

Retinal Vessel Segmentation Using Minimum Spanning Superpixel Tree Detector

Bin Sheng, Ping Li, Shuangjia Mo, Huating Li, Xuhong Hou, Qiang Wu, Jing Qin, Ruogu Fang,
and David Dagan Feng, *Fellow, IEEE*

Abstract—Retinal vessel is one of the determining factors in ophthalmic examination. Automatic extraction of retinal vessels from low-quality retinal images still remains a challenging problem. In this paper, we propose a robust and effective approach that qualitatively improves the detection of low-contrast and narrow vessels. Rather than using pixel grid, we use a superpixel as the elementary unit of our vessel segmentation scheme. We regularize this scheme by combining geometrical structure, texture, color, and space information in the superpixel graph. And the segmentation results are then refined by employing the efficient minimum spanning superpixel tree (MSST) to detect and capture both global and local structure of the retinal images. Such effective and structure-aware tree detector significantly improves the detection around the pathologic area. Experimental results have shown that the proposed technique achieves advantageous connectivity-area-length (CAL) scores of 80.92% and 69.06% on two public datasets, namely, DRIVE and STARE, thereby outperforming state-of-the-art segmentation methods. In addition, the tests on the challenging retinal image database have further demonstrated the effectiveness of our method. Our approach achieves satisfactory segmentation performance in comparison with state-of-the-art methods. Our technique provides an automated method for effectively extracting vessel from fundus images.

Index Terms—Vessel segmentation, retinal image, superpixel, feature extraction, MSST.

Manuscript received July 4, 2016; revised September 30, 2017, March 29, 2018, and April 29, 2018; accepted May 3, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61572316 and Grant 61671290, in part by the Research Grants Council of Hong Kong under Grant 28200215, in part by the National HighTech Research and Development Program of China (863 Program) under Grant 2015AA015904, in part by the Key Program for International S&T Cooperation Project of China under Grant 2016YFE0129500, in part by the Science and Technology Commission of Shanghai Municipality under Grant 16DZ0501100 and Grant 17411952600, and in part by the Interdisciplinary Program of Shanghai Jiao Tong University under Grant 14JCY10