Statistical Pattern Analysis of Blood Vessel Features on Retina Images and Its Application to Blood Vessel Mapping Algorithms

Huajun Ying[1], Xing Wang[2] and Jyh-Charn Liu[3]

***Abstract*— Computer based modeling and analysis of blood vessel (BV) networks is essential for automated detection and tracking of anomalies and structural changes in retina images. Among many published techniques for automated BV mapping, optimal selection of thresholds to delineate BV pixels from their background pixels remains an open problem. In this paper we propose a novel representation of a BV pixel feature, daisy graph, using rotational contrast transform (RCT), and two feature descriptors energy E_p and symmetry difference S_p of the daisy graph. Non-BV pixels are separated from BV and boundary pixels based on E_p. Fitness of the lognormal distribution to S_p of BV pixels with negative E_p has been tested extensively for images in the STARE and DRIVE databases. Based on statistical pattern analysis in the feature space, we propose a fast self-calibrated BV mapping algorithm which achieve comparable and statistically sound performance as contemporary solutions.**

I. INTRODUCTION

The blood vessel (BV) network is the most prominent organelle in retina images. Precise BV network mapping is the key for accurate assessment of anatomic and pathological conditions. Despite a large number of BV mapping algorithms published in the literature, selection of optimal parameters to reflect differences between individuals remains an open problem and most algorithms used emperical results to choose parameters. Besides, BV segments at different generation levels [1] also have different characteristics in their width, length and transitional properties on their boundaries. It is relatively easier to capture the ridge of a BV, but the challenge level increases rapidly when we try to delineate BV vs. non-BV pixels near the boundary of a BV. A good solution needs to be able to incorporate individuality and the differences between large vs. small BV segments into the modeling process. To advance the state of art, there is a great need for self-calibrated parameter selection algorithms, which needs to be based on a sound theoretical model, so that one can reliably reflect individual differences in parameter selection to achieve the targeted sensitivity and error rate goals.

In this paper, we first propose two novel feature descriptors symmetry-difference S_p and energy E_p based on RCT. E_p is a powerful indicator to seperate BV and non-BV pixels. S_p of BV and boundary pixels with negative E_p fits to lognormal distribution, which provides sound statistical principles for parameter tradeoff. More importantly, the strong correlation between statistical measures of BV pixels and boundary pixels provides a way to personalize parameter as we can easily get boundary pixels using edge detector. Using increased threshold values in S_p derived from the fitted curve at each stage, we grow the vascular network from the ridge BV to BV boundary and small BV branches. Candidates are judged by continuity which avoids the problem of parameter difference of large and small branches. The new BV mapping algorithm achieves comparable performance as state-of-the-art algorithms, with additional advantages of low computational costs, and ease of parameter selection.

The rest of the paper is organized as follows. Related work on BV mapping is presented in section II. In section III, we discuss in details the feature extraction method on retina images and the statistical attributes. Then we introduce our BV mapping algorithm in section IV. In section V, we compare our BV mapping results with the state-of-the-art algorithms using conventional evaluation metric.At last, we conclude our work in section VI.

