



Published in final edited form as:

Arch Ophthalmol. 2005 February ; 123(2): 200–206. doi:10.1001/archopht.123.2.200.

Automated Detection of Macular Drusen Using Geometric Background Leveling and Threshold Selection

R. Theodore Smith, MD, PhD, Jackie K. Chan, MS, Takayuki Nagasaki, PhD, Umer F. Ahmad, MD, Irene Barbazetto, MD, Janet Sparrow, MD, Marta Figueroa, MD, and Joanna Merriam, MD, PhD

Edward S. Harkness Eye Institute, New York Presbyterian Medical Center, New York, NY (Drs Smith, Nagasaki, Ahmad, Barbazetto, Sparrow, and Merriam and Mr Chan); and Department of Ophthalmology, Ramon Y. Cajal Hospital, Madrid, Spain (Dr Figueroa)

Abstract

Background—Age-related macular degeneration (ARMD) is the most prevalent cause of visual loss in patients older than 60 years in the United States. Observation of drusen is the hallmark finding in the clinical evaluation of ARMD.

Objectives—To segment and quantify drusen found in patients with ARMD using image analysis and to compare the efficacy of image analysis segmentation with that of stereoscopic manual grading of drusen.

Design—Retrospective study.

Setting—University referral center.

Patients—Photographs were randomly selected from an available database of patients with known ARMD in the ongoing Columbia University Macular Genetics Study. All patients were white and older than 60 years.

Interventions—Twenty images from 17 patients were selected as representative of common manifestations of drusen. Image preprocessing included automated color balancing and, where necessary, manual segmentation of confounding lesions such as geographic atrophy (3 images). The operator then chose among 3 automated processing options suggested by predominant drusen type. Automated processing consisted of elimination of background variability by a mathematical model and subsequent histogram-based threshold selection. A retinal specialist using a graphic tablet while viewing stereo pairs constructed digital drusen drawings for each image.

Main Outcome Measures—The sensitivity and specificity of drusen segmentation using the automated method with respect to manual stereoscopic drusen drawings were calculated on a rigorous pixel-by-pixel basis.

Results—The median sensitivity and specificity of automated segmentation were 70% and 81%, respectively. After preprocessing and option choice, reproducibility of automated drusen segmentation was necessarily 100%.

Conclusions—Automated drusen segmentation can be reliably performed on digital fundus photographs and result in successful quantification of drusen in a more precise manner than is traditionally possible with manual stereoscopic grading of drusen. With only minor preprocessing

Correspondence: R. Theodore Smith, MD, PhD, Edward S. Harkness Eye Institute, New York Presbyterian Medical Center, Suite 314, 635 W 165th St, New York, NY 10032 (rts1@columbia.edu).

Financial Disclosure: None.

requirements, this automated detection technique may dramatically improve our ability to monitor drusen in ARMD.

Extensive drusen area as seen on the fundus photograph is a strong risk factor for the progression of age-related macular degeneration (ARMD).¹⁻⁸ However, there is difficulty in obtaining interobserver agreement in drusen identification. For example, interobserver agreement on the presence of soft drusen only was 89% and on the total number of drusen was 76% in one study.⁹ Studies have been based on the current standard for drusen grading of digital fundus photographs in ARMD: manual grading of stereo pairs at the light box.¹⁰⁻¹¹ Examiners are asked to mentally aggregate the amount of drusen in the field occupying the macular region. Then, lesion quantification using the international classification assigns broad category intervals of 0% to 10%, 10% to 25%, and so forth.¹¹ Clearly, there is a pressing need for the development of techniques that allow for more precise grading and thereby result in significant improvement in the quality of data being gathered in clinical trials and epidemiological studies.

Known for precision, computers have the computational power to solve this problem. However, digital techniques have not as of yet gained widespread acceptance, despite progress,¹²⁻¹⁸ for several reasons. First, the inherent nature of the reflectance of the normal macula is nonuniform. There is less reflectance centrally and increasing reflectance moving out toward the arcades. Local threshold approaches to drusen segmentation have been attempted with only partial success because the background variability limits the extent to which purely histogram-based methods can succeed. This has increased the need for operator intervention and has been the main obstacle to automating drusen segmentation. We had previously developed an interactive method to correct the macular background globally by taking into account the geometry of macular reflectance.^{19,20} This method required subjective user choice of background input and final threshold. Here we combine automated histogram techniques and the analytic model for macular background to give a completely automatic measurement of macular area occupied by drusen.

The second major obstacle to drusen identification has been that of object recognition. A computer must ultimately learn to differentiate drusen from areas of retinal pigment epithelial hypopigmentation, exudates, and scars. Goldbaum et al²¹ have suggested subtleties of coloration and shape as modes of automated recognition. However, this subject has not been developed further. At present, in our hands, the complete attention of the operator during the preprocessing phase is required to exclude such confounders in approximately 20% of images.^{20,22}

The third major obstacle to drusen identification is that of boundary definition: soft, indistinct drusen have no precise boundary, and therefore the solution to their segmentation, by definition, cannot be precise. The central color fades into the background peripherally, and on stereo viewing there is no well-defined edge. Practical segmentation of drusen then requires that areas of drusen determined by a digital method agree, in aggregate, with the judgments of a qualified grader. This approach was adopted by Shin et al¹² for validation of their method. However, expert manual drawings themselves are necessarily variable. For some of the images reported here, expert manual drawings varied as much as digital segmentation methods. Indeed, specificity and sensitivity calculations for expert manual drawings of 2 retinal experts demonstrated significant interobserver differences. Therefore, achieving comparable accuracy in automated drusen segmentation relative to an acceptable stereo viewing standard represents an advance.

