Segmentation of the Optic Nerve Head Combining Pixel Classification and Graph Search

Michael B. Merickel Jr.[a], Michael D. Abràmoff[a,b,c], Milan Sonka[a,b], Xiaodong Wu[a,d]

[a]The University of Iowa, Dept. of Electrical and Computer Eng., Iowa City, IA, USA
[b]The University of Iowa, Dept. of Ophthalmology and Visual Sciences, Iowa City, IA, USA
[c]VA Medical Center, Ophthalmology Service, Iowa City, IA, USA
[d]The University of Iowa, Dept. of Radiation Oncology, Iowa City, IA, USA

ABSTRACT

Early detection of glaucoma is essential to minimizing the risk of visual loss. It has been shown that a good predictor of glaucoma is the cup-to-disc ratio of the optic nerve head. This paper presents an automated method to segment the optic disc. Our approach utilizes pixel feature selection to train a feature set to recognize the region of the disc. Soft pixel classification is used to generate a probability map of the disc. A new cost function is developed for maximizing the probability of the region within the disc. The segmentation of the image is done using a novel graph search algorithm capable of detecting the border maximizing the probability of the disc. The combination of graph search and pixel classification enables us to incorporate large feature sets into the cost function design, which is critical for segmentation of the optic disc. Our results are validated against a reference standard of 82 datasets and compared to the manual segmentations of 3 glaucoma fellows.

Keywords: image segmentation, feature selection, pixel classification, graph search, optic disc

1. INTRODUCTION

Glaucoma affects approximately 2-3% of the US population and is the second leading cause of blindness in the US. The risk of visual field loss due to glaucoma is minimized by early diagnosis and optimal treatment methods [1]. The optic nerve head is a three-dimensional structure characterized by a peripheral 'disc' and a central depression called the 'cup'. Certain characteristics of the optic nerve head facilitate early detection of glaucoma.

Currently, the gold standard for diagnosis and treatment follow-up of glaucoma is optic nerve planimetry [1]. The method determines the extent of the disc and cup through manual evaluation of stereo 2-D retinal images of the optic nerve head by an ophthalmologist. The technique is time-consuming and tedious and introduces large variability due to the need for human interpretation [2]. It is clear that an automated, quantitative method is necessary for analyzing the optic nerve head in stereo photographs.

In this paper, we focus on the segmentation of the optic disc. Current methods, though approaching those of human experts [3], are influenced by poor image quality and lack well-defined features for use in the segmentation or classification of the disc [3–7]. The algorithms generally encompass three different techniques: pixel classification, graph search and active contour. The pixel classification [3, 4] tends to yield disjoint regions that are difficult to use in later processing. Graph search techniques and active contours have difficulty incorporating a large number of features into the segmentation because they need to be combined in some intelligent way. Features are generally selected using prior knowledge and is widely considered an art [8]. To formulate the cost terms in the segmentation the features need to be combined, and this is usually done using a linear combination which may not be optimal (e.g. it is difficult to compare edge strength with texture) [4]. The selection of features and their combinations involves many different variables and it can become difficult to generate a robust result.

Techniques exist in level sets and pattern recognition for incorporating statistics into image segmentation. For example, Paragios *et al.* analyze the image histogram to generate Gaussian distributions approximating the different regions of interest [9]. In Bayesian decision theory [10], *a priori* knowledge about different regions is

Contact information: michael-merickel@uiowa.edu

incorporated into the analysis. We investigate the incorporation of different statistics further by attempting to generate probabilities for the different regions for use in a graph search. We also investigate a scheme for learning the probabilities based on patterns.

Our algorithm was validated against manual segmentation by three glaucoma experts in 82 datasets from patients with a diagnosis of glaucoma. We also compare the results to three glaucoma fellows. The theoretical advantages of the algorithm were proven by experiments.

2. METHODS

The optic disc segmentation is formulated as a region detection problem where the optimal region is described by the closed set of pixels that maximizes the total probability of belonging to the disc. Computing the probabilities is done by incorporating a set of low-level image features into a soft classification. The set of features can be optimized by using feature selection methods to minimize the error in the borders from the reference standard. The computational feasibility of this approach is accomplished by transforming the region detection to a problem of computing an optimal closed path in a 2-D geometric graph. The novelty of this method lies in the ability to automatically optimize the cost function for use in the graph search, as well as the formulation of the problem in terms of an optimal border detection for segmenting a region-based objective function.

2.1. Preprocessing and Feature Computation

Simple grayscale information can be extracted from the color fundus images, however it has been shown [3] that different color spaces such as the red-green and blue-yellow color planes can also provide useful information. We take advantage of this by computing various low-level features (e.g. Gaussians at different scales, edge detectors, etc) on the different planes, yielding a large set of features that describe the original image.

