

Quantitative depth analysis of optic nerve head using stereo retinal fundus image pair

Toshiaki Nakagawa*

Gifu University
Graduate School of Medicine
1-1 Yanagido, Gifu
Gifu 501-1194, Japan

Takayoshi Suzuki

Kowa Company, Ltd.
1-3-1 Shinmiyakoda, Kitu-ku, Hamamatsu
Shizuoka 431-2103, Japan

Yoshinori Hayashi

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

Yutaka Mizukusa

Kowa Company, Ltd.
3-3-1 Chofugaoka, Chofu
Tokyo 182-0021, Japan

Yuji Hatanaka

The University of Shiga Prefecture
2500 Hassaka-cho, Hikone
Shiga 522-8533, Japan

Kyoko Ishida

Takeshi Hara

Hiroshi Fujita

Tetsuya Yamamoto

Gifu University
Graduate School of Medicine
1-1 Yanagido, Gifu
Gifu 501-1194, Japan

1 Introduction

Glaucoma is the second leading cause of vision loss in the world.¹ Moreover, the number of people with glaucoma is estimated to be 60.5 million in 2010 and 79.6 million in 2020.² In a population-based prevalence survey of glaucoma in Tajimi City, Japan, one in 20 people aged over 40 was found to have the disease.^{3,4} Glaucoma is a group of eye diseases causing optic nerve damage; the exact causes of this damage are not fully understood, but involve mechanical compression and/or decreased blood flow to the optic nerve.⁵ Although incurable, glaucoma can be treated if diagnosed early. Mass screening of glaucoma using retinal fundus images is simple and effective.⁶

Abstract. Depth analysis of the optic nerve head (ONH) in the retinal fundus is important for the early detection of glaucoma. In this study, we investigate an automatic reconstruction method for the quantitative depth measurement of the ONH from a stereo retinal fundus image pair. We propose a technique to obtain the depth value from the stereo retinal fundus image pair, which mainly consists of five steps: 1. cutout of the ONH region from the stereo retinal fundus image pair, 2. registration of the stereo image pair, 3. disparity measurement, 4. noise reduction, and 5. quantitative depth calculation. Depth measurements of 12 normal eyes are performed using the stereo fundus camera and the Heidelberg Retina Tomograph (HRT), which is a confocal laser-scanning microscope. The depth values of the ONH obtained from the stereo retinal fundus image pair were in good accordance with the value obtained using HRT ($r=0.80\pm 0.15$). These results indicate that our proposed method could be a useful and easy-to-handle tool for assessing the cup depth of the ONH in routine diagnosis as well as in glaucoma screening. © 2008 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3041711]

Keywords: retinal fundus image; stereo image; depth measurement; optic nerve head; glaucoma.

Paper 08348 received Oct. 1, 2008; accepted for publication Oct. 8, 2008; published online Dec. 15, 2008. This paper is a revision of a paper presented at the SPIE conference on Medical Imaging 2007: Physiology, Function, and Structure from Medical Images, February 2007, San Diego, California. The paper presented there appears (unrefereed) in SPIE Proceedings Vol. 6511.

The cup/disk (C/D) ratio, which is the ratio of the diameter of the depression (cup) to that of the optic nerve head (ONH, disk), is one of the important parameters for an early diagnosis of glaucoma. The C/D ratio is generally used in clinical practice, because its value is greater in the case of glaucoma. The interpretation of the ONH (which actually has a 3-D structure) by using a 2-D image is subjective, and there is a wide variation between the examinations of the ONH by different observers and even between the examinations by the same observer.⁷ A more quantitative alternative is to use the Heidelberg Retina Tomograph (HRT), which is a confocal laser scanning microscope, for the acquisition and analysis of the 3-D measures of the ONH.^{8,9} It has been revealed that an HRT is capable of ONH imaging, and it is an established technique for detecting glaucomatous structural changes.

Address all correspondence to: Toshiaki Nakagawa, Graduate School of Medicine, Gifu University, 1-1 Yanagido, Gifu, Gifu 501-1194, Japan. Tel: 81-58-230-6519; E-mail: nakagawa@jti.info.gifu-u.ac.jp

*Current address: Kowa Company, Ltd., 3-3-1 Chofugaoka, Chofu, Tokyo 182-0021, Japan. Tel: 81-42-483-4128; E-mail: tsnakagw@kowa.co.jp

A computerized technique for the qualitative estimation of the depth of the ONH from the stereoscopic pairs of retinal fundus images has been suggested as another objective method for the 3-D analysis of the depression of the ONH.¹⁰⁻¹³ It has been shown that this technique is useful for the investigation of the 3-D measures of the ONH. Corona et al.¹² calculated 2-D and 3-D C/D ratios, and the results showed good correlation between clinical measures and computer-generated measures. Xu and Chutatape¹³ described an automatic reconstruction method from a stereo image pair by the use of a new approach that included sparse disparity measurement, subpixel modification, searching range autoadjustment, and piece-wise cubic interpolation and smoothing operations. The subpixel modification performed reconstruction from a low-resolution stereo retinal fundus image pair. Xu et al.¹⁴ also compared the C/D vertical ratio generated from the stereo image pair with the results from the HRT and an experienced ophthalmologist for evaluating the reconstruction results. The correlations of the results with the ophthalmologist's measurement were 0.71 and 0.67, respectively, for the proposed method and the HRT. However, the relative depth value was expressed as the number of pixels in these techniques, and the experimental results regarding the quantitative depth value calculated from the stereo image pair of the ONH have not been reported. Although Xu et al.¹⁴ reported the real depths, the values were calibrated by using the maximum cup depth obtained from the HRT results. Moreover, there have been no studies in which the depth value calculated from the stereo image pair has been compared with the HRT outputs.

In this study, an automatic method for reconstructing the 3-D structure of the ONH by using the stereo retinal fundus image pair has been proposed. To evaluate the accuracy of our method, the shape of the depression of a disk depth model, which had a circular dent when generated from the stereo image pair and which was used to model the ONH, was compared with its true depth value. The true depth value was measured by use of a vernier caliper with accuracy level of ± 0.01 mm. Moreover, the depth value of the ONH obtained from the stereo image pair was compared with the HRT measurement results.