We hope to demonstrate the ability of our automated method to more accurately segment drusen using an algorithm based on the geometry of macular reflectance. We believe the methodology

described may gain widespread acceptance as a useful tool in studying problems of clinical relevance with respect to ARMD. This method is speedy, reproducible, and cost-effective in drusen segmentation, and we believe it will be applicable for use in clinical trials.

METHODS

SUBJECTS

A group of 20 stereo pair slides from 17 patients was chosen randomly from the Columbia University Macular Genetics Study, a study approved by the institutional review board of New York Presbyterian Hospital, New York, NY. All patients were white and older than 60 years. One slide from each pair was digitized (CoolScan LS-2000; Nikon Corp, Tokyo, Japan) at 2700 pixels-per-inch tagged image file format (TIFF) files (8 bits per channel, RGB color mode, with a gray scale range of 0 to 255 per color channel).

IMAGE PREPROCESSING

All image preprocessing and analysis were performed with commercially available software (Photoshop 7.0; Adobe Systems Inc, San Jose, Calif; and Matlab; Mathworks, Natick, Mass) on a desktop personal computer. The region studied was the central 3000- μ m-diameter circle (the combined central and middle subfields defined by the Wisconsin grading template: central subfield, the circle of diameter 1 mm; middle subfield, the annulus of outer diameter 3 mm). All area measurements are stated as percentages of the entire 3000- μ m-diameter circle.

All images were resized in Photoshop so that the distance from the center of the macula to the temporal disc edge was 500 pixels. This macula-disc distance (3000 μ m) is established as the constant of reference in clinical macular grading systems. Although this distance varies anatomically, it does not affect area measurements calculated as percentages.

By methods described previously, we next corrected the large-scale variation in brightness found in most fundus photographs²³ (photographic variability not intrinsic to retinal reflectance). This shading correction was carried out independently on each color channel, and results were combined as the RGB channels of a new standardized, color-balanced image. Each image was also paired with a contrast-enhanced version (Au-to levels command in Photoshop) for ease of lesion visualization. All further image analysis was carried out on this preprocessed image, which we call the standardized image.

IMAGE ANALYSIS

Stereo Viewing and Manual Tracing Method for Drusen—We obtained manual digital segmentations of drusen as follows. On a graphic tablet (Intuos; Wacom Corp, Vancouver, Wash), the user drew the boundaries of all lesions identified in the contrast-enhanced image. As the user drew, the 1-pixel pencil tool in Photoshop outlined the lesions in a transparent digital layer. Reference was also made as needed to the stereo fundus photographs to determine the exact boundary. The lesion outlines were then filled and their areas calculated. The same technique was used to segment possible confounding lesions such as geographic atrophy in 3 images as a preprocessing step before automated segmentation. We also refer to the stereo viewing method as the ground truth method in image analysis terminology.

Automated Method of Drusen Measurement

Luteal Compensation: The first step is a luteal pigment correction applied to the green channel of the standardized image. The ratio of the median values of the histograms of the green channel in the middle and central subfields was calculated. This ratio was applied to a Gaussian distribution centered on the fovea and having a half-maximum at 600-micron diameter. The green channel was multiplied by this Gaussian distribution to produce the luteal compensate

dimage. This compensation is a variable version of a fixed luteal compensation described previously.²³ All further processing and segmentation were carried out on this image.

Two-Zone Math Model: Zone 1 is the central subfield, and zone 2 is the annulus of inner and outer diameters of 1000 and 3000 microns, respectively. The pixel gray levels were considered to be functions of their pixel coordinates (x, y) in the x - y plane. The general quadratic $q(x, y) = ax^2 + bxy + cy^2 + dx + ey + \text{constant}$ in 2 variables was fit by custom software employing least-squares methods to any chosen background input of green-channel gray levels to optimize the 6 coefficients ($a, b, c, d, e,$ and constant).²⁴ In this case, the model consists of a set of 2 quadratics, 1 for each zone, with cubic spline interpolations at the boundary.²⁴

Initial Background Selection by Otsu Method: We employed the automatic histogram-based thresholding technique known as the Otsu method²⁵ in each zone to provide initial input to the background model. Briefly, let the pixels in the green channel be represented in L gray levels $[1, 2, \dots, L]$. Suppose we dichotomized the pixels into 2 classes, C_0 and C_1 , by a threshold at level k . C_0 denotes pixels with levels $[1, \dots, k]$ and C_1 denotes pixels with levels $[k+1, \dots, L]$. Ideally, C_0 and C_1 would represent background and drusen. A discriminant criterion that measures class separability was used to evaluate the goodness of the threshold (at level k). The Otsu method uses the criterion of between-class variance and selects the threshold k that maximizes this variance.²⁵ The Otsu method can be generalized to the case of 2 thresholds k and m , where there are 3 classes, $C_0, C_1,$ and C_2 , defined by pixels with levels $[1, \dots, k], [k+1, \dots, m],$ and $[m+1, \dots, L]$, respectively. In a given image, these classes might represent background, objects of interest, and other objects (eg, retinal vessels), in some permutation. The criterion for class separability is the total between-class variance

