Blood Vessel Segmentation in Retinal Images

J. O'Meara T. Miller P. Echevarria

March 13, 2004

Abstract

Segmentation of blood vessels in retinal images allows early diagnosis of disease; automating this process provides several benefits including minimizing subjectivity and eliminating a painstaking, tedious task. We applied two methods to this problem. First we developed local threshold probing, an efficient innovation on the traditional filter and threshold methods. Using this approach we achieved results with greater sensitivity and lower specificity than basic thresholding. Secondly we pioneered the use of a machine learning technique on retinal images. Using the technique, we have achieved a 70.3% true positive result with only 5.1% false positives.

1 Introduction

Examination of blood vessels in the eye allows detection of eye diseases such as glaucoma and diabetic retinopathy. Traditionally, the the vascular network is mapped by hand in a time-consuming process that requires both training and skill. Automating the process allows consistency, and most importantly, frees up the time that a skilled technician or doctor would normally use for manual screening.

While some success has been achieved on normal retinal images, on abnormal or diseased images - for which accuracy is more crucial than ever - the algorithms frequently fail. For instance, popular convolution approaches suffer from variable retinal background and low contrast between vessels and surrounding pixels. Tracking algorithms fail in special cases on abnormal images; they are often sidetracked by light objects and sometimes experience difficulty locating starting points. Current levels of success are still frequently inadequate for wide-scale implementation.

We implemented two methods for solving this problem on the publicly available STARE database [2]. The first method is an innovation on the

basic filtering and thresholding approach. Initially, to improve contrast in the image, we apply a variable width Gaussian filter at twelve different angles and take the maximum output at each pixel. Next, we implement our new technique of localized threshold probing in which the filtered image is locally thresholded based on seed points and then flood-filled. This technique was both simple and fast.

After realizing results only comparable to attempts by other researchers by variations of filtering and thresholding, we chose to experiment with supervised computer learning. Data about each of the image pixels was gathered and fed to a learning algorithm. A decision tree was generated and applied to test data. The output image was finally subjected to minor post-processing. Promising results were achieved in this new approach to retinal segmentation.

2 Background

The problem we are trying to solve is to segment images of retinas into portions that are blood vessel and not blood vessel. There has been extensive previous work on this problem and related problems, such as artery mapping in angiograms. This work can be divided into several techniques.

One technique is a window based method. This involves looking at a small region of the image and extracting possible blood vessel pixels based on local image characteristics. This was done by [7] in the mapping of arteries in angiograms. A model was created for blood vessel densitometry and matched to regions of angiograms to determine blood vessel location as well as diameter and cross-sectional area. Work done by [1] attempts to match the gray-scale intensities of regions of a retinal image to a Gaussian profile and thereby locate the blood vessels. Edge detection was done by [8], and parallel edges were connected and identified as blood vessels.

Another methodology involves exploiting the known structure of veins and arteries to evaluate potential vessels. Pattern recognition is utilized to eliminate spurious blood vessels [11]. More recent methods typically utilize a cross-section of the three major categories of retinal image segmentation. Generally, some type of filtering and thresholding is used in the preprocessing stage. The image is then processed in one of several ways, and finally a post-processing step detects misclassified pixels and patches together labelled segments. Li et. al. [4] reject the traditional single Gaussian filter for a double Gaussian filter which more accurately models the vessel profile. They use this piecewise Gaussian model to differentiate between arteries and

veins. Lowell et. al. [5] use a localization method utilizing a correlation filter to tackle the problem of locating the optic disk, frequently a starting point for tracking algorithms. Lalonde et. al. [4] use the Canny operator in order to achieve optimal edge placement and to calculate an accurate estimate of the normal vector the the edge. They follow the edge detection with thinning and then a tracking algorithm. Their tracking algorithm tracks each edge of each vessel individually before combining found walls as vessels by exploiting knowledge of the parallel character of vessel borders. Post-processing displays the centerline of the vessels. Hunter et. al. [3] developed an algorithm relying primarily on image filtering with secondary reliance on exploiting global connectivity. Order statistic filtering at multiple orientations is used for its robustness against noise and insensitivity to distractors. The results are cleaned up for noise and thinned. Finally, a Gaussian model is fit locally to the image at each pixel in order to measure vascular diameter and estimate beading.

