

Efficient Reconstruction-Based Optic Cup Localization for Glaucoma Screening

Yanwu Xu¹, Stephen Lin², Damon Wing Kee Wong¹, Jiang Liu¹, and Dong Xu³

¹ Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore

² Microsoft Research Asia, P.R. China

³ School of Computer Engineering, Nanyang Technological University, Singapore

Abstract. We present a reconstruction-based learning technique to localize the optic cup in fundus images for glaucoma screening. In contrast to previous approaches which rely on low-level visual cues, our method instead considers the input image as a whole and infers its optic cup parameters from a codebook of manually labeled reference images based on their similarity to the input and their contribution towards reconstructing the input image. We show that this approach can be formulated as a closed-form solution without any search, which leads to highly efficient and 100% repeatable computation. Our tests on the *ORIGA* and *SCES* datasets show that the performance of this method compares favorably to those of previous techniques while operating at faster speeds. This suggests much promise for this approach to be used in practice for screening.

1 Introduction

Glaucoma is a disease of the optic nerve and a leading cause of blindness worldwide. Due to the lack of visual symptoms or discomfort in the early stages, studies have shown that more than 90% [1] of those with glaucoma are unaware that they have the disease. Since visual loss in glaucoma is irreversible and permanent, there is a strong need to detect this “silent thief of sight” as early as possible for immediate intervention.

During glaucoma progression, the death of ganglion nerve cells often leads to changes in the appearance of the optic nerve head, also known as the optic disc. In particular, there is a deepening and expansion of the excavation in the optic disc, known as the optic cup, as glaucoma advances, leading to an increase in the cup-to-disc ratio (CDR) as illustrated in Fig. 1. The CDR is a major consideration of clinicians in assessing glaucoma from direct observation [2]. However, manual assessment is labor-intensive, highly reliant on the training and expertise of the examiner, and not cost-efficient for screening.

These practical issues have motivated the development of computer-aided techniques for glaucoma screening using images from digital fundus cameras, which have become widely available in healthcare settings. Various automated methods based on CDR estimation have been presented, with the optic cup localized either by classifying pixels as part of the cup or rim (the disc area outside the cup) [3][4], or through an analysis of sub-regions (*e.g.*, defined by sliding windows [5] or superpixels [6]). These methods typically perform their classification or analysis based on explicit low-level visual cues such as intensity [4][5] or vessel kinks [7], but the reliance on low-level cues can make a method less robust to imaging noise and low contrast edges.

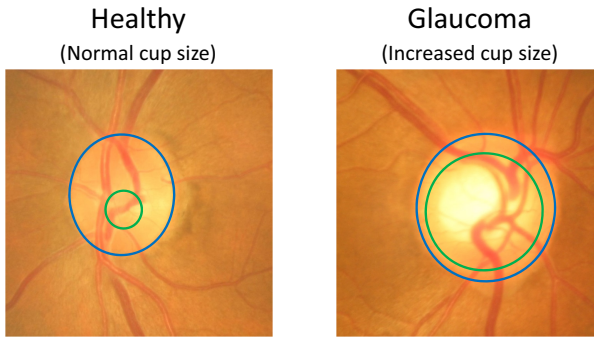


Fig. 1. Effect of glaucoma on cup-to-disc ratio (CDR): optic disc (blue) and cup (green)

In contrast to these methods, our technique identifies the optic cup within an input image in a holistic, reconstruction-based manner. From a codebook of reference images with expert-labeled cup regions, our method computes an optimal linear reconstruction of the input image from the reference images in a manner that gives greater emphasis to references more similar to the input. The linear reconstruction weights are then applied to the cup parameters of the corresponding reference images to estimate the cup region in the input image. Through this use of manually-labeled cup localization examples from reference images, our reconstruction-based technique can directly take advantage of clinician knowledge without having to rely on low-level feature processing.

