Journal of Medical Imaging

SPIEDigitalLibrary.org/jmi

Automatic nuclear cataract grading using image gradients

Ruchir Srivastava
Xinting Gao
Fengshou Yin
Damon W. K. Wong
Jiang Liu
Carol Y. Cheung
Tien Yin Wong

SPIE

Automatic nuclear cataract grading using image gradients

Ruchir Srivastava,^{a,*} Xinting Gao,^a Fengshou Yin,^a Damon W. K. Wong,^a Jiang Liu,^a Carol Y. Cheung,^b and Tien Yin Wong^b

^aInstitute for Infocomm Research, 138632 Singapore

^bSingapore Eye Research Institute, Singapore 168751, Singapore

Abstract. This paper deals with automatic grading of nuclear cataract (NC) from slit-lamp images in order to reduce the efforts in traditional manual grading. Existing works on this topic have mostly used brightness and color of the eye lens for the task but not the visibility of lens parts. The main contribution of this paper is in utilizing the visibility cue by proposing gray level image gradient-based features for automatic grading of NC. Gradients are important for the task because in a healthy eye, clear visibility of lens parts leads to distinct edges in the lens region, but these edges fade as severity of cataract increases. Experiments performed on a large dataset of over 5000 slit-lamp images reveal that the proposed features perform better than the state-of-the-art features in terms of both speed and accuracy. Moreover, fusion of the proposed features with the prior ones gives results better than any of the two used alone. © 2014 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.1.1.014502]

Keywords: biomedical image processing; nuclear cataract; automatic grading; image gradients.

Paper 13032R received Dec. 30, 2013; revised manuscript received Mar. 12, 2014; accepted for publication May 9, 2014; published online Jun. 4, 2014.

1 Introduction

Cataract is the leading cause of blindness in the world, accounting for almost half of the blindness.¹ Cataract can be categorized into three types based on the location of opacity within the lens structure: nuclear, cortical, and posterior subcapsular (PSC).² Nuclear cataract (NC) begins at the center of the lens and spreads toward the surface. Cortical cataract begins at the outer rim of the lens and moves toward the center. PSC forms at the back of the lens. Of the three types, this paper focuses on NC, which is the most common of the three types.

Due to lack of proven preventive strategies, NC patients have to resort to surgery, for which to be effective, timely diagnosis of NC is essential. While diagnosing NC, it is helpful to quantify the severity of NC. This quantification is performed based on scientific grading systems, such as the Lens Opacities Classification System III (LOCS III),³ Oxford Clinical Cataract Classification and Grading System (OCCGS),⁴ Wilmer system,⁵ and the Wisconsin grading system.⁶ As an example, considering the Wisconsin grading system, it grades the level of NC based on opacity and color of lens nucleus as observed in a slit-lamp image (Fig. 1). The standard grades for NC are from 0 to 5 denoting an increasing severity.

Traditionally, grading is conducted manually through visual examination, which, however, is prone to be subjective and time consuming. Due to the extensive labor involved, the grader may also lose focus and cause variations even within his/her own grading.⁷ Considering these limitations of manual grading of NC, methods have been proposed for automating the process. Automatic grading of cataract is also helpful in monitoring the progression of NC, which assists clinical studies.

The considerations for automatic NC grading are:

- *Which cue to use for the task?* For NC grading, two cues have been identified:⁸ (1) brightness and color of the eye lens and (2) visibility of parts (or landmarks) of the lens [Fig. 2(b)]. These cues indicate that the eye lens is the region of interest (ROI) for automatic grading.
- *How to accurately localize the lens automatically?* For automation of NC grading, the lens should be automatically localized, but until recently it was manually marked.
- *Which part of the lens should be used for the task?* Ref. 9 uses the full lens while Ref. 10 defines a circular ROI within the lens nucleus and extracts features from that region.
- *Which features to use?* Different features have been tried out (Sec. 2), but most of the works neglected the visibility of lens landmarks (visibility cue) while extracting features.

In an attempt to address the above-mentioned considerations, this paper proposes a system called ACASIA-NC_v0.10 (Automatic Cataract Screening from Image Analysis-Nuclear Cataract, Version 0.10), which attempts to address the issues identified above for automatically grading NC. Given a raw grayscale image, the eye lens and its nucleus are automatically localized. To enhance the visibility of the lens landmarks, the localized nucleus is preprocessed (Sec. 4.1). Features are derived based on image gradients to utilize the visibility cue (Sec. 4.2). NC grade prediction is formulated as a supervised regression problem, and support vector regressor (SVR) is used (Sec. 4.3) for regression.