2 Methodology

In our technique, the depth value is obtained from the stereo retinal fundus image pair, which is a digital color image; the technique mainly consists of five processes. The flowchart of the main procedure is shown in Fig. 1. A stereo image pair consists of a left image and a right image captured from different perspectives. The stereo image pair can be generated by taking two shots with a parallel shift using a single-lens retinal camera, or by taking a single shot using a stereo retinal camera. In the first step, the images of the ONH region are cut out from the original retinal fundus images in the first step. In the second step, the registration process of the stereo ONH image pair is performed to remove any displacements. The displacement between the stereo retinal fundus image pair is caused by movement of the eye and differential refraction of light due to astigmatic eye. It is necessary for precise measurement of the depth value to reduce displacement. In the third step, the "corresponding points" in each stereo ONH

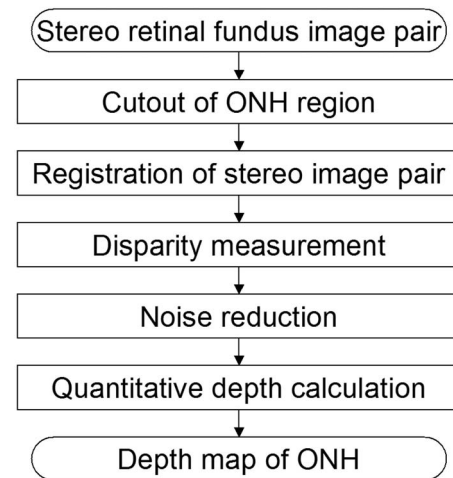


Fig. 1 Flowchart of the procedure for depth calculation of stereo retinal image pair.

image are detected, as discussed in Sec. 2.3. In the fourth step, the fluctuations in the disparity are reduced by both medium and smoothing filters. The fluctuations of disparity are caused by failure in the disparity measurement process. In the last step, the depth values of the 3-D structures are calculated from the results of the disparities measured for the configuration of the corresponding points.¹⁵

2.1 Cutout of Optic Nerve Head Region

The images of the ONH region were cut out from the original stereo image pair to reduce the processing area to expedite the subsequent steps, namely the registration of the stereo pair and disparity detection. In this processing step, the retinal fundus images were cut out to form quadrates at the position of the ONH region that was extracted automatically.

The ONH region has relatively high pixel values in three channels (red, green, and blue channels of the image) in the color stereo retinal fundus image pairs. *P*-tile thresholding¹⁶ can be applied to define a threshold for an approximate extraction of the ONH region, because individual variations of the ONH do not vary significantly. *P*-tile thresholding was performed on three channel images; subsequently, the region that was extracted in more than two channel images was determined as the extraction result. When more than two regions were extracted, the region with the maximum area was selected as the ONH region. The value of *P* in the *P*-tile thresholding operation was experimentally set to a value slightly greater than the average area of the ONH.

The blood vessels (BVs) running on the surface of the ONH interfered with the correct extraction of the ONH region in the *P*-tile thresholding operation. To solve this problem, the extraction of the ONH region was performed by using the images in which the BVs were erased. These erased pixels were then interpolated by using the RGB values of the pixels in the surrounding region. The pixel value used in the interpolation was calculated as

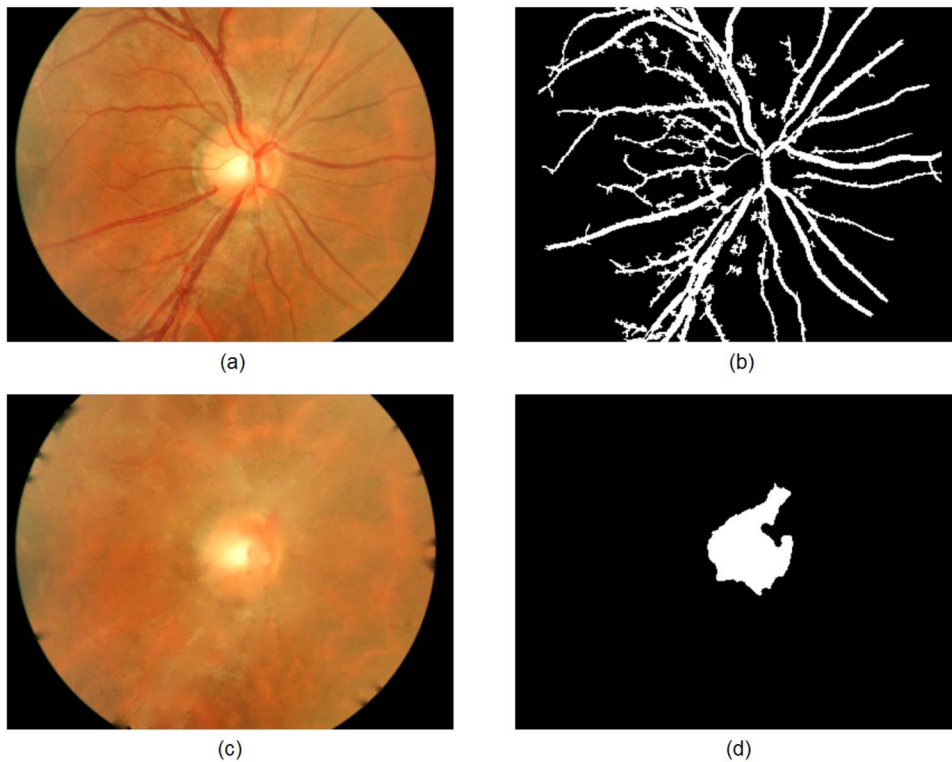


Fig. 2 Result of locating optic nerve head: (a) original left image, (b) blood vessel extraction, (c) blood vessel erasing, and (d) rough-segmented region of optic nerve head by P -tile thresholding operation.

$$P = \frac{\sum_{k=1}^n \frac{P_k}{l_k}}{\sum_{k=1}^n \frac{1}{l_k}}, \quad (1)$$

where P_k denotes the values of the pixels in the surrounding region, n is the number of surrounding pixels, and l_k is the distance between the interpolated pixel and each surrounding pixel. An example of the blood-vessel-erased image is shown in Fig. 2(c).

The BVs were extracted by using the black-top-hat transformation, which is a type of grayscale morphological operation,^{17,18} from the green channel of the color stereo retinal fundus image pairs. This grayscale morphological transformation was well suited to the task of segmenting BVs from the retinal fundus images. The black-top-hat transformation is defined as the residue between the image processed by morphological closing, which is a dilation followed by an erosion operation, and the original image. The “structure element” used in this transformation was a disk whose diameter was set to the same level as the thickness of the largest BVs in the ONH region. There are differences in the diameters of BVs among individuals. However, the differences are not significant. The thickness of the largest BV was determined in advance from experiments. The regions containing BVs were extracted after applying the Otsu thresholding technique¹⁹ to the black-top-hat transformed image. An example of the blood vessel image is shown in Fig. 2(b).

The center of the quadrate of the region of interest (ROI) around the ONH region was the gravity point of the ONH region extracted from the images in which the BVs were erased, as shown in Fig. 2(d).

