Optimal Modeling of Corneal Surfaces with Zernike Polynomials

D. Robert Iskander*, *Member, IEEE*, Michael J. Collins, and Brett Davis

***Abstract*—Zernike polynomials are often used as an expansion of corneal height data and for analysis of optical wavefronts. Accurate modeling of corneal surfaces with Zernike polynomials involves selecting the order of the polynomial expansion based on the measured data. We have compared the efficacy of various classical model order selection techniques that can be utilized for this purpose, and propose an approach based on the bootstrap. First, it is shown in simulations that the bootstrap method outperforms the classical model order selection techniques. Then, it is proved that the bootstrap technique is the most appropriate method in the context of fitting Zernike polynomials to corneal elevation data, allowing objective selection of the optimal number of Zernike terms. The process of optimal fitting of Zernike polynomials to corneal elevation data is discussed and examples are given for normal corneas and for abnormal corneas with significant distortion. The optimal model order varies as a function of the diameter of the cornea.**

***Index Terms*—Bootstrap, cornea, model order selection, Zernike polynomials.**

I. INTRODUCTION

THE cornea is the major refracting component of the human eye, contributing approximately two thirds of the eye's optical power. The ideal optical shape of the anterior surface of the cornea is a prolate ellipsoid. However, there are wide variations in shape producing common aberrations such as astigmatism. Corneal conditions such as keratoconus, in which the cornea thins and distorts, produce significant amounts of asymmetric aberrations that cannot be corrected with traditional spectacles.

Accurate measurement and modeling of the corneal surface is important from a number of perspectives. Corneal refractive surgery requires accurate modeling of corneal shape prior to surgery to ensure good optical and visual outcomes. Contact lens design and fitting can also be based upon corneal topography characterization.

To establish the contribution that the cornea makes to vision, one can take the measurement of the anterior surface of the cornea using noninvasive instruments such as the videokeratoscope [1], and then apply geometrical and wave optics to determine the wavefront aberration error [2]. For this purpose, the corneal data from the videokeratoscope is often given a functional representation in terms of a Zernike polynomial expansion [3].

The anterior surface of the cornea can be modeled by a finite series of Zernike polynomials [4]

$$C(\rho,\theta)=\sum_{p=1}^{P}a_p Z_p(\rho,\theta)+\varepsilon \tag{1}$$

where

$C(\rho,\theta)$	corneal surface;
the index p	polynomial-ordering number;
$Z_p(\rho,\theta)$, $p=1,\ldots,P$,	pth Zernike polynomial;
a_p, $p=1,\ldots,P$,	coefficient associated with $Z_p(\rho,\theta)$;
P	order;
ρ	normalized distance from the origin;
θ	angle;
ε	measurement and modeling error (noise).

It is assumed that this noise is an independent and identically distributed random variable with zero-mean and finite variance. We shall consider the general case where the distribution of the additive noise is unknown.

In such modeling, a fundamental problem arises of how many Zernike terms one should use. Traditionally, vision researchers have chosen to use the first 15 Zernike terms, which seems to be a legacy from the previous techniques of fitting Taylor series to the wavefront data to include the spherical aberration component.

Strictly speaking, modeling corneal surfaces with (1) involves *selection* of the model and *conditional estimation* of the parameters. By selecting the model we mean choosing an appropriate set of Zernike terms. The conditional estimation of the model parameters refers to estimation conditioned on the chosen model. In most practical cases, it is sufficient to choose only the model order of the polynomial expansion, P, rather than a particular model being a subset of $\{Z_1(\rho,\theta),\ldots,Z_P(\rho,\theta)\}$.

One way of selecting the number of Zernike terms is to minimize the residual variance (it will always decrease when more parameters are estimated) and determine a suitable cutoff threshold value. If this threshold is too small, however, we would over-parameterise our model and start modeling the measurement error rather than the surface.

An alternative approach is to use some suitable penalty function that increases with the number of parameters. Adding this penalty function to residual variance leads to model order selection criterion.

Manuscript received March 16, 2000; revised September 13, 2000. *Asterisk indicates corresponding author.*

*D. R. Iskander is with the Centre for Eye Research, Queensland University of Technology, Victoria Park Rd, Kelvin Grove Q4059, Australia (e-mail: d.iskander@qut.edu.au).

M. J. Collins and B. Davis are with the Centre for Eye Research, Queensland University of Technology, Kelvin Grove Q4059, Australia.

Publisher Item Identifier S 0018-9294(01)00138-0.

Many model selection procedures can be applied to find the optimal (in some sense) fit to the corneal surface with Zernike polynomials, because the model can be formulated as a linear regression in the unknown parameters. Some of the more commonly used model selection procedures include Akaike's information criterion (AIC) [5], the Rissanen's minimum description length (MDL) criterion [6], Hannan and Quinn (HQ) criterion [7], and the corrected Akaike information criterion (AIC_C) [8]. To the best of our knowledge, none of these model selection criteria have been used in the context of selecting the number of terms in Zernike polynomial expansion. As mentioned earlier, in most reports in the literature concerning aberrations of the human eye, authors choose to use the first 15 Zernike terms [3]. It should be also noted that even in the general context of linear regression, experimental as well as theoretical results indicate that these model selection criteria do not yield definitive results [8]. They may also perform poorly when the sample size is small.

In this paper, we first compare various methods for selecting the optimal number of terms of a Zernike polynomial expansion and propose a new method using the bootstrap [9]–[11]. With the bootstrap, the selection of the optimal number of Zernike terms can be performed in an objective manner. The bootstrap is a statistical technique for assessing the accuracy of a parameter estimator in situations where conventional techniques are not valid. As noted by Zoubir [11], "the bootstrap does with a computer what the experimenter would do in practice, if it were possible", that is repeat the experiment. The bootstrap randomly reassigns the observations and recomputes the estimates. The main benefit in using the bootstrap is that knowledge of the distribution of the measurement and modeling error is not necessary.

The paper is organized as follows. In Section II, we present the problem of modeling of the corneal surface with Zernike polynomials using videokeratoscopic measurements. In Section III, we introduce the bootstrap-based procedure for selecting the optimal number of Zernike terms. Section IV is devoted to simulation and experimental results. We first show in simulations that the proposed bootstrap approach outperforms the traditional model selection techniques. Then, we use the proposed technique to determine the optimal Zernike expansion for videokeratoscopic measurements from normal and distorted corneas.

