# Toward Glaucoma Classification with Moment Methods

A. R. McIntyre[1], M. I. Heywood[1], P. H. Artes[2], S. S. R. Abidi[1]

[1]Faculty of Computer Science
Dalhousie University
Halifax, NS Canada

[2]Department of Ophthalmology
Dalhousie University
Halifax, NS Canada

## Abstract

This paper presents a series of experiments testing the feasibility of employing image-processing techniques for the feature extraction stage in the implementation of a basic optic nerve image classifier. Such a scheme completely removes the need for manually identifying the edge of the optic nerve. In this work, Zernike moments are extracted from Confocal Scanning Laser Tomography images of optic discs for the purposes of classifying the disc as healthy or damaged using a linear discriminant function derived from a linear perceptron. Our preliminary results, when compared with the performance of conventional feature sets, demonstrate the appropriateness of this approach.

## 1 Introduction

It is estimated that approximately 300,000 Canadians suffer from glaucoma, an age-related disease that slowly and painlessly damages the optic nerve, causing loss of vision and potentially blindness. Glaucoma rarely causes symptoms until the later stages of the disease, and epidemiology surveys in North America and Europe have shown that approximately 50% of cases are undetected. As yet there is no single diagnostic test that provides both high sensitivity and specificity.

Medical decision support systems, leveraging an assortment of intelligent techniques, are routinely used by healthcare practitioners for diagnostic support in a variety of clinical scenarios. Medical image processing, vis-à-vis the application of sophisticated image processing techniques to medical images, provides a viable mechanism to not only highlight the salient aspects of the image but also to automatically classify the medical image into pre-defined categories of diagnostic relevance.

At present, glaucoma is diagnosed by visual evaluation of the optic nerve (which requires highly trained specialists) and examination of the visual of vision (which is time-consuming and difficult for inexperienced patients). However, the process has recently been augmented by the availability of *Confocal Scanning Laser Tomography* (CSLT) technology. Such systems provide mean topography and reflectance images from scans of the optic disc.

The availability of CSLT image capture systems naturally provides the potential for the automated classification of optic nerve images. However, historical practice has been to base any diagnostic analysis on *optic disc profiles* or *contours* that require human intervention to accurately define. Thus for each image, a trained professional is required to define the margins of the optic disc profile by manually outlining the optic disc of the patient (a process that is highly subjective in nature). This contour line is then used to extract features (stereometric indices) that are sensitive to the profile provided on which the current classification is based.

As a first step to the fully automatic diagnosis of optic nerve damage, an investigation is made into the applicability of *Moment Methods*. A study is therefore made between features extracted using common practice and those provided by the orthogonal Method of Zernike Moments. A clear preference is demonstrated for the Moment Method under a simple linear classification rule.

## 2 Objective and Methodology

The objective of this work was to qualify the appropriateness of Zernike Moment features derived directly from CSLT topography and reflectance images. Moment Methods are utilized specifically on account of their orthogonal properties and support for invariance's. Moreover, the application domain provides "well framed" image content. The implication being that a feature method based on global image properties facilitated, providing the basis for a single step process for extracting appropriate features (no separate segmentation stage).

At the classification stage, the basic objective is to separate control from glaucoma patient data. The ensuing study considers the significance of feature selection under two classifier systems: linear perceptron, and C5, as follows,

1. Conventional features from HRT software
2. Zernike features extracted from topography images
3. Zernike features extracted from reflectance images
4. Combined topography and reflectance Zernike features

Performance on these feature sets will be measured in terms of sensitivity and specificity for the classifier in question. Pre-processing, based on quartile and Wilcoxon hypothesis testing is used to identify subsets of features appropriate for the linear classifier. A simple pruning technique is used to further identify features that are of most discriminant value to the classifier.

The significance of this study lies in the departure from previous practices in which little or no use of features derived from image processing techniques has been made. Moreover, reflectance images have not been previously used for these purposes.

# 3    Optic Nerve Imaging

The optic nerve consists of ganglion cell fibers which relay the information from the retinal photoreceptors to the visual cortex. Glaucomatous damage manifests as loss of neural tissue at the optic disk where the nerve fibers collectively exit the eye. This damage is often difficult to detect because there is tremendous biological variation in the size and shape of the optic disk even between healthy subjects.