1.1 Contribution

The main contribution of this paper is in using gray level image gradients to utilize the visibility cue for computerized grading of

*Address all correspondence to: Ruchir Srivastava, E-mail: srivastavar@i2r.a-star.edu.sg

Fig. 1 Standard slit-lamp photographs of the Wisconsin cataract grading system. The numbers from 0 to 5 indicate the grade of NC severity. A lower number means lower severity. Note that although there are four standard images corresponding to grades 1 to 4, the grades are in the range 0 to 5 (Sec. 5.1). This grading system has been used to mark the ground truth.

NC. Although image gradients have been used in other applications (Sec. 2), their application to NC grading is novel.

The core idea behind ACASIA-NC_v0.10 is that the visibility of landmarks introduces prominent edges in the lens region, which can be captured using image gradients (referred to henceforth as gradients for simplicity). ACASIA-NC_v0.10 uses gradient-based features (Sec. 4.2), which have been found to perform better (Sec. 5.4) than the state-of-the-art features proposed in Ref. 8 (referred to henceforth as prior features) in terms of both speed and accuracy. In addition, fusion of the proposed gradient-based features with the prior features performs better than either of the two features used alone (Sec. 5.9).

2 Related Works

The works covered in this section are on two lines: (1) usage of image gradients and (2) automatic NC grading.

Image gradients have been widely used in applications, such as object recognition, three-dimensional modeling, gesture recognition, and video tracking. Popular feature descriptors using image gradients are scale-invariant feature transform (SIFT),¹¹ histogram of oriented gradients (HoG),¹² gradient location and orientation histogram (GLOH),¹³ speeded up robust features (SURF),¹⁴ and the machine-optimized gradient-based descriptors.¹⁵ Although these descriptors have been shown to be successful in various applications, the “application” of image gradients to automatic NC grading is novel.

Proposed gradient-based features (referred to henceforth as proposed features) use orientation histograms, which are similar to that in Ref. 12 but without illumination normalization since it was not considered as necessary for our problem. Reference 12 dealt with pedestrian detection, in which illumination variations are more prominent. Images of pedestrians can be in shadow, sunlight, night, indoors, or outdoors. They are usually shot with different cameras with possibly different settings. In our case, all the images were shot in a controlled environment with the same camera and same settings (details are in Sec. 5.1).

Considering automatic NC grading, Ref. 16 proposed intensity-based features for the grading task. Grades were predicted by a regressor trained with sulcus intensity and intensity ratio between anterior and posterior lentils [Fig. 2(b)] as features. Reference 10 computed the longitudinal profile of luminance in a circular ROI located in the nuclear region. Nuclear mean gray level, slope at the posterior point of profile, and the fractional residual of the polynomial least-squares fit were extracted as features and used to train a neural network to predict the grade.

Reference 17 located lens automatically for the first time and extracted mean intensity of the lens as the feature. Subsequent works by the same group^{8,18} proposed novel color, gray level intensity, and entropy-based features extracted from the nuclear region. The features consist of mean values of intensity, color (in each channel of HSV space), entropy, and neighborhood standard deviation in the lens contour. These features were also extracted from inside the lens nucleus, which is a part of the lens. Other features were intensity ratio between nucleus and lens, sulcus intensity, intensity ratio between sulcus and nucleus, intensity ratio between anterior and posterior lentils, strength of nucleus edge, and color on posterior reflex. Support vector machines regression was used to predict the NC grade.

Color and intensity-based features were also used in Ref. 19, in which a new algorithm based on ranking method was proposed for NC grading. NC grade for a test slit-lamp image was predicted using its neighboring labeled images in a ranked image list constructed using a learned ranking function. To the best of our knowledge, existing works on computerized diagnosis and/or grading of NC have only explored color or gray level information. Visibility cue has not received much attention.

Huiqi et al.^{8,18} have attempted to utilize the visibility cue by proposing features, such as entropy inside both the lens contour and lens nucleus contour in the slit-lamp image. Features based on intensity of the sulcus region are also used. However, we believe that the utilization of the visibility cue can be explored further by using image gray level intensity information as proposed in this work. The gray level image gradient-based features used in this work make it different from the previous related works,^{8,17,18} which did not explore image gradients.

3 Importance of Gradient for Nuclear Cataract Grading

This section explains the intuition behind using gradients for NC grading.

As per the visibility cue (Sec. 1), nuclear landmarks [Fig. 2(b)], especially the sulcus is clearly visible for grade 1 and form edges in the nuclear region. The visibility of nuclear landmarks diminishes as the NC grade increases (Fig. 3). Consequently, the edges are also indistinct in higher grades of NC. This can be understood with an example presented in Fig. 2(a). The idea illustrated in this example was experimentally tested as explained in detail in Sec. 5.3.

