Published in final edited form as:

Invest Ophthalmol Vis Sci. 2007 April ; 48(4): 1665–1673. doi:10.1167/iovs.06-1081.

Automated segmentation of the optic nerve head from stereo color photographs using physiologically plausible feature detectors

Michael D. Abràmoff, MD, PhD^{1,2,3}, Wallace L. M. Alward, MD^{1,2}, Emily C. Greenlee, MD^{1,2}, Lesya Shuba, MD, PhD¹, Chan Y. Kim, MD, PhD¹, John H. Fingert, MD, PhD¹, and Young H. Kwon, MD, PhD¹

¹ Department of Ophthalmology and Visual Sciences, University of Iowa Hospitals and Clinics, 200 Hawkins Drive, Iowa City, IA 52242, USA

² Department of Veterans Affairs, Iowa City VA Medical Center, 601 Highway 6 West, Iowa City, IA 55242, USA

³ Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242, USA

Abstract

Purpose—To evaluate a novel automated segmentation algorithm for cup-to-disc segmentation from stereo color photographs of glaucoma patients for glaucoma progression measurement.

Methods—Stereo color photographs of the optic disc were obtained using a fixed stereo-base fundus camera (Nidek 3DX) from 58 eyes of 58 patients with glaucoma suspect or open-angle glaucoma. Manual planimetry by three glaucoma faculty was performed to delineate a reference standard rim and cup segmentation of all stereo pairs, and by three glaucoma fellows as well. Pixel feature classification was evaluated on the stereo pairs and corresponding reference standard, using feature computation based on simulation of photoreceptor color opponency and visual cortex simple and complex cells. An optimal subset of twelve features was used to segment all pixels in all stereo pairs, and the percentage of pixels assigned the correct class and linear cup-to-disc ratio (LCDR) estimates of the glaucoma fellows and the algorithm were compared to the reference standard.

Results—The algorithm was able to correctly assign cup, rim or background to 88% of all pixels. Correlations of the LCDR estimates for glaucoma fellows with the reference standard were 0.73 (95% CI, 0.58 – 0.83), 0.81 (95% CI, 0.70 – 0.89) and 0.86 (95% CI, 0.78 – 0.91), respectively, while correlation of the algorithm with the reference standard was 0.93 (95% CI, 0.89 – 0.96) (n=58).

Conclusions—The pixel feature classification algorithm allows objective segmentation of the optic disc from conventional color stereo photographs automatically without human input. The performance of the disc segmentation and LCDR calculation of the algorithm was comparable to glaucoma fellows in training, and is promising for objective evaluation of optic disc cupping.

Introduction

Glaucoma is the second leading cause of blindness in the US, characterized by gradual damage to the optic nerve and resultant visual field loss ¹. Early diagnosis and optimal treatment have been shown to minimize the risk of visual loss due to glaucoma ². The hallmark of glaucomatous progression is cupping of the optic nerve head. One way of determining the amount of cupping is planimetry by experienced glaucoma specialists from stereo color photographs of the optic nerve head ^{3, 4}.

Disclosure:

Michael D. Abràmoff: P, Young H. Kwon: P, others N

Pixel feature classification is a machine learning technique able to assign a class to the pixels in an image, in this case either cup, rim, or background^{5–8}. Pixel feature classification uses multiple pixel features, which are numeric properties of a pixel and its surroundings. Originally, pixel intensity was used, and later, its contrast with the surrounding region, its proximity to an edge and so forth were added. As a supervised algorithm, it has two stages: a training stage, where the algorithm ‘statistically learns’ to correctly classify pixels from known classifications, and a testing or classification stage where the algorithm classifies a new previously unseen image. The number of features for any pixel is essentially infinite, and it is easy to prove that one or more subsets of this infinite set are optimal for classifying the image according to some reference standard. Hundreds of features for a pixel can be calculated in the training stage to cast as wide a net as possible. Still, because all potential features cannot possibly be taken into account, such a selection is still arbitrary to some extent and carries a bias. One of the advantages of the pixel feature classification method is that it avoids the potentially much larger bias of an imaging expert developing the algorithm selecting one or two feature detectors. This is common in unsupervised image classification, such as edge-based, deformable model or template-based algorithms, that have been used in segmentation of retinal images and optic disc in non-stereo images⁹. Pixel feature classification has so far outperformed all other approaches for the segmentation of vessels, hemorrhages and microaneurysms, exudates, infarcts and drusen in retinal images^{7, 8} (Niemeijer, et al. *IOVS* 2005;45:ARVO E-Abstract 3468).

Another advantage of pixel feature classification is its potential to mimic human experts’ visual performance utilizing parallel visual ‘channels’¹⁰. We have designed pixel features for color mimicking the output of light-dark, red-green, and blue-yellow center surround opponency cells, with different receptive field sizes and orientation, and features for stereo disparity that mimic the output of binocular cells in area V1 of the visual cortex^{11, 12}.