$\sigma_B^2 = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 + \omega_2(\mu_2 - \mu_T)^2$ where ω_i and μ_i are the zero-order and the first-order normalized cumulative moments of the histogram for class C_i as defined above for $i=1, 2, 3,$ and μ_T is the image mean. The solution is found by the finite search on k for $k=1, \dots, L-1$ and m for $m=k+1$ to L for the maximum of σ_B . The Otsu method may also be performed sequentially to subdivide a given class. That is, if a given class, C , is already defined (by Otsu or otherwise), then C may be treated as the initial histogram (setting other histogram values to zero), and one can apply an Otsu method to subdivide C into 2 (or 3) classes.

Operator Options: We found by trial and error on a large variety of images that on most images (15/20 in the present series) the 2-threshold Otsu method performed well in zone 2 to provide an initial segmentation by thresholds k and m into 3 desired classes: C_0 (dark nonbackground sources, eg, vessels and pigment), C_1 (background), and C_2 (drusen). In zone 1 where vasculature was not present, the single-threshold Otsu method was used to divide the region into 2 classes labeled C_1 (background) and C_2 (drusen). In particular, for each region, we then had an initial choice of background, C_1 , for input to the mathematical background model. These were the default settings, or option 0, used in 15 images. If multiple large, soft, ill-defined drusen were present, we found that the upper (drusen) thresholds tended to capture the brighter central portions of these drusen and miss the fading edges. Option 1, which allowed the operator to reduce all initial thresholds by 4, was used in 3 such images. Finally, when drusen load was small (5% or less estimated range), we found that their statistical power was insufficient to be recognized by the initial Otsu subdivision, which instead would pick out a larger subset C_2 that included brighter background and the drusen themselves. It was on further subdivision of this C_2 by the single-threshold Otsu method, as described earlier, that small groups of drusen were recognized. The higher pixel values became the new C_2 , and the remainder was included in C_1 . This option (option 2) was used on 2 images. These were the only operator decisions needed to determine C_1 (the background) for input to the model. The rest of the algorithm up to final segmentation was completely automatic. These initial subdivisions of the image are illustrated in Figure 1 A and B.

Sequential Automated Background Leveling and Thresholding: Let Z be the luteal corrected image data and let Q be the model fit to the background data C_1 determined by the Otsu method specified previously. The first leveled image Z_1 is defined as $Z_1 = Z - Q + 125$. The constant offset 125 maintains an image with an approximate mean of 125. The process can now be iterated, with Z_1 the input to the Otsu background segmentation, resulting in a new background choice C_1 from Z_1 . If Q_1 is the model fit to the new background data, the next leveled image is $Z_2 = Z_1 - Q_1 + 125$, and so forth. The process terminates after a predetermined number of steps or when the net range of the model Q reaches a set target (that is, the range is sufficiently small, indicating that the new background is nearly flat). The final drusen segmentation is then obtained by applying the specified Otsu method to the final leveled image and removing any confounding lesions identified in manual preprocessing. In practice, we found our final results changed little after 2 iterations of the leveling process. This sequence is illustrated in Figure 1C through F.

MEASUREMENTS

We compared the automated digital method with the stereo viewing method. Two retinal expert graders (R.T.S., I.B.) each used the stereo viewing method on 10 images, and the results were compared. Total drusen areas were measured. In cases in which the experts disagreed by more than 5%, the 2 graders collaborated to redraw to consensus. Expert drawings were also made of the remaining 10 images. On a total of 20 images, the drusen areas were also measured by the automated method (R.T.S., J.K.C.) and compared with a stereo viewing drawing of an expert grader. As described earlier, the only choices made in the automated method were the options chosen to guide the Otsu method in background selection. The 95% limits of agreement were calculated. False-positive pixels (drusen areas found by the automated method but not selected by the retinal expert) and false-negative pixels (drusen areas selected by the retinal expert but not selected by the automated method) were also identified. Specificity and sensitivity of the automated method were calculated accordingly.

RESULTS

In 10 images, the difference in drusen area measurements between 2 expert graders ranged from -0.2% to 7.0% (stereo viewing measurements by grader I.B. were 3.4% higher on average). The 95% limits of agreement were from -2.0% to 8.8% .²⁶

The segmentations created by the automated method were then compared on a pixel-by-pixel basis with the respective manual drawings with sensitivity from 0.42 to 0.86 (median, 0.70) and specificity from 0.53 to 0.98 (median, 0.81).

Comparison of these methods for a representative patient is illustrated in Figure 2. Two expert stereo viewing drawings of the same image are also compared in Figure 2. They demonstrate variability similar to the differences between the automated method and a stereo drawing.