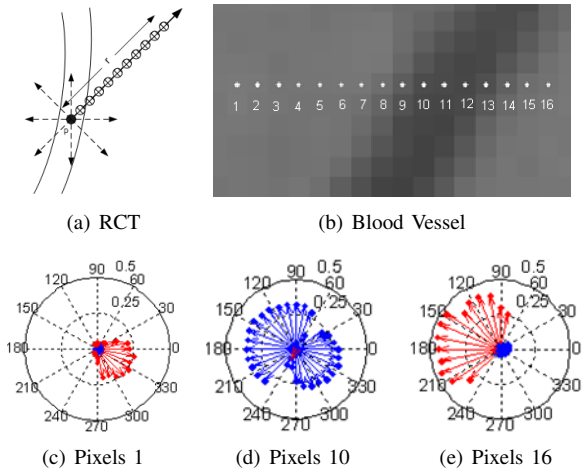
II. RELATED WORK

A detailed survey of blood vessel segmentation methodologies are presented in survey [2]. Blood vessel mapping algorithms in the literature are mainly based on shape matching filters, morphological operators, and linearity-continuity tracking. The Gaussian distribution, second order derivative Gaussian, or Gabor have been widely proposed for shape based filtering [3], [4], [5], [6]. A main challenge of this approach is selection of filter parameters for a broad range of BV shapes and sizes. The second major type of algorithms is based on mathematical morphology operators such as erosion, opening, closing, etc. [7], [8]. Typically, BV segments were first enhanced by a sum of top-hats morphological operators, and then BV pixels are detected based on linear coherent curvature calculation. Algorithms based on this approach work well on angiography images, but they tend to overestimate the BV width in the produced BVs maps due to quantization effects. Similar to the filter-based designs, they also face the challenge of parameter calibration. Most BVs have continuous flows with small intensity change. The continuity of the flow in the spatial domain is widely use for tracking algorithm [4], [9]. Our BV algorithm checks the candidates BV pixels based on this property.

With the blooming of machine learning, both unsupervised and supervised learning techniques have been proposed for pixel classification [11], [12], [13], where BV pixel properties are captured in high dimensional feature vectors for analysis. However, the results are hard to interpret and features extraction and pre-processing highly influence the performance.

[1]Department of Computer Science and Engineering, Texas A&M University, College Station, TX 77843 USA liu@cse.tamu.edu

978-1-4244-7929-0/14/$26.00 ©2014 IEEE



(a) RCT (b) Blood Vessel

(c) Pixels 1 (d) Pixels 10 (e) Pixels 16

Fig. 1. Rotation Contract Transform

III. Feature Analysis of RCT values of retina image pixels

A. Daisy graph and its feature descriptors

Rotational contrast transform (RCT) is a contrast based filter applied to the green channel of a color retina image, which obtains readings along different orientations [14]. The contrast C_p^θ [15] of a pixel p along a direction θ is defined as $C_p^\theta = (I_p - \hat{I_p^\theta})/\hat{I_p^\theta}$,where I_p is the intensity of p, $\hat{I_p^\theta} = 1/r\sum_{q\in N_p^{\theta,r}} I_q$ is the average intensity of p's neighbors along direction θ, and r the distance span of the neighborhood. The neighborhood of p along direction θ with size r is defined as: $N_p^{\theta,r} = \{(x_q,y_q)|x_q = \lceil x_p + kcos\theta \rceil, y_q = \lceil y_p + ksin\theta \rceil, k = 1..r\}$. Neighborhood size r is about twice the width of the vessel. We set $r = 21$ and $N = 32$.

In Figure 1, we illustrate how RCT is calculated and RCT (daisy graph) of some pixels. For pixels near the BV, there are large positive contrast values on directions towards the BV. For pixels inside the BV, the contrast values are axisymmetric along the flow direction and are negative values. The daisy graph representation provides a qualitative representation of the morphological property of the image pixel. To explore and quantify the relationship between shapes of daisy graph patterns with respect to BV, we propose two feature descriptors, energy E_p and symmetry-difference S_p, for pattern and statistics analysis of BV, BV boundary, and the background pixels.

For a pixel p, its energy descriptor $E_p = \sum_{m=0}^{N-1} C_p^{\theta_m}$, where $C_p^{\theta_m}$ is the RCV values along θ_m, $\theta_m = 2\pi m/N$. Based on Figure 1, we know that E_p for BV pixels is mostly negative while it is usually non-negative for non-BV pixels. The magnitude of E_p of larger BVs is usually larger than that of smaller/shallower BVs.