In this study, the size of the original retinal fundus image was 1600×1200 pixels, the angle of view was 27 deg, and the size of the ROI around the ONH region was 600×600 pixels, as shown in Fig. 3. The size of the ROI was determined by considering the vertical diameter of the ONH in advance from experiments. The extraction process of the ONH region was implemented by using an image was reduced to one-sixth of its original left retinal fundus image.

2.2 Registration of Stereo Image Pair

The disparity, which is defined as the difference in the position of the corresponding points in the stereo image pair, de-

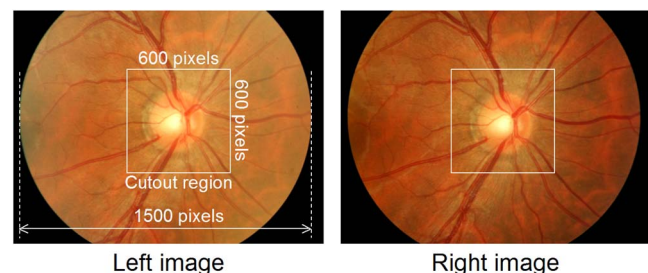


Fig. 3 An example of an original color retinal image pair and ROI around the ONH region. (Color online only.)

depends on the change in the position of not only the camera but also the subject (subject's motion). The disparity due to the subject's motion affects the calculated result of the depth value. To accurately measure the depth value, it is necessary to rectify the disparity due to the subject's motion. However, it is difficult to determine the motion that induces the disparity only on the basis of observations.

In the retinal fundus, the bell-shaped ONH has a dent on the opposite side, that is, the side facing the camera. Therefore, the cup region of the ONH exists at a distant position from the camera, and the disparity is small. Theoretically, it is possible to use the pixels in the ONH region, which have small disparities, for image registration. However, the region of the retina around the ONH is more suitable for the image registration task, because the blood vessels in the retina exist even on the curved surface. In this region, the right and left images exhibit a parallel shift. Moreover, the 3-D structure of the retinal region is simpler than that of the ONH. From the previously mentioned description, the image registration for rectifying the disparity due to a subject's motion was performed by using the pixels from regions other than the ONH region.

To exclude the ONH region from the registration process, the pixels of a retinal fundus image were allocated to two regions: the retina and ONH region. The pixels in the retina region were used for registering the stereo image pair. The boundary between the two regions was obtained by automatically extracting the ONH region.

In the first step, a contour of the ONH region was extracted from two stereo fundus images. The ONH region has a tendency to have a higher pixel value than the other regions. Furthermore, the contour of the ONH region can be expressed with a smooth closed curve in many cases. Therefore, the contour that exhibits high edge intensity, which is defined as a change in the brightness, was extracted as the smooth closed curve by using an active contour model.²⁰ Some techniques for the extraction of the ONH region have been previously reported.²¹⁻²⁴ The ONH extraction would be relatively easy, as long as the contour of the ONH is clear without any obscured region. Therefore, details pertaining to the definition of the energy and parameters of the active contour model used in the ONH extraction in this study are briefly explained here.

In the next step, the registration of the stereo image pair was performed by using all the pixels of the images from regions other than the ONH region. If the positional error is minimum, the sum of all the differences between the pixel values of the two images will be minimum. Therefore, the right image was translated and rotated until the sum of all the differences was minimum. This registration procedure used the cross correlation between the two images, calculated as

$$r = \frac{\sum_{i=0}^W \sum_{j=0}^H \{L(i,j) - \bar{L}\} \{R(i,j) - \bar{R}\}}{\left[\sum_{i=0}^W \sum_{j=0}^H \{L(i,j) - \bar{L}\}^2 \right]^{1/2} \left[\sum_{i=0}^W \sum_{j=0}^H \{R(i,j) - \bar{R}\}^2 \right]^{1/2}}, \quad (2)$$

where L and R are the feature values in the coordinate system (i, j) of the left and right images, respectively; W and H are the width and height of the image, respectively; and \bar{L} and \bar{R} are the average pixel values in the left and right images, re-

spectively. The information used in the registration was the pixel values of the three channels and its edge images created by the Sobel filter, whose algorithm detects the horizontal and vertical qualities of edges.¹⁶ The cross correlation coefficients were calculated between the red, green, and blue channels of the images and the edge images of these three channel images; subsequently, the average cross correlation coefficient was used as the index of positional error.

2.3 Disparity Measurement

The required disparity for obtaining the depth value was calculated from the location differences between the corresponding points. The detection of the corresponding points was performed using the pixels of the area within the registered ONH image pair. The detection of the corresponding points comprised the search for a point on the right image that corresponded to the reference point on the left image. The search was performed by setting up regions of interest (ROIs), including the pixels around the reference point and the candidate point separately. Two points on the left and right images, having a similar texture in their respective ROIs, were regarded as the corresponding points. The similarity was measured by the cross correlation coefficient, defined as

$$r = \frac{\sum_{i=-W/2}^{W/2} \sum_{j=-H/2}^{H/2} \{L(x_L + i, y_L + j) - \bar{L}\} \{R(x_R + i, y_R + j) - \bar{R}\}}{\left(\left[\sum_{i=-W/2}^{W/2} \sum_{j=-H/2}^{H/2} \{L(x_L + i, y_L + j) - \bar{L}\}^2 \right]^{1/2} \times \left[\sum_{i=-W/2}^{W/2} \sum_{j=-H/2}^{H/2} \{R(x_R + i, y_R + j) - \bar{R}\}^2 \right]^{1/2} \right)}, \quad (3)$$

where L and R are the feature values of the pixels in the ROIs set in the ONH image pair, \bar{L} and \bar{R} are the average feature values of the ROIs, (x_L, y_L) is the coordinate of the reference point in the left image, (x_R, y_R) is the coordinate of the candidate point in the right image, and W and H are the width and height of the ROI, respectively.

The information used in the detection of the corresponding points was the pixel values of the three channels of the images and its edge images created by the Sobel filter¹⁶ using the pixel value in the three channel images. In other words, the cross correlation coefficients were calculated between the red-, green-, and blue-channel images and the edge images of these three images; subsequently, the average cross correlation was used as an index of similarity.

