

# A Depth Based Approach to Glaucoma Detection Using Retinal Fundus Images

Akshaya Ramaswamy<sup>1</sup>, Keerthi Ram<sup>2</sup>, Mohanasankar Sivaprakasam<sup>1,2</sup>

<sup>1</sup> Indian Institute of Technology (IIT) Madras, India

<sup>2</sup> Healthcare Technology Innovation Centre (HTIC), India

**Abstract:** Qualitative evaluation of stereo retinal fundus images by experts is a widely accepted method for optic nerve head evaluation (ONH) in glaucoma. The quantitative evaluation using stereo involves depth estimation of the ONH and thresholding of depth to extract optic cup. In this paper, we attempt the reverse, by estimating the disc depth using supervised and unsupervised techniques on a single optic disc image. Our depth estimation approach is evaluated on the INSPIRE-stereo dataset by using single images from the stereo pairs, and is compared with the OCT based depth ground truths. We extract spatial and intensity features from the depth maps, and perform classification of images into glaucomatous and normal. Our approach is evaluated on a dataset of 100 images and achieves an AUC of 0.888 with a sensitivity of 83% at specificity 83%. Experiments indicate that our approach can reliably estimate depth, and provide valuable information for glaucoma detection and for monitoring its progression.

## 1 Introduction

Optic nerve head (ONH) assessment for glaucoma detection is commonly done by observing stereo optic disc images for structural abnormalities. Stereoscopic photographs of the optic nerve allow a perception of the cup depth, which is an important cue for glaucoma. A few works in literature have quantified the changes in the optic nerve by computing the ONH depth map using stereo matching methods [1] [2]. Cup-to-disc ratio (CDR) is an important parameter determined from the optic disc and optic cup and commonly used in identifying elongation of optic cup and loss of neuro-retinal rim. CDR estimation by thresholding depth maps calculated from stereo images has been implemented in [3]. However, due to the huge cost and high level of expertise required for stereoscopic evaluation, the use of stereo imaging is limited in glaucoma screening.

In [4], depth estimation and cup extraction is done on a single image using a coupled sparse dictionary based supervised method. CDR estimation has been widely performed using monocular retinal fundus images. Various methods such as deformable model based [5], superpixel based [6] and supervised classifier based methods [7] have been implemented on monocular images to segment the optic cup and compute CDR. Very few works have directly used CDR for classification of normal and glaucomatous discs [8]. Moreover the use of CDR as a discriminatory measure for glaucoma prediction is limited due to the considerable overlap between glaucomatous and normal cases [9].

The idea of our approach is to get an estimate of the depth of the optic disc, by analysing the intensity variations in the ONH region, and to extract discriminative features for glaucoma prediction. Optic disc cupping is a common risk factor for glaucoma, and its spatial variation is an important indicator to identify the extent of glaucoma. Our approach is based on the assumption that the actual depth map of the disc is related to the intensity variations inside the disc. We perform superpixel-based supervised soft segmentation of optic cup from the retinal fundus image [10], and build a depth map of the optic disc ROI. We also implement a fast marching technique to estimate the unsupervised depth map, and combine the two to compute our final depth estimate of the optic disc. This is used to compute spatial features and intensity features that characterize the variations in the depth, for the detection of glaucoma.

## 2 Methodology

Our proposed approach consists of four main steps: Supervised depth estimation, unsupervised depth estimation and integration, feature extraction, and glaucoma prediction. The steps of our approach are illustrated in figure 1.

### 2.1 Supervised Depth Estimation

**Cup Segmentation:** We implement the superpixel-based cup segmentation method presented in our previous work [9]. This method is similar to the one in [6]. First, we generate superpixels on the optic disc ROI  $I$  using simple linear iterative clustering (SLIC). The optic disc boundary is assumed to be given, and we perform feature extraction and classifier training on the superpixels within the optic disc. From each superpixel within the optic disc, histogram based features, center-surround difference based features, and a distance feature is extracted. These superpixel features are used to train a Classifier  $C_I$  for cup-rim classification.  $C_I$  outputs a cup confidence map  $M_I$ , assigning a cup probability score for each superpixel in the disc.

