A Computer-Aided Diagnosis System of Nuclear Cataract

Huiqi Li*, *Member, IEEE*, Joo Hwee Lim, Jiang Liu, Paul Mitchell, Ava Grace Tan, Jie Jin Wang, and Tien Yin Wong

***Abstract*—Cataracts are the leading cause of blindness worldwide, and nuclear cataract is the most common form of cataract. An algorithm for automatic diagnosis of nuclear cataract is investigated in this paper. Nuclear cataract is graded according to the severity of opacity using slit lamp lens images. Anatomical structure in the lens image is detected using a modified active shape model. On the basis of the anatomical landmark, local features are extracted according to clinical grading protocol. Support vector machine regression is employed for grade prediction. This is the first time that the nucleus region can be detected automatically in slit lamp images. The system is validated using clinical images and clinical ground truth on >5000 images. The success rate of structure detection is 95% and the average grading difference is 0.36 on a 5.0 scale. The automatic diagnosis system can improve the grading objectivity and potentially be used in clinics and population studies to save the workload of ophthalmologists.**

***Index Terms*—Automatic grading, computer-aided diagnosis, nuclear cataract, slit lamp image.**

I. Introduction

A CATARACT is a loss of the normal transparency of human lens due to opacity. The lens is mostly made of water and protein. A cataract is formed when the protein clumps together, which clouds the lens and reduces the light that reaches the retina [1]. Cataracts are the leading cause of blindness worldwide, accounting for half of global blindness [2]. Cataract is largely related to the aging process. For example, it was estimated that 17.2% Americans older than 40 years have cataracts and more than half of Americans over 75 years old have either a cataract or had cataract surgery [3]. Population studies in other countries showed similar trends [4], [5]. Because of the population growth and increasing longevity, blindness caused by age-related cataract is expected to increase. Risk factors for cataract have been identified, but there are no current proven preventative strategies to prevent cataract progression [6]. Thus, accurate diagnosis and timely surgical treatment of cataract is essential to prevent vision loss and improve quality of life.

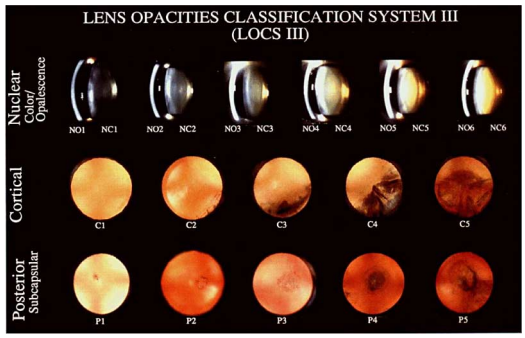

Fig. 1. Standard photographs in LOCSIII system [8].

There are mainly three types of age-related (senile) cataract defined by their clinical appearance: nuclear, cortical, and posterior subcapsular [7]. Nuclear cataract starts at the center of the lens and progresses toward the surface. Cortical cataract begins in the lens's outer rim and grows toward the center. Posterior subcapsular cataract usually forms from the back of the lens. Nuclear cataract is diagnosed using slit lamp images, while cortical cataract and posterior subcapsular cataract are examined using retroillumination images. Nuclear cataract is the most common type and the focus of this paper.

Clinically, nuclear cataract is diagnosed via comparing the picture observed during slit lamp biomicroscopy with a set of standard photographs with increasing opacity of lens nucleus (see Fig. 1). In the lens opacities classification system III (LOCS III) [8] shown in Fig. 1, NO1–NO6 and NC1–NC6 are the standards for nuclear opacity and nuclear color, respectively. C1–C5 are the standards for cortical cataract and P1–P5 are the standards for posterior subcapsular cataract. These standard images exemplify the boundaries of scaling intervals. The average opacity of the entire nucleus in the lens is compared with that in each of the standards visually. A decimal grade is assigned to reflect the position of the unknown within the appropriate standard interval, for example, a grade of 2.5 means that the severity of the opacity was judged to be midway between standards 2 and 3. For

Manuscript received September 2, 2009; revised November 9, 2009; accepted December 28, 2009. Date of publication February 17, 2010; date of current version June 16, 2010. This work was supported in part by Singapore Agency for Science, Technology and Research Singapore Bioimaging Consortium Grant of "Singapore Retinal Image Archival and Analysis Network for Disease Prediction" and by the National Medical Research Council under Grant 0796/2003. *Asterisk indicates corresponding author*.

*H. Li is with the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore 138632, Singapore (e-mail: huiqili@i2r.a-star.edu.sg; li_huiqi@yahoo.com).

J. H. Lim and J. Liu are with the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore 138632, Singapore.

P. Mitchell and A. G. Tan are with the Centre for Vision Research, Department of Ophthalmology, University of Sydney, Sydney, N.S.W. 2001, Australia.

J. J. Wang is with the Centre for Vision Research, University of Sydney, Sydney, N.S.W. 2001, Australia, and also with the Centre for Eye Research Australia, University of Melbourne, Vic. 3002, Australia.

T. Y. Wong is with Singapore Eye Research Institute, National University of Singapore, Singapore 119077, Singapore, and also with the Centre for Eye Research Australia, University of Melbourne, Vic. 3002, Australia.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2010.2041454

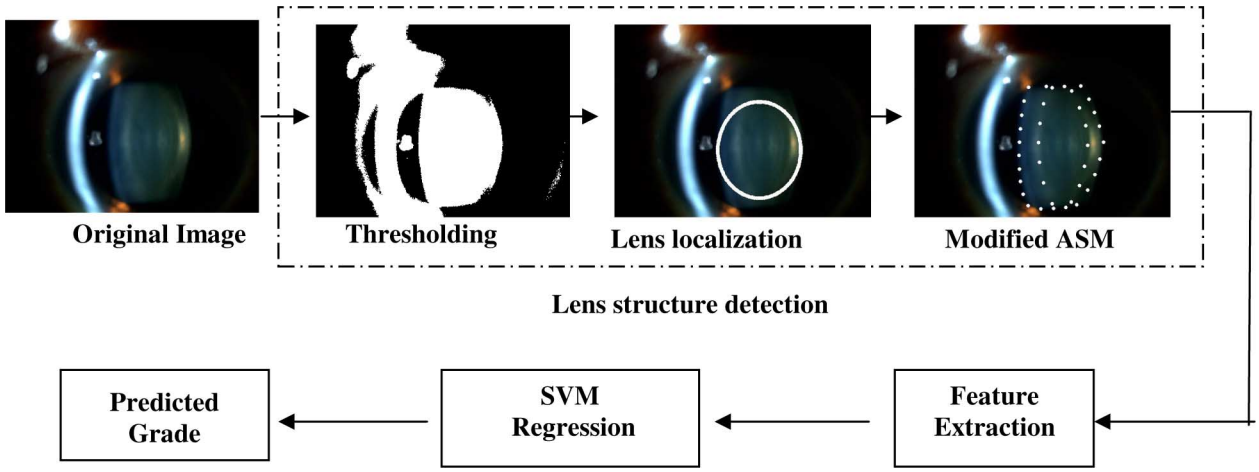

Fig. 2. Schematic diagram of the AGNC system.