The vessels tend to negatively influence the position of the segmented border, thus before computing the features we have incorporated a preprocessing step to remove the vessels from the images. The vessels can be detected using a simple kNN classifier trained on some small training set [10, 11]. The vessel pixels are then treated similarly to convolution with a kernel that exceeds the bounds of the image. An invalid pixel is mirrored (flipped) across the closest N valid pixels, and then replaced by the average of the mirrored pixels. This technique helps to minimize the effect of the vessels on the computed features, allowing for a more consistent segmentation.

2.2. Probability Maps

The computation of the probability maps is done for a given set of features using soft classification. A feature set F contains features $f_1, f_2, \cdots, f_n$. Each pixel (i, j) is represented by a feature vector containing the sampled features $f_1(i, j), \cdots, f_n(i, j)$. These feature vectors can then be input to a kNN classifier to compute the k nearest neighbors for each pixel. The probability $\mathrm{P}_c(i, j)$ for pixel (i, j) in class $c \in \{disc, bg\}$ is calculated as

$$\mathrm{P}_c(i, j) = \frac{n_c}{k}$$

where n_c is the number of nearest neighbors of class c and k is the total number of nearest neighbors. This approach can provide an approximation of the probability for each pixel based on the nearest neighbors. The kNN classifier can be trained using known samples where the class is given in the reference standard. Efficient implementations of approximate kNN classifiers exist, vastly improving the performance of the probability map computation [12].

2.2.1. Feature Selection

We are interested in optimizing the feature set for use in the graph search. However, the set of possible features can grow to become unmanageable. Our algorithm uses a one-time feature selection routine to optimize the feature set for our particular application, the optic disc. This is done using the Sequential Floating Forward Selection (SFFS) algorithm [13]. This is an efficient, iterative search through the feature space which involves two loops working against one another to improve the performance at each step. There is an inclusion loop which performs Sequential Forward Selection (SFS) to try and add the best feature to the set. After the best feature

is added, Sequential Backward Selection (SBS) is run to attempt to remove any feature that may degrade the performance, thus improving the overall performance of the selected set while reducing its cardinality.

The SFFS requires an error metric to measure the performance of the algorithm at each step. Because feature set is used to compute probability maps which directly influence the final segmentation it's logical to try and optimize the feature set for our application by measuring the performance in terms of the absolute border positioning errors. The Euclidean distances are calculated from points on the reference standard to points on the calculated borders, and vice versa. The final metric is then representative of the total error in the dataset. We calculate distances *from* and *to* the reference standard as a method for minimizing the effect that oddly shaped borders may have on the error. The borders are computed using the segmentation process described in the following sections.

2.3. Disc Segmentation

The disc segmentation is the core of the algorithm, utilizing the probability maps to provide a segmented border. Segmentation involves transforming the original probability map P_{disc} into polar coordinates, modifying the closed border segmentation into a min-cost path problem. From the smoothness parameters a graph is constructed, and the cost function is computed to assign a cost to each vertex within the graph. Finally, the optimal path is computed and transformed into a 2-D spline representing the final border of the detected region.

2.3.1. Polar Transformation

The original probability map P_{disc} is transformed into polar coordinates using a similar algorithm to that described by Chen, Wang, Wu [14]. A kernel point Q from which the optic disc is unwrapped is determined as the approximate center of the disc. Q must be within the disc for the transformation to be valid. The optic disc is sampled from Q radially outward with I rays of radius R and J samples per ray. The result of the transformation to polar coordinates is a new 2-D image $\mathcal{P}(i, j)$ where $0 \leq i < I$ and $0 \leq j < J$.

The polar transformation we use is similar to this, however it is possible that a sample could lie within vessel region which we have removed earlier. To minimize the effect of these regions, each sample is computed from an average of the values, excluding vessel pixels, along an arc whose endpoints bisect the surrounding two rays. An example of this is shown in Figure 1. Sample 3, ray 0 and sample 5, ray 0 are generated by averaging across their respective arcs.

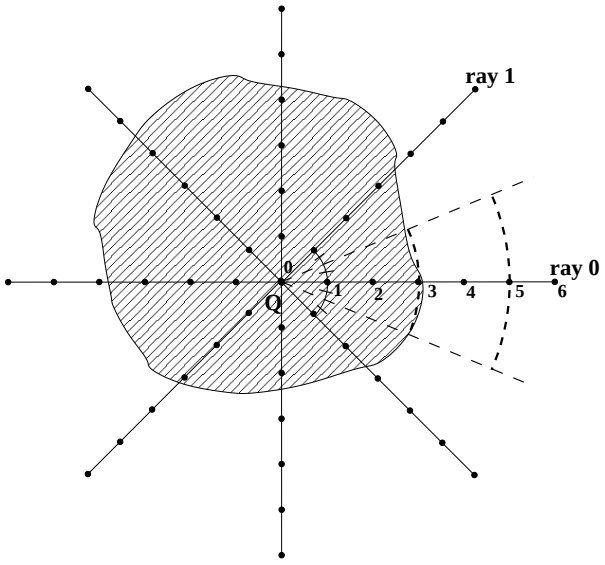


Figure 1. An example of the polar transform where the samples along each ray are interpolated along arcs between the rays.