The point having the maximum cross correlation coefficient was considered to be the corresponding point. When the maximum value of the cross correlation coefficient was smaller than a preset threshold value, it was assumed that the corresponding point of the reference point was not found. Additionally, when the texture amount in the ROI was low, the detection of the corresponding points was skipped because the reliability of the results of the detection from such regions was particularly low. The texture amount was estimated by considering the contrast of the ROI. The contrast was defined as the difference between the maximum and minimum pixel values. The disparity of the point that did not have a corre-

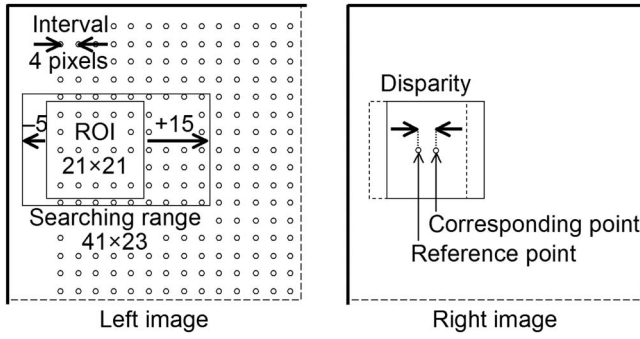


Fig. 4 Parameters in corresponding point detection of left image and right images.

sponding point was interpolated by the average of the disparities of the surrounding reference points. The threshold values were 0.5 for the cross correlation coefficient and 10 for the texture. These values were determined by a trial-and-error method. Therefore, the similarity of the ROI was computed as

$$\text{similarity} = \frac{\left(C_R \cdot r_R + C_G \cdot r_G + C_B \cdot r_B + C_{R\text{-edge}} \cdot r_{R\text{-edge}} \right) + C_{G\text{-edge}} \cdot r_{G\text{-edge}} + C_{B\text{-edge}} \cdot r_{B\text{-edge}}}{C_R + C_G + C_B + C_{R\text{-edge}} + C_{G\text{-edge}} + C_{B\text{-edge}}},$$

$$C = 1 \text{ if } r \geq 0.5 \text{ and } [\max\{L\} - \min\{L\}] \geq 10,$$

$$C = 0 \text{ otherwise.} \quad (4)$$

The disparity variance in the stereo retinal image pair is not very large, because the depth of the ONH is not over 1 mm. To distinguish the disparity difference, subpixel measurement was carried out by using expanded images. If the disparity measurement is implemented by using a double-sized (expanded) image, the disparity whose fineness is half of the original pixel size is obtained by dividing the result of the disparity measurement by two. The expanded images were generated from the original images by employing a bilinear interpolation technique. The expansion was performed in the horizontal direction of the image, and the expansion ratio was set to 3.

The parameters of this process are shown in Fig. 4. The size of the ROI was set to 21×21 pixels, and the searching range was set to 41×23 pixels ($[x_L - 5 \text{ pixels}, x_L + 15 \text{ pixels}]$). The reference points arranged at the equally spaced positions and the intervals were set to 4 pixels. Note that all the parameters about the disparity measurement using expanded images in conjunction with the expansion procedure, such as the size of the ROI and the searching range, were multiplied by three.

2.4 Noise Reduction

Sharp changes in the disparities were observed during the disparity detection. The depth of the ONH varies smoothly in the ROI. Therefore, if depth varies considerably in the ROI, the singular depth value assumes an incorrect result (these are called “noise” in this study). To reduce this type of noise, five iterations of a 5×5 median filter and a 3×3 moving average filter¹⁶ were applied to the disparity matrix.

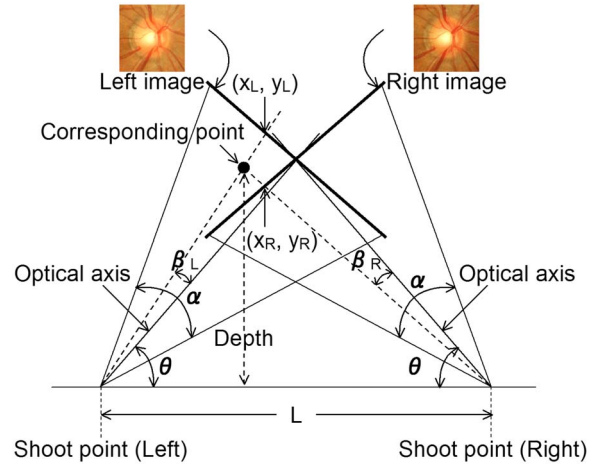


Fig. 5 Convergent visual system for depth calculation of stereo image pair.

2.5 Quantitative Depth Calculation

In general, there are two typical vision systems for implementing stereo vision—convergent and parallel (nonconvergent) systems. The depth value of the 3-D position was determined according to the value of the disparity at each location of the reference point in both the systems. Our stereo fundus camera has a parallel stereo configuration, in which there is no rotation between two optical axes. However, parallel light rays from the camera are refracted by the cornea and the crystal lens inside the eye. Hence, the parallel visual system is not suitable for the retinal fundus examination. The convergent visual system (Fig. 5) was selected in this study, and the depth value was calculated as

$$\text{Depth} = \frac{L \times \tan(\theta - \beta_L) \times \tan(\theta + \beta_R)}{\tan(\theta - \beta_L) + \tan(\theta + \beta_R)}, \quad (5)$$

$$\beta_L = \tan^{-1} \left\{ x_L \times \tan \left(\frac{\alpha}{2} \times \frac{\pi}{180} \right) \times \frac{2}{W} \right\}, \quad (6)$$

$$\beta_R = \tan^{-1} \left\{ x_R \times \tan \left(\frac{\alpha}{2} \times \frac{\pi}{180} \right) \times \frac{2}{W} \right\}, \quad (7)$$

and

$$x_R = x_L + \text{disparity}, \quad (8)$$

where x_L and x_R are the horizontal coordinates of the corresponding points in the left and right images, respectively. The original points in the coordinate system were arranged on the optical axis of the left and right viewpoints. α is the angle of view of the images; β is the angle between the position of the corresponding point and the optical axis; W is the width of the images; and L is the length of the baseline, which is the distance between the optical centers of the camera. θ is the angle of the optical axis. The depth value calculated by this method is the distance from the point at which the light rays are refracted.

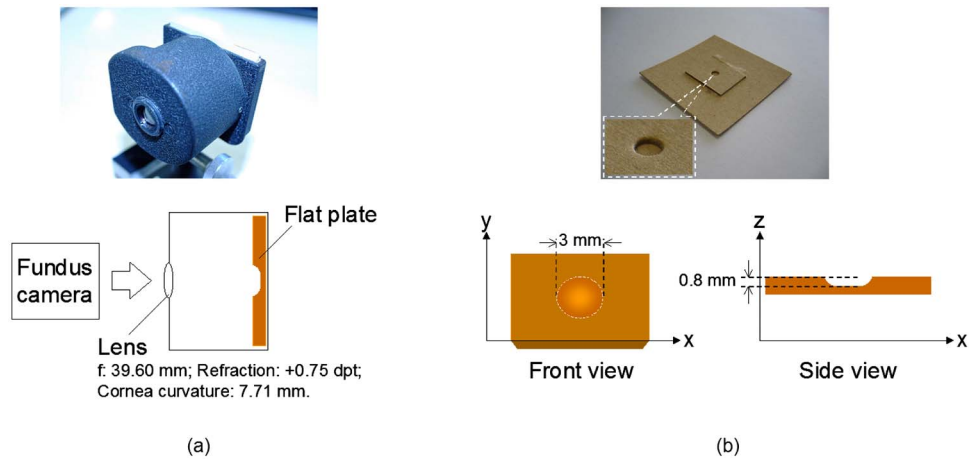


Fig. 6 Depth measurement of the disk depth model: (a) disk depth model that was built in-house for the estimation of the depth measurement method, and (b) structure of the flat plate that models the ONH.

