Automated determination of cup-to-disc ratio for classification of glaucomatous and normal eyes on stereo retinal fundus images

Chisako Muramatsu
Toshiaki Nakagawa
Akira Sawada
Yuji Hatanaka
Tetsuya Yamamoto
Hiroshi Fujita
Automated determination of cup-to-disc ratio for classification of glaucomatous and normal eyes on stereo retinal fundus images

Chisako Muramatsu, a Toshiaki Nakagawa, b Akira Sawada, c Yuji Hatanaka, d Tetsuya Yamamoto, e and Hiroshi Fujita a

aGifu University, Graduate School of Medicine, Department of Intelligent Image Information 1-1 Yanagido, Gifu 501-1194, Japan
bKowa Company, Ltd., Electronics and Optics Division, Research & Development Section, Hamamatsu Factory, 1-3-1 Shin-Miyakoda, Kita-ku, Hamamatsu-shi, Shizuoka 431-2103, Japan
cGifu University, Graduate School of Medicine, Department of Ophthalmology, 1-1 Yanagido, Gifu 501-1194, Japan
dUniversity of Shiga Prefecture, School of Engineering, Department of Electronic System Engineering, 2500 Hassakacho, Hikone, Shiga 522-8533, Japan

Abstract. Early diagnosis of glaucoma, which is the second leading cause of blindness in the world, can halt or slow the progression of the disease. We propose an automated method for analyzing the optic disc and measuring the cup-to-disc ratio (CDR) on stereo retinal fundus images to improve ophthalmologists' diagnostic efficiency and potentially reduce the variation on the CDR measurement. The method was developed using 80 retinal fundus image pairs, including 25 glaucomatous, and 55 nonglaucomatous eyes, obtained at our institution. A disc region was segmented using the active contour method with the brightness and edge information. The segmentation of a cup region was performed using a depth map of the optic disc, which was reconstructed on the basis of the stereo disparity. The CDRs were measured and compared with those determined using the manual segmentation results by an expert ophthalmologist. The method was applied to a new database which consisted of 98 stereo image pairs including 60 and 30 pairs with and without signs of glaucoma, respectively. Using the CDRs, an area under the receiver operating characteristic curve of 0.90 was obtained for classification of the glaucomatous and nonglaucomatous eyes. The result indicates potential usefulness of the automated determination of CDRs for the diagnosis of glaucoma. © 2011 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.3622755]

Keywords: stereo retinal fundus images; glaucoma; cup-to-disc ratio; computer-aided diagnosis.

Paper 11035RR received Jan. 20, 2011; accepted for publication Jul. 18, 2011; published online Sep. 2, 2011.

1 Introduction

Glaucoma is the second leading cause of blindness in the world, and its incidence rate is expected to increase because of the aging population, affecting approximately 80 million people in 2020. Once loss of a visual field occurs, it cannot be regained, therefore early diagnosis and treatment are key to minimizing the chance of significant visual impairment. However, because of the slow progressive nature of glaucoma, patients remain unaware of the visual disturbance until the disease reaches advanced stages. In a population-based prevalence survey of glaucoma in Tajimi, Japan, it was found that 93% of patients diagnosed with primary open-angle glaucoma (POAG) were unaware of their disease, i.e., previously undiagnosed. On the basis of the community-based descriptive study involving 295 patients, whom a majority (96%) were white, with the mean follow up of 15 years, the cumulative probability of glaucoma patients’ becoming blind in at least one eye at 20 years is estimated to be 27%. Although an elevated intraocular pressure (IOP) is a major risk factor for most glaucoma patients, its measurement fluctuates, and there is variation in the normal ranges between patients. In the Tajimi study, 92% of POAG patients had an IOP within the normal range. For diagnosis of glaucoma, the results from several different tests are considered, in addition to the direct examination by ophthalmologists, which may include fundus photography, Heidelberg retinal tomography (HRT), scanning laser ophthalmoscopy, and optical coherence tomography. One of the glaucoma-associated findings is the deformation of the optic nerve head (ONH), i.e., enlargement of the depression (called cupping), or rim thinning caused by the decrease of retinal nerve fibers. Although it is not yet in widespread use, a stereo fundus camera can be an effective tool for examining the three-dimensional (3D) structure of the ONH. Without the stereoscopic view, ophthalmologists must infer the cup deformation on the basis of the vessel directions (bending points) and faint color change inside the disc. One of the advantages of the stereo fundus photograph is that it can be interchangeably used with regular fundus photographs, which are frequently obtained for the purpose of diagnostic record, screening, and diagnosis of glaucoma and other eye diseases such as diabetic retinopathy.

