

LESION BOUNDARY SEGMENTATION USING LEVEL SET METHODS

Elizabeth M. Massey, James A. Lowell

University of Lincoln, Brayford Pool, Lincoln LN6 7TS U.K.

Foster Findlay Associates Limited, Newcastle Technopole Kings Manor, Newcastle Upon Tyne NE1 6PA U.K.

Andrew Hunter, David Steel

University of Lincoln, Brayford Pool, Lincoln LN6 7TS U.K.

Consultant Ophthalmologist, Sunderland Eye Infirmary, Queen Alexandra Road, Sunderland SR2 9HP U.K.

Keywords: Computer vision, Retinal lesion segmentation, Segmentation, Level set methods.

Abstract: This paper addresses the issue of accurate lesion segmentation in retinal imagery, using level set methods and a novel stopping mechanism - an *elementary features* scheme. Specifically, the curve propagation is guided by a gradient map built using a combination of histogram equalization and robust statistics. The stopping mechanism uses elementary features gathered as the curve deforms over time, and then using a *lesionness* measure, defined herein, 'looks back in time' to find the point at which the curve best fits the real object. We implement the level set using a fast upwind scheme and compare the proposed method against five other segmentation algorithms performed on 50 randomly selected images of exudates with a database of clinician marked-up boundaries as ground truth.

1 INTRODUCTION

The diagnosis of diabetic retinopathy is based upon visually recognizing various clinical features. Retinal lesions are among the first visual indicators suggestive of diabetic retinopathy. To enable early diagnosis, it is therefore necessary to identify both frequency and position of retinal lesions.

This paper focuses on the segmentation of retinal lesions and presents an application of level set methods and a novel elementary features scheme for ensuring an accurate boundary detection solution. We present a novel stopping mechanism which uses elementary features gathered over time as the curve deforms and then a calculated *lesionness* measure to find the point in time at which the curve best fits the lesion candidate.

This paper is presented as follows: Sections 2 provides background information and discusses the current literature on region growing schemes as a basis for comparison. Section 3 describes the level set method used followed by a description of the algorithm and the process framework. Section 4 discusses the evaluation results and provides comparison and

observations about the proposed method. Section 5 concludes the paper.

2 PREVIOUS WORK

2.1 Segmentation Algorithms

Retinal exudates are an interesting challenge for segmentation algorithms as they vary in appearance, conforming to one of three structures: dot exudates, fluffy exudates and circumscribed plaques of exudate. Dot exudates consist of round yellow spots lying superficially or deep in the sensory retina (Porta and Bاندello, 2002). Exudates are usually reflective and may appear to have a rigid, multifaceted contour, ranging in color from white to yellow (Chen, 2002).

With varying shapes, sizes, patterns and contrast, exudate segmentation is a demanding problem, complicated by lighting variation over the image, natural pigmentation, the intrinsic color of the lesion, and decreasing color saturation at lesion boundaries (Goldbaum et al., 1990).

Several authors have presented algorithms for the segmentation of exudates in fundus images, attaining varied results. Ward et al., (1989) introduced a semi-automated exudate detection and measurement method, in which an operator selected a threshold value to segment exudates from a shade-corrected retinal background.

Sinthanayothin et al., (2002) presented a recursive region-growing algorithm applied to a contrast enhanced image. To reduce the effects of uneven illumination over the fundus, images were pre-processed to enhance local contrast. The intensity component of the IHS (Intensity Hue Saturation) model was decoupled from color and the fundus images converted from RGB (Red Green Blue) and normalized to IHS. Sinthanayothin stated that the algorithm would not detect faint exudate regions, nor distinguish between other similar colored lesions.

Wang et al., (2000) defines a feature space \mathbf{F} to include color and exposure information and represents the red, green and blue (R,G,B) channels as spherical coordinates. A training set for each of two groupings is obtained by selecting small sub-windows inside exudate and background regions. The means of each sub-window are calculated and stored as feature centers. For each pixel in the fundus image the illumination and color information are extracted and a minimum distance discriminant is calculated to determine lesions from background regions.