II. Modeling with Zernike Polynomials

The pth-order Zernike polynomial is defined as [4]

$$Z_p(\rho,\theta) = \begin{cases} \sqrt{2(n+1)}R_n^m(\rho)\cos(m\theta), & \text{even } p, m \neq 0 \\ \sqrt{2(n+1)}R_n^m(\rho)\sin(m\theta), & \text{odd } p, m \neq 0 \\ \sqrt{n+1}R_n^0(\rho), & m = 0 \end{cases}$$

where n is the radial degree, m is the azimuthal frequency, and

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

The radial degree and the azimuthal frequency are integers which satisfy $m \leq n$ and $n - |m| =$ even. The radial degree n and the azimuthal frequency m can be evaluated from the polynomial-ordering number p using $n = \lfloor q_1 \rfloor - 1$, and $m = q_2 + (n + (q_2 \bmod 2) \bmod 2)$ respectively, where $q_1 = 0.5(1+\sqrt{8p-7})$, $p = 1,\ldots,P$, $q_2 = \lfloor (n+1)(q_1 - n + 1) \rfloor$, $\lfloor \cdot \rfloor$ is the floor operator and mod denotes the modulus operator.

In some cases the Zernike polynomials $Z_p(\rho,\theta)$ are denoted as a two–dimensional expansion, $Z_n^{\pm m}(\rho,\theta)$ in terms of radial, n, and azimuthal, m, parameters [3]. This is mathematically equivalent to Noll's notation and has been widely adopted by vision researchers. We have chosen the Zernike polynomials proposed by Noll as they are convenient for subsequent statistical analysis of the coefficient estimates and for finding the optimal model order (one-dimensional search for the minimum of the bootstrap mean-square error). Other classifications of Zernike terms are possible [12], and care should be taken when comparing the results of different authors.

The discrete data from the corneal surface, denoted in polar coordinates as $C(\rho_d,\theta_d)$, $d = 1,\ldots,D$, can be modeled by a finite series of discrete Zernike polynomials

$$C(\rho_d,\theta_d) = \sum_{p=1}^{P} a_p Z_p(\rho_d,\theta_d) + \varepsilon_d, \quad d = 1,\ldots,D \tag{2}$$

where $Z_p(\rho_d,\theta_d)$, $p = 1,\ldots,P$, is the pth discrete Zernike polynomial sampled from $Z_p(\rho,\theta)$ at discrete points (ρ_d,θ_d), $d = 1,\ldots,D$. Such sampling may require further orthogonalization using the Gram-Schmidt procedure.

Using a set of such orthogonalised discrete Zernike polynomials, we can form a linear model

$$\mathbf{C} = \mathbf{Z}\mathbf{a} + \varepsilon \tag{3}$$

where

- $\mathbf{C}$ — D-element column vector of corneal surface evaluated at discrete points (ρ_d,θ_d), $d = 1,\ldots,D$;
- $\mathbf{Z}$ — $(D \times P)$ matrix of discrete, orthogonalised Zernike polynomials $Z_p(\rho_d,\theta_d)$;
- $\mathbf{a}$ — P-element column vector of Zernike coefficients;
- ε — D-element column vector of the measurement and modeling error.

For such a model, the coefficient vector $\mathbf{a}$ can be easily estimated using the method of least-squares, i.e.,

$$\hat{\mathbf{a}} = (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{C} \tag{4}$$

where T denotes the transposition and provided that the inverse exists.

In some commercially available videokeratoscopes one may have access to the slope rather than the height data. In such cases, we can form the following linear model [similar to (3)]

$$\begin{bmatrix} \dfrac{\partial\mathbf{C}}{\partial\rho} \\ \dfrac{\partial\mathbf{C}}{\partial\theta} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial\mathbf{Z}}{\partial\rho} \\ \dfrac{\partial\mathbf{Z}}{\partial\theta} \end{bmatrix} \mathbf{a} + \begin{bmatrix} \varepsilon_\rho \\ \varepsilon_\theta \end{bmatrix}$$

where $\partial\mathbf{C}/\partial\rho$ and $\partial\mathbf{C}/\partial\theta$ are D-element column vectors of the partial derivatives of optical surfaces with respect to ρ and θ, while $\partial\mathbf{Z}/\partial\rho$ and $\partial\mathbf{Z}/\partial\theta$ are $(D \times P)$ matrices of partial derivatives of the discrete, orthogonalised Zernike polynomials

$Z_p(\rho_d, \theta_d)$ with respect to ρ and θ. For such a model, the coefficient $\mathbf{a}$ can be easily estimated using

$$\hat{\mathbf{a}} = \begin{bmatrix} \frac{\partial \mathbf{Z}}{\partial \rho} \\ \frac{\partial \mathbf{Z}}{\partial \theta} \end{bmatrix}^{\dagger} \begin{bmatrix} \frac{\partial \mathbf{C}}{\partial \rho} \\ \frac{\partial \mathbf{C}}{\partial \theta} \end{bmatrix}$$

where $\mathbf{X}^{\dagger} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ denotes the Moore–Penrose pseudo inverse of $\mathbf{X}$. Because slope measures provide twice as much data, this leads to better parameter estimates. Note that a similar approach may be taken if the slope data are in the Cartesian coordinates, if desired.

Note that the model based on the slopes is also linear in parameters. Thus, from the point of view of model order selection, a model based on the slope data can also be used for finding the optimal number of Zernike terms.

III. Model Order Selection

We can formulate the problem of finding the unknown model order of the Zernike polynomial expansion in the following manner: given discrete values of the corneal surface $C(\rho_d, \theta_d)$, $d = 1, \ldots, D$, estimate P.

Let model β be a subset of $\{1, \ldots, P\}$ corresponding to the model in (2) of order β. Under β, we have

$$\mathbf{C} = \mathbf{Z}_\beta \mathbf{a}_\beta + \varepsilon,$$

where $\mathbf{Z}_\beta$ is a matrix containing the first β columns of $\mathbf{Z}$ and $\mathbf{a}_\beta$ is a column vector containing the first β elements of $\mathbf{a}$. The proposed methodology is based on minimizing bootstrap estimates of the prediction error [10].

Let us assume that $\beta = P$. The *optimal* model is

$$\beta_0 = \max\{\beta : 1 \le \beta \le P, a_\beta \ne 0\}.$$

Let ε_d^*, $d = 1, \ldots, D$, be an independent and identically distributed random variable drawn with replacement from the empirical distribution of

$$\sqrt{\frac{D}{L_D}}\left(\hat{r}_d - \frac{1}{D}\sum_{d=1}^{D}\hat{r}_d\right), \quad d = 1 \ldots, D$$

where

$$\hat{r}_d = C(\rho_d, \theta_d) - \sum_{p=1}^{P} \hat{a}_p Z_p(\rho_d, \theta_d)$$

with $\hat{a}_p$, $p = 1, \ldots, P$, being elements of $\hat{\mathbf{a}}$ in (4) is the dth residual under the largest model $\beta = P$, and L_D is a scaling parameter. By multiplying the residuals by the factor $\sqrt{D/L_D}$ we increase the variability among the bootstrap observations to achieve consistency [10], i.e.,

$$\lim_{D\to\infty} \mathsf{Pr}\left\{\hat{\beta}_{D,L_D} = \beta_0\right\} = 1$$

provided that L_D is such that $\lim_{D\to\infty} L_D/D = 0$ and $\lim_{D\to\infty} L_D = \infty$.

