34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Efficient Optic Cup Localization Using Regional Propagation Based on Retinal Structure Priors*

Yanwu Xu[1], Jiang Liu[1], Jun Cheng[1], Fengshou Yin[1], Ngan Meng Tan[1], Damon Wing Kee Wong[1], Mani Baskaran[2], Ching Yu Cheng[2,3] and Tien Yin Wong[2,3]

Abstract— **We present a regional propagation approach based on retinal structure priors to localize the optic cup in 2D fundus images, which is the primary image component clinically used for identifying glaucoma. This method provides three major contributions. First, it proposes processing of the fundus images at the superpixel level, which leads to more descriptive and effective features than those employed by pixel based techniques, without additional computational cost. Second, the proposed approach does not need manually labeled training samples, but uses the structural priors on relative cup and disc positions. Third, a refinement scheme that utilizes local context information is adopted to further improve the accuracy. Tested on the *ORIGA-light* clinical dataset, which comprises of 325 images from a population-based study, the proposed method achieves a 34.9% non-overlap ratio with manually-labeled ground-truth and a 0.104 absolute cup-to-disc ratio (CDR) error. This level of accuracy is much higher than the state-of-the-art pixel based techniques, with a comparable or even less computational cost.**

I. Introduction

Glaucoma affects about 60 million people [1] and is responsible for approximately 5.2 million cases of blindness (15% of world total) [2]. It is the second leading cause of blindness (behind cataracts), with a mean prevalence of 2.4% for all age groups and 4.7% for ages 75 years and above [3]. Glaucoma unfortunately cannot be cured because the damage to the optic nerve cannot be reversed. Thus it is critical to detect this degeneration of the optic nerve as early as possible in order to stall its progression and and prevent the deterioration of vision [4]; however, more than 90% of the afflicted were unaware of their optical neurodegeneration [5], [6]. To facilitate widespread testing, much recent work has focused on computer-assisted glaucoma diagnosis techniques based on inexpensive and widely used digital color fundus images.

Among the structural image cues studied for glaucoma diagnosis, those based on the optic disc and cup are of particular importance. The optic disc is located where the ganglion nerve fibers congregate at the retina. The depression inside the optic disc where the fibers leave the retina via the optic nerve head (ONH) is known as the optic cup. The boundaries of the cup and disc structures need to be identified as it facilitates evaluation of glaucoma cues such as cup and disc asymmetry and large cup-to-disc ratio (CDR), defined as the ratio of the vertical cup diameter to the vertical disc diameter [7]. Typically, the CDR value is determined from a manually outlined optic disc and cup. But since manual annotation is labor intensive, researchers have sought automatic methods for disc and cup segmentation.

In previous work, researchers have mainly focused on automated segmentation of the optic disc [8], using various techniques such as intensity gradient analysis, Hough transforms, template matching, pixel feature classification, vessel geometry analysis, deformable models and level sets [9][10]. In this paper, we address the challenging problem of cup detection [8][11][12], using a large clinical dataset called *ORIGA-light* [13] in which the ground-truth optic discs and cups are assessed and annotated by a team of expert graders.

Previous cup segmentation algorithms are based on classifying pixels as part of the cup or rim (the disc area outside the cup) [8][12]. Unlike these methods, our approach identifies a cup via visual similarity-based label propagation at a local region scale (*i.e.*, superpixels), taking advantage of prior information on retinal structure to infer confident cup/rim labels of partial superpixels directly from the test image, without the need of a pre-learned classifier and training samples. Furthermore, the local context is utilized to refine the cup/rim labels for each superpixel.

This optic cup localization approach achieves significant improvement on cup localization accuracy, with a comparable or even less computational cost comparing with the current state-of-the-art optic cup segmentation methods. This work indicates much promise for developing practical automated/assisted glaucoma diagnosis systems with low-cost and widespread digital fundus cameras.