In addition, several tracking methods have been implemented in the past. Previous tracking methods proceeded by first determining start points and then tracking the vessels from those points. One principal strategy involves locating the fundus, usually the brightest part of the eye, and locating starting points from there. For instance, Tamura et. al. [10] use this strategy and then incrementally locate pixels within the center of vessels by incrementally probing a vessel with a Gaussian profile. Tolias and Panas [12] also began at the fundus but propagated their probe forward using a fuzzy pixel-classification profile rather than a Gaussian profile. Others chose alternative means to locate tracking starting points. Hoover et. al. [2] skeletonize a thresholded image and start from the endpoints of the line segments they obtained. Subsequently, they probe the image and test segments for region-based properties to determine whether they are vessels or not.

One of the landmark works in this field was done by [2] and uses a hybrid method. These authors use matched filters (as in [1]) to extract possible vessels from the background. This filtered image is then segmented using iterative thresholding. One of the reasons the work is considered so crucial is that they hand labelled the blood vessels on 20 images and used these as the truth data. This dataset has been referenced and used by many subsequent researchers and was vital to our machine learning segmentation method.

3 Approach

As shown in Section 2, many researchers have worked on the problem of segmenting blood vessels in retinal images. To this work we add two new methods: localized threshold probing and segmentation by machine learning.

3.1 Variable Width Matched Filter

Initially, all images are filtered using two-dimensional matched filters in the manner of Chaudhuri et. al. [1] with the innovation of variable width. The filter employs three primary principles. First, pairs of blood vessel walls may be approximated as linear due to typically small curvature. Second, vessels appear darker than the background as a result of low reflectance in comparison to other retinal surfaces. The profile may be modelled as a Gaussian. Finally, the width of blood vessels is variable [1]. Precisely, the algorithm is as follows. A kernel is generated in the following fashion: $P = [x\ y]$ is a point in the kernel and θ_i is the orientation. The kernel is centered around the origin, [0 0]. The rotation matrix,

$$r_i = \begin{bmatrix} \cos\theta_i & \sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix}$$

is used, and $p_i = Pr_i$. Twelve different kernels are generated for an angular resolution of 15 degrees and five variations in the width of the Gaussian are considered, resulting in sixty total kernels. Each kernel is convolved with the image, and at each point within the image, we take the maximum response [1].

The Gaussian is extended to three standard deviations. The neighborhood N is defined as follows: $N = \{(u,v)||u| = 3\sigma, |v| = L/2\}$. Then within N, $K_i(x,y) = -\exp(-u^2/2\sigma^2)$. The mean value of each kernel, m_i, is determined by averaging across N, and for all p_i in N, $K_i' = K_i(x,y) - m_i$. Finally, the kernel is scaled by a factor of 10 [1].

3.2 Localized Threshold Probing

The first new method used to segment retinal images takes place in two parts. A matched filtered retinal image is first thresholded locally, and then regions of this image that are likely to contain blood vessels are filled in.

To perform the local thresholding the matched filtered images are divided into rectangular regions such that each image has 25 regions on a