2.3.2. Graph Construction

The unwrapped image $\mathcal{P}(i,j)$ can be viewed as a 2-D weighted, directed graph $G = (V, E)$ (see Figure 2) such that each pixel $\mathcal{P}(i,j)$ corresponds to a vertex in V and the edges of E correspond to the connections between pixels to form a feasible border for the optic disc.

Since the boundary is smooth, the segmented borders should be sufficiently "smooth", that is, any two adjacent pixels on the border should not be too far apart. Precisely, a smoothness constraint M exists in which a vertex V at the point (i,j) is said to have a directed edge from itself to every point $(i+1, j \pm q)$ where $0 \leq q \leq \lfloor M/2 \rfloor$ and $j - q \geq 0$ and $j + q < J$.

The cost of a vertex in V is inversely proportional to the likelihood that it is located on the desired border. These constraints allow us to build a cost function for the disc and the problem is reduced to tracing an optimal path p_{disc} in G, where $0 \leq p_{disc}(i) < J$ for every $0 \leq i < I$. The optimality of the path is defined with respect to the total cost of the vertices on it.

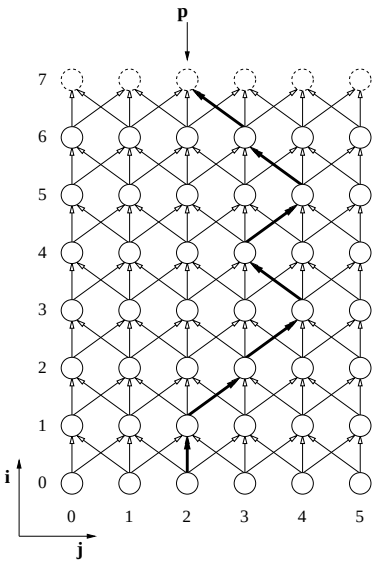


Figure 2. A 2-smoothness graph constructed from the unwrapped image. Path p represents a possible closed path through the graph.

2.3.3. Cost Function Design

Our algorithm relies on the ability to maximize the probability of the detected region which can be seen as maximizing the total probability

$$\mathrm{P}_{total} = \sum_{(i,j) \in R_{disc}} \mathrm{P}_{disc}(i,j) + \sum_{(i,j) \in R_{bg}} \mathrm{P}_{bg}(i,j) \tag{1}$$

where $\mathrm{P}_{disc,bg}$ are the probabilities of pixel (i,j) belonging to the disc and background, respectively, and $R_{disc,bg}$ are the regions containing the disc and background, respectively.

The relationship between the unfolded image and the original probability map lies in the fact that the unfolded image is a sampled version of the original probabilities with some interpolation to fill the invalid vessels. The actual content of the unfolded image is still consistent with the original image, however, allowing us to perform the optimization in polar coordinates. The optimization problem can be viewed similarly to Equation 1 in the unfolded graph as minimizing the total cost of the path p_{disc}.

The optimal path is computed in polar coordinates where $j = 0$ represents the center pixel, thus the disc grows outward as j increases (e.g. in Figure 3 the segmented disc is on the left of the optimal path). This implies that the cost $\mathcal{C}(i,j)$ represents the probability for the entire region $j' \in [0,j]$ on row i. For a vertex (i,j) representing the region $(i, \leq j)$ we assign a cost $\mathcal{C}(i,j)$ to the cost function $\mathcal{C}$ where $0 \leq i < I$ and $0 \leq j < J$.

$$\mathcal{C}(i,j) = \sum_{0 \leq j' < j} \mathrm{P}_{disc}(i,j') - \mathrm{P}_{bg}(i,j') \tag{2}$$

Note that it is equivalent to view the total cost as the difference of probabilities within the disc, or the sum of the probabilities within their respective regions. The computation can be further optimized by realizing that $\mathrm{P}_{disc}(i,j) + \mathrm{P}_{bg}(i,j) = 1$, thus the final cost function is described in Equation 3.

$$\mathcal{C}(i,j) = \sum_{0 \leq j' < j} 2\mathrm{P}_{disc}(i,j') - 1 \tag{3}$$

Finally, the the segmentation algorithm is focused on minimizing the cost function, and so it is necessary to invert the costs which translates to finding the region with the lowest probability of being background.

