35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Self-Assessment For Optic Disc Segmentation

Jun Cheng, Jiang Liu, Fengshou Yin, Beng-Hai Lee, Damon Wing Kee Wong,
Tin Aung, Ching-Yu Cheng and Tien Yin Wong

***Abstract*— Optic disc segmentation from retinal fundus image is a fundamental but important step in many applications such as automated glaucoma diagnosis. Very often, one method might work well on many images but fail on some other images and it is difficult to have a single method or model to cover all scenarios. Therefore, it is important to combine results from several methods to minimize the risk of failure. For this purpose, this paper computes confidence scores for three methods and combine their results for an optimal one. The experimental results show that the combined result from three methods is better than the results by any individual method. It reduces the mean overlapping error by 7.4% relatively compared with best individual method. Simultaneously, the number of failed cases with large overlapping errors is also greatly reduced. This is important to enhance the clinical deployment of the automated disc segmentation.**

I. Introduction

The optic disc (in short, disc) is the location where ganglion cell axons exit the eye to form the optic nerve. Segmentation of disc is very important in many applications such as automated glaucoma diagnosis. The process of segmentation is to find the disc boundary. In this paper, we focus on this problem for automated glaucoma diagnosis from 2D retinal fundus images.

Optic disc segmentation is a challenging task due to blood vessel occlusions, pathology around disc, imaging conditions, etc. Previous approaches for disc segmentation can be generally classified as template based approaches [1][2], deformable model based approaches [3][4] and classification based approaches [5]. In [1], circular Hough transform is used to model the disc because of its computational efficiency. However, clinical studies have shown that the disc has a slightly oval shape with vertical diameter being about 7%-10% larger than horizontal one [6]. Therefore an elliptical Hough transform (EHT) based approach is proposed [2]. Deformable models are sensitive to poor initialization. Very often, the deformation cannot exclude peripapillary atrophy (PPA) from the segmented OD if it has been included in the initialization. The template and deformable model based methods are based on edge characteristics. Their performance very much depend on the differentiation of edges from the OD and other structures especially the PPA. Classification based methods use various features such as intensity, texture, etc. from each pixel and its surroundings to get the disc.

This work was supported in part by the Agency for Science, Technology and Research, Singapore, under SERC grant 092-148-00731.

J. Cheng, J. Liu, F. Yin, B.-H. Lee and D. W. K. Wong are with iMED Ocular Imaging Programme, Institute for Infocomm Research, A*Star, Singapore. T. Aung, C. Y. Cheng and T. Y. Wong are with Singapore Eye Research Institute and National University of Singapore.

Fig. 1. Disc Segmentation.

Their performance were similar to deformable model based methods [7]. Recently, we proposed superpixel classification based approach [8]. It is a combination of superpixel classification and deformable models. Experimental results show some improvement compared with prior art. However, it might fail for images with irregular appearance.

Although many methods have been proposed for disc segmentation, there is still a large room for improvement because: 1) the appearance of disc varies largely from one to another 2) each method has a limitation on certain types of images and there is no method to take the advantages of all. In practice, it is difficult to have one method covering all scenarios. Motivated by this, we propose confidence based optic disc segmentation to get an optimal result from several methods. For this purpose, we compute self-assessment scores as confidence for different segmentation methods and combine them empirically. Self-assessment is an important index that has seldom been discussed previously. In practice, one method may work well for certain types of images and work poorly for other types. Since different methods usually have different set of poor results, it is possible to combine results from several methods for an optimal one.

The rest of the paper is organized as follows. In Section II, we introduce confidence scores for self-assessment of three different disc segmentation methods together with a brief review of the disc segmentation methods. Section III introduce a simple method to combine the results from several method for an optimal result. Section IV shows the experimental results by the proposed method. Conclusions are given in the last section.