The purpose of this study was to compare the performance of pixel feature classification to segment the optic disc into cup, rim and background from color stereo photographs and the performance of glaucoma fellows to the results of 3 glaucoma faculty experts, for glaucoma progression measurement.

Specifically, we compared the performance of the pixel classification algorithm to segment the optic disc photographs into cup, rim, and background to the planimetric performance of the fellows against the reference standard planimetric segmentation by the glaucoma experts.

Methods

Stereo color images of the disc

The study was approved by our institute’s Institutional Review Board, and adhered to the tenets of the declaration of the Declaration of Helsinki. For the test set, stereo photographs of the optic disc were obtained from 58 consecutive patients from the Glaucoma Service at the University of Iowa. Patients were included if they met the definition of glaucoma suspect or open-angle glaucoma; those with a history of angle closure or combined mechanism glaucoma, or any non-glaucomatous optic neuropathy were excluded. The diagnosis of ‘glaucoma suspect’ was based on suspicious optic nerve appearance (enlarged cupping on clinical examination) with normal visual field and intraocular pressure (IOP) ≤ 21 mmHg; it also included those with normal optic disc appearance on biomicroscopy and normal visual field, but with elevated IOP (> 21 mmHg). The ‘open-angle glaucoma’ diagnosis included patients with both primary and secondary open-angle glaucoma, and was based on the presence of an open iridocorneal angle, glaucomatous optic disc and/or nerve fiber layer defects on biomicroscopy, and visual field changes (regardless of IOP level). Glaucomatous optic discs were identified as those with either diffuse or focal thinning of the neuroretinal rim. Visual

field abnormalities were considered to be glaucomatous if they were consistent with the optic nerve examination and had either a (1) typical nerve fiber layer distribution, or (2) glaucoma hemifield test outside normal limits. The diagnoses were made by the treating faculty glaucoma specialist according to the above definitions.

Color slide stereo photographs centered on the optic disc of both eyes were acquired using a fixed stereo-base Nidek 3Dx stereo retinal camera (Nidek, Newark, NJ); after development, slides were scanned at 4096×4096 pixel resolution, 24 bit depth with a Kodak slide scanner (Kodak Inc., Rochester, NY). Only one eye of each patient was selected for the study based on the quality of the stereo image. To limit the number of pixels, over 16 million for each left-right stereo pair, that would otherwise have to be processed by the algorithm, the left and right stereo images of the selected stereo pairs were cropped to 512×512 pixels each with the optic nerve head in the center, by automatically locating the optic disc in the 4096×4096 images, and cropping to 256 pixels outwards from the center of the disc based on its brightness¹³. The cropped images included the optic disc in its entirety in all images. See Figure 1.

Computer-aided planimetry of the stereo pairs using a hand-held stereoscope was performed by 3 faculty glaucoma specialists at the University of Iowa on the 58 stereo pairs using a custom-made annotation software¹⁴. Additionally, three glaucoma fellows in our Glaucoma Fellowship performed planimetry on all stereo pairs in the same manner. The human graders were masked from the patient identifier, the clinical information associated with the photographs, and the results of all other graders, at the time of grading. Graders were instructed to outline the disc margin while viewing the photographs using a stereo viewer. Then, they were instructed to draw the cup to the best of his or her ability. When the cup margin was difficult to identify (usually at the temporal slope of the disc), they were instructed to identify and draw the cup margin on the slope that is half way between the bottom of the cup and top of the rim.

A *reference standard* grading with a unique class (cup, rim, background) assigned to each pixel was then created by combining the grading by each faculty grader in a “majority-win” manner: each pixel was assigned the class that received the majority of votes (e.g. if two graders voted a pixel to be rim and one cup, the pixel assigned was rim). In the case of a draw, the pixel was assigned the background class. See Figure 2.

Pixel feature classification

Rationale for inclusion of specific features

Originally, pixel intensity and simple edge operators were used as pixel features⁵. Once it was realized that a pixel’s feature can also incorporate the adjacent pixels, i.e. its surround, and the Gaussian filterbank concept had been established mathematically¹⁵, it was possible to use filterbank outputs to create a larger set of features, from which an optimal combination of features could be selected using feature selection, initially in pulmonary image analysis¹⁶. Gaussian filterbank features are sensitive to edges and textures in the grayscale intensity image obtained from a color image, at different scales and orientations. To more optimally use the color information in fundus color images we added the outputs of a Gaussian steerable, i.e., multiple orientations, filterbank in hue, saturation, brightness space as features and also the variance in the R,G and B channels in a small region around the pixel⁸.

Such features are obviously different from the color opponency model of primate color vision, and for this study we have added the output of a Gaussian steerable filterbank in color opponency space for dark-bright, red-green and blue-yellow center-surround color opponency. See Appendix A for the mathematical details. Preliminary experiments showed that the color opponency filterbank improved the performance. See Figure 3.

