Model Based Segmentation for Retinal Fundus Images*

Li Wang and Abhir Bhalerao

Department of Computer Science, University of Warwick, UK
{liwang, abhir}@dcs.warwick.ac.uk

Abstract. This paper presents a method for detecting and measuring the vascular structures of retinal images. Features are modelled as a superposition of Gaussian functions in a local region. The parameters i.e. centroid, orientation, width of the feature are derived by a minimum mean square error (MMSE) type of spatial regression. We employ a penalised likelihood test, the Akakie Information Criteria (AIC), to select the best model and scale for vessel segments. A maximum-cost spanning tree (MST) algorithm is then used to perform the neighbourhood linking and infer the global vascular structure. We present results of evaluations on a set of twenty digital fundus retinal images.

1 Introduction

The detection and measurement of blood vessels in fundus images has significant importance for many reasons. It allows a quantitative measurement of the geometrical changes of arteries, in diameter, tortuosity or lengths and can provide the localisation of landmark points (such as bifurcations) needed for image registration [1]. Furthermore, it can be used to measure the size of the optic disc and fovea and estimate the leakage of blood into the retina (exudate) that can indicate retinal disease and diabetes [2].

The collection of algorithmic tools developed for blood vessel detection can be, broadly speaking, classified into two major categories: tracking based approaches and template or model based approaches. Tracking based approaches work by first locating an initial point and then exploiting local image properties to trace the vascularture recursively [3] but require user intervention and appear to have a proclivity for termination near branch points. The 'matched filter response' method is a widely used template-based technique introduced by Chauduri [4] and further developed by Hoover [5]. A set of 2D Gaussian kernels with fixed length and orientation to enhance the vessels. A local threshold is then set to differentiate them from the retinal background. The method has some interesting features like handling bifurcations and obtaining fairly robust separation of vessels. However it is computationally intensive and the threshold selection may be critical.

* This work is funded by UK EPSRC (GR/M82899/01)

J. Bigun and T. Gustavsson (Eds.): SCIA 2003, LNCS 2749, pp. 422–429, 2003.

We propose a model based estimation that operates directly on the image intensities. Regions of the image are modelled locally as a superposition of Gaussian functions and an iterative, ML estimation used to derive their parameters. A penalised likelihood information measure, the Akakie Information Criteria, is employed to choose the best *model* and *scale* for vessel segments. A Kruskal type of maximum-cost spanning tree algorithm is then used to perform the neighbourhood linking and infer the global vascular structure. After a description of the methodology, results are presented on a sample retinal fundus image. To provide a quantitative evaluation of the algorithm, in terms of sensitivity and specificity, a receiver-operating characteristic (ROC) curve is also produced by comparing the detection results of the entire data set (20 images) with hand-label 'gold-standard' images from two expert observers. The paper is concluded by a discussion of the results and some proposals for further work.

2 Feature Modelling and Estimation

The global shape of the retinal vessel structure (figure 1(a)), in general, cannot be modelled by a single primitive form, therefore the model can only represent the local shape of a vessel in a small region (figure 1(b), (e)). Here, we exploit the symmetry and translational invariance properties of the Fourier domain to model linear and branching structures.

