



ELSEVIER

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Automated segmentation of optic disc region on retinal fundus photographs: Comparison of contour modeling and pixel classification methods

Chisako Muramatsu^{a,*}, Toshiaki Nakagawa^{b,2}, Akira Sawada^{c,3}, Yuji Hatanaka^{d,4}, Takeshi Hara^{a,1}, Tetsuya Yamamoto^{c,3}, Hiroshi Fujita^{a,1}

^a Department of Intelligent Image Information, Graduate School of Medicine, Gifu University, 1-1 Yanagido, Gifu 501-1194, Japan

^b Research & Development Section, Electronics & Optics Division, Kowa Company, Ltd., Hamamatsu Factory, 1-3-1 Shin-Miyakoda, Kita-ku, Hamamatsu-shi, Shizuoka 431-2103, Japan

^c Department of Ophthalmology, Graduate School of Medicine, Gifu University, 1-1 Yanagido, Gifu 501-1194, Japan

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

ARTICLE INFO

Article history:

Received 3 December 2009

Received in revised form

12 April 2010

Accepted 19 April 2010

Keywords:

Retinal fundus images

Glaucoma

Cup-to-disc ratio

Active contour model

Pixel classification

ABSTRACT

The automatic determination of the optic disc area in retinal fundus images can be useful for calculation of the cup-to-disc (CD) ratio in the glaucoma screening. We compared three different methods that employed active contour model (ACM), fuzzy c-mean (FCM) clustering, and artificial neural network (ANN) for the segmentation of the optic disc regions. The results of these methods were evaluated using new databases that included the images captured by different camera systems. The average measures of overlap between the disc regions determined by an ophthalmologist and by using the ACM (0.88 and 0.87 for two test datasets) and ANN (0.88 and 0.89) methods were slightly higher than that by using FCM (0.86 and 0.86) method. These results on the unknown datasets were comparable with those of the resubstitution test; this indicates the generalizability of these methods. The differences in the vertical diameters, which are often used for CD ratio calculation, determined by the proposed methods and based on the ophthalmologist's outlines were even smaller than those in the case of the measure of overlap. The proposed methods can be useful for automatic determination of CD ratios.

© 2011 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Glaucoma is the second leading cause of vision loss worldwide, and its prevalence is expected to increase [1,2]. The

risk of significant visual impairment due to glaucoma can be minimized by early diagnosis and optimal treatment, including a screening examination by use of retinal fundus photographs. One of the glaucomatous changes observed in the retinal fundus images of affected individuals is enlarge-

* Corresponding author. Tel.: +81 58 230 6519; fax: +81 58 230 6514.

E-mail address: chisa@fjt.info.gifu-u.ac.jp (C. Muramatsu).

¹ Tel.: +81 58 230 6519; fax: +81 58 230 6514.

² Tel.: +81 53 428 5712; fax: +81 53 428 5719.

³ Tel.: +81 58 230 6284; fax: +81 58 230 6285.

⁴ Tel.: +81 749 28 9556; fax: +81 749 28 9576.

Table 1 – The summary of the image datasets used for the training and testing.

	Number of patients	Number of glaucoma images	Number of normal images	Camera system	Disc outlines by ophthalmologist
Training set	56	25	55	Stereo	All by A, partially by B, C
Testing set I	98	60	38	Stereo	A
Testing set II	30	8	22	2D	A, B and D

ment of depression, called cupping, of optic disc and/or local rim thinning. Although the ratio of the diameters of the cup and disc (CD ratio) is a widely accepted index for the assessment of cupping, its determination is often subjective and may include considerable intra- and inter-reader variations [3,4]. To assist consistent evaluation and facilitate the comparison of disc features in successive examinations, we have been investigating a method for the automated determination of the cup and disc regions in stereo retinal fundus images [5] and regular fundus images. In this study, we focused on the determination of the disc regions.

A number of studies have reported on automated localization of optic discs; several studies have also reported on segmentation of optic discs [6–14]. The method by Walter and Klein [6] is based on the morphological operation; they applied the watershed transformation to the gradient image. Lalonde et al. [7] detected optic disc edge using the Canny edge detector, and the optic disc region was determined by matching the edge map with a circular template. Li and Chutatape [8] proposed an active shape model, which consists of building a model with training cases and iteratively matching the landmark points on the disc edges and main vessels inside the disc. Osareh et al. [9] and Lowell et al. [10] employed the active contour model, which consists of finding the optimal points based on the external energy corresponding to the image gradient and the internal energy corresponding to the smoothness of the contour. In the method by Merickel et al. [11], the optimal path corresponding to the disc border was determined based on the cost function consisting of edge, texture, and priori border information. Xu et al. [12] also employed the deformable model technique that includes disc margin determination through minimization of the energy function defined by image intensity, image gradient, and boundary smoothness. The method proposed by Wong et al. [13] was based on the level-set technique followed by ellipse fitting in order to smooth the disc boundary. The method employed by Abramoff et al. [14] was different from other methods: they employed a pixel classification method using the feature analysis and nearest neighbor algorithm. The final output of their scheme comprised the classification of each pixel to a group that belongs to rim, cup, or background.

It can be seen that the majority of these methods were based on the contour modeling on the basis of the edge characteristics, and the result of the study by Abramoff et al. indicates the pixel classification method is also potentially useful. However, these methods are tested on the different databases; therefore, it is difficult to compare the different methods and determine their efficacy with the new database. In addition, the images captured by different camera systems may possess different characteristics. In this study, we investigated and compared three different methods for the determination of the disc regions on retinal fundus images: a method using

the active contour model (ACM) as an example of favored disc segmentation algorithms and methods using the fuzzy c-mean (FCM) clustering and an artificial neural network (ANN) as samples of pixel classification algorithms based on unsupervised and supervised learning. The result of each method was tested on the new database that included images captured by two different camera systems.

2. Image databases

Three sets of retinal fundus images were prepared and used as training and testing sets according to the time they were obtained. The details of the datasets are summarized in Table 1. The distributions of the disc diameters in pixels are shown in Fig. 1. The average disc diameters of the test cases are slightly larger than that of the training cases.