Similarly, because expert human graders require stereo photographs for optimal segmentation of the optic nerve head, we decided to add horizontal stereo disparity maps as features. In the literature, many known algorithms use edge and texture information and some form of energy minimization to solve this depth from stereo problem in color or video images¹⁷. These algorithms are fundamentally different from what is known about stereo disparity computation in primates¹⁸. In preliminary studies, we found that the stereo disparity algorithms that have been found to be optimal for natural scene stereo disparity, namely the graph search algorithms first described by Kolmogorov and Birchfield, did not perform as well as expected on the low contrast/high noise images in this study^{17, 19}. Therefore, we decided to add features calculated by simulating the output of binocular complex cells in V1 visual cortex as first described by Qian, et al.¹². The output of binocular simple cells in the mammalian V1 has been shown to closely correspond to the convolution of its receptive field with Gabor basis kernels with the same spatial frequency and scale but different phases for the left and right field. Stereo disparity sensitive complex cells can then be modeled by summing the squared outputs of two simple cells with different phases, but the same size receptive field. See Appendix B for the mathematical background.

An informal taxonomy of pixel features obtainable from stereo color images may be:

- ‘less physiologically plausible’: the outputs of edge detectors such as those described by Canny and Sobel, or graph search stereo disparity algorithms, as there is currently no evidence that these are expressed by any neural circuit^{17, 20}.
- ‘somewhat physiologically plausible’: the outputs of Gaussian filterbanks in intensity, or RGB, space, because there is evidence that a Gaussian profile is expressed by certain neural circuits²¹. No Red/Green/Blue or Hue/Saturation/Intensity sensitive receptive fields have been described to our knowledge.
- ‘more physiologically plausible’: Gabor wavelets in color opponency space, and the output of simulations of binocular complex cells, as there is evidence that these are expressed in certain neural circuits^{22, 23}

We have assumed that a combination of retinal and cortical image processing as found in biological systems would be effective in classifying the stereo color images of the optic disc, and thus have implemented these physiologically plausible features and added them to the set of potential features from which an optimal set of features is then selected. Whether these features indeed increase the performance of the system can be studied (implicitly) by running feature selection on sets with and without the ‘more physiologically plausible’ features and comparing the accuracy.

A priori expectation feature

In pixel feature classification, non-pixel based information is not naturally included. However, it is easy to see that in stereo disc color image of this type, the likelihood of a pixel in the center of the image being cup is higher than it being rim or background. To capture this a priori expectation of the class of a pixel based on its location in the image, an a priori rim/non-rim feature was calculated by performing Principal Component Analysis on all reference standard images and taking the first principal component²⁴. This feature expresses the expectation of any pixel being rim or either background or cup based on its position. Preliminary studies without the a priori feature did not attain an accuracy over 73% (see Results).

A total of $108 + 3 \times 48 + 1$ features were thus calculated, for all 15,204,352 pixels in all 58 color stereo pairs. Because these features can be calculated independently in parallel, as is the case in the primate visual system, this is a naturally parallel computation. See figure 4 for an example of all values of the twelve optimal pixel features that were selected for a specific stereo pair.

Classifier implementation

In pixel feature classification, a classifier is trained on a subset of all data, the train set, and then repeatedly tested on another subset of the data, the test set. By repeatedly testing with different subsets of all of the features, and comparing the performance of the classifier with these different feature subsets, the optimal subset of features is selected. In preliminary studies, the k-nearest neighbor (kNN) classifier outperformed Support Vector Machines and Linear Discriminant Analyzer classifiers, as was found in other studies^{8, 16, 25, 26}. A kNN classifier has the advantage that no elaborate training is required and that only one parameter needs to be adjusted, the k. We have interfaced the publicly available knn library, available at www.cs.umd.edu/~mount/ANN, to Java²⁷.

Data analysis

The classifier was optimized on the accuracy as explained in Appendix C to find the set of optimally discriminant features. A clinically more relevant measure was also used to compare the performance of the pixel feature classification and the three glaucoma fellows to the reference standard. For this purpose, all pixels in all stereo pairs were classified in a leave-one-out fashion²⁵. The resulting gradings contained the algorithm's estimate for each pixel as to whether that pixel was cup, rim or background, similar to the gradings by the glaucoma faculty and the glaucoma fellows. A linear cup-to-disc ratio was then computed for each graded image as a square root of area cup-to-disc ratio:

$$cd_{r,linear} = \sqrt{\frac{n_{c,p}}{n_{r,p} + n_{c,p}}}$$

with $n_{r,p}$ the number of pixels in each image that pixel feature classification classified as rim²⁸.

Optimal feature selection

The algorithm had been previously trained on a larger set of 101 stereo disc pairs of glaucoma patients (Abràmoff, et al., *IOVS* 2006;46:ARVO E-Abstract 3635). The training set was used to determine which combination of features gives the best accuracy ("feature selection", see Appendix C). A hard classification using 'majority win' of the k nearest neighbors was used as the classification for each pixel. Each feature column in the training set was scaled to have zero mean and unit variance over all pixels. The most discriminant features were then selected in a Sequential Forward Floating Search²⁵, by adding all features to an initially empty feature set one by one and retaining only the new feature that showed the highest gain in classification performance. Preliminary studies showed that $k=501$ gave slightly better performance over a range of feature vector lengths, but sensitivity to changing from 1 to 1001 was low. The number of samples was 1200 per image, per class. The pixel feature classification was implemented in Java, and it currently takes 15 minutes to calculate twelve features and completely grade a stereo pair on a standard (2005) 2GHz 2GBytes PC under Windows XP.