**Depth Estimation from Cup Segmentation Outputs:** We compute our depth estimate  $D_I$ , by first taking a complement of  $M_I$  and then smoothing it. We invert  $M_I$  so that we have low intensity values around the disc center indicating a deeper region, and higher intensity values as we move away. We consider  $M_I$  to be the initial estimate of the depth map. Due to the segmentation at the superpixel level,  $M_I$  is not smoothly varying. We perform smoothing on  $M_I$  by first extracting the locations of centroids of the generated superpixels and their corresponding cup confidence values. We use this to perform cubic interpolation to get the interpolated cup confidence values on the rest of the pixels. This gives us a reasonably smooth map  $D_I$ . Further smoothing is performed by using a large sized moving average filter to get the final smooth depth map  $D_f$  using the cup confidence map. Figure 2 shows the intermediate outputs, the final depth map and the OCT depth ground truth for an input image ROI.

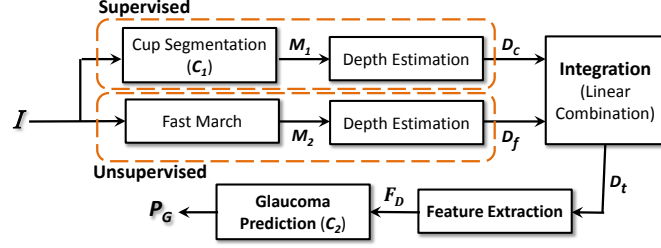


Figure 1: Depth based approach framework: intensity-based depth estimation, followed by feature extraction and glaucoma prediction

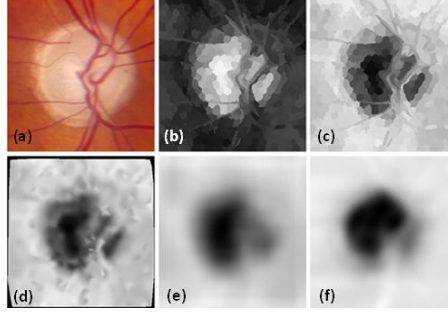


Figure 2: Depth estimation using cup segmentation: (a) original optic disc ROI, (b) Cup Confidence map  $M_1$  (c)  $M_1'$ , (d)  $D_1$ , after cubic spline interpolation, (e) Final depth map  $D_f$ , (f) the depth map ground truth  $D_{gt}$

## 2.2 Unsupervised Depth Estimation

We perform an unsupervised depth estimation using fast marching to capture the intensity variations from the boundary of the optic disc to the disc center, and characterize the depth in terms of the geodesic distance.

**Cost Image Generation:** We generate an appropriate cost image for the fast march algorithm that can describe the shape of cupping. We need a cost image that enhances the cup-rim contrast. In the RGB colorspace, the green channel image is known to be suitable to extract the cup. Here along with the green channel image  $I_G$  we also compute the luminance channel image  $I_L$  in the  $L^a*b^*$  colorspace. Our final cost image  $W$  is the pixel-wise product of  $I_G$  and  $I_L$  (refer figure 3(b)).

**Depth Estimation using Fast March:** The depth map of the optic nerve head using  $W$  is estimated by means of the fast marching method. For every image  $I$ , we have  $C_{od}$  as the optic disc center, and  $p_1, p_2 \dots p_K$  as  $K$  points lying on a circle, with diameter a little greater than that of the optic disc (Refer figure 3(c)). For every point  $p_k$  in the image  $W$ , we perform the following steps:

i) We implement fast marching with  $p_k$  as the start point, and  $C_{OD}$  as the end point. The output is a geodesic map  $G_k$ .

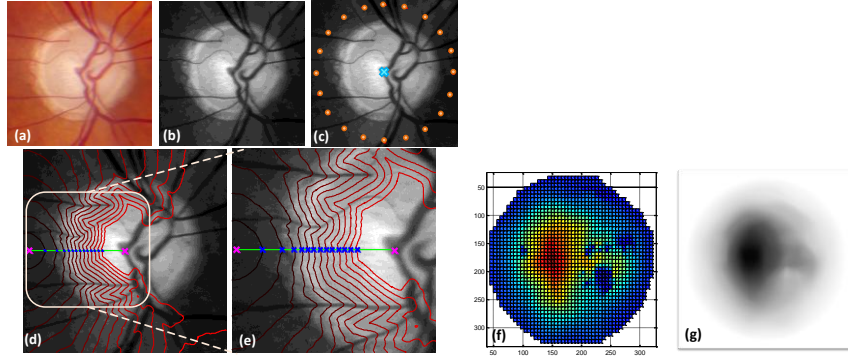
ii) From  $G_k$ , we extract  $N$  contours. Each contour  $C_n$  consists of a set of pixel locations  $\{X_n\}$ , and the corresponding contour height  $\{H_n\}$ .

iii) We also compute the line  $L_k$  joining  $p_k$  and  $C_{OD}$ .

iv) For each contour  $C_n$  we extract the points of intersection between  $C_n$  and  $L_k$ , and the corresponding contour height (refer figure 3(d)).