each cataract type, higher scores indicate greater severity. There are multiple grading systems available, which are based on similar principle. Other examples are the Oxford clinical cataract classification and grading system (OCCCGS) [9]. In order to diagnose the lens opacity more objectively, cataracts are graded based on photographs or digital images in some grading systems [10] rather than a direct clinical assessment. Regardless, these manual grading systems are subjective and vary substantially between observers and within observers over time. For nuclear cataract grading, the interobserver agreement is around 65% and the intragrader agreement is between 70% and 80% in one study [10]. It is also time-consuming and costly to assess and grade cataracts in large number of subjects for population studies.

With potential advantages of improved objectivity and reproducibility, computer-aided diagnosis of cataract can facilitate clinical diagnosis and study. There have been some previous efforts to investigate automatic diagnosis of nuclear cataract. A group from the University of Wisconsin [11], [12] extracted anatomical structure on the visual axis. In one system [11], the spatial relationship between cornea and anterior cortex was used to detect the visual axis. The location of anatomical features such as cornea, sulcus, and lentil was further identified on the visual axis. The features to assess nuclear cataract and the algorithm for automatic analysis were not investigated yet in this study. In another system [12], two features on the visual axis (sulcus intensity and intensity ratio between anterior and posterior lentil) were selected and linear regression was performed for automatic grading. Another group from Johns Hopkins [13] also identified anatomical structure on visual axis. A longitudinal intensity profile was calculated within a circle in nucleus region and three features were extracted: nuclear mean gray level, slope at the posterior point of profile, and the fractional residual of the polynomial least-square fit. Neural network was trained to determine the grade of nuclear opacity. Because of low SNR, none of these studies could reliably detect the entire region of lens nucleus and extract features in it, even though the entire nucleus region is analyzed in the clinical diagnosis [8], [10]. Importantly, none of these systems was validated with a large amount of clinical data.

In this paper, combining both clinical and image analysis expertise, we developed a new system for automatic grading of nuclear cataract (AGNC) in this paper. Compared with the existing work, it is the first time that the nucleus region can be automatically detected in slit lamp images. More clinically, meaningful features are extracted and the proposed automatic diagnosis system has been validated using >5000 slit lamp images with their clinical grading results.

II. Approach

The computer-aided diagnosis system of AGNC has three components: structure detection, feature extraction, and grade prediction. Robust feature extraction, which depends on the accurate structure detection, is always the key for a successful computer-aided diagnosis. Bottom-up and top-down strategies are combined to detect lens structure in this paper. Features are extracted based on the lens structure, and support vector machine (SVM) regression [14] is applied to predict the grades of nuclear cataract. The schematic diagram of AGNC system is shown in Fig. 2. Model-based approach is adopted in each step to apply domain knowledge. Some of our initial work has been published in [15] and [16]. The improvement of the AGNC system over these previous works, as described in this paper, includes the following aspects. 1) A new 38 point shape model is proposed to detect the nucleus region in the lens. Twenty-four points were used in the previous model to describe the contour of the lens. The new model comprises a segment around a boundary of nucleus region, which is usually the only region in which nuclear cataract is normally assessed. 2) More meaningful and accurate features are extracted for the grading. One improvement is that the intensity, color, and texture within the nucleus region can be measured in AGNC system. In the previous works, the intensity is averaged in the entire lens, which is not consistent with clinical grading. Another improvement is the intensity measurement of sulcus (see Fig. 3). Whether the sulcus of the nucleus is well defined throughout the nucleus region is one of the most important factors in grading the severity of nuclear cataract. In previous works, the intensity of the sulcus is only measured on the visual axis. With the proposed lens structure detection

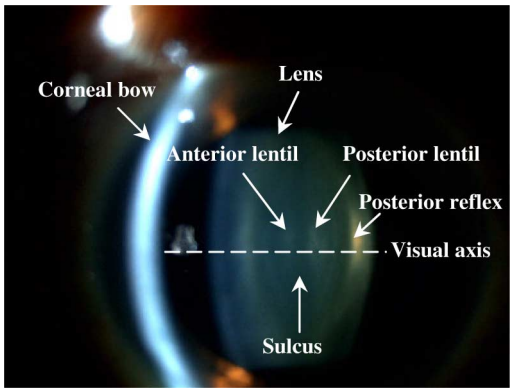

Fig. 3. Slit lamp lens image of right eye.

approach, the intensity of sulcus can be measured in the whole length of the nucleus region to reflect the clinical grading better. 3) The AGNC system is validated using >5000 images acquired in the clinic. The size of test data is less than 1000 in the previous works.

A. Anatomical Lens Structure Detection

There are basically two strategies for image understanding. One is bottom-up processing, which starts from the raw data, works upward to object representation and higher level analysis or decision. Another is top-down processing, which makes hypotheses about what is happening, and then uses the image data to validate or reject those hypotheses. For medical images, it is not easy to achieve robust structure detection using bottom-up processing only due to noise, uneven illumination, and variations across individual images. In the lens structure detection, bottom-up strategy and top-down strategy are combined in order to obtain robust detection. The first step is to locate lens position via bottom-up strategy. This localization is utilized to initialize the active shape model (ASM) [17], which belongs to top-down strategy, for lens structure detection because ASM is a local searching method. All the structure detection algorithms perform on the gray slit lamp image.