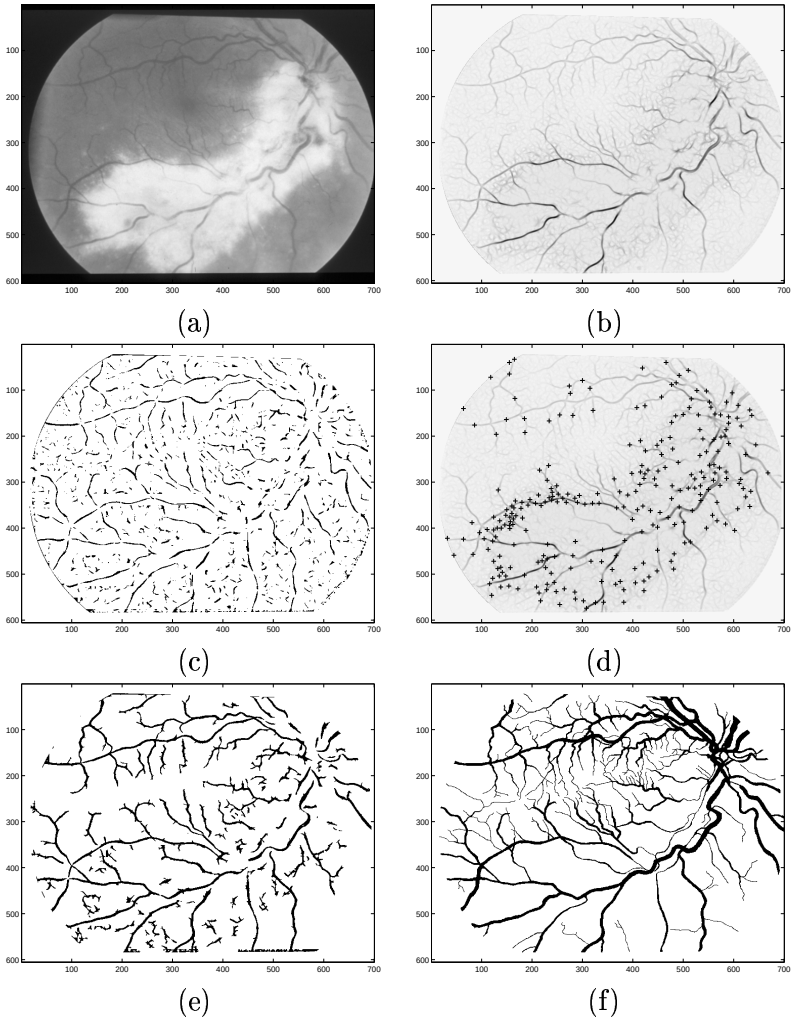


Figure 1: (a) An image of a retina, (b) the same image matched filtered, (c) locally thresholded with $P_{THRESH} = 8$, (d) starting points chosen, (e) the final segmentation using seed points from (d) and $P_{THRESH} = 20$, and (f) the hand labelled "truth" image. [2]

side. Within each block the image is then thresholded, keeping the brightest P_{THRESH}% of the pixels. This percentage is varied to produce an ROC curve. (ROC curves are described further in Section 4.1.) One sample result is shown in Figure 1(c) for $P_{THRESH} = 8$.

The second step of the algorithm requires initial seed points. These are calculated using another thresholding technique, followed by several transformations. The thresholding is done using a double thresholding technique. The matched filtered image is globally thresholded at two levels - such that 8 and 11 percent of all pixels are selected. These values were chosen experimentally to produce the best results. Connected regions of the 11 percent image that did not also appear in the 8 percent image are then eliminated. The thresholded image then does not include the minor regions that only appeared in the 8 percent image.

Having created a binary version of the image, it is then altered with a series of morphological operators. The image is skeletonized, spurs are removed, and isolated pixels are eliminated. Any intersections in the image are detected and subtracted from the skeleton of the image. Finally, the center points of all remaining line segments are saved to be used as seed points. These are saved in order of the length of the line segment they came from, from longest to shortest. This makes the first seed points more likely to come from blood vessels, as skeletonized blood vessels are composed of long line segments and typical image aberrations are not. An example of a set of starting points is shown in Figure 1(d).

The starting points chosen are then used to flood fill regions of the locally thresholded images. This ensures that those regions of the locally thresholded images which are retained are those most likely to belong to blood vessels. An example of a final image is shown in Figure 1(e). This can be compared with the hand-segmented image shown in Figure 1(f).

It should be noted that this method, while not as good as some results shown in literature, is quite fast and simple. Given a matched filtered image, calculation of the retinal segmentation takes about 30 seconds (on a 1.13 GHz AMD processor) and 75 lines of MATLAB code. However, the performance is still improved by nearly a factor of two over the most basic of algorithms.

3.3 Segmentation by Machine Learning

An approach completely novel to the world of retinal image segmentation was applied as an alternative to more traditional tracking or filtering and thresholding schemes. Supervised computer learning was applied to the images of the Hoover dataset. Features were developed on a pixel by pixel basis on retinal images, and the C4.5 computer learning algorithm was applied [9]. The Kouznetsova segmentations from the STARE data set were taken as truth images.

First, we applied the local threshold probing technique at 95% sensitivity. Information gained from filtering and thresholding designated some points as clearly not vessel and these points were considered for neither learning nor testing. In addition, points outside the retina were disregarded.