The cup-to-disc ratio (CDR), defined as the ratio of the diameters of the cup and the optic disc, is one of the indices for evaluation of cupping. It may be easier to evaluate CDR on the stereo images than on conventional fundus images. However, manual quantification of CDR is time-consuming and prone to intra- and inter-reader variations. Several research groups have investigated automated and semi-automated methods to analyze the ONH for diagnosis of glaucoma on stereo and regular fundus images.
Nayak et al. investigated an automated method for diagnosis of glaucoma on regular fundus images by extracting three features. These features include the CDR and two features characterizing a shift of retinal vessels due to cupping. Their system is trained with 46 images and tested on 15 images. The effectiveness of these features needs to be evaluated on a larger database. Wong et al. proposed an improved scheme in determining the CDR on retinal fundus images. They developed an automated segmentation method for the disc and two segmentation methods for the cup on the basis of a level set algorithm and a histogram analysis. It was suggested that no single combination works best for all cases. CDRs are determined with mean errors of 0.11 for 40 test cases by using a neural network and a support vector machine for combining these results.

Corona et al. investigated an automated method for generating a 3D map of an optic disc by using stereo image pairs. Disparities of the stereo pair are found by the cross-correlation of edge-enhanced images for creating the 3D map. Using the 3D map, the cup and disc are segmented semi-automatically, and the length and volumes are measured. Glaucoma progression was evaluated in three patients, and the results agreed well with manual evaluations. Abramoff et al. proposed an automated cup and disc segmentation method on stereo fundus images by a pixel classification method. The k-nearest neighbor algorithm with 12 features is employed for the classification of pixels into cup, rim, or background. The CDRs are determined by taking the ratio of the numbers of pixels classified as the cup to cup-and-rim, which correlate well with the reference standard on 58 cases of glaucoma or suspected glaucoma. Another method for automated assessment of ONH using stereo fundus images was proposed by Xu et al. In their method, the disparity is determined by two matching methods, i.e., cross-correlation and minimum feature difference, for 3D reconstruction of the ONH. A deformable model is employed for segmenting the optic disc, and the cup margin is located at a certain distance posterior to the disc margin. Although automated measurement of the disc area agreed well with manual measurements on 54 test cases, correlations for the rim area and the CDR were moderate because of a strong dependence of cup segmentation on the results of disc segmentation.

In our previous study, an automated method for reconstructing a depth map by using a stereo fundus image pair was investigated. In this study, the cup region is automatically segmented using the depth map. With an automatic disc segmentation method developed previously the CDRs are determined and compared with those determined on the basis of manual segmentations by an ophthalmologist. The method is applied to a new test database, and the usefulness of the CDR for classification of glaucomatous and nonglaucomatous cases is evaluated.

2 Image Data

2.1 Training Dataset

The initial database consists of 80 stereo retinal fundus image pairs, which are used for developing our automated scheme. The images were obtained from volunteers and from patients who visited the ophthalmology department at Gifu University Hospital, Japan from March 2006 to April 2007 using stereo retinal fundus cameras (prototype of the WX-1, Kowa Company, Ltd., Tokyo, Japan). The images were saved in JPEG format with a 1600×1200 pixel matrix. They were captured with an optic angle of 27 deg and the majority were approximately centered at the ONH; in 4 cases, the ONH was off-centered about 1 to 1.5 disc diameter. An ophthalmologist (ophthalmologist A) who specializes in glaucoma diagnosis reviewed all 80 cases and provided manual outlines of the cup and disc on a stereo display. The 80 cases include 25 cases with a glaucomatous optic disc and 55 cases without signs of glaucoma assessed by the ophthalmologist. For a subset of 44 cases including 24 glaucomatous and 20 nonglaucomatous eyes, two other ophthalmologists, who also specialize in glaucoma diagnosis, provided the outlines of the cup and disc for evaluation of inter-reader variability.

2.2 Test Dataset

For evaluating the usefulness of the automated measurement of the CDR in distinguishing between glaucomatous and normal eyes, additional image data were collected. Images were obtained from patients who visited Gifu University Hospital from May 2007 to June 2008 by using the same stereo camera. For each patient, a pair of images with one eye of better image quality was selected. The diagnoses and outlines of cup and disc were provided by ophthalmologist A. This test dataset consists of 98 cases, including 60 images with signs of glaucoma and 38 images without signs.

3 Methodology

The automated scheme for determination of the CDR consists of segmentation of the disc by using the canny edge detector and the active contour method, reconstruction of a depth map based on the stereo disparity, segmentation of cup by polar boundary searching, and measurement of the CDR. The flowchart of the overall scheme is shown in Fig. 1.