In patients with glaucoma, the spectrum of optic disk appearances is still larger. There are several distinct morphological patterns of optic disk damage, all of which probably carry different prognoses for the progression of the disease and have implications for clinical management [1].

Traditionally, ophthalmologists have relied on visual assessment of the optic disk (by in-vivo microscopy and stereo-photography) to observe the optic nerve. These procedures are time-consuming and highly subjective, giving rise to substantial disagreement even between highly trained experts. This applies to the detection of optic disk damage [2], to the evaluation of its progression [3], and to the distinction between different patterns of damage [4].

## 3.1    Confocal Scanning Laser Tomography

Confocal scanning laser tomography (CSLT) is a new imaging technology that is rapidly establishing its role in clinical practice. Using a low-intensity laser beam to scan the back of the eye, the instrument rapidly acquires a series of images from 32 consecutive focal planes. After an image series has been acquired, the software of the CSLT device (in this case the Heidelberg Retina Tomograph, or HRT, Heidelberg Engineering, Dossenheim, Germany) pre-processes the images using standard alignment algorithms, compensating for shifts, tilt and rotation that may result from a patient's inadvertent eye movements during image acquisition, and stores the series in a proprietary format. Within each image series, the peak of the luminance distribution occurs where the focal plane corresponds closest to the imaged structure and can therefore be used to infer the relative height of that structure (topography image).

The sum of the luminances within each image series relates to the reflectance (i.e. brightness and color) of the structure. Owing to the use of monochromatic laser light, the reflectance image is similar to a filtered black-and-white photograph. After several series have been acquired and aligned, the software of the instrument computes mean reflectance and topography images.

As it is reproducible and objective, confocal laser scanning tomography bears great promise to improve diagnostic decisions in glaucoma. Statistical methods have already resulted in previously unattainable precision in detecting change over time [5].

## 3.2    Current CSLT Analysis Tools

The software of the commercial CSLT device supports several different diagnostic algorithms to distinguish between images of healthy and glaucomatous optic disks. With all of these methods, the user is required to manually outline the edges of the optic disk on the reflectance image. Identifying the contour of the optic disk is a skilled and highly subjective task, thus contour lines differ significantly and systematically even between expert users.

After the contour line has been drawn and stored, the software calculates the location of a 'reference plane' on which the subsequent computations of stereometric parameters (which aim to summarize the structural features of the optic disk) are based. Although the reference plane is essentially arbitrary, its location is dependent on the contour line and therefore has similar variability. However, current tools for analysis of CSLT data suffer from several shortcomings:

1. Although even healthy optic disks show tremendous biological variation in their size and shape, current diagnostic analyses extract mainly area- and volume parameters from the topographic image and largely disregard valuable information on the shape of the optic disk and the spatial pattern of the glaucomatous tissue loss;
2. Current CSLT analyses use the topographic (3-D) image only. It is likely that valuable information is

also contained in the reflectance image, which is disregarded;

3. Current area- and volume-parameters depend largely on the essentially arbitrary definition of a reference plane which may vary significantly from one examination to the next;

4. A trained operator has to manually outline the contours of the optic disk, a process that is subjective and introduces variability and scope for human error.

# 4   Data Set

Confocal Scanning Laser Tomography data have been collected at the Department of Ophthalmology since the end of 1991, as part of an ongoing longitudinal study on glaucoma. To date, 111 healthy controls and 115 patients with open-angle glaucoma have been participating in this study for up to ten years. The glaucoma patients were enrolled consecutively from clinics at the department, and the healthy control subjects were recruited from patient's relatives, hospital employees and from 2 private companies in Halifax so that the entire group constitutes a mixture of cases representative of a typical glaucoma practice in Nova Scotia. All of these subjects have had concurrent visual field examinations which contribute independent diagnostic information. Almost all glaucoma patients have also undergone optic disk photography, and these images have already been classified (according to subtype of damage and confidence rating) independently by two glaucoma specialists.

## 4.1   CSLT Images

At present our study database contains approximately 4300 optic disk images, i.e. aligned series of 32 images (256 by 256 pixels) acquired from consecutive focal planes, from which topography- and reflectance-images can be calculated using standard algorithms. The image series are saved in a proprietary format, which can be exported to the standard bitmap format using the software of the CSLT device.