II. Regional Propagation Based Cup Localization

In this work, we localize the optic cup in a given disc image which may be obtained using segmentation methods such as [9]. As illustrated in Fig. 1, our method segments the input disc image into local regions (*i.e.*, superpixels), removes superpixels that corresponds to blood vessels, labels superpixels as the cup or rim based on structure priors, propagate the labels to remaining superpixels, refines the superpixel labels, and then determines a cup location by ellipse fitting.

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

[1]Y. Xu, J. Liu, J. Cheng, F. Yin, N.-M. Tan and D.W.K. Wong are with the Institute for Infocomm Research, Agency for Science, Technology and Research, 138632, Singapore {yaxu, jliu, jcheng, fyin, nmtan, wkwong} at i2r.a-star.edu.sg

[2]M. Baskaran, C.-Y. Cheng and T.-Y. Wong are with the Singapore Eye Research Institute, 168751, Singapore baskaran.mani at seri.com.sg, ching-yu_cheng at nuhs.edu.sg, tien_yin_wong at nuhs.edu.sg

[3]C.-Y. Cheng and T.-Y. Wong are also with the Department of Ophthalmology, National University of Singapore, 119074, Singapore

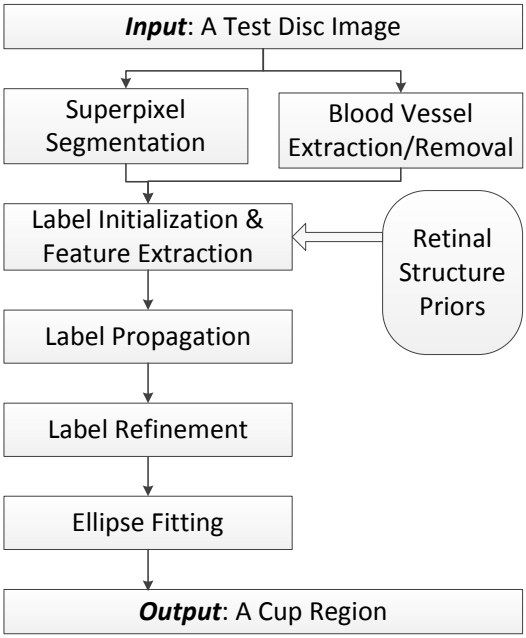

Fig. 1. Flowchart of the proposed cup localization approach using regional propagation based on retinal structure priors.

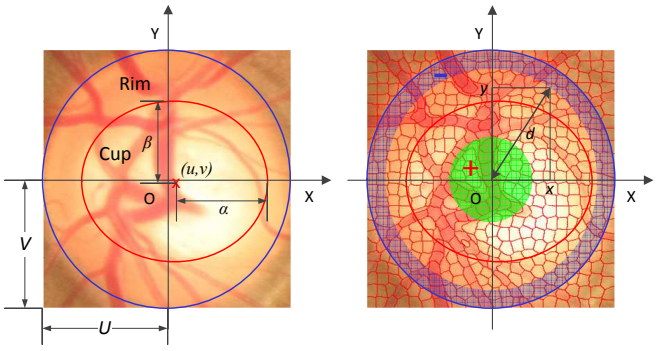

Fig. 2. Illustration of a fundus disc image segmented into 512 superpixels. Left: outline of rim and cup in the original disc image. Right: segmentation into superpixels.

A. Superpixel segmentation

The first step of the proposed approach is to divide the disc into local regions. Unlike dividing an image into a grid of regular patches, we instead segment the disc into superpixels based on the following considerations. First, superpixels are becoming increasingly popular in computer vision applications because of improved performance over pixel based methods. Second, superpixels have the important property of preserving local boundaries, as exemplified by the typical segmentation result shown in Fig. 2. Third, this segmentation can be processed rapidly. In this work, we utilize the state-of-the-art SLIC (Simple Linear Iterative Clustering) algorithm [14] to segment the fundus disc image into compact and nearly uniform superpixels. It takes only 21ms for a GPU implementation or 354ms for a CPU implementation on a 640×480 image [15].

