Published in final edited form as:
Conf Rec Asilomar Conf Signals Syst Comput. 2008 October 26; 42: 651–654. doi:10.1109/ACSSC.2008.5074487

Interactive Image Analysis in Age-related Macular Degeneration (AMD) and Stargardt Disease (STGD)

Prof. R. Theodore Smith,
Ophthalmology Department, Columbia University, New York, NY 10027 USA

Noah Lee,
Biomedical Engineering Department, Columbia University, New York, NY 10027 USA (phone: 212-854-5996; fax: 212-854-5995)

Jian Chen, **Mihai Busuioc**, and **Prof. Andrew F. Laine**
Department of Biomedical Engineering, Columbia University, New York, NY 10027 USA

R. Theodore Smith: rts1@columbia.edu; Noah Lee: nl2168@columbia.edu; Andrew F. Laine: laine@columbia.edu

Abstract

The literature of the last three decades is replete with automatic methods for retinal image analysis. Acceptance has been limited due to post-processing or tuning requirements that may be just as time consuming as the original manual methods. The point of view herein is that by taking advantage of the human visual system and expert knowledge from the outset, the promised efficiencies of digital methods can be achieved in practice as well as in theory. Thus, simple labeling of regions of interest that is accepted and easily performed in a few moments by the human can provide enormous advantage to an already well-developed algorithm. Three examples are provided: drusen segmentation, image registration, and geographic atrophy segmentation, with applications to disease understanding.

Index Terms

Image Analysis; Interactive Segmentation; Age-related Macular Degeneration (AMD); Autofluorescence; Stargardt Disease (STGD)

I. Introduction

Clinical medical retinal research, in particular, and visual science in humans, in general, is based on minimally invasive testing with imaging serving as the surrogate for biopsy. Given the transparency of ocular tissue, retinal images are able to provide large amounts of valuable information. Image analysis of the retina can be performed in a variety of settings ranging from the standard digital fundus photograph to autofluorescence imaging and optical coherence tomography, all of which provide unique information to the viewer. Combined analysis of imaging data from multiple methods can reveal heretofore-unexpected relationships.

Herein we summarize the development and application of several digital tools for retinal image analysis, particularly the images of age-related macular degeneration (AMD) and its juvenile form, Stargardt disease (STGD).

Extensive drusen area, as seen on the fundus photograph, is the greatest risk factor for the progression of AMD (1). When examiners are asked to mentally aggregate the amount of drusen occupying a given macular subfield (2), as in the International System, where drusen

areas were estimated to within 10% to 25% or 25 to 50%, and so on (3), these semi-quantitative estimates prove difficult for human observers. Clearly, there is a pressing need to implement more precise techniques to improve the quality of data being gathered in clinical trials and epidemiological studies.

In addition to standard fundus photography, autofluorescence (AF) imaging with the scanning laser ophthalmoscope (SLO) has played a greater role in understanding AMD. It is already clear that the autofluorescence of RPE lipofuscin, which contains known fluorophores including A2E, is related to AMD (4,5), and focally increased AF (FIAF) is demonstrated in a broad range of AMD patients (6,7). Lipofuscin is also imaged by AF as it accumulates in the flecks of juvenile macular degeneration (4). or Stargardt disease (STGD) Fundus AF images are recorded with a confocal SLO (model HRA/HRA2; Heidelberg Engineering, Dossenheim, Germany), which uses blue laser light at 488 nm for illumination and a barrier filter at 500 nm to limit the captured light to autofluorescent structures.

Digital image analysis techniques face significant obstacles in drusen identification (8–11). First, the inherent nature of the reflectance of the normal macula is non-uniform. There is less reflectance centrally and increasing reflectance moving out towards the arcades. Local threshold approaches to drusen segmentation met with only partial success because the background variability limited the extent to which purely histogram-based methods could succeed. This increased the need for operator intervention and has been the main obstacle to automating drusen segmentation.