For these experiments, we are relying on a relatively small amount of locally available CSLT image data acquired from the HRT device. In total, the data set consists of 127 unique patients that have been labeled *a priori* by specialists as control or glaucoma, outlined in table 1.

**Table 1**: Composition of Control and Glaucoma Dataset

| Label | Total | Age | Male | Female |
|----------|-------|-----|------|--------|
| Control | 46 | 61 | 39 | 7 |
| Glaucoma | 81 | 74 | 26 | 55 |

Each image plane is a 256 by 256 digital grayscale image of intensity values in the range of [0, 255]. In the work presented here, the topography and reflectance images are extracted from the proprietary HRT format and pre-processed using our own MATLAB image processing tools [6].

## 4.2   Partitioning

Patient image data was separated into three sets, referred to as training, validation and test partitions to be used in conjunction with the classification stage. For each patient, only the baseline examination is used, while repeat (follow-up) examination data is discarded.

For this study, ten different partitions were generated for each of training, validation and test sets. Each partitioning of the data is created by randomly assigning patients to one of training, test or validation such that 70% of patterns appear in training while 15% appear in each of test and validation. Control and glaucoma data are stratified in order to achieve proportional representation from each class within the three partitions. Note that the test and validation partitions represent disjoint sets ('unseen' patients) from the classifier's perspective (the training set). This property of the partitioning scheme affords analysis of the classifier's generalization ability across the patient population.

# 5   Image Analysis

Machine vision systems typically consist of five generic stages [7]: image capture; pre-processing; segmentation; extraction of relevant features; and classification. The first two stages (capture of the digital topographic optic disk image and pre-processing by compensation for shifts, tilt and rotation) are carried out by the HRT device. These present a significant advance over the traditional approach based on visual inspection. The latter three stages (segmentation, feature extraction, and classification) are of particular interest to this research.

## 5.1   Image Segmentation and Pre-processing

As indicated earlier, segmentation of the CSLT image currently relies on the subjective definition of a contour line and is therefore subject to random and systematic error. However, for the purposes of this study, a two-part pre-processing step is applied to each image in composing the feature set in order to control for degenerate border pixels (due to the HRT device's alignment process) and outlier points over the three image samples that are available per patient scan:

1. Each digital image is taken as the median image (implemented through the application of a 2D me-

dian filter, size 4 by 4) of the three image samples available.

2. The image border of width 25 pixels is replaced by the median image's average intensity value.

## 5.2 The Moment Method

Features are extracted from CSLT images using an image processing technique referred to as the Moment Method. Orthogonal moments have been used extensively in image processing and are well known for their superior performance and robust representation properties under noisy image conditions [8-10]. An attractive property of this analysis is that these moments can be made invariant to shifts, rotations and magnification changes, which can affect CSLT images when patients are unable to maintain steady eye position during imaging. Moreover, the optic disc is centered in the image, thus avoiding the requirement for an independent segmentation stage in which the object is explicitly identified.

The moments employed here are the orthogonal moments of Zernike [9], the set of which are defined by the projection of the image function, in terms of the image pixel intensities, onto the orthogonal bases of Zernike functions. Each Zernike moment describes image properties and is used as an element of the feature set representing the digital image [8-10].

### 5.2.1 General Zernike Algorithm

The Zernike moments are defined by their order ($n$) and repetition ($m$) with respect to a digital image, where the low order moments capture gross shape information and high order moments incrementally resolve high frequency information (representing detail) of the digital image.

For a given order, $n$, there are a total of $(n / 2 + 1)(n + 1)$ Zernike moments, each of which is invariant to rotation in the digital image and which may be further normalized to effect invariance to image translation and scale.

Given a set of Zernike moments (of order $n$) it is possible to approximate (or reconstruct) the original digital image, where the reconstructed approximation approaches the original image function as $n$ approaches infinity. In this way, Zernike moments can be easily calculated to arbitrary order sufficient to characterize the desired level of detail in the reflectance and topography images. As suggested by the orthogonality property, particularly descriptive Zernike moments of a given image can be easily isolated based on their contribution to the reconstruction process.