2.6 Instrument

2.6.1 Retinal fundus camera

A new prototype stereo fundus camera (prototype of the WX-1, Kowa Company Limited, Tokyo, Japan) was used in this study. The camera can capture two images sequentially by the parallel movement of the aperture diaphragm in 0.14 s. Thus, a stereo image pair of retinal fundus is obtained without any mydriatic drug. The length of the base line, which is the length between the two cameras in a stereo vision system, equals the amount of parallel movement and is 2.3 mm. The

size of the image is 1600×1200 pixels; the diameter is 1500 pixels; and the angle of view is 27 deg.

2.6.2 Heidelberg retina tomograph

The HRT is a confocal scanning laser ophthalmoscope that uses a 670-nm wavelength diode laser to scan the retinal surface in three dimensions.^{25,26} The HRT used in this study was the first model marketed in 1991. The HRT provides a topographic image of the ONH, which is derived from multiple optical sections at 32 consecutive focal depth planes.

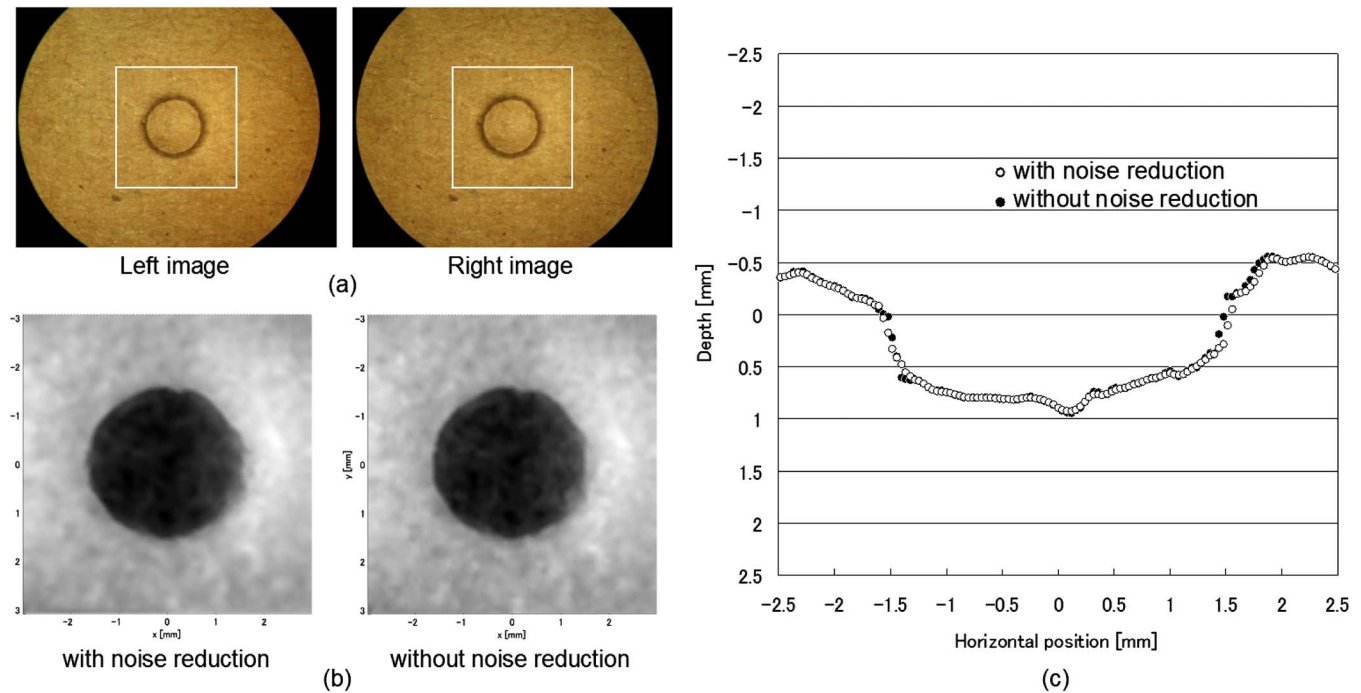


Fig. 7 Results of depth measurement of the disk depth model: (a) stereo image pair of the disk depth model (white squares indicate ROIs around the ONH regions), (b) depth maps obtained from the stereo image pair with and without noise reduction, and (c) depth distributions in the vicinity of the center line.