2.3.4. Optimal Path Computation

The optic disc is a closed border, thus in polar coordinates $p_{disc}(0) = p_{disc}(I)$. The Chen, Wang, Wu [14] algorithm facilitates this fact with an efficient graph search algorithm. The graph search takes advantage of the fact that two optimal paths starting at two different points, $(0, j')$ and $(0, j'')$ $(j' \neq j'')$, can be found that do not cross each other. Following from this property is a *divide-and-conquer* algorithm in which the optimal path p_0 beginning at the point $(0, \lfloor \frac{J}{2} \rfloor)$ is computed. The graph is then divided into two subgraphs G_1 and G_2 along p_0 and the algorithm is recursively called on G_1 and G_2 to find the optimal closed paths p_1 and p_2 in G_1 and G_2, respectively. The optimal closed path in G is the path with a minimum cost among p_0, p_1 and p_2. This divide-and-conquer paradigm yields an improvement of nearly an order of magnitude over previous algorithms [15]. A dynamic programming algorithm is used to compute the optimal path beginning at a specific point $(0, j)$. The optimal closed path is then back-traced from the point (I, j) thus beginning and ending at the same point. In this way, we obtain the optimal path p_{disc} for the optic disc.

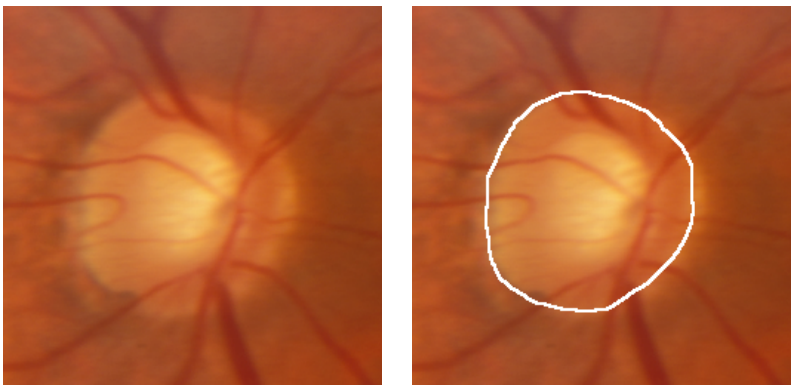
The optimal path, computed in polar coordinates, consists of samples along the final border in the original image. The final border is calculated by transforming the samples using the inverse of the polar transform, and then fitting an interpolating 2-D closed spline to the data [16].

3. EXPERIMENTAL METHODS

3.1. Data

Stereo photographs were obtained from 82 patients with a diagnosis of glaucoma. Color slide stereo photographs centered on the optic disc were acquired using a fixed geometry Nidek 3Dx stereo retinal camera. This camera takes simultaneous left and right stereo photographs of the optic disc on slide film; after development, the slides were scanned at 4096×4096 pixel resolution, 24 bit color depth, with a Kodak slide scanner (Kodak Inc., Rochester, NY).

The Nidek 3Dx camera projects two alignment marks onto the retina that are photographed simultaneously. Mutual information-based affine registration was used to align the right and left stereo pairs. The image was then cropped to 512×512 pixels keeping the optic disc in the center (see Figure 3).



(a) (b)

Figure 3. (a) the right image from a stereo pair. (b) is the same image overlayed with the reference standard. The border is the boundary of the optic disc.

3.2. Independent Standard

Three ophthalmologists, widely respected glaucoma specialists, carefully marked all pixels of the disc on each of the left images of the stereo pair. Planimetry of the stereo pairs using a stereo viewer was utilized. The experts were requested to include blood vessels into their classification of the surrounding tissue; in other words a part of a vessel surrounded by disc was classified as disc. Two-class reference standards were obtained by a hard winner-take-all threshold of the three expert classifications. The two classes are disc and background.

3.3. Error Indices and Data Analysis

To objectively compare computer-detected borders against the reference standard, absolute and signed mean border positioning errors and root-mean-square border positioning errors are computed and expressed in pixels (1 pixel $\approx 8 \times 8 \mu m^2$). Corresponding points are defined as pairs of points where the first point p_1 is taken from the computer-detected border and the second point p_2 is the closest point to p_1 on the reference standard border using the Euclidean distance.

As the borders are defined by 2-D splines, the nearest point p_2 with respect to p_1 can be calculated quickly using a gradient descent approach [10]. This can be optimized using a combination of Newton's method with quadratic minimization [17]. Because of the length of each spline and the possible points in our discrete space, we achieved comparable results by simply testing a predetermined number of samples ($N = 10$) along each interval of the 2-D spline.

A classification accuracy is also reported, reflecting the number of correctly classified pixels. The accuracy is based on a balanced sampling of the two classes from the reference standard. The accuracy is computed using

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

where statistics are collected from the sampled pixels for the number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN).