Table 1 shows the twelve features that were found to be most dominant. Figure 5 shows an example stereo pair and those pixel features that were most discriminant features. After feature selection, the algorithm was ready for testing using this optimal set of features.

Results

The stereo pairs of 58 eyes of 58 patients were included to test the algorithm. The age of these patients was 61.2 ± 17.8 years (mean $\pm$ SD), and there were 33 females. There were 22 glaucoma

suspect and 36 open-angle glaucoma patients. Of the total of 58 eyes, 27 eyes underwent Humphrey visual field testing (24-2 SITA Standard, HFA II, Humphrey-Zeiss, Dublin, CA) within 1 year of the disc photographs. The remainder (n=31) had either Goldmann visual fields (n=28, Goldmann Perimeter 940, Haag Streit, Liebefeld, Switzerland) or no fields (n=3). We performed Goldmann visual fields in patients whose Humphrey visual field was not reliable and/or had moderate to severe visual field defect (typically -15dB MD or worse). In these patients, we have found the Goldmann visual fields to be more reliable. Few patients (n=3) could not perform any visual fields at all (e.g. Alzheimer's disease). The Humphrey visual field mean defect ($\text{dB} \pm \text{SD}$) in eyes with glaucoma suspicion (n=11) and open-angle glaucoma (n=16) were -1.1 ± 2.3 and -4.1 ± 5.2 respectively. The Humphrey visual field pattern standard deviation ($\text{dB} \pm \text{SD}$) in eyes with glaucoma suspicion and open-angle glaucoma were 2.8 ± 3.6 and 4.5 ± 4.5 respectively. The Goldmann visual fields of the glaucoma suspect group (n=8) were within normal limits. The Goldmann visual fields of the open-angle group (n=20) had a variety of defects consistent with glaucomatous nerve fiber layer loss including paracentral scotoma, nasal step, and arcuate defect(s).

Average disc size (cup and rim area as determined from the reference standard) was 14085 pixels ($\pm 1 \text{ SD } 3467$ pixels). All pixels in all stereo pairs were graded by the algorithm. Accuracy was 0.876; the algorithm was able to assign 88% of all pixels out of the 58 stereo pairs to the correct class. Figure 5 shows how the gradings by the glaucoma faculty, the glaucoma fellows and the algorithm compare to each other visually. Correlations of the linear cup to disc ratio for each of the glaucoma fellows with the reference standard were 0.73 (95% Confidence Interval (CI), 0.58 – 0.83), 0.81 (95% CI, 0.70 – 0.89) and 0.86 (95% CI, 0.78 – 0.91), respectively, while correlation of the algorithm with the reference standard was 0.93 (95% CI, 0.89 – 0.96) (n=58).

Subgroup analysis showed no significant differences in the correlation between the glaucoma suspect and open-angle glaucoma groups for each of the fellows or the algorithm, and there was no dependency on disc size. ($p > 0.05$). Figure 6 shows scatter-plots of the estimates for the linear CDR for each stereo pair for the glaucoma fellows and the algorithm versus the reference standard.

Discussion

Our results show that pixel feature classification with biologically inspired features can classify a set of optic disc stereo images with an accuracy that is comparable to that of ophthalmologists in training. Other studies have also shown agreement of pixel classification algorithms for the segmentation of retinal vessels and the detection of diabetic retinopathy lesions with human experts^{7, 8, 29}. In fact, Niemeijer, *et al.*, found that a pixel classification based approach for retinal vessel segmentation outperformed other supervised and unsupervised methods⁷. The same group of authors used a pixel classification approach to detect red diabetic retinopathy lesions and reached a sensitivity of 100%, a specificity of 87%, and an area under the ROC curve of 98.5% in a study of 100 patients⁸.

We are aware of only one previous study of classification of the optic nerve head using stereo photographs³⁰. The authors utilized a matching cost approach, based on pyramidal cross correlation. Recent depth from stereo algorithms, mostly based on graph searches, have been shown to outperform 'matching cost computation' algorithms on real world depth scenes^{17, 19, 31}. As shown by our results, these graph search algorithms in their turn can be outperformed by an algorithm that mimics binocular cells in visual cortex in optic disc stereo images¹².

On a computational level, the 'role' of most features is how much a patch around the pixel resembles the filter kernel of that feature. The results indicate that the features selected for

optimal accuracy of the system were predominantly of the ‘somewhat’ and ‘more’ physiologically plausible categories, and not for example, the familiar Sobel or Canny intensity edge detectors. On an image understanding level, it is difficult to explain the ‘role’ of the selected features. A second order Gaussian derivative of a specific orientation and blue-yellow color opponency can be said to correspond to “striped oriented patches” of the pixel surround, but is more than a simple edge detector, and also different from a texture detector. This situation is analogous to the difficulty in describing the role of different receptive field properties for scene analysis by the primate visual cortex.