Observing a lens image, it can be seen that the corneal bow is usually the leftmost (for right eye) or rightmost (for left eye) brightest vertical curve in a well-taken slit lamp image (see Fig. 3). The lens is the largest part in the foreground, which occupies 20%–30% of the whole image, and it appears in the center of an image. The corneal bow is identified first, and then the lens is located by horizontal and vertical profile clustering. The lens localization scheme is described in Fig. 4. The 10% brightest pixels in the gray image are selected by thresholding to locate the corneal bow as shown in Fig. 4(b). The foreground is obtained by thresholding 20%–30% of the brightest pixels in the gray image as in Fig. 4(c). An initial thresholding value is first set. The value of thresholding is increased or decreased iteratively until the area of the detected foreground is between 20% and 30% of the total area of the image. Five horizontal profile clustering and five vertical profile clustering are performed to locate the center of the lens. The profiles are illustrated in Fig. 4(c). The horizontal profile through the median line of the image [see line A in Fig. 4(c)] and four parallel profiles are analyzed. On each profile, clustering is performed on the right-hand side of the corneal bow for right eye (left-hand side for left eye). The clustering is based on the connectedness of the white pixels on each profile and adjacent white pixels are assigned to the same cluster. The centroid of the largest cluster is calculated. The number of pixels in the largest cluster is termed cluster size. Excluding the largest cluster size and the smallest cluster size among the five profiles, the mean of the centroid and cluster size on the other three profiles are calculated as the horizontal coordinate of the lens center and horizontal diameter, respectively. The vertical profile through the estimated horizontal coordinate [see line B in Fig. 4(c)] and four parallel profiles are analyzed similarly as the horizontal profiles. The lens can be estimated as an ellipse based on the lens center and horizontal and vertical diameters. The coordinate of the estimated lens center is denoted as (L_x, L_y) and its semiaxes radii, which are halves of the diameters, are denoted as (R_x, R_y).

The contour of lens and its nucleus are detected using a modified ASM method [17]. Model-based approach constructs a model from training data and uses it to fit the test data to find the modeled object. ASM introduced by Cootes and Taylor [18] is a parametric deformable model that combines point distribution model and iterative refinement procedure. It can describe the shape of a nonrigid object, and is efficient in many applications, especially in the extraction of anatomical structures, which are complex and variable among individuals and where a high degree of *priori* knowledge is available.

A point distribution model (PDM) has to be estimated before ASM can be applied. It is a method for building models by learning patterns of variability from a training set of correctly annotated images. The model allows deformation in certain ways that are consistent with the training set. The shape of the lens structure is described by the position of 38 landmark points, as shown in Fig. 5. Besides the lens contour described in our previous model [15], [16], the contour of the lens nucleus is included in this new model. The opacity in the nucleus region is the main area for the assessment of nuclear cataract. Therefore, the new model is more meaningful for feature extraction. Ten images with well-defined lens structures were selected as the training set, and the landmark points of the shape model were annotated manually. Among the ten training images, five images are from left eyes and five images are from right eyes. The shapes from the right eyes are mirrored horizontally to get their symmetrical shape so that they can be calculated together with the shapes from the left eyes and *vice versa*. The exemplar shapes are alignment using affine transformation by a routine least square approach. Principal component analysis (PCA) is applied to the aligned exemplar shape to obtain the shape model. A lens shape, i.e., PDM, can be described as

$$x = \bar{x} + \Phi b \tag{1}$$

where $\bar{x}$ is the mean shape of the aligned training set, $b = (b_1, b_2, \ldots, b_t)^T$ is a vector of shape parameters, and $\Phi = (\Phi_1, \Phi_2, \ldots, \Phi_t) \in R^{2n \times t}$ is the set of the eigenvectors corresponding to the largest t eigenvalues of the covariance matrix of the aligned training set. In our application, ten lens shapes are

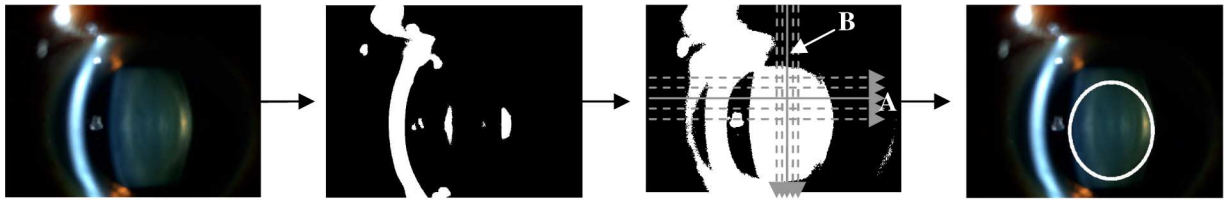

Fig. 4. Lens localization scheme. (a) Original Image. (b) Corneal detection. (c) Horizontal and vertical profile clustering. (d) Lens estimation.

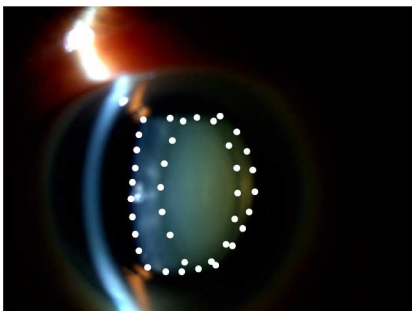

Fig. 5. Shape model fitted on a slit lamp image.

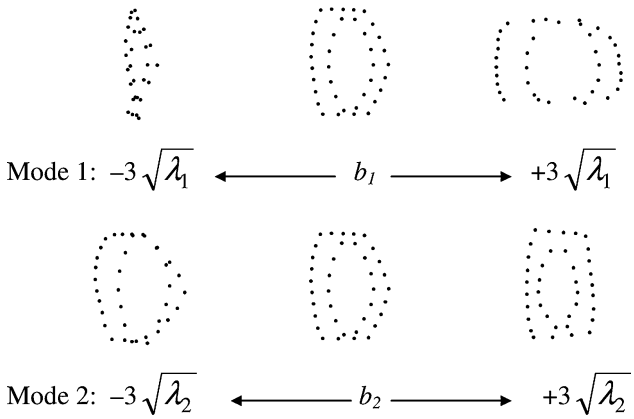

Fig. 6. First two modes of variation from –3 to +3 standard deviation for right eye. The middle column shows the mean shape, where λ_i is the *i*th largest eigenvalue.

used as the training set, $n = 38$, and t is set to 4. These first four eigenvectors represent 95% of the total variance of the shapes in the training set. The model obtained for the right eye and the effect of varying each of the first two shape parameters in turn between ± 3 standard deviations are shown in Fig. 6.