II. Disc Segmentation with Self-Assessment

A. Active Shape Model based Approach

We first give a brief introduction of the active shape model (ASM) [4] approach before we introduce the proposed confidence score for this approach. In this method, Canny edge is first extracted from the image. Then an initial boundary is then determined by a circular Hough transform over the Canny edges in the image. After that, a series of m landmark points, $x_i, i = 1, \cdots, m$, are sampled evenly from the initial boundary. The mean and variance of the 2D

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

shapes from training samples are computed as

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1}$$

$$\sigma_x = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})^T \tag{2}$$

By applying principal component analysis, a shape can be approximated by the mean shape $\bar{x}$ and major eigenvectors ϕ:

$$x \approx \bar{x} + \phi b \tag{3}$$

Starting from the mean shape, the models are fitted in an iterative manner. Each model point is moved toward the direction perpendicular to the contour. The new landmark position can be obtained by minimizing the Mahalanobis distance:

$$f(g_i) = (g_i - \bar{g})^T S_g^{-1}(g_i - \bar{g}) \tag{4}$$

where g_i is the normalized derivative profiles of the i^{th} landmark point, $\bar{g}$ is the mean profile and S_g is the covariance matrix. The updated segmentation can be obtained after all the landmarks are moved to new positions. This process is repeated by a specified number of times at each resolution, in a coarse-to-fine fashion. Due to the influence of the blood vessel and PPA, the boundary detected is often non-accurate. A direct least squared ellipse fitting is used to smooth the boundary of the contour.

In order to compute a self-assessment confidence score, we compute the distance between the final landmark points $\hat{x}_i$ and the edge points in the image:

$$d = \frac{1}{m}\sum_{i}^{m} \inf\{d(\hat{x}_i, y)|y \in Y\} \tag{5}$$

where inf represents the infimum, $d(x, y)$ the Euclidean distance between points x and y and Y the set of all edge points. Intuitively, the larger distance d, the less accurate the result. The self-assessment of the ASM approach is computed as

$$r_1 = 1 - \alpha \cdot d, \tag{6}$$

where α is set to be the reciprocal of the maximum d.

B. Superpixel Classification

Superpixel Classification (SC) [8] replaces circular Hough transform based initial contour with a smoothed contour obtained from superpixel classification. In the approach, each image is divided into superpixels through over-segmentation using a simple linear iterative clustering approach [9]. Then two types of feature are extracted. The first type is the histogram extracted from each superpixel from five channels including red, green, blue in RGB space and hue, saturation in HSV space. The second type of feature is the center surround statistics computed by center surround operation over red, green and blue channels. A support vector machine (SVM) is used as the classifier to classify each superpixel from test images as disc or non-disc. Instead of directly using the binary classification results from SVM, the output values from the SVM decision function are used. The output value for each superpixel is used as the decision values for all pixels in the superpixel. A smoothing filter is then applied on the decision values to achieve smoothed decision values. The smoothed decision values are then used to obtain the binary decisions for all pixels. The largest connected object is obtained and its boundary is used as the raw estimation of the disc boundary. The active shape model in Section II-A is used to fine tune the boundary.

Fig. 2. Confidence Score Computation: the raw estimation (red) and the fitted estimation (white)

The self-assessment for the SC approach is computed based on the raw estimation. An ellipse fitting using elliptical Hough transform [2] is applied to get a fitted estimation. As the disc is often close to an ellipse, the obtained boundary before and after ellipse fitting should be close if the superpixel based segmentation gets a result close to actual boundary. Inspired by this, we compute the self-assessment score based on the difference between the raw and fitted estimations. Define the set of points from the raw estimation as X and the set of points from the fitted estimation as $Y = f(X)$, e.g., the red and white lines in Fig. 2, respectively. For each point x in X, we find its nearest point in Y and their distance is computed as

$$d_f(x) = \inf\{d(x, y)|y \in Y\} \tag{7}$$

Then, the self-assessment score is computed as the ratio of the number of x with $d_f(x) < T$ to the total number of x, i.e,

$$r_2(X) = \frac{Card(\{x|d_f(x) < T, x \in X\})}{Card(X)}, \tag{8}$$

where $Card(Z)$ is the cardinality of the set Z, and T is a threshold empirically set as five pixels based on the typical size of disc with disc diameter around 350 pixels.