The symmetry-difference $S_p = \frac{\sum_{m=0}^{N/2-1} |C_p^{\theta_m} - C_p^{\theta_m+N/2}|/|\hat{C_p}|}{|Max(C_p)-Min(C_p)|}$,where $C_p^{\theta_m}$ and $C_p^{\theta_m+N/2}$ are the RCT values in opposite directions along θ_m. $\hat{C_p}$ is the mean of contrast magnitudes. $|Max(C_p) - Min(C_p)|$ denotes the maximal variation in the contrast along different directions. Taking it as the denominator will make the S_p of large BV smaller.

B. Statistical properties of E_p and S_p for BV and boundary

1) Human annotation: We use the twenty hand-labeled images in DRIVE [16] by two human experts as the reference sets to explore the statistical properties of different types of image pixels. The two hand-labeled BV maps (for each image) have some subtle differences, especially for BV boundaries and small/shallow vessels. In order to evaluate the robustness of statistical distribution against different detection sensitivity, we will use the double-marked BV maps, i.e., BV pixels marked by both experts and single-marked BV maps, i.e., BV pixels marked only by one expert. Non-BV pixels are those recognized as non-BVs by both experts. Our analysis shows that the ratio between single-marked to double-marked BV pixels is roughly 1:2, which indicates the significant discrepancies between two human experts due to different interpretation of BV boundaries.

To explore boundary properties of BV, we obtain the BV boundary maps from both reference sets. Here, BV boundary pixels are generated by applying the Sobel edge detector to the binary hand-labeled BV pixel maps(0 for non-BV pixel and 255 for BV pixel). The kernel size of the edge detector is 3x3 and a threshold of 130 is used to pick up the BV boundary pixels.

TABLE I
Ratio of pixels with negative E_p

	Dobule Marked	Single Marked	Expert 1	Expert 2	non-BV
Ratio(%)	94	68	80	84	27

2) The E_p descriptor for different types of image pixels: The result suggests that a threshold near the zero-crossing of E_p values is a reasonable first indicator to eliminate a large number of non-BV pixels. We can control the false negative value by adjusting the threshold on the E_p. Any potentially false positive BV pixels would then be eliminated by subsequent analysis of their S_p values.

3) The S_p descriptor for different types of image pixels: Using statistical model fitting techniques [17], we find that both the BV (blue color) and BV-boundary (red color) histogram plots can be fitted to the lognormal distribution functions as shown in Figure 2. We exploit six statistical measures on these lognormal distribution curves [17]: Ts_1: mode; Ts_2: median; Ts_3: mean; Ts_4: 1σ upper bound; Ts_5: 2σ upper bound; Ts_6: 3σ upper bound. Important statistical information can be inferred from these measures, such as the proportion of the area covered under the curve, curve shape, etc. Let $\{Ts_{BV}\}$ and $\{Ts_B\}$ respectively denote six measures for fitted distribution curves of BV and BV boundary pixels. We then compared the paired measures of $\{Ts_1,...,Ts_6\}$ from these two sets for all the images in DRIVE. Result based on groudtruth of expert 1 is shown in Figure 3, and result for expert 2 is similar. We observe that the paired measures from Ts_1 to Ts_5 values are highly concentrated