ASM is an iterative refinement procedure to fit the point distribution model in a new image to find the modeled objects. It consists of five main steps: initialization, matching point detection, transformation, model update, and convergence evaluation. The shape model in the shape space is denoted by the lowercase and shape instance in the image space is represented by uppercase in this paper. The transform between the shape space and image space can be described as

$$X = T(x) = \begin{pmatrix} s\cos\theta & -s\sin\theta \\ s\sin\theta & s\cos\theta \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (2)$$

where the shape models in the shape space and the image space are denoted as x and X respectively, and x_i, y_i denotes the position of the *i*th landmark point of the shape model in shape space. The transform T comprises translation (t_x, t_y), rotation θ, and scaling s.

1) Initialization: The mean shape $\bar{x}$ in (1) is employed as the initial shape model in shape space, and it is transformed to image space according to (2). The initialization is fully automatic, which is based on the lens localization, where (t_x, t_y) is set to the center of lens (L_x, L_y) and $\theta = 0$. The scaling factor s is determined using the semiaxes radii (R_x, R_y) of the ellipse estimation.

2) Matching Point Detection: The matching points are searched along the profiles normal to the model boundary through each model point. The first derivative of the intensity distribution on the normal profile is utilized to find the edge of lens or nucleus. A matching point cannot be detected successfully if the maximum or the minimum of the first derivative cannot meet a set threshold. It will be estimated from the neighboring matching points using linear interpolation.

3) Transformation: The transformation defined in (2) is used to transform between shape space and image space. In order to reduce the effect of the misplaced matching points, a weighted least square approach is employed to obtain the optimal transformation parameters. In particular, parameter $\tau(s, \theta, t_x, t_y)$ is selected so as to minimize a weighted sum of point difference between shape model and its matching points. In each iteration, the transformation for alignment is performed twice: once for the initialized weight factor and once for the adjusted weight factor. The weight factor is initialized according to how its corresponding matching point is obtained. Larger value will be given to those having their matching points detected directly along the profile. Smaller value will be assigned to the points having their matching points estimated from the nearby matching points. The weight factor is adjusted to be piece-wise reciprocal ratio of the Euclidean distance between the matching point and its updated landmark points of shape model in the image space, which is a negative feedback. By this modification, the misplaced matching point will not affect the transformation as much as the correct matching points, so that a better pose parameter τ is obtained.

4) Model Update: The detected matching points are transferred back to shape space, which is denoted as y. The shape parameter vector b in (1) is updated by projecting the matching points y onto the shape model space. When the Euclidean distance between the matching points and the updated shape model is larger than a constant value, the matching point is

considered to be a misplaced matching point. The misplaced matching points are proposed to be excluded in the projection. The shape parameter vector *b* is therefore determined only by the matching points that match the boundary well. The constriction is applied to b_i as

$$|b_i| \leq 3\sqrt{\lambda_i} \tag{3}$$

where λ_i is the *i*th largest eigenvalue, *i* =1–4.

5) Convergence Evaluation: The convergence of the shape model is evaluated according to the following formula to stop the iteration:

$$E_X = ||X^n - X^{n-1}|| < \varepsilon_T \tag{4}$$

where X^n and X^{n-1} denote the derived shape model of the *n*th iteration and the (*n*–1)th iteration in the image space, respectively, and ε_T is a small constant value, which is set to 10 in this paper, based on empirical studies.

Compared with the original ASM method [18], two improvements have been proposed in this modified ASM method [17] to fit the shape model in a new image more robustly. The improvement includes adding the self-adjusting weight in the transform between shape space and image space, and excluding the outlier in the projection of shape parameter. The main purpose of these improvements is to avoid the influence of the misplaced matching points in the detection. In slit lamp lens images, edges, especially edges of the nucleus region, are quite weak in some images. Some matching points are misplaced in such cases, in which the modified ASM method [17] can help to obtain more robust detection of lens structure than original ASM method [18].

The anatomical structure of the lens can be obtained using the modified ASM method, and the regions inside the lens and inside the nucleus can be obtained using the landmark points. The accurate detection of the anatomical structure is essential to extract robust feature for the grading.

B. Feature Extraction

On the basis of the detection of anatomical lens structure, features are extracted for diagnosis. The features were carefully selected according to clinical grading protocol of nuclear cataract [8], [10] as well as clinicians' opinion. The features to assess nuclear opacity in AGNC system are presented in Table I. All the features are normalized to be in the range [0, 1]. Normalization to zero mean and unit standard deviation was also tested, and there is not much difference of grade prediction accuracy for our dataset.

For all the features related to color, the hue, saturation, and value (HSV) color space is chosen to represent the color information. The lens contour and the contour of lens nucleus can be obtained based on the landmark points (see Fig. 5) of the lens structure. The inner contour is the contour of nucleus region and the outer contour is the lens contour. For features 1–6, the measurement is averaged within the lens contour. Similarly, the measurement is averaged within the region of lens nucleus for features 7–12. For local entropy (features 5 and 11), the probability distribution p_i of intensity i in the 9 × 9 neighborhood

TABLE I
FEATURE TABLE FOR NUCLEAR CATARACT DIAGNOSIS

Feature	Description
1	Mean intensity inside lens contour
2-4	Mean color inside lens contour
5	Mean entropy inside lens contour
6	Mean neighborhood standard deviation inside lens contour
7	Mean intensity inside nucleus contour
8-10	Mean color inside nucleus contour
11	Mean entropy inside nucleus contour
12	Mean neighborhood standard deviation inside nucleus contour
13	Intensity ratio between nucleus and lens
14	Intensity of sulcus
15	Intensity ratio between sulcus and nucleus
16	Intensity ratio between anterior lentil and posterior lentil
17-18	Strength of nucleus edge
19-21	Color on posterior reflex

of the corresponding pixel is calculated according to intensity histogram. The local entropy is further calculated as

$$\text{Local Entropy} = -\sum_{i=1}^{256} p_i \log p_i. \tag{5}$$

The neighborhood standard deviation in features 6 and 12 is calculated within the 3 × 3 neighborhood around the corresponding pixel.

