

# Automated decision tree classification of keratoconus from videokeratography

Michael D. Twa, OD, MS, Srinivasan Parthasarathy, PhD, Mark A. Bullimore, MCOptom, PhD, Cynthia Roberts, PhD, Ashraf M. Mahmoud, Tom Raasch, OD, PhD, and David J. Schanzlin, MD  
College of Optometry The Ohio State University: [twa.1@osu.edu](mailto:twa.1@osu.edu)

## 1 INTRODUCTION

Classification and prediction are common tasks in the biomedical sciences. Several machine learning classification methods exist including Bayesian classifiers, decision trees, neural networks, statistical regression, and others.<sup>1</sup> Each approach has unique strengths that make it more or less appropriate for a particular classification problem. Decision trees have emerged as one of the most versatile and robust classification methods,<sup>2</sup> and have been widely applied to medical diagnosis problems.

We have developed a quantitative method of corneal shape classification to discriminate keratoconus from normal videokeratography.<sup>3,4</sup> This decision tree induction method can be applied to the raw data output from any videokeratography instrument platform, and could be used to analyze both group and longitudinal data. In this study we describe an alternative method of corneal shape classification based upon Zernike polynomials and compare performance of this classification method with existing classifiers for a sample of normal eyes and eyes diagnosed with keratoconus.

## 2 METHODS

### 2.1 Patient Selection

All keratoconus patients selected for this study were examined between August 1998 and January 2000. We first identified potential patients from ICD-9 diagnosis codes and then confirmed their diagnostic classifications by chart review, applying diagnostic criteria established by the Collaborative Longitudinal Evaluation of Keratoconus (CLEK) study.<sup>5</sup> The ocular findings that defined keratoconus were: (1) an irregular cornea determined by distorted keratometry mires, distortion of the retinoscopic or ophthalmoscopic red reflex (or a combination of these); (2) at least one of the following biomicroscopic signs: Vogt's striae, Fleischer's ring of  $>2$  mm arc, or corneal scarring consistent with keratoconus.

We excluded eyes that did not have videokeratography examination data or cases where the videokeratography examination data clearly contained artifacts due to obstruction of the corneal surface by the eyelids. All other eligible cases were included. Normal eyes were selected sequentially from healthy patients that presented for surgical correction of refractive error during the same time interval. We included cases that were myopic with or without astigmatism, had videokeratography examination data, and had no other documented ocular disease. No normal eyes were excluded because of missing or poor quality videokeratography data.

### 2.2 Data Preprocessing

We exported videokeratography data from a single image consisting of elevation, angle and radial position coordinates from the Keratron corneal topographer (v3.49, Optikon 2000, Rome, Italy). The data from each examination consisted of 6,900 individual points in a 3-dimensional polar coordinate grid. In addition to the elevation data we recorded the axial curvature valued associated with each individual point.

To model the corneal surface we computed a 10th order expansion of Zernike polynomials over the central 5 mm from the exported elevation data using methods described by Schweigerling et al.<sup>6</sup> We limited our area of analysis to the central 5mm to insure complete data over the area of analysis for all eyes in this sample.

### 2.3 Decision Tree Classification

A decision tree is a collection of if-then conditional rules for assignment of class labels to instances of a data set. Decision trees consist of nodes that specify a particular attribute of the data, branches that represent a test of each attribute value, and leafs that correspond to the terminal decision of class assignment for an instance in the data set (Figure 1).

We used C4.5, a decision tree induction algorithm, to construct a classifier for our data set. The

data were structured as rows of records for each eye (records) and columns of Zernike coefficients (attributes). We selected Zernike polynomial coefficients to represent corneal surface features and used the magnitude of these coefficients as a basis to discriminate between normal and keratoconus records in our record set. Each record of the data set was labeled with a known class assignment from chart review.

The C4.5 algorithm is a supervised machine learning method where information about true class assignments are known during construction of the classifier. The algorithm begins with all records in a single pool, a heterogeneous collection of both normal and keratoconus records. The C4.5 algorithm then iteratively selects individual attributes and their values to determine a best separation of the records. The benefit of using the C4.5 decision tree induction algorithm is automated selection of the most relevant Zernike coefficients and discovery of the coefficient value necessary for optimal class separation.

