Detection of retinal nerve fiber layer defects on retinal fundus images for early diagnosis of glaucoma

Chisako Muramatsu
Gifu University
Graduate School of Medicine
Department of Intelligent Image Information
1-1 Yanagido
Gifu 501-1194
Japan

Yoshinori Hayashi
TAK Co., Ltd.
4-35-12 Kono
Ogaki, Gifu 503-0803
Japan

Akira Sawada
Gifu University
Graduate School of Medicine
Department of Ophthalmology
1-1 Yanagido, Gifu 501-1194
Japan

Yuji Hatanaka
The University of Shiga Prefecture
School of Engineering
Department of Electronic Systems Engineering
2500 Hassakacho
Hikone, Shiga 522-8533
Japan

Takeshi Hara
Gifu University
Graduate school of Medicine
Department of Intelligent Image Information
1-1 Yanagido
Gifu 501-1194
Japan

Tetsuya Yamamoto
Gifu University
Graduate School of Medicine
Department of Ophthalmology
1-1 Yanagido
Gifu 501-1194
Japan

Hiroshi Fujita
Gifu University
Graduate School of Medicine
Department of Intelligent Image Information
1-1 Yanagido
Gifu 501-1194
Japan

Abstract. Retinal nerve fiber layer defect (NFLD) is a major sign of glaucoma, which is the second leading cause of blindness in the world. Early detection of NFLDs is critical for improved prognosis of this progressive, blinding disease. We have investigated a computerized scheme for detection of NFLDs on retinal fundus images. In this study, 162 images, including 81 images with 99 NFLDs, were used. After major blood vessels were removed, the images were transformed so that the curved paths of retinal nerves become approximately straight on the basis of ellipses, and the Gabor filters were applied for enhancement of NFLDs. Bandlike regions darker than the surrounding pixels were detected as candidates of NFLDs. For each candidate, image features were determined and the likelihood of a true NFLD was determined by using the linear discriminant analysis and an artificial neural network (ANN). The sensitivity for detecting the NFLDs was 91% at 1.0 false positive per image by using the ANN. The proposed computerized system for the detection of NFLDs can be useful to physicians in the diagnosis of glaucoma in a mass screening. © 2010 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3322388]

Keywords: retinal fundus images; glaucoma; retinal nerve fiber layer defects; computer-aided diagnosis.

Paper 09254R received Jun. 18, 2009; revised manuscript received Dec. 7, 2009; accepted for publication Dec. 18, 2009; published online Feb. 23, 2010.

1 Introduction

Glaucoma is the second leading cause of vision loss in the world.1 Because of its slow progressive nature, however, many patients stay unaware of it until the disease reaches to advanced stages. In a population-based prevalence survey of glaucoma in Tajimi, Japan (“Tajimi” Study), the prevalence of the primary open-angle glaucoma in the population 40 years or older was 3.9%, of which 93% were previously undiagnosed.2,3 It is important to detect glaucomatous changes as early as possible, thus minimizing the chance to cause sig-
significant visual disability including permanent loss of vision. Retinal examination by use of fundus photographs is simple, relatively inexpensive, and suited for a mass screening. However, ophthalmologists must review volumes of cases for detecting a number of ocular diseases in a limited time. Moreover, in the mass screening, physicians may not be specialized in glaucoma diagnosis. Therefore, for improving diagnostic efficiency in busy screening, our interest has been the development of a computerized system for detecting glaucomatous changes in retinal fundus images.\textsuperscript{4,5,6}

Two major signs of glaucomatous changes are enlarged cupping, or rim thinning, and retinal nerve fiber layer (RNFL) defect (NFLD). It is well known that NFLD development precedes an occurrence of morphologic change in optic nerve head and development of visual field defects in many cases, and is considered as one of the earliest glaucomatous changes.\textsuperscript{6–9} There have been studies about the analysis of RNFL using fundus photographs,\textsuperscript{9,10} scanning laser polarimetry,\textsuperscript{10–13} and optical coherence tomography\textsuperscript{13–17} for diagnosis of glaucoma. Although more studies were focused because of its routine use.