Totally, for all the  $K$  start points, we get a set of  $V(x,y)$  pixel intersection locations  $r_1, r_2 \dots r_v$  and corresponding contour heights  $h_1, h_2 \dots h_v$ . Using these pixel locations and height values, we perform a cubic spline interpolation to get the interpolated height values at the rest of the pixel locations. This gives us  $D_I$ , a crude complement of the depth map (refer figure 3(f)). We perform further processing by inverting it and applying median filter, to get the final unsupervised depth map  $D_f$  (refer figure 3(f)).

Our final depth estimate  $D_t$  is a linear combination of the supervised depth map  $D_c$  and the unsupervised depth map  $D_f$ .



**Figure 3: Unsupervised depth map estimation: (a) Original optic disc ROI; (b) the cost image  $W$ ; (c)  $W$  showing  $K$  start points in orange, and the end point  $C_{OD}$  in blue, for fast marching; (d)  $W$  showing  $p_k$  and  $C_{OD}$  in magenta, the contours of a  $p_k$  overlayed in red, the line  $L_k$  joining  $p_k$  and  $C_{OD}$  in green, and the points of intersection in blue; (e) a zoomed version of the points of intersection in (d); (f) Geodesic map  $D_2$  after cubic spline interpolation; (g) Final depth map  $D_f$**

### 2.3 Feature extraction

The extent of cupping can be visualized from the depth map  $D_t$  computed using supervised and unsupervised techniques. From the depth map, we extract spatial and intensity features that describe the local dissimilarities on the surface. All the depth maps, whether normal or not, have approximately the same structure, therefore it is necessary to compute features that extract spatial information and capture local differences. Many works in literature emphasize

the importance of considering disc size in detection of glaucoma [9]. So, along with depth map features, we also extract shape features from the optic disc.

**Spatial features:** We utilize the polar map template, introduced in [10] to extract features that capture the spatial information from the depth maps. We construct the template here, by fixing the number of concentric shells around the optic disc center to be 4, and the number of sectors in each shell as 8. We calculate two features-mean depth and mean gradient of depth- within each of the 32 bins, to get a total of 64 spatial features.

**Intensity features:** We threshold the depth map using eight values between 0 and 1, and compute the number of thresholded pixels, and the orientation of the extracted binary map. This gives us 16 intensity features.

**Global features:** We also compute three global features- major axis length, minor axis length, and orientation- of the optic disc, and thresholded optic cup from the depth map. We set the threshold as 0.5 heuristically, to approximately extract the optic cup from the depth map. This gives us 6 global features.

Totally, we extract 86 features ( $F_D$ ) from the  $D_t$  for glaucoma prediction.

#### 2.4 Glaucoma Prediction

From depth map  $D_t$ , we extract 86 features to train a random forest classifier  $C_2$ .  $C_2$  is trained using these features to classify images as normal or glaucomatous. The output of  $C_2$  is a probability value  $P_G$  which indicates the risk of presence of glaucoma.

### 3 Experimental results and discussion

We evaluate our depth estimation approach on 30 stereo images ( $T_1$ ) from INSPIRE dataset [1]. We consider only one image from the stereo pair and evaluate our approach with respect to the available depth reference based on spectral domain Optical Coherence Tomography (OCT). From the OCT volume of the ONH, surfaces of the retinal layer are detected by 3D segmentation. Depth is recorded as intensities and registered manually with the reference stereo images to create depth ground truth for evaluation. Our glaucoma prediction approach is evaluated on a set of 100 images ( $T_2$ ) created using publicly available RIM-ONE R1, RIM-ONE R2 and RIM-ONE R3 datasets [11].  $T_2$  consists of 53 normal images, and 47 glaucoma and glaucoma suspicious images. We manually crop images in  $T_1$  and  $T_2$  to retain optic disc ROI. We perform three main experiments: i) comparison of our depth estimation approach with other state-of-art methods on  $T_1$ , ii) evaluation of glaucoma prediction performance of our approach on  $T_2$ , and iii) evaluation of class separation of the glaucoma probability outputs for glaucoma and normal classes.

#### 3.1 Depth Estimation

**Depth Estimation from Cup Segmentation Output:** We use the classifier model generated by  $C_1$  in our previous work [9] to perform cup segmentation on the images in  $T_1$  and carry out further smoothing to compute  $D_c$ .