An example of the reconstruction process for a topography images is illustrated in figure 1. In this example the upper left image is the original; the upper right is

the result of an edge filter applied to extract the outline (or shape) of the original image; the lower left cell is the reconstructed image from Zernike moments (up to and including order 35 for figure 1); the lower right cell plots the original edge image overlaid on the reconstructed Zernike image for comparative purposes.
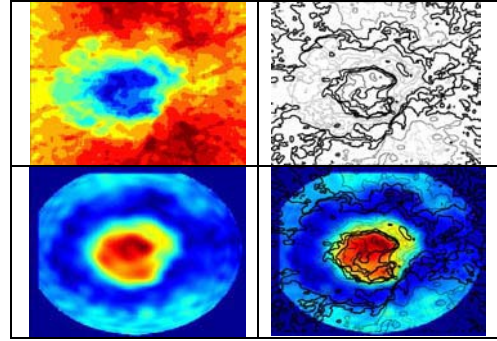


**Figure 1**: Example topography and reconstruction using Zernike Moment Method.

### 5.2.2 Moment Calculation

Invariance to translation, rotation and scale of overall image geometry may be achieved through the use of invariant Moment features [8, 10]. These features are termed Moment invariants and may be defined as a set of non-linear functions on the regular, geometric moments of an image as follows. Geometric (or regular) moments map the image function $f(x, y)$ onto the monomial $x^p y^q$. The $(p + q)^{th}$ order of geometric moment for an $N \times M$ image function, $f(x, y)$, is defined as:

$$m_{pq} = \sum_{x=1}^{M} \sum_{y=1}^{N} x^p y^q \cdot f(x, y) \quad (1)$$

$$\text{where } p, q \in Z^+$$

Furthermore, the $(p + q)^{th}$ central geometric moment is defined so as to normalize regular moment calculations with respect to the image centroid, thus yielding moments invariant to object translation:

$$\mu_{pq} = \sum_{x=1}^{M} \sum_{y=1}^{N} \left(x - \bar{x}\right)^p \cdot \left(y - \bar{y}\right)^q \cdot f(x, y) \quad (2)$$

where $\bar{x}, \bar{y}$ are coordinates of the image centroid obtained from:

$$\bar{x} = \frac{m_{10}}{m_{00}} \text{ and } \bar{y} = \frac{m_{01}}{m_{00}} \quad (3)$$

Similarly it is possible to scale-normalize the target image through the use of low-level moments by defining a scale factor $a = \sqrt{\beta / m_{00}}$ [10] (where $\beta$ is a predetermined scaling constant), and mapping the original image function, $f(x, y)$ into $g(x, y)$ as follows:

$$g(x, y) = f\left(\frac{x}{a} + \bar{x}, \frac{y}{a} + \bar{y}\right) \quad (4)$$

Unfortunately the basis set for geometric moments is not orthogonal and the resulting moments, therefore, lack many desirable properties in the context of feature selection. The complex-valued Zernike polynomials [9], however, form an orthogonal basis set over the unit circle, $x^2 + y^2 \leq 1$.

Orthogonal Zernike moments are defined by the projection of the image function $f(x, y)$ within the unit circle onto the complex Zernike polynomials. The Zernike invariants are the magnitudes (the hypotenuse) of the real and imaginary components of the resulting moments.

When composed in terms of scale and translation normalized images or in terms of the regular, low-level central geometric moments, the Zernike invariants have the additional property of being invariant to rotation as well as invariance to translation and scale, given an image function of sufficient resolution within the unit disc.

The set of Zernike polynomials is denoted by $\{Z_{nm}(x, y)\}$, or equivalently in their polar form by $\{Z_{nm}(\rho, \vartheta)\}$. The general form of these polynomials being:

$$Z_{nm}(x, y) = Z_{nm}(\rho, \vartheta) = R_{nm}(\rho)\exp(jm\vartheta)$$

where $x, y; \rho, \vartheta$ correspond to Cartesian and polar coordinates respectively $n \in Z^+$, $m \in Z$, constrained to $n - |m|$ even, $|m| \leq n$