Let us define the bootstrap analog of $\hat{\mathbf{a}}_\beta$ as

$$\hat{\mathbf{a}}_\beta^* = \left(\mathbf{Z}_\beta^T \mathbf{Z}_\beta\right)^{-1} \mathbf{Z}_\beta^T \mathbf{C}^*$$

TABLE I
BOOTSTRAP-BASED SELECTION PROCEDURE FOR THE MODEL ORDER OF ZERNIKE POLYNOMIAL EXPANSION

1. Select $\beta = \beta_{\max}$, find the estimate $\hat{a}_\beta$ of a_β using least-squares method and compute

$$\hat{C}(\rho_d, \theta_d) = \sum_{p=1}^{\beta_{\max}} \hat{a}_p Z_p(\rho_d, \theta_d)$$

2. Compute the residuals $\hat{r}_d = C(\rho_d, \theta_d) - \hat{C}(\rho_d, \theta_d)$, $d = 1, \ldots, D$.
3. Rescale the empirical residuals

$$\tilde{r}_d = \sqrt{\frac{D}{L_D}}\left(\hat{r}_d - \frac{1}{D}\sum_{d=1}^{D}\hat{r}_d\right).$$

4. For all $1 \le \beta \le \beta_{\max}$
(a) calculate $\hat{a}_\beta$ and $\hat{C}(\rho_d, \theta_d)$ as in step 1.
(b) Using a pseudo-random number generator, draw independent bootstrap residuals $\tilde{r}_d^*$ with replacement, from the empirical distribution of $\tilde{r}_d$.
(c) Define the bootstrap surface

$$C^*(\rho_d, \theta_d) = \hat{C}(\rho_d, \theta_d) + \tilde{r}_d^*.$$

(d) Using $C^*(\rho_d, \theta_d)$ as the *new* surface, compute the least-squares estimate of a_β, $\hat{a}_\beta^*$, and calculate

$$\hat{C}^*(\rho_d, \theta_d) = \sum_{p=1}^{\beta} \hat{a}_p^* Z_p(\rho_d, \theta_d)$$

and

$$\mathsf{SSE}_{D,L_D}^*(\beta) = \frac{1}{D}\sum_{d=0}^{D}\left(C(\rho_d, \theta_d) - \hat{C}^*(\rho_d, \theta_d)\right)^2.$$

(e) Repeat steps (b)–(d) a large number of times (e.g. 100) to obtain a total of B bootstrap statistics $\mathsf{SSE}_{D,L_D}^*(\beta)_1, \ldots, \mathsf{SSE}_{D,L_D}^*(\beta)_B$, and estimate the bootstrap mean-square error

$$\hat{\Gamma}_{D,L_D}(\beta) = \frac{1}{B}\sum_{b=1}^{B}\mathsf{SSE}_{D,L_D}^*(\beta)_b.$$

5. Choose β for which $\hat{\Gamma}_{D,L_D}(\beta)$ is a minimum.

where

$$\mathbf{C}^* = \mathbf{Z}_\beta \mathbf{a}_\beta + \varepsilon^*$$

with ε^* being a column vector containing elements ε_d^*, $d = 1, \ldots, D$. The model order selected by the bootstrap, denoted by $\hat{\beta}_{D,L_D}$, is then the minimizer of

$$\begin{aligned}\hat{\Gamma}_{D,L_D}(\beta) &= \mathsf{E}_*\frac{1}{D}\sum_{d=1}^{D}\left(C(\rho_d, \theta_d) - \sum_{p=1}^{\beta}\hat{a}_p^* Z_p(\rho_d, \theta_d)\right)^2 \\ &= \mathsf{E}_*\frac{1}{D}\|\mathbf{C} - \mathbf{Z}_\beta \hat{\mathbf{a}}_\beta^*\|^2\end{aligned}$$

over $\beta = 1, \ldots, P$, where E_* is the asymptotic expectation conditioned on the measured data [9]. A detailed procedure for selecting the model order is given in Table I.

The first step in the procedure is to choose an arbitrary large number of terms $\beta_{\max}$ and perform a fit of Zernike polynomials to the surface data using a least-squares procedure. Next, we

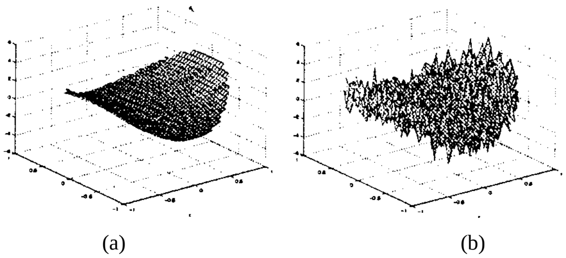

Fig. 1. The model used in performance analysis: pure surface $C_1(\rho,\theta)$ representing (a) Seidel's astigmatism and (b) surface and measurement noise with $\sigma^2 = 1$.

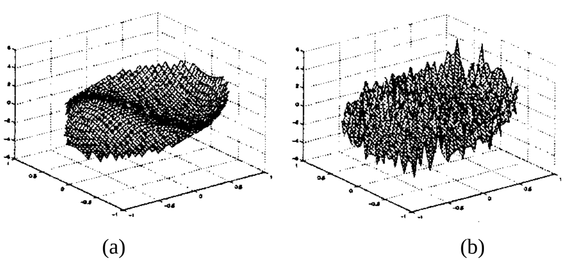

Fig. 2. The models used in performance analysis: pure surface $C_2(\rho,\theta)$ representing (a) coma and (b) surface and measurement noise with $\sigma^2 = 1$.

calculate the residuals by subtracting the fitted surface from the original one. In the third step, we detrend and rescale the residuals. After that, we perform a bootstrap order selection procedure (steps 4(a)–4(e)) for all orders starting from 1 and ending at $\beta_{\max}$ in which we obtain the bootstrap mean-square error as a function of the model order. The last step is to choose that model order which results in the minimum bootstrap mean-square error.

A Matlab code of the bootstrap procedure for calculating the optimal number of Zernike terms can be obtained at no cost by contacting the authors.