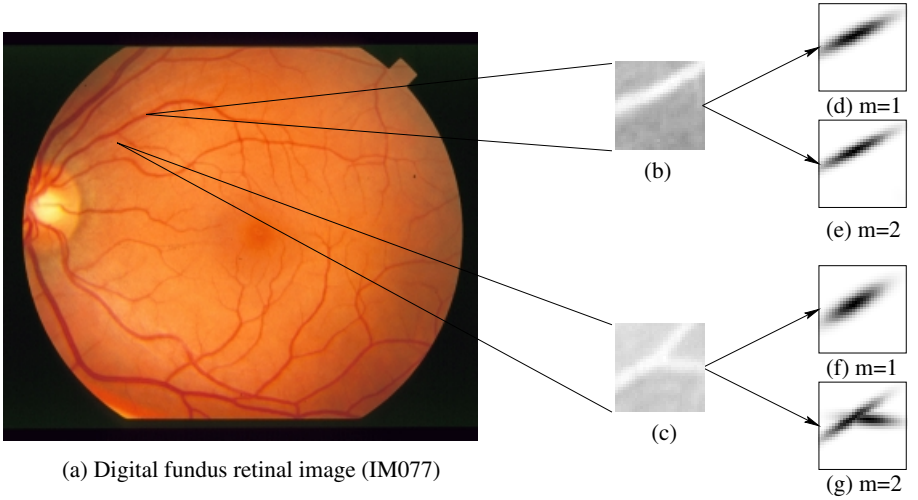


Fig. 1. (a) Sample of digital fundus retinal image (605×700). (b)-(c): Part of inverted grey level image. (d)-(e): Estimation linear structure shown in (b) using single and multiple Gaussian intensity models (f)-(g): Estimation of branching structure shown in (c) using single and multiple models. Block size is 32×32.

In the spatial frequency domain, a single linear feature (figure 1(b)) can be approximated by a 2-dimensional Gaussian function $G()$, whereas multiple linear features within the same region (figure 1(c)) are modelled as a superposition of Gaussian models and the spectrum is approximated as a sum of component spectra:

$$G(\boldsymbol{\omega}) = \sum_{m=1}^{M} |G_m(\boldsymbol{\omega})| \exp[-j\phi_m(\boldsymbol{\omega})] \tag{1}$$

where $G_m()$ is the mth Gaussian feature and $\phi_m()$ is the corresponding phase spectrum.

In particular, the phase-spectrum of a component exhibits a phase variation which is related to the feature centroid. We assume that this phase variation is independent for each of the M components of the model [6]:

$$\phi(\boldsymbol{\omega}) = \sum_{m=1}^{M} \phi_m(\boldsymbol{\omega}_m) = -\frac{B_l}{2\pi} \sum_{m=1}^{M} \boldsymbol{\mu}_m \cdot \boldsymbol{\omega}_m, \tag{2}$$

According to the Fourier shift theorem, the centroid μ_m can therefore be estimated within the block, B_l, by taking the average pairwise correlations between neighbouring coefficients along each axes.

Feature components are estimated by first separating out the component centroids using K-means clustering. Using the M spatial frequency coordinate partitions, the orientations and the widths (as variances) of the Gaussian components can be estimated by performing PCA on the covariances matrix, C_m, which is calculated from the inertia tensor of the spectral energy [7].

From the initial estimation, Θ_0 which is obtained from the spectrum of the windowed data, $f_B()$, we calculate the sample statistics Θ weighted by $w^t(\boldsymbol{x}) = f_B(\boldsymbol{x})G_m(\boldsymbol{x}|\Theta_t)$ using a set of iterative equations:

$$A_m^{t+1} = \sum_{\boldsymbol{x}} w^t(\boldsymbol{x}) / \sum_{\boldsymbol{x}} G^2{}_m(\boldsymbol{x}|\Theta_t), \qquad \boldsymbol{\mu}_m^{t+1} = \sum_{\boldsymbol{x}} w^t(\boldsymbol{x})\boldsymbol{x} / \sum_{\boldsymbol{x}} w^t(\boldsymbol{x})$$

$$C_m^{t+1} = 2\sum_{\boldsymbol{x}} w^t(\boldsymbol{x})(\boldsymbol{x} - \boldsymbol{\mu}_t)(\boldsymbol{x} - \boldsymbol{\mu}_t)^T / \sum_{\boldsymbol{x}} w^t(\boldsymbol{x}) \tag{3}$$

where t denotes iteration number. The initial estimation (eqn. 2) is performed in the spatial frequency domain whereas the following optimization is caculated in spatial domain (eqn. 3). The regression is a MMSE estimation and similar to an EM type approach. Unlike EM however, the weights here are not probabilities and do not sum to one. More importantly, the estimation implicitly takes into account the spatial arrangement of the data $f_B(\boldsymbol{x})$ relative to the intensity model $G(\boldsymbol{x})$ whereas EM estimates the underlying distribution from which $f_B(\boldsymbol{x})$ are drawn (see [7] for details). Convergence is achieved rapidly in 3-5 iterations. Figure 1(d)-(g) illustrates estimates of two type of linear features in part of a 2D retinal fundus image for single and multiple Gaussian models. Note that the parametric description Θ has information about feature amplitude and width together with position and orientation.

