

# A Robust Lesion Boundary Segmentation Algorithm using Level Set Methods

E.M. Massey<sup>1</sup>, A. Hunter<sup>1</sup>, J. Lowell<sup>2</sup>, and D.H. Steel<sup>3</sup>

<sup>1</sup> University of Lincoln, Brayford Pool, Lincoln, LN6 7TS UK

<sup>2</sup> Foster Findlay Associates Limited, Newcastle Technopole Kings Manor, Newcastle Upon Tyne NE1 6PA UK

<sup>3</sup> Consultant Ophthalmologist, Sunderland Eye Infirmary, Queen Alexandra Road, Sunderland SR2 9HP UK

**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 compare the proposed method against five other segmentation algorithms performed on 50 randomly selected images of exudates with a database of clinician demarcated boundaries as ground truth.

**Keywords**— Retinal Image Analysis, Segmentation, Level Sets

## I. 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. The threat to visual loss increases with the frequency of retinal lesions combined with their encroachment into the macula (one optic disc diameter around the fovea). To enable early diagnosis, it is therefore necessary to identify both frequency and position of retinal lesions in relation to the fovea. This paper focuses on the segmentation of retinal lesions and presents a novel elementary features scheme as a level set stopping mechanism for ensuring an accurate boundary detection solution. While most applications of level set methods have yielded excellent results, many assume a fairly noise free surface. We propose to apply level set methods to retinal images, which are noisy and have a slight surface curve especially near the edges. 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. Sections II and III provide background and discuss the current literature, respectively, on region growing schemes as a basis for comparison. Section IV describes the level set method and technique used. Section V discusses the evaluation results and provides comparison and observations about the proposed method. Section VI gives conclusions.

## II. BACKGROUND

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 Bandello, 2002). Fluffy exudates are paler than dot exudates and tend to lie more superficially in the sensory retina. Plaque exudates vary in size more than the other two groups and represent a more diffuse accumulation of lipoprotein. In addition to their various appearances, exudates may have various arrangement patterns. Exudates may surround leaking capillaries and microaneurysms in a circinate pattern or be randomly scattered. 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).

## III. SEGMENTATION ALGORITHMS

Lesion segmentation algorithms vary widely along with their results. The five chosen for comparison are discussed here for context. Ward et al., (1989) introduced a semi-automated exudate detection and measurement method, in which an operator selects 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 Gaussian smoothed, contrast enhanced image, where all but the faintest, or regions of similar color were correctly distinguished. Wang et al., (2000) defines a feature space to include color and exposure information and represents the red, green and blue (R,G,B) channels as spherical coordinates. Features, including pixel color and illumination, were used to perform regional segmentation. 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. The images are converted (RGB to HIS), normalized and locally contrast enhanced. The algorithm finds all but the faintest exudate regions; however, it has a high false positive rate 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. A seed point is determined using a small  $5 \times 5$  sub-window morphologically run over the fundus image, applying a maximum filter within each sub-window, producing peak points of the highest intensity pixel. The core region is then grown by appending (selecting) the brightest boundary pixels on each iteration. Once the growing process halts 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.

Since the pioneering work of Osher and Sethian (Osher and Sethian, 1988) Geometric Deformable Models, or Level Sets, have had a significant impact on the imaging community due to their capability to preserve the topological information in an image. However, 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) shows 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 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).

#### IV. LEVEL SET METHOD

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

$$\phi_t + F_0 |\nabla \phi| + \bar{U}(x, y, t) \nabla \phi = \varepsilon K |\nabla \phi| \quad (9)$$

where:  $\phi_t$  is the propagating function at time  $t$ ,  $F_0 |\nabla \phi|$  is the motion of the curve in the direction normal to the front,  $\bar{U}(x, y, t) \nabla \phi$  is the term that moves the curve across the surface, and  $\varepsilon K |\nabla \phi|$  is the speed term depending upon curvature. For our purposes,  $\bar{U}(x, y, t) \nabla \phi$  is the gradient map,

described in the next section and  $\varepsilon K |\nabla \phi|$  is approximated using a central differencing scheme.

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 *Iorig* this maximum rate of change can be exploited. Optimally, what we want is to propagate a curve 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 smoothes out the noise but preserves the edges. Isotropic filters (such as Gaussians) smooth the image, but also lose important detail. Anisotropic filters, however, address the issue of edge preservation. Foundational work in anisotropic diffusion (Perona and Malik, 1990) gives the following classical description:

$$\frac{\partial I(x, y, t)}{\partial t} = \text{div} (g(|\nabla I|, \sigma) \nabla I) \quad (10)$$

where:  $|\nabla I|$  is the gradient magnitude, and  $g(|\nabla I|)$  is an edge-stopping function and  $\sigma$  is a scale parameter. The  $g$  function is chosen to satisfy  $g(\mathbf{x}, \sigma) \rightarrow \mathbf{0}$  when  $\mathbf{x} \rightarrow \infty$ , so that diffusion is 'stopped' across the edges; see also (Black and Sapiro, 1999). We apply the following function from (Perona and Malik, 1990) to create our gradient map:

$$g_I(x, y) = \frac{2 * (I_n)}{(2 - (I_n)^2)} \quad (11)$$

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

#### V. PROCESS AND ALGORITHM

A single channel, 59x59 pixel image *Iorig* is used to generate a gradient map and 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. 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*

( $c = p2/a$ ), 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. We are looking for measurements that can give indicators of how well-formed a region is as a candidate lesion. The elementary features gathered, then, 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:  $chp$ ; 4) the maximum gradient contrast:  $gc$ . At each change in the curve shape two morphological operators (dilation and erosion) are used to calculate ‘rings’ about the curve. The contrast (difference) between these two rings is taken:

$$D = C_0 \oplus CE \quad (12)$$

$$E = C_0 \ominus CE \quad (13)$$

$$gc = \sum_{p \in D} g_I(p) - \sum_{p \in E} g_I(p) \quad (14)$$

Where:  $C_0$  is the infilled curve,  $CE$  is a 3x3 structuring element,  $g_I$  is the gradient map.