3.1 Determination of Disc Outline

The automated schemes for the determination of the disc boundary were investigated and described in detail in our previous study. A brief description is provided here. In this study, we
employed the active contour model, namely Snakes. Only the first image (left image) of a pair was used for disc segmentation because the image quality is usually better in the first image. First, an approximate location of the disc was detected by a p-
tile thresholding method applied on the red channel of RGB images. The Canny edge detector was applied to a region of interest (ROI) of 600x600 pixels in size extracted about the detected location. Using the edge information, the disc outline was initialized and adjusted by using the Snakes algorithm. Finally, the outline was fitted by an ellipse for potentially reducing errors on cases with unclear margins and peripapillary chorioretinal atrophy (PPA).

3.2 Reconstruction of Depth Map
A detailed description for stereo disparity determination can be found elsewhere. The process of creating the depth map is shown in Fig. 2. First, the right and left images are registered to remove disparity due to patients’ motion and to extract ROIs from the right images [Fig. 2(a)]. The corresponding ROI is found by cross-correlation of the images calculated without the disc region. The disc region is excluded because the “real” disparity due to depth exists in this region. Subsequently, at every four pixels in horizontal and vertical directions of the left image, matching points on the right image are identified on the basis of the highest cross-correlation coefficients in regions of 21x21 pixels [Fig. 2(b)]. The cross-correlation is calculated in three color channels if the contrast, which is defined as the difference between the highest and lowest pixel values in a 21x21 pixel region, is larger than 10. The disparities found in the three channels are averaged if the cross-correlation coefficients are larger than 0.5. A search for the corresponding point is performed within the range of 41x23 pixels, i.e., +15 and -5 pixels in the horizontal direction and ±1 pixel in the vertical direction. If the corresponding point is not found because of the low contrast or low cross-correlation coefficient, the disparity is interpolated by the surrounding points. The depth image is created on the basis of the disparity and smoothed using the median filter and averaging filter for reducing the noise.

3.3 Determination of Cup Outline
In our previous study, a median filter of 5x5 pixels and an averaging filter of 3x3 pixels were used for noise reduction on the depth map on the basis of a comparison with the HRT topographic images. In this study, larger smoothing kernels of 11x11 and 5x5 pixels for the median and averaging filters, respectively, are employed for determining smooth contours. The median and averaging filters are applied three and two times, respectively. The contrast of the depth maps is linearly enhanced between the deepest and shallowest points. Figure 3 shows a left image, the HRT topographic image of the same eye, and the depth maps before and after the noise reduction.

The cup margin is determined by searching for edges in radial directions on the depth maps. The algorithm starts by determining the center of the cup. First, a binary image is created with a threshold of depth at 0.9 of the deepest depth inside the disc region. The centroid of the binarized region, which includes the deepest point, is determined as the center of the cup. By setting the center pixel as an origin, the depth map is

![Fig. 2 Process of reconstructing the depth map. (a) Identification of corresponding ROI position by scanning the left image. Disc region (black circle) was disregarded in order to account for disparity due to motion and not for real disparity due to depth. (b) Determination of disparity by finding the corresponding points at every four pixels (white dots) based on the correlation coefficient for 21x21 pixel ROIs (squares) within the searching range (dotted rectangle).](https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics)

![Fig. 3 Depth maps of the optic nerve head. (a) Original left image, (b) topographic image obtained by Heidelberg retina tomograph, (c) depth map reconstructed on the basis of the stereo disparity without noise reduction, and (d) depth map after the noise reduction process.](https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics)
transformed to a polar coordinate system, as shown in Fig. 4(b), where the horizontal axis corresponds to the radial angles with an increment of about 0.45 deg and the vertical axis corresponds to the distance from the origin up to 255 pixels. Note that the pixel values are inverted in Fig. 4 so that the deeper the points, the whiter the pixels. The pixels outside the disc region are disregarded (black pixels in Fig. 4).

A radial edge image is produced simply by subtracting the pixel values at a distance of 10 pixels away in the vertical (radial) direction, as shown in Fig. 4(c). In each column, the position of the maximum radial gradient point with the edge strength above a prespecified value, if it exists, is searched and considered as the cup boundary. The threshold edge strength is experimentally set at 50. For columns where a boundary is not found, it is interpolated using 5 columns from the first boundary found in each direction. For interpolation, the average of the radial distances weighted by the edge strengths is employed to account for more reliable boundary points. Finally, the boundary is smoothed by taking the running average. Figures 4(d)–4(f) show the initial boundary points, the boundary after the interpolation, and the smoothed boundary, respectively. The cup boundary determined in the polar coordinate is translated back to the Cartesian coordinate system, as shown in Fig. 4(g).