Muramatsu et al. attempted to quantify RNFL striations by comparing the variations in pixel values along and across the RNFL.\textsuperscript{18} Lee et al. introduced an automated method for quantifying nerve fiber layer based on the intensity profile around the optic disk.\textsuperscript{19} By taking the first derivative of the intensity plot, the edges of NFLDs can be found; however, the method was influenced by the presence of blood vessels. Song et al. suggested a computer-based glaucoma diagnosis system, a part of which includes image processing techniques for enhancement of NFLDs.\textsuperscript{20} Kolar and Vacha employed the texture analysis to distinguish RNFL and RNFL loss.\textsuperscript{21} The performance of these methods needs to be evaluated on larger databases. We investigated a computerized method for detection of NFLDs in fundus photographs\textsuperscript{18–23} can be important because of its routine use.

Peli et al. attempted to quantify RNFL striations by comparing the variations in pixel values along and across the RNFL.\textsuperscript{18} Lee et al. introduced an automated method for quantifying nerve fiber layer based on the intensity profile around the optic disk.\textsuperscript{19} By taking the first derivative of the intensity plot, the edges of NFLDs can be found; however, the method was influenced by the presence of blood vessels. Song et al. suggested a computer-based glaucoma diagnosis system, a part of which includes image processing techniques for enhancement of NFLDs.\textsuperscript{20} Kolar and Vacha employed the texture analysis to distinguish RNFL and RNFL loss.\textsuperscript{21} The performance of these methods needs to be evaluated on larger databases. We investigated a computerized method for detection of NFLDs in retinal fundus images. Our system was tested on a larger database and requires only manual identification of an optic disk center, which can be replaced by an automated scheme in the future. If regions of NFLDs can be automatically detected and presented to physicians, it may be useful for early diagnosis of glaucoma.

2 Methodology

2.1 Image Database

Color photographs of retinal fundus were obtained from the Eye Health Care Project in Tajimi, Japan,\textsuperscript{24} which includes clinical data and eye examination results, such as visual field test, intraocular pressure, and retinal fundus image; however, this database excludes the cases subjected for the Tajimi Glaucoma Epidemiological Study, or the Tajimi Study.\textsuperscript{2,3} The retinal fundus was photographed by using IMAGEitet digital fundus camera system (TRC-NW6S, Topcon, Tokyo, Japan) with an image resolution of 768 × 576 pixels, which was saved as a JPEG image.

One hundred forty-four images with reports indicating the presence of NFLD in left eyes by an expert ophthalmologist were reviewed independently by two other ophthalmologists specialized in glaucoma diagnosis, who marked the locations of NFLDs by two lines specifying the extent of defects. Ten images were excluded because of the poor image quality. On the basis of the visual field test using a frequency-doubling technique, 19 of 134 patients (14\%) had visual field defects in their left eyes. The average intraocular pressure in the left eyes was 14.3 mmHg. Figure 1 shows retinal fundus images with a NFLD marked by the two ophthalmologists. The regions marked by both readers were considered as the “gold standard” NFLDs in this study. For each image with at least one identified NFLD, an age and gender-matched normal image was randomly selected from the Tajimi Database. As the result, our database consisted of 81 images, including 99 NFLDs and 81 images without major eye diseases. For avoiding the inclusion of both eyes from the same patient, only images of left eyes were employed in this study, although our method can easily be applied to images of right eyes.

