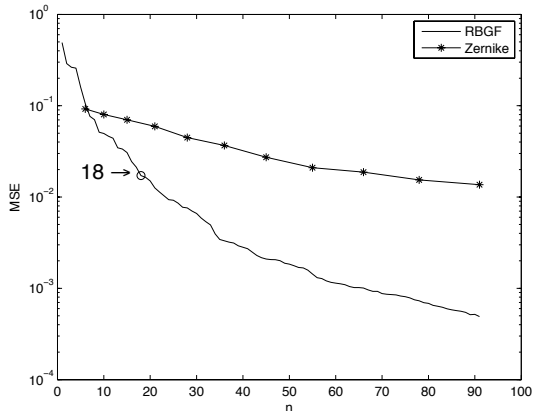


Figure 8: The MSE_j for corneal power data reconstruction by Zernike polynomials and by the A-RBGF algorithm.

leads to a PSF that more closely resembles the original PSF (capturing some finer features) than that derived from the Zernike polynomials.

4. Discussion and summary

In this work, we develop an adaptive fitting method for corneal data, combining modal simplicity with advantages of zonal reconstruction. It consists of a preliminary fit of the data with some global function (sphere or a combination of Zernike polynomials of a low order) and an iterative procedure that adds terms to the analytic representation of the corneal data. Each term consists of a scaled anisotropic radial basis gaussian function. The coefficients are computed dynamically and allow to fit the data in each iteration independently of the scale. The method comprises also a filtering procedure that discards the outliers (data clearly corresponding to measurement noise) and a stopping criterium that chooses the final number of functions in the analytic representation in accordance with the evolution of the residual error.

The numerical implementation of this algorithm in a standard personal computer is very fast (execution time below 2 seconds). Experimental results allow us to draw the following conclusions:

- the least square approximation by a linear combination of Zernike polynomials of a radial order up to 6 (which is the standard in modern aberrometers [34]) fits adequately the altimetric data in the case of a normal cornea. It can be used also to capture the major features of the shape of the surface. However, for strongly aberrated corneas the Zernike-based procedure saturates relatively early, and we need a high number of terms to achieve the desired accuracy at regions of localized steepening, at the price of overparametrizing the model and fitting the measurement noise. Last but not least, the complexity of each individual term in the functional representation increases with its index.
- in contrast, the iterative method presented here exhibits a steady exponential error decay, independently of complexity of the cornea. Its actual rate is basically influenced by the residue distribution: the fast decrease in the first iterations, when all salient features are reconstructed,

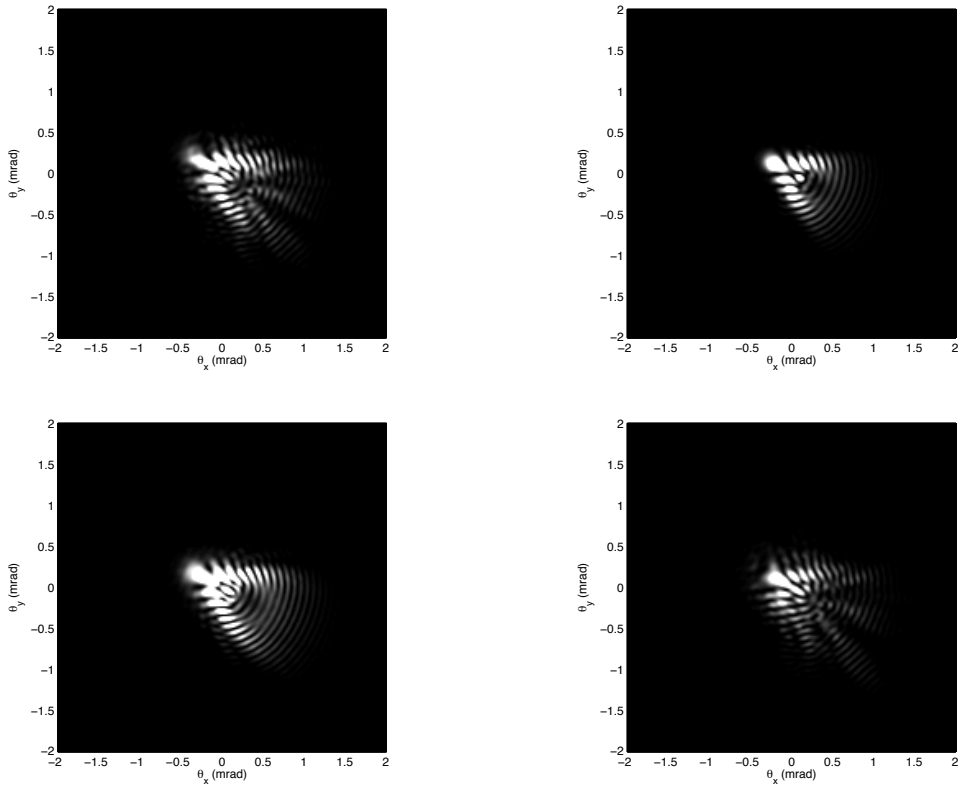


Figure 9: Original PSF for a simulated cornea with a pupil size of 8 mm of diameter and high wavefront error (upper left), and the corresponding PSFs from a reconstruction with first 21 (order ≤ 5) Zernike polynomials (upper right), first 55 (order ≤ 9) Zernike polynomials (lower left), and with 21 A-RBGF functions (lower right).