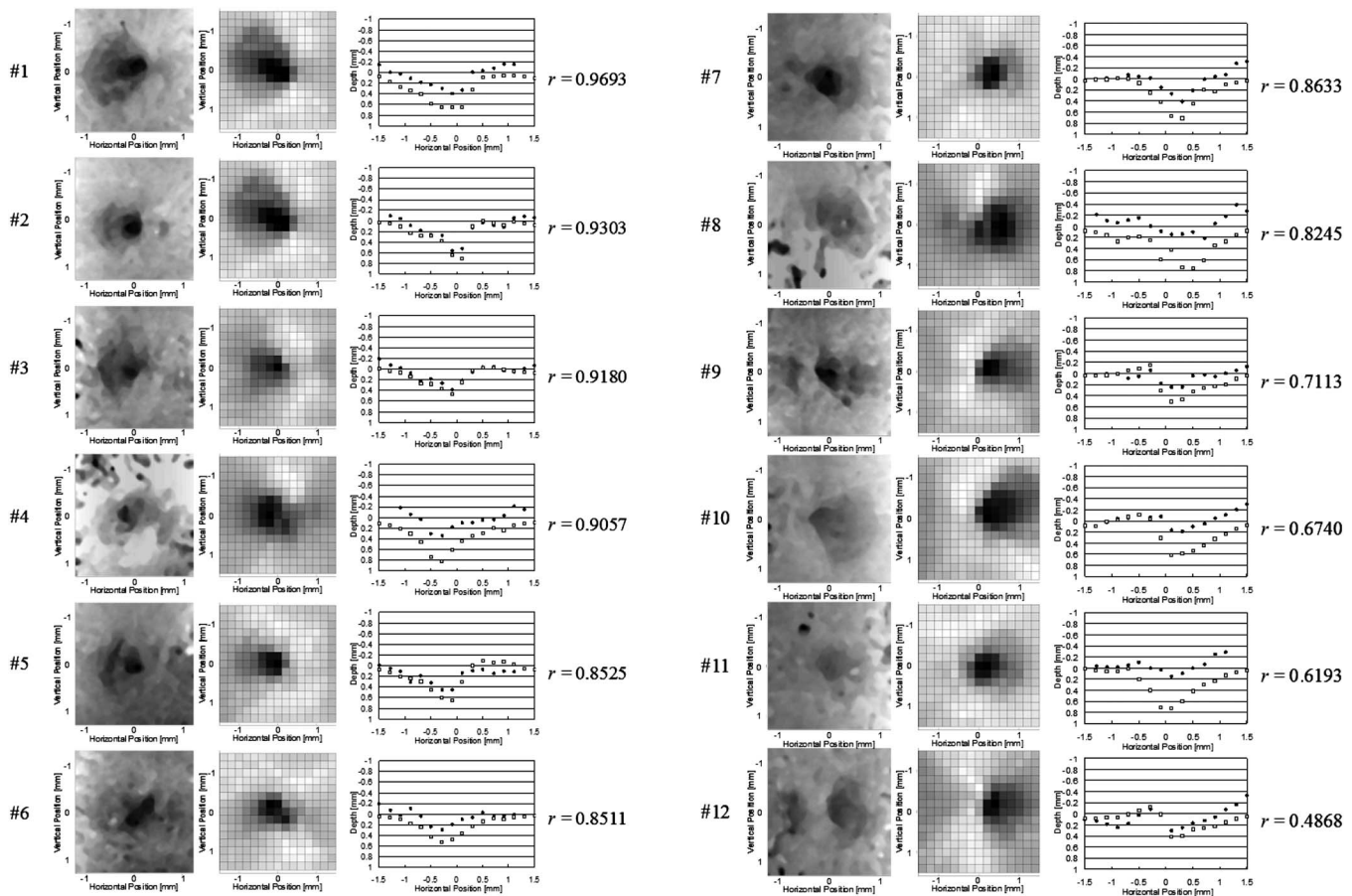


Fig. 8 The first column shows the depth maps generated from stereo fundus image pairs. The second column shows the topographic images obtained using the HRT. The third column shows the plots of the depth values in the depth maps and the topographic images along the midline across the ONH region. The results of the stereo fundus image pairs and the HRT are indicated by white circles and black squares, respectively.

3 Results

3.1 Depth Measurement of the Disk Depth Model

A basic experiment was conducted to estimate the accuracy of the proposed method. The test object is a disk depth model, as shown in Fig. 6(a), that was built in-house. It comprises a flat plate made of paper and a lens. The flat plate was arranged on the focal plane of the lens; it has a circular dent at the center to model the ONH, as shown in Fig. 6(b). The diameter of the dent is 3 mm and its depth is 0.80 mm. The focal length of the lens is 39.60 mm. Therefore θ , which is related to the angle of the optical axis, was set to 88.3 deg and the resolution of the images is 10.8 μm per pixel.

The accuracy of the proposed method was tested using the disk depth model described. The stereo image pair of the disk depth model is shown in Fig. 7(a). The registration of the stereo image pair was not performed because the subject did not move. The depth maps generated from the stereo image pair are shown in Fig. 7(b). The left image depicts the depth maps obtained with noise reduction and the right-hand image obtained shows the depth maps without noise reduction. The noise reduction improved the uniformity of the depth distribution.

Figure 7(c) shows the depth profile curve sampled at the center of the dent region in the model. The vertical axis indi-

cates the depth from the base plane, and the horizontal axis indicates the location. The focal length of the disk depth model (39.60 mm) is defined in the base plane. The white dots in this figure refer to the depth calculated with the noise reduction, and black dots refer to the depth calculated without noise reduction. The difference between these two depth distributions was small. The test result yielded a value of approximately 0.8 mm for the depth of the circular dent. The average depth distribution in the range from -1.0 to 1.0 mm was 0.767 ± 0.102 mm. The measured results showed good accordance with the actual value.

3.2 Depth Measurement of the Optic Nerve Head

The proposed technique was evaluated using 12 stereo image pairs of the ONH. All the subjects had normal eyes exception for refractive errors. The mean age of the subjects was 26.3 ± 3.94 yr (range 22 to 33 yr). The mean refractive error of the subjects was $-2.27 \text{ D} \pm 1.48$ (range 0 D to -4.25 D). This experiment assumed a focal length of 17 mm for the subjects' eyes. Therefore, θ was set to 83.4 deg and the resolution of the images was 5.44 μm per pixel. The focal length value of 17.0 mm was obtained from the data of the simplified Gullstrand eye model.²⁷

Table 1 Correlation coefficient between depth values obtained from stereo fundus camera and HRT. (RE)=right eye; (LE)=left eye; D=dioptr.

Oculus	Age	Refractive power	<i>r</i>	<i>p</i> value
1 (RE)	24	-1.25 D	0.9693	<0.0001
2 (RE)	33	-1.00 D	0.9303	<0.0001
3 (RE)	22	-3.75 D	0.9180	<0.0001
4 (RE)	28	-3.50 D	0.9057	<0.0001
5 (RE)	23	-4.25 D	0.8525	0.0001
6 (RE)	28	-1.50 D	0.8511	0.0001
7 (LE)	33	-4.00 D	0.8633	<0.0001
8 (LE)	28	-2.00 D	0.8245	0.0002
9 (LE)	23	-3.25 D	0.7113	0.0043
10 (LE)	24	-2.50 D	0.6740	0.0082
11 (LE)	28	-0.25 D	0.6193	0.0138
12 (LE)	22	0.00 D	0.4868	0.0558
Mean±SD	26.3±3.94	-2.27±1.48	0.80±0.15	

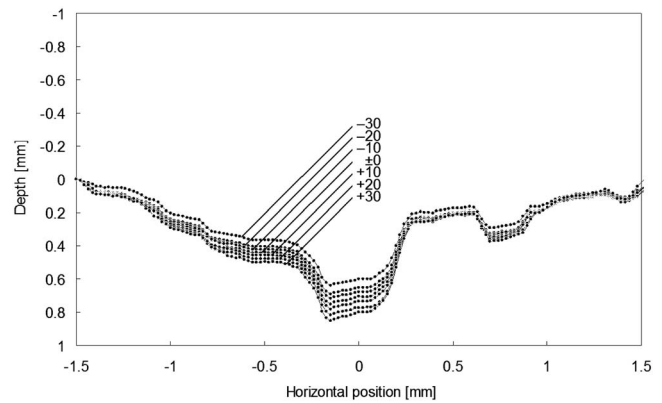
The depth value of the same ONH was measured using the HRT. Three topographic images were obtained for each eye, and the mean topographic image was generated. Each image consisted of 16×16 pixels, with each pixel corresponding to the retinal height at its location. The disk margin was determined by using a contour line of the ONH that was drawn by an ophthalmologist.