$R_{nm}(\rho, \vartheta)$ is the radial polynomial:

$$R_{nm}(\rho, \vartheta) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s [(n-s)!]\rho^{n-2s}}{s!\left(\frac{n+|m|}{2} - s\right)!\left(\frac{n-|m|}{2} - s\right)!} \quad (5)$$

The complex, orthogonal Zernike moments are then defined by:

$$A_{nm} = \frac{n+1}{\pi} \sum_{x=1}^{M} \sum_{y=1}^{N} f(x, y) \cdot Z_{nm}^*(\rho, \vartheta) \quad (6)$$

## 5.5 Features and Subset Selection

Zernike moments were extracted to order 30 to provide us with 254 features for each reflectance and topography image, for a total of 508 Zernike features per image. The unit circle normalization, required for the extraction process, corresponds to the largest circle fitting completely within the image. The resulting feature sets were converted to UCI format (a widely used machine learning format) for future use.

Normalization of features, relative to the maximums and minimums across all feature values in the training set, is performed in each instance prior to training, validation and test runs.

This work compares results of classification using eight feature sets in total, which are summarized in Table 2. Six sets are based on Zernike moments extracted from Topography and / or Reflectance images accordingly, while the remaining two sets are based on the conventional stereometric features extracted by HRT support software using contours defined by a trained technician as described in Section 3.2.

Feature sets are described as 'all' when every feature available is used in the training and classification process, while the 'selected' sets indicate preprocessing for feature subset selection based on non-parametric Wilcoxon rank sum hypothesis testing, as described below.

**Table 2**: Exemplar Feature Dimensions

| ID | Feature Set | Size |
|----|-------------|------|
| SC | Selected Conventional Indices | [50-78] |
| AC | All Conventional Indices | 100 |
| ST | Selected Zernike Topography | [55-115] |
| AT | All Zernike Topography | 254 |
| SR | Selected Zernike Reflectance | [4-34] |
| AR | All Zernike Reflectance | 254 |
| T+R | AT Combined with AR | 508 |
| ST+R | ST Combined with SR | [59-149] |

In selecting a feature subset, we consider the difference between the glaucoma and control class sample distributions of the training set: the more that a given feature differs between these two classes, the simpler the decision rule necessary to distinguish the two classes. Here an appropriate indicator is the variance normalized inter-class separability (included on median feature plots).

To isolate the features that best discriminate between classes, we consider a summary of all features for each class for both topography and reflectance feature types; to this end, we have plotted the median of each glaucoma and control feature in the training sets with their respective first and third quartile values as error bars (figure 2). For each feature, we plot the result of a two-tailed, two sample Student's T-test (H0: $\bar{x}_{control} = \bar{x}_{glaucoma}$; Ha: $\bar{x}_{control} \neq \bar{x}_{glaucoma}$) along with the non-parametric Wilcoxon rank sum test (H0: Equal population medians; Ha: Different population medians). In each instance, lower values indicate better strength to reject the null hypothesis.

An assumption implicit in the use of the Student's T-test is that both samples come from normal population distributions. Although we have no clear evidence to support the assumption in this case, the t-test result is only meant to serve as a guide in locating promising

candidate features. The median and quartile bars along with the variance normalized inter-class separability were ultimately used to validate the separability of any features that are selected with this statistic, although the results were invariably consistent with the Wilcoxon tests. Here we simply defined a threshold of 0.01, below which a feature is selected.
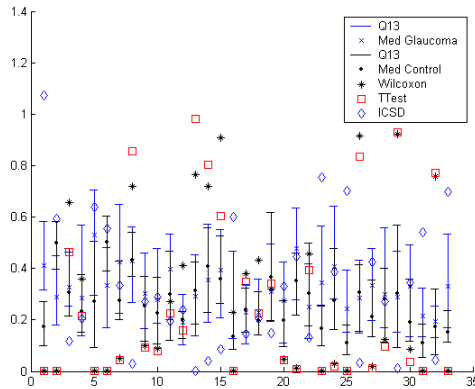


**Figure 2**: Quartiles, Inter-class separation distance, T-test and Rank Sum analysis for Zernike Moments 1 to 32

## 5.6  Learning and Classification

### 5.6.1  C5 Decision Tree

C5 is a set of programs, based on the ID3 algorithm for decision trees, that inductively construct classification models based on known classification instances [11]. C5 is the latest available incarnation of Ross Quinlan's ID3 algorithm and is widely accepted as a standard in decision tree classifiers.