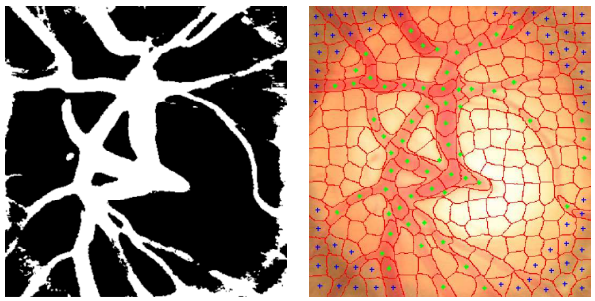

Fig. 3. Blood vessel extraction and removal. Left: blood vessel mask. Right: superpixels on blood vessels (marked with green dots) or out of disc (with blue dots).

B. Blood vessel removal

Since blood vessels appear approximately the same in both the rim and cup regions, many algorithms have been proposed for blood vessel extraction in retina images to avoid their effects on rim/cup labeling accuracy. In our case, extraction results need not be very precise, since images are processed at the level of superpixels. So we use the bottom-hat filtering algorithm [16], which trades off some precision for speed, to rapidly generate a rough blood vessel mask, and then identify superpixels that overlap the mask by at least 75% (Fig. 3). These superpixels, as well as those that lie outside the disc, are eliminated from further processing.

C. Feature representation for superpixels

The main step of the proposed approach is the superpixel label propagation. To assess the visual similarities among superpixels, we have to extract certain visual features from each superpixel. Various features have been used for modeling superpixels, including shape, location, texture, color, and thumbnail appearance [17].

In our application, only location and color information are relevant for cup/rim classification. For the i-th superpixel, we extract a feature vector $\mathbf{f}_i$ that consists of position information (denoted by (x_i, y_i, d_i) as shown in Fig. 2), mean RGB colors (r_i, g_i, b_i) and a 256-bin histogram (h_i^r, h_i^g, h_i^b) for each color channel. To avoid magnitude differences among the features, they are each normalized to the range of $[0, 1]$, with L_1-normalization of each histogram.

It is worth mentioning that no additional processing is needed to reduce the influence of illumination change among the images, since no learning process is involved in the proposed approach.

D. Superpixel label initialization and propagation

To label each superpixel as the cup or rim without learning procedure, we first initialize the labels of some superpixels by capitalizing on prior knowledge of retinal structure, and then propagate the labels to the remaining superpixels based on visual similarity.

For superpixel label initialization, as illustrated in the disc images of Fig. 2, a superpixel is essentially certain to lie in the rim region if it is very close to the disc outer boundary.

On the other hand, we can be assured that a superpixel exists within the cup region if it is very close to the disc center. Based on discussions with the expert graders, we have conservatively modeled this structural prior such that superpixels within 1/5 of the disc radius from the disc center (green area in Fig. 2) are considered to be definitely in the cup, while superpixels beyond 9/10 of the disc radius from the center (blue area) are definitely in the rim region.

With these labeled superpixels, the remaining superpixels are labeled by similarity based propagation. For the i-th superpixel which is unlabeled, its label l_i is defined as

$$l_i = \frac{1}{N}\sum_{j=1}^{N} s_{i,j} - \frac{1}{M}\sum_{k=1}^{M} s_{i,k}, \tag{1}$$

where $s_{i,j}$ denotes the similarity between i-th superpixel and the j-th superpixel labeled as cup (+1), similarly, $s_{i,k}$ denotes the similarity between i-th superpixel and the k-th superpixel labeled as rim (-1), and the similarity function is defined as

$$s_{i,j} = e^{-\frac{(f_i - f_{i,j})^2}{2\sigma_f^2}}, \tag{2}$$

in which σ_f controls the sensitivity to feature noise.

This technique not only avoids classifier training with manually labeled training samples, but also avoids any ad hoc image manipulations to conform the test data to the training data (*e.g.* illumination normalization), since the labeled and unlabeled superpixels are from *within the test image*.