This holistic approach to optic cup localization leads to higher accuracy than the current state-of-the-art optic cup detection method [6]. Furthermore, our formulation of this method has a one-step closed-form solution which results in significant computational savings. Our approach is inspired by medical image segmentation techniques that employ linear reconstruction [8] but require additional processing steps (*e.g.*, nearest neighbor search) and are less effective for sparse sets of reference images. The accuracy and efficiency of our reconstruction-based technique holds much promise for the development of practical automated/assisted glaucoma diagnosis systems.

2 Methodology

In our proposed localization framework, we adopt a left-eye frame of reference, with fundus images from right eyes horizontally flipped to be aligned with this reference frame. The optic disc is first segmented by using existing methods such as Active Shape Model (ASM) [9]. After normalization of the segmented disc, we localize the optic cup with our proposed reconstruction-based method. The risk of glaucoma is finally assessed using the calculated CDR value.

2.1 Disc Segmentation and Normalization

To segment the optic disc from a retinal fundus image, we employ either ASM [9] trained on codebook images or Template Matching [10]. The purpose of using two

different disc segmentation methods is so that we can compare the effect of different disc segmentation results on our optic cup localization. Following [9] and [10], we use only information from the green channel to localize the optic cup. Each segmented disc is normalized to a standardized circle with a radius of 128 pixels, and then each pixel value is linearly scaled to $[0, 1]$. Finally, the mean value over all the pixels in the optic disc is subtracted from each pixel to remove the influence of illumination variation among images.

2.2 Codebook Generation

In this work, our codebook is generated by random sampling from the manually-labeled images. To prevent bias, these samples are selected with an even distribution over the range of CDR values, which is divided for our dataset into three intervals: large-cup ($\text{CDR} > 0.6$, 72 images), mid-cup ($0.4 < \text{CDR} \leq 0.6$, 418 images), and small-cup ($\text{CDR} \leq 0.4$, 160 images). One-third of the codebook is thus generated by taking random samples from each of the three CDR intervals.

2.3 Closed-Form Reconstruction

From the codebook, we have n reference (ground truth) disc images $X = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^{k \times n}$ and their corresponding ground truth cups $C = \{c_1, c_2, \dots, c_n\} \in \mathbb{R}^{4 \times n}$, where $c_i = \{u_i, v_i, a_i, b_i\}^T$ denotes the cup descriptors of the i -th reference disc, u_i, v_i are the coordinates of the cup center, and a_i is the cup radius and b_i is the associated CDR after disc normalization to a radius of 1. For a given test disc $y \in \mathbb{R}^{k \times 1}$, we want to recover optimal linear reconstruction coefficients $w \in \mathbb{R}^{n \times 1}$, $|w| = 1$, that minimize the reconstruction error $\|y - Xw\|^2$. After w is obtained, the cup localization of y is estimated as Cw .

Since reconstruction of cup parameters is more accurate from images more similar to the test disc, we also include a cost term in the objective function that penalizes the use of references that are less similar to the test image. Let us denote the costs for the reference discs in X as the vector $d = \{d_1, d_2, \dots, d_n\}^T \in \mathbb{R}^{n \times 1}$, where d_i is the cost of using x_i for reconstruction. The overall cost term can then be expressed as $\|d \odot w\|^2$ where \odot denotes the Hadamard product. Combining this cost term with the reconstruction error gives the following objective function:

$$\arg \min_w \|y - Xw\|^2 + \lambda \|d \odot w\|^2, \quad \text{s.t. } |w| = 1. \quad (1)$$

This objective can be minimized in closed form using the Lagrange multiplier method, without the need for iterations:

$$w = \frac{1}{\mathbf{1}^T (\hat{X}^T \hat{X} + \lambda D^T D) \mathbf{1}} (\hat{X}^T \hat{X} + \lambda D^T D)^{-1} \mathbf{1}, \quad \hat{X} = (\mathbf{1} \otimes y - X), \quad (2)$$

where $D = \text{diag}(d)$ and \otimes denotes the Kronecker product. For simplicity, we define in our implementation the cost d_i as the Gaussian distance between the test disc y and the i -th reference disc x_i , *i.e.*,

$$d_i = e^{-\frac{\|y - x_i\|^2}{2\sigma_f^2}}, \quad (3)$$

where σ_f is a parameter that accounts for imaging noise.