We have previously described our decision tree induction methods and results from optimization experiments.<sup>3,4</sup> In summary, we used WEKA (version 3.2.3) a Java based implementation of C4.5 release 8. The algorithm was set to overfit the data generating trees of maximal complexity with pruning criteria set to a minimum of two records per leaf. We used 10-fold cross validation to estimate the generalization error of our decision tree classifier.

## 2.4 Other Videokeratography Classifiers

Using this same set of videokeratography records, we computed five additional keratoconus classification indices including: Modified Rabinowitz-McDonnell Index (*RM*),<sup>7</sup> *KISA%*,<sup>7</sup> Keratoconus Prediction Index (*KPI*),<sup>8</sup> *Z3*,<sup>6</sup> and Cone Location and Magnitude Index (*CLMI*).<sup>9</sup> These indices were calculated from published formulas and computed from the appropriate curvature or elevation videokeratography data.

## 2.5 Statistical Methods

Sensitivity, specificity, and accuracy are the metrics most commonly used to evaluate the quality of a classifier. In the context of this study, sensitivity is defined as the number of correct identifications of keratoconus from all eyes with keratoconus. Specificity is the proportion of keratoconus eyes correctly identified by the classifier; accuracy is the proportion of total correct classifications of both classes out of all eyes in the sample.

We represent the tradeoff between sensitivity and specificity by Receiver Operating Characteristic

(ROC) Plots. The ROC curve shows the tradeoff between sensitivity and specificity over the entire range of possible cut point criteria. The overall quality of a classifier can be judged by the area under the curve, and this area can be used for statistical comparisons between classifiers.

We compared each classifier considered in this study under two different conditions. First, we computed each classification index as published and compared the area under the ROC curve, accuracy, sensitivity, and specificity using our decision tree method (*C4.5*) as the standard for comparison. Second, whenever possible, we normalized each classification index for the data in our sample and then compared the area under the ROC curve, accuracy, sensitivity and specificity as before, using *C4.5* as the standard for comparison.

## 3 RESULTS

Our sample consisted of 82 normal eyes from 46 patients and 99 eyes diagnosed with keratoconus from 54 patients.

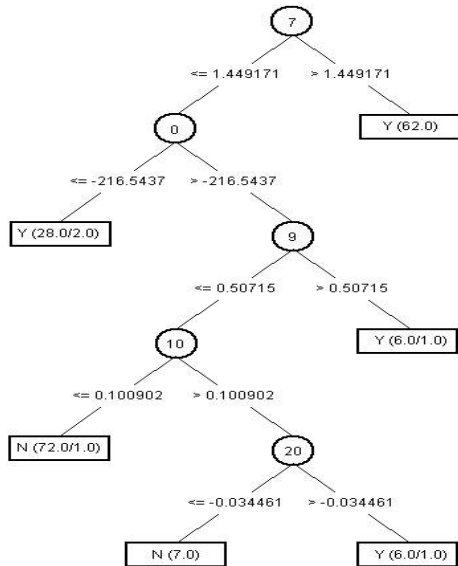
### 3.1 Standard Classifier Performance

We compare the accuracy sensitivity and specificity of our decision tree classifier (Figure 1) to other Keratoconus classification methods in Table 1.

**Table 1: Comparison of Standard Classifiers**

Classifier	Cut point ( $\geq$ )	Sensitivity (%)	Specificity (%)	Accuracy (%)
<i>C4.5</i>	1	97.98	91.46	95.03
<i>CLMI</i>	3	79.80	98.78	88.4
<i>RM (D)</i>	47.2 (1.4)	90.91	82.93	87.29
<i>KISA%</i>	300	90.91	75.61	83.98
<i>Z3 (<math>\mu m</math>)</i>	2.33	48.48	100	71.82
<i>KPI</i>	0.23	94.95	28.05	64.64

We found the best sensitivity and accuracy for our decision tree classification method. Our *CLMI* classification method also had very good accuracy. Although it had greater specificity, the sensitivity was less than the decision tree classifier. The *RM* index and *KISA%* classification methods performed well on this data. The *Z3* index had excellent specificity, but much worse sensitivity than other methods. The *KPI* index had excellent sensitivity but very poor specificity and was least accurate of any of the classifiers judged.