2.2 NFLD Detection Method

2.2.1 Removal of blood vessels

The visibility of NFLDs is known to improve by using red-free or green-channel images. In the images used in this study, the contrast of NFLDs was maximized using green channel, and not much information was added by blue channel; therefore, the detection of NFLDs was proceeded using green-channel images. However, the contrast of retinal blood vessels is also enhanced in the green-channel images. In order to exclude regions of the blood vessels from potential false positives (FPs), the major blood vessels were detected. The method for extracting the blood vessels is explained elsewhere.\textsuperscript{2} Briefly, the black-top-hat transformation was applied to the green channel of color retinal fundus images. The structure element used in this operation was a circle with the diameter of 11 pixels, which was determined empirically for the detection of the blood vessels near the optic disk. A p-tile thresholding method\textsuperscript{25} was employed on the resulted images for determination of the blood vessel regions.

The pixels identified as parts of the blood vessels were interpolated by the surrounding pixels for creating “blood-vessel-erased” images.\textsuperscript{4} For each blood-vessel pixel, the neighbor pixels corresponding to the retinal regions were searched in 16 directions and the pixel value was replaced by the weighted sum of those of the neighbor pixels with the weights inversely proportional to the distance between the two pixels. The original image and its blood-vessel-erased image are shown in Fig. 2.

2.2.2 Modified polar transformation of images

Localized NFLDs appear as dark curved bands or fans. Retinal nerves run almost straight from the center of fovea to the
optic disk and relatively straight from superior, inferior, and nasal retinal regions to the optic disk. On the other hand, temporal nerves above and below fovea are curved. For facilitating the detection of NFLDs in the subsequent process, the images were transformed so that the courses of the nerves, and thus NFLD regions, become relatively straight in the transformed images. It was based on the polar transform with the center of optic disk as an axis; however, instead of straight lines radiating from the axis, a set of elliptic lines were employed. The parameters of the ellipses were determined experimentally and varied with angle, $\theta$, as follows:

$$a = 175 + 2.5|\theta|$$

$$b = \frac{a|\theta|}{90},$$

where $a$ and $b$ are the major and minor semi-axes of the ellipses and $\theta$ is the angle with respect to a horizontal line connecting the optic disk and the temporal edge. All the elliptic lines were passing through the center of the optic disk, which was identified manually in this study, and therefore, the center coordinates of the ellipses were moved with respect to the parameter $a$. In the regions with angles of $<90$ deg or smaller than $-90$ deg (i.e., nasal side of the optic disk), the ellipses were rotated about the center of the optic disk. Some of the elliptic lines are shown in Fig. 3. Because NFLDs are most visible in the upper and lower temporal regions, a part of the nasal regions was not included in the transformed image.

Figure 4(a) shows the green-channel image of a transformed blood-vessel-erased image.

### 2.2.3 Background brightness correction

Because of photographic conditions, the brightness in an image may be uneven. In order to facilitate the detection of the NFLDs, the brightness of the images was corrected. The image brightness is mainly represented by the red channel of a color image. Therefore, we estimated the background brightness trend by use of the red-channel images, which were smoothed by pixel averaging to reduce an effect of noise. The estimated background image was inverted, and a fraction of the inverted background value was simply added to the corresponding pixel value in the green-channel image. In this study, the fraction was set to 50%, which was selected empirically. Subsequently, the image contrast was normalized by linear enhancement technique using the maximum and minimum pixel values in each image to handle images with different illumination. The brightness-corrected image is shown in Fig. 4(b).

### 2.2.4 Enhancement of NFLDs by using the Gabor filter

After the image transformation, the NFLD regions are likely to be stretched vertically. In order to enhance contrast of NFLDs, the Gabor filtering was applied. The Gabor function consists of a multiplication of the Gaussian function and cosine function. For enhancing the vertical edges, the orientation of the Gabor function was set at 0 deg. Without the image transformation, the filter orientation had to be changed at different locations in an image and the filter shape might need to be changed for detection of curved regions. Although our main target was to detect early, localized NFLDs, our database included the NFLDs with various widths ranging from approximately $5 \text{--} 70$ pixels. Therefore, three filters with dif-
different sizes and parameters were employed. The sizes of the three filters were 41 × 61, 41 × 61, and 81 × 61 pixels. The standard deviations of the Gaussian functions in \( x \) and \( y \) directions were set 6 and 15, 12 and 20, and 24 and 36 pixels, respectively, whereas the wavelengths of the cosine functions were varied according with the standard deviations as 20, 40, and 100 pixels, respectively. Figure 5 shows the filter kernels and the corresponding output images by convoluting the filters with the transformed images.