3 Model/Scale Selection and Feature Linking

The goal of the feature estimation and local structure type classification is to produce a succinct global object representation.

We firstly estimated the image (e.g. figure 1(a)) using multiple feature models $m = 1, 2$ at different block sizes, B_l. If the block sizes of each scales, l, are chosen to increase by a factor of two, ${B_l}^2 = 2^l$, then each four 'child' blocks, $\mathbf{c_0} = (2i, 2j, l)$, $\mathbf{c_1} = (2i+1, 2j, l)$, $\mathbf{c_2} = (2i, 2j+1, l)$, $\mathbf{c_3} = (2i+1, 2j+1, l)$, will correspond to one 'parent' block $\mathbf{p} = (i, j, l+1)$ at next level and forming a quad-tree feature representation of the data.

Once the parameters of each model at each scale have been estimated, the most fit model and 'natural' scale for the features is determined for a given image region by using a penalised distance measure, the Akakie information criteria (AIC), to bias the residual fit error which is expected to fall with increasing m. In minimum mean-square estimation, AIC is calculated using residual sums of squares (RSS) [8], $\chi^2 = \sum_x (f_{B_l}(x) - G(x|\theta))^2$.

$$AIC_m = {B_l}^2 \, ln(\chi^2 / {B_l}^2) + 2P \tag{4}$$

where ${B_l}^2$ is the number of data points and P is the number of parameters in the model. The recursive process to finding the optimal model m_{opt} for each block at each scale is summarised below:

1. Estimate the initial parameters for each model $m = 1$ for some starting block size. (In this paper we set $B_l = 16$.)
2. Use Eqns. 3 to improve the first estimate for each model m.
3. Compute AIC_m from RSS_m (Eq. 4).
4. Repeat steps $(1) - (3)$ for $m + 1 \leq M$.
5. $m_{opt} = \arg_m \min(AIC_m)$ (Eq. 5).
6. Repeat steps $(1) - (5)$ for $l + 1$ until m_{opt} is calculated at all levels.

According the quad-tree structure, the 'natural' scales of the features can also be determined by taking the smaller value of the AIC summations of child blocks across scales and comparing the information number with their parents:

$$\sum_{k=0}^{3} AIC_m(\mathbf{c}_k) \longleftrightarrow AIC_m(\mathbf{p}) \tag{5}$$

The block(s) whit the minimum AIC value is then picked to represent the data in a given region.

For a complete segmentation of the image, a neighbourhood linking strategy is also been employed to extract the global vascular topology. The data are modelled as a tree-like structure and each node of the tree are associated with a Gaussian kernel G. Adjacent blocks are linked to form a region adjacency graph where each vertex is a feature from the scale selection. The probability that two adjacent blocks i and j are part of the same vessel is estimated by calculating a link weight W_{ij} in an n-neighbour system by integrating the product of the two

Gaussian feature models along a line $\boldsymbol{y}(t) = \boldsymbol{\mu_i} + (\boldsymbol{\mu_j} - \boldsymbol{\mu_i})t,\ 0 \leq t \leq 1$, joining their centroids. The number of neighbours n vary according to the block sizes.