Due to the relative sensitivity of the C5 algorithm to the partitioning of the data into training and test sets, we use identical partitions of data patterns as for the Perceptron (rather than employing the f-fold cross validation function in C5) in order to obtain a more reliable and comparable estimate of predictive accuracy under C5.

In the interest of clarity these experiments use only the basic C5 algorithm. As such, no boosting, pruning, or misclassification weights are used to further augment trial runs.

### 5.6.2  Perceptron

The perceptron is a widely studied supervised learning algorithm capable of forming linear discriminant functions [12]. Cost function takes the form of the Sum Square Error (SSE), and updates to the free parameters are made exemplar wise. The basic algorithm takes the following form:

1. Initialize a network of $P$ inputs, $x_i$, $O$ outputs, $y_j$, and corresponding matrix of initial free parameters $w_{ij}$ $i \in \{0,...,P\}$, $j \in \{1,...,O\}$ defined by a uniform p.d.f. over the unit interval and $x_0 = 1$.
2. Present input exemplar vector, $x(t)$, and desired output classification vector, $d(t)$;
3. Calculate network output(s), $y_j(t)$,

$$y_j(t) = \sum_{i=0}^{P} w_{ij}(t)x_i(t) \quad (7)$$

4. Update all free parameters,

$$w_{ij}(t+1) = w_{ij}(t) + \eta[d_j(t) - y_j(t)]$$

5. IF ($t$ mod (# training exemplars) == 0) THEN (estimate Sum Square Error over validation data)
6. IF (SSE(validation) increases) OR (t == 500)) THEN (END) ELSE (return to point (2))

Thus, early stopping is identified by an independent validation set and a test set (also independent) used for assessing generalization of the model.

Following training, a simple pruning rule was implemented to identify features of most discriminant value to the classifier. Pruning proceeds while there is some positive improvement in overall performance on the validation set as follows:

1. Iteratively remove each feature input, one at a time;
2. Eliminate the input that results in the largest overall improvement (across the validation set);
3. IF (improvement increases) THEN repeat ELSE stop.

# 6  Results and Discussion

The following scatter plots of sensitivity versus specificity present the mean test results for each of the eight feature sets over the 10 partitions, with Perceptron runs prior to and following pruning shown in figures 3 and 4, respectively, while C5 results are given in figure 5. Standard errors (of sensitivity and specificity) for each result are reported in figures 6 to 8.

## 6.1  Perceptron Classification Results

Both perceptron models have a distinct preference for the moment feature 'AT' with respect to sensitivity and specificity (in terms of both mean and variance). This establishes the best performance irrespective of feature set or learning algorithm over all results. For the perceptron models, AR and SR (both reflectance only sets) always return the worst mean performance.
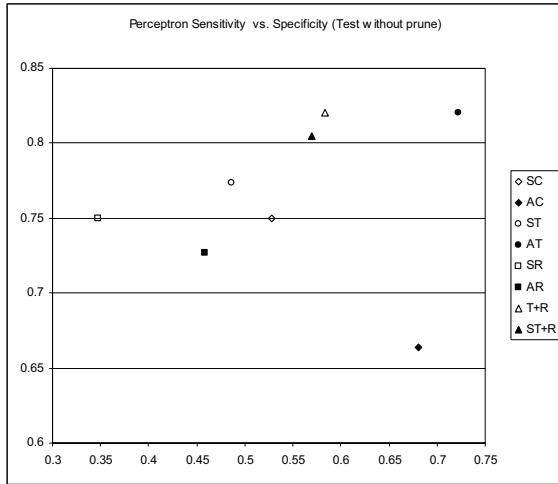
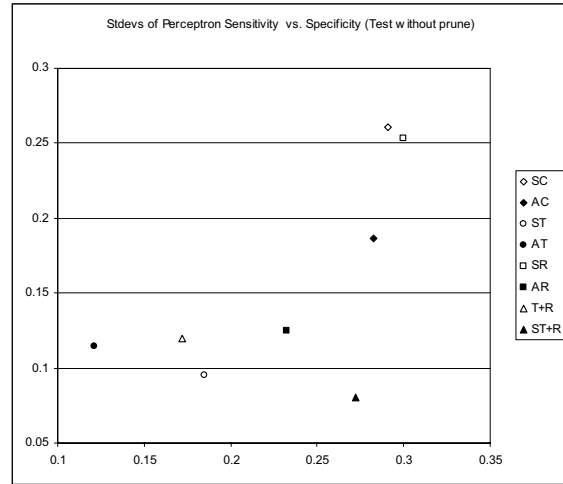**Figure 3**: Perceptron Without Feature Pruning – Mean.
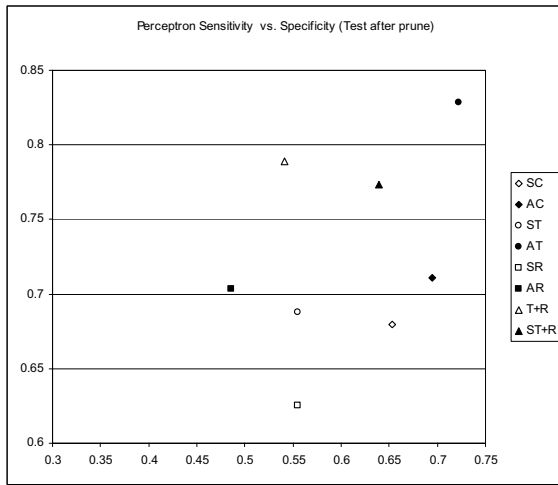