2.1. Training database

The disc determination methods were developed and trained by use of 80 retinal fundus images. These images were obtained from ophthalmology patients at Gifu University Hospital, Gifu, Japan and normal volunteers from March 2006 to April 2007 using a stereo retinal fundus camera (prototype of the WX-1, Kowa Company, Ltd., Tokyo, Japan) [15]. Images were saved in 1600 × 1200 pixel JPEG format. Although each case consists of a pair of images captured from two directions in a split second, data on the depth of the optic disc were not required for disc determination; therefore, the image captured from one direction, i.e., the image that was captured first, was

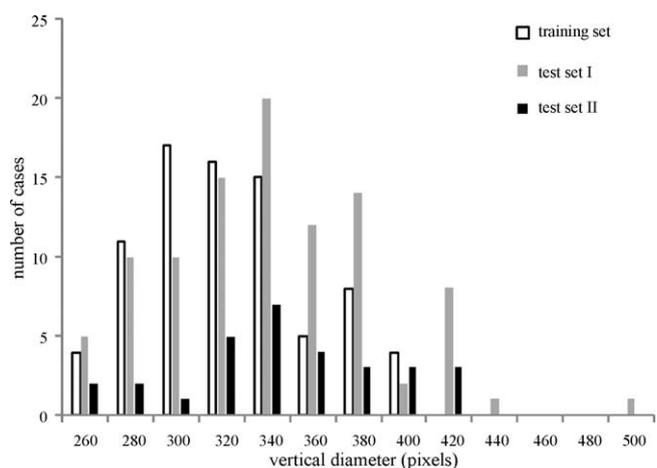


Fig. 1 – Distributions of vertical disc diameters in pixels for the training and test cases. The diameters of the optic discs in test set II are adjusted by the factor of 1.144 to account for the smaller optic angle.

used for disc determination. Because of the limited number of cases available, some of the included images were those of right and left eyes of the same patients. These 80 images included 25 images of eyes with a glaucomatous optic disc and 55 images of eyes without the signs of glaucoma as assessed by an ophthalmologist specializing in the diagnosis of glaucoma.

An ophthalmologist (ophthalmologist A) determined the outlines of the discs by providing several points on the border of a disc on a computer monitor. These points were connected by using the spline interpolation method, and the ophthalmologist could adjust the points if required. The disc regions determined by the ophthalmologist were considered as the gold standard for this study. For a subset of the cases, including 23 glaucomatous and 20 normal cases, the disc outlines were determined by two other ophthalmologists (ophthalmologists B and C), in addition to ophthalmologist A, in order to evaluate inter-reader variability. One of the latter two drew the outlines in one stroke using of a digital pen, while the other marked several points, as done by ophthalmologist A.

Each method for the determination of the disc regions was trained and optimized by using the 80 training images. The training parameters for the unsupervised training methods, i.e., contour modeling and clustering methods, were selected on the basis of the results of the substitution test which achieved the best results. In the case of the ANN method, randomly selected pixels from each case were employed for the training for reducing the effect of overtraining and time spent on training. The set of ANN parameters and the combination of input features that provided the best result by using all of the pixels for the testing were selected.

2.2. Testing set I

For an objective evaluation of efficacy on unseen cases, new images were obtained. The images were obtained from ophthalmology patients at Gifu University Hospital from May 2007 to June 2008 using the same stereo camera system. Images of one eye per patient were randomly selected and included in the database, if the image quality was adequate. The diagnoses and disc outlines were determined by ophthalmologist A as previously described. For simplicity, images of eyes that exhibited other major ocular diseases were excluded; the database consisted of 98 cases including 60 cases with signs of glaucoma and 38 cases without.

2.3. Testing set II

In order to investigate whether the proposed methods can be applicable to images obtained by a different camera, another database was prepared and employed for the evaluation of the methods. These images were obtained from ophthalmology patients prior to May 2005 and normal volunteers from April 2006 to May 2007; images from the volunteers whose images were included in the training set were excluded from this dataset. Images were captured by a regular retinal fundus camera (nonmyd7, Kowa Company, Ltd., Tokyo, Japan), and saved in the JPEG format with a matrix size of 3008×2000 pixels. The stereo fundus camera is generally used to observe optic discs; therefore, a smaller optic angle

of 27° was employed. In contrast, with the regular camera, a larger retinal field is captured for the diagnosis of any abnormality in the retina and optic discs using a 45° optic angle.

Three ophthalmologists, A, B, and D, determined the disc outlines and also provided the diagnosis for these cases. If at least one of them considered that it was not possible to make a reliable diagnosis, such as the images of eyes with cataract, those images were excluded. The cases in which glaucoma was diagnosed by at least two ophthalmologists were considered as glaucoma cases, while those diagnosed by only one ophthalmologist as glaucoma were considered to be suspected cases. Only one eye per patient was randomly selected and included in the dataset. Thus, testing set II consisted of 30 cases including 7 glaucoma, 1 suspected glaucoma, and 22 non-glaucoma cases.

3. Methods

3.1. Localization of optic disc

The optic disc can be identified as a bright region on a retinal fundus image. In order to reduce computational time, the approximate locations of optic discs were identified, and regions of interest (ROIs) that included the optic discs were extracted from the images. Initially, the image size was reduced by a factor of 4 by pixel averaging. The potential location of an optic disc was identified through the p-tile thresholding method [16] using the red channel of RGB colored images. A fixed threshold value of 75,000 pixels (approximately 3.9%) was applied; this was determined experimentally. If an island with the maximum area in the thresholded image had a degree of circularity larger than 0.6, the island was considered to be the approximate disc region, and an ROI of 600×600 pixels was extracted about the center of the island; otherwise, the next brightest region was searched by repeating the process. The degree of circularity was defined as the fraction of the area of an island that overlaps with a circle that has an area equivalent to the island and is centered at the centroid of the island.

Because the images in the test database II were captured by the different camera with the different optic angle, the size of the discs appeared to be different. The pixel sizes of the images were not known; therefore, the images of an eye that were captured with both cameras were used to estimate the ratio of the pixel sizes in two images. The area of the disc in the image captured by the stereo camera was 80,615 pixels, whereas that in the image captured by the regular camera was 69,046 pixels. It was estimated that the disc would appear approximately 14.4% smaller in the image captured by the regular camera. Based on this result, the threshold value of the p-tile method for the images in test database II was set to 64,200 pixels, and the ROI size was set to 520×520 pixels. These were the only parameters that were changed in the testing procedures. In the future, if the camera specifications are known, these parameters can be adjusted easily and automatically for images captured with different cameras.