Then, data regarding image intensity, gradient magnitude, and entropy was collected for each pixel of images designated for learning. We chose eight features:

- Local image intensity from the green channel of the RGB image
- Local image intensity from the filtered image
- Gradient magnitude on the green channel of the RGB image using a 3x3 kernel
- Gradient magnitude on the green channel of the RGB image using a 5x5 kernel
- Gradient magnitude on the filtered image using a 3x3 kernel
- Gradient magnitude on the filtered image using a 5x5 kernel
- Log energy entropy of the green channel of the RGB image
- Log energy entropy of the filtered image

We collected intensity and log energy entropy data on a local basis. A window was considered around each point and the intensity data within that window normalized to values between 0 and 1. We noted the comparative intensity of the relevant pixel on both the green component of the original RGB image and on the image after a matched filter had been applied. We calculated the entropy for a window at each point. For both intensity data and entropy data, the size of the window was varied to maximize the disparity between vessel data and other data. Entropy can be calculated using several methods; by a process of trial and error, it was determined that calculating entropy based on log energy was most effective on both the green image and the filtered image. Entropy based on the logarithm of the energy is calculated as follows: $E3(s) = \sum_i \log(s_i^2)$ In general, feature data was collected for data points, and parameters were adjusted to maximize the disparity between the mean of data at vessel and not vessel points while minimizing the standard deviation with respect to the mean.

The data was summarized into a matrix of eight vectors containing complete feature data set for one image. This training data set contained approximately 225000 data points per feature. We inputted this matrix along with

a vector containing the corresponding vessel or not vessel indications into the C4.5 algorithm. ID3, developed by Ross Quinlan, utilizes the method of top-down induction of decision trees. At each node, a test is chosen to determine what best separates the learning data into the desired categories. The tree is biased by the information gain measure. The information gain measure is based on entropy is defined as follows. Given a set of classified examples E and partition $P = E_1, ..., E_n$,

$$ig(E, P) := entropy(E) - \sum_{i=1,...,n} entropy(E_i) * |E_i|/|E|$$

At each step, the test with the highest information gain is chosen. A corresponding set of successors is generated, and the algorithm continues recursively on the successors [6].

The C4.5 algorithm is an extension of the base algorithm ID3. It supports post-pruning after the tree is inducted to increase accuracy. We found our results to be only slightly impacted by pruning and chose to use an unpruned tree.

After applying the decision tree to classify all the image pixels of our test image, we performed post-processing on the image. First, a morphological close operation was performed. Next, we generated an image using local threshold probing with greater than 98% specificity. We added this thresholded image to the processed image classified by the C4.5 algorithm, thus increasing sensitivity with slight decrease in specificity.

4 Results

4.1 Local Threshold Probing

One example of graphical results for the local threshold p is shown in Figure 1(e). One can see that for the threshold chosen ($P_{THRESH} = 20\%$) the algorithm tends to miss very small and very large blood vessels. Missing very small vessels is a common problem with all known segmentation algorithms. It is caused by the vessels' very low contrast with the background of the image. Missing large vessels is a less common problem. In our results this arises from allowing the local thresholder to keep only the brightest 20% of the pixels when the large blood vessels may fill the entire local region.

Figure 2 shows a Receiver Operating Characteristic (ROC) curve to illustrate the success of the Local Threshold Probing method of retinal segmentation. ROC curves are generally used in medical literature to assess the success of a given test for diagnosis. A threshold in the test is varied

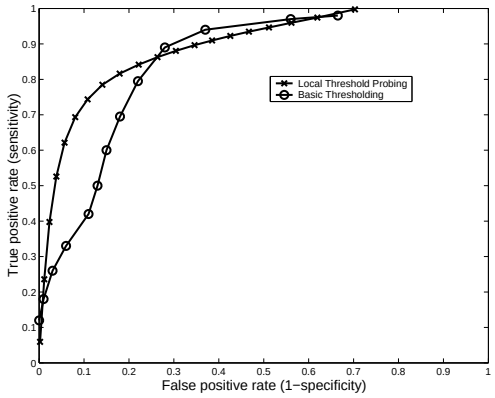


Figure 2: ROC curve for the Local Threshold Probing segmentation method. Compare with the curve for basic thresholding (from [2])

and the resulting number of false positives and true positives is plotted. For this particular test the value of P_{THRESH} (described in section 3.2) is varied. This may be compared to the results for another basic thresholding technique (from [2]).