IV. Simulation and Experimental Results

We now demonstrate the power of the proposed methodology using the following surface models

$$\begin{aligned} C_1(\rho,\theta) &= \frac{1}{2} + \frac{1}{2}Z_4(\rho,\theta) + Z_5(\rho,\theta) + \varepsilon \\ C_2(\rho,\theta) &= Z_7(\rho,\theta) + \varepsilon \end{aligned} \tag{5}$$

where $\rho \in [0,1]$ and $\theta \in [0,2\pi]$. The first surface represents Seidel's regular astigmatism while the second surface represents horizontal coma.

The measurement zero-mean and unit variance Gaussian noise process ε 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 the bootstrap algorithm. In Figs. 1 and 2, the model surfaces are shown together with a realization of the measurement noise added to each surface. For illustration purposes, the surfaces in (5) have been sampled at four rings corresponding to radius $\rho_1 = 0.25$, $\rho_2 = 0.5$, $\rho_3 = 0.75$, and $\rho_4 = 1$, and at 36 equally spaced semimeridians leading to a sample of $D = 144$ data points. Such sampling is equivalent to placido disk instruments which also adopt the polar coordinates, though the sample size in the latter is usually much larger. We then applied the

TABLE II
Empirical Probability (in Percent) of Selecting the Model Order of the Surface $C_1(\rho,\theta)$ for $D = 144$ and $L_D = 12$

β	$\hat{\Gamma}_{D,L_D}$	AIC	MDL	HQ	AIC_C
5	**99.7**	**68.3**	**96.3**	**51.1**	**74.3**
6	0.3	7.8	3.1	7.8	8.3
7	0.0	6.4	0.6	8.0	6.2
8	0.0	5.2	0.0	6.7	4.5
9	0.0	1.7	0.0	3.0	1.3
10	0.0	1.4	0.0	3.2	0.8
11	0.0	2.7	0.0	4.0	1.7
12	0.0	2.0	0.0	4.3	0.9
13	0.0	2.0	0.0	4.1	1.2
14	0.0	1.2	0.0	3.7	0.2
15	0.0	1.3	0.0	4.1	0.6

$\hat{\Gamma}_{D,L_D}(\beta)$—Bootstrap, AIC($\beta$)—Akaike information criterion, MDL(β)—minimum description length, HQ(β)—Hannan and Quinn criterion, $\text{AIC}_C(\beta)$—corrected Akaike information criterion.

TABLE III
Empirical Probability (in Percent) of Selecting the Model Order of the Surface $C_2(\rho,\theta)$ for $D = 144$ and $L_D = 12$

β	$\hat{\Gamma}_{D,L_D}$	AIC	MDL	HQ	AIC_C
7	**99.2**	**65.5**	**97.2**	**52.2**	**73.4**
8	0.8	10.3	2.3	10.0	9.0
9	0.0	6.3	0.4	7.2	6.4
10	0.0	2.1	0.0	4.2	2.1
11	0.0	4.9	0.1	6.3	3.1
12	0.0	3.2	0.0	4.5	2.2
13	0.0	2.8	0.0	5.6	1.5
14	0.0	2.8	0.0	5.4	1.2
15	0.0	2.1	0.0	4.6	1.1

$\hat{\Gamma}_{D,L_D}(\beta)$—Bootstrap, AIC($\beta$)—Akaike information criterion, MDL(β)—minimum description length, HQ(β)—Hannan and Quinn criterion, $\text{AIC}_C(\beta)$—corrected Akaike information criterion.

bootstrap algorithm from Table I to estimate the model order in each case. The maximum model order and the number of bootstrap replications were chosen to $\beta_{\max} = 15$ and $B = 200$, respectively.

In Tables II and III, we show the empirical probabilities of selecting a particular model (evaluated over 1000 independent runs) together with the results obtained by using classical model selection techniques.

It should be noted that none of the methods underestimates the true model order. In the test we ran, it was clear that the proposed bootstrap-based technique performs very well. Over 1000 replications, the empirical probability that the method selects the true order was very close to one. The only traditional technique that provided comparable performance to the bootstrap was the MDL. All other model order selection criteria tend to over-parameterization (i.e., over-estimated the true model order).

One may argue that the choice of L_D (in our case $L_D = 12$) may be a problem in practical situations. We noted that for a very high noise level ($\sigma^2 > 10$) this parameter needs to be increased as the variability of bootstrap residuals is already high. For more guidelines as to the choice of the parameter L_D in linear regression the reader is referred to Shao [10].

In many videokeratoscopic applications, the level of the measurement noise may not be known. The scaling parameter L_D can be tuned in such applications in the following manner. First,

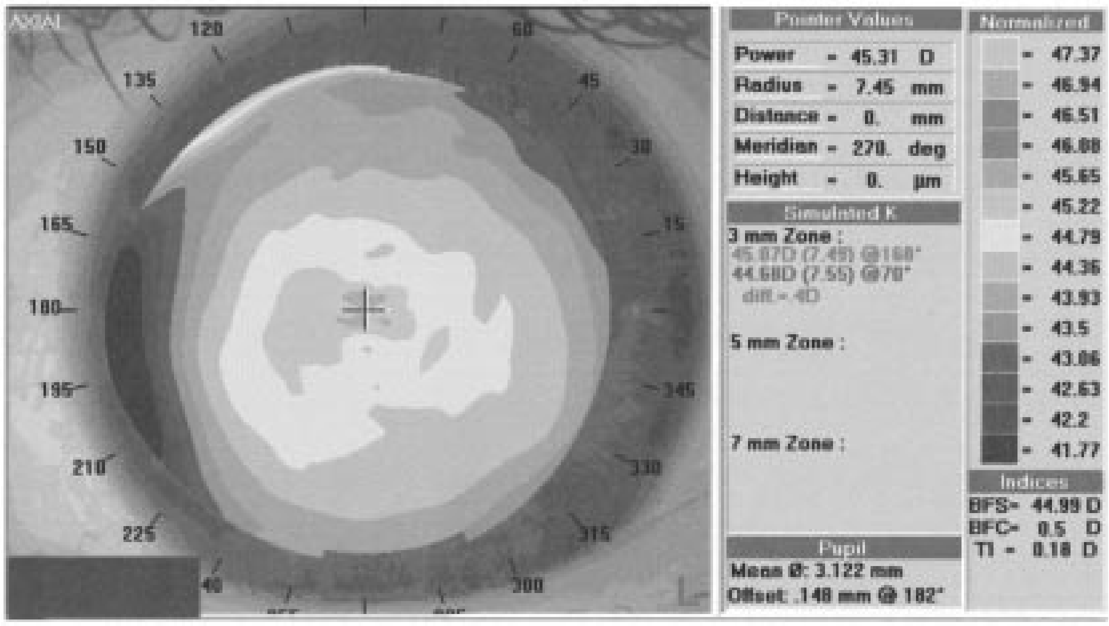

Fig. 3. Corneal axial power for subject A (normal topography).

we obtain height data from a known surface, typically it is a paraboloid. The optimal number of Zernike terms for a paraboloid should be four. We then tune the parameter L_D in our procedure until we achieve the desired number of terms.