is followed by a stable linear decay, when essentially errors are being fit. This can be used as a stopping criterium for the iterations. In this way, the minimal amount of functions in the analytic representation of the cornea is used.

- unlike in the case of Zernike polynomials, the complexity of each term in the functional representation with A-RBGF is the same, but their parameters vary to fit the current scale. This scale is determined only by the residual errors and not by the number of the iteration.
- due to the localized character of A-RBGF, the position and clustering of their centers, as well as the size of the shape parameters, provides an additional spatial information about the regions of higher irregularity. These ideas were actually used in elaboration of new cornea irregularity indices that are currently under study.
- Zernike-based reconstruction, being center-oriented, is also very sensitive to rings with incomplete data. In clinical practice this is usually motivated by eye lashes obstruction or a tear film disruption. Still, a large portion of data of each incomplete mire is available and used by the iterative reconstruction with the A-RBGF.

The iterative adaptive algorithm for the cornea modeling proposed here provides a method of obtaining a compact mathematical description of the shape or power map of the cornea. All information is ultimately encoded in the following set of values: center and radius of the best-fit sphere, plus the center locations, shape parameters and scaling factors. This description can be used for global visualization of the cornea or of its portions, capturing smaller details than with standard procedures. It can serve also as the input data for resampling and computation of some other relevant values via ray tracing, numerical integration and others.

Acknowledgments

A.M.-F. is supported in part by Junta de Andalucía grants FQM-229 and P09-FQM-4643, and by the Ministry of Science and Innovation of Spain (project code MTM2008-06689-C02-01). Both A.M.-F. and D.R.-L. are also supported in part by Junta de Andalucía grant P06-FQM-01735.

Bibliography

References

- [1] J. T. Schwiegerling, J. E. Greivenkamp, Keratoconus detection based on videokeratoscopic height data, *Optometry and Vision Science* 73 (1989) 721–728.
- [2] P. Artal, E. Berrio, A. Guirao, P. Piers, Contribution of the cornea and internal surfaces to the change of ocular aberrations with age, *J Opt Soc Am A* 19 (2002) 137–143.
- [3] N. Maeda, S. D. Klyce, M. K. Smolek, Comparison of methods for detecting keratoconus using videokeratography, *Archives of Ophthalmology* 113 (7) (1995) 870–874.
- [4] J. Nemeth, B. Erdelyi, B. Csakany, P. Gaspar, A. Soumelidis, F. Kahlesz, Z. Lang, High-speed videotopographic measurement of tear film build-up time, *Invest. Ophthalmol. Vis. Sci.* 43 (6) (2002) 1783–1790.