**Figure 6**: Perceptron Without Feature Pruning – SD.
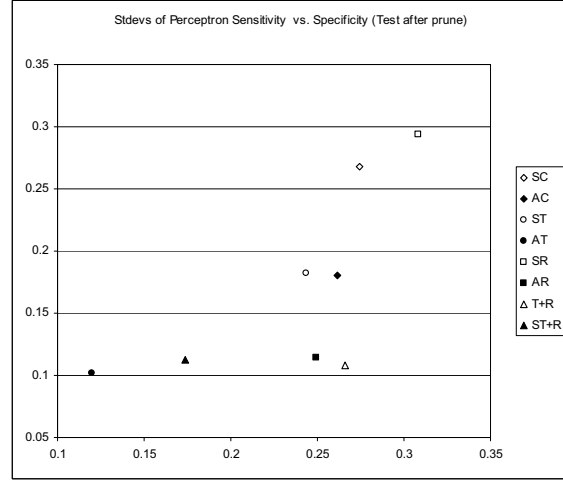


**Figure 4**: Perceptron With Feature Pruning – Mean.



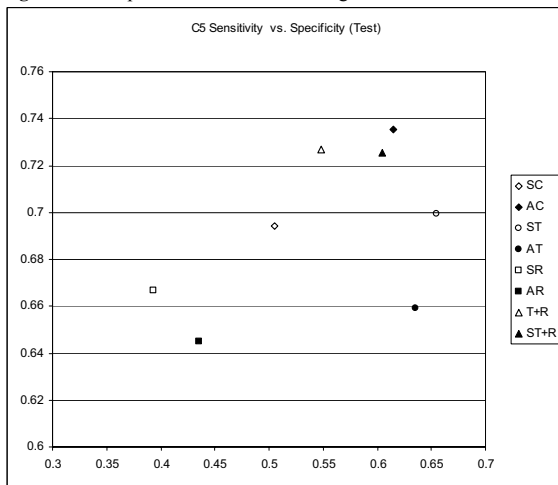**Figure 7**: Perceptron With Feature Pruning – SD.



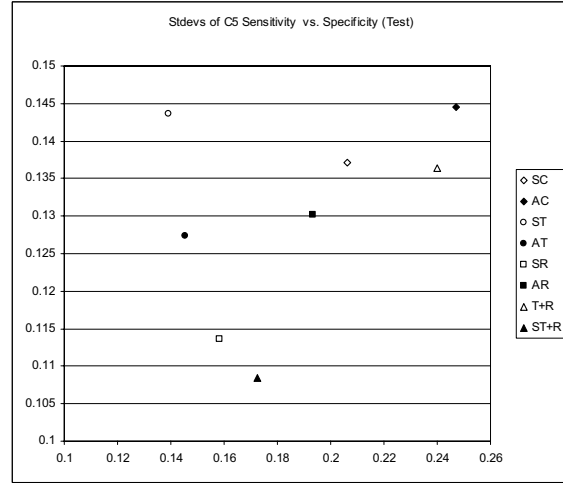**Figure 5**: C5 Classifier–Mean.