A. Fitting Zernike Polynomials to Corneal Elevation

In the following we provide an example of fitting Zernike polynomials to the corneal elevation data measured by a videokeratoscope (Optikon Keratron). The Keratron data are in the format of 26 rings and 256 semimeridians.

First, we chose a subject with a normal cornea. From the videokeratoscope data we select a certain portion of the cornea around the instrument axis, for example an 8-mm diameter. The axial power corneal map for subject A from the Keratron videokeratoscope is shown in Fig. 3. Axial power is a common form of videkeratoscope data presentation and is calculated by finding a normal to the surface and intersecting this normal with the instrument axis. The reciprocal of this distance multiplied by the refractive index minus one is the axial dioptric power. Subject A's axial power map shows a typical decrease in axial power toward the periphery of the cornea with only minor asymmetry.

We ran the proposed bootstrap algorithm and the other information criteria to estimate the required number of Zernike terms for the corneal elevation. The maximum model order and the number of bootstrap replications were chosen to $\beta_{\max} = 40$ and $B = 200$, respectively. The scaling parameter L_D was estimated in this case to be $L_D = 0.05$. This has ensured that the bootstrap residuals have some variability because the level of the measurement noise is quite small (less than 3 microns).

In Fig. 4, we show the results of the model order selection criteria for fitting Zernike polynomials to corneal elevation for subject A. The minimum of $\hat{\Gamma}_{D,L_D}(\beta)$ is denoted by a dashed vertical line which also indicates the optimal number of terms. It should be noted that all the classical model selection techniques cannot find a minimum as long as the number of Zernike terms is

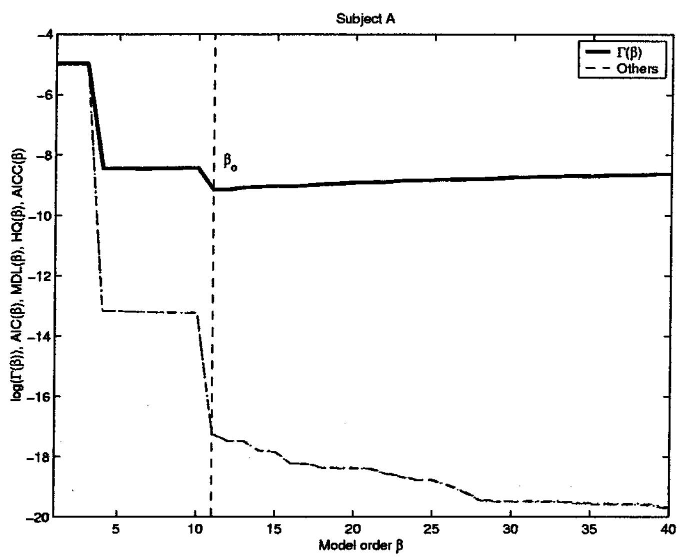

Fig. 4. The results of the model order selection criteria for fitting Zernike polynomials to corneal elevation for subject A. The minimum of $\hat{\Gamma}_{D,L_D}(\beta)$ is denoted by a dashed vertical line which also indicates the optimum number of terms.

less than 600. This is because the penalty part of each of the criterion is too small compared to the logarithm of the residual variance. However, the bootstrap-based technique, in contrast to traditional methods, has the ability to select an optimal model order much earlier by appropriately setting the scaling parameter L_D. As noted in Fig. 4, the optimal model-order for the subject A was found to be 11, a result that is expected for a normal cornea, which is a prolate ellipsoid with small levels of aberrations.

Next, we repeated the above procedure for several subjects with different corneal aberrations, in particular:

1) subject B who has a significant amount of astigmatism;
2) subject C who has a decentred corneal apex;
3) subject D who has keratoconus;

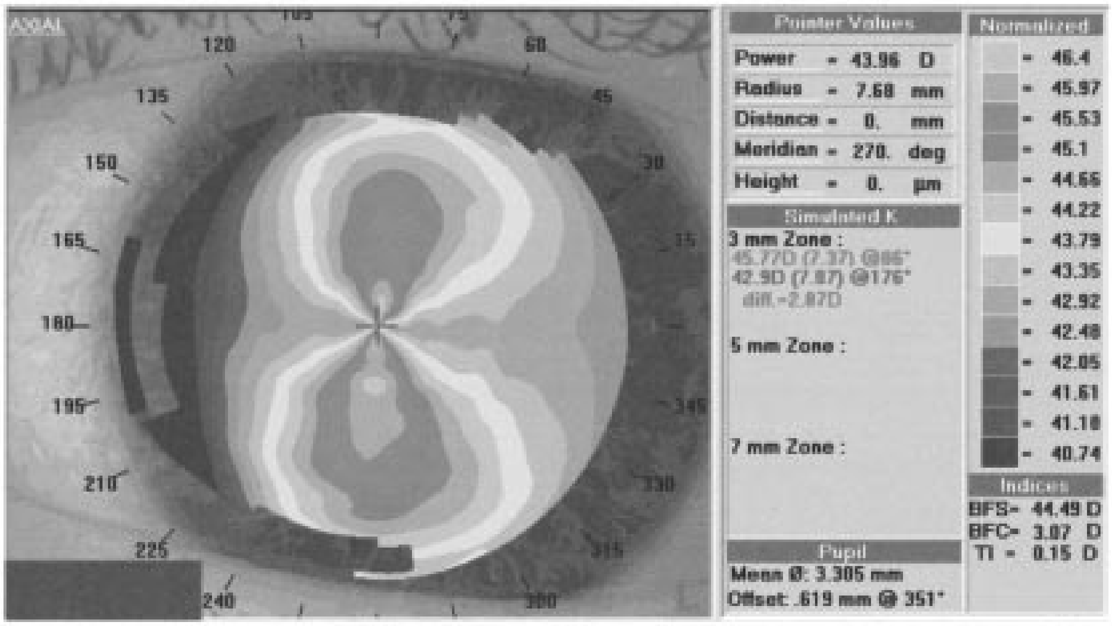

Fig. 5. Corneal axial power for subject B (astigmatism).