Sensitivity and specificity of drusen quantification by the automated method compared with the stereo viewing method is detailed in the Table. The lowest sensitivity of 0.42 occurred in measurements of small quantities of drusen (patient 7), for which small false-negative errors of 4.2% caused large decrements in the sensitivity.

Comparing the automated with the ground truth method showed the difference in drusen area measurements of the 20 images ranged from -6.7% to 12.7% (ground truth measurements were 3.4% higher on average). The 95% limits of agreement between the 2 methods were -7.1% to 13.5% Figure 3.

The 3 images requiring supervision in the form of manual segmentation of confounding lesions (geographic atrophy, retinal pigment epithelial hypopigmentation, and photographic dust spots) are shown in Figure 4. In these complex images, combining the manual and automated techniques produced a more accurate segmentation. In contrast to the multiple and poorly defined drusen, the smoother contours of these few lesions were easily traced. In each case, the specificity improved.

COMMENT

This article combines 2 approaches to the digital analysis of macular drusen: a stereo viewing method with manual tracing on a graphic tablet and an automated method with automatic threshold selection. Shin et al¹² introduced stereo viewing with tracing for drusen only as an adjunct for validation of their digital measurements, without determining reproducibility of the stereo viewing method itself. We have addressed reproducibility, added the ergonomic superiority of the graphic tablet for smooth tracing, and improved lesion visualization in the contrast-enhanced image to lessen user fatigue and improve accuracy. This method as a validation tool is also improved herein by explicitly depicting false positives and false negatives for specificity and sensitivity calculations. The automated drusen method offers efficiency in the tedious task of drusen segmentation (requiring about 10 seconds per slide) and provides results comparable with those of stereo viewing.

A limitation of any drusen measurement method (human or automated) is optimum boundary definition. There is no absolute correct choice for indistinct, soft drusen. If photograph quality is suboptimal, the difficulty is compounded. Highly reflectant lesions, such as retinal pigment epithelial hypopigmentation, geographic atrophy, exudates, and scars as well as photographic dust spots, would more likely be mistaken for drusen by the automated method than by an expert grader. In this study, such lesions were present in 3 cases and were manually segmented in the preprocessing step. We felt it was important to include these cases to demonstrate an important limitation of the completely automated method, as well as to illustrate that a straightforward solution was available.

The sensitivity (median, 0.70) of the automated method was less than the specificity (median, 0.81) with respect to the stereo viewing method. A partial explanation lies in our finding that lowering the threshold for drusen identification past critical levels to try to improve drusen recognition usually resulted in an unacceptable increase in false positives. However, we also found that stereo drawing measurements of the same macula by 2 retinal experts could vary comparably, with similarly limited sensitivity and specificity.

Further testing should include application of these drusen measurements to serial images over a number of years for sensitivity, specificity, and reliability. Another potential application, with appropriate modification of the statistical methods, would be screening of normal to near-normal images for the presence or absence of age-related maculopathy.

In summary, we have demonstrated a digital drusen measurement method that reproduces expert stereo drawings with an accuracy rivaling that of the expert stereo gradings themselves. When combined with easily implemented expert drawing for other lesions, such as geographic atrophy, this method also handles important categories of more complex images. This efficiency and accuracy may become useful in clinical studies.

References

1. Smiddy WE, Fine SL. Prognosis of patients with bilateral macular drusen. *Ophthalmology* 1984;91:271–277. [PubMed: 6201789]