4. RESULTS

We compare our algorithm against the manual segmentations by three glaucoma fellows to provide a realistic benchmark of current segmentations. Also, we compare the method to standard graph search approaches, involving costs computed from simple edge features. The three methods we use are

1. Gaussian smoothing ($\sigma = 5.6$) of the grayscale intensity followed by Canny edges ($r = 8$ pixels).
2. Canny edges ($r = 8$ pixels) of the blue-yellow color plane.
3. Canny edges ($r = 16$ pixels) of the blue-yellow color plane.

The blue-yellow plane is used because it tends to emphasize the optic disc [3]. These three costs were selected for performing the best over our feature set, including features computed on the green plane which is much more widely used.

An expert determined that 5 of the 82 datasets had resulted in a complete failure. The statistics are computed on the remaining 77 datasets to more accurately display the performance of the algorithm in normal operation. A numerical comparison of the different methods is in Table 1 along with some example results in Figure 7.

A list of the most significant features can be seen in Figure 4. The features were calculated from a large pool of features including different scale smoothness operators, edge detectors and first and second order derivatives [18]. Each feature is computed on the intensity, red-green and blue-yellow color planes. The probability map computed from the selected features is shown along with the final segmentation in Figure 5. Intermediate results demonstrating the steps of our algorithm are shown in Figure 6.

5. DISCUSSION

This paper presents a novel and superior method to generate a cost function for incorporating region-based information into a graph search. Utilizing a kNN classifier which can be initialized from a training set, the input feature vectors are combined in an intelligent way using the soft classification technique to create probability maps. We have developed a new feature selection scheme to optimize the cost function design for segmenting closed objects. The kNN classifier enables the use of standard feature selection techniques to optimize the feature set for the given application in an automated way (e.g. SFFS) [10]. The cost function described in §2.3.3 maximizes the probability of the disc using graph search to find a min-cost shortest path.

		Border Distances (pixels)		
Method	Accuracy	Signed Error	Absolute Error	RMSE
Fellow 1	94.02%	-3.14 ± 5.06	4.84 ± 4.22	5.94
Fellow 2	97.39%	1.92 ± 3.55	3.64 ± 2.76	4.02
Fellow 3	95.60%	-1.57 ± 3.22	3.77 ± 1.97	3.56
Our Algorithm	93.89%	9.59 ± 8.78	11.27 ± 7.17	12.96
Alt. Method 1	75.35%	-15.99 ± 18.91	24.04 ± 10.59	24.67
Alt. Method 2	74.51%	-17.02 ± 20.59	25.31 ± 11.80	26.62
Alt. Method 3	78.54%	-15.78 ± 24.81	23.02 ± 18.69	29.26

Table 1. Performance results from different segmentation methods. Note that our algorithm performs considerably better than the standard methods and shows potential in approaching the accuracy of the fellows.