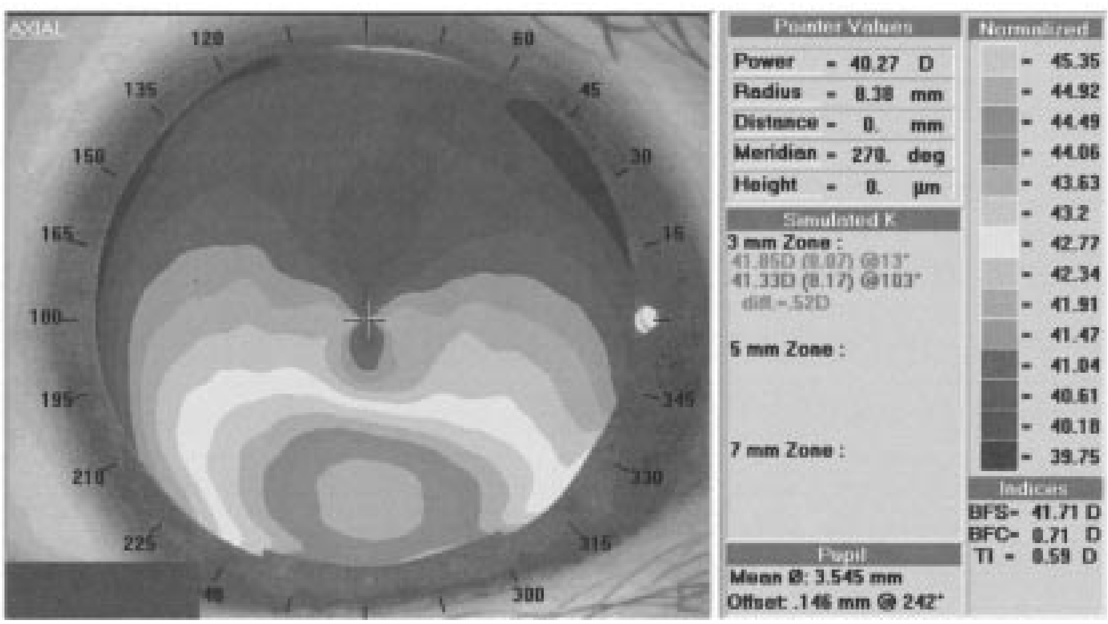

Fig. 6. Corneal axial power for subject C (decentred corneal apex).

4) subject E who has undergone a poorly centred refractive surgery procedure.

Axial power maps of the corneas of these subjects, from the Keratron are shown Figs. 5–8. The corresponding number of optimal Zernike terms was found to be 11, 14, 8, and 12, respectively. As previously observed, the classical model selection techniques could not find a minimum as long as the number of Zernike terms was less than several hundred.

The axial power map for subject B (Fig. 5) shows classical with-the-rule astigmatism, where the vertical meridian of the cornea is steeper (higher axial power) than the horizontal meridian. Corneal astigmatism gives rise to a characteristic bow-tie pattern of axial power. The astigmatism is modeled theoretically by the fifth and sixth Zernike terms.

Subjects C and D (Figs. 6 and 7) show axial power maps which are characteristic of a decentred corneal apex. In the case of Subject D, this is caused by a degenerative thinning of the cornea leading to a protrusion of the corneal surface in the inferotemporal region (the condition called keratoconus).

The corneal topography of subject E's eye is unusual. The subject underwent a poorly centred refractive surgery procedure (laser in situ keratomileusis) which resulted in regions of very low axial power in the superior cornea, compared to relatively steeper normal axial powers in the inferior cornea.

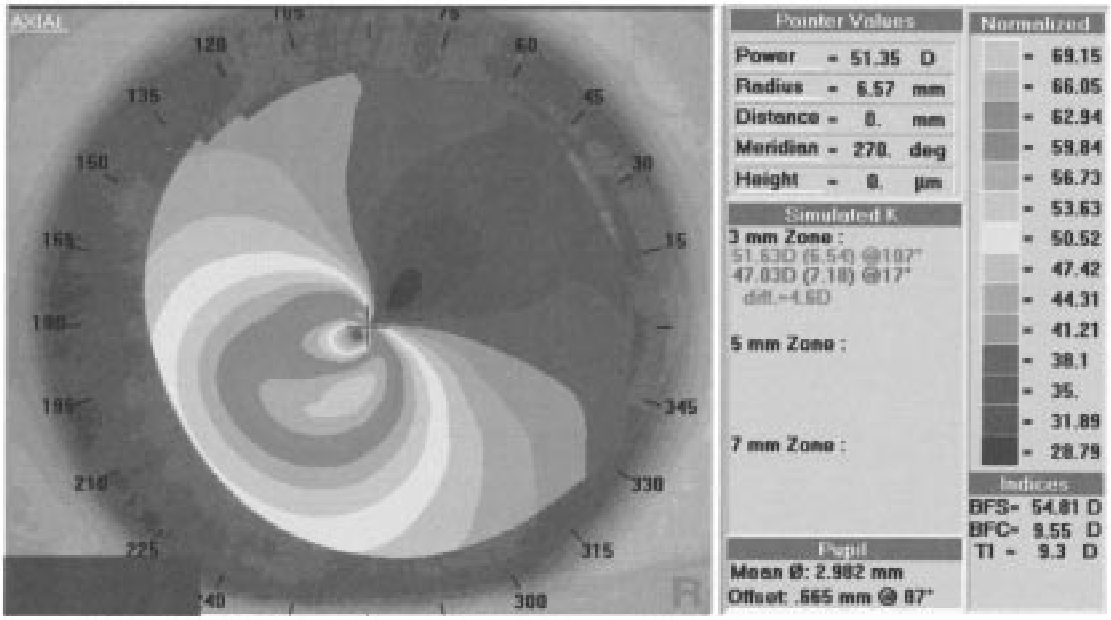

Fig. 7. Corneal axial power for subject D (keratoconus).