One accepted method for evaluating the quality of a test is by evaluating the area under the ROC curve. One can see that in this case the Local Threshold Probing technique performs better than basic thresholding for the given images.

4.2 Segmentation by Machine Learning

We have implemented a computer learning algorithm for the first time to the problem of retinal image segmentation. We successfully combined this technique with more commonly used filtering and thresholding methods to achieve high sensitivity and low specificity.

Figure 3(a) displays an image to be segmented. Figure 3(b) shows the segmentation of the previous image by Kouznetsova. By applying the learning algorithm alone, we achieved a 60.3% true positive rate with a 4.15% false positive rate. Finally, Figure 3(c) displays the image segmented using the machine learning algorithm with post-processing applied. Comparing this result to Figure 3(b), one can see that once again our algorithm detected all principal vessels but failed to classify some small veins and arteries. After applying post-processing, we achieved a 70.3% true positive rate with only 5.07% false positives. Figure 4 displays a comparison of an image with and

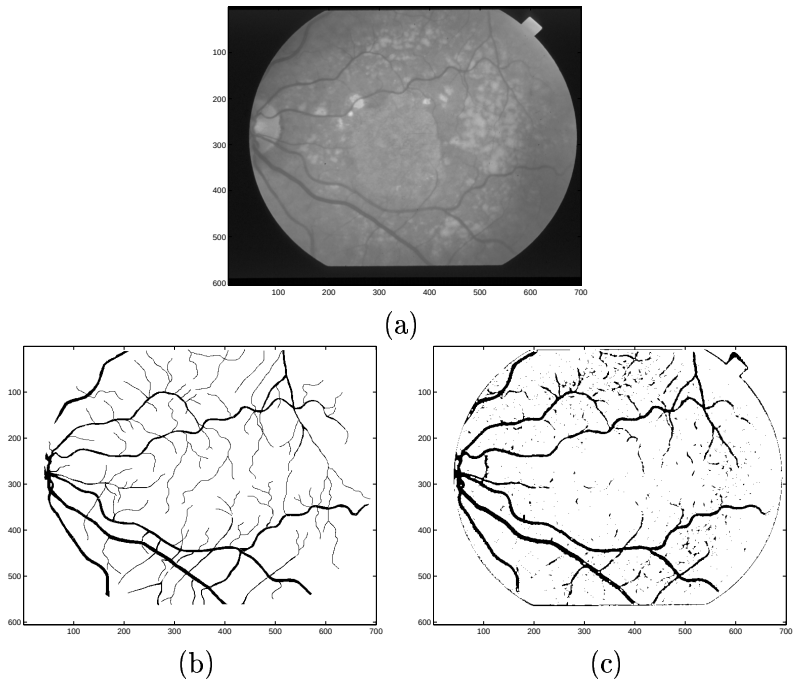
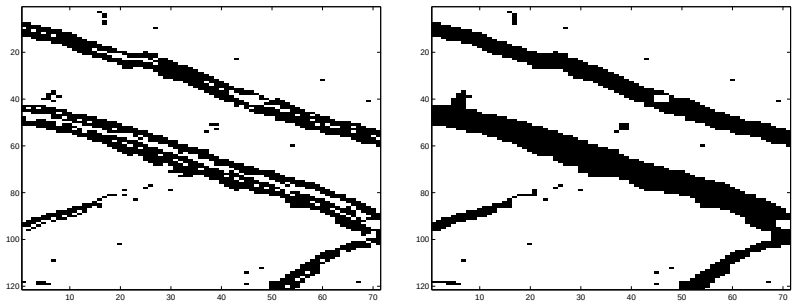


Figure 3: (a) A retinal image (b) The retinal image in (a) segmented by hand (c) The retinal image in (a) segmented by a machine learning algorithm

without post-processing. Figure 4(a) displays an image segmented using the machine learning algorithm without post-processing applied. Figure 4(b) shows the previous image with post-processing applies. The post-processing algorithm fills in gaps in the detected vessels at little cost to specificity. Applying post-processing more aggressively results in increased sensitivity at a cost to specificity.

We have generated an ROC curve by implementing the local threshold probing with varying values of P_{THRESH} in the post-processing stage. FigureřefROCcomputerlearning displays the curve overlaid with the curve achieved using local threshold probing. The results here were generated by training on data from a single image.