2. Bressler SB, Maguire MG, Bressler NM, Fine SL. Relationship of drusen and abnormalities of the retinal pigment epithelium to the prognosis of neovascular macular degeneration: The Macular Photocoagulation Study Group. *Arch Ophthalmol* 1990;108:1442–1447. [PubMed: 1699513]
3. Bressler NM, Bressler SB, Seddon JM, Gragoudas ES, Jacobson LP. Drusen characteristics in patients with exudative versus non-exudative age-related macular degeneration. *Retina* 1988;8:109–114. [PubMed: 3420311]
4. Holz FG, Wolfensberger TJ, Piguet B, et al. Bilateral macular drusen in age-related macular degeneration: prognosis and risk factors. *Ophthalmology* 1994;101:1522–1528. [PubMed: 8090455]
5. Zarbin MA. Current concepts in the pathogenesis of age-related macular degeneration. *Arch Ophthalmol* 2004;122:598–614. [PubMed: 15078679]
6. Little HL, Showman JM, Brown BW. A pilot randomized controlled study on the effect of laser photocoagulation of confluent soft macular drusen. *Ophthalmology* 1997;104:623–631. [PubMed: 9111254]
7. Frennesson IC, Nilsson SE. Effects of argon (green) laser treatment of soft drusen in early age-related maculopathy: a 6 month prospective study. *Br J Ophthalmol* 1995;79:905–909. [PubMed: 7488578]
8. Bressler NM, Munoz B, Maguire MG, et al. Five-year incidence and disappearance of drusen and retinal pigment epithelial abnormalities: Waterman study. *Arch Ophthalmol* 1995;113:301–308. [PubMed: 7534060]
9. Bressler SB, Bressler NM, Seddon JM, Gragoudas ES, Jacobson LP. Interobserver and intraobserver reliability in the clinical classification of drusen. *Retina* 1988;8:102–108. [PubMed: 3420310]
10. Bird AC, Bressler NM, Bressler SB, et al. An international classification and grading system for age-related maculopathy and age-related macular degeneration: The International ARM Epidemiological Study Group. *Surv Ophthalmol* 1995;39:367–374. [PubMed: 7604360]
11. Klein R, Davis MD, Magli YL, Segal P, Klein BE, Hubbard L. The Wisconsin age-related maculopathy grading system. *Ophthalmology* 1991;98:1128–1134. [PubMed: 1843453]
12. Shin DS, Javornik NB, Berger JW. Computer-assisted, interactive fundus image processing for macular drusen quantitation. *Ophthalmology* 1999;106:1119–1125. [PubMed: 10366080]
13. Sebag M, Peli E, Lahav M. Image analysis of changes in drusen area. *Acta Ophthalmol (Copenh)* 1991;69:603–610. [PubMed: 1776413]
14. Morgan WH, Cooper RL, Constable IJ, Eikelboom RH. Automated extraction and quantification of macular drusen from fundal photographs. *Aust N Z J Ophthalmol* 1994;22:7–12. [PubMed: 8037918]
15. Kirkpatrick JN, Spencer T, Manivannan A, Sharp PF, Forrester JV. Quantitative image analysis of macular drusen from fundus photographs and scanning laser ophthalmoscope images. *Eye* 1995;9:48–55. [PubMed: 7713250]
16. Peli E, Lahav M. Drusen measurement from fundus photographs using computer image analysis. *Ophthalmology* 1986;93:1575–1580. [PubMed: 3808617]
17. Rapantzikos K, Zervakis M, Balas K. Detection and segmentation of drusen deposits on human retina: potential in the diagnosis of age-related macular degeneration. *Med Image Anal* 2003;7:95–108. [PubMed: 12467724]
18. Ben Sbeh Z, Cohen LD, Mimoun G, Coscas G. A new approach of geodesic reconstruction for drusen segmentation in eye fundus images. *IEEE Trans Med Imaging* 2001;20:1321–1333. [PubMed: 11811832]
19. Chan JWK, Smith R, Nagasaki T, Sparrow JR, Barbazetto I. A method of drusen measurement based on reconstruction of fundus background reflectance. *Invest Ophthalmol Vis Sci* 2004;45:E-2415.
20. Smith RT, Chan JK, Nagasaki T, Sparrow JR, Barbazetto I. A method of drusen measurement based on reconstruction of fundus reflectance. *Br J Ophthalmol*. In press.
21. Goldbaum MH, Katz NP, Nelson MR, Haff LR. The discrimination of similarly colored objects in computer images of the ocular fundus. *Invest Ophthalmol Vis Sci* 1990;31:617–623. [PubMed: 2186008]
22. Sivagnanavel V, Smith RT, Chong NHV. Digital drusen quantification in high-risk patients with age related maculopathy. *Invest Ophthalmol Vis Sci* 2003;44:E-5002.
23. Smith RT, Nagasaki T, Sparrow JR, Barbazetto I, Klaver CCW, Chan JK. A method of drusen measurement based on the geometry of fundus reflectance. *Biomed Eng Online* 2003;2:10. [PubMed: 12740042]

24. Smith RT, Nagasaki T, Sparrow JR, Barbazetto I, Koniarek JP, Bickmann LJ. Patterns of reflectance in macular images: representation by a mathematical model. *J Biomed Opt* 2004;9:162–172. [PubMed: 14715069]
25. Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* 1979;9:62–66.
26. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–310. [PubMed: 2868172]