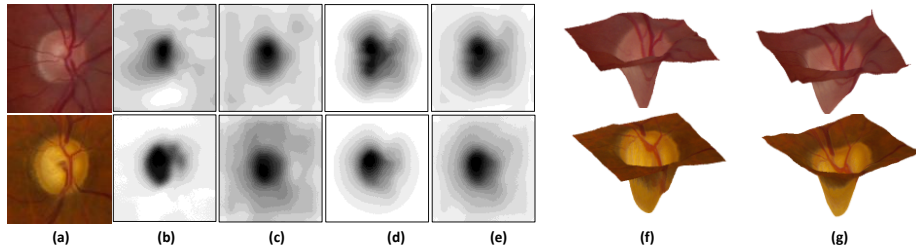
**Depth Estimation from Fast March:** We perform pre-processing on  $T_1$  to get the cost image  $W$ . On this, we implement our fast marching algorithm. We set the number of start points  $K$  to be 20, chosen at an interval of  $18^\circ$  from each other on the circle boundary with diameter 1.2 times the vertical diameter of the optic disc (refer figure 3 (c)). We set the number of contours  $N$  from each point as 15. Setting these values to the parameters, we compute  $D_f$ .

Using  $D_c$  and  $D_f$ , we also compute our final depth map  $D_t$  by taking the linear combination of the two, with equal weights of 0.5 for both. Figure 5 shows the ground truth  $D_{gt}$ , the computed depth maps  $D_c$ ,  $D_f$  and  $D_t$ , and the surface visualization, for two images from the INSPIRE-stereo dataset.

**Performance Evaluation:** To evaluate the performance of our methods, we compute the parameters - Correlation, Root mean square error, and the standard deviation- between the estimated depth maps and ground truth, and compare it with state of art methods. Table 1 compares the values of the evaluation metrics for different depth estimation methods.

### 3.2. Evaluation of Glaucoma Prediction

We perform feature extraction on our depth estimate  $D_t$  to compute  $F_D$ .  $F_D$  is used to train a random forest classifier  $C_2$  to output the glaucoma probability values  $P_G$ . We perform leave-one-out validation on 86 features extracted from the 100 images, to evaluate the performance of  $C_2$ . We set the number of trees as 1000, and the number of randomly selected variables as  $9(\sqrt{N})$ . Figure 4 shows the glaucoma detection performance for leave-one-out validation on  $C_2$ . Our depth-based approach using  $D_t$  achieves a maximum AUC of 0.888. We compare this performance with the best performing polar map based approach, and the CDR-based approach [9]. We observe that our depth based approach performs better than the other two, achieving a sensitivity of 83% at specificity 83%.



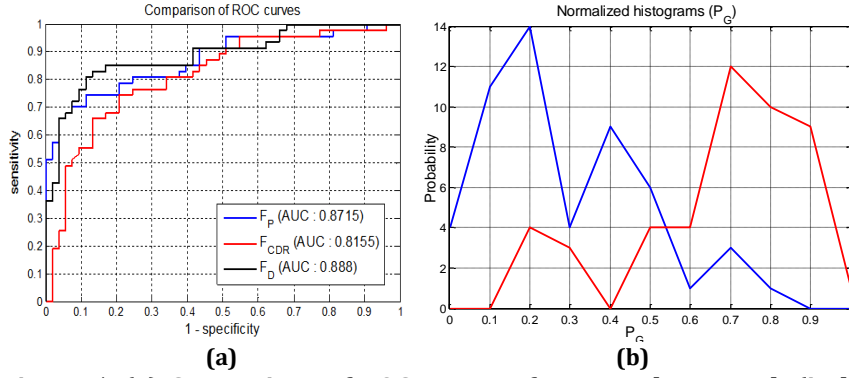
**Figure 3:** (a) Original image; (b) OCT Depth map  $D_{gt}$ , (c) Depth from cup confidence map,  $D_c$ , (d) Depth from fast marching,  $D_f$ , (e) Combined depth map  $D_t$ , (f) Visualized surface using  $D_{gt}$ , (g) Visualized surface using  $D_t$

**Table 1: Comparison of performance of various depth estimation methods**

Method	Evaluated parameters		
	<i>Correlation</i>	<i>RMSE</i>	<i>Std Dev.</i>
Multi-scale stereo matching [1]	-	0.1592	0.0879
Coupled Sparse dictionary based (monocular) [4]	0.8	-	-
Proposed supervised ( $D_c$ )	0.7727	0.1825	0.1302
Proposed unsupervised ( $D_f$ )	0.7852	0.1488	0.1393
Proposed Combined ( $D_t$ )	0.8225	0.1532	0.1206