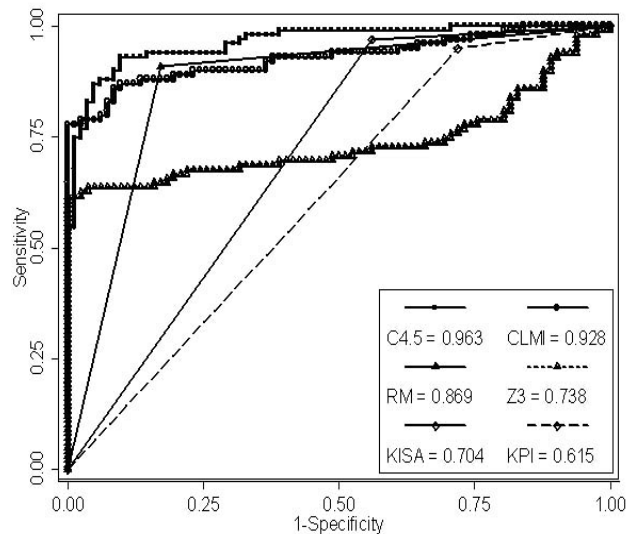
**Figure 1:** Decision Tree Classification of Keratoconus Classification attributes (circled numbers) are OSA single index Zernike polynomial coefficients. Split criteria label each branch of the decision tree. Terminal nodes (boxes) are labeled with class assignments (Y = Keratoconus, N = Normal) and number of records assigned to this class (correct / incorrect).

In Figure 2, we plot the ROC curves for each of these classifiers. We make a statistical comparison of the differences between classification methods by comparing the area underneath each ROC curve to our standard, the *C4.5* decision tree classifier (Table 1). Our *CLMI* classifier was not statistically significantly different from our standard method. Each of the other classification methods, *RM* index, *KPI* index, *Z3* index, and *KISA%* had significantly less area underneath the ROC curve. (all  $P < .001$ )

### 3.2 Normalized Classifier Performance

We modified the threshold criteria for each keratoconus classification index selecting cut points that optimized classification accuracy. In Table 2, we show optimal cut points determined for our sample. After this optimization, we continued to find differences in sensitivity and specificity among the classifiers. However accuracy was more homogenous and was improved for all methods. Our *C4.5* classifier continued to have the best sensitivity and accuracy. Our other index, *CLMI* had the best specificity. The only observed decrease after cut point optimization was a decrease in sensitivity of the *KPI* and *KISA%* indices.

There was a significant increase in area underneath



**Figure 2:** Figure 2-Keratoconus Classifier ROC Curve Comparisons. Area under the ROC curve for classifier is given in the legend

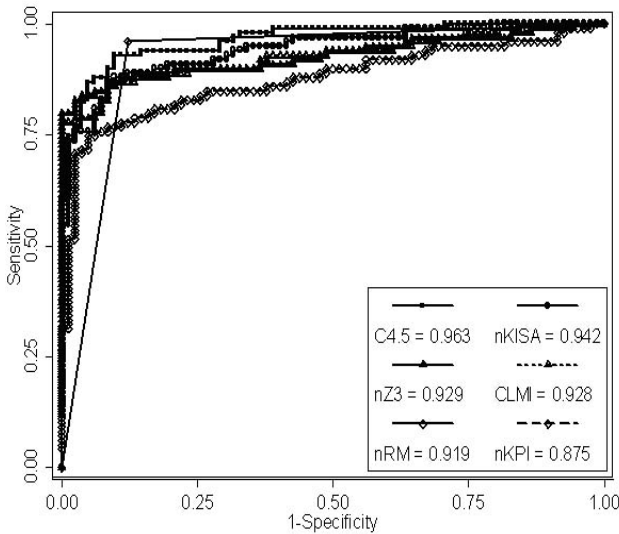
each ROC curve after a point optimization (all  $p < .006$ ). Individual ROC curves (Figure 3) were very similar after optimization. As before, we compared area under each curve to our standard, the decision tree classifier. We found no statistically significant difference in areas underneath each ROC curve when compared with our standard except that the *KPI* index was significantly less.