Fig. 2 (a) Prominent edges are marked in the lens nuclear region in slit-lamp images for (A) beginning stage and (B) severe stage of NC. Observe the difference in the smoothness in lens nucleus images. This is the motivation behind using gray level intensity gradients (related to smoothness) for automatic grading of NC (Sec. 3). (b) Parts of the right eye as viewed in a slit-lamp image.⁸ Nucleus (demarcated by a boundary) is taken as the region of interest (ROI) in the proposed approach (Sec. 4.1).

Fig. 3 Examples of images of each grade. Observe the disappearance of the landmarks within the nucleus, which is a prominent difference between grades 1 and 5 (Sec. 3).

The above-mentioned characteristic of the nuclear region makes gradient an important measure in distinguishing between different grades of NC.

4 Methodology

A flowchart of the proposed system is shown in Figs. 4(a) and 4(b). Major steps in the flowchart are described below.

4.1 Preprocessing

As discussed in Sec. 1, since the difference between the NC grades lies in the eye's nucleus region, only the nucleus is taken as the ROI for the proposed system.

First, the lens region is localized automatically and then structure is detected using a modified version of active shape model (ASM).⁸ Lens localization refers to just an initial estimate of lens's location, which is obtained by first locating the corneal bow. In a slit-lamp image of the eye lens, the corneal bow is usually the leftmost (for right eye) or rightmost (for left eye) brightest vertical curve [Fig. 2(b)]. Corneal bow is located by selecting the 10% brightest pixels in the grayscale image. Next, the lens is located as the largest part in the foreground, which

occupies 20%–30% of the whole image and which appears in the center of an image. Horizontal and vertical profile clustering is used to locate the lens.

Lens structure detection is finding the exact contour of the lens and the lens nucleus. This is achieved using ASM. The lens location is used to initialize ASM. For ASM, the lens along with nucleus contour are modeled using 38 points. ASM learns this model and then tries to fit the model in a test image to detect lens structure. More details are in Ref. 8. Henceforth, ROI detection refers to lens nucleus detection and not just lens localization. Once the nucleus is extracted as the ROI, in order for further analysis it is resized to a standard size, which we refer to as ROI size say of $N \times N$ pixels.

The resized ROI is histogram equalized to enhance visibility of edges. This technique increases the global contrast of an image and is especially useful when the grayscale intensity values of the image lie within a narrow range. Histogram equalization was performed using MATLAB function “histeq” and the associated threshold was chosen using Otsu's method (“gray-thresh” function in MATLAB). After image enhancement, features are extracted from the ROI as explained in the next section.

Fig. 4 (a) Flowchart for the proposed approach (Sec. 4). (b) Flowchart for preprocessing of the raw image (Sec. 4.1). (c) Feature extraction is performed by subdividing the enhanced ROI into a 10×10 grid numbered in the shown order. Partial feature vector (represented by a colored block) extracted from each subimage has a unique position in the final feature vector. This captures the location information (Sec. 4.2.1).

4.2 Feature Extraction

Gradients have three aspects: location, orientation, and magnitude. Features are extracted to capture each of these aspects as explained below.

4.2.1 Capturing location information

To capture the location information, the extracted ROI is divided into a square grid and features extracted from each subimage of the grid are concatenated together to form the final feature vector [Fig. 4(c)]. To define the division of ROI, we use the term “grid,” such that when a grid of size $G \times G$ is used to divide the ROI, it divides the ROI into $G \times G = G^2$ subimages. Here, $G = N/m$ if ROI is of size $N \times N$ pixels and each subimage is of size $m \times m$ pixels. Since different subimages correspond to different positions in the feature vector, the location information is inherently captured in the feature vector. This scheme has been used in the computer vision community.²⁰

4.2.2 Features related to magnitude

Once location of a subimage is captured, number of prominent edge pixels in the subimage is computed by counting the number of pixels where horizontal gradient magnitude exceeds a predefined threshold (chosen as mentioned in Sec. 5.4). A similar count is made for vertical gradient. Image gradient is computed using the MATLAB function “gradient.” Mathematically, if $C(x, y)$ is the image representing a square subimage of size $m \times m$ and its gradient is given by

$$\nabla C(x, y) = \frac{\partial C(x, y)}{\partial x} \hat{i} + \frac{\partial C(x, y)}{\partial y} \hat{j}, \quad (1)$$

the number of prominent edge pixels along x and y directions (p_x and p_y , respectively) is given by

$$p_x = \sum_{j=1}^m \sum_{i=1}^m H \left[\left. \frac{\partial C(x, y)}{\partial x} \right|_{x=j, y=i} - \theta_x \right]$$

$$p_y = \sum_{j=1}^m \sum_{i=1}^m H \left[\left. \frac{\partial C(x, y)}{\partial y} \right|_{x=j, y=i} - \theta_y \right], \quad (2)$$

where i and j denote pixel coordinates along y and x directions, respectively; θ_x and θ_y are gradient thresholds along x and y directions, respectively; and H is the Heaviside step function given by

$$H[n] = \begin{cases} 0, & n < 0, \\ 1, & n \geq 0. \end{cases} \quad (3)$$

p_x and p_y along with mean and standard deviation of gray level intensity inside the subimage constitute the magnitude-related features.