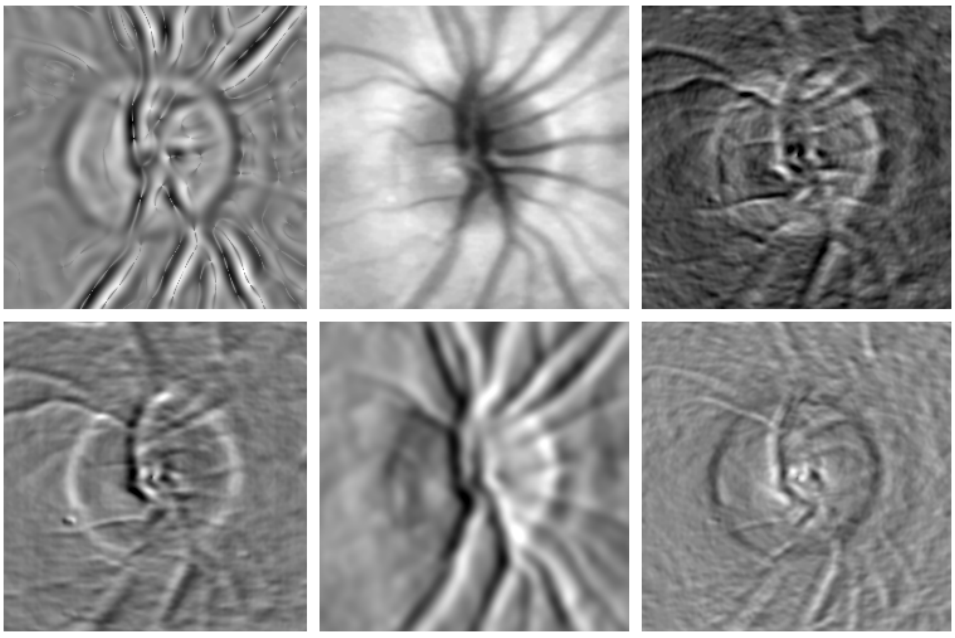


Figure 4. Features selected from the SFFS algorithm. Listed features are calculated on different color planes: DB, RG, BY (intensity, red-green and blue-yellow, respectively). **Top row** (left to right): BY - Invariant Gaussian derivative Lww $\sigma = 5.6$, RG - Gabor smoothing $\sigma = 1.0, \theta = 0.0°, \phi = 0.0$, $\omega = \pi$, RG - Radial Gaussian derivative $\sigma = 1.4$. **Bottom row** (left to right): DB - Radial Gaussian derivative $\sigma = 1.4$, RG - Gaussian derivative $\sigma = 4.0$, BY - Radial Gaussian derivative $\sigma = 1.4$.

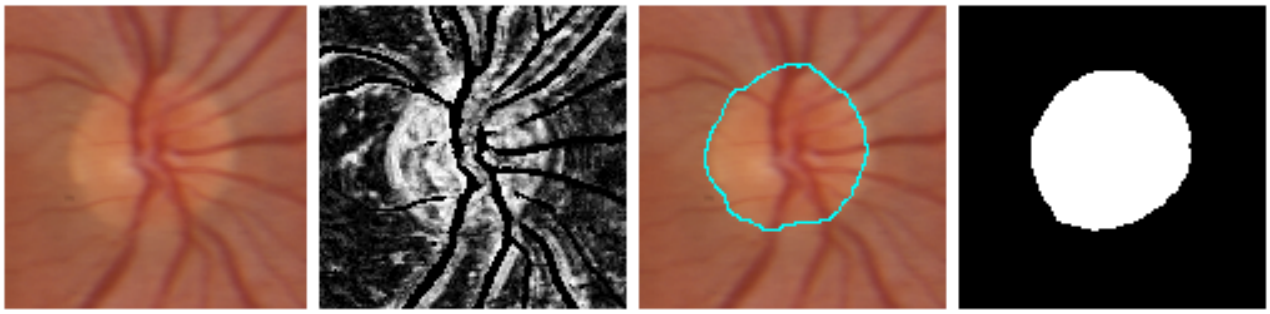


Figure 5. Results for selected features. From left to right, the left image of the original stereo pair, the resulting probability map, the final border and the reference standard. The probability map is computed from the selected features shown above.

Figure 6. Intermediate results. On the left is the probability map after unfolding into polar coordinates. In the middle is the cost function, and on the right is the computed border overlayed on the cost function. The images are sampled with $I = 72, J = 128$ and a radius of 128 pixels. Note, without loss of generality, the original stereo images are resized to 256×256.

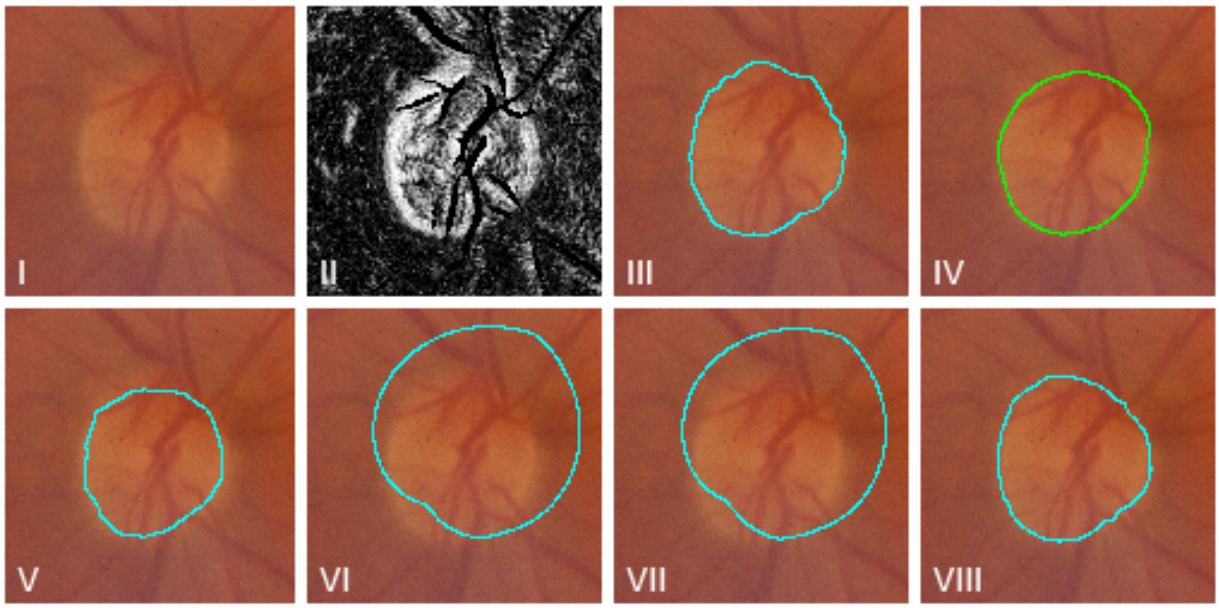
The segmentation tends to depend on the quality of the classification within the disc. For example, in Figure 7b, the lower right portion of the border is drawn into the disc because of the poor classification within that particular area. However, the segmentation provides a good approximate region that can be used for further processing of the disc such as determining the location of the 'cup' [3, 4]. The cupping of the optic nerve head is an important motivation for our investigations, as it can be helpful in determining the onset of glaucoma in patients [1]. While the algorithm provides a good approximation, there are 5 complete failures within the dataset that reduce the accuracy. Even with these difficulties the performance approaches that of the trained fellows.