Osareh et al., (2001) introduced a fuzzy C-Means clustering algorithm based on the work of (Lim and Lee, 1990) to segment a color retinal image into homogeneous regions. To compensate for the wide variation of color in the fundus, the images are converted from RGB to IHS, normalized and finally locally contrast enhanced (described above in Sinthanayothin et al., (2002)). Osareh et al., (2001) state that the segmentation by FCM is a conservative process finding all but the faintest (ambiguous) exudate regions. False positive non-exudate segmented regions were also found by the algorithm caused by cluster overlapping, noise, and uneven color distribution.

Contrast Gradient Region Growing (CG), introduced in (Lowell, 2005), uses a traditional region growing method employing a pixel intensity aggregation scheme for region growth, while using a Gaussian smoothed gradient image to iteratively calculate a gradient contrast between a grown (core) inner boundary and a dilated outer boundary. The algorithm starts by using a small 5×5 sub-window morphologically applied to the fundus image, and then applying a maximum filter within each sub-window, producing peak points. The core region is then grown by appending (selecting) the brightest neighboring (bound-

ary) pixels on each iteration. This growing process continues, halting when the grown region loses its *compactness*. The final boundary is then located by using a combination of *diameter* and *contrast* to determine the point of growth at which the object's contrast gradient is most significant.

The literature on retinal image object segmentation using level sets focuses mainly on segmenting structures rather than pathologies. Excellent work by Wang et al., (2004) show the power of evolving a curve to map prominent structures in an image. Deschamps et al., (2004) used level sets combined with embedded boundary methods to simulate blood flow and segment major vessels. Lowell et al., (2004) used active contours, the fore-runner to level sets, to find the optic nerve head. The work described herein is based on the seminal paper from (Osher and Sethian, 1988) and the numerical implementation takes insights from Sapiro, chap. 2, (Sapiro, 2001).

3 LEVEL SET METHOD

For our work in lesion segmentation, level set methods provide the capability to determine not just the coarse shape of an object, but are extremely useful to tease out the fine delicate boundary fissures and curves that give a deeper look into the overall shape of a lesion candidate.

3.1 Curve Propagation

Beginning with the definition of level sets from (Osher and Sethian, 1988)

$$\phi_t + F |\nabla\phi| = 0, \text{ given } \phi(\bar{x}, t = 0) \quad (1)$$

then,

$$\frac{\partial\phi}{\partial t} = F |\nabla\phi| \quad (2)$$

and

$$\phi_t + F_0 |\nabla\phi| + \vec{U}(x, y, t) \nabla\phi = \epsilon K |\nabla\phi| \quad (3)$$

where: ϕ_t is the propagating function at time t ,

$F_0 |\nabla\phi|$ is the motion of the curve in the direction normal to the front,

$\vec{U}(x, y, t) \nabla\phi$ is the term that moves the curve across the surface,

$\epsilon K |\nabla\phi|$ is the speed term dependent upon curvature.

For our purposes, $\vec{U}(x, y, t) \nabla\phi$ is the gradient map, described in section 3.3 and $\epsilon K |\nabla\phi|$ is approximated using a central differencing scheme.

3.2 Numerical Implementation

We consider curve movement of the form:

$$\frac{\partial C}{\partial t} = \beta \tilde{N} \quad (4)$$

where $\beta = \beta(k)$, that is, β is a function of the Euclidean curvature. For simplicity we use $\beta(k) = 1 + \varepsilon k$ as our velocity function.