C. Elliptical Hough Transform based Approach

The elliptical Hough transform (EHT) method [2] approximates the disc as an ellipse with parametric representation given by:

$$\begin{aligned} x(t) &= x_c + a\cos t\cos\phi - b\sin t\sin\phi \\ y(t) &= y_c + a\cos t\sin\phi + b\sin t\cos\phi \end{aligned}, \tag{9}$$

where $t \in [0, 2\pi]$, (x_c, y_c) is row and column coordinate of the center, a and b are the vertical and horizontal radius respectively, and ϕ is the rotation angle of the ellipse. The main procedures are summarized as follows:

1) Set parameter (a, b, ϕ) for ellipse.
2) For each edge point (x_e, y_e), draw an ellipse centered at (x_e, y_e) with (a, b, ϕ) and increment all coordinates

that the perimeter of the ellipse passes through in the accumulator A corresponding to the parameters.

3) Update $(a,\ b,\ \phi)$ and repeat step 2 for all $(a,\ b,\ \phi)$ from the parameter space.
4) Find the maximum value in A to get an ellipse centered at (x_1, y_1) and corresponding parameters (a_1, b_1, ϕ_1).

The self-assessment of the EHT based approach is computed based the maximum value in the accumulator A:

$$r_3 = \max(A) \tag{10}$$

III. Fusion

In this paper, we propose a simple algorithm to combine several results for an optimal one. Assuming that we have K algorithms with output disc D_i and confidence score r_i, $i = 1, 2, \cdots K$. The fusion algorithm is as follows:

Let $i = 0, r = 0, T_0 = +\infty$;
while $i + 1 \leq K$ and $r < T_i$ **do**
  $i = i + 1$;
  Compute D_i and r_i;
  Let $D = D_i, r = r_i$;
end while
if $r < T_i$ **then**
  Let D be the one obtained by the most reliable method;
end if

In the above, T_i is the confidence level to get an confident output for algorithm i. Therfore, the disc is determined by the first method with confident output ($r_i \geq T_i$) or the most reliable one otherwise.

IV. Experimental Results

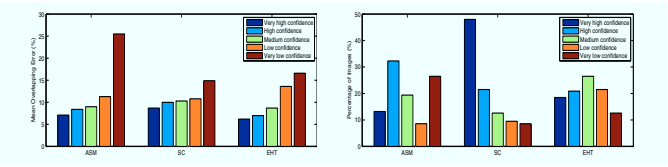
In this paper, a database with 650 images is used for evaluation. PPA are present in 230 of the 650 images. The disc boundaries of these images have been marked by trained professionals manually. The 650 images are randomly divided into 325 images for training and 325 images for testing.

The overlapping error E is computed as the evaluation metric:

$$E = 1 - \frac{Area(S \cap G)}{Area(S \cup G)}, \tag{11}$$

where S and G denote the segmented and the manual ground truth disc respectively.

First, we compute the accuracy of images at different confidence levels to show the effectiveness of the scores. In this paper, we empirically divided the confidence scores into five intervals: very high confidence, high confidence, medium confidence, low confidence and very low confidence. The mean overlapping error of the images at these confidence levels as well as their numbers are computed and shown in Fig. 3 As performed tests that show the SC method gives the best overall performance, we use it as the most reliable method. It is important to set thresholds for different methods to define confident output. In this paper, the acceptable confidence levels for ASM, SC and EHT are empirically set as: very high confidence, medium confidence, and high confidence, respectively based on the fact that SC performs



(a) Mean overlapping error (b) Percentages of images

Fig. 3. Performance of individual methods at different confidence levels

TABLE I
Mean overlapping error μ_E

	ASM [4]	EHT [2]	SC [8]	Proposed
w/o PPA	10.1%	9.3%	9.1%	8.2%
with PPA	13.5%	12.2%	10.3%	9.8%
All	11.3%	10.3%	9.5%	8.8%

better than EHT and EHT performs better than ASM. Table I shows the mean overlapping error for images with and without PPA as well as the overall performance by three individual methods and the proposed method. The results show that the proposed method achieves a better result than any individual method. In addition, the number of large errors is also an important criterion as it reflects the number of failed cases. According to our experience, a reasonable threshold to determine acceptable result can be set at around $T = 20\%$. Table II shows the numbers of cases greater than T with T set at 15%, 20%, and 25%. It shows that the proposed method reduces the number of large errors by 28.6%, 21.7%, and 9.8% relatively at the three different levels compared with best individual method. Fig. 4 shows five sample results by various methods.