**Figure 8**: C5 Classifier – Standard Deviation.

SC and SR always demonstrate most variation in classification performance; while AT is the most consistent – returning least variance in sensitivity and specificity and matching the low level of classification performance variation provided by C5. Features T+R / ST+R / AR all provide low sensitivity variance at the expense of specificity variance.

Overall, the perceptron results (other than those for AT) show a greater degree of variation across all features than for those of C5.

### 6.2   C5 Classification Results

C5 does not indicate a dominant classification performance for any particular feature set. The best mean sensitivity value was AC; while the highest sensitivity pair is T+R/ST+R. The highest specificity value was the ST/AT pair and worst sensitivity and specificity are the SR/AR pair, where this is the only result in common with the perceptron. Overall C5 is able to better contain the variance in mean classification performance than the perceptron.

## 7   Conclusions

The direct classification of optic nerve images has been established using the Method of Moments. The Method of Moments is particularly appropriate as features are constructed 'globally'. This improves the robustness properties of the approach, whilst being facilitated by the "well framed" nature of the image content (not necessary to explicitly separate object from background). Moreover, the Moment Methods employed are orthogonal – enabling the incremental addition of increasingly more specific features – whilst also supporting translation, rotation and shift invariances [8, 9, 10].

The overall scheme removes the remaining subjective process of optic disk interpretation. Although preliminary in nature, the results do establish the potential for using moment methods directly on topography and reflectance images of patient records. More data will be necessary to provide the basis for robustly training more complex classifiers or additional testing of the performance on unseen data.

Further improvements are expected following further analysis of features utilized by the classifiers. Moreover, pseudo-Zernike features may provide further improvements to the effectiveness of the system [8, 10]. Other opportunities also available include the analysis for topological orderings in the glaucoma cases through the use of appropriate unsupervised learning algorithms such as Self Organizing Feature Maps.

## Acknowledgements

## References

[1] Nicolela M.T., Drance S.M., "Various Glaucomatous nerve appearances: Clinical Correlations," Ophthalmology. 103(4) pp 640-649, 1996.

[2] Varma R., Steinmann W.C., Scott I.U., "Expert agreement in evaluating the optic disc for glaucoma," Ophthalmology. 99(2) pp 215-221, 1992.

[3] Tielsch J.M., et al., "Intraobserver and interobserver agreement in measurement of optic disc characteristics," Ophthalmology. 95(3) pp 350-356, 1988.

[4] Nicolela M.T., et al., "Visual field and optic disc progression in patients with different types of optic disc damage in glaucoma," Ophthalmology. In press.

[5] Chauhan B.C., et al., "Test-retest variability of topographic measurements with confocal scanning laser tomography in patients with glaucoma and control subjects," Am. J. Ophthalmology. 118, pp 9-15, 1994.

[6] Matlab Image Processing Toolbox, Version 3, The Mathworks. http://www.mathworks.com, 2004.

[7] Castleman, K. R., Digital Image Processing, Prentice Hall, Upper Saddle River, New Jersey 1996.

[8] C. H. Teh and R. T. Chin, "On image analysis by the methods of moments," IEEE Trans. Pattern Anal. Machine Intell., 10(4), pp. 96-513, July 1998.

[9] M. Teague, 1980, "Image analysis via the general theory of moments," J. Opt. Soc. Amer., 70(8), pp. 920-930.

[10] Heywood, M. I., Noakes P., "Fractional central moment method for movement-invariant object classification," IEE Proceedings on Vision, Image and Signal Processing, 142(4), pp 213-219, 1995.

[11] J. Ross Quinlan, C4.5: Programs for machine learning, Morgan Kaufmann Publishers, Inc., San Mateo, CA, 1993.

[12] Widrow B., Lehr M.A., "Adaptive Neural Networks and Their Applications," International Journal of Intelligent Systems. 8 pp 453-507, 1993.