In the stereo color images of the optic nerve there is poor contrast between the background and the disc. This problem can be seen clearly in the example results. To reduce the effects of poor contrast in the segmentation there are some potential methods to investigate. Currently the segmentation is based on features derived from the left image in the stereo pair. Incorporation of the stereo information present in the images could result in a better signal to noise ratio and improved contrast. While there are stereo and disparity features [19] within our feature set, they are not found to be significant using the SFFS feature selection. It is apparent from the results that segmentation of the disc requires the use of better stereo features. Also, it would be beneficial to the algorithm to incorporate the general shape of the disc into the segmentation.

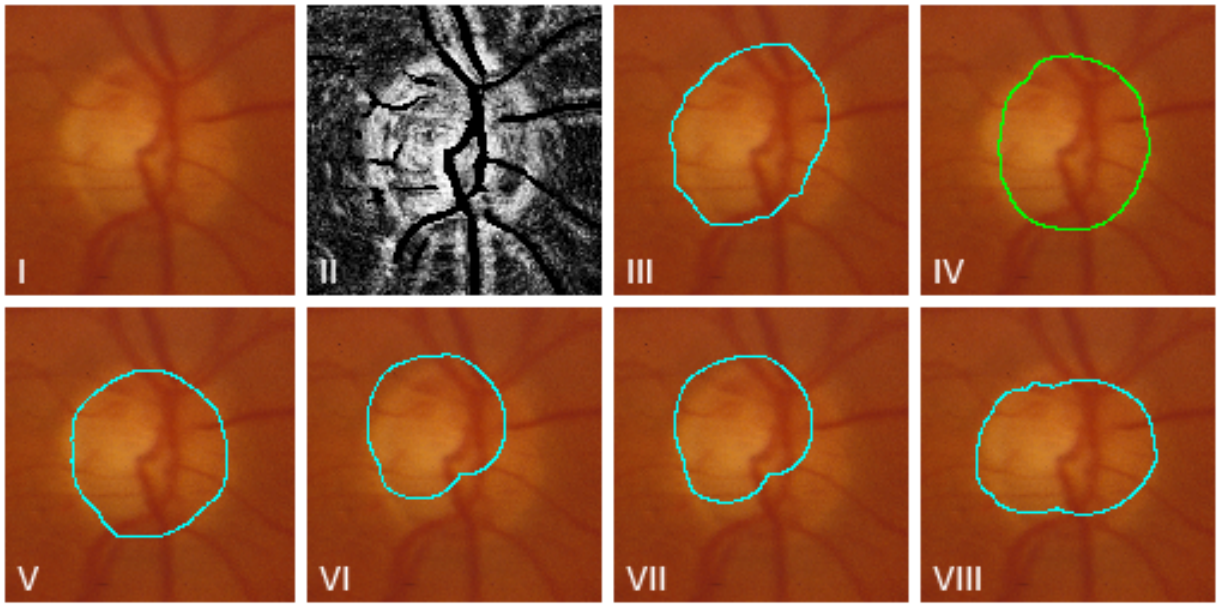
The optic disc segmentation is only one application of our general algorithm. The use of feature selection to optimize a cost function is useful when dealing with objects that do not exhibit strong edge or region characteristics. The combination of pixel classification and graph search automatically combines the features optimally depending on the training data. Classification greatly reduces the dependence on determining linear combinations of features which may not be optimal when trying to compare region-based features with edge-based features.