Objective classification of optic disc cup and rim from color stereo photographs is promising, because it allows optic disc stereo photographs to be evaluated objectively and rapidly. However, the ultimate goal of an algorithm such as this is the objective measurement of glaucoma progression, or enlargement of the optic cup, over time. The results of the present study are foremost intended to serve as a proof of concept, and many aspects can potentially be improved, including adding features which mimic physiology even more, and adding vessel segmentation.

Nevertheless, several serious issues remain to be addressed before the present algorithm can be thus used. First, the algorithm is deterministic and therefore will give the same classification if offered the same stereo pair repeatedly. However, whether it also gives very similar classifications on two stereo pairs of the same patient of the same eye taken minutes apart, with subsequent patient movement, vignetting, and refocusing, has not been determined.

Second, there is room for substantial improvement as is evident from the noisy visual appearance of the classification, compared to the smooth boundaries of the reference standard, and the percentage of correctly classified pixels, which is only 88%. For example, compared to optic disc images obtained with the Zeiss fundus camera, the Nidek 3Dx gives decidedly darker and less crisp images, which may have impacted the quality of the performance negatively. We have adopted several different post-processing techniques that are showing promising results (Merickel, M.B. jr, et al., Optimal segmentation of the optic nerve head from stereo retinal images, SPIE Medical Imaging, Vol. 6143).

Third, the algorithm performs well on stereo pairs obtained with the fixed base Nidek 3DX, which may not be widely available, and it has not been established whether the algorithm will perform similarly on non fixed base stereo pair cameras.

Fourth, the algorithm’s performance is highly dependent on a single feature, the prior PCA. One of its roles is likely be to indicate that the disc is in the center of the image, and the cup is in the center of the disc, and surrounded by rim. Because this is an assumption about the expected size and location of the disc, the algorithm’s accuracy may be lower on discs that are smaller or larger than in our sample, and also if the disc is not properly centered in the cropped image. Therefore, we are currently engaged in removing this feature from the selected features while maintaining performance.

Finally, the algorithm is slow. It is important to understand that the current implementation of the algorithm was not written for speed but for ease of use and modifiability, and can easily be optimized for speed. In our previous optimized algorithms, classification times of 15s for a whole image were reached⁸. Additionally, calculation of the features and feature classification is naturally parallel, so that with appropriate hardware, the features for all pixels can be calculated in parallel.

Contrary to expectations, stereo disparity was not the most discriminant feature for the performance of the algorithm. In Figure 1 it can be appreciated that the stereo photographs are not very crisp. Possibly, using a pair of photographs improves the quality of the fused image

by suppressing these irregularities to some extent. In addition, human graders may be unaware of the relative importance of color and texture of the tissues compared to depth-from-stereo as the features they use to segment the optic nerve head. This may imply that a similar algorithm may also perform reasonably on non-stereo color photographs.

In summary, the results of this study have shown that pixel feature classification with biologically inspired features of stereo color images is a good starting point for the classification of the optic nerve head and other three-dimensional structures on the retina, and also that features based on physiological process outperform standard pixel features on segmentation of the optic nerve head.

Acknowledgments

MDA: NEI R01-EY017066; Research to Prevent Blindness, NY, NY; Wellmark Foundation; Department of Defense A064-032-0056; US Department of Agriculture; University of Iowa; Netherlands Organization for Health Related Research;