3.4 Measurement of CDR

In early stages of glaucoma, thinning of the rim often appears in the upper or lower parts of an optic disc. Therefore, ophthalmologists usually evaluate the CDR in a vertical direction. In this study, the CDR is defined as a ratio of vertical distances between pixels at the highest and lowest vertical positions inside the cup and disc regions as shown in Fig. 5. The CDRs are measured in the same manner using the outlines determined by the ophthalmologists for the training dataset (“reference standard”) and by the automated scheme for the training and test datasets. Therefore, the reference standard CDRs are not directly measured by the ophthalmologists, but rather are indirectly determined from the outlines. The diagnostic ability of the CDR is tested for classifying the glaucomatous and nonglaucomatous eyes. The classification performance is evaluated using the receiver operating characteristic (ROC) analysis by using the software (PROPROC, The University of Chicago).

4 Result

4.1 Inter-reader Variability on Cup and Disc Determination

For 44 cases in the training set, manual outlines of the cup and disc were provided by three ophthalmologists. The agreement of the cup and disc regions was evaluated by an overlap measure, which is defined as a ratio of the area of intersection to the area of union. Table 1 summarizes the inter-reader variability. The agreements for the disc regions between the three ophthalmologists were high and equivalent for the glaucomatous and nonglaucomatous cases. On the other hand, the agreements for the cup regions were lower, especially for the nonglaucomatous cases.

<table>
<thead>
<tr>
<th>Ophthalmologist</th>
<th>A vs B</th>
<th>B vs C</th>
<th>A vs C</th>
<th>A vs B</th>
<th>B vs C</th>
<th>A vs C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glaucoma</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
<td>0.73</td>
<td>0.84</td>
<td>0.71</td>
</tr>
<tr>
<td>Non-glaucoma</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
<td>0.52</td>
<td>0.76</td>
<td>0.48</td>
</tr>
<tr>
<td>All</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
<td>0.63</td>
<td>0.80</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Table 2. Average differences in vertical cup-to-disc ratios (CDRs) by three ophthalmologists.

<table>
<thead>
<tr>
<th></th>
<th>CDR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ophthalmologist A vs B</td>
</tr>
<tr>
<td>Glaucoma</td>
<td>0.09</td>
</tr>
<tr>
<td>Non-glaucoma</td>
<td>0.20</td>
</tr>
<tr>
<td>All</td>
<td>0.14</td>
</tr>
</tbody>
</table>

The average differences in the CDR were determined using the three ophthalmologists’ outlines which are shown in Table 2. The same trend was observed in these results. The Pearson’s correlation coefficients between the CDRs determined by ophthalmologist A and B, B and C, and A and C are 0.90, 0.83, and 0.87, respectively.

4.2 Results of Cup and Disc Segmentation on the Training Dataset

The results for the automatic segmentation of the cup and disc regions on the 80 training cases are summarized in terms of the average overlap measures and errors in CDRs with respect to the reference standard in Table 3. The average overlap measure for the disc regions was slightly lower than those between the three ophthalmologists, whereas the average overlap measure for the cup regions was comparable to those between ophthalmologists A and B, and A and C. Note that while the number of cases for these results are different, the similar results were obtained for the subset of 44 cases. Figure 6 shows the cup and disc outlines provided by the three ophthalmologists and the automated scheme for the glaucomatous and nonglaucomatous eyes. The correlation coefficient between the CDRs determined by ophthalmologist A and the automated scheme was 0.74. The results of automated and semi-automated methods by the other groups in terms of correlation coefficients between manual CDRs and computer-determined CDRs ranged from 0.67 and 0.99 however, the results cannot be directly compared because the cases used are different, and the difficulties of the cases are unknown. Note that in one of the studies, the correlation coefficient was determined for each patient for the three longitudinal changes, which could apparently yield a higher correlation. When the CDRs are inputted, the software calculates the ROC curve by varying the threshold on the likelihood ratio, which is the function of the CDR. The threshold is gradually changed to create smooth curves, and the areas under the ROC curve (AUC) for distinction between glaucomatous and nonglaucomatous eyes was determined, which was 0.91 for the training dataset.

4.3 Classification between Glaucomatous and Nonglaucomatous Eyes

Automated segmentation techniques were applied to the test dataset. Because the reference standard for the cup and disc outlines in the test dataset does not exist, the segmentation accuracy cannot be evaluated. Figure 7 shows the box plot of the automatically determined CDRs for the glaucomatous and nonglaucomatous eyes. The result indicates the CDRs for most glaucomatous cases are higher than those for nonglaucoma cases. Using the CDRs, an AUC of 0.90 was obtained. With a threshold of 0.7 according to the Japanese guideline, a sensitivity of 87% and a specificity of 82% were obtained.