6. CONCLUSION

A new and innovative algorithm has been developed and validated against clinical data capable of segmenting the optic disc in stereo retinal images. Pattern classification techniques provide a flexible method for creating a



(a)



(b)

Figure 7. Example results for two different datasets (a) and (b). **Top row:** (I) the left image of the stereo pair, (II) the probability map, (III) our segmentation result and (IV) the ground truth. **Bottom row:** (V) the segmentation from fellow 3 and (VI - VIII) alternative methods 1, 2, and 3.

cost function in an automated way from the combination of a large set of features. We have also demonstrated the usefulness of feature selection to greatly improve the results of our method. A region-based cost function is also created in polar coordinates for maximizing the probability of a region by reformulating it as a border detection problem. Results are represented as 2-D splines which make further processing steps easier.

ACKNOWLEDGMENTS

This work was supported by grants to MDA: National Eye Institute R01-EY017066; Department of Defense A064-032-0056; Research to Prevent Blindness, NY, NY; Wellmark Foundation.

REFERENCES

1. A. Heijl, M. C. Leske, B. Bengtsson, L. Hyman, and M. Hussein, "Reduction of intraocular pressure and glaucoma progression: results from the early manifest glaucoma trial," *Archives of Ophthalmology* **120**, pp. 1268–1279, Oct. 2002.
2. J. M. Tielsch, J. Katz, H. A. Quigley, N. R. Miller, and A. Sommer, "Intraobserver and interobserver agreement in measurement of optic disc characteristics," *Ophthalmology* **95**, pp. 350–356, Mar. 1998.
3. M. D. Abràmoff, W. L. M. Alward, E. C. Greenlee, L. Shuba, C. Y. Kim, J. H. Fingert, and Y. H. Kwon, "Automated segmentation of the optic nerve head from stereo color photographs using biologically plausible feature detectors." Investigative Ophthalmology and Visual Sciences [in press], 2007.
4. M. B. Merickel, X. Wu, M. Sonka, and M. D. Abràmoff, "Optimal segmentation of the optic nerve head from stereo retinal images," *Proceedings of SPIE Medical Imaging* **6143**, 2006.
5. F. Mendels, C. Heneghan, P. D. Harper, R. B. Reilly, and J.-P. Thiran, "Extraction of the optic disc boundary in digital fundus images," in *Proceedings of the 1st Joint BMES/EMBS Conference*, p. 1139, 1999.
6. F. Mendels, C. Heneghan, and J. Thiran, "Identification of the optic disk boundary in retinal images using active contours," *Proceedings of the Irish Machine Vision and Image Processing Conference* , pp. 103–115, 1999.
7. J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, and E. Fletcher, "Optic nerve head segmentation," *IEEE Transactions on Medical Imaging* **23**(2), pp. 256–264, 2004.
8. M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, Brooks/Cole Publishing Company, second ed., 1999.
9. N. Paragios and R. Deriche, "Coupled geodesic active regions for image segmentation: A level set approach," *Proceedings of the European Conference on Computer Vision* , pp. 224–240, June 2000.
10. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Wiley-Interscience, second ed., 2001.
11. M. Niemeijer, J. S. Staal, B. van Ginneken, M. Loog, and M. D. Abràmoff, "Comparative study of retinal vessel segmentation on a new publicly available database," in *Proceedings of the SPIE*, pp. 5370–5379, 2004.
12. S. Arya, D. Mount, N. Netanyahu, R. Silverman, and A. Wu, "An optimal algorithm for approximate nearest neighbor searching in fixed dimensions," *Journal of the ACM* **45**, pp. 891–923, 1998.
13. P. Pudil, F. J. Ferri, J. Novovicova, and J. Kittler, "Floating search methods for feature selection with non-monotonic criterion functions," *International Conference on Pattern Recognition* **2**, pp. 279–283, Oct. 1994.
14. D. Z. Chen, J. Wang, and X. Wu, "Image segmentation with asteroidality/tubularity and smoothness constraints," *International Journal of Computational Geometry and Applications* **12**, pp. 413–428, Oct. 2002.
15. D. R. Thedens, D. J. Skorton, and S. R. Fleagle, "Methods of graph searching for border detection in image sequences with applications to cardiac magnetic resonance imaging," *IEEE Transactions on Medical Imaging* **14**(1), pp. 42–55, 1995.
16. K. Atkinson and W. Han, *Elementary Numerical Analysis*, John Wiley and Sons, Inc., third ed., 2004.
17. H. Wang, J. Kearney, and K. Atkinson, "Robust and efficient computation of the closest point on a spline curve," in *Proceedings of the 5th International Conference on Curves and Surfaces*, pp. 397–406, June 2002.
18. B. M. ter Haar Romeny, *Front-End Vision and Multi-Scale Image Analysis*, Computational Imaging and Vision, Kluwer Academic Publishers, 2003.
19. N. Qian and R. A. Andersen, "A physiological model for motion-stereo integration and a unified explanation of pulfrich-like phenomena," *Vision Research* **37**(12), pp. 1683–1698, 1997.