3 Experiments

We evaluated the cup localization accuracy and CDR-based diagnosis performance of our proposed method. Comparisons are presented of our reconstruction-based approach to state-of-the-art pixel [4], sliding window [5] and superpixel [6] based methods, as well as to other reconstruction-based approaches. In addition, we compare our system against the current clinical standard for glaucoma detection using intra-ocular pressure (IOP) and to CDR values from expert graders. The effects of different optic disc segmentation algorithms and codebook sizes on our system are also examined.

3.1 Evaluation Criteria

For cup localization, we use the two evaluation criteria in [5] and [6], *i.e.*, non-overlap ratio (m_1) and absolute CDR error (δ), defined as

$$m_1 = 1 - \frac{\text{area}(E_{dt} \cap E_{gt})}{\text{area}(E_{dt} \cup E_{gt})}, \quad \delta = \frac{|D_{dt} - D_{gt}|}{2} \quad (4)$$

where E_{dt} denotes a detected cup region, E_{gt} denotes the ground-truth cup ellipse, D_{dt} is the vertical diameter of the detected cup, and D_{gt} is the vertical diameter of the ground-truth cup.

For glaucoma diagnosis, the area under the ROC (receiver operation characteristic) curve (AUC) is used for performance evaluation. The ROC is plotted as a curve which shows the tradeoff between sensitivity (P_+) and specificity (P_-), defined as

$$P_+ = \frac{TP}{TP + FN}, \quad P_- = \frac{TN}{TN + FP}, \quad (5)$$

where TP and TN denote the number of true positives and true negatives, respectively, and FP and FN denote the number of false positives and false negatives, respectively.

3.2 Experimental Setup

Two datasets, namely ORIGA [11] and SCES [12], were used in the experiments to validate both cup localization and glaucoma diagnosis accuracy. The *ORIGA* dataset, with ground truth disc and cup labelings as well as clinical glaucoma diagnoses, is comprised of 168 glaucoma and 482 normal images from studies of a Malay population. It is used in our work to validate both cup localization and glaucoma diagnosis. The Chinese population-based *SCES* dataset consists of 1676 images, of which 46 ($\sim 3\%$) are glaucoma cases. Since the *SCES* dataset provides only clinical diagnoses, it will be used only to assess the diagnostic performance of our system.

3.3 Effect of Different Disc Segmentation Methods on Our Method

We investigated the effect of different disc segmentation methods in our framework. The same parameters (codebook size $n = 99$, regularization weight $\lambda = 1000$, and Gaussian distance parameter $\sigma_f = 4$) were used for all the methods, and the results are

Table 1. Performance comparisons using different disc segmentation algorithms

Segmentation algorithms		Evaluation criteria	
Reference discs	Test discs	m_1	δ
Template matching	Template matching	0.225	0.071
ASM	ASM	0.227	0.073
Manual	Manual	0.229	0.074
Manual	Template matching	0.230	0.075
Manual	ASM	0.231	0.075
ASM	Template matching	0.229	0.074
Template matching	ASM	0.232	0.076

listed in Table 1. From the results, we observe that the proposed framework is robust and not very sensitive to the disc segmentation method. Slightly better performance was achieved when the reference discs and test discs were obtained using the same method. We note that using the same automatic segmentation method gives better results than manual labeling, likely because of greater consistency in the automatic methods.