By use of three filtered images, dark bandlike regions were detected as initial candidates of NFLDs. If an average pixel value in a rectangular region was smaller than the average minus two times the standard deviation in the surrounding regions, then the center pixel of the rectangular region was considered as a candidate pixel. The height of the rectangular regions was determined as 15 pixels, and the width was varied according with the filter sizes as 5, 11, and 23 pixels. The surrounding regions were defined as two 3 × 15 regions that were two pixels away to the right and left of the rectangular region. The selected pixels were then expanded by a region growing method for determination of the initial candidate regions. The range of the region growing was limited within the size of each filter. If a region was <100 pixels in area or if the contrast of a region was <5 in pixel value, the region was excluded from the candidates.

The macular regions appear as a dark spot in the green-channel images and sometimes misdetected as NFLDs. For removing such FP, approximate locations of fovea were detected. In this study, a relatively simple method was employed; a circular region with the smallest average pixel value in the red and green-channels was determined as the fovea. The size of the circular region was predefined as 41 pixels in diameter. If a NFLD candidate was overlapped with the detected fovea region, it was considered as a FP and removed from the initial candidates.

2.2.5 Feature determination and FP reduction

For each candidate, six image features were determined. Six features included the area, the vertical length, the average pixel values in the original and filtered images, and the contrasts in the original and filtered images. The contrast is defined as the difference in the average pixel values in a candidate region and its surrounding background region. The surrounding background region included the pixels with the distances of 3–6 pixels away from the outer edge of the candidate region.

With these six features, the likelihood scores of the candidates being true NFLDs were determined by using two classifiers: a linear discriminant analysis (LDA)\(^{28}\) and an artificial neural network (ANN).\(^{29}\) For ANN, three-layered feed-forward network with a backpropagation algorithm and sigmoidal output function was employed. Because of the limited number of cases, a leave-one-out test method was employed. In the leave-one-out method, the candidates corresponding to the same case was removed from the training samples at once. The candidates detected by using the three filters were entered to the classifiers independently. For the final results, the remaining candidates by the three filters were combined by OR operation (i.e., taking union of the candidate regions). The results of the detection methods were evaluated by use of the free-response receiver operating characteristic (FROC) analysis.

3 Results

By applying the Gabor filters, the NFLD regions were enhanced as shown in Fig. 5; however, at the same time, many FPs were also detected. The average number of initial candidates was 11.8 per case, of which 82% were detected using the filter in Fig. 5(a). In this study, if the region of a candidate was overlapped with that of the gold standard, it was considered as a true-positive detection. At the initial detection, one NFLD was not able to be detected (the sensitivity of 99%).

The initially detected candidates were classified into final candidates of NFLDs and FPs by use of the LDA and ANN. Because there were high correlation (Pearson’s correlation coefficient, \( r=0.99 \)) between the average pixel values and relatively high correlation \( (r=0.68) \) between the contrasts determined with the original images and filtered images, the number of image features used in the classifiers was varied from four to six. For the ANN, the number of hidden units was varied from two to five according with the number of input units.