## 4 DISCUSSION

In this study we described a new method of quantitative videokeratography analysis. Using a Zernike polynomial representation of corneal surface features, we demonstrate that it is possible to automate classification of corneal shape based on surface features modeled by Zernike polynomials. Attributes selected during classifier induction have empirical validity and corroborate familiar clinical findings associated with keratoconus, namely inferior-temporal steepening.

**Table 2: Comparison of Normalized Classifiers**

Classifier	Cut point ( $\geq$ )	Sensitivity (%)	Specificity (%)	Accuracy (%)
<i>C4.5</i>	1	97.98	91.46	95.03
<i>nRM (D)</i>	45.97 (1.77)	95.96	87.80	92.27
<i>nZ3 (<math>\mu\text{m}</math>)</i>	0.88	83.84	96.34	89.50
<i>nKISA%</i>	575	87.88	90.24	88.95
<i>CLMI</i>	3	79.80	98.78	88.40
<i>nKPI</i>	11.83	78.79	86.59	82.32



**Figure 3:** Figure 2-Normalized Keratoconus Classifier ROC Curve Comparisons. Area under the ROC curve for classifier is given in the legend

This method has excellent sensitivity, specificity, and accuracy. Since this method is based on decision tree induction, it has additional advantages including faster computational speed, scalability for analysis of large data sets, and better interpretability than neural networks or other machine learning methods.<sup>2</sup> Since our classifier is based upon Zernike polynomials calculated from videokeratography elevation data, this classification method is applicable to data from most clinical videokeratography instruments.

Neural network methods described by Smolek et al. have not been widely adopted.<sup>10</sup> Advantages of greater classification accuracy by this approach are subordinate to several clinical concerns: simplicity, interpretability, and portability. Success of any machine learning method will depend on providing useful additional information to extend knowledge and analytic power for users. In this increasingly complex analysis we were pleased to find the modified Rabinowitz-McDonnell index performed very well even without normalization for our specific data set.

Much of what we learn by classification of videokeratography data is easily learned from other clinical tests. Application of these methods to keratoconus classification is a test bed to demonstrate the value of this approach. However, a more important implication of these results lies in their application to larger quantitative videokeratography data analysis problems. What we have shown with these results is successful application of a method that we will extend this larger and more varied clinical data sets.

## References

- [1] Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning : Data mining, inference, and prediction. Springer series in statistics., New York, Springer, 2001; xvi, 533
- [2] Mitchell TM. Machine learning. New York, McGraw-Hill, 1997; xvii, 414
- [3] Twa MD. Spatial pattern recognition of videokeratography by decision tree classification of Zernike polynomials. Masters Thesis In Vision Science. Columbus, The Ohio State University. 2002; 96
- [4] Twa MD, Parthasarathy S, Raasch TW, Bullimore MA. Decision tree classification of spatial data patterns from videokeratography using zernike polynomials. In: SIAM International Conference on Data Mining. 2003. San Francisco, CA.
- [5] Zadnik K, Barr JT, Edrington TB, et al. Baseline findings in the collaborative longitudinal evaluation of keratoconus (CLEK) study. Invest Ophthalmol Vis Sci 1998; 39:2537-2546.
- [6] Schwiegerling J, Greivenkamp JE. Keratoconus detection based on videokeratoscopic height data. Optom Vis Sci 1996; 73:721-728.
- [7] Rabinowitz YS, Rasheed K. Kisa% index: A quantitative videokeratography algorithm embodying minimal topographic criteria for diagnosing keratoconus. J Cataract Refract Surg 1999; 25:1327-1335.
- [8] Maeda N, Klyce SD, Smolek MK, Thompson HW. Automated keratoconus screening with corneal topography analysis. Invest Ophthalmol Vis Sci 1994; 35:2749-2757.
- [9] Mahmoud AM, Roberts C, Herderick EE, et al. The cone location and magnitude index (CLMI) [ARVO abstract]. Invest Ophthalmol Vis Sci 2001; 42:S898 Abstract nr 4825
- [10] Smolek MK, Klyce SD. Current keratoconus detection methods compared with a neural network approach. Invest Ophthalmol Vis Sci 1997; 38:2290-2299.