4.2.3 Features related to orientation

To capture the edge orientation in a subimage, a histogram of gradient orientations is computed for the subimage and added to the feature set. Implementation details are given in Sec. 5.4.

4.3 Nuclear Cataract Grade Prediction

Since grades need to be predicted as decimal numbers instead of whole numbers, NC grading is formulated as a supervised regression problem. SVR is chosen as the tool. Details about training and testing datasets are described in Sec. 5.1.1.

5 Experimental Results and Discussion

5.1 Dataset

The dataset used for evaluating the proposed method is called the ACHIKO-NC dataset. It consists of 5378 slit-lamp images of one or both eyes of 2961 subjects, which were Malays aged

40–80 living in Singapore. These images were collected as a part of the Singapore Malay Eye Study^{21,22} conducted by Singapore Eye Research Institute. The 5378 images were formed after removing irrelevant images from total 5850 images (Refer to Sec. 5.2 for details).

The images were captured by passing a slit beam through dilated pupil and filming it using a digital slit-lamp camera (Topcon, Japan DC-1 with FD-21 flash attachment). The beam was focused at the sulcus of the lens and bisected the lens from 12:00 to 6:00 at a 45-deg angle. Captured images are of size 1356×2048 pixels and stored in jpeg format. They are converted into grayscale (using MATLAB function “rgb2gray”) for experimentation.

The ground truth is marked by an expert based on the Wisconsin cataract grading system.^{6,7} There are four standard images in the Wisconsin system, which have been assigned grades from 1 to 4 corresponding to four progressive stages of NC (Fig. 1). However, the assigned grades range from 0 to 5 when marked using the following process: a sample image was visually compared with each of these standard photographs and assigned a grade, up to first decimal place, indicating the severity of NC in the image. For example, a grade of 2.1 means that the NC severity is a little more than that in the standard image of grade 2. The assigned grades range from 0.1 to 5, with 5 corresponding to the most severe case of NC, even more than the standard grade 4.

5.1.1 Training and testing sets

All the 5378 images were divided into five groups based on their grading, i.e., 0–1, 1.1–2, 2.1–3, 3.1–4, and 4.1–5. Twenty images from each of these groups were selected to form the training set. All the remaining images (5278 in total) constitute the testing set. Since samples in the range 0–1 and 4.1–5 were few, more samples could not be selected for training while maintaining equal number of samples from each group.

5.2 Region of Interest Detection

This section discusses the performance of ROI detection. Accuracy in lens localization and ROI detection is important since it is a step prior to feature extraction. This accuracy was assessed by manual examination of the overlap between predicted and actual ROIs. The detection was considered successful when this overlap was greater than 95%. An overlap of 80%–95% was considered as slight deviation, and a lesser overlap was considered as a wrong detection. ROI detection was run on 5850 images and the lens structure was successfully detected with a rate of 95%. Images with wrong or failed localization were removed. In addition, images marked as ungradable by the grader were also removed, reducing the final number of images to 5378, which constitute the ACHIKO-NC dataset.

Figures 5(a) and 5(b) show sample results of successful and unsuccessful ROI localization, respectively. In Fig. 5(b), the top row shows slight deviations in localization while the bottom row shows cases of wrong localization. It was observed that the ROI localization was affected sometimes by artifacts produced during the imaging process, such as a bright patch.

After localization, the ROIs were resized to 200×200 pixels before further feature extraction.

Fig. 5 (a) Examples showing results of successful lens localization for both low (top row) and high (bottom row) grades of NC. The dots outline the detected lens and nucleus contours. (b) Examples of unsuccessful lens localization. Top row shows slight deviations in localization, while the bottom row shows cases of wrong localization.