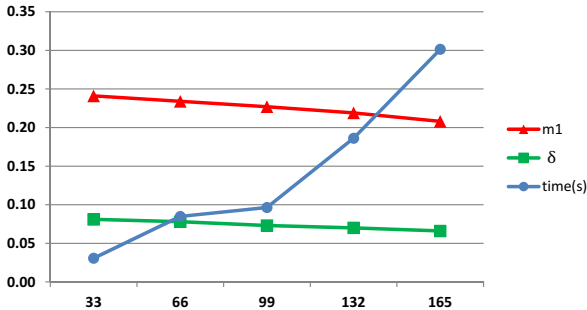
3.4 Cup Localization Comparisons

We compared our reconstruction-based approach to state-of-the-art pixel, window and superpixel based methods. Listed in Table 2, the results show that the proposed method achieves significant improvements in cup localization accuracy, which indicates the advantage of a holistic, reconstruction-based approach over techniques that analyze low-level visual cues. Substantial gains are made in terms of speed as well. With a codebook of 99 reference images, the proposed method locates cups in 0.1s, more than ten times faster than the comparison methods: pixel-based [4] (1.5s), window-based [5] (360s), and superpixel-based [6] (1.7–20.2s).

To support our particular reconstruction-based formulation, we also compare it to other possible reconstruction approaches, based on sparse coding [13], Locally Linear Embedding (LLE) [14], and Locality Preserving Projections (LPP) [15]. The codebook used for our method (with 99 reference images) is also used for these other reconstruction techniques. The sparse coding regularization weight was set to 100; LLE was implemented using the 11 nearest neighbors, and LPP was applied with an 11-dimensional subspace. These settings were found to approximately maximize the performance of these techniques, which is also reported in Table 2. The performance of the sparse coding method indicates limitations in using only a small number of reference images and not taking their similarity to the test image into account. Its optimization runs at 6.4s per image. Like our method, LLE considers both reference image similarity and reconstruction error, but does so in separate steps (first finding nearest neighbors, then reconstructing from them). This is shown to give lower accuracy than optimizing them jointly in our method. As with sparse coding, a reason for this is the sparse reference data, which makes it difficult to determine a priori which or how many reference images are needed for accurate reconstruction. A benefit of LLE, however, is its computational efficiency, running at 0.02s per image. The LPP method is also very fast

Table 2. Cup localization performance comparisons on *ORIGA* dataset

Method	m_1	δ
reconstruction based	0.225	0.071
<i>pixel based [4]</i>	0.474	0.149
<i>window based [5]</i>	0.284	0.096
<i>superpixel based [6]</i>	0.267	0.081
<i>error reduction relative to [4]</i>	52.5%	52.3%
<i>error reduction relative to [5]</i>	20.8%	26.0%
<i>error reduction relative to [6]</i>	15.7%	12.3%
<i>sparse coding</i>	0.408	0.158
<i>LLE</i>	0.247	0.081
<i>LPP</i>	0.377	0.150
<i>reconstruction-based without similarity cost</i>	0.271	0.091

**Fig. 2.** Cup localization accuracy and speed comparison with different codebook sizes

(0.012s per image), but loses much information in the dimensionality reduction and is very sensitive to alignment errors in the training and testing samples.

Table 2 also shows that leaving out the similarity-based cost term from our objective function (by setting $\lambda = 0$) yields a result that is worse (with significance tests of m_1 : $p < 0.001$, cdr : $p < 0.001$). This indicates that giving higher weights to more similar references is indeed better.

3.5 Influence of Codebook Size on Accuracy and Processing Time

For reconstruction-based approaches, codebook size is a key performance factor in terms of both accuracy and speed. We examined the cup localization accuracy and processing time of the proposed method with different codebook sizes $N \in \{33, 66, 99, 132, 165\}$ and the remaining images used for testing. As shown in Fig. 2, accuracy in terms of m_1 and δ improves with larger codebook sizes. The consequent increases in computation time still leave this method much faster than previous cup localization techniques, and in the following subsection we use all 650 images from *ORIGA* as reference images in evaluating glaucoma diagnostic performance on the *SCES* dataset.