The intensity distribution on a horizontal median line of the nucleus region is obtained to analyze visual axis profile. Low-pass Chebyshev filter [19] is applied to smooth the profile. An example of smoothed visual axis profile is demonstrated in Fig. 6. The first derivative of the profile is analyzed and the positions of the edges of nucleus region are identified using the maximum and minimum of the first derivatives. The edges of nucleus are detected as point A and point B, as illustrated in Fig. 7. In a clear slit lamp image (see Fig. 3), a dark band can be seen running vertically through the center of lens, which is termed sulcus. Bordering the sulcus are two bright bean-shaped bands, which are termed anterior lentil and posterior lentil, respectively. The horizontal position of the sulcus (point C in Fig. 7) is defined as the median point of nucleus edges. The positions of posterior lentil (PL in Fig. 7) and anterior lentil (AL in Fig. 7) are further defined as the central half of the nucleus region. The intensity ratio between anterior lentil and posterior lentil on the visual axis profile is selected as feature 16. The gradients of both nucleus edges are calculated based on the visual axis profile and defined as features 17 and 18.

The intensity of sulcus is a very important feature to decide the grade of nuclear cataract clinically. The horizontal position of sulcus is obtained using the aforementioned visual axis profile analysis. A narrow vertical band (20 pixels wide) within nucleus range is defined using the horizontal sulcus position. The intensity of pixels in the narrow band is averaged to obtain the intensity of sulcus as feature 14. The intensity ratio between nucleus and lens (feature 13), and the intensity ratio between sulcus and nucleus (feature 15) are calculated as feature 7/feature 1 and feature 14/feature 7, respectively.

The posterior subcapsular reflex is the best location to judge the quality of opacity color [8]. The position of central posterior subcapsular reflex (see Fig. 3) can be obtained based

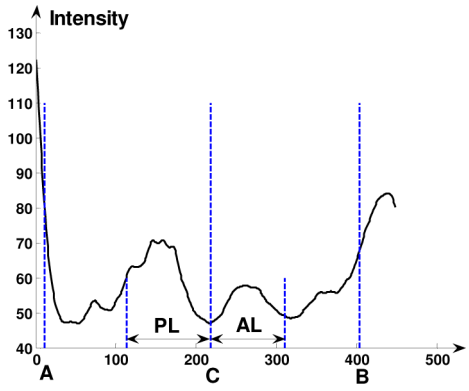

Fig. 7. Visual axis profile analysis.

on the landmarks of the ASM. The mean values of HSV three channels in a 60 × 120 pixels region of central posterior reflex are calculated as the features 19–21. The region is vertically centered on the central posterior reflex and horizontally defined from the central reflex to the inside of the lens.

These features are extracted from both training images and test images. Features from training images together with their grades are utilized to train a grading model. The grading model can be applied to feature from a test image to predict an automatic grade.

C. Automatic Grading via SVM Regression

SVM regression [14] is a popular regression method, and a promising empirical performance has been reported. It is employed here for the purpose of grade prediction. The training procedure of SVM can be described as the following optimization problem:

$$\min \frac{1}{2} w^T w + C \sum_{i=1}^{N} \xi_i + C \sum_{i=1}^{N} \xi_i^* \tag{6}$$

with the condition of

$$\begin{aligned} & l_i - w^T \phi(f_i) - b \le \varepsilon + \xi_i \\ & w^T \phi(f_i) + b - l_i \le \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* \ge 0 \end{aligned} \tag{7}$$

where f_i denotes the feature vector of training sample i, l_i represents its associated label (grade), $\phi()$ denotes the kernel function, w is the vector of coefficients, $C > 0$ is a regularization constant, b is an offset value, and ξ_i and ξ_i^* are the slack variables for pattern f_i. In our application, radial basis function (RBF) kernel is applied.

For grade prediction, the feature vector f is extracted first for a testing image. Then the grade of the testing image is predicted using the optimized parameters w and b as

$$g(f) = w^T \phi(f) + b. \tag{8}$$

In order to follow the clinical grading protocol, all the features listed in Table I are used in the SVM regression for automatic grade prediction. Further research will be carried out to investigate the key features in the grading to improve automatic grading accuracy and benefit clinical grading as well.

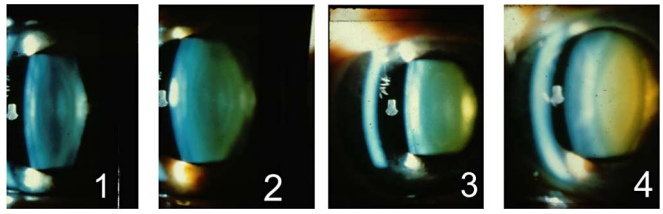

Fig. 8. Standard photographs of Wisconsin nuclear cataract grading system.

III. Results and Discussion

A. Data Description

The AGNC system was tested using the slit lamp images from a population-based study, the Singapore Malay eye study [20], [21]. The sampling frame consisted of all Malays aged 40–80 living in designated study areas in Singapore. The study was conducted at Singapore Eye Research Institute, a clinical research facility. A digital slit lamp camera (Topcon DC-1 with FD-21 flash attachment) was used to photograph the lens through dilated pupil. The slit beam was adjusted to completely fill the pupil, bisecting the lens from 12:00–6:00 at a 45° angle. Focus was placed on the sulcus of the lens. The images were saved as 24 bit color images with the size of 1536 × 2048 pixels. In total, 5850 images from right and left eyes from 3280 subjects were tested.

Human graders graded nuclear cataract based on slit lamp images using the Wisconsin cataract grading system, as previously described [5], [10], and these grades are treated as ground truth in our testing. The standard photographs for Wisconsin cataract grading system are shown in Fig. 8. Grading of nuclear cataract is accomplished by comparing each subject's photographs with the four standard photographs. A decimal grade is then assigned to the opacity using 0.1-unit intervals. For example, a grade of 3.5 means that the severity of the cataract is judged to be midway between standards 3 and 4. The range of the grade is from 0.1 to 5, in which the grade of 5 indicates the most serious case of nuclear cataract.

B. Experimental Result and Discussion

Our AGNC system was validated using 5850 slit lamp images. For the accuracy of feature extraction, lens structure detection results are evaluated. Accurate feature extraction will only be possible if the structure detection is correct. Some examples of the lens estimation and lens structure detection results are shown in Fig. 9. From our empirical observations reported later and as illustrated by these examples, we can see that the lens localization and structure detection are satisfactory given the variations of lens size and location in the images.