Another major obstacle to drusen identification is that of boundary definition: soft, indistinct drusen have no precise boundary, and thereby the solution to their segmentation, by definition, cannot be precise. Indeed, expert manual drawings themselves are necessarily variable, and in such cases we have found that specificity and sensitivity calculations for expert manual drawings of two retinal experts can demonstrate significant inter-observer differences (12). See Fig 1. Therefore, accuracy in digital drusen segmentation relative to expert graders can only reasonably be held to a comparable standard.

II. Mathematical Models of Macular Images

Concept of Background Leveling

The non-uniform reflectance of the normal background of the fundus and photographic illumination cause the apparent intensity of drusen to differ with location. While human observers can compensate, computers cannot, limiting the use of digital methods for drusen measurement.

If the variability in reflectance could be converted to a uniform background, global threshold selection could be applied for uniform object identification, avoiding the obstacles that have foiled local thresholding. This is the concept behind background leveling, which involves correction of the macular image in multiple regions, exploiting the specific geometry of macular reflectance. This concept is quite general, with applications beyond the analysis of ophthalmic images, and appears to be original in both the medical and imaging literature [US Patent # 7,248,736 B2, July, 2007]. The first step in the application of background leveling was to demonstrate that the mathematical model, consisting of quadratic polynomials in several zones with cubic spline interpolation in blending regions between the zones, could approximate the global macular image background of a normal photograph or autofluorescence image with sufficient accuracy to allow its reconstruction and leveling (13–17). The next step was to show that the model, operating on user-defined subsets of background data in abnormal images, was still capable of accurately leveling the

background for reliable segmentation of drusen (16,17). Details of the model with interactive background input selection follow.

The general quadratic in 2 variables is fit by least-squares methods to any chosen background input of green-channel gray levels. Thus, the model consists of quadratics, 1 for each zone, with cubic spline interpolations at the boundaries. The model covers the full 6000 micron diameter macular grading zone with 4 zones in the 3000 micron diameter fovea/parafovea: a 600-micron central disc, three annular zones (600–1000, 1000–2000, and 2000–3000 micron diameter), and two outer annular zones (3000–4500 and 4500–6000 micron). The two outer zones are each subdivided into four quadrants, giving 8 outer zones, and thus 12 zones in all. Local models are obtained in each region by fitting the local background, as determined by the two-threshold Otsu method (see *infra*). The resultant global model is formed from the 12 local models with appropriate radial and angular cubic spline interpolations at interfaces.

To provide initial input to the background model, the automatic histogram-based thresholding technique known as the Otsu method (18) is employed in each zone. Briefly, let the pixels in the green channel be represented in L gray levels [1, 2,...., 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,…, k] and C_1 denotes pixels with levels [k+1,…, L]. Ideally, C_0 and C_1 would represent background and drusen. The Otsu method uses the criterion of between-class variance and selects the threshold k that maximizes this variance. 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,…, k], [k+1,…, m], and [m+1,…, L], respectively. The criterion for class separability is the total between-class variance giving three desired classes: C_0 (dark, non-background sources, e.g., vessels and pigment), C_1 (background) and C_2 (drusen). In particular, for each region there is an initial choice of background, C_1, for input to the background model. Before final background selection, however, the operator interactively modifies the Otsu method by drawing regions of interest for two other options: region 1, where multiple large, soft, ill-defined drusen are present, the upper (drusen) thresholds were each reduced by three gray levels; region 2, where few drusen (less than 5% estimated) are present, the drusen class C_2 is subdivided again by the single threshold Otsu method, with the higher values becoming the new C_2 and the lower values included in C_1. These brief drawings determine C_1 (the background) for input to the model. The background is then leveled by subtracting the model, with appropriate offset, and segmentation results from a uniform threshold. (see Fig 2 for application to a normal AF scan).

Autofluorescence Image Analysis

In order to make quantitative assessments of abnormal AF relative to the image background in the setting of significant background variability, the AF image can be “leveled” to an image with a uniform background with a mathematical model similar to that just described for fundus photographs (Fig 2). Because a consistently small fraction of pixels in a normal leveled image fell 2.0 σ above the image mean, we used this as a working definition of focally increased AF (FIAF).