3.2. Removal of blood vessels

Large blood vessels extending from optic discs may interfere with the precise determination of the optic disc region. In order to reduce such effect, the pixels corresponding to blood vessels were detected, and their pixel values were interpolated by those of the surrounding pixels for creating a “blood-vessel-erased” image. A detailed description of this method has been provided elsewhere [5]. Briefly, the blood vessels were extracted by using the black top-hat transformation [17] on green-channelled images. After applying the Otsu thresholding technique [18], the detected pixels were replaced by the weighted average of the surrounding non blood vessel pixels, whose weights were inversely proportional to the distance between the two pixels.

3.3. Disc extraction using three methods

3.3.1. Disc extraction by ACM

One of the methods investigated for the determination of optic disc regions was based on the active contour modeling [19]. In this method, expected disc borders were searched in radial directions from the center of an ROI. The initial candidate points of the borders were selected on the basis of the edges identified by applying the Canny edge detector [17] on the red channel of the blood-vessel-erased images. Assuming that the optic discs are approximately round in shape, the edges with a gradient directed perpendicularly to the radial direction were suppressed by,

$$A' = A \cos(\theta - \varphi),$$

where A is the amplitude of the gradient, and θ and φ are the radial and edge directions, respectively, so that the edges in the radial direction are sustained, while the edges perpendicular to the radial direction are eliminated. The edges of the remaining blood vessels can be suppressed using this technique.

Based on the detected edges, a median distance from the center to the closest edges was determined in 48 directions. In each direction, a point on the detected edges that was close to the median value and also close to the edge candidates in the previous and next directions was selected as the initial candidate point. These candidate points were adjusted by energy minimization based on,

$$E = W_{\text{dist}}E_{\text{dist}} + W_{\text{curv}}E_{\text{curv}} + W_{\text{edge}}E_{\text{edge}}$$

where E_{dist} and E_{curv} represent the contour smoothness while E_{edge} accounts for the edge strength with the corresponding weight factors, which were determined empirically using the training cases. E_{dist} was determined by the average distance between candidate points in the current and the next directions in both sides, and E_{curv} was based on the distance between the candidate point in the current direction and the middle point of the candidates in the next directions. At each iteration, eight neighborhood pixels of the candidate pixel in the previous iteration were searched; if the cumulative energy in all directions became lower than that of the previous iteration, the candidate points were updated.

3.3.2. Disc extraction by FCM method

Two kinds of pixel classification methods were investigated: FCM and ANN. One of the advantages of unsupervised learning can be that it can minimize the overtraining effect so that the result is generalizable to new cases. In FCM, the membership degree, u_{ij} , and the cluster centers, c_j , are updated by,

$$u_{i,j} = \frac{1}{\sum_{k=1}^C ((X_i - C_j)/(X_i - C_k))^{2/(m-1)}}, \quad c_j = \frac{\sum_{i=1}^N u_{ij} \cdot X_i}{\sum_{i=1}^N u_{ij}}$$

at each iteration [20]. In this study, we selected the index of fuzziness, m , as 2 and the optimization parameter, ϵ , which is the threshold of maximum change in the membership degrees, as 0.2. For each pixel of interest, features such as the original pixel values (RGB components), pixel values in the blood-vessel-erased image, pixel values in the surrounding pixels, local contrast, and edge information were employed as potential input data to the classifier. The surrounding pixels were varied from 3×3 to 15×15 pixels around the pixel of interest. The local contrast was defined as the maximum difference in pixel values between the surrounding pixels and the pixel of interest, and the edge information corresponded to the pixel value in the result of the Canny edge detection or Sobel filtering. The algorithm was trained to classify pixels into two to four clusters, and a resulting cluster that contained the center pixel of an ROI was regarded as the disc cluster. After the morphological opening operation, an isolated region in the center of the ROI was determined as the disc region.

3.3.3. Disc extraction by using ANN

ANN generally has the advantage of being able to solve relatively complex problems; however, the training of an ANN takes time, especially when the numbers of training cases and input units are large, and there is a risk of overtraining. In order to complete the training in a reasonable amount of time and to reduce the effect of overtraining, not all the pixels in the training set were used for the training of the ANN. A feed forward, three layered network with a backpropagation algorithm was employed.

The ANN was trained with the selected features mentioned in the previous section as input data and the gold standard based on the disc outlines determined by ophthalmologist A as teacher data. The number of the hidden units was varied depending on the number of input units. The number of output units was set to two or three, corresponding to the likelihood of a pixel belonging to the disc and background regions or to the disc, disc border, and background regions, respectively. Based on the number of output units, 20 pixels per group were randomly selected from each case and employed for training. The remaining pixels and those used for training were used for testing of the trained ANN to select the optimal parameters. If the output unit corresponding to the disc likelihood had the highest value, those pixels were classified as part of the potential disc region. After the morphological opening operation, an isolated region at the center of the ROI was determined as the final disc region.

Table 2 – Agreement between the disc regions determined by three ophthalmologists (A–C) for the training database.

	Ophthalmologists A vs B	Ophthalmologists A vs C	Ophthalmologists B vs C
Glaucoma	0.940	0.935	0.942
Normal	0.945	0.939	0.940
All	0.942	0.937	0.941

Table 3 – Agreement between the disc regions determined by three ophthalmologists (A, B, and D) for the test database II.

	Ophthalmologists A vs B	Ophthalmologists A vs D	Ophthalmologists B vs D
Glaucoma	0.940	0.851	0.855
Normal	0.935	0.904	0.912
All	0.937	0.896	0.903

4. Results

4.1. Agreement on disc regions between ophthalmologists

The region agreement for 43 cases in the training dataset was evaluated by the ratio of the area of intersection to the area of union,

$$M = \frac{A \cap B}{A \cup B}$$

where A and B are the disc regions determined by two ophthalmologists. The result is summarized in Table 2. The intra-reader agreement for the disc regions in a subset of five glaucoma and five normal cases by ophthalmologist A was 0.95; therefore, the agreement between different ophthalmologists can be considered to be comparable to the intra-reader agreement. Garway-Heath et al. [21] reported that the inter-observer variation in disc area in 30 fundus photographs

assessed by six observers was 8.1% in terms of a coefficient of variation. In our study, the average coefficient of variation in disc area assessed by three observers for the training set was 2.6%, which was very small. There was no apparent difference between the agreement for the glaucoma cases and the normal cases, nor was there any specific tendency among the three ophthalmologists.