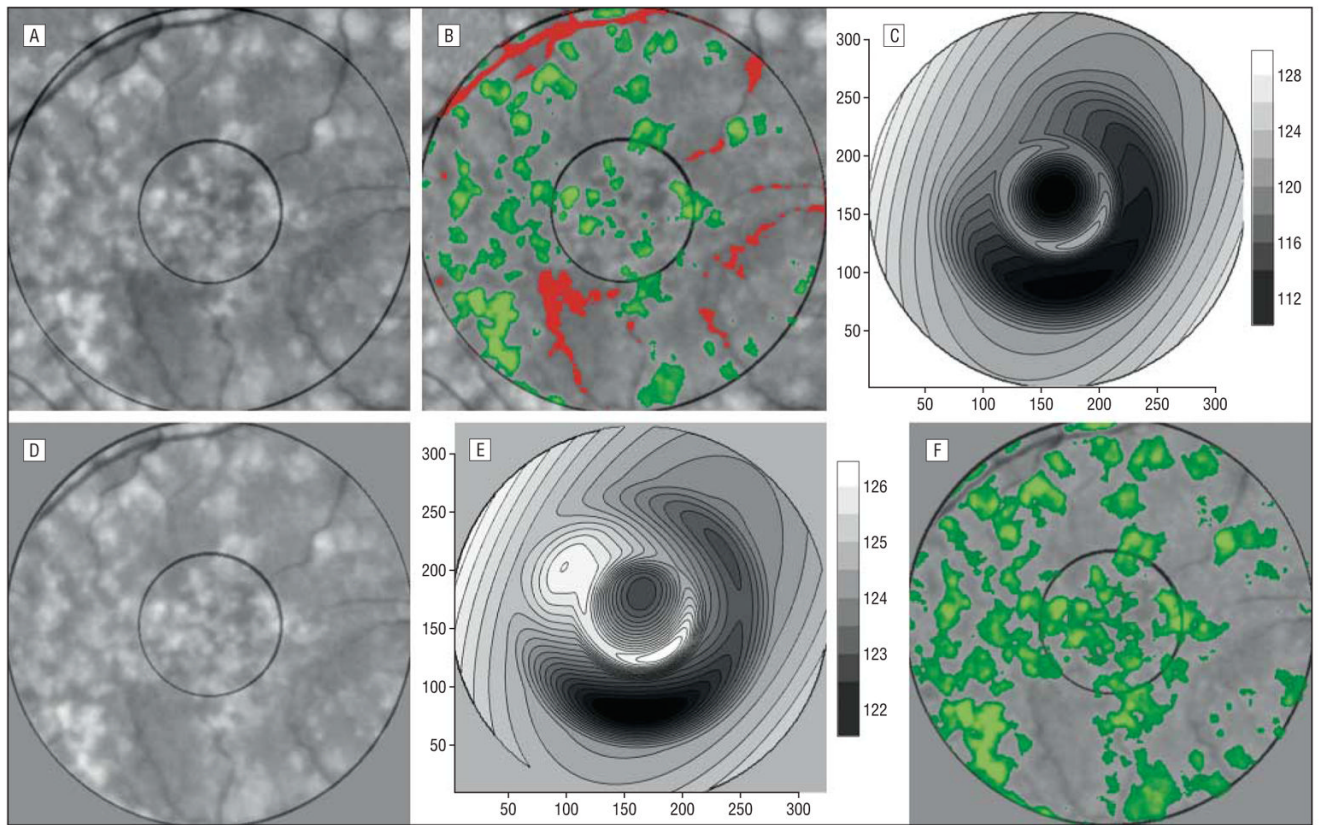


Figure 1.

Sequential automated background leveling and thresholding. A, A standardized image, green channel, gray scale, slightly contrast enhanced for visualization. B, The initial Otsu double thresholds in each region have provided estimates for vessels and some darker perivasculature inferiorly (pixel values below the lower threshold, red), background (pixel values between the 2 thresholds, gray), and drusen (pixel values above the higher threshold, green). C, The mathematical model fit to the background in B, displayed as a contour graph. D, The image in A has been leveled by subtracting the background variability of the math model in C (result slightly contrast enhanced). E, The new mathematical model fit to the background in D has a significantly smaller range than the initial model in C, indicating that the background in D is more nearly uniform than in A. F, The process is repeated (leveling by subtracting the model and Otsu segmentation). This is the final iteration and drusen segmentation, showing a significant improvement.