5 Discussion

In this study, an automated scheme for measuring the CDR on stereo retinal fundus images and its ability to discriminate eyes with and without glaucoma-associated signs were investigated. For the measurement of the CDR, the cup and disc regions were automatically segmented using a depth map based on stereo disparity, and the brightness and edge information on the plain photographs. The manual outlines of the cup and disc regions were provided by an expert ophthalmologist for...
the training dataset, which were used as the reference standard for tuning the automated scheme. For the subset of 44 cases, outlines were also provided by two other experts. There were good agreements for the disc regions determined by the three ophthalmologists, while the agreements for the cup regions were not as high. When manual outlines were obtained, there was no specific guideline as to where the outlines should be placed. Therefore, one could delineate the bottom of the cup, while another could select the entrance of the cup as the margin. It may have caused the observed differences especially for cases with “sloping cup” whose depth changes gradually. This is a probable reason for the lower agreements for the nonglaucomatous cases than for the glaucomatous cases. The results summarized in Tables I and II suggest that the judgment by ophthalmologist A is slightly different from those of the other two ophthalmologists. However, the inter-reader correlation coefficients on the CDRs determined by ophthalmologists A and B and by ophthalmologists A and C were higher than that by ophthalmologists B and C, indicating consistent diagnostic assessments between the ophthalmologists.

For the 80 training cases, good agreement for the disc outlines determined by ophthalmologist A and the automated scheme was obtained. For some cases with PPA, no clear edge was detected at the disc margin. A relatively high agreement for the disc regions determined by the ophthalmologist and the automated scheme was obtained, whereas the agreement for the cup regions was moderate, but within the range of the inter-reader agreement. The successes in reconstructing the depth maps and segmenting the disc region partly account for the success in the cup segmentation. When there is no apparent landmark for matching ROIs because of saturation of pixel values or an absence of noticeable blood vessels, an accurate disparity measurement becomes difficult. Unfortunately, in the absence of reference depth data such as HRT data, quantitative evaluation of the depth map reconstruction was not possible. The over-extraction of PPA in the disc segmentation often causes over-extraction in the cup segmentation. In the future, detection of the PPA region should be considered for improving segmentation accuracy. For cases with a small cup and/or disc, a small shift of segmented regions with respect to the reference standard can strongly degrade the overlap measure, and a small difference in diameters can result in a large error in the CDR. This is one of the reasons that the average overlap measure was lower for the nonglaucoma cases.

The proposed method was applied to the test cases, and the ability of automatically-determined CDRs to distinguish glaucomatous and nonglaucomatous eyes was evaluated. A high AUC of 0.90 was obtained, indicating the potential utility of computer analysis in screening glaucoma when a glaucoma specialist is absent. Computer analysis may increase the diagnostic efficiency and improve intra- and inter-reader consistency for follow-up monitoring by providing the quantitative data. However, the actual practicality and diagnostic utility of the proposed method must be assessed by the observer performance study and clinical study. The comparable AUCs obtained for the training and test cases indicate that consistent results can be expected in segmenting cup and disc regions in images with similar quality. Note that distinction of glaucomatous and nonglaucomatous eyes may be more difficult for the test cases than for the training cases, because the training cases include images acquired from the student volunteers. One limitation of this study is that one ophthalmologist determined the diagnosis based only on the image findings, rather than on more concrete data such as those from visual field tests. However, the purpose of the computer analysis is not an automatic diagnosis, but to provide quantitative data or to suggest a possible disease condition to doctors for cases in which an expert ophthalmologist would be alerted to. In the future, an acquisition of definite diagnosis or diagnosis by multiple experts is desirable.

6 Conclusion
An automated analysis of ONH on stereo fundus images was investigated. The disc region was segmented on the basis of the image brightness and edge information. The segmentation of the cup region was performed using the depth map, which was reconstructed using the disparity in the stereo image pairs. A relatively high agreement for the disc regions determined by the ophthalmologist and the automated scheme was obtained, whereas the agreement for the cup regions was moderate, but within the range of the inter-reader agreement. When the method was applied to the test dataset, most images of glaucomatous eyes could be correctly distinguished from those without signs of glaucoma using the CDR. The automated determination of the CDR can be useful for improving the diagnostic efficiency by providing the quantitative data especially to doctors with limited experience in the diagnosis of glaucoma.

References
19. Available at: [http://metz-roc.uchicago.edu/MetzROI/software](http://metz-roc.uchicago.edu/MetzROI/software)