For testing dataset II, outline data from ophthalmologist B for two glaucoma cases were missing; therefore, agreement was evaluated by using 28 cases. Table 3 shows the agreement for the disc regions determined by the three ophthalmologists. The average agreements were slightly lower for testing dataset II, especially between ophthalmologist A and D and between ophthalmologist B and D. Fig. 2 shows the disc outlines determined by the three ophthalmologists. The agreement was high (0.952, 0.956, and 0.967) for the case in the top row while it was relatively low (0.917, 0.619, and 0.632) for the case in the bottom row. For the bottom case, the disc region determined by ophthalmologist D was larger than those by the other two, and

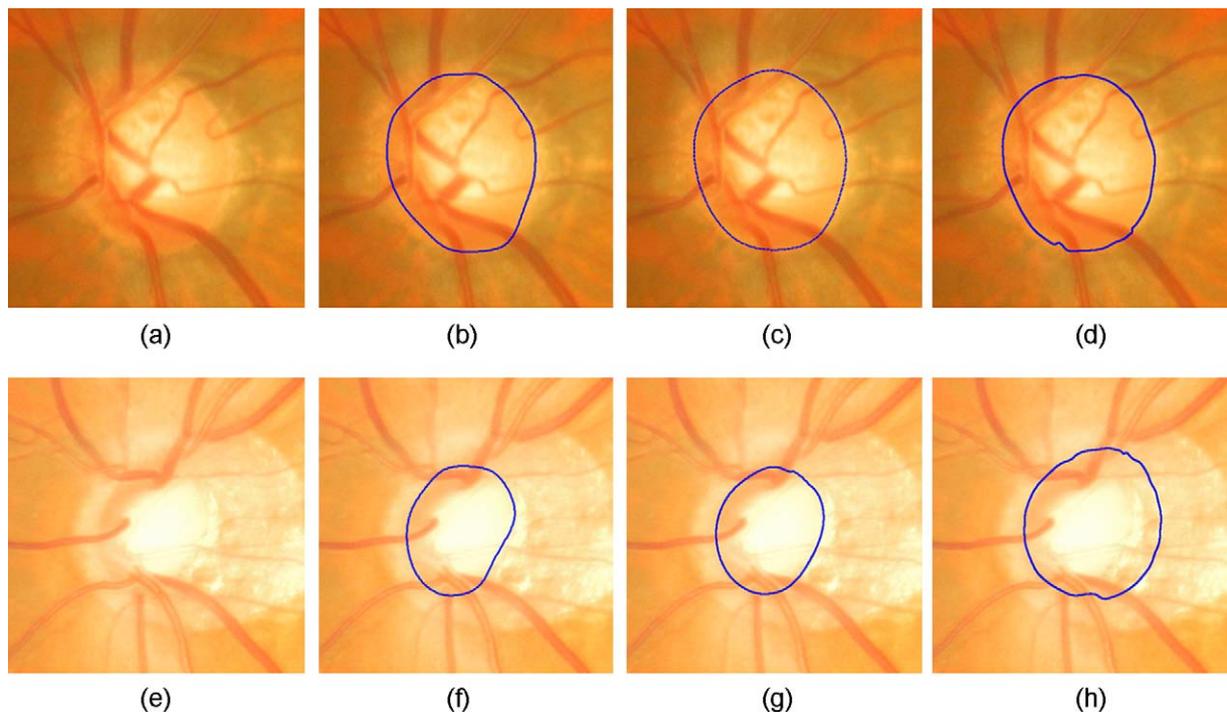


Fig. 2 – Comparison of manual outlines of the optic discs determined by three ophthalmologists: (a and e) original images; (b–d) disc outlines by three ophthalmologists for (a); (f–h) disc outlines by three ophthalmologists for (e). Both images are from the test dataset II.

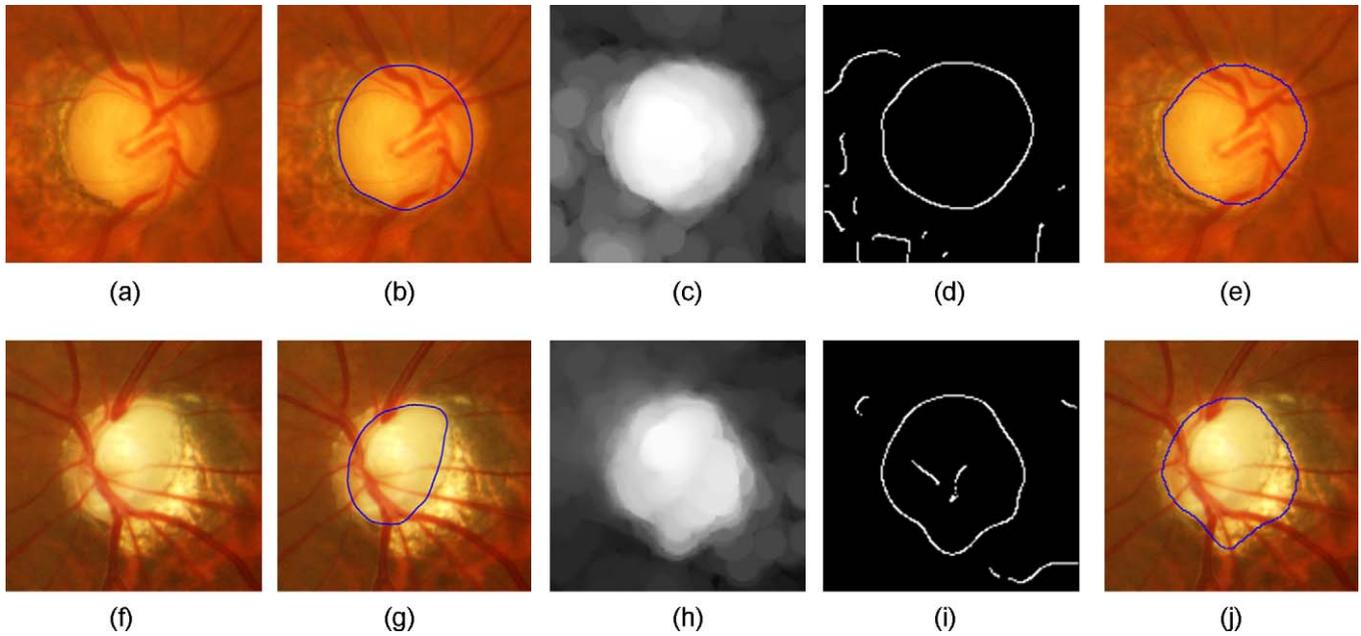


Fig. 3 – Disc outlines determined by the ACM method: (a and f) original images; (b and g) images with the gold standard of disc outlines; (c and h) red channel of the blood-vessel-erased images; (d and i) the edges detected by the Canny edge detector; (e and j) images with the disc outlines determined by the ACM method.