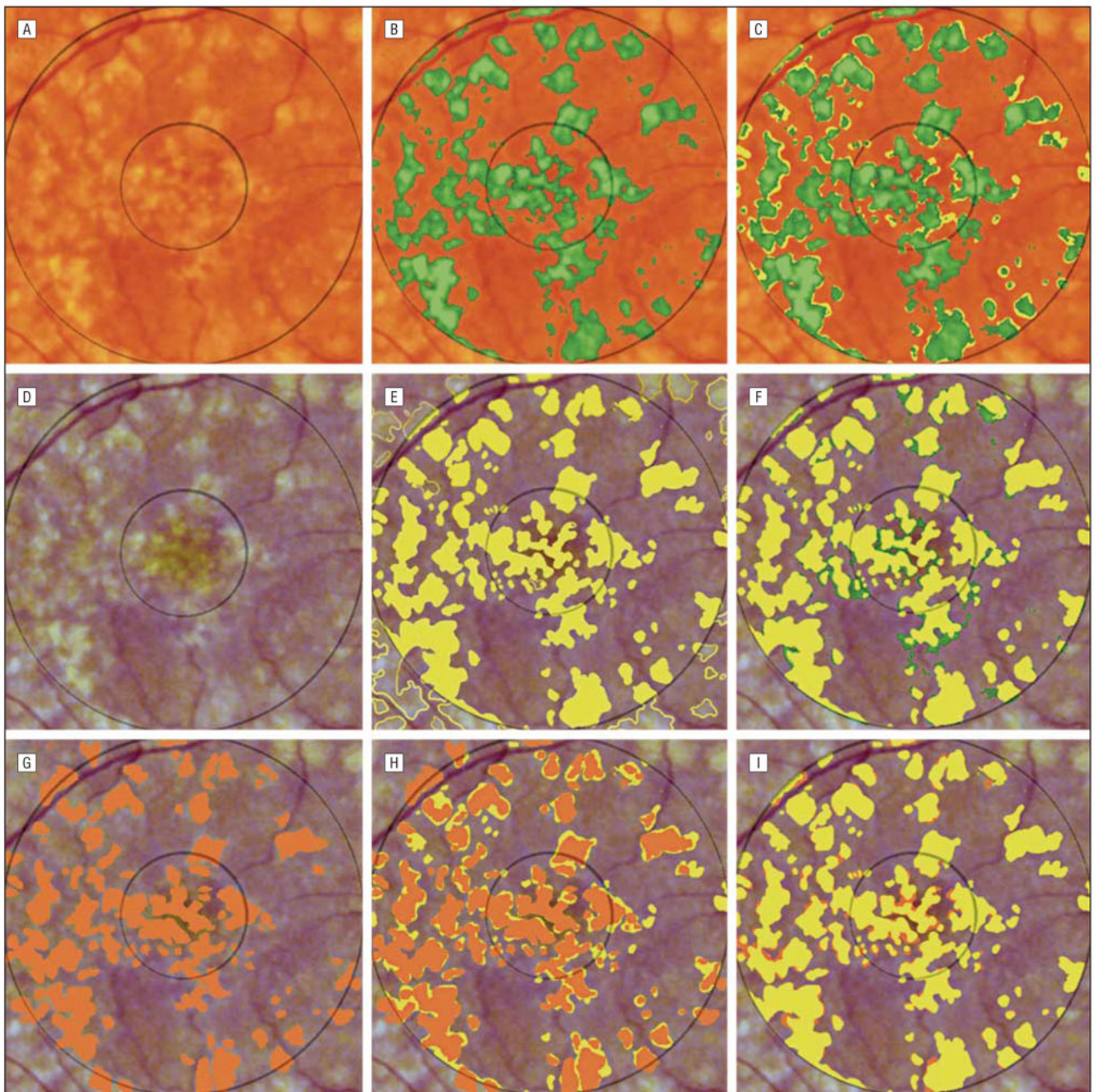


Figure 2.

Comparison of drusen segmented by the automated method and 2 expert drawings for patient 13. The drusen in the original image (A) were identified by the automated method (green) in B. The drusen in the contrast-enhanced version of the image (D) were traced manually in E by a retinal expert (I.B.), who also viewed the original stereo slide pair as needed. The drusen tracings in the central and middle subfields in E have been filled (yellow), while other partial tracings have been left unfilled for illustration. C shows the results of the automated method digitally overlaid on the retinal expert's drawing. The remaining yellow regions in C (not selected by the automated method) are the false negatives (7.0%). F shows the retinal expert's drawing digitally overlaid on the results of the automated method. The remaining green regions

(not selected by the retinal expert) are the false positives (6.5%). Both types of errors are rather randomly distributed. There is almost exact agreement in total drusen area (34.6% vs 34.1%). Pixel-by-pixel agreement, however, is not exact (sensitivity, 0.80; specificity, 0.81). The drusen were also traced manually in G by a second retinal expert (R.T.S.); the drawing (orange) was slightly more conservative than that of the first expert (total drusen area 26%). These results were overlaid on those of the first expert's drawing in H and vice versa in I, highlighting areas of disagreement. If the first expert drawing is taken as truth, then the second expert drawing has a sensitivity of 0.70 and a specificity of 0.91, roughly comparable to the accuracy of the automated method.

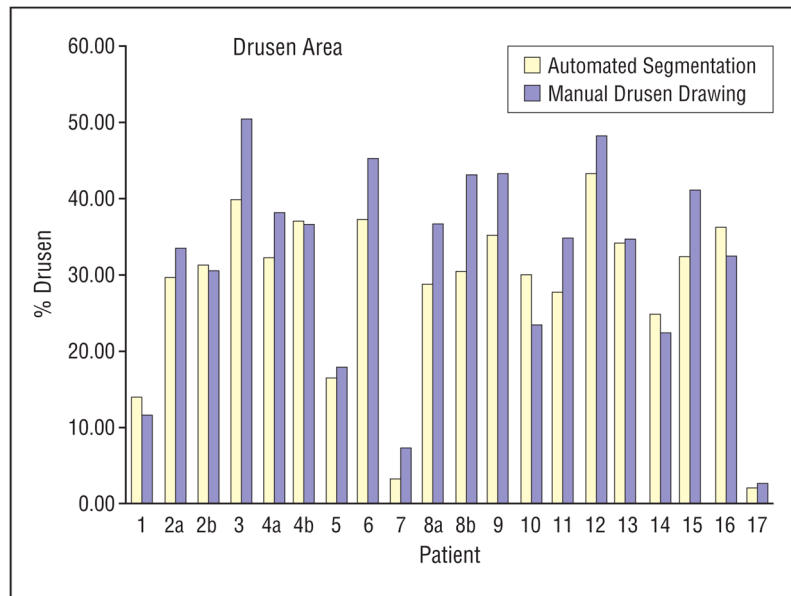


Figure 3. A comparison of the automated segmentation method with the manual stereo drawing by a retinal expert grader. Areas are given as a percentage of the 3000- μ m-diameter circle.