5.3 Testing the Effectiveness of Gradients

In order to test the effectiveness gradients, five sets were created based on their grading, i.e., 0–1, 1.1–2, 2.1–3, 3.1–4, and 4.1–5. For each set, the number of prominent edge pixels [p_x and p_y from Eq. (2)] is counted for both horizontal and vertical directions (Sec. 4.2) and the mean of p_x and p_y is recorded. The mean number of prominent edge pixels for each set is shown in Fig. 6(a). This value for the first set was 1033 per image, which means that there were many prominent edge pixels for images with lower grades. The value decreases as the grade increases and for the last set the same number was around 20 per image, which is much lower than that for the first set. This indicates that there are actually more prominent edge pixels in images of lower grade. This observation supports the idea that the gradient information can be used to distinguish between higher and lower grades of NC.

5.4 Nuclear Cataract Prediction

After resizing, the ROI was subdivided into 100 subimages, such that each subimage is of size 20×20 pixels. Features

Fig. 6 (a) Mean number of edge pixels for each range of NC grade. It can be seen that the images of lower grade have more edge pixels making gradient information important for predicting NC grade. (b) Effect of image enhancement for ROIs for different grades. For each pair from (A) to (D), left and right images are those before and after enhancement.

are extracted from each subimage as described in Sec. 4.2. Gradient threshold for computing the number of prominent edge pixels was chosen as 0.1. For computing orientation-related features, orientation values were binned into 13 bins from -180 to $+180$ deg, such that each bin, except the last one, spanned 30 deg. The last bin contained number of orientation values equal to $+180$ deg. These yielded 13 features per subimage. In addition, there are four magnitude-related features per subimage (Sec. 4.2). The features extracted from each subimage are:

- Features 1–13: Histogram counts for 13 bins.
- Features 14,15: Number of prominent edge pixels along x and y directions, p_x and p_y Eq. (2).
- Feature 16: Mean gray level intensity inside the subimage.
- Feature 17: Standard deviation of the gray level intensities of pixels inside the subimage.

Features from different subimages are concatenated as shown in Fig. 4(c). The total number of features from 100 subimages add up to 1700. The above-mentioned parameters will be used for all the experiments reported henceforth, unless different values are stated.

Gradient threshold was chosen based on a cross validation on the training set. Other parameters were found to give the best results as reported in the upcoming sections. Note that we did not intend to choose the best parameters, instead we wish to show the effect of varying these parameters on the results for the test set. Therefore, while varying these parameters (except gradient threshold), experiments were conducted on the test data and no cross validation on training data was performed.

To select the most relevant features, feature selection was performed using the Fisher ratio test.²³ This test computed a ratio for each feature, which indicates the discriminatory power of the feature. Note that Fisher ratio test is applicable only to the two-class problem, considering which the training dataset was divided into two classes with assigned grades in

the ranges 0–2.5 and 2.6–5, respectively. However, this division was made only for feature selection and NC prediction was still a five-class problem.

An SVR²⁴ (of type nu-SVR) was used to predict grades for test images. Prediction for a few sample images is shown in Fig. 7(a) with ground truth grades lying in each range of the grades. Prediction performance was assessed using the average difference between predicted grades and the ground truth (average grading difference) as the performance metric (referred to, henceforth, as “error”). This measure has been used in past works.^{8,16} Note that for each test image, the absolute difference between the predicted grade and ground truth was used before averaging this absolute difference over all test images.

Our results are compared with the results using prior features⁸ since it is closest to our work. Moreover, Ref. 8 already reported better performance than Ref. 16. As compared to Ref. 8 on the same dataset, it can be seen from Fig. 8(a) that ACASIA-NC_v0.10 achieves more than 8.5% reduction in the error. Note that ACASIA-NC_v0.10 just uses the proposed gradient-based features, while fusion uses both the prior features and the gradient-based features. Results for fusion are discussed in Sec. 5.9.

The reduction in error effectively means that as compared to prior works, the predicted grades are closer to actual grade. Figure 9(a) shows a scatter plot of predicted grade versus actual grade. For an ideal prediction, $P = A$ and all the plotted points would lie on the bold straight line. It can be observed that most predictions lie within an error of 0.5. The exact statistics are presented in Fig. 9(b), which shows the fraction of images for which the error lies in each of the three ranges. These results show that for ACASIA-NC_v0.10, there are fewer cases with error greater than 1 and more cases with error less than 0.5.

Performance of ACASIA-NC_v0.10 for each range of grades is presented in Fig. 10(b). The error was found to be very large for the extreme range of grades, i.e., grade ≤ 1 and grades lying between 4.1 and 5. Since the number of images belonging to these extreme grades is very small (1.7% of all images) as evident from Fig. 10(a), the poor performance on these ranges is not reflected

Fig. 7 (a) Result of NC grade prediction on a few test images. Each row shows images with ground truth of the grades lying in one of the five ranges defined in Sec. 5.1.1. First row corresponds to the range 0–1, second corresponds to 1.1–2, and so on. In each row, the last column shows the result with error >1, which is considered as a poor prediction. (b) A drawback of image enhancement: artifacts already present in the extracted ROI can get enhanced after image enhancement. On the top is the ROI before and bottom is the ROI after image enhancement. In this case, artifacts are in the form of horizontal lines in the ROI.

in the final result. This analysis is important in cases of unbalanced distribution of test data across different classes.