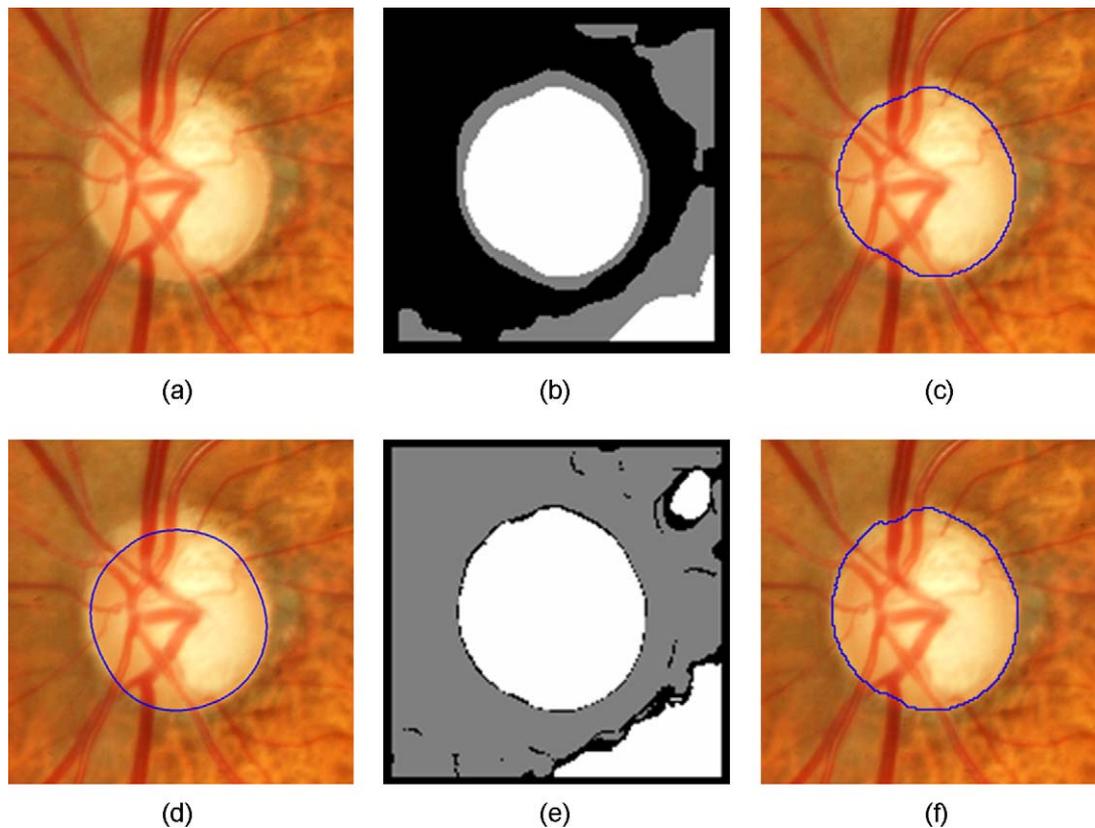


Fig. 4 – Disc outlines determined by the FCM method: (a) original image; (d) image with the gold standard; (b and e) clustering results by using two and three features, respectively; (c and f) the corresponding disc outlines determined based on (b) and (e).