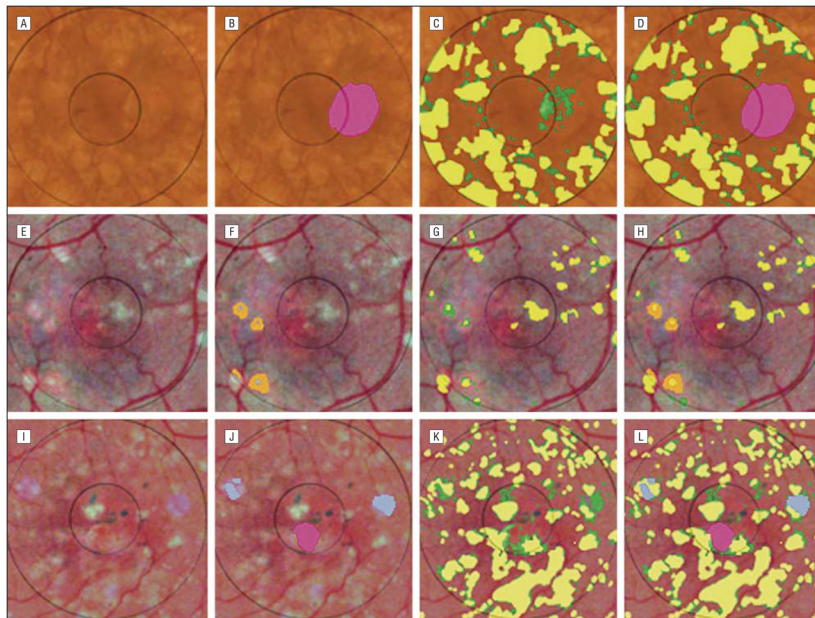


Figure 4.

Combined manual and automated segmentation of complex images. A is image 16 in the Table, showing ill-defined drusen and a lesion of geographic atrophy (GA). The latter was manually segmented in B (pink). The expert drusen drawing (yellow) was overlaid on the initial automated drusen segmentation in C to show the false positives (green), which include part of the GA. Removing the pixels identified as GA from the automated segmentation, shown in D, reduced the false positives by 3.4% and improved the specificity from 0.70 to 0.78. The combined method also provided segmentation into drusen and GA. E is the contrast-enhanced version of image 7. Fading drusen are surrounded by areas of retinal pigment epithelial hypopigmentation, which were manually segmented in F (orange). The expert drusen drawing (yellow) was overlaid on the initial automated drusen segmentation in G to show the false positives (green), which include part of the retinal pigment epithelial hypopigmentation. Removing the pixels identified as retinal pigment epithelial hypopigmentation from the automated segmentation, shown in H, reduced the false positives by 0.4% and improved the specificity from 0.87 to 0.98. I is the contrast-enhanced version of image 2a, showing ill-defined drusen, as well as a central lesion of GA and 2 bluish photographic artifacts that are segmented manually in J (pink and blue, respectively). The expert drusen drawing (yellow) was overlaid on the initial automated drusen segmentation in K to show the false positives (green), which include part of the GA and the photographic artifacts. Removing these confounding pixels from the automated segmentation, shown in L, reduced the false positives by 2.0% and improved the specificity from 0.73 to 0.78.

Table

Statistical Analysis of Manual Drawing and Automated Drusen Segmentation Results*

Photo Identification No.	Eye	Manual Drawing, %	Automated, %	False Positives, %	False Negatives, %	Sensitivity	Specificity
1	Right	11.4	13.9	5.7	3.3	0.71	0.59
2a	Left	33.5	29.5	6.5	10.5	0.69	0.78
2b	Right	30.5	31.3	8.7	7.9	0.74	0.84
3	Right	50.8	39.8	6.8	17.30	0.66	0.83
4a	Left	38.0	32.1	10.1	15.7	0.59	0.69
4b	Right	36.5	36.9	11.2	10.5	0.71	0.70
5	Right	17.8	16.4	5.2	6.2	0.65	0.69
6	Left	45.2	37.2	5.3	13.3	0.71	0.86
7	Left	7.2	3.1	0.1	4.2	0.42	0.98
8a	Left	36.6	28.6	5.5	13.4	0.63	0.81
8b	Right	43.0	30.4	2.5	15.2	0.65	0.92
9	Left	43.2	35.1	4.7	12.9	0.70	0.87
10	Left	23.3	29.9	14.1	7.5	0.68	0.53
11	Left	27.6	34.8	1.8	8.9	0.74	0.94
12	Left	48.1	43.2	1.7	6.7	0.86	0.96
13	Right	34.6	34.1	6.5	7.0	0.80	0.81
14	Left	22.3	24.7	6.5	4.1	0.81	0.74
15	Left	41.1	32.3	1.1	9.9	0.76	0.97
16	Right	32.4	36.1	7.0	8.2	0.75	0.78
17	Right	2.5	1.9	0.4	1.1	0.58	0.78

* Percentages refer to drusen area as a percentage of the 3000- μ m-diameter circle. False positives, false negatives, sensitivity, and specificity are calculated on a pixel-by-pixel basis for the automated segmentation compared with the expert manual drawing.