- [5] D. R. Iskander, M. J. Collins, Applications of high-speed videokeratometry, *Clin. Exp. Optom.* 88 (4) (2005) 223–231.
- [6] M. Zhu, M. J. Collins, D. R. Iskander, Dynamics of ocular surface topography, *Eye* 21 (2007) 624–632.
- [7] D. R. Iskander, M. J. Collins, B. Davis, Evaluating tear film stability in the human eye with high-speed videokeratometry, *IEEE Trans. Biomed. Eng.* 52 (11) (2005) 1939–1949.
- [8] M. Schneider, D. R. Iskander, M. J. Collins, Modeling corneal surfaces with rational functions for high-speed videokeratometry data compression, *IEEE Trans. Biomed. Eng.* 56 (2) (2009) 493–499.
- [9] M. A. Halstead, B. Barsky, S. A. Klein, R. B. Mandell, A spline surface algorithm for reconstruction of corneal topography from a videokeratographic reflection pattern, *Optometry and Vision Science* 72 (11) (1995) 821–827.
- [10] X. Liu, Y. Gao, B-spline based wavefront reconstruction for lateral shearing interferometric measurement of engineering surfaces, *Adv. Abrasive Technol.* 238–239 (2003) 169–174.
- [11] M. Ares, S. Royo, Comparison of cubic B-spline and Zernike-fitting techniques in complex wavefront reconstruction, *Applied Optics* 45 (2006) 6945–6964.
- [12] S. D. Klyce, M. D. Karon, M. K. Smolek, Advantages and disadvantages of the Zernike expansion for representing wave aberration of the normal and aberrated eye, *J. Refractive Surgery* 20 (2004) S537–S541.
- [13] M. K. Smolek, S. D. Klyce, Zernike polynomial fitting fails to represent all visually significant corneal aberrations, *Invest. Ophthalmol. Vis. Sci.* 44 (11) (2003) 4676–4681.
- [14] L. A. Carvalho, Accuracy of Zernike polynomials in characterizing optical aberrations and the corneal surface of the eye, *Invest. Ophthalmol. Vis. Sci.* 46 (2005) 1915–1926.
- [15] L. A. Carvalho, Preliminary results of neural networks and Zernike polynomials for classification of videokeratography maps, *Optometry and Vision Science* 82 (2) (2005) 151–158.
- [16] M. K. Smolek, S. D. Klyce, Current keratoconus detection methods compared with a neural network approach, *Invest. Ophthalmol. Vis. Sci.* 38 (11) (1997) 2290–2299.
- [17] M. K. Smolek, S. D. Klyce, Goodness-of-prediction of Zernike polynomial fitting to corneal surfaces, *J. Cataract Refract. Surg.* 31 (2005) 2350–2355.
- [18] R. Iskander, M. J. Collins, B. Davis, Optimal modeling of corneal surfaces with Zernike polynomials, *IEEE Trans. Biomed. Eng.* 48 (1) (2001) 87–95.
- [19] J. Espinosa, D. Mas, J. Pérez, C. Illueca, Optical surface reconstruction technique through combination of zonal and modal fitting, *Journal of Biomedical Optics* 15 (2) (2010) 026022.
- [20] A. Martínez-Finkelshtein, A. M. Delgado, G. M. Castro, A. Zarzo, J. L. Alió, Comparative analysis of some modal reconstruction methods of the shape of the cornea from the corneal elevation data, *Invest. Ophthalmol. Vis. Sci.* 50 (12) (2009) 5639–5645.

- [21] R. Schaback, H. Wendland, Adaptive greedy techniques for approximate solution of large RBF systems, *Numer. Algorithms* 24 (3) (2000) 239–254.
- [22] G. E. Fasshauer, *Meshfree Approximation Methods with MATLAB*, World Scientific, 2007.
- [23] A. Martínez-Finkelshtein, D. Ramos-López, G. M. Castro-Luna, J. L. Alió, A new efficient method of corneal reconstruction as an alternative to zernike polynomials fit, *Invest. Ophthalmol. Vis. Sci.* 51 (5) (2010) ARVO E-Abstract 5690.
- [24] S. J. Ahn, W. Rauh, H.-J. Warnecke, Least-squares orthogonal distances fitting of circle, sphere, ellipse, hyperbola, and parabola, *Pattern Recognition* 34 (12) (2001) 2283–2303.
- [25] W. Tang, M. J. Collins, L. G. Carney, B. Davis, The accuracy and precision performance of four videokeratoscopes in measuring test surfaces, *Optometry and Vision Science* 77 (9) (2000) 483–491.
- [26] A. A. Jamshidi, M. J. Kirby, Towards a black box algorithm for nonlinear function approximation over high-dimensional domains, *SIAM Journal on Scientific Computing* 29 (3) (2007) 941–963.
- [27] S. De Marchi, R. Schaback, H. Wendland, Near-optimal data-independent point locations for radial basis function interpolation, *Adv. Comput. Math.* 23 (3) (2005) 317–330.
- [28] W. Alkhaldi, D. R. Iskander, A. M. Zoubir, M. J. Collins, Enhancing the standard operating range of a Placido disk videokeratoscope for corneal surface estimation, *IEEE Trans. Biomed. Eng.* 56 (3) (2009) 800–809.
- [29] L. N. Trefethen, D. Bau III, *Numerical Linear Algebra*, SIAM, 1997.
- [30] R. Iskander, M. R. Morelande, M. J. Collins, B. Davis, Modeling of corneal surfaces with radial polynomials, *IEEE Trans. Biomed. Eng.* 49 (4) (2002) 320–328.
- [31] D. Alonso-Caneiro, R. Iskander, M. J. Collins, Estimating corneal surface topography in videokeratoscopy in the presence of strong signal interference, *IEEE Trans. Biomed. Eng.* 55 (10) (2008) 2381–2387.
- [32] D. R. Iskander, W. Alkhaldi, A. M. Zoubir, On the computer intensive methods in model selection, in: *Proc. 33rd IEEE Int. Conf: Acoust., Speech Signal Process. (ICASSP)*, 2008, pp. 3461–3464.
- [33] C. R. Rao, Y. Wu, A strongly consistent procedure for model selection in a regression problem, *Biometrika* 76 (2) (1989) 369–374.
- [34] J. J. Rozema, D. E. M. Van Dyck, M.-J. Tassignon, Clinical comparison of 6 aberrometers. part 1: technical specifications, *J. Cataract Refract. Surg.* 31 (2005) 1114–1127.