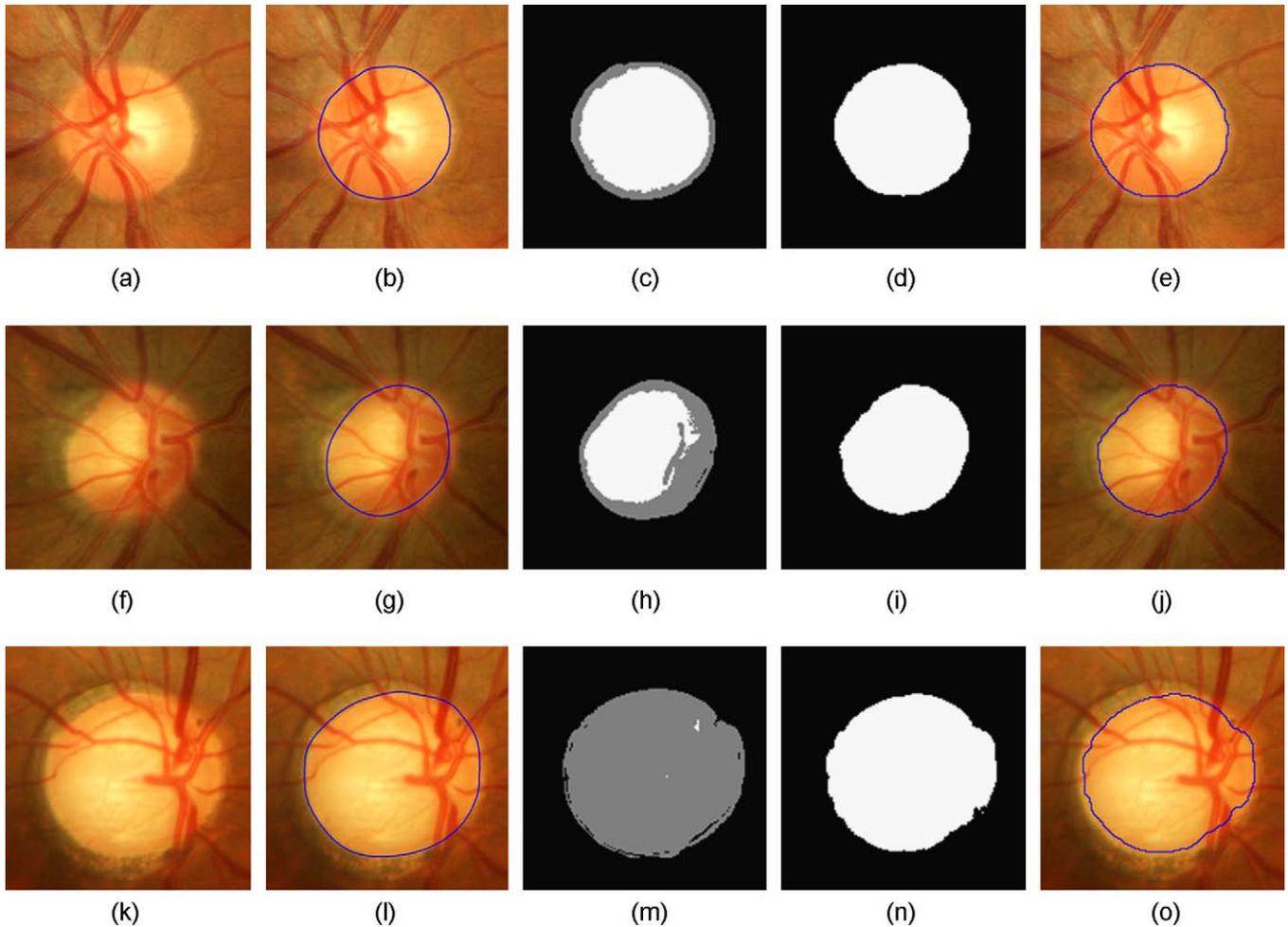


Fig. 5 – Disc outlines determined by the ANN method: (a, f and k) original images; (b, g and l) image with the gold standard; (c, h and m) classification results when the number of output units was three; (d, i and n) classification results when the number of output units was two; (e, j and o) images with the disc outlines determined by the ANN method.

may include a region of peripapillary chorioretinal atrophy (PPA). It can be seen that although this case was considered as diagnosable, the image is over-exposed, so that the texture inside the cup region cannot be observed, and therefore not of desirable quality.

4.2. Resubstitution test results on the training database

In most cases, the contrast in the red component was high near the disc border. Therefore, the approximate disc border was detected by the Canny edge detector. However, for some cases, especially those with PPA, there was no clear edge, and the PPA regions were often included as the part of the disc region. Fig. 3 shows the ROIs of original images, images with the gold standard, the red channel of the blood-vessel-erased images, the edges detected by the Canny edge detector, and the images with the outlines of the optic disc determined by the ACM. The average agreement of the disc regions based on the gold standard and the segmentation result by the ACM for the 80 training cases was 0.882.

By using the FCM method, it was found that the best segmentation result for the 80 training cases was obtained when the pixels were classified into three clusters. A number of different combinations of the features were tested. As the result, two features were employed for the input data: the median pixel value in the red component of the original images and the mean pixel value in the red component of the blood-vessel-erased images in the surrounding 15×15 pixels. The inclusion of the edge information was useful in some cases, however detrimental in others; therefore, the overall performance was not improved by the inclusion of the edge feature. Fig. 4(b) shows an image specifying the pixels classified into the three clusters (black, gray and white) using the proposed method. For comparison, the result when the edge information is added to the two features is shown in Fig. 4(e). Inclusion of the other color components (green or blue) and contrast did not improve the classification result. The average disc region agreement for the 80 training cases with the FCM method was 0.837. In the FCM method, the pixels corresponding to PPA regions were often classified into the disc cluster. Therefore, the disc regions tend to be overestimated for cases with PPA.

Table 4 – Disc determination results on the training dataset by the three methods.

	ACM method		FCM method		ANN method	
	Disc region agreement	Error in vertical length (%)	Disc region agreement	Error in vertical length (%)	Disc region agreement	Error in vertical length (%)
Glaucoma	0.853	4.2	0.812	6.9	0.836	4.9
Normal	0.895	4.7	0.848	6.7	0.879	5.3
All	0.882	4.5	0.837	6.8	0.866	5.2

ACM: active contour model; FCM: fuzzy c-means; ANN: artificial neural network.

On the basis of the results for the 80 training cases, three features were selected as input data for the ANN: the pixel values in the red component of the original and blood-vessel-erased images and the presence of the edges in the surrounding 3×3 pixel area detected by the Canny edge detector. ANN was better trained with two output units than with three output units. When the ANN was trained with three output units, the classification results were varied as shown in Fig. 5. In some cases as in Fig. 5(c), pixels were classified correctly to those corresponding to the disc, disc border, and retina. However, in some cases as in Fig. 5(h), pixels corresponding to the rim were classified to the disc border group, while, in some other cases as in Fig. 5(m), pixels corresponding to the disc were classified to the disc border group. Therefore, the result was stable when the number of output units was two as shown in Fig. 5(d), (i), and (n). The number of hidden units for this combination of the input and output data was set at three; however, the results were comparable when the number of hidden units was changed to two or four. The results were similar in the way that the regions corresponding to the PPA were often misclassified as disc regions because

of their high-intensity values. The average disc agreement for the 80 cases was 0.866, which was slightly higher than that of the FCM method but slightly lower than that of the ACM method.

The resubstitution test results for the training dataset by using the three methods are summarized in Table 4. For the determination of the CD ratio, vertical lengths of the cup and disc regions are usually measured; therefore, average errors in the largest vertical lengths of the disc regions are calculated and are shown in Table 4. It should be noted that although the agreement is lower, the error rate may be smaller for glaucoma cases than for normal cases because PPA is often present in the temporal side of the optic disc and may not affect the measurement of the vertical length. In most cases, especially those without PPA, the results of the three methods were comparable. However, in some cases, the results of the ACM method were better because the edges between the disc and PPA were partially detected and utilized effectively by this method. On the other hand, the presence of edges could provide inappropriate results. The disc outlines determined by the three methods for cases without and with PPA are shown in Fig. 6.