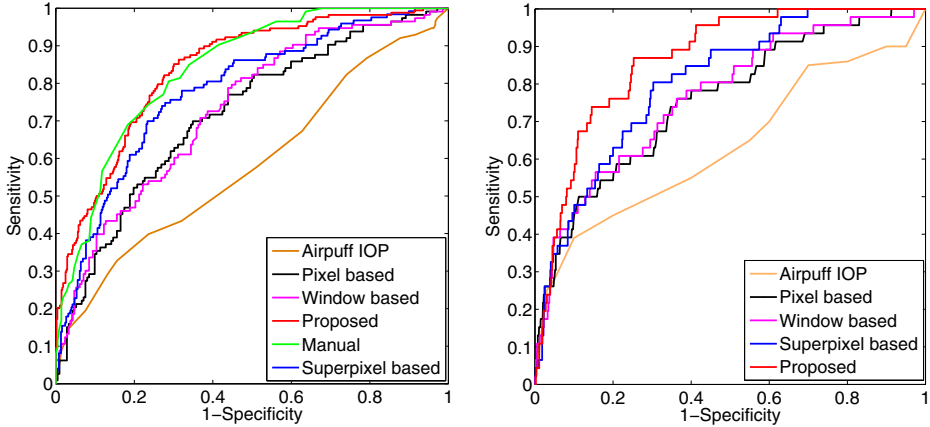


Fig. 3. Glaucoma diagnosis performance. Left: *ORIGA* dataset. Right: *SCES* dataset.

3.6 Glaucoma Diagnosis

To assess the glaucoma diagnosis accuracy of our method, we compare its predictions to the ground truth diagnoses of ophthalmologists and to state-of-the-art techniques. In addition, we compare to the current standard of care for glaucoma detection using IOP, as well as CDR grading results from an expert grader. For testing on *ORIGA*, a random selection of $N = 99$ images is used to generate the codebook and the remaining 551 images are used for testing. Testing on *SCES* was carried out using a codebook trained using the 650 images from *ORIGA*. We note that the reference images from *ORIGA* are of a different ethnicity from that of *SCES*.

The results are shown in Fig. 3. One can observe that the proposed reconstruction-based method outperforms previous automatic methods and IOP. The results are also close to those of the expert grader on the *ORIGA* dataset. The AUC values of our method on *ORIGA* and *SCES* are 0.823 and 0.860, respectively. However, the corresponding sensitivity values are only 58.0% and 73.9% on *ORIGA* and *SCES*, respectively, at an observation specificity of 85%. This indicates that other clinical diagnosis cues besides just CDR should be explored to further improve accuracy.

4 Conclusion

For CDR-based cup localization in glaucoma diagnosis, we have proposed a one-step reconstruction-based method that is computationally efficient and relatively insensitive to the disc segmentation method used. On tests with the *ORIGA* dataset to evaluate cup localization accuracy, our system achieves a 22.5% non-overlap ratio (m_1) against the ground truth and a 0.071 absolute CDR error (δ), a significant improvement over current state-of-the-art methods. Tests on the *SCES* dataset show that our method achieves an AUC value of 0.86, a sensitivity of 73.9% at a specificity of 85%. In future work, we plan to improve this reconstruction technique by introducing a learning algorithm to generate more effective codebooks for larger datasets, and also will investigate using data from multiple modalities to improve system performance.