$$W_{ij} = e^{-(A_i - A_j)^2/(2\sigma_A^2)} \int_0^1 G_{\{\boldsymbol{A_i}, \boldsymbol{\mu_i}, \boldsymbol{y}(t)\}} \cdot G_{\{\boldsymbol{A_j}, \boldsymbol{\mu_j}, \boldsymbol{y}(t)\}} dt \tag{6}$$

where $e^{-(A_i - A_j)^2/(2\sigma_A^2)}$ is a coefficient which is formed by a normal distribution of the difference of the amplitude and models the likelihood of changes in amplitude between features. σ_A^2 is estimated from the data. The nodes can then be connected by using a modified Kruskal (mKruskal) method for finding an MST of a weighted graph, W_{ij}, as the cost along arcs and which constrains the resulting tree to be binary i.e vertices can have degree 3 or less. The algorithm proceeds as follows

1. Pre-sort the edge set, E, into descending order of weight
2. Examine each edge and add to the tree T if it satisfy:
   a) It connects two vertices from different component.
   b) The degree of vertices at each end is currently less than 3.
3. Stop when the weight of the current edge is less than some user defined threshold.

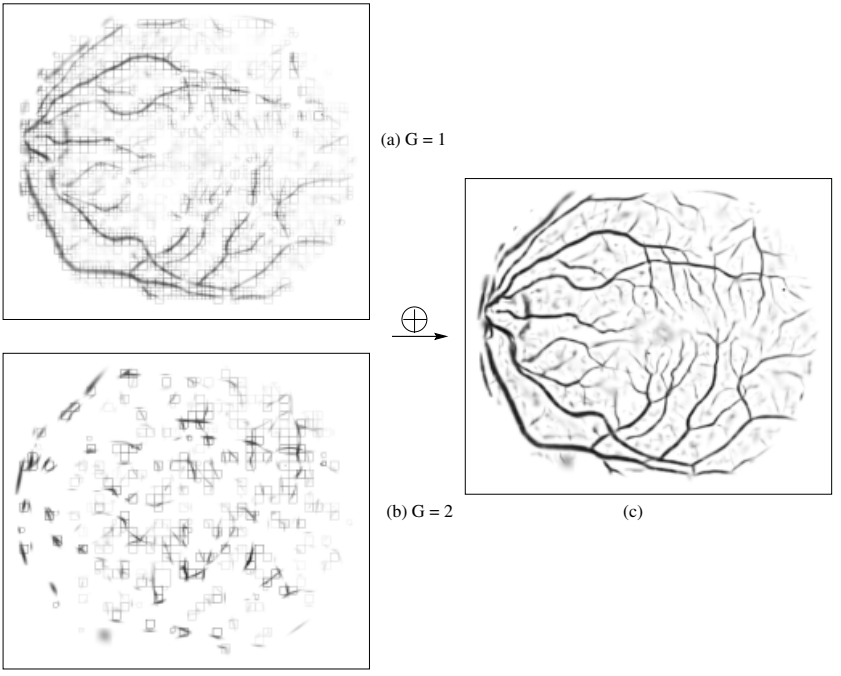


Fig. 2. (a) Reconstructions showing regions where single Gaussian model, $m = 1$, is selected. (b) Reconstructions showing regions where multiple Gaussian model, $m = 2$,is selected. (c) Combined reconstruction result after model/scale-selection

4 Experiments and Evaluations

The results of the described method were tested by using twenty retinal fundus images (size 605×700) obtained from the website of STARE project[1]. Prior to analysis the images are converted to grey scale by simple taking and inverting the green channel of the original RGB images.

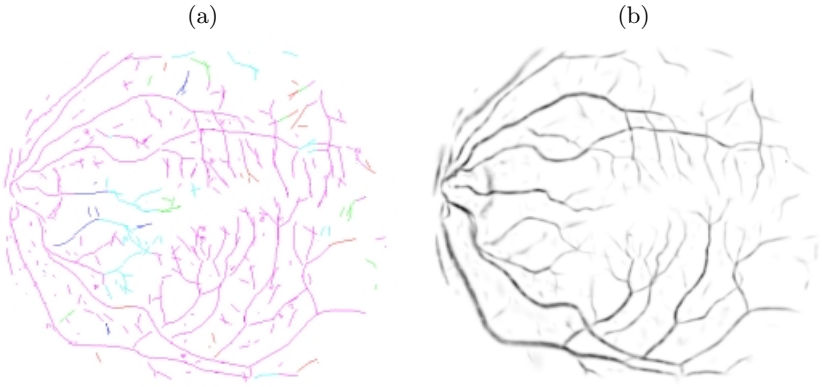


Fig. 3. (a) Feature linking result. (b) Reconstruction result.