Of the elements tracked during propagation,  $shp$  and  $chp$  are indicators of curve stabilization (slowing down). Let  $q$  be the iteration number and let  $h(q)$  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  $M < q_c < N$ , and  $q_{gc}$  if  $M < q_{gc} < N$ . Thus, the set  $Z$  includes the strongest stabilizing points and any other critical iteration(s) 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 X shows an example of the plotted elemental feature points and the final curve.

$$SV = \frac{\sum_{q \in Z^*} q}{\#Z^*} \quad (15)$$

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

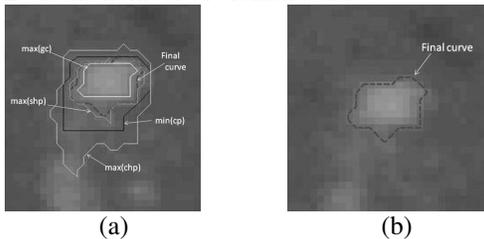


Figure X. (a) Plots of elemental features; (b) final curve.

## VI. EVALUATION

A comparison is made between the presented ELS algorithm and five other segmentation approaches. Table 1 shows the results of our evaluation.

Table 1 Algorithm Performance Metrics

Model	Sens.	Spec.	Accuracy	Error	Timing (secs)
ELS	96.94	98.97	98.87	29.35	561.26
CG	96.24	98.71	98.59	36.59	196.64
AR	91.13	92.53	92.45	196.15	69.70
Fuzzy	88.29	94.18	93.89	158.95	98.39
RRG	47.72	90.99	88.85	290.1	82.30
DC	64.67	75.77	75.21	644.75	483.28

ELS – Elementary Features Scheme; Fuzzy – Fuzzy C-means;  
CG – Contrast Gradient; RRG – Recursive Region Grow;  
AR – Adaptive Recursive; DC – Color Discriminant.

All algorithms were implemented and evaluated against a reference standard dataset of 50 randomly selected lesion images. Images were taken with a Canon EOS 20D attached to a Canon fundus camera and demarcated with boundary markups by an expert ophthalmologist. The images are provided courtesy of the Sunderland Eye Infirmary with permission to be used in this research. The benchmark comparison was achieved by measuring the number of common pixels shared between the reference standard and the algorithm’s segmented output. For each reference standard region  $R$ , true positive (pixels matched to reference standard)  $TP$ , false negative (pixels missed in reference standard)  $FN$ , false positive (pixels added over the reference standard)  $FP$  and true negative (background pixels in reference standard)  $TN$ . Statistics were calculated for each segmentation approach. The values in Table 1 were measured using pixel-wise sensitivity, specificity, accuracy and error-rate:

$$Sens = \frac{TP}{TP + FN} \quad (16)$$

$$Spec = \frac{TN}{TN + FP} \quad (17)$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (18)$$

$$Error = FN + FP \quad (19)$$

The following observations are made on the performance of the ELS algorithm:

**Accuracy:** As shown in table 1 the ELS method outperforms the CG algorithm in all areas and especially in showing a reduction in error. Experiments show that the CG algorithm tends to underestimate the lesions in general, as denoted by the sensitivity measure. This underestimation is

due to the smoothed gradient image used to determine the boundary contrast. As such, low contrast pixels get merged into the retinal pigmentation.

**Robustness:** The ELS algorithm does not depend on a single criterion, such as compactness to find a solution; rather, several measurements are taken as the curve propagates. Since the measurements are not dependent on specific thresholds, the true measures of the data can be taken into account during the initial value calculation and reassessment phases.

**Geometric:** The ELS algorithm also is not dependent on a single pixel value at a specific point in time, rather the curve moves in relation to curvature and direction of the normal. Thus, global as well as local information is used during curve propagation. Tracking the zero level set, as we do here, overcomes topological problems (such as discontinuities) that would hamper, even halt, traditional curve propagation algorithms.

## VII. CONCLUSIONS

A novel algorithm for the automated segmentation and classification of candidate lesions has been presented and compared against other well-known algorithms. Due to marginal color and intensity differences between lesion and background pixels, algorithms which depend on color and illumination are severely limited. The results shown in Table 1 demonstrate the advantage of allowing a curve to propagate past an optimal boundary point to ‘peek ahead’ to adjacent areas. Then use gathered features to ‘look back in time’ to determine the best fitting curve and thus accurate segmentation.

## ACKNOWLEDGMENT

The authors would like to thank Dr. Bashir Al-Diri for his contribution to this work.

## REFERENCES

Black, M. and Sapiro, G. (1999). Edges as outliers: Anisotropic smoothing using local image statistics. *Scale-Space Theories in Computer Vision*, pages 259–270.

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.

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 *ICPR04*, 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