References

1. Shen, S., Wong, T.Y., Foster, P., Loo, J., Rosman, M., Loon, S., Wong, W., Saw, S.M., Aung, T.: The Prevalence and Types of Glaucoma in Malay People: the Singapore Malay Eye Study. *Invest. Ophthalmol. Vis. Sci.* 49(9), 3846–3851 (2008)
2. Jonas, J., Budde, W., Panda-Jonas, S.: Ophthalmoscopic Evaluation of the Optic Nerve Head. *Survey of Ophthalmology* 43, 293–320 (1999)
3. Abramoff, M., Alward, W., Greenlee, E., Shuba, L., Kim, C., Fingert, J., Kwon, Y.: Automated Segmentation of the Optic Disc from Stereo Color Photographs Using Physiologically Plausible Features. *Invest. Ophthalmol. Vis. Sci.* 48(4), 1665–1673 (2007)
4. Wong, D.W.K., Lim, J.H., Tan, N.M., Zhang, Z., Lu, S., Li, H., Teo, M., Chan, K., Wong, T.Y.: Intelligent Fusion of Cup-to-Disc Ratio Determination Methods for Glaucoma Detection in ARGALI. In: *Int. Conf. Engin. in Med. and Biol. Soc.*, pp. 5777–5780 (2009)
5. Xu, Y., Xu, D., Lin, S., Liu, J., Cheng, J., Cheung, C.Y., Aung, T., Wong, T.Y.: Sliding Window and Regression based Cup Detection in Digital Fundus Images for Glaucoma Diagnosis. In: Fichtinger, G., Martel, A., Peters, T. (eds.) *MICCAI 2011, Part III. LNCS*, vol. 6893, pp. 1–8. Springer, Heidelberg (2011)
6. Xu, Y., Liu, J., Lin, S., Xu, D., Cheung, C.Y., Aung, T., Wong, T.Y.: Efficient Optic Cup Detection from Intra-image Learning with Retinal Structure Priors. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) *MICCAI 2012, Part I. LNCS*, vol. 7510, pp. 58–65. Springer, Heidelberg (2012)
7. Wong, D.W.K., Liu, J., Tan, N.M., Yin, F., Wong, T.Y.: Automatic Detection of the Optic Cup Using Vessel Kinking in Digital Retinal Fundus Images. In: *IEEE Int. Symp. Biomed. Imaging*, pp. 1647–1650 (2012)
8. Artaechevarria, X., Munoz-Barrutia, A., Ortiz-de-Solorzano, C.: Combination Strategies in Multi-Atlas Image Segmentation: Application to Brain MR Data. *IEEE Trans. Med. Im.* 28(8), 1266–1277 (2009)
9. Yin, F., Liu, J., Ong, S.H., Sun, D., Wong, D.W.K., Tan, N.M., Baskaran, M., Cheung, C.Y., Aung, T., Wong, T.Y.: Model-based Optic Nerve Head Segmentation on Retinal Fundus Images. In: *IEEE Int. Conf. Engin. in Med. and Biol. Soc.*, pp. 2626–2629 (2011)
10. Cheng, J., Liu, J., Wong, D.W.K., Yin, F., Cheung, C.Y., Baskaran, M., Aung, T., Wong, T.Y.: Automatic Optic Disc Segmentation with Peripapillary Atrophy Elimination. In: *IEEE Int. Conf. Engin. in Med. and Biol. Soc.*, pp. 6224–6227 (2011)
11. Zhang, Z., Yin, F., Liu, J., Wong, D.W.K., Tan, N.M., Lee, B.H., Cheng, J., Wong, T.Y.: Origa-Light: An Online Retinal Fundus Image Database for Glaucoma Analysis and Research. In: *IEEE Int. Conf. Engin. in Med. and Biol. Soc.*, pp. 3065–3068 (2010)
12. Sng, C.C., Foo, L.L., Cheng, C.Y., Allen Jr., J.C., He, M., Krishnaswamy, G., Nongpiur, M.E., Friedman, D.S., Wong, T.Y., Aung, T.: Determinants of Anterior Chamber Depth: the Singapore Chinese Eye Study. *Ophthalmology* 119(6), 1143–1150 (2012)
13. Olshausen, B.A., Field, D.J.: Emergence of Simple-Cell Receptive Field Properties by Learning a Sparse Code for Natural Images. *Nature* 381(6583), 607–609 (1996)
14. Roweis, S., Saul, L.: Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science* 290(5500), 2323–2326 (2000)
15. Cao, Y., Yuan, Y., Li, X., Turkbey, B., Choyke, P.L., Yan, P.: Segmenting Images by Combining Selected Atlases on Manifold. In: Fichtinger, G., Martel, A., Peters, T. (eds.) *MICCAI 2011, Part III. LNCS*, vol. 6893, pp. 272–279. Springer, Heidelberg (2011)