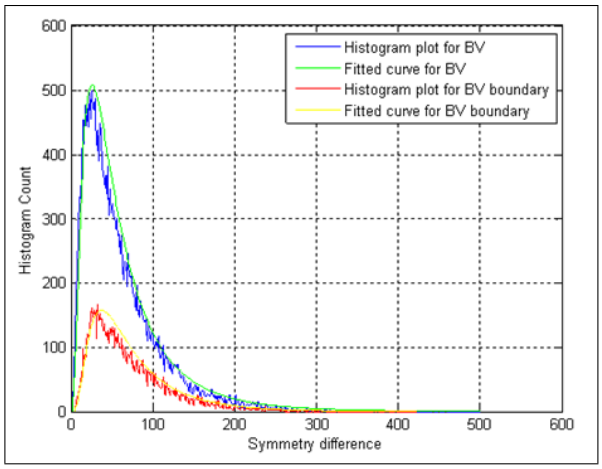


Fig. 2. S_p descriptor fitted to lognormal distribution

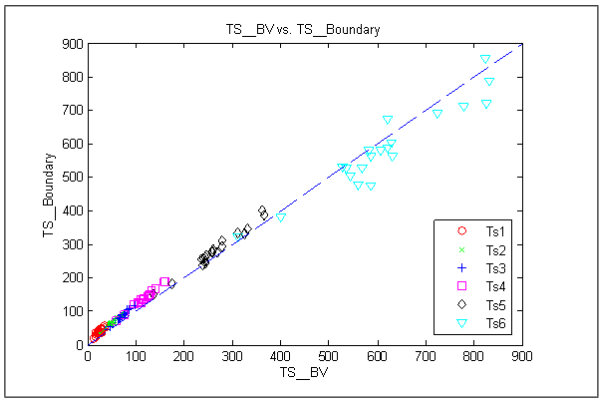


Fig. 3. paired measures from BV and Boundary pixels sets

along the $y = x$ line. Only a few pairs of Ts_6 deviate from each other slightly. We repeat this checking on STARE [18] database and results show the same conclusion. As a result, with high statistics confidence we can use personalized $\{Ts_B\}$ which is easier to obtain from edge pixels to predict $\{Ts_{BV}\}$.

IV. BV MAPPING ALGORITHM

From line 2-4, we can get personalized BV measure based on boundary pixels using Sobel operator based on the stable relationship between statistical measures of BV and BV boundaries. In line 5, ridge of the vessel is constructed using threshold of $Ts_1$and ratio of non-BV pixels includes is very low. In line 6,7, we get more candidates for BV pixels near boundary and smaller BV. Then from line 9-13, we use the continuity of the blood vessel to check whether candidates pixels are BVs. Firstly, there must be a flow direction at current pixel which is check in Line 9-10. C_{min} is based on the minimum contrast that is discernible by human eyes [19] which is set to be -0.02. Secondly, there must be a neighbor pixel within a distance γL along the flow direction in the current BV set as shown in Line 11-13. L is determined by the shortest BV segments which are approximately 10 pixels in length. γ is the error tolerance for BV continuity. As we increase γ, we will import more pixels of smeared

Algorithm 1 Blood Vessels Mapping

1: **procedure** BVMAPPING
2: Use the Sobel operator to get edge pixels set ES
3: Filter out pixels with negative E_p in ES
4: Fit S_p of pixels in ES to log-normal distribution curve, derive the $\{Ts_B\}$ set for the fitted distribution
5: Intialize BV set $BVS = \{p : S_p \leq Ts_1 E_p \leq 0\}$,$Ts_1 \in \{Ts_B\}$
6: **for** $\tau \in \{Ts_2, ..., Ts_6\}$ **do**
7: $CS = \{p : S_p \leq \tau, E_p \leq 0\}$ ▷ Candidate set
8: **for** $p \in CS$ **do**
9: $FS \leftarrow \{\theta : C_p^{\theta \pm \pi/2} \leq C_{min}\}$
10: **if** $FS == \emptyset$ **then** continue
11: $NS = \bigcup_{\theta \in FS} N_{\gamma L,\theta}$ ▷ $N_{\gamma L,\theta}$ is set of γL pixels in the neighborhood of p along the direction θ
12: **if** $NS \cap BVS! = \emptyset$ **then**
13: $BVS \leftarrow BVS \cup \{p\}$

and fragmented BVs. We set $\gamma = 0.8$, $L = 10$.

Let N denote the number of image pixels, the computing cost for RCT is r*N, where r is the number of neighbor pixels. It takes 2N+2 steps to compute S_p and E_p. The computing cost of the Sobel edge detection is O(N). The cost of Lognormal fitting is O(N), and the cost for continuity tracking O(N). Adding these costs together, we conclude that the asymptotic cost of the BV mapping algorithm is O(N).