III. User Interactive Tools for Clinical Studies

Image Registration Tool

The purpose of image registration is to spatially align two or more retinal images for simpler clinical review of disease progression by researchers and physicians. Since these images come from separate screening events and are often taken at changing fields of view, accurate image registration becomes essential. Our completely automated method developed for retinal image registration (19), involves first detecting corner points by a Harris detector

(20) and then assigning a main orientation for each corner point. After this, a local neighborhood from each corner point is chosen to extract intensity invariant feature descriptors (IIFD), which undergo bilateral matching. Incorrect matches are removed, the location of each match is refined, and finally, the transformation mode is selected. Details of these steps can be found in the IEEE International Symposium manuscript by Chen et al. (19). If the characteristics of the two images are too diverse, then the user can direct the algorithm to consider only the vasculature.

Drusen and Autofluorescence Segmentation Tools

User interactive background leveling as just described produces an image in which drusen or AF abnormalities appear on a uniform background for global thresholding. This is implemented in a GUI (21).

Geographic Atrophy Segmentation Tool

GA segmentation in AF images via our interactive operator selection tool (watershed method) is described in a companion paper (Interactive segmentation for geographic atrophy in retinal fundus images, N Lee et al., these proceedings).

IV. Applications

Age-related Macular Degeneration (AMD)

Many correlations have been made between pathologic changes on autofluorescence images and stages of AMD, with markedly decreased autofluorescence over large areas having been associated with geographic atrophy (GA). Furthermore, excessive lipofuscin accumulation in the RPE, seen in focally increased autofluorescence (FIAF), has been proposed to be a marker of RPE disease and photoreceptor cell degeneration (5). Through image registration and automated techniques for drusen identification, extended to analysis of FIAF lesions, improved correlation of such lesions with clinical stages can be obtained (7).

We specifically studied the relationship between FIAF and subsequent development of GA in AF images 2 to 3 years apart (22). GA was first segmented by our tool and masked from the image to be leveled by the model. After leveling the background of the remaining initial image, the mean and SD σ of the resultant leveled image were used to define the threshold for FIAF. The threshold was set at 2.0 σ above the mean to determine the total FIAF in the image. The FIAF contained within the 250-micron border zone of the initial GA lesion was determined. The initial and final AF images were registered with our tool. Because they were precisely superimposed, the area of GA from the initial image could be seen as a core of GA within the GA in the final image.

The positive predictive value (PPV) of FIAF, the probability that any pixel with FIAF would become part of the new GA (NGA) in the final image, was calculated. Negative predictive value (NPV), which is the chance that pixels without FIAF would not become atrophic, was also calculated. For both PPV and NPV, relative values were determined, which more accurately reflected the strength of the association by calculating it relative to chance (19). Compared with the relative PPV of chance of 1.0, the mean relative PPV of increased FAF was 1.15 +/− 0.28. The mean relative NPV was 1.00 +/− 0.02. Because the relative predictive values of FIAF were generally no greater than chance, these results suggested that FIAF is not a strong focal risk factor for development or extension of GA.

Stargardt Disease (STGD)

We evaluated STGD progression & lipofuscin levels via our automated autofluorescence (AF) image analysis. We analyzed the relationship between focally increased

autofluorescence (FIAF) and progression to geographic atrophy (GA) and focally decreased autofluorescence (FDAF) in serial, registered autofluorescence (AF) scans of 10 patients with STGD (20 eyes 40 scans, mean follow-up 2.0 years). We found that GA progressed uniformly and centrifugally in a transition zone with minimal FIAF. Few (mean 4.0%) excess lipofuscin deposits (FIAF) progressed to GA or FDAF, despite significant progression of GA (median 30%/yr) and FDAF (mean 25%/yr). There was no correlation between total initial FIAF load and subsequent rate of total GA or FDAF progression. As a spatial predictor, when normalized to the relative positive predictive value (PPV) of chance (1.0), the mean relative PPV of increased FAF for progression of FDAF was 0.56 +/− 0.40, and for GA 0.30 +/− 0.27.