WLMA: NEI, Lew R Wasserman awardee from Research to Prevent Blindness

CYK: Unrestricted educational grant from Allergan Corp

References

1. Tielsch JM, Sommer A, Katz J, Royall RM, Quigley HA, Javitt J. Racial variations in the prevalence of primary open-angle glaucoma. The Baltimore Eye Survey. *JAMA* 1991;266:369–374. [PubMed: 2056646]
2. Heijl A, Leske MC, Bengtsson B, Hyman L, Bengtsson B, Hussein M. Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Arch Ophthalmol* 2002;120:1268–1279. [PubMed: 12365904]
3. Greaney MJ, Hoffman DC, Garway-Heath DF, Nakla M, Coleman AL, Caprioli J. Comparison of optic nerve imaging methods to distinguish normal eyes from those with glaucoma. *Invest Ophthalmol Vis Sci* 2002;43:140–145. [PubMed: 11773024]
4. Tielsch JM, Katz J, Quigley HA, Miller NR, Sommer A. Intraobserver and interobserver agreement in measurement of optic disc characteristics. *Ophthalmology* 1988;95:350–356. [PubMed: 3174002]
5. Panda DP, Rosenfeld A. Image Segmentation by Pixel Classification in (Grey Level, Edge Value) Space. *IEEE Trans on Computers* 1978;27:875–879.
6. Spence, CD.; Sajda, P. Role of feature selection in building pattern recognizers for computer-aided diagnosis. Hanson, KM., editor. 2006. p. 1434-1441.
7. Niemeijer M, Staal JS, van Ginneken B, Loog M, Abràmoff MD. Comparative study of retinal vessel segmentation on a new publicly available database. *Proceedings of the SPIE* 2004;5370–5379.
8. Niemeijer M, vanGinneken B, Staal J, Suttorp-Schulten MS, Abramoff MD. Automatic detection of red lesions in digital color fundus photographs. *IEEE Trans Med Imaging* 2005;24:584–592. [PubMed: 15889546]
9. Lowell J, Hunter A, Steel D, et al. Optic nerve head segmentation. *IEEE Trans Med Imaging* 2004;23:256–264. [PubMed: 14964569]
10. Tailor DR, Finkel LH, Buchsbaum G. Color-opponent receptive fields derived from independent component analysis of natural images. *Vision Research* 2000;40:2671–2676. [PubMed: 10958917]
11. Serre, T.; Riesenhuber, M. CBCL Paper #239/AI Memo #2004-004. Massachusetts Institute of Technology; Cambridge, MA: 2004. Realistic Modeling of Simple and Complex Cell Tuning in the HMAX Model, and Implications for Invariant Object Recognition in Cortex.
12. Qian N, Zhu Y. Physiological Computation of Binocular Disparity. *Vision Res* 1997;37:1811–1827. [PubMed: 9274767]
13. Abràmoff, MD.; Niemeijer, M. Automatic detection of the optic disc location in retinal images using optic disc location regression. Vol. 2006. New York, NY: IEEE EMBC; 2006. p. 4432-4435.

14. Kwon YH, Kim YI, Pereira ML, Montague PR, Zimmerman MB, Alward WL. Rate of optic disc cup progression in treated primary open-angle glaucoma. *J Glaucoma* 2003;12:409–416. [PubMed: 14520149]
15. ter Haar Romeny, BM. *Front End Vision and Multi-scale Image Analysis*. Dordrecht: Kluwer Academic Publishers; 2003.
16. van Ginneken B, Haar Romeny BM. Automatic segmentation of lung fields in chest radiographs. *Med Phys* 2000;27:2445–2455. [PubMed: 11099215]
17. Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts? *IEEE Trans Pattern Anal Mach Intell* 2004;26:147–159. [PubMed: 15376891]
18. Marr, D. *Vision*. New York: Freeman and Company; 2005.
19. Birchfield, S.; Tomasi, C. Depth discontinuities by Pixel-to-pixel stereo. Bombay, India. *Proceedings 1998 IEEE International Conference on Computer Vision*; 1998.
20. Sonka, M.; Fitzpatrick, JM. *Handbook of Medical Imaging - Volume 2, Medical Image Processing and Analysis*. Wellingham, WA: The International Society for Optical Engineering Press; 2000.
21. Segev R, Goodhouse J, Puchalla J, Berry MJ. Recording spikes from a large fraction of the ganglion cells in a retinal patch. *Nat Neurosci* 2004;7:1154–1161. [PubMed: 15452581]
22. Lee TW, Wachtler T, Sejnowski TJ. Color opponency is an efficient representation of spectral properties in natural scenes. *Vision Res* 2002;42:2095–2103. [PubMed: 12169429]
23. Olshausen BA, Field DJ. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* 1996;381:607–609. [PubMed: 8637596]
24. Abràmoff MD, Kwon YH, Ts'o DY, Li H, Barriga ES, Kardon R. A spatial truncation approach to the analysis of optical imaging of the retina in humans and cats. *Proc IEEE International Symposium on Biomedical Imaging* 2004 2004;2:1115–1118.
25. Duda, RA.; Hart, PE.; Stork, DG. *Pattern classification*. Vol. 2. New York: Wiley-Interscience; 2001.
26. Staal JS, Abramoff MD, Niemeijer M, Viergever MA, van Ginneken B. Ridge based vessel segmentation in color images of the retina. *IEEE Trans Med Imaging* 2004;23:501–509. [PubMed: 15084075]
27. Arya S, Mount D, Netanyahu N, Silverman R, Wu A. An Optimal Algorithm for Apprximate Nearest Neighbor Searching in Fixed Dimensions. *Journal of the ACM* 1998;45:891–923.
28. Parkin B, Shuttleworth G, Costen M, Davison C. A comparison of stereoscopic and monoscopic evaluation of optic disc topography using a digital optic disc stereo camera. *Br J Ophthalmol* 2001;85:1347–1351. [PubMed: 11673304]
29. Niemeijer M, Russell SR, van Ginneken B, Abràmoff MD. Automated Detection and Differentiation of Drusen, Exudates, and Cotton-wool Spots in Digital Color Fundus Photographs for Early Diagnosis of Diabetic Retinopathy. *Invest Ophthalmol Vis Sci*. 2007[in press]
30. Corona E, Mitra S, Wilson M, Krile T, Kwon YH, Soliz P. Digital stereo image analyzer for generating automated 3-D measures of optic disc deformation in glaucoma. *IEEE Trans Med Imaging* 2002;21:1244–1253. [PubMed: 12585706]
31. Scharstein D, Szeliski R. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *ICJV* 2002;47:7–42.
32. Calkins DJ, Tsukamoto Y, Sterling P. Microcircuitry and mosaic of a blue-yellow ganglion cell in the primate retina. *J Neurosci* 1998;18:3373–3385. [PubMed: 9547245]