V. EXPERIMENTAL RESULTS AND EVALUATIONS

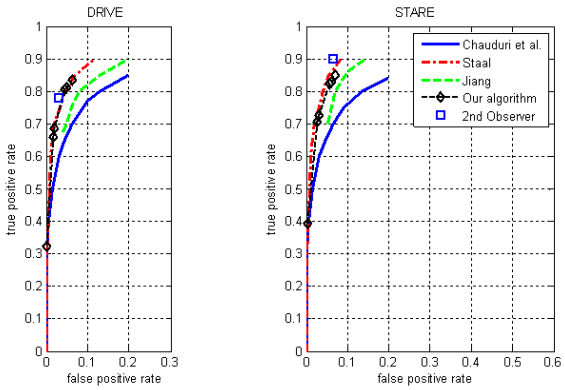
A. Experiment setup

We test the performance of BV detection scheme on all images of both DRIVE and STARE databases. A quantile threshold QT of 0.92 is posed on the Sobel edge detector results. Empirical results show that it serves well for extracting the BV edges with minimal side effects caused by severe noise or artifacts in highly diseased retinas. Experiment results show that $\{Ts_1, ..., Ts_5\}$ is not affected when QT changes in range $[0.9, 0.95]$. Only Ts_6 differs which might serve different purpose for small and shallow BV detection.

We adopt the most common conventional method to evaluate BV segmentation by selecting one human segmented result as the ground truth set and measure the algorithm performance using the receiver operating characteristic (ROC) curves. ROC curve is a graphical plot of the true positive rate (TPR) vs. false positive rate (FPR). For conventional binary classification problems, the closer a ROC curve approaches the top left corner, the better the performance of the method is and the area under the ROC curve is to be the largest equal to 1.

B. Results and analysis

Figure 4 shows the ROC curves for different methods on DRIVE and STARE databases using the ground truth set from the first human expert. Paired values of the TPR vs. FPR for our detection algorithm are obtained under different threshold values (from Ts1 to Ts6) on F for each stage. Paired measures for other detection algorithms were obtained



(a) DRIVE (b) STARE

Fig. 4. ROC for experiment on two datasets

based on different parameter settings reported. Our algorithm achieves similar performance with the Staals method [13] and outperforms Chauduri et al.s [3] and Jiangs algorithm [10]. Staals method is among the state-of-the-art algorithms that produce high area under the ROC curves. However, our detection scheme has much less computational cost as compared with Staals method. (10∼15 seconds for our algorithm compared with 15 minutes for Staals method on a 1GHz CPU)

Like many others, we also include the hand-labeled result from another observer into our discussion and use it as the human observer performance shown as the blue square in Figure 4. Due to the disagreement in human judgments on "single-marked" BV pixels, the human observer performance is not located at the top-left corner on the plot. Human observer performance is important in setting up the criterion for evaluating the BV detection performance, which can be taken as the upper boundary. And the curves of our algorithm are very close to human observer performances.

Although our algorithm is fast and accurate, there are still some limitation in our work. Firstly, some objects that strongly resemble BV segments may still be mistaken for BV pixels. Secondly, a closer look at retina images on DRIVE and STARE shows that few of them have vascular diseases such as neovascularizations, for which the ratio of large to small BV pixels may differ significantly.

VI. Conclusions and Future Work

In this paper, we proposed two novel descriptors and show that the distribution of BV and BV boundary pixels can both be fitted to the lognormal distribution with stable correlated measures. We further designed our BV mapping algorithm to capture the dynamics in BV generations that grows from "ridge to boundary, large to small branches". The algorithm is simple, fast, self-calibrated, and its ROC performance results are highly competitive against existing algorithms.

We would like to explore more descriptors and distribution to model other types of objects such as vessel pixels in curvatures, gaps, bends, ets. Our ongoing study also aims to capture the change in the distribution that may correlate to the vascular disease patterns. It is not known if any pre-disease conditions, such as mild hypertension would trigger change of the BV boundary patterns in the feature space. We also would like to conduct experiments on datasets with other image resolution to confirm that our algorithm is independent of the resolution.