We conclude that Stargardt disease manifestations of GA and focally decreased AF are not spatially predicted by elevated lipofuscin levels imaged with autofluorescence, suggesting alternate mechanisms to the generally accepted theory of lipofuscin toxicity. Instead, FIAF and FDAF tend to undergo focal remodeling, or even transition of FDAF back to FIAF. Indeed, FDAF tends to develop, not coincident with, but adjacent to initial FIAF (Fig 3). Similarly, as reviewed previously, we found for AMD that increased autofluorescence does not predict the spatial progression of geographic atrophy. Taken together, our results challenge the lipofuscin based theories of pathogenesis of two important retinal degenerations. These results could only have been obtained by applying precise registration and segmentation tools such as described herein.

Acknowledgments

This work was funded by NEI (R01 EY015520), the NYC Community Trust (RTS), and unrestricted funds from Research to Prevent Blindness.

This work was supported in part by NEI (R01 EY015520), the NYC Community Trust (RTS), and unrestricted funds from Research to Prevent Blindness.

References

1. Bressler NM, et al. Drusen characteristics in patients with exudative versus non-exudative age-related macular degeneration. Retina 1998;8(2):109–14. [PubMed: 3420311]
2. Bird AC, et al. An international classification and grading system for age-related maculopathy and age-related macular degeneration. The International ARM Epidemiological Study Group. Survey of Ophthalmology 1995;39(5):367–74. [PubMed: 7604360]
3. Bressler SB, et al. Interobserver and intraobserver reliability in the clinical classification of drusen. Retina 1988;8(2):102–8. [PubMed: 3420310]
4. Sparrow J, et al. A2E, a byproduct of the visual cycle. Vision Research 2003;43(28):2983–90. [PubMed: 14611934]
5. Lois N, et al. Fundus autofluorescence in patients with age-related macular degeneration and high risk of visual loss. American Journal of Ophthalmology 2002;133(3):341–9. [PubMed: 11860971]
6. von Ruckmann A, Fitzke FW, Bird AC. Fundus autofluorescence in age-related macular disease imaged with a laser scanning ophthalmoscope. Investigative Ophthalmology & Visual Science 1997;38(2):478–86. [PubMed: 9040481]
7. Smith RT, et al. Autofluorescence Characteristics of Early, Atrophic, and High-Risk Fellow Eyes in Age-Related Macular Degeneration. Invest Ophthalmol Vis Sci 2006;47(12):5495–5504. [PubMed: 17122141]
8. Shin DS, Javornik NB, Berger JW. Computer-assisted, interactive fundus image processing for macular drusen quantitation [see comments]. Ophthalmology 1999;106(6):1119–25. [PubMed: 10366080]
9. Sebag M, Peli E, Lahav M. Image analysis of changes in drusen area. Acta Ophthalmologica 1991;69(5):603–10. [PubMed: 1776413]

10. Ben Sbeh, Z.; Cohen Laurent, D. IEEE Computer Society DL. 1997. An adaptive contrast method for segmentation of drusen; p. 255-7.
11. 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.
12. Smith RT, et al. Automated detection of macular drusen using geometric background leveling and threshold selection. Archives of Ophthalmology 2005;123:200–207. [PubMed: 15710816]
13. Smith RT, et al. Digital filters for Correcting Common Sources of Error in Automated Macular Image Analysis. ARVO. 2006 Abstract.
14. Smith RT, et al. A method of drusen measurement based on the geometry of fundus reflectance. BioMedical Engineering OnLine 2003;2:10. [PubMed: 12740042]
15. Smith RT, et al. Photographic Patterins in Macular Images: Representation by a Mathematical Model. Journal of Biomedical Optics 2004;9(1):162–172. [PubMed: 14715069]
16. Smith RT, et al. Autofluorescence characteristics of normal foveas and reconstruction of foveal autofluorescence from limited data subsets. Invest Ophthalmol Vis Sci 2004;45(8):2940–2946.
17. Smith RT, et al. A method of drusen measurement based on reconstruction of fundus reflectance. British Journal of Opthalmology 2005;89(1):87–91.
18. Otsu N. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 1979;9(1):62–66.
19. Chen, J.; Smith, RT., et al. A Novel Registration Method for Retinal Images Based on Local Features. IEEE International Symposium on Engineering in Medicine and Biology; 2008.
20. Harris, C., et al. A combined corner and edge detector. Alvey Vision Conference; 1988. p. 147-152.
21. Uy K, Smith RT, et al. Automating Drusen Analysis With User-Friendly Graphical Interfaces With and Without Artifact Correction. ARVO. 2005 Abstract.
22. Hwang JC, Smith RT, et al. Predictive Value of Fundus Autofluorescence for Development of Geographic Atrophy in Age-Related Macular Degeneration. Invest Ophthalmol Vis Sci 2006;47(6):2655–2661. [PubMed: 16723483]