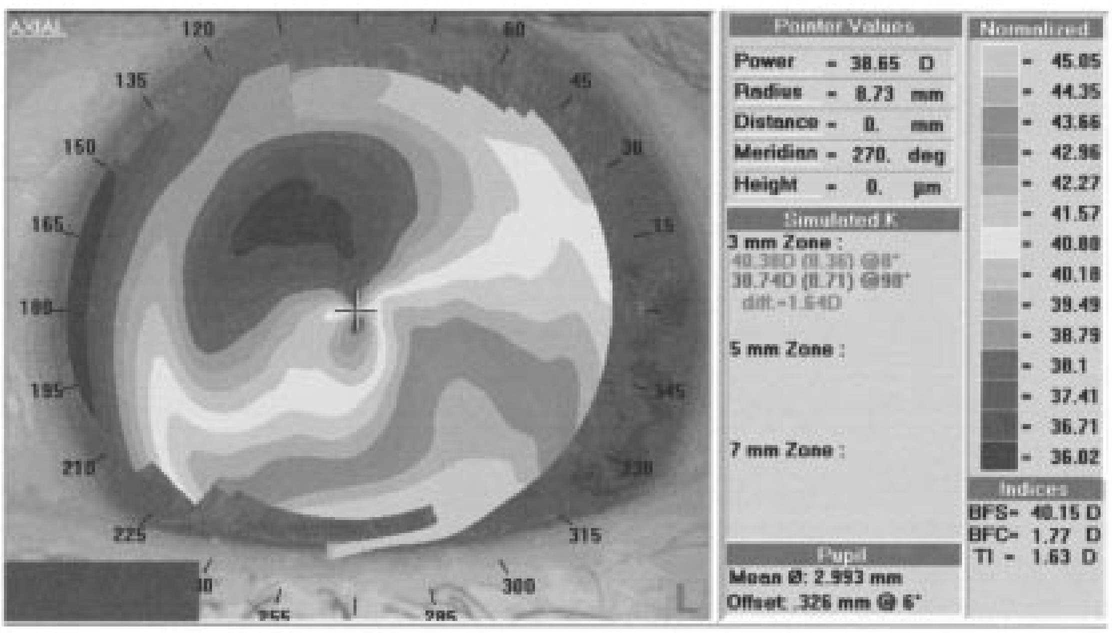

Fig. 8. Corneal axial power for subject E (decentred laser in situ keratomileusis).

B. Discussion

It can be noticed that in both the normal and astigmatic cornea the number of optimal Zernike terms was found to be eleven. This is consistent with the common view that the cornea can be approximated by a conic (elliptical) surface. For distorted corneas, the number of optimal Zernike terms varies from subject to subject.

Having determined the optimal model order, β_o, one can proceed to determine the optimal model, that is the set of Zernike terms, being a subset of $\{Z_1(\rho,\theta),\ldots,Z_{\beta_o}(\rho,\theta)\}$. To do so, we can use a similar bootstrap procedure to the one described in [13] and [14]. The computational intensity of such a procedure is quite significant, especially when β_o is large. For completeness, however, and for the purpose of illustration we have used such a bootstrap procedure to determine the best Zernike expansion model for each of the considered corneas. In Table IV, we list the sets of Zernike terms for each of the corneas.

We can see from the data in Table IV that as expected, when the shape of the cornea becomes more irregular, the total number of Zernike terms required for modeling the cornea becomes higher. For example, the normal cornea of subject A requires only five Zernike terms whereas the cornea of subject E requires ten terms. This result indicates that by using arbitrary number of Zernike terms, for example 15, leads in most cases to over-parameterization of the model.

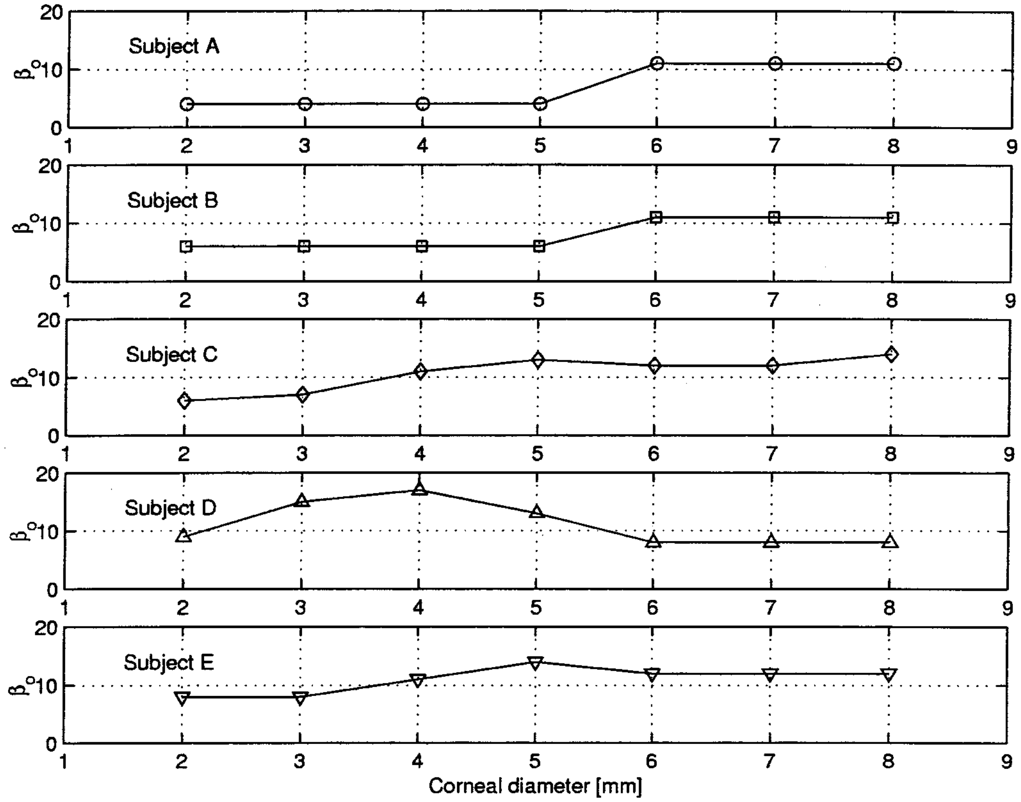

Fig. 9. The results of the model order selection criteria for fitting Zernike polynomials to corneal elevation as a function of pupil diameter for subjects A–E. The y-axis β_o is the optimal model order for the corresponding corneal diameter.

TABLE IV
THE OPTIMAL SET OF ZERNIKE TERMS AS DETERMINED BY THE BOOTSTRAP PROCEDURE FOR SUBJECTS A–E AND FOR CORNEAL DIAMETER OF 8 mm. THE TICK AND THE CROSS DENOTE A SIGNIFICANT AND AN INSIGNIFICANT TERM IN THE ZERNIKE POLYNOMIAL EXPANSION, RESPECTIVELY

Zernike term	Subject A	B	C	D	E
Z_1	√	√	√	√	√
Z_2	×	√	√	√	√
Z_3	√	√	√	√	√
Z_4	√	√	√	√	√
Z_5	×	√	×	√	×
Z_6	×	√	√	√	×
Z_7	√	√	√	√	√
Z_8	×	√	×	√	√
Z_9	×	×	√		√
Z_{10}	×	×	×		√
Z_{11}	√	√	√		√
Z_{12}			×		√
Z_{13}			×		
Z_{14}			√		
Total	5	9	9	8	10
MSE_{15}	2.8e-7	7.1e-7	4.4e-7	1.5e-6	4.4e-6
MSE_β	2.1e-6	1.1e-6	2.7e-6	2.1e-6	6.0e-6

It is known that using more terms in polynomial expansion will reduce the MSE. However, the amount of increase in error induced by choosing less Zernike terms is of interest. We have calculated the MSE based on fitting the first 15 Zernike terms, MSE_{15}, to the surface as well as the optimal number of terms, MSE_β (see Table IV). It could be seen that reduction in the MSE is small and that the MSE is smaller than the instrument error, which is often estimated at 1–5 micrometers.