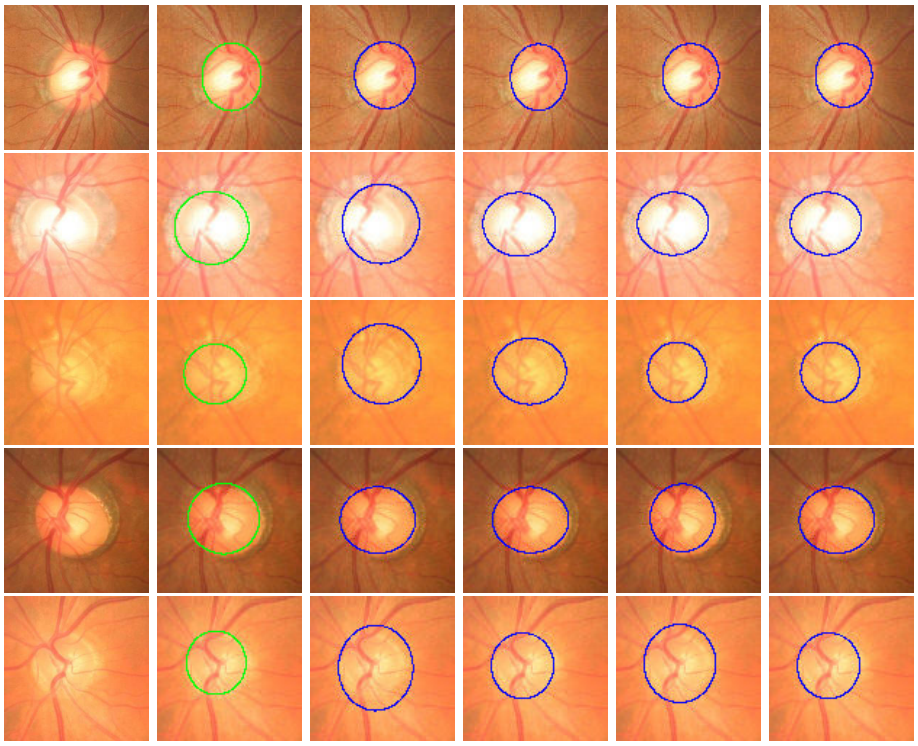
As one of our objectives of disc segmentation is to compute the cup to disc ratio (CDR) for glaucoma diagnosis, we further evaluate how it benefits the CDR computation. The ASM, EHT, SC and the proposed method are used to obtain the disc and compute the vertical disc diameter (VDD). Then, the optic cup segmentation method in [10] is used to segment the cup from the segmented disc to compute the vertical cup diameter (VCD). The CDR is computed as $CDR = VCD/VDD$. The CDR error δ is computed as

$$\delta = |CDR_{GT} - CDR|, \tag{12}$$

where CDR_{GT} denotes the manual CDR from trained professionals. Table III shows the percentage of images at different δ intervals as well as the mean CDR error μ_δ when different disc segmentation methods are used. By using discs determined by the fusion, the CDR error drops by (0.094-0.090)/0.094=4.3% compared with the best prior.

TABLE II
Number of large errors above different levels

	ASM [4]	EHT [2]	SC [8]	Proposed
$E > 25\%$	34	19	14	10
$E > 20\%$	43	28	23	18
$E > 15\%$	62	50	41	37



(a) Original (b) Manual (c) ASM (d) EHT (e) SC (f) Proposed

Fig. 4. Sample results. From left to right: the original images, manual ground truth, outlines by the ASM [4], EHT [2], SC [8] and the proposed method.