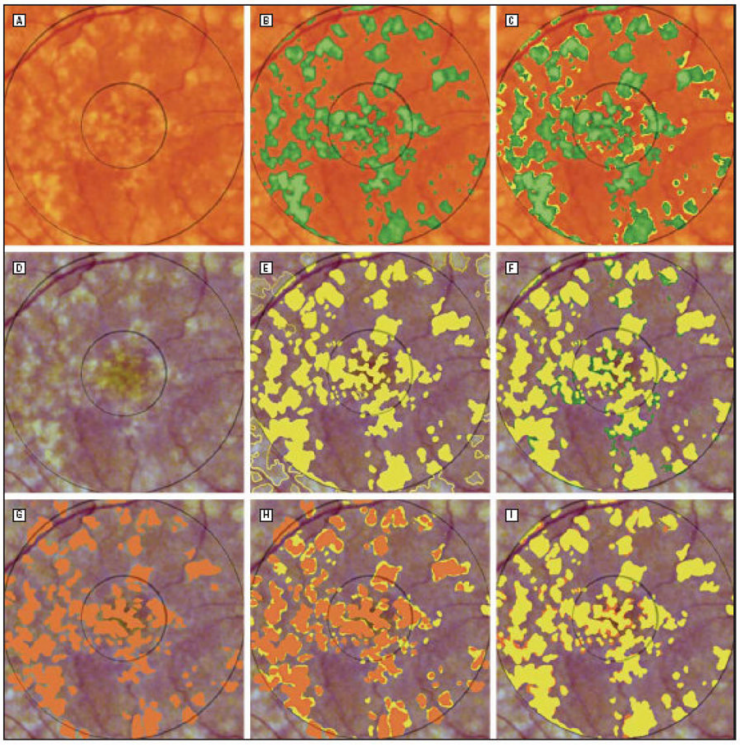

Fig. 1. Comparison of Drusen Segmentation by the Interactive Method vs. 2 Expert Drawings

The drusen in the original image (**A**) were identified by the interactive method (green) in (**B**), a retinal expert manually in (**E**) (yellow), and traced manually in (**G**) by a second retinal expert (orange). (**C**) shows the interactive results overlaid on the first expert's. The remaining yellow regions are false negatives (7.0%). (**F**) The overlays are reversed. The remaining green regions are the false positives (6.5%). Thus, although there is almost exact agreement in total drusen area (34.6% vs. 34.1%), pixel-by-pixel agreement is not exact (sensitivity, 0.80; specificity, 0.81%). Comparing experts' drawings in (**H**) and (**I**), disagreements are comparable. The second expert drawing has a sensitivity of 0.70 and a specificity of 0.91 relative to the first expert drawing