Let ϕ_i^n be the value of ϕ at a point (pixel) i at the time n . An algorithm to describe the evolution of the curve over a given time step is

$$\phi_{ij}^{n+1} = \phi_{ij}^n - \Delta t [\max(-\beta_{ij}, 0) \Delta^+ + \min(-\beta_{ij}, 0) \Delta^-] \quad (5)$$

where u_{ij}^n is the 'current' level set zero, Δt is the time step (or scaling factor) and the $[\max \dots \min]$ describes the *normal* component, and where

$$\Delta^+ = [\max(D_x^-, 0)^2 + \min(D_x^+, 0)^2 + \max(D_y^-, 0)^2 + \min(D_y^+, 0)^2]^{1/2} \quad (6)$$

$$\Delta^- = [\max(D_x^+, 0)^2 + \min(D_x^-, 0)^2 + \max(D_y^+, 0)^2 + \min(D_y^-, 0)^2]^{1/2} \quad (7)$$

and $D_x^-, D_x^+, D_y^-, D_y^+$ are the forward and backward difference approximations in the x and the y direction, respectively.

3.3 Gradient Map

The boundary of a lesion can be characterized by the point of strongest intensity contrast between itself and the background retina. By determining the gradient of image I_{orig} , this maximum rate of change can be exploited. Equation 5 propagates the curve ϕ over the surface u . Optimally, what we want is to propagate to an object edge and then stop when the curve has correctly formed to the (correct) perimeter pixels. To do this we must provide an edge stopping function. Since the retinal images are inherently noisy, and the edge pixels of retinal lesions can look very much like background pixels, we want a mechanism that smooths out the noise but preserves the edges. Isotropic filters (such as Gaussians) smooth the image, but also lose important detail. Anisotropic filters address the issue of edge preservation.

Perona and Malik suggested the following edge-preserving g function (Perona and Malik, 1990)

$$g(x) = \frac{2x}{2 + \frac{x^2}{\sigma^2}} \quad (8)$$

The function $g(x)x$ acts as a 'weighted' function in that, small gradient values x will receive high

weight and high gradient values will have low influence on the diffusion solution. In other words, areas of high gradient will be 'smoothed' less, thus preserving edges. We applied the function to create our gradient map

$$g_I(x,y) = \frac{2 * (I_n)}{(2 - (I_n)^2)} \quad (9)$$

where: I_n is a histogram equalized, normalized gray-scale (green channel) image $I(x,y)$ and $\sigma = 1$.

3.4 Stopping Criteria

Once the gradient map is generated from the original (gray-scale) image the curve propagates for a given number of iterations. Finding the 'best' stopping point for the curve is relative to the object boundary. In cases such as figure 1 the boundary is not well defined, even with a properly contrasted gradient map, and especially in the case of bright lesions, the 'boundary' can be much the same color as the background. It is for these reasons that we need to use a mechanism that is robust to conditions of noise and illumination variance.

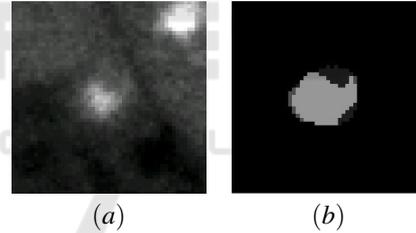


Figure 1: Curve Fitting: a) Gradient Map b) Match Results.

A traditional use of level sets is to track a curve to an object's boundary. In our case, it is more interesting to 'peek ahead' by allowing the curve to move past the optimal boundary and then 'look back' and measure how well-formed the accumulated region is as a lesion. We define the term *lesionness* as a combination of compactness ($c = p^2/a$), where p is the perimeter and a is the area (Gonzalez and Woods, 2001) and perimeter size constancy shp and use it as our 'stopping' mechanism. These measurements and others are explained in detail in section 3.5.2.

3.5 Process and Algorithm

The elementary features algorithm encapsulated in a three phase framework: 1) Pre-processing 2) Processing and Measurements 3) Best Fit Value Determination.

3.5.1 Pre-processing

The single channel, 59x59 pixel image I_{orig} is used to generate a gradient map as discussed in section 3.3. The initial level set begins as a small circle of $radius = 1$ and propagates outward according to equation 5 and as the curve deforms measurements at each change are taken.

3.5.2 Processing and Measurements

The starting point of the curve is determined using the simple peak detection algorithm described in Contrast Gradient Region Growing (above). The curve is then allowed to propagate past the optimal point (boundary) of the object. The purpose of this is to avoid the underestimation problem inherent in traditional region growing methods, and take advantage of ‘forward/backward looking’ measures.