(a) (b)

Figure 4: (a) Section of a retina segmented by machine learning (b) The retinal image in (a) with post-processing applied.

5 Conclusions

In this paper we have presented to novel methods for segmentation of retinal images. The first, local threshold probing, is both fast and simple, yet still outperforms basic thresholding techniques. The second is the first application of machine learning techniques to the problem of retinal image segmentation. This method yielded improved results over local threshold probing.

The computer learning algorithm has great scope for future improvement. This problem has never previously been addressed, so our work has introduced first results as well as a realm of possibilities. The retinal images are a wealth of possible training data. We chose image intensity, gradient, and entropy as the focus of our learning algorithm. Other possibilities include more elaborate texturing data. In addition, one may generate learning data from not only the original image and a filtered image as we did, but an image processed in any manner of ways.

While the ready availability of data is an advantage, it can also create problems. The matrices we used for training were based on only one image at a time, yet they still contained 1.8 million data points. Obtaining software and hardware more suited to the large data processing needs of this problem would allow learning on large numbers of images at one time and theoretically could create a better decision tree. In addition, we chose to run the learning algorithm on at least half the points in our image. A more selective data-set or using a random selection of points from the image might yield a more manageable matrix of training data.

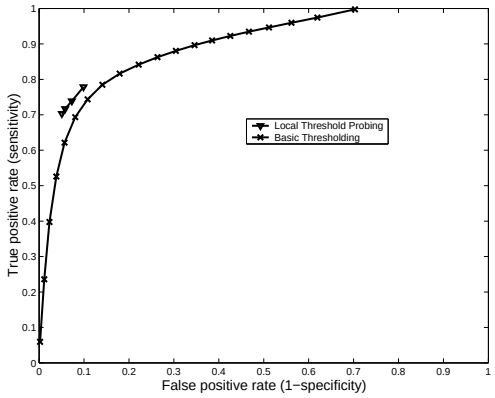


Figure 5: ROC curve for the computer learning algorithm compared with the local threshold probing algorithm result.

The fact that this initial foray into utilizing a computer learning algorithm on segmented retinal images yielded such promising results indicates that this avenue is indeed promising. We have left a benchmark of 70.3% true positives and 5.07% false positives for others to improve on.

References

[1] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Tran. on Medical Imaging*, 8(3):263–269, September 1989.

[2] A. Hoover, V. Kouznetsova, and M. Goldmaum. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. In *AMIA Annual Symposium*, 1998.

[3] A. Hunter, J. Lowell, D. Steel, A. Basu, and R. Ryder. Non-linear filtering for vascular segmentation and detection of venous beading. Technical report, University of Durham, 2003.

[4] M. Lalonde, L. Gagnon, and M.-C. Boucher. Non-recursive paired tracking for vessel extraction from retinal images. In *Proc. of Conference Vision Interface*, 2000.

[5] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, and E. Fletcher. Optic nerve head segmentation. Technical report, University of Durham, 2002.

[6] MLNet. Training information server: C4.5. Internet, March 2004. http://kiew.cs.uni-dortmund.de:8001/mlnet.

[7] T. Pappas and J. Lim. A new method for estimation of coronary artery dimensions. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 36(9):1501–1513, September 1988.

[8] A. Pinz, S. Bernogger, P. Datlinger, and A. Kruger. Mapping the human retina. *IEEE Trans. on Medical Imaging*, 17(4):606–619, August 1998.

[9] Ross Quinlan. Ross quinlan, ai group, cse. Internet, March 2004. http://www.cse.unsw.edu.au/ quinlan/.

[10] S. Tamura, Y. Okamoto, and K. Yanashima. Zero-crossing interval correction in tracing eye-fundus blood vessels. *Pattern Recognition*, 21(3):227–233, 1988.

[11] S. Tamura, K. Tanaka, S. Ohmori, K. Okazaki, A. Okada, and M. Hoshi. Semiautomatic leakage analyzing system for time series fluorescein ocular fundus angiography. *Pattern Recognition*, 16(2):149–162, 1983.

[12] Y. Tolias and S. Panas. A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering. *IEEE Trans. on Medical Imaging*, 17(2):263–273, April 1988.