Figure 2 illustrates results of the ML estimation and model/scale selection method. Overall, after the ML estimation, the orientation, position and width of the features both along blood vessels and near bifurcations are accurately modelled. Based on the AIC model/scale selection scheme, we then attempted to select the best model and scale for each feature to represent the data. Figure 2(a), (b) shows the reconstruction using different type of models, whereas a combined results is presented at figure 2(c). The highlighted block edges indicates that the features were located at different scales.

To infer global structure we have used a mKruskal linking strategy. In figure 3(a) the different colour is used to show the ability of the algorithm to successfully track and characterize much of the vessel topology, however, the deterministic linking approach is prone to becoming trapped in local maximum and unable to explore less certain, alternative explanations of the data. The reconstruction results in figure 3(b) shows that most of the noises apear in figure 2(c) is eliminated after using the linking algorithm. The binarised vascular structure reconstructions using the above method and hand-labelled ground truth images for two sample images from the data set are also presented at figure 4.

In order to further verify the algorithmic performance, the binarisation is produced by thresholding each feature block at 2 standard deviation ($t_\sigma = 2$) of

[1] `www.ces.clemson.edu/~ahoover/stare`

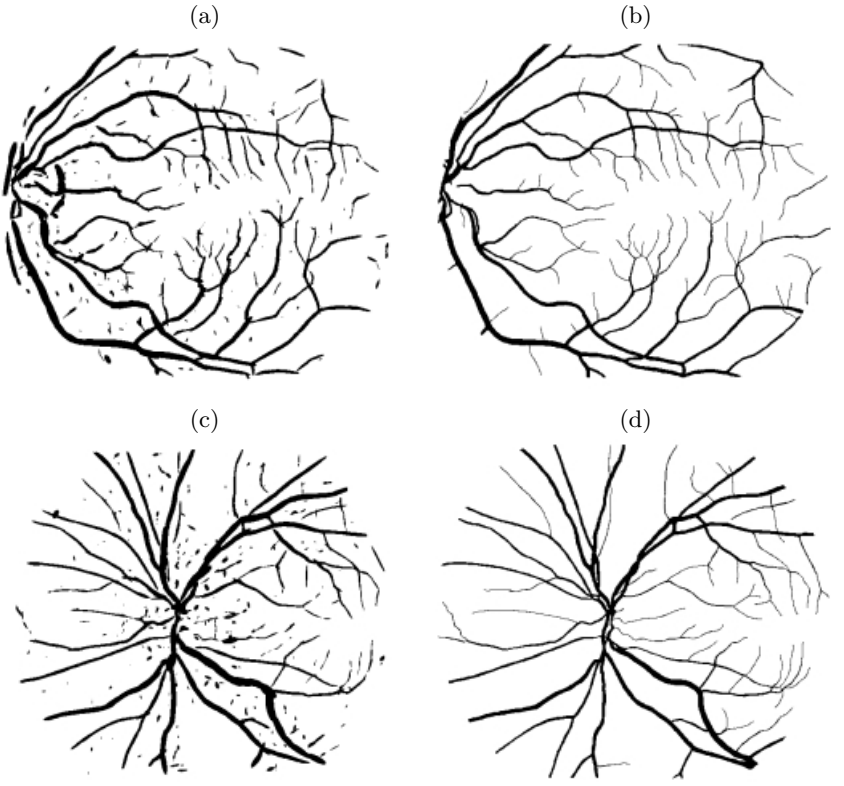
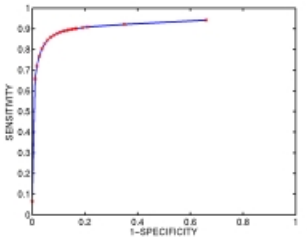


Fig. 4. (a)&(c) Binarised vessel segmentation for (IM077 & IM136). (b)&(d) Hand-labelled ground truth images.