We are looking for measurements that can give indicators of how well-formed a region is as a candidate lesion. Thus, elementary features include 1) the number of iterations the curve held its perimeter size: shp ; 2) the minimum compactness value: c ; 3) the number of iterations the curve held that compactness value: chp ; and 4) the gradient contrast: gc . Using morphological operations of dilation, equation 10, and erosion, equation 11, two ‘rings’, an inner and an outer ring, are generated about the curve. The contrast (difference) between these two rings is calculated.

$$D = C_0 \oplus CE \quad (10)$$

$$E = C_0 \ominus CE \quad (11)$$

$$gc = \sum_{p \in D} g_I(p) - \sum_{p \in E} g_I(p) \quad (12)$$

where C_0 is the infilled curve, CE is a 3×3 structuring element, g_I is the gradient map.

After the curve has moved for a number of iterations (we use $P = 180$) it is possible that the curve has evolved past the optimal point describing the object boundary. Because of this possibility, the gathered measurement values are then used to ‘look back in time’ to find the point at which the curve best fit the object boundary.

3.5.3 Best Fit Value Determination

Elementary features calculated are: shp, c, chp and gc . Of these, the two measurements that indicate curve stabilization (slowing down) are shp and chp . When the curve reaches an ‘edge’ its propagation rate slows down and over a number of counted iterations shp and chp remain constant. We track these stabilizing points and find that they tend to coincide with the other important features.

Let q be the iteration number and $h(q)$ be the count of the number of iterations for which the values of both chp and shp have held up to and including iteration q . Let q_M, q_N be the iterations with the two largest values of $h(q)$, $M < N$. Let q_c be the iteration with the smallest value of compactness c , and q_{gc} be the iteration with the largest contrast. Let Z be the set of critical iterations including q_M and q_N , and q_c if $M \leq q_c \leq N$, and q_{gc} if $M \leq q_{gc} \leq N$. Thus, the set Z includes the strongest stabilizing points and any other critical iterations between them. Sometimes there may be outlying critical iterations. For this reason we determine the largest gap between successive critical iterations and discarding those after the largest gap form the set Z^* , where $Z^* \subset Z$. We define the best fit point, SV , as the average of these critical iterations. Figure 2 shows the curve plots at the various elemental values.

$$SV = \frac{\sum_{q \in Z^*} q}{\#Z^*} \quad (13)$$

where $\#Z^*$ is the number of elements used.

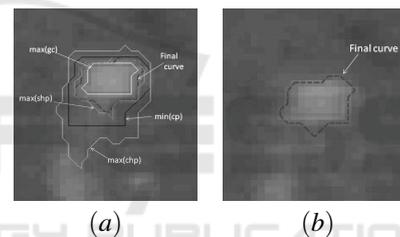


Figure 2: a) Plots of various elemental points, b) Final curve.

4 EVALUATION

A comparison is made between the presented algorithm and five segmentation approaches - fuzzy C-Means clustering, recursive region growing, adaptive recursive region growing, contrast gradient region growing and a color discriminant function. Table 1 shows the results of our evaluation.

Table 1: Algorithm Performance Metrics.

Model	Sens.	Spec.	Accuracy	Error
ELS	96.94	98.97	98.87	29.35
CG	96.24	98.71	98.59	36.59
AR	91.13	92.53	92.45	196.15
Fuzzy	88.29	94.18	93.89	158.95
RRG	47.72	90.99	88.85	290.1
DC	64.67	75.77	75.21	644.75

Where:

ELS - Elementary Features Scheme;

CG - Contrast Gradient;
 AR - Adaptive Recursive;
 Fuzzy - Fuzzy C-means;
 RRG - Recursive Region Grow;
 DC - Color Discriminant.

All algorithms were implemented and evaluated against a reference standard dataset of 50 randomly selected lesion images. Each image is provided with boundary markups by an expert ophthalmologist using custom designed software. The images are provided by the Sunderland Eye Infirmary with permission to be used in this research.