### 3.3. Evaluation of Class Separation

We evaluate the glaucoma probability output  $P_G$  by measuring the class separation of  $P_G$ . For this, we compute a distance metric, Bhattacharya distance  $\rho_D$  which measures how good a parameter is at discriminating between two or more classes. A higher value of  $\rho_D$  indicates better class separation. Here, we compute  $\rho_D$  to measure the glaucoma-normal separation of  $P_G$ . Figure 5(b) shows the normalized histograms of  $P_G$  for glaucoma and normal classes. The Bhattacharya distance  $\rho_D$  is calculated to be 0.6141, which is comparable with that for the CDR-based approach and polar map based approach. This shows glaucoma probability outputted by  $C_2$  using features extracted from  $D_t$  has high class discrimination and can be a dependable indicator of the presence of glaucoma.



**Figure 5: (a) Comparison of ROC curves of proposed approach (in black), CDR-based approach (in red) and polar map approach (in blue); (b) normalized histograms of  $P_G$  for normal (blue) and glaucoma (red) classes; Bhattacharya distance  $\rho_D = 0.6141$ .**

## 4 Conclusion

In this paper, we implement a depth-based approach for glaucoma detection. We perform supervised and unsupervised techniques on monocular fundus images to compute a depth map  $D_t$ . From  $D_t$ , we compute spatial features, intensity features and global features, and train a classifier  $C_2$  to output a glaucoma probability value  $P_G$ . We evaluate the performance of our depth estimation, glaucoma prediction and the class separation of the glaucoma probability output. The glaucoma-normal classification using our approach achieves an AUC of 0.888, with sensitivity 83% at specificity 83%. Experiments indicate that our approach can be used to reliably estimate depth from a single image, and extract features that capture local dissimilarities in the depth. These features that characterize the depth are useful in predicting the glaucoma risk with reasonable class separation. Future work aims to automate ROI extraction, and apply this method on both images in the stereo pair to analyse the results. From the depth maps, more informative geometric features can be extracted to monitor glaucoma progression, and also identify the stages of glaucoma.

## References

1. L. Tang, et al., "Robust Multiscale Stereo Matching from Fundus Images with Radiometric Differences," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 11, pp. 2245–2258, Nov. 2011.
2. T. Nakagawa, et al., "Quantitative depth analysis of optic nerve head using stereo retinal fundus image pair," J. Biomed. Opt., vol. 13, no. 6, pp. 064026–064026–10, 2008.
3. C. Muramatsu, et al., "Determination of cup-to-disc ratio of optical nerve head for diagnosis of glaucoma on stereo retinal fundus image pairs," 2009, vol. 7260, p. 72603L–72603L–8.
4. A. Chakravarty and J. Sivaswamy, "Coupled Sparse Dictionary for Depth-Based Cup Segmentation from Single Color Fundus Image," in MICCAI 2014, Springer, 2014, pp. 747–754.
5. D. W. K. Wong, et al., "Level-set based automatic cup-to-disc ratio determination using retinal fundus images in ARGALI," in IEEE EMBS 2008, pp. 2266–2269.
6. J. Cheng, et al., "Superpixel Classification Based Optic Disc and Optic Cup Segmentation for Glaucoma Screening," IEEE Transactions on Medical Imaging, vol. 32, no. 6, pp. 1019–1032, 2013.
7. D. W. K. Wong, et al., "Intelligent fusion of cup-to-disc ratio determination methods for glaucoma detection in ARGALI," in IEEE EMBC 2009, pp. 5777–5780.
8. J. Nayak, S. Bhat, N. Shetty, and T.-C. Lim, "Automated Diagnosis of Glaucoma Using Digital Fundus Images," J Med Syst, vol. 33, no. 5, pp. 337–346, Oct. 2009.
9. J.B. Jonas, et al., "Optic disc, cup and neuroretinal rim size, configuration and correlations in normal eyes", Invest Ophthalmol Vis Sci. 1988;29:1151–8
10. A. Ramaswamy, et al., "A polar map based approach using retinal fundus images for glaucoma detection," Second Ophthalmic Medical Image Analysis International Workshop, OMIA 2015
11. F. Fumero, et al., "RIM-ONE: An open retinal image database for optic nerve evaluation," in Intl. Symposium on Computer-Based Medical Systems (CBMS), 2011.