The result of lens localization and structure detection experiments is shown in Table II. The overlap between the automatic detection and the real lens structure is evaluated visually. The lens structure contour is assessed according to how well the automatically defined lens structure contour matches the actual lens structure contour in the image. When the overlap is between 80% and 95%, the overlap is categorized as a slight deviation.

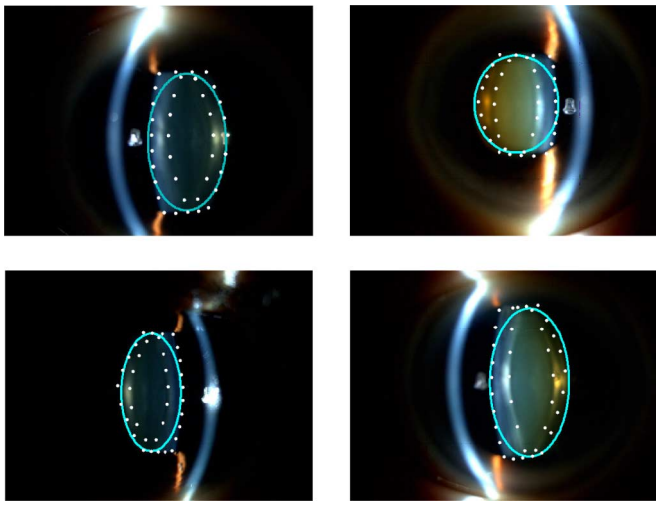

Fig. 9. Examples of lens estimation and structure detection. The ellipse represents the estimation of lens localization and the white dots are meant to outline the detected contour of lens and nucleus.

TABLE II
STATISTICAL ANALYSIS OF FEATURE EXTRACTION

	Lens Localization	Lens Structure Detection
Testing images	5850	5850
Wrong detection	23	69
Slight deviation	161	222
Success detection	5666	5559
Success rate	96.9%	95%

If the overlap is less than 80%, the overlap is categorized as a wrong detection. Successful detections are defined as those overlaps that are not partial detections or wrong detections. The AGNC system can achieve 96.9% success rate for lens localization and 95% success rate for lens structure detection. As ASM is a local searching method, all the wrong localization of lens will lead to wrong structure detection. For some images with the slightly deviated lens estimation, the modified ASM method can converge to the contour of lens structure. For the images with wrong structure detection (69 images) and slight deviation (222 images), the user intervention function of the AGNC system is provided for the user to correct the detection so that the feature extracted for these images are still valid. The experimental results show that lens structure detection is successful in 95% of the total images and only 5% of the images need user interaction, which will save a lot of workload, especially for mass screening or population study.

Among the 5850 slit lamp images, 200 images were excluded from the evaluation of automatic grading. These 200 images were marked by the clinical grader as not gradable and their ground truth is not available. One hundred images were used as the training set in the SVM regression. The images were classified into five groups according to their clinical grading results (0–1, 1–2, 2–3, 3–4, 4–5), and 20 images from each group were selected as the training set. The severity of nuclear opacity in the test images (5550 images) is automatically diagnosed using the SVM regression to predict the grades. The comparison between the automatic grades from AGNC system and grader's grading

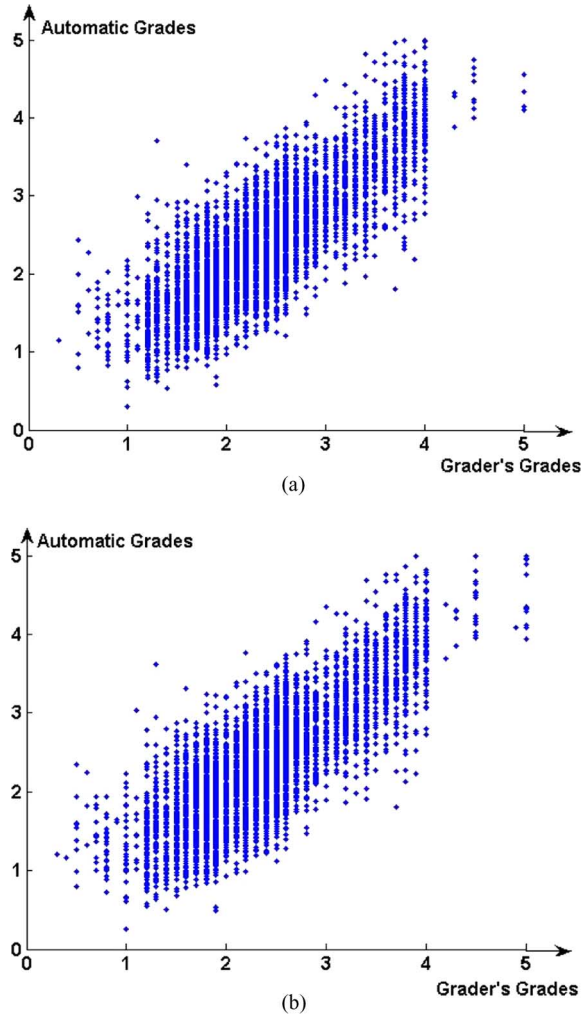

Fig. 10. Validation of automatic diagnosis. (a) Test set A (5278 images), where average difference = 0.355. (b) Test set B (5551 images), where average difference = 0.359.

results were performed. The comparison results were illustrated in Fig. 10. Two set of test results are analyzed: test set A (5278 images) excluding all images that automatic structure detection is unsuccessful or slightly deviated, and test set B (5551 images) including these images whose structure detection results were corrected using the user intervention function. Taking the grader's grading as the ground truth, the average differences of the automatic grades are 0.355 (test set A) and 0.359 (test set B), respectively, on a 5.0 scale. The difference of mean for test sets A and B is 0.018 and 0.017, respectively, and normal distribution of difference is valid for both test sets. A paired *t*-test between automatic grades and grader's grades is performed, and the results are described in Table III. It shows that there is a good agreement between AGNC automatic grades and clinical grades. The difference between automatic grades and grades of grader were further analyzed, as shown in Table IV. The grading difference for at least 97.5% of the test images was less than one grade, which is the acceptable difference in clinical diagnosis. The experimental result shows that the proposed AGNC system is promising and it is viable to apply it to clinical diagnosis later.