Appendix

Appendix A

Color opponency Gaussian filterbank

By careful selection of the steerable Gaussian kernels in a Gaussian filter bank, second order invariant features can be calculated, so that all other features up to second order are combinations of the features in this invariant set¹⁵. color opponency Gaussian filterbank³². L_{db} , L_{rg} , L_{by} , for dark-bright, red-green and blue-yellow opponency images, are computed from each color image as follows:

$$\begin{aligned}L_{db} &= \frac{L_r + L_g + L_b}{3} \\L_{rg} &= L_r - L_g \\L_{by} &= (L_r + L_g) - 2L_b\end{aligned}$$

with L_r, L_g, L_b the red, green and blue channels of the left image of the stereo pair respectively.

L_{db}, L_{rg}, L_{by} are then each convolved with Gaussian filterbank kernels to obtain the following color-opponency Gaussian derivative features

$$L_0(\sigma), L_1^{0^\circ}(\sigma), L_1^{90^\circ}(\sigma), L_2^{0^\circ}(\sigma), L_2^{60^\circ}(\sigma), L_2^{120^\circ}(\sigma)$$

with $L_{n,op}^\alpha(\sigma)$ the feature obtained by computing $L_n^\alpha(\sigma) = L_l \otimes G_n^\alpha(\sigma)$ for order n , orientation α (0° and 90° orientation to the local gradient for first order and $0^\circ, 60^\circ$, and 120° to the local gradient for second order derivatives), scale

$\sigma \in (2, 4, 8, 16, 32, 64)$ and opponency image $L_{op} \in (L_{db}, L_{rg}, L_{by})$.

$$\begin{aligned}L_0 &= L_l \\L_1^\alpha &= \cos\alpha L_x + \sin\alpha L_y\end{aligned}$$

$\otimes$ denotes convolution. L_n^α is computed as follows: $L_2^\alpha = \cos^2\alpha L_{xx} + \cos\alpha\sin\alpha L_{xy} + \sin^2\alpha L_{yy}$

where the lettered (xx, yy, xy) subscripts to L denote differentiation of L_{op} to x and/or y (after Haar Romeny, 2003 811/id). This results in 108 color opponency Gaussian filterbank features.

Appendix B

Stereo features

Gabor features $G_{\sigma,op}^\alpha$ were calculated by convolution of the three opponency images $op \in (L_{db}, L_{rg}, L_{by})$ with Gabor wavelet kernel at scales $\sigma = 1, 2, 4, 8, 16, 32$ pixels, at orientation $\alpha = 0^\circ, 45^\circ, 90^\circ$, and 135° :

$$G_{\sigma,op}^\alpha(x, y) = e^{-\frac{x^2 + y^2}{2\sigma^2}} \cos(\omega x + \phi)$$

with phase $\phi = 0$ and with spatial frequency $\omega = \frac{\pi}{\sigma}$ constant, for a total of 82 features. See Figure 3 for an example of these kernels.

Stereo disparity features were computed by first convolving the left and right color opponency images with left and right Gabor kernels k_l and k_r

$$\begin{aligned}k_l(x) &= e^{-\frac{x^2}{2\sigma^2}} \cos(\omega x + \phi_l) \\k_r(x) &= e^{-\frac{x^2}{2\sigma^2}} \cos(\omega x + \phi_r)\end{aligned}$$

where ϕ_l , ϕ_r are the left and right phases of a Gabor wavelet, and ω its preferred spatial frequency, which was kept constant at $\omega = \frac{\pi}{\sigma}$. The linear combination of convolutions of the left and right images with these respective kernels mimics the output of a simple cell:

$$S_{\omega}(\sigma, \phi_l, \phi_r) = L_{l,l} \otimes k_l + L_{l,r} \otimes k_r$$

with $L_{l,l}$, $L_{l,r}$ the left and right intensity images respectively. The sum of the squares of two $S_{1,\omega}(\sigma, \phi_l, \phi_r)$, $S_{2,\omega}(\sigma, \phi_l + \frac{\pi}{2}, \phi_r + \frac{\pi}{2})$ formed quadrature pairs:

$$C_{\omega}(\sigma, \Delta\phi) = (S_{1,\omega}(\sigma, \phi_l, \phi_r))^2 + \left(S_{2,\omega}(\sigma, \phi_l + \frac{\pi}{2}, \phi_r + \frac{\pi}{2})\right)^2$$

With $\Delta\phi$ the phase difference $\Delta\phi = \phi_r - \phi_l$ between the right and left eye receptive fields.