The benchmark comparison with the aforementioned techniques was achieved by measuring the number of common pixels shared between the reference standard and the algorithm's segmented area. The values in Table 1 were measured using pixel-wise sensitivity, specificity, accuracy and error-rate.

5 CONCLUSIONS

Algorithms for the automated segmentation and classification of candidate lesions have been presented. Although a number of algorithms have been published for lesion segmentation, many are unreliable due to marginal color and intensity difference between diabetic lesions and the background retina. This limited contrast has an adverse effect on alternate algorithms causing poor lesion boundary estimations.

Experimental comparisons have been conducted on five segmentation approaches - Contrast Gradient, Fuzzy C-Means clustering, recursive region growing, adaptive recursive region growing, and a color discriminant function. All algorithms were evaluated against a randomly-selected image set with ophthalmic lesion boundary demarcation. The results shown in Section 4 demonstrate the advantage of allowing the curve propagation (region growing) to run past the optimal boundary point, thus providing a 'peek ahead' to adjacent areas. Then using gathered elementary features to 'look back in time' to determine the best fitting curve.

REFERENCES

Chen, H.-C. (2002). *Vascular Complications of Diabetes; current issues in pathogenesis and treatment*, chapter 10, pages 97–108. Blackwell Publishing.

Deschamps, T., Schwartz, P., Trebotich, D., Colella, P., Saloner, D., and Malladi, R. (2004). Vessel segmentation and blood flow simulation using level-sets and

embedded boundary methods. *Computer Assisted Radiology and Surgery. Proceedings of the 18th International Congress and Exhibition*, 1268:75–80.

Goldbaum, M., Katz, N., Nelson, M., and Haff, L. (1990). The discrimination of similarly colored objects in computer images of the ocular fundus. *Investigative Ophthalmology & Visual Science*, 31:617–623.

Gonzalez, R. C. and Woods, R. E. (2001). *Digital Image Processing*. Prentice Hall, Upper Saddle River, NJ.

Lim, Y. W. and Lee, S. U. (1990). On the color image segmentation algorithm based on the thresholding and the fuzzy c-means technique. *Pattern Recognition*, 23:935–952.

Lowell, J. (2005). *Automated Retinal Analysis*. PhD thesis, University of Durham.

Lowell, J., Hunter, A., Steel, D., Basu, A., Ryder, R., Fletcher, E., and Kennedy, L. (2004). Optic nerve head segmentation. *IEEE Transactions on Medical Imaging*, 23(2):256–264.

Osareh, A., Mirmehdi, M., Thomas, B., and Markham, R. (2001). Automatic recognition of exudative maculopathy using fuzzy c-means clustering and neural networks. In Claridge, E. and Bamber, J., editors, *Medical Image Understanding and Analysis*, pages 49–52. BMVA Press.

Osher, S. and Sethian, J. A. (1988). Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79:12–49.

Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639.

Porta, M. and Bandello, F. (2002). Diabetic retinopathy a clinical update. *Diabetologia*, 45(12):1617–1634.

Sapiro, G. (2001). *Geometric Partial Differential Equations and Image Analysis*. Cambridge University Press.

Sinthanayothin, C., Boyce, J., Williamson, T., Cook, H., Mensah, E., and Lal, S. and Usher, D. (2002). Automated detection of diabetic retinopathy on digital fundus images. *Diabetic Medicine*, 19:105–112.

Wang, H., Hsu, W., Goh, K., and Lee, M. (2000). An effective approach to detect lesions in color retinal images. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 181–186.

Wang, L., Bhalerao, A., and Wilson, R. (2004). Robust modelling of local image structures and its application to medical imagery. In *ICPRO4*, pages III: 534–537.

Ward, N., Tomlinson, S., and Taylor, C. J. (1989). Image analysis of fundus photographs: the detection and measurement of exudates associated with diabetic retinopathy. *Ophthalmology*, 96(1):80–86.