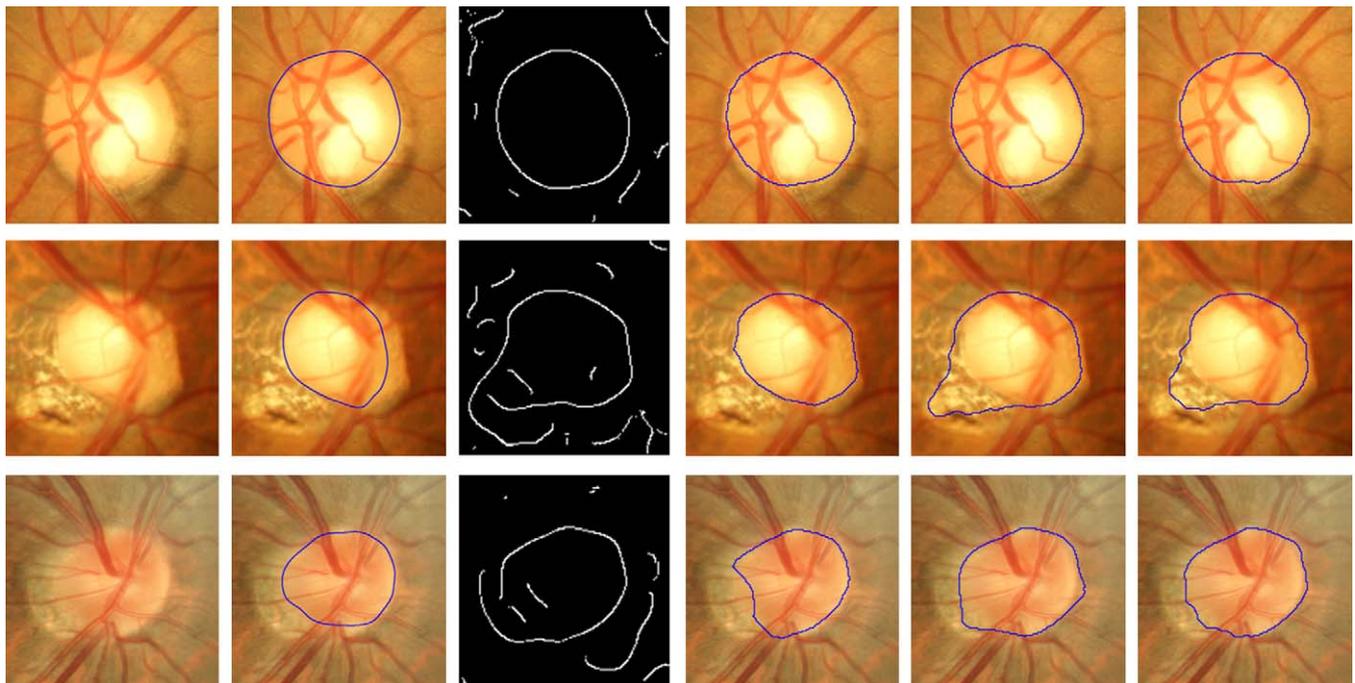


Fig. 6 – Comparison of the disc outlines determined by the three methods. From left to right columns: the original images, images with the gold standard, edges detected by the Canny edge detector, outlines determined by using the ACM, FCM, and ANN.

Table 5 – Disc determination results on the test datasets by the three methods.

	ACM method		FCM method		ANN method	
	Disc region agreement	Error in vertical length (%)	Disc region agreement	Error in vertical length (%)	Disc region agreement	Error in vertical length (%)
Test dataset I						
Glaucoma	0.871	3.9	0.852	6.3	0.868	4.8
Normal	0.901	4.6	0.883	4.8	0.898	4.4
All	0.882	4.2	0.864	5.7	0.879	4.7
Test dataset II						
Glaucoma	0.798	5.2	0.752	15	0.850	4.5
Normal	0.885	5.7	0.880	5.7	0.895	4.9
All	0.871	5.5	0.858	7.2	0.887	4.7

ACM: active contour model; FCM: fuzzy c-means; ANN: artificial neural network.

4.3. Validation test results on the test datasets

The methods trained by using the training database were tested on the test datasets without any modification or adjustment of the parameters. The approximate localization of the optic disc was successful, i.e., the centroid of the detected disc region was located inside the gold standard disc region, for all cases (100%) for test dataset I. However, the localization completely failed in one case included in test dataset II. The failure in disc localization was likely due to overexposure, so that the pixel values in the red component were saturated around the disc region (Fig. 2(e)). There were two such cases in test dataset II. Although the approximate disc region could be identified to some extent accidentally in the one of these cases, they were excluded for the evaluation of the segmentation methods. In another case, the identified centroid was at the disc border; this was due to the presence of PPA. This case was included for the subsequent process; therefore, the disc segmentation methods were evaluated using 28 cases for test dataset II (93%).

The agreements of the gold standard disc regions and the regions determined by the three methods were comparable to those of the resubstitution test results in the training dataset, and are shown in Table 5. The average agreement was slightly lower for the glaucoma cases in test dataset II by the FCM method. In one case, the disc region was overestimated on the side of PPA, and the agreement was 0.50; this affected strongly on the average agreement because it was based only on five cases. The results indicate that the accuracies of disc determination of the ACM and the ANN methods were comparable and slightly higher than that of the FCM method. Although the PPA regions were often misclassified as part of the disc region by all three methods, the agreements were high for those cases without PPA. Overall, the methods can be applied successfully for new unknown cases including those imaged by different camera systems.

5. Discussion

The segmentation of the optic disc regions using the three different methods was tested on the independent datasets, which included images captured by the different camera systems. An optic disc region can be usually identified as a bright region in a retinal fundus image. However, PPA regions also appear with high-intensity pixel values and are located next

to the optic disc. As a result, in the present study, they were often included in the disc regions determined by the three methods. Previous studies [6–9,11–14] on the segmentation of disc regions have not discussed PPA; however, it may be one of the risk factors for glaucoma and can often be seen in the target age groups for glaucoma screening [22]. Therefore, segmentation methods must be improved to accommodate such cases.

In this study, there were 21, 24, and 4 cases with mild to severe PPA in the training dataset, test dataset I, and test dataset II, respectively. Disc segmentation by ACM method was performed on the basis of the disc edges detected by the Canny edge detector. When the edges between the disc and PPA regions were detected successfully, the disc region could be determined adequately, and the result may be better for ACM method than those obtained by the other two methods. Some PPA regions could be recognized by their mesh-like texture in the green or blue component as shown in Fig. 7 (green arrow). By using ANN and including features such as the standard deviation in pixel values or the presence of fine edges in different color components, it was expected that some of the PPA regions could be distinguished from the disc regions; how-



Fig. 7 – The green channel of the image with the PPA region.

ever, these features were not useful for selectively excluding PPA regions from the disc regions. This may be due to the fact that the pixels used in the training of the ANN were selected randomly and the number of pixels corresponding to the PPA regions was small. In the future, if the gold standard of the PPA regions is established, and the pixels corresponding to such regions are trained, the result may improve.