With $\Delta\phi \in \left\{ \pi\frac{6}{8}\pi\frac{4}{8}\pi, \frac{2}{8}\pi, 0, -\frac{2}{8}\pi, -\frac{4}{8}\pi, -\frac{6}{8}\pi \right\}$, $C_{\omega}(\sigma, \Delta\phi)$ was then tuned to stereo disparities of $\{-4, -3, -2, -1, 0, 1, 2, 3\}$ pixels respectively, calculated at scales $\sigma=1, 2, 4, 8, 16, 32$ pixels, and for each of the $(L_{op}, R_{op}) \in \{(L_{db}, R_{db}), (L_{rg}, R_{rg}), (L_{by}, R_{by})\}$ color opponency pairs, resulting in 48 features for each of the color opponency spaces¹².

Appendix C

Balanced accuracy

Performance was calculated in the form of *balanced* accuracy (a combination of sensitivity and specificity in a single number) by calculating the number of correctly classified pixels in equally sized samples from the rim, the cup and the background

$$\pi = \frac{N_{r,T} + N_{c,T} + N_{b,T}}{N_r + N_c + N_b}$$

with $N_{r,T}$ the number of pixels correctly classified in a sample of rim pixels, $N_{c,T}$ the number of pixels classified correctly in a sample of cup pixels, $N_{b,T}$ the number pixels classified correctly as background in a sample of cup pixels, and $N_r = N_c = N_b$ the total number of pixels in each sample. The reason to do this is that otherwise because of the predominance of background pixels in the image, an algorithm that is ‘good’ in classifying background but ‘bad’ in classifying rim would have an unfair advantage.

Fig. 1.

Stereo color pair of the optic nerve head of the left eye after alignment and cropping to 512×512pixels. The left and right stereo images can be fused by holding the page at about 30cm (12") distance from the eye and gazing into infinity.

Fig. 2.

A, B, C: Gradings of an optic disc stereo pair by three faculty glaucoma specialists. Rim in grayish, cup in whitish, shown on the left image of the stereo pair from Figure 1. D: reference standard created from A, B, C with white, cup, gray, rim, and black, background

Figure 3.

Example of the variety of color opponency steerable Gaussian filterbank kernels. Kernels oriented as if the local gradient was horizontal. First row, from left to right dark-bright opponency kernels for zeroth order, first order 0° to the local gradient, first order 90° to the local gradient, second order 0° to the local gradient, second order 60° to the local gradient and second order 150° to the local gradient, at scale σ 32 pixels. Second row, same for scale σ 64 pixels, and third row, 128 pixels. Next three rows same for blue-yellow opponency kernels, and last three rows, same for red-green kernels. Smaller scales not shown because they are difficult to see. These kernel images also represent the response of each of the features to an impulse function.

Fig. 4.

Stereo pair and its most discriminant pixel features. A,B: the left and right stereo image of the stereo pair. Rows below that show images of the most dominant twelve features in rank order (see Table 1 for a description of these features).

Fig. 5.

Classification of stereo pairs by glaucoma faculty, three glaucoma fellows, and pixel feature classification. Rows 1,2,3,4: small, medium and large disc excavation and excavation with inferior notching. From left to right: left and right images of the stereo pair; reference standard by three glaucoma faculty; grading by fellow A, by fellow B and by fellow C; and grading by pixel feature classification.

Fig. 6.

Scatterplots of linear cup-to-disc ratio estimates from all 58 stereo pairs by three glaucoma fellows and the pixel feature classification algorithm (vertical axes), against the reference standard grading by three glaucoma faculty (horizontal axes). Clockwise, fellow A, B, the algorithm and fellow C respectively. 'r' is the correlation coefficient for each of these estimates with the reference standard.

Table 1

Twelve features selected as the most dominant, ranked by their contribution to cumulative accuracy on the training set. Second column is the description of the feature, third column is the scale of the feature in pixels, if applicable, and fourth column shows cumulative accuracy.

Rank	Description	Scale (pixels)	Cumulative accuracy on training set
1	A priori probability of pixel being cup	1	0.810
2	Red-green opponency (opp.) second order Gaussian derivative transversal to local gradient	32	0.825
3	A priori probability of pixel being rim	1	0.829
4	Red-green opp. Gaussian blurred (zeroth order)	8	0.832
5	Complex cell output tuned to a disparity of -12 pixels	16	0.845
6	Blue-yellow opp. 2 nd order Gaussian derivative	32	0.854
7	Red-green opp. 1 st order Gaussian derivative perpendicular to local gradient	16	0.869
8	Bright dark opp. Gaussian blurred (zeroth order)	16	0.872
9	Red-green opp. 1 st order Gaussian derivative radial to local gradient	4	0.874
10	Blue-yellow Gaussian blurred (first-order)	4	0.874
11	Red-green opp. 1 st order Gaussian derivative parallel to local gradient	4	0.875
12	Red-green opp. Gaussian blurred (zeroth order)	32	0.876