Worse performance for extreme grades, especially for severe cases of cataract, is an issue of concern for the automatic grading of NC since immediate action may be required at the severe stages.

5.4.1 Computational effort

It was observed that in ACASIA-NC_v0.10, computing features is more than 80 times faster than computing the prior features [Fig. 8(b)]. As a drawback of having more features, training and testing times are much more in ACASIA-NC_v0.10 as

Fig. 8 (a) Comparison of the accuracy of ACASIA-NC_v0.10 and prior work.⁸ Results of fusing the proposed and prior features are also presented. (b) Comparison of proposed features with the prior features in terms of computation time. Experiments were run on Octave 3.6.1 on a Windows-7 platform with 4GB RAM, 32-bit, 3.4 GHz processor.

Fig. 9 (a) Scatter plot of predicted and actual grades for the 5278 test images. It can be seen that most of the predicted grades lie within an error of 0.5 (Sec. 5.4). (b) Statistics of grading difference. The numbers shown on the pie chart correspond to the number of images for which grading difference lies in each of the three ranges, 0–0.5, 0.6–1, and >1 (Sec. 5.4).

compared to Ref. 8. However, for testing a new image with an already trained regressor, feature extraction time will play the major role.

5.4.2 Cross validation

For evaluating the consistency of the two methods, a cross-validation experiment was performed using both systems. Twenty images were randomly selected from each of the five groups mentioned in Sec. 5.1.1 to form the training set of 100 images. The remaining images were used for testing. This process was repeated 10 times. Mean errors with standard deviations over the 10 runs were found to be 0.33 ± 0.007 and 0.37 ± 0.029 for ACASIA-NC_v0.10 and Ref. 8, respectively. ACASIA-NC_v0.10 consistently performed better than Ref. 8 and also had a lower standard deviation.

5.5 Effect of Image Enhancement

The improvement in visibility after enhancement is seen in Fig. 4(b), in which the edge due to the sulcus is more distinct after enhancement. Note that the enhancement was performed using histogram equalization (Sec. 4.1). Figure 6(b) shows

the effect of enhancement for cases with different grades of NC. As shown in the figure, it was observed that the enhancement has greater effect on images of lower grades than on images of higher grades. This is possible because extracted ROIs have dark extraneous parts at the corners, which do not belong to the eye lens. For lower grades, the lens region is darker, making the overall contrast low, and thus image enhancement was more effective in enhancing the sulcus. Eye lens for higher grades are brighter as compared to those for lower grades. The dark corners result in an overall high contrast for the ROIs of higher grades. However, it is the sulcus that needs enhancement to become visible. Higher-grade images do not have a visible sulcus. Hence, images of higher grade may not need much enhancement in the first place. Hence, even if there is not much difference after enhancing image of higher grade, it may not affect the final NC grade prediction results. In fact, the average grading difference was found to reduce from 0.35 to 0.32 after image enhancement.

5.6 Effect of Varying Region of Interest Size

ROIs extracted from different images are of different sizes and need to be resized to a standard size before feature extraction.

Fig. 10 (a) Number of test images for each range and (b) performance of ACASIA-NC_v0.10 for each range of grades.

Fig. 11 (a) Prediction performances for some of the possible standard ROI sizes. The size of 200×200 pixels gave the best results, hence all ROIs were resized to this size. (b) Prediction performances for different grid sizes. The error and the computation time were almost the same for all sizes. The size of 10×10 was chosen for experiments.

Choosing the standard size of all ROIs is important since reducing the actual size may lead to loss of edge details while enlarging the ROI involves interpolation, which may add noise.

Experiments were performed on the full dataset with ROI sizes of 100×100 , 200×200 , 500×500 , 1000×1000 , and 1200×2000 pixels. As shown in Fig. 11(a), best prediction results were obtained for the size of 200×200 pixels.

Fig. 12 (a) Locations of the 150 topmost features (about top 10%) after feature selection. The locations correspond to those in the enhanced ROI shown in (b). Brighter subimage means more features come from that region of the ROI.