In this study, the overall disc segmentation results using ACM and ANN were comparable and slightly better than that of FCM. An open-access database for the evaluation of segmentation algorithms is desirable, through which the different methods by different research groups could be evaluated; however, in the absence of such a database, we evaluated three different methods using unseen images captured by different camera systems. Comparable results for the test datasets and those on the training dataset indicated that the proposed methods can be applied successfully to unknown cases. Although the segmentation of optic disc regions should be improved to the level of the inter-reader agreement, especially for cases with PPA regions, the proposed methods can be useful for the measurement of vertical lengths, which can then be used for the determination of CD ratios. The usefulness of the proposed methods must be further evaluated when an automated cup determination method is established.

Acknowledgements

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, M.D., K. Ishida, M.D., and R. Shiraki, M.D. for their contribution.

REFERENCES

- [1] S. Resnikoff, D. Pascolini, D. Etya'ale, L. Kocur, R. Pararajasegaram, G.P. Pokharrel, S. Mariotti, Global data on visual impairment in the year 2002, *Bull. WHO* 82 (2004) 844–851.
- [2] H.A. Quigley, A.T. Broman, The number of people with glaucoma worldwide in 2010 and 2020, *Br. J. Ophthalmol.* 90 (2006) 262–267.
- [3] R. Varma, G.L. Spaeth, W.C. Steinmann, L.J. Katz, Agreement between clinicians and an image analyzer in estimating cup-to-disc ratios, *Arch. Ophthalmol.* 107 (1989) 526–529.
- [4] J.M. Tielsch, J. Katz, H.A. Quigley, N.R. Miller, A. Sommer, Intraobserver and interobserver agreement in measurement of optic disc characteristics, *Ophthalmology* 95 (1988) 350–356.
- [5] T. Nakagawa, T. Suzuki, Y. Hayashi, Y. Mizukusa, Y. Hatanaka, K. Ishida, T. Hara, H. Fujita, T. Yamamoto, Quantitative depth analysis of optic nerve head using stereo retinal fundus image pair, *J. Biomed. Opt.* 13 (2008), 064026-1-10.
- [6] T. Walter, J.-C. Klein, Segmentation of color fundus images of the human retina: detection of the optic disc and the vascular tree using morphological techniques, *Int. Symp. Med. Data Anal.* 2199 (2001) 282–287.
- [7] M. Lalonde, M. Beaulieu, L. Gagnon, Fast and robust optic disc detection using pyramidal decomposition and Hausdorff-based template matching, *IEEE Trans. Med. Imaging* 11 (2001) 1193–1200.
- [8] H. Li, O. Chutatape, Automated feature extraction in color retinal images by a model based approach, *IEEE Trans. Biomed. Eng.* 51 (2004) 246–254.
- [9] A. Osareh, M. Mirmehdi, B. Thomas, R. Markham, Comparison of colour spaces for optic disc localization in retinal images, in: *Proc. Int. Conf. Pattern Recog.*, 2002, pp. 743–746.
- [10] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, E. Fletcher, Optic nerve head segmentation, *IEEE Trans. Med. Imaging* 23 (2004) 256–264.
- [11] M.B. Merickel, X. Wu, M. Sonka, M. Abramoff, Optimal segmentation of the optic nerve head from stereo retinal images, *Proc. SPIE Med. Imaging* 6243 (2006), 61433B-1-2.
- [12] J. Xu, H. Ishikawa, G. Wollstein, R.A. Bilonick, K.R. Sung, L. Kagemann, K.A. Townsend, J.S. Schuman, Automated assessment of the optic nerve head on stereo disc photographs, *Invest. Ophthalmol. Vis. Sci.* 49 (2008) 2512–2517.
- [13] D.W.K. Wong, J. Liu, J.H. Lim, X. Jia, F. Yin, H. Li, T.Y. Wong, Level-set based automatic cup-to-disc ratio determination using retinal fundus images in ARGALI, *Proc. IEEE Eng. Med. Biol. Soc.* 2008 (2008) 2266–2269.
- [14] M.D. Abramoff, W.L.M. Alward, E.C. Greenlee, L. Shuba, C.Y. Kim, J.H. Fingert, Y.H. Kwon, Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features, *Invest. Ophthalmol. Vis. Sci.* 48 (2007) 1665–1673.
- [15] H. Fujita, Y. Uchiyama, T. Nakagawa, D. Fukuoka, Y. Hatanaka, T. Hara, G.N. Lee, Y. Hayashi, Y. Ikedo, X. Gao, X. Zhou, Computer-aided diagnosis: the emerging of three CAD systems induced by Japanese health care need, *Comput. Methods Prog. Biol.* 92 (2008) 238–248.
- [16] J.R. Parker, Basics of grey-level segmentation, in: *Algorithms for Image Processing and Computer Vision*, Wiley Computer Publishing, New York NY, 1997, pp. 116–127.
- [17] J. Serra, Introduction to mathematical morphology, *Comput. Vis. Graph Image Process* 35 (1986) 283–305.
- [18] N. Otsu, A threshold selection method from gray-level histogram, *IEEE Trans. Syst. Man Cybern.* 9 (1979) 62–66.
- [19] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, *Int. J. Comput. Vis.* 1 (1988) 321–331.
- [20] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Kluwer Academic Publishers, Norwell, MA, 1981.
- [21] D.F. Garway-Heath, D. Poinosawmy, G. Wollstein, A. Viswanathan, D. Kamal, L. Fontana, R.A. Hitching, Inter- and intraobserver variation in the analysis of optic disc images: comparison of the Heidelberg retina tomography and computer assisted planimetry, *Br. J. Ophthalmol.* 83 (1999) 664–669.
- [22] G. Tomais, G. Georgopoulos, C. Koutsandrea, M. Moschos, Correlation of central corneal thickness and axial length to the optic disc and peripapillary atrophy among healthy individuals, glaucoma and ocular hypertension patients, *Clin. Ophthalmol.* 2 (2008) 981–988.