It is also interesting to see how the optimal number of Zernike terms varies with corneal diameter. This is particulary important when determining the corneal aberrations at high and low levels of light (i.e., corresponding to small and large pupil diameters). In Fig. 9, we show the optimal number of Zernike terms, determined by the bootstrap procedure, as a function of pupil diameter for all considered subjects.

Subject A and B show systematic change from 4 and 6 Zernike terms (respectively) for a small corneal diameter to 11 Zernike terms for a large one. This trend reflects the elliptical nature of normal corneas. However, for distorted corneas there is no clear trend in the optimal number of Zernike terms as the diameter varies. For these subjects the interactions between optimal number of Zernike terms and corneal diameter is less predictable than in the case of regular corneas.

The proposed bootstrap methodology is computationally intensive. To calculate the optimal number of Zernike terms for the given videokeratoscope data ($D = 256 \times 26 = 6656$), takes about 10 min on a Pentium II, 450-MHz computer, for $\beta_{\max} = 40$ and $B = 200$. This corresponds to approximately 40 Gigaflops. On the other hand, calculating the optimal order took only 6.5 s (1.2 Megaflops) for $D = 144$, $\beta_{\max} = 15$, and $B = 200$.

V. SUMMARY

We have proposed a procedure for determining the model order of Zernike polynomial expansion using the bootstrap. The method is based on minimizing bootstrap estimates of the mean-square prediction error. The method achieves a very high probability of selecting the true model, irrespective of the statistical distributions of the measurement noise, while the sample size is small. It outperforms the classical model selection techniques such as the AIC, MDL, HQ, and AIC_C. The proposed method was applied to estimate the number of Zernike terms when fitting them to corneal elevations. However, it could be easily adopted to perform the fit to other optical surfaces or wavefronts.

ACKNOWLEDGMENT

The authors would like to thank Prof. A. M. Zoubir from the Australian Telecommunications Research Institute, Curtin University of Technology, Perth, Australia, for encouragement to use bootstrap techniques.

REFERENCES

[1] R. A. Applegate and H. C. Howland, "Noninvasive measurement of corneal topography," *IEEE Eng. Med. Biol. Mag.*, vol. 14, no. 1, pp. 30–42, 1995.
[2] V. N. Mahajan, *Optical Imaging and Aberrations: Ray Geometrical Optics*. Bellingham, WA: SPIE Opt. Eng., 1998.
[3] J. Schweigerling and J. E. Greivenkamp, "Using corneal height maps and polynomial decomposition to determine corneal aberrations," *Optometry Vision Sci.*, vol. 74, no. 11, pp. 906–916, 1997.
[4] R. J. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Amer.*, vol. 66, no. 3, pp. 207–211, 1976.
[5] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automatic Contr.*, vol. AC-19, pp. 716–723, Dec. 1974.
[6] J. Rissanen, "Modeling by shortest data decription," *Automatics*, vol. 14, pp. 465–471, 1978.
[7] E. J. Hannan and B. G. Quinn, "The determination of the order of an autoregression," *J. Roy. Stat. Soc.*, ser. Ser. B, vol. 41, pp. 190–195, 1979.
[8] C. M. Hurvich and C. L. Tsai, "Regression and time series model selection in small samples," *Biometrika*, vol. 76, pp. 297–307, 1989.
[9] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*. London, U.K.: Chapman & Hall, 1993.
[10] J. Shao, "Bootstrap model selection," *J. Amer. Statist. Assoc.*, vol. 91, pp. 655–665, 1996.
[11] A. M. Zoubir and B. Boashash, "The bootstrap and its application in signal processing," *IEEE Signal Processing Mag.*, vol. 15, no. 1, pp. 56–76, 1998.
[12] C. J. Kim and R. R. Shannon, "Catalog of Zernike polynomials," in *Applied Optics and Optical Engineering*, R. R. Shannon and J. C. Wyant, Eds. New York: Acedemic, 1980, vol. X, pp. 193–221.
[13] A. M. Zoubir, "Model selection: A bootstrap approach," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, ICASSP 99*, vol. III, Phoenix, AZ, 1999, pp. 1853–1856.
[14] A. M. Zoubir and D. R. Iskander, "Bootstrap modeling of a class of nonstationary signals," *IEEE Trans. Signal Processing*, vol. 48, pp. 399–408, Feb. 2000.

D. Robert Iskander (M'98) received the Magister Inżynier degree in electronic engineering from the Technical University of Lodz, Poland, in 1991, and the Ph.D. degree in signal processing from Queensland University of Technology, Australia, in 1997.

From 1996 to 1998, he was a Post-Doctoral Fellow at the Signal Processing Research Centre and the Cooperative Research Centre for Satellite Systems, QUT, Australia. In 1998, he joined the Centre for Eye Research, QUT, where he is currently a Research Fellow. His research interests include statistical signal processing, radar and radio navigation, information theory and visual optics.

Michael J. Collins received the Dip App Sc (Optom), M App Sc, and Ph.D degrees from Queensland University of Technology, Brisbane, Brisbane, Australia, in 1977, 1988, and 1996 respectively.

He is an Associate Professor in the School of Optometry at the Queensland University of Technology, Brisbane, Australia. His research laboratory, the Contact Lens and Visual Optics Laboratory, specialises in the visual and optical characteristics of the cornea and contact lenses.

Dr. Collins is a member of the Optometrists Association of Australia, Fellow of the American Academy of Optometry, a Fellow of the Contact Lens Society of Australia and a member of the International Society for Contact Lens Research.

Brett Davis is a Senior Research Assistant in the Centre for Eye Research at the Queensland University of Technology, Brisbane, Australia (QUT). received the Bachelor of Applied Science degree in physics from Queensland University of Technology, Brisbane, Australia, in 1990.

Since then, he has been involved in various research projects within the Contact Lens and Visual Optics Laboratory, Centre For Eye Research, QUT. His interests include visual optics and optical design.