The first and second columns of Fig. 8 show the depth map generated from the stereo retinal fundus image pair and the topographic images obtained using the HRT, respectively. The results of the depth measurements obtained by using the stereo image pair were in accordance with the results of the HRT at some level.

The correlation coefficient between two depth distributions on the line running through the point with the maximum depth value was calculated for a comparison of the stereo fundus camera and the HRT. For this comparison experiment, 16 equally spaced points were selected from a total of 140 points in the horizontal direction of the depth map.

The third column of Fig. 8 shows the horizontal cross section of the depth distributions along with the correlation coefficients of the comparison results. In the point diagram, the white dots refer to the depths obtained by using the stereo fundus camera, and the black dots refer to the depth obtained using the HRT. The depth values obtained from the stereo fundus camera were horizontally shifted for comparison with the results of the HRT through manual procedures.

Table 1 presents the age and refractive power of the subjects, the calculated correlation coefficients, and the *p*-values obtained in the analysis results. There were four cases with significant correlation with $r > 0.9$ ($p < 0.001$), and four cases

**Fig. 9** The difference in depth measurement results affected by the amount of horizontal shift in the registration of a stereo image pair.

with good correlation with $r=0.8$ to 0.9 ($p < 0.001$). The mean of the correlation coefficients was $0.80 (\pm 0.15)$. There were three cases with poor correlation with $r < 0.7$; however, in these three cases, the distribution patterns were similar to each other. The relation between the refractive power of the eyes and the correlation coefficients are shown in Table 1. There was no correlation between the refraction and correlation coefficients ($r = -0.4586$, $p = 0.1337$).

4 Discussion

The displacements of stereo image pairs could be compensated by employing the registration technique in which pixels from regions other than the ONH region were used. It is difficult for actual eyes to determine the magnitude of displacements. Instead of estimating the accuracy of the registration, the effects of the displacements on the calculation results of the depth values were simulated by changing the amount of the displacement. Figure 9 shows different profiles of the depth values with varying horizontal displacement in the stereo image ranging from +30 pixels (further separation between the stereo image pair) to -30 pixels. The reference point (marked by ± 0) is the location having the maximum cross correlation coefficient in the proposed registration technique. The depth values of the ONH region were increased by $0.020 \text{ mm} (\pm 0.01)$, owing to the increase in the amount of displacement by 10 pixels (0.054 mm). The maximum amount of horizontal displacement in the stereo image pair used in this study was 134 pixels. From these facts, it is inferred that the registration is an indispensable procedure to obtain accurate depth values. The amount of displacement presumed from the pixel values in the stereo image pair will be different for the registration approach. However, it is assumed that the differences in the displacement calculated by various methods are a few pixels. Thus, the differences between the depth values are considered to be small.

In the disparity detection step, the thresholds of the cross correlation coefficient and the texture were set with the viewpoint of eliminating low-trust results. If the thresholds are set to a high value, only the high-trust corresponding points are determined. However, the number of corresponding points determined is decreased in such cases. Thus, the undetermined points cannot be interpolated by using the disparities of the

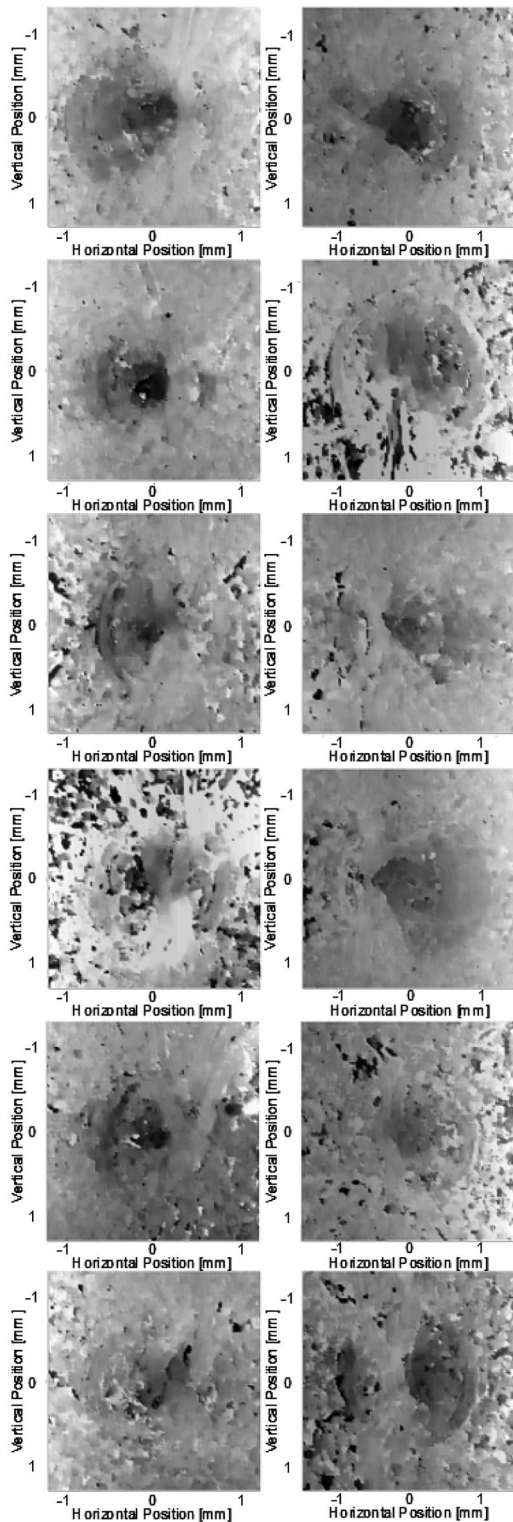


Fig. 10 Depth maps obtained from stereo fundus image without noise reduction.

surrounding reference points. The threshold values should be automatically adjusted according to the condition of the stereo image pair.

The median filter was effective in reducing noise in the depth maps without substantial deformation of the curve

shape. In Fig. 10, some incorrect depth values can be observed; these are labeled as “noise” in the depth maps without the noise reduction. It is considered that the inaccuracy was caused by an error while searching for the corresponding points. The noise tended to appear in regions in which the amount of textures was larger than the threshold, but differences in the color between the left and right image were large. In the proposed method, the amount of texture, in the ROI used for the disparity detection was presumed by using the contrast defined as the difference between the maximum and minimum pixel values in the ROI. However, for the case where the color gradient is smooth in the ROI, there may not be a considerable amount of texture even if the contrast is high. Therefore, it might be better to use only the edge-extracted image to evaluate the amount of texture in the ROI. If a property of a subject can be presumed, the range of admissible disparity values will be limited to drop a sharply changed disparity as a wrong result. Moreover, because the searching range can also be limited, not only the appearance of noises but also the calculation time will be reduced.