TABLE III
STATISTICS OF VALIDATION EXPERIMENT

	Difference of Mean	Standard Deviation	t-test	p-value
Test Set A (5278 images)	0.018	0.453	2.89	0.004
Test Set B (5551 images)	0.017	0.455	2.74	0.006

TABLE IV
STATISTICS OF GRADING DIFFERENCE

Difference	Test Set A (5278 images)		Test Set B (5551 images)	
	No. of Images	Percentage	No. of Images	Percentage
0~0.5	4106	77.8%	4262	76.8%
0.5~1	1041	19.7%	1154	20.8%
>1	131	2.5%	135	2.4%

TABLE V
COMPARISON STUDY OF AUTOMATIC DIAGNOSIS SYSTEMS

	Average grading difference
Wisconsin group[12]	0.541
AGNC system	0.359
Improvement	33.6%

The performances of AGNC system and the Wisconsin automatic grading system [12] were compared. The same training set (100 images, 20 images from each grading group) is used in the experiment to train the linear regression model. Using linear regression of the features of sulcus intensity and intensity ratio between anterior and posterior lentil, the grade prediction model in [12] is obtained as follows:

$$\text{Grade} = 0.0238\text{Sulcus} + 1.8101\text{anterior\_posterior\_ratio} - 0.7152. \quad (9)$$

Using the same test set B (5551 images), the comparison results are shown in Table V. This comparison study shows that our AGNC system can obtain 33.6% of improvement to reduce the grading difference, which indicates that the extraction of nucleus region and more meaningful features are important for automatic grading of nuclear cataract.

C. System Interface

The user interface of AGNC system is illustrated in Fig. 11. The main menu includes: load image, feature extraction, and predict grade. The interface is simple and straightforward for usage in clinics. For a large amount of images, the AGNC system has a batch processing mode to process all the test images sequentially.

One of the drawbacks of automatic disease diagnosis is that it is difficult to handle exceptional cases, whereas accuracy is the first priority in clinical diagnosis. Keeping this in mind, a user intervention function is added in the GUI so that inaccurate detection of anatomical structure can be corrected. The user can use the mouse to input four points to fine-tune the results of lens structure detection until it is satisfactory.

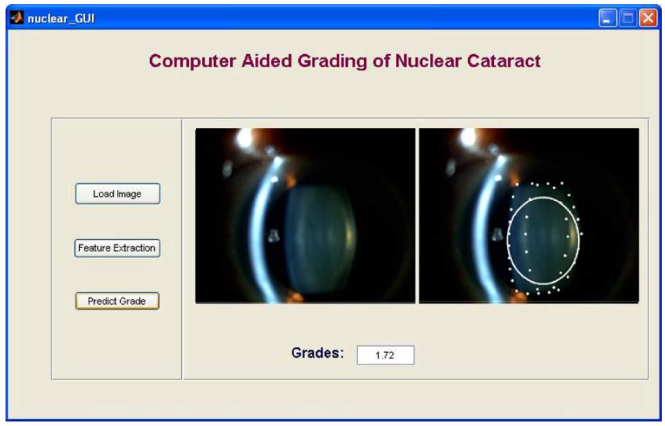

Fig. 11. User interface of the AGNC system.

IV. CONCLUSION

An automatic diagnosis system of nuclear cataract, AGNC system, is presented in this paper. To the best of our knowledge, it is the first system that can automatically detect the nucleus region satisfactorily in slit lamp images, as tested in our database of 5000 over images. The nucleus region is usually the only area that nuclear cataract is assessed clinically. Only with the robust nucleus region detection, the features such as intensity, color, and texture within the nucleus region can be measured accurately for automatic grading. In AGNC system, the intensity of sulcus can be measured throughout the nucleus region rather than on the visual axis only, which is consistent with clinical grading. This is also the first automatic nuclear cataract diagnosis system that has been validated using a large amount of clinical data. As high as 95% of the images can be diagnosed fully automatically without any user intervention, and the average error of automatic grading is 0.36 on a 5.0 scale. The user intervention function is provided so that the AGNC system can handle the special images due to inaccurate focus, small pupil, or dropping upper lid. The promising experimental results show that the AGNC system can be utilized to facilitate clinical diagnosis, and the system is currently in the stage of clinical validation.

REFERENCES

[1] *Cataract: What You Should Know*, Nat. Eye Inst. Nat. Inst. Health, Bethesda, MD, NIH Publication No: 03-201, Sep. 2003.
[2] *State of the World's Sight: VISION 2020: The Right to Sight: 1999–2005*, World Health Organization, Geneva, Switzerland, 2005.
[3] The Eye Diseases Prevalence Research Group, "Prevalence of cataract and pseudophakia/aphakia among adults in the united states," *Arch. Ophthalmol.*, vol. 122, pp. 487–494, 2004.
[4] T. Y. Wong, S. C. Loon, and S. M. Saw, "The epidemiology of age related eye diseases in Asia," *Br. J. Ophthalmol.*, vol. 90, pp. 506–511, 2006.
[5] P. Mitchell, R. G. Cummings, K. Attebo, and J. Panchapakesan, "Prevalence of cataract in Australia: The Blue Mountains eye study," *Ophthalmology*, vol. 104, pp. 581–588, 1997.
[6] P. J. Foster, T. Y. Wong, D. Machin, G. J. Johnson, and S. K. L. Seah, "Risk factor for nuclear, cortical and posterior subcapsular cataracts in the Chinese population of Singapore: The Tanjong Pagar survey," *Br. J. Ophthalmol.*, vol. 87, pp. 1112–1120, 2003.
[7] P. A. Asbell, I. Dualan, J. Mindel, D. Brocks, M. Ahmad, and S. Epstein, "Age-related cataract," *Lancet*, vol. 365, pp. 599–609, 2005.
[8] L. T. Chylack, J. K. Wolfe, D. M. Singer, M. C. Leske, M. A. Bullimore, I. L. Bailey, J. Friend, D. McCarthy, and S. Y. Wu, "The lens opacities