5.7 Effect of Varying Grid Size

Grid resolutions of 4×4 , 6×6 , 8×8 , 10×10 , and 12×12 were tried on the full dataset. Prediction results [Figure 11(b)] show only a minor variation in the error for all the sizes except the size of 4×4 , which performed the worst. The best results correspond to a resolution of 10×10 . Since the computation time for all the pixels sizes was almost the same, the grid size of 10×10 was chosen for reporting the final results. But other sizes also could have been chosen.

Please note that the values of ROI size and grid size are specific for this dataset and the observations reported here may not generalize for all datasets.

5.8 Analysis of Feature Selection

Feature selection was performed based on the value of Fisher ratio (Sec. 5.4). A higher Fisher ratio reflects better discriminatory power.

Locations (in the ROI) yielding the top 150 (approximately 10%) features, based on values of Fisher ratios, are shown in Fig. 12(a) and an ROI is shown in Fig. 12(b) for reference. The figure shows the contribution of each subimage (in the 10×10 grid) to the top 150 features. A brighter subimage reflects a higher contribution means more features from that the subimage

Fig. 13 (a) Distribution of top features into prior and proposed. It can be seen that the proposed features form a majority of the top 20 features. (b) Variation of error with number of features. The lowest error corresponds to 774 features.

constitutes the top 150 features. It is observed that most of the top features come from the sulcus region, mainly its boundary. This is consistent with our proposition of using gradients since gradients are higher near the sulcus boundary. A similar finding was reported for human detection using gradient-based features.¹²

Figure 13(b) shows how the performance varies with the number of features used. The lowest error was obtained with 774 features.

5.9 Fusion of Proposed and Prior Features

A feature level fusion of prior features with the proposed ones was tried and results using SVR are also presented in Fig. 8(a). Proposed and prior features were pooled together and an analysis of the Fisher ratios for the feature pool showed that the proposed features form majority of the most discriminating features out of the feature pool [Fig. 13(a)]. The total number of prior features is only 21, hence more than 20 top features were not considered to avoid bias toward new features, which are much more in number (numbering 1700).

6 Conclusion

The proposed system, ACASIA-NC_v0.10, uses gradient information to diagnose NC automatically. Automating the grading saves enormous manual effort and avoids subjectivity involved in traditional grading. Results indicate that the proposed features outperform the prior color- and intensity-based features for the grading task and reduce the error by more than 8.5%. This indicates utility of the gradient information for NC grading task.

The proposed features form a large pool of 1700 features. These features were fused with the prior ones, and relevant features were selected using the Fisher ratio test. It was found that the majority of the relevant features belong to the proposed ones. Moreover, fusion was observed to perform better than using either of the two alone. A feature level fusion was performed in this work. However, there is a scope of exploring other fusion techniques, such as late fusion.

Automatic grading depends a lot on accurate lens localization. In ACASIA-NC_v0.10, the localization was 95% accurate, which supports automatic grading. However, grading performance can be further improved by automatically removing the samples involving failed localization.

Proposed features are still sensitive to the data quality. Image enhancement using histogram equalization can also enhance image artifacts as shown in Fig. 7(b). The figure shows a case where image artifacts produce extraneous edges in the ROI. Image enhancement will further increase the artifacts. This problem can be avoided by enhancing the image based on the current properties of the image, such as intensity, contrast level, and image quality. This is a part of our future work.

Currently, experiments were performed with only 100 training samples since number of samples for some of the classes was very low. In order to have a balanced sampling from all the classes, the number of samples from each class was small. Future works will investigate the effects of using a larger training set with unbalanced sampling.

ACASIA-NC_v0.10 was found to perform unsatisfactorily for very low or very high grades of NC. This issue will also be investigated in future. Future work also includes exploring gradient information further to extract even better features.