(a)



True positive (TP) True negative (TN)
False positive (FP) False negative (FN)
SP = SPECIFICITY = FN/(FP+FN)
SE = SENSITIVITY = TP/(TP+TN)

(b)

Image	SE(%)	SP(%)	Image	SE(%)	SP(%)
1	74.6	90.5	11	85.2	96.5
2	76.6	92.5	12	88.2	95.5
3	76.1	90.1	13	78.8	95.0
4	87.2	89.5	14	81.7	95.0
5	70.4	94.2	15	85.5	94.1
6	80.5	92.4	16	85.4	93.5
7	90.6	92.7	17	87.2	92.5
8	88.5	93.0	18	76.2	96.5
9	91.5	93.2	19	71.5	94.6
10	79.2	92.2	20	85.0	92.9

Fig. 5. (a) ROC curve on the sample image by varying the binarisation threshold $t_\sigma = (0.5 - 3.5)$ (b) The table of SE/SP result across 20 images ($t_\sigma = 2$)

the Gaussian model which approximates well the local width of the vessel. By varying this thresholding, we plotted a ROC curve, (figure 5(a)).

The binarised vessel segmentation results of the whole data set (20 images) were also produced by using threshold value at $t_\sigma = 2$ and compared to the hand-labelled 'gold-standard' images. The results are listed at figure 5(b). For these experiments, our algorithm yields an average sensitivity of approximately 82% and a specificity of approximately 93%. These figures compare well with the results of match-filter response based algorithms [5].

5 Conclusions

In the paper, we have presented a vessel segmentation algorithm. The estimator can explicitly model bifurcations and is cheap to compute and robust in the presence of uncertainties in the image due to low signal to noise ratios. We have validated the approach against observer studies and shown it to have good accuracy and reliability. We are currently investigating the use of a randomised linking strategy [9] in place of the mKruskal to improve the vascular structure inference.

References

1. H. Shen, C. V. Stewart, B. Roysam, G. Lin, and H. L. Tanenbaum, "Frame-rate spatial referencing based on invariant indexing and alignment with application to online retinal image registration," *IEEE Trans. on PAMI*, vol. 25, pp. 379–384, Mar. 2003.
2. A. Pinz, S. Bernogger, P. Datlinger, and A. Kruger, "Mapping the human retina," *IEEE Transactions on Medical Imaging*, vol. 17, no. 4, pp. 606–619, 1998.
3. A. Can, H. Shen, J. N. Turner, J. L. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms," *IEEE Transactions on Information Technology in Biomedicine*, vol. 3, no. 2, pp. 125–137, June 1999.
4. S. Chardhuri, S. Chatterjee, N. katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Transcations on Medical Imaging*, vol. 8, no. 3, pp. 263–269, 1989.
5. A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical Imaging*, vol. 19, no. 3, pp. 203–210, 2000.
6. A. R. Davies and R. Wilson, "Curve and corner extraction using the multiresolution fourier transform," in *Image Processing and its Applications*. 4h IEE Conf., 1992.
7. L. Wang and A. Bhalerao, "Detecting branching structures using local gaussian models," in *International Symposium on Biomedical Imaging (ISBI)*. IEEE, July 2002.
8. K. P. Burnham and D. R. Anderson, *Model Selection and Inference*, Springer-Verlag, 1998.
9. A. Bhalerao, E. Thönnes, W. Kendall, and R. Wilson, "Inferring vascular structure from 2d and 3d imagery," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2001.