References

[1] T. L. McKay, D. J. Gedeon, M. B. Vickerman, A. G. Hylton, D. Ribita, H. H. Olar, P. K. Kaiser, and P. Parsons-Wingerter, "Selective inhibition of angiogenesis in small blood vessels and decrease in vessel diameter throughout the vascular tree by triamcinolone acetonide," Investigative Ophthalmology and Visual Science, vol. 49, pp. 1184-1190, 2008

[2] M.M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, S.A. Barman, Blood vessel segmentation methodologies in retinal images A survey, Computer Methods and Programs in Biomedicine, Volume 108, Issue 1, Pages 407-433, 2012.

[3] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, Detection of blood vessels in retinal images using two-dimensional matched filters, IEEE Transactions on Medical Imaging, vol. 8, pp. 263-269, 1989.

[4] O.Chutatape, L. Zheng, and S. M. Krishnan, Retinal blood vessel detection and tracking by matched Gaussian and Kalman filters, Proceedings of the 20th Annual Conference of the IEEE Engineering in Medicine and Biology Society, vol. 6, pp. 3144-3149, 1998.

[5] M. Sopfka and C. V. Stewart, Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures, IEEE Transactions on Medical Imaging, vol. 25, pp. 1531-1546, 2006.

[6] M. Sopfka and C. V. Stewart, Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures, IEEE Transactions on Medical Imaging, vol. 25, pp. 1531-1546, 2006.

[7] F. Zana and J. Klein, Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation, IEEE Transactions on Image Processing, vol. 10, pp. 1010-1019, 2001.

[8] A. M. Mendonca and A. Campilho, Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction, IEEE Transaction on Medical Imaging, vol. 25, pp. 1200-1213, 2006.

[9] K. H. Fritzsche, A. Can, H. Sheng, C. Tsai, J. N. Turner, H. L. Tanenbaum, C. V. Stewart, and B. Roysam, Automated model-based segmentation, tracing, and analysis of retinal vasculature from digital fundus images, in Angiography and Plaque Imaging in Advanced Segmentation Techniques, S. Laxminarayan and J. S. Suri, Eds., London: CRC Press, 2003, pp. 225-290.

[10] X. Y. Jiang and D. Mojon, "Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, pp. 131-137, 2003.

[11] A. Bhuiyan, B. Nath, J. Chua, and R. Kotagiri, Blood vessel segmentation from color retinal Images using unsupervised texture classification, ICIP, vol. 5, pp. 521-524, 2007.

[12] E. Ricci and R. Perfetti, Retinal blood vessel segmentation using line operators and support vector classification, IEEE Transactions on Medical Imaging, vol. 26, pp. 1357-1365, October 2007.

[13] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. V. Ginneken, "Ridge-based vessel segmentation in color images of the retina," IEEE Transactions on Medical Imaging, vol. 23, pp. 501-509, 2004.

[14] M. Zhang and J. Liu, Directional local contrast based blood vessel detection in retinal images, ICIP, vol. 4, pp. 317-320, 2007.

[15] P. Zingaretti and A. Carbonaro, "Route following based on adaptive visual landmark matching," Robotics and Autonomous Systems, vol. 25, pp. 177-184, 1998.

[16] DRIVE Project, (2004) The drive databases, Image Sciences Institute, University Medical Center Utrecht. Available: http://www.isi.uu.nl/Research/Databases/DRIVE/.

[17] D. C. Boes, F. A. Graybill, and A. M. Mood, Introduction to the Theory of Statistics, 3rd ed, New York: McGraw-Hill, 1974.

[18] STARE Project, (Nov.2000) The stare databases, University of California, San Diego. Available: http://www.ces.clemson.edu/ ahoover/stare/

[19] A. B. Watson, H. B. Barlow, and J. G. Robson, "What does the eye see best," Nature, vol. 302, pp. 419-422, 1983.