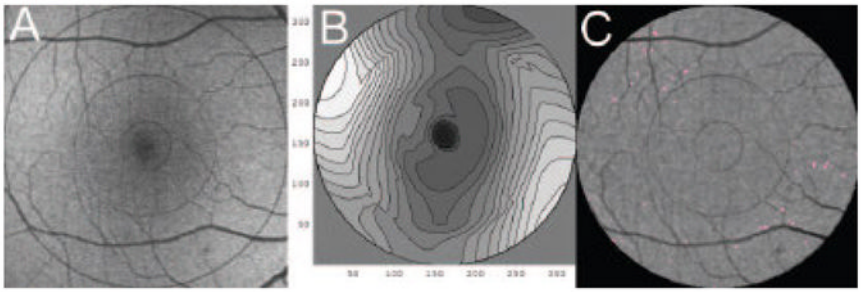

Fig 2. Mathematical Model and Segmentation of a Normal AF Scan

(**A**) Right eye showing significant background/illumination variability and foveal decreased fluorescence due to luteal pigment. (**B**) Twelve-zone mathematical model of the AF background in (**A**), presented as a contour graph. Note how the model captures the background variability of the original scan. It is essentially smooth throughout. The contour lines are closer together in the fovea where the background is more highly variable. (**C**) The image in (**A**) leveled by subtracting the model in (**B**). The background of the leveled image is now homogenous, with a mean gray level of 126 +/− 11.6 (SD). The global threshold of 2.0 standard deviations above the mean defining increased was applied to the entire leveled image and yielded the increased FAF shown in pink (0.28% of the 6000-micron zone), a reasonable selection. Thresholding the unleveled image (**A**) would cause major errors, due to the illumination variability.

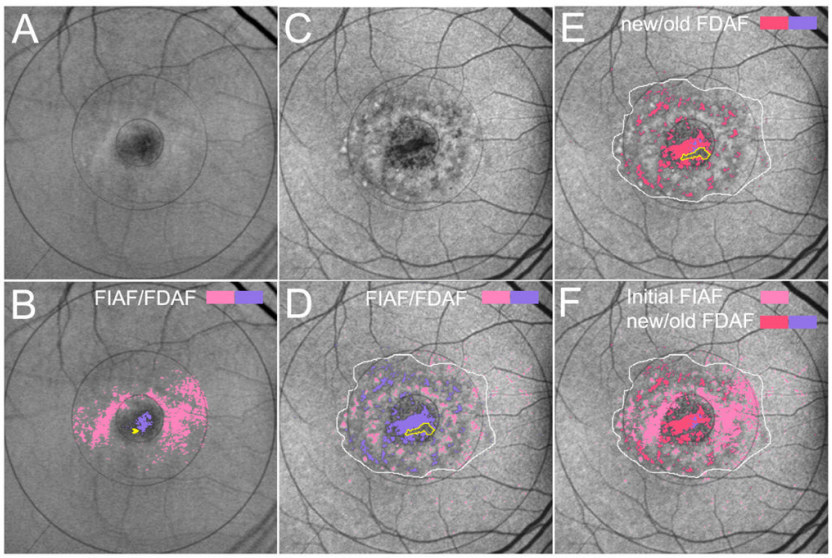

Fig 3. Stargardt Disease. Progression/remodeling of autofluorescence

A. AF image 2002, original. **B**. Focally increased autofluorescence (FIAF) and FDAF detected by the model. **C**. AF image 2007, original, registered to the image from 2002.. **D**. FIAF and FDAF in 2007, both more scattered and peripheral in the later image, with remodeling of the FDAF centrally. In particular there is some decreased autofluorescence (yellow arrowhead) in 2002 which now falls in the yellow outlined area in D, an island of mottled but not decreased autofluorescence within the larger new FDAF. **E**. FDAF is divided into new (not present in 2002) and old (present in 2002). **F**. To see the relation of FIAF to new FDAF formation, the FIAF from 2002 is overlaid on the new FDAF in 2007 New FDAF arises both inside and outside the initial FIAF ring, but little coincides. Despite significant growth of FDAF, only 4.8% of FIAF pixels turn to FDAF in 5 years, relative to 7.1% of randomly chosen pixels. New FDAF appears *adjacent* to initial FIAF but is not *coincident* with it: an FIAF lesion is *less* likely to turn to FDAF than another randomly chosen point.