classification system III," *Arch. Ophthalmol.*, vol. 111, pp. 831–836, 1993.
[9] J. M. Sparrow, A. J. Bron, N. A. Brown, W. Ayliffe, and A. R. Hill, "The oxford clinical cataract classification and grading system," *Int. Ophthalmol.*, vol. 9, no. 4, pp. 207–225, 1986.
[10] B. E. K. Klein, R. Klein, K. L. P. Linton, Y. L. Magli, and M. W. Neider, "Assessment of cataracts from photographs in the Beaver Dam eye study," *Ophthalmology*, vol. 97, no. 11, pp. 1428–1433, 1990.
[11] N. J. Ferrier, "Automated identification of the anatomical features in slit lamp photographs of the lens," *Invest. Ophthalmol. Vis. Sci.*, vol. 43, p. 435, 2002.
[12] S. Fan, C. R. Dyer, L. Hubbard, and B. Klein, "An automatic system for classification of nuclear sclerosis from slit-lamp photographs," in *Proc. 6th Int. Conf. Med. Image Comput. Comput.-Assist. Intervention*, 2003, pp. 592–601.
[13] D. D. Duncan and O. B. Shukla, "New objective classification system for nuclear opacification," *Opt. Soc. Amer.*, vol. 14, no. 6, pp. 1197–1204, 1997.
[14] A. J. Smola and B. Scholkopf, "A tutorial on support vector regression," *Stat. Comput.*, vol. 14, pp. 199–222, 2004.
[15] H. Li, J. H. Lim, J. Liu, T. Y. Wong, A. Tan, J. Wang, and M. Paul, "Image based grading of nuclear cataract by SVM regression," in *Proc. SPIE Med. Imag.*, 2008, pp. 691536-1–691536-8.
[16] H. Li, J. H. Lim, J. Liu, and T. Y. Wong, "Towards automatic grading of nuclear cataract," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2007, pp. 4961–4964.
[17] H. Li and O. Chutatape, "Boundary detection of optic disk by a modified ASM method," *Pattern Recognit.*, vol. 36, no. 9, pp. 2093–2104, 2003.
[18] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-Their training and application," *Comput. Vis. Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.
[19] A. B. Williams and F. J. Taylors, *Electronic Filter Design Handbook*. New York: McGraw-Hill, 1988.
[20] A. W. Foong, S. M. Saw, J. L. Loo, S. Shen, S. C. Loon, M. Rosman, T. Aung, D. T. H. Tan, E. S. Tai, and T. Y. Wong, "Rationale and methodology for a population-based study of eye diseases in Malay people: The Singapore Malay eye study (SiMES)," *Ophthalmic Epidemiol.*, vol. 14, no. 1, pp. 25–35, 2007.
[21] T. Y. Wong, E. W. Chong, W. L. Wong, M. Rosman, T. Aung, J. L. Loo, S. Shen, S. C. Loon, D. T. H. Tan, E. S. Tai, and S. M. Saw, "Prevalence and causes of low vision and blindness in an urban Malay population: The Singapore Malay eye study," *Arch. Ophthalmol.*, vol. 126, no. 8, pp. 1091–1099, 2008.

Huiqi Li (S'03–M'04) received the B.E. and M.E. degrees from Harbin Institute of Technology, Harbin, China, and the Ph.D. degree from Nanyang Technological University, Singapore, both in electrical and electronic engineering.

She is currently a Senior Research Fellow at the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore. She has more than ten years of research experience on medical image processing. Her research work has been published in leading journals and international conferences. Her research interests include medical image processing, computer-aided disease analysis, pattern recognition, and computer vision.

Joo Hwee Lim received the B.Sc. (Hons. I) and M.Sc. (by research) degrees in computer science from the National University of Singapore, Singapore, and the Ph.D. degree in computer science and engineering from the University of New South Wales, Kensington, N.S.W., Australia.

Since October 1990, he has been at the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore, where he is currently the Head of the Computer Vision and Image Understanding Department. He is also the Co-Director of the Image and Pervasive Access Laboratory, a French-Singapore Joint Laboratory. He has authored or coauthored more than 140 international refereed journal and conference papers. He has also coauthored 15 patents (awarded and pending).

Jiang Liu received the B.E. degree from the University of Science and Technology China, Hefei, China, and the M.S. and Ph.D. degrees from the School of Computing, National University of Singapore, Singapore.

During the past 22 years, he was working with industries, where he holds technical development and management positions. In the past six years, his research interests have been in medical imaging. He is currently the Head of the Feature and Segmentation Group and Intelligent Medical Imaging Group, Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore. He is the Adjunct Research Scientist in Singapore Eye Research Institute, Singapore General Hospital. His research interests include computed tomography, retina image processing, as well as medical-image-based computer-assisted diagnosis/computer-aided surgery.

Paul Mitchell, photograph and biography not available at the time of publication.

Ava Grace Tan received the B.Sc. (with Honors) degree and the Master's of Applied Information Technology degree in 2002 and 2004, respectively, from the University of Sydney, Sydney, N.S.W., Australia, where she is currently working toward the Master's of Public Health degree.

She is currently a Research Assistant at the Centre for Vision Research, Department of Ophthalmology, University of Sydney.

Ms. Tan is a member of the Association for Research in Vision and Ophthalmology.

Jie Jin Wang received the B.S. and M.S. degrees in medicine from Sun Yat-Sen University of Medical Sciences, China, the Master's and Ph.D. degree in epidemiology from the University of Sydney, Sydney, N.S.W., Australia, and the Master's degree of applied statistics from Macquarie University, North Ryde, N.S.W., Australia.

She is currently an Associate Professor at the Centre for Eye Research Australia, University of Melbourne, Melbourne, Vic., Australia, and the Centre for Vision Research, University of Sydney. She is a Chief Investigator of the Blue Mountains Eye Study with 300 publications, and a recipient of the Australian National Health and Medical Research Council Senior Research Fellowship since 2005.

Dr. Wang joined the Association for Research in Vision and Ophthalmology in 1997, where she is a member of the Program Committee of Clinical/Epidemiologic Section (2002–2005).

Tien Yin Wong graduated in medical from the National University of Singapore, Singapore, and received the Ph.D. degree from Johns Hopkins University, Baltimore, MD.

He was with Singapore National Eye Centre as a Clinical Ophthalmology Trainee. He is currently a Professor and the Director of Singapore Eye Research Institute, National University of Singapore, Singapore. He is also a Professor and the Director of the Retinal Vascular Imaging Centre, Centre for Eye Research Australia, University of Melbourne, Melbourne, Vic., Australia. He leads a research program on the use of novel retinal imaging techniques to predict cardiovascular disease. He has authored or coauthored more than 400 papers and given more than 100 lectures around the world.