Figure 6 shows the FROC curves determined by the results of the two methods with four and six features. Although the curves were almost overlapped, the results by using six features were slightly better than those by using four features for both LDA and ANN based on the jackknife FROC analysis\(^ {30} \) \((p=0.004 \) for the LDA and \( p=0.003 \) for the ANN). In addition, with six features, the figure-of-merit by using ANN was slightly higher than that by using LDA: 0.827 and 0.794 for the ANN and LDA, respectively \((p=0.02)\).
In 162 cases, including 81 cases with 99 NFLDs, the sensitivity of the proposed method was relatively high, with both the lesion and case sensitivities of 91% using the ANN when the FP per image was 1.0. Many of the FPs were due to the retinal blood vessels and could be easily identified as FPs by ophthalmologists. We attempted to identify and remove the blood vessels by the image interpolation. However, some of them were remained and detected as the candidates of NFLDs. Improvement in the blood vessel detection method is work in progress; when the new method is incorporated, the number of FPs may be reduced.

The results of this study seem promising; however, there are some limitations in this study. First, because of the limited number of cases, the proposed method was evaluated by use of the leave-one-out test method, and some of the threshold values were determined empirically. The method must be evaluated with the independent database in the future. Second, in this study, the centers of the optic disks were identified manually. Although the manual identification of the optic disk...
is a simple task, a fully automated system would be desired. In addition, it was assumed that the center of the fovea was located directly horizontal to the optic disk for the ellipse-based image transformation purpose, although it is commonly found slightly lower than the optic disk. In most of cases, the identification of the optic disk and fovea would not be very difficult; however, determination of the exact centers could be difficult. Although, it is expected that the moderate mislocation of the centers would not affect the result severely, the result must be validated with the fully automated system.

In this study, the gold standard was defined as the NFLDs that were identified by both ophthalmologists. As the result, the NFLDs identified in this study may include relatively obvious to moderately difficult defects. One of the ophthalmologists (ophthalmologist A) generally identified more NFLDs than the other (ophthalmologist B). Out of 134 images reviewed, ophthalmologist A identified 173 NFLDs, whereas ophthalmologist B identified 104 NFLDs. In the 81 cases used in this study, there were 16 and 4 regions identified only by ophthalmologists A and B, respectively. In Fig. 7(c), it can be seen that one of the FPs detected by the proposed method was identified only by ophthalmologist A. At the FP of 1.0 per case, nine “false” candidates detected by the method were such regions in which eight of them identified only by ophthalmologist A and one by ophthalmologist B. Some of these regions could be considered as suspicious regions, and the patients might be followed.

In addition, the determination of the extent of NFLD on the fundus images is very difficult. Therefore, some NFLDs, although identified by both ophthalmologists, span different ranges. Figure 8 shows the cases with NFLDs that one of the ophthalmologists identified a part of the regions identified by the other. Moreover, some FPs detected by our method were extremely difficult to be distinguished from the gold standard NFLDs as shown in Fig. 9. In future study, it may be desirable to include a third reader or have a consensus reading.

Finally, although three sizes of Gabor filters were employed, it was more difficult to detect wide defects. For some NFLDs, parts of the regions were detected by the proposed method and they were counted as the true detection in this study. The purpose of this study was to develop a system that draws physicians’ attention to the suspicious NFLD regions. Identification of only a part of the defect may be adequate. Therefore, if the candidates were overlapped with the gold standard, then they were considered as the true-positive detection. However, the effect of partial detections and usefulness of the computerized detection system must be evaluated by ophthalmologists in the observer performance study. Although our main target was the detection of localized, early defects, it may be desirable to improve sensitivity for the semidiffused defects. For detection of relatively wide defects, different filters or a new detection method may be needed.

5 Conclusion

An automated method for detection of NFLDs in retinal fundus photographs was investigated. The proposed methods achieved relatively high sensitivities with a small number of FP per case. Such computerized system may be useful for the early diagnosis of glaucoma in the screening.

Acknowledgments

This work was partly supported by a grant for the Knowledge Cluster Creation Project from the Ministry of Education, Culture, Sports, Science and Technology, Japan. The authors are grateful to A. Aoyama for his contribution.

References

5. Y. Hayashi, T. Nakagawa, Y. Hatanaka, A. Aoyama, M. Kakogawa,