TABLE III

PERCENTAGE OF IMAGES PER CDR ERROR δ INTERVAL AS WELL AS THE MEAN ERROR μ_δ

Method	$\delta \leq 0.05$	$\delta \leq 0.10$	$\delta \leq 0.15$	$\delta \leq 0.20$	μ_δ
ASM [4]	30%	56%	76%	89%	0.104
EHT [2]	34%	62%	79%	89%	0.096
SC [8]	35%	64%	82%	90%	0.094
Proposed	37%	66%	82%	90%	0.090

V. CONCLUSIONS

Automated disc segmentation is a fundamental but necessary step for developing automated glaucoma diagnosis systems. Because of large variation in appearance and noise affecting the OD segmentation, it is difficult to have a single method that works well for all images. In this paper, we compute self-assessment scores for three different methods and combine them for an optimal result. The experimental results show a significant reduction of both the mean overlapping errors and the number of large errors compared with any individual method. More importantly, the fusion algorithm can also be extended for future algorithms for further improvement. The limitation of the approach is that it requires some heuristics to determine the confidence level of a method.

REFERENCES

[1] A. Aquino, M. Gegundez-Arias, and D. Marin, "Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques," *IEEE Trans. Med. Imag.*, vol. 29, pp. 1860–1869, 2010.

[2] J. Cheng, J. Liu, D. W. K. Wong, F. Yin, C. Cheung, M. Baskaran, T. Aung, and T. Y. Wong, "Automatic optic disc segmentation with peripapillary atrophy elimination," *Int. Conf. of IEEE Eng. in Med. and Bio. Soc.*, pp. 6624–6627, 2011.

[3] J. Xu, O. Chutatape, E. Sung, C. Zheng, and P.C.T. Kuan, "Optic disk feature extraction via modified deformable model technique for glaucoma analysis," *Pattern Recognition*, vol. 40, pp. 2063–2076, 2007.

[4] F. Yin, J. Liu, S. H. Ong, Y. Sun, D. W. K. Wong, N. M. Tan, C. Cheung, M. Baskaran, T. Aung, and T. Y. Wong, "Model-based optic nerve head segmentation on retinal fundus images," *Int. Conf. of IEEE Eng. in Med. and Bio. Soc.*, pp. 2626–2629, 2011.

[5] M. D. Abramoff, W. L. M. Alward, E. C. Greenlee, L. Shuba, C. Y. Kim, J. H. Fingert, and Y. H. Kwon, "Automated segmentation of theoptic disc from stereo color photographs using physiologically plausible features," *Invest. Ophthalmol. Vis. Sci.*, vol. 48, pp. 1665–1673, 2007.

[6] J. B. Jonas, W. M. Budde, and S. Panda-Jonas, "Ophthalmoscopic evaulation of the optic nerve head," *Surv. Ophthalmol.*, pp. 293–320, 1999.

[7] C. Muramatsu, T. Nakagawa, A. Sawada, Y. Hatanaka, T. Hara, T. Yamamoto, and H. Fujita, "Automated segmentation of optic disc region on retinal fundus photographs: Comparison of contour modeling and pixel classification methods," *Computer Methods and Programs in Biomedicine*, vol. 101, pp. 23–32, 2011.

[8] J. Cheng, J. Liu, Y. Xu, F. Yin, D. W. K. Wong, N. M. Tan, C. Y. Cheung, Y. C. Tham, and T. Y. Wong, "Superpixel classification based optic disc segmentation," *Asian Conference on Computer Vision*, 2012.

[9] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pat. Ana. Mach. Intel.*, vol. 34, no. 11, pp. 2274–2281, 2012.

[10] F. Yin, J. Liu, D. W. K. Wong, N. M. Tan, C. Cheung, M. Baskaran, T. Aung, and T. Y. Wong, "Automated segmentation of optic disc and optic cup in fundus images for glaucoma diagnosis," *IEEE Int. Symp. on Computer-Based Medical Systems*, pp. 1–6, 2012.