The axis length of the eye was assumed as 17.0 mm for all cases in this study. However, there are individual variations in the focal lengths of actual eyes. The refraction index of the eye also varies from place to place on the surface of the subject’s cornea in astigmatic eyes. These are considered the causes for the differences in the obtained depth values between the stereo fundus camera and the HRT in absolute level. Moreover, the camera lens distortion should be connected to obtain more accurate quantitative values.

5 Conclusion

In this study, we conduct quantitative measurements of the depth value of an ONH region from stereo retinal images. The measurement results obtained when the disk depth model is used are approximately consistent. The depth values obtained from the stereo image pair are in accordance with the results of the HRT. These depth values might be useful as assisting parameters for ophthalmologists in the diagnosis of degrees of glaucoma.

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 would like to acknowledge the contribution of R. Shiraki of Gifu University for the acquisition of the HRT data.

References

1. H. A. Quigley, “Number of people with glaucoma worldwide,” *Br. J. Ophthalmol.* **80**(5), 389–393 (1996).
2. H. A. Quigley and A. T. Broman, “The number of people with glaucoma worldwide in 2010 and 2020,” *Br. J. Ophthalmol.* **90**(3), 262–267 (2006).
3. A. Iwase, Y. Suzuki, M. Araie, T. Yamamoto, H. Abe, S. Shirato, Y. Kuwayama, H. K. Mishima, H. Shimizu, G. Tomita, Y. Inoue, and Y. Kitazawa, “The prevalence of primary open-angle glaucoma in Japanese: the Tajimi Study,” *Ophthalmology* **111**(9), 1641–1648 (2004).
4. T. Yamamoto, A. Iwase, M. Araie, Y. Suzuki, H. Abe, S. Shirato, Y. Kuwayama, H. K. Mishima, H. Shimizu, G. Tomita, Y. Inoue, and Y. Kitazawa, “The Tajimi Study report 2: prevalence of primary angle closure and secondary glaucoma in a Japanese population,” *Ophthalmology* **112**(10), 1661–1669 (2005).
5. University of Michigan Kellogg Eye Center, “Glaucoma,” <http://>

- www.kellogg.umich.edu/patientcare/conditions/glaucoma.html (2007).
6. M. Detry-Morel, T. Zeyen, P. Kestelyn, J. Collignon, and M. Goethals, "Screening for glaucoma in a general population with the non-mydiatic fundus camera and the frequency doubling perimeter," *Eur. J. Ophthalmol.* **14**(5), 387–393 (2004).
 7. J. M. Tielsch, J. Katz, H. A. Quigley, and A. Sommer, "Intraobserver and interobserver agreement in measurement of optic disc characteristics," *Ophthalmology* **95**(3), 350–356 (1988).
 8. F. S. Mikelberg, C. M. Parfitt, N. V. Swindale, S. L. Graham, S. M. Drance, and R. Gosine, "Ability of the Heidelberg retina tomograph to detect early glaucomatous visual field loss," *J. Glaucoma* **4**(4), 242–247 (1995).
 9. M. Iester, F. S. Mikelberg, and S. M. Drance, "The effect of optic disc size on diagnostic precision with the Heidelberg retina tomograph," *Ophthalmology* **104**(3), 545–548 (1997).
 10. V. R. Algazi, J. L. Keltner, and C. A. Johnson, "Computer analysis of the optic cup in glaucoma," *Invest. Ophthalmol. Visual Sci.* **26**(12), 1759–1770 (1985).
 11. M. Okutomi and G. Tomita, "Color stereo matching and its application to 3-D measurement of optic nerve head," *Proc. 11th IAPR Intl. Conf. Patt. Recog.*, pp. 509–513 (1992).
 12. E. Corona, S. Mitra, M. Wilson, T. Krile, Y. H. Kwon, and P. Soliz, "Digital stereo image analyzer for generating automated 3-D measures of optic disc deformation in glaucoma," *IEEE Trans. Med. Imaging* **21**(10), 1244–1253 (2002).
 13. J. Xu and O. Chutatape, "Auto-adjusted 3-D optic disk viewing from low-resolution stereo fundus image," *Comput. Biol. Med.* **36**(9), 921–940 (2006).
 14. J. Xu, O. Chutatape, C. Zheng, and P. C. T. Kuan, "Three dimensional optic disc visualisation from stereo images via dual registration and ocular media optical correction," *Br. J. Ophthalmol.* **90**(2), 181–195 (2006).
 15. T. Nakagawa, Y. Hayashi, Y. Hatanaka, A. Aoyama, T. Hara, M. Kakogawa, H. Fujita, and T. Yamamoto, "Comparison of the depth of an optic nerve head obtained using stereo retinal images and HRT," *Proc. SPIE*, **6511**, 65112M-1–65112M-9 (2007).
 16. J. R. Parker, *Algorithms for Image Processing and Computer Vision* Wiley Computer Publishing, New York, (1997).
 17. J. Serra, "Introduction to mathematical morphology," *Comput. Vis. Graph. Image Process.* **35**(3), 283–305 (1986).
 18. S. R. Sternberg, "Grayscale morphology," *Comput. Vis. Graph. Image Process.* **35**(3), 333–355 (1986).
 19. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (1979).
 20. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.* **1**(4), 321–331 (1987).
 21. J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, E. Fletcher, and L. Kennedy, "Optic nerve head segmentation," *IEEE Trans. Med. Imaging* **23**(2), 256–264 (2004).
 22. M. B. Merickel, X. Wu, M. Sonka, and M. Adramoff, "Optimal segmentation of the optic nerve head from stereo retinal images," *Proc. SPIE* **6143**, 1031–1038 (2006).
 23. M. B. Merickel, M. D. Adramoff, M. Sonka, and X. Wu, "Segmentation of the optic nerve head combining pixel classification and graph search," *Proc. SPIE* **6512**, 651215-1–651215-10 (2007).
 24. M. D. Adramoff, W. L. Alward, E. C. Greenlee, L. Shuba, C. Y. Kim, J. H. Fingert, and Y. H. Kwon, "Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features," *Invest. Ophthalmol. Visual Sci.* **48**(4), 1665–1673 (2007).
 25. Heidelberg Engineering, see <http://www.heidelbergengineering.de>.
 26. S. Miglior, M. Casula, M. Guareschi, I. Marchetti, M. Iester, and N. Orzalesi, "Clinical ability of Heidelberg retinal tomograph examination to detect glaucomatous visual field changes," *Ophthalmology* **108**(9), 1621–1627 (2001).
 27. Y. L. Grand and S. E. Hage, *Physiological Optics*, Springer-Verlag, New York (1980).