References

1. D. Pascolini and S. Mariotti, "Global estimates of visual impairment: 2010," *Br. J. Ophthalmol.* **96**(5), 614–618 (2010).
2. P. Asbell et al., "Age-related cataract," *The Lancet* **365**(9459), 599–609 (2005).
3. L. Chylack et al., "The lens opacities classification system III," *Arch. Ophthalmol.* **111**, 1506–1506 (1993).
4. J. Sparrow et al., "The Oxford clinical cataract classification and grading system," *Int. Ophthalmol.* **9**(4), 207–225 (1986).
5. S. West et al., "Use of photographic techniques to grade nuclear cataracts," *Invest. Ophthalmol. Visual Sci.* **29**(1), 73–77 (1988).
6. P. Mitchell et al., "Prevalence of cataract in Australia: the Blue Mountains eye study," *Ophthalmology* **104**(4), 581 (1997).
7. B. Klein et al., "Assessment of cataracts from photographs in the beaver dam eye study," *Ophthalmology* **97**(11), 1428 (1990).
8. H. Li et al., "A computer-aided diagnosis system of nuclear cataract," *IEEE Trans. Biomed. Eng.* **57**(7), 1690–1698 (2010).
9. H. Li et al., "Image based grading of nuclear cataract by SVM regression," *SPIE Proc. Med. Imaging* **6915**, 691536–691536 (2008).
10. D. Duncan et al., "New objective classification system for nuclear opacification," *J. Opt. Soc. Am. A, Opt Image Sci.* **14**(6), 1197–1204 (1997).
11. D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. Seventh IEEE International Conference on Computer Vision*, IEEE, Kerkyra, Greece (1999).
12. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. Computer Vision and Pattern Recognition, CVPR, IEEE Computer Society Conference on*, Vol. 1, pp. 886–893, IEEE, San Diego, California (2005).
13. K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(10), 1615–1630 (2005).
14. H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *Proc. European Conference on Computer Vision, ECCV*, Springer Berlin Heidelberg, Graz, Austria (2006).
15. S. Winder, G. Hua, and M. Brown, "Picking the best daisy," in *Proc. Computer Vision and Pattern Recognition, CVPR, IEEE Computer Society Conference on* (2009).
16. S. Fan et al., "An automatic system for classification of nuclear sclerosis from slit-lamp photographs," in *Proc. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003*, pp. 592–601, Springer Berlin Heidelberg, Montreal, Canada (2003).
17. H. Li et al., "Towards automatic grading of nuclear cataract," in *Proc. Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 4961–4964, IEEE, Lyon, France (2007).
18. H. Li et al., "An automatic diagnosis system of nuclear cataract using slit-lamp images," in *Proc. Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pp. 3693–3696, IEEE, Minneapolis, Minnesota (2009).
19. W. Huang et al., "A computer assisted method for nuclear cataract grading from slit-lamp images using ranking," *IEEE Trans. Med. Imaging* **30**(1), 94–107 (2011).
20. S. Moore and R. Bowden, "The effects of pose on facial expression recognition," in *Proc. British Machine Vision Conference*, pp. 1–11, BMVA Press, London (2009).
21. A. Foong et al., "Rationale and methodology for a population-based study of eye diseases in Malay people: the Singapore Malay eye study (SiMES)," *Ophthalmol. Epidemiol.* **14**(1), 25–35 (2007).
22. T. Wong et al., "Prevalence and causes of low vision and blindness in an urban Malay population: the Singapore Malay eye study," *Arch. Ophthalmol.* **126**(8), 1091 (2008).
23. S. Weiss and N. Indurkha, *Predictive Data Mining: a Practical Guide*, Morgan Kaufmann, Burlington, Massachusetts (1998).
24. C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Trans. Intell. Sys. Technol.* **2**, 27:1–27:27 (2011).

Ruchir Srivastava is a research scientist in the Institute for Infocomm Research, Singapore. He received his PhD degree from National University of Singapore in 2012 and BTech from Indian Institute of Technology, Roorkee, India, in 2007. His research interests include computer vision, image processing, pattern recognition, and multimedia applications.

Xinting Gao received her BE degree from Northwestern Polytechnical University, China, and her PhD degree in EEE from Nanyang Technological University, Singapore. She is currently a research scientist at the Institute for Infocomm Research, A*STAR, Singapore. Her research interests include medical image processing, computer vision, and machine learning.

Fengshou Yin received his BS degree in electrical engineering and MS degree in computer engineering from National University of Singapore. He is currently a research engineer at the Institute for Infocomm Research, A*STAR, Singapore. His research interests include medical image processing and computer-aided diagnosis.

Damon W. K. Wong is a senior computer scientist and deputy program manager of the Ocular Imaging Programme at I2R under A*STAR. He pursued his PhD studies under the A*STAR Graduate Scholarship, and has research experience in both optics and image processing. He has been committed to medical imaging research and development in the past few years.

Jiang Liu is the manager of the ocular program at Institute for Infocomm Research. He has dedicated much of his effort to medical image-based CAD research in the past years. He had completed and is managing many research grants on ocular image processing.

Carol Y. Cheung is a research scientist with a focus on ocular imaging as a biological approach to study human diseases. Her research may improve the understanding of early disease changes, and may allow the prediction, early detection, and diagnosis of eye diseases via ocular imaging, leading to a better targeted and more effective screening.

Tien Yin Wong is a clinician scientist with dual training in ophthalmology and epidemiology. He is the principal investigator of the Singapore Epidemiology of Eye Disease (SEED) program. He directs a large multidisciplinary research program combining imaging, epidemiology, and clinical trials on retinal diseases and the use of ocular imaging technologies to predict eye and cardiovascular diseases.