E. Superpixel label refinement

To reduce superpixel labeling errors, we employ a refinement scheme that accounts for local context, *i.e.*, superpixel labels are filtered with respect to feature similarity among superpixels within a certain range (*i.e.*, the propagation range R_p, *e.g.*, 1/10 of the disc radius). This yields the final label l' of a superpixel

$$l'_i = l_i + \frac{1}{N}\sum_{t=1}^{N} l_{i,t} \cdot s_{i,t},\ \forall\ Dis(i,t) \le R_p, \tag{3}$$

where $l_{i,t}$ denotes the label of the t-th neighbor of the i-th superpixel, $s_{i,t}$ defined in Eq. (2) and $Dis(i,t)$ denotes the similarity and distance between the i-th superpixel and its t-th neighbor. In our experiments, $Dis(i,t)$ is defined as the Euclidean distance

$$Dis(i,t) = \sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}. \tag{4}$$

After obtaining the final labels of all superpixels, the minimum ellipse that encompasses all the superpixels with positive labels is computed to produce the detection result, represented by ellipse center/elongation parameters $(\hat{u}, \hat{v}, \hat{\alpha}, \hat{\beta})$.

III. Experiments

In this section, we describe the evaluation criteria, then evaluate our regional propagation based cup localization approach through an experimental comparison to state-of-the-art pixel based segmentation methods [9][12], using the *ORIGA-light* dataset comprised of 325 images. We also report how the algorithm parameters affect performance.

TABLE I

Performance comparison of our method and pixel based segmentation methods

Method	m_1	m_2	δ
regional propagation	**0.349**	**0.336**	**0.104**
Level-set [9]	0.495	0.847	0.162
Level-set+Hist-analysis [12]	0.476	0.702	0.140
Relative error reduction to [9]	29.5%	60.3%	35.8%
Relative error reduction to [12]	**26.7%**	**52.1%**	**25.7%**

A. Cup localization evaluation criteria

Three evaluation criteria are commonly used for cup localization/segmentation, namely non-overlap ratio (m_1), relative absolute area difference (m_2) [18] and absolute cup-to-disc ratio (CDR) error (δ), defined as:

$$\begin{aligned} m_1 &= 1 - \frac{area(E_{dt} \bigcap E_{gt})}{area(E_{dt} \bigcup E_{gt})}, \\ m_2 &= \frac{|area(E_{dt}) - area(E_{gt})|}{area(E_{gt})}, \\ \delta &= \frac{|D_{dt} - D_{gt}|}{DD} \end{aligned} \tag{5}$$

where E_{dt} denotes a detected cup region, E_{gt} denotes the ground-truth ellipse, D_{dt} is the vertical diameter of the detected cup, D_{gt} is the vertical diameter of the ground-truth cup, DD is the vertical diameter of optic disc which is normalized to 2, thus $0 < D_{dt}, D_{gt} \le 2$.

In fact, m_1 is most related to cup localization accuracy and δ is most related to glaucoma diagnosis; thus these two criteria are relatively more important.

B. Comparison to pixel based segmentation methods [9][12]

We compared our regional propagation based approach to state-of-the-art pixel based segmentation methods. One of the few methods for both cup and disc segmentation is the level-set method of [9] (referred to as *Level-set*), which first identifies the pixels that belong to the cup region, then uses a convex hull method to generate an ellipse. In [12] (referred to as *Level-set+Hist-analysis*), histogram based analysis of the color pixel intensity together with multiple method fusion are also employed to further improve cup detection accuracy.

Table I compares our method to these two level-set approaches. Compared with the more advanced approach [12], our method is shown to significantly improve cup localization accuracy in terms of m_1, m_2 and CDR error (δ), which are reduced by 26.7%, 52.1% and 25.7%, respectively. The results indicate the advantage of superpixel based features over pixel based methods, as also observed in previous work [14].

C. Influence of parameter settings on accuracy and speed

We examined the performance and stability of the proposed regional propagation based method with different parameter settings. For this approach, the number of superpixels (SP) is a key parameter that affects performance in terms of both accuracy and speed, and the propagation range (R_p) may also affect the performance. As shown in Fig. 4 and Table II, one can make the following observations:

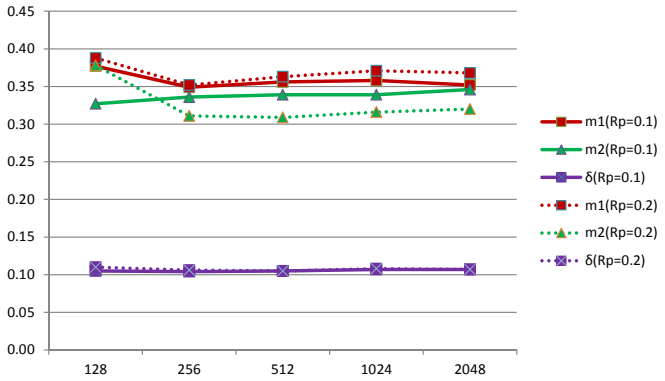

Fig. 4. Cup localization errors with different parameters.

TABLE II

PERFORMANCE COMPARISON OF OUR METHOD WITH DIFFERENT PARAMETERS

Parameters		**Evaluation criteria**		
SP	R_p	m_1	m_2	δ
128	0.1	0.377	0.327	0.105
128	0.2	0.388	0.378	0.110
256	0.1	**0.349**	0.336	**0.104**
256	0.2	0.352	0.311	0.106
512	0.1	0.356	0.339	0.105
512	0.2	0.363	**0.309**	0.105
1024	0.1	0.358	0.339	0.107
1024	0.2	0.371	0.316	0.108
2048	0.1	0.352	0.346	0.107
2048	0.2	0.368	0.320	0.107

- The errors become almost stable when $SP \geq 256$, with the lowest errors at $SP = 256$, in terms of m_1 and δ.
- When superpixels are too large (*i.e.*, smaller SP), the resulting under-segmentation can lead to ambiguously-labeled boundary superpixels that require further segmentation. When superpixels are too small (*i.e.*, larger SP), the resulting features computed from the over-segmented regions may become somewhat less distinctive, making it more difficult to infer the correct labels.
- In terms of m_1 and δ, little change in accuracy was found when R_p varies from 0.1 to 0.2; however, the errors become larger when R_p is set much larger (*e.g.*, 0.5). This can be expected, as long range context is less applicable for a superpixel and may introduce error.

Computation speed was also evaluated, using a four-core $3.4GHz$ PC with $12GB$ RAM. The overall processing time increases almost linearly with the number of superpixels (SP), since the most time consuming part is feature extraction for each superpixel. For 128, 256, 512, 1024 and 2048 superpixels, the computation takes 1.0, 1.7, 3.2, 6.5 and 20.2 seconds per image, respectively. When $SP \leq 256$, the speed is comparable to or even faster than that of pixel based segmentation [12], which costs about 1.5 seconds per image; however, our method is much more accurate.

IV. CONCLUSION

For cup localization and CDR assessment in glaucoma diagnosis, we proposed a regional propagation approach based on retinal structure priors. Tested on a large clinical dataset with three evaluation criteria, it achieves a 34.9% non-overlap ratio (m_1) with manually-labeled ground-truth, a 33.6% relative absolute area difference (m_2) and a 0.104 absolute CDR error (δ), deducting the errors significantly comparing with pixel based segmentation methods. In future work, we plan to elevate the accuracy by introducing other features and learning methods.

REFERENCES

[1] H.A. Quigley and A.T. Broman. The number of people with glaucoma worldwide in 2010 and 2020, *British Journal of Ophthalmology*, vol. 90, no. 3, pp. 262–7, 2006.

[2] B. Thylefors, and A. Negrel. The global impact of glaucoma, *Bull World Health Organ*, vol. 72, no. 3, pp. 323–6, 2006.

[3] EMBC2012 B. Klein., R. Klein, W. Sponsel, T. Franke, L. Cantor, J. Martone, and M. Menage. Prevalence of glaucoma: the beaver dam eye study, *Ophthalmology*, vol. 99, no. 10, pp. 1499–504, 1992.

[4] G. Michelson, S. Wärntges, J. Hornegger, and B. Lausen. The papilla as screening parameter for early diagnosis of glaucoma, *Dtsch. Arztebl. Int.*, vol. 105, no. 34-35, pp. 583–9, 2008.

[5] P. Foster, F. Oen, D. Machin, T. Ng, J. Devereux, G. Johnson, P. Khaw, and S. Seah. The prevalence of glaucoma in Chinese residents of Singapore: a cross-sectional population survey of the Tanjong Pagar district, *Arch Ophthalmology*, vol. 118, no. 8, pp. 1105–11, 2000.

[6] S. Shen, T.Y. Wong, P. Foster, J. Loo, M. Rosman, S. Loon, W. Wong, S.M. Saw, and T. Aung. The prevalence and types of glaucoma in Malay people: the Singapore Malay eye study, *Invest. Ophthalmol. Vis. Sci.*, vol. 49, no. 9, pp. 3846–51, 2008.

[7] J. Jonas, W. Budde, and S. Panda-Jonas. Ophthalmoscopic evaluation of the optic nerve head, *Survey of Ophthalmology*, vol. 43, pp. 293–320, 1999.

[8] M. Abramoff, W. Alward, E. Greenlee, L. Shuba, C. Kim, J. Fingert, and Y. Kwon. Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features, *Invest. Ophthalmol. Vis. Sci.*, vol. 48, no. 4, pp. 1665–73, 2007.

[9] J. Liu, D.W.K. Wong, J.H. Lim, H. Li,. N.M. Tan, Z. Zhang, T.Y. Wong, and R. Lavanya. ARGALI: an automatic cup-to-disc ratio measurement system for glaucoma analysis using level-set image processing, In *Int. Conf. IEEE Engin. in Med. and Biol. Soc.*, 2008.

[10] C. Li, C. Xu, C. Gui, and M. Fox. Level set evolution without re-initialization: A new variational formulation, In *IEEE Conf. Computer Vision and Pattern Recognition*, 2005, pp. 430–6.

[11] M. Merickel, X. Wu, M. Sonka, and M. Abramoff. Optimal segmentation of the optic nerve head from stereo retinal images, In *Medical Imaging: Physiology, Function, and Structure from Medical Images*, 2006.

[12] D.W.K. Wong, J.H. Lim, N.M. Tan, Z. Zhang, S. Lu, H. Li., M. Teo, K. Chan, and T.Y. Wong. Intelligent fusion of cup-to-disc ratio determination methods for glaucoma detection in ARGALI, In *Int. Conf. Engin. in Med. and Biol. Soc.*, 2009, pp. 5777–80.

[13] Z. Zhang, F. Yin, J. Liu, D.W.K. Wong, N.M. Tan, B.H. Lee, J. Cheng, and T.Y. Wong. Origa-light: An online retinal fundus image database for glaucoma analysis and research, In *Int. Conf. IEEE Engin. in Med. and Biol. Soc.*, 2010, pp. 3065–8.

[14] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. SLIC Superpixels, *Technical report*, 2010.

[15] C.Y. Ren and I. Reid. gSLIC: a real-time implementation of SLIC superpixel segmentation, *Technical report, University of Oxford, Department of Engineering Science*, 2011.

[16] D. Onkaew, R. Turior, B. Uyyanonvara, N. Akinori, and C. Sinthanayothin. Automatic Vessel Extraction with combined Bottom-hat and Matched-filter, In *Int. Conf. Information and Communication Technology for Embedded Systems (ICICTES)*, 2011, pp. 101–5.

[17] J. Tighe and S. Lazebnik. SuperParsing: Scalable Nonparametric Image Parsing with Superpixels, In *European Conference on Computer Vision*, 2010, pp. 352–65.

[18] X. Deng and G. Du. Liver tumor segmentation, In *3D Segmentation in the Clinic: A Grand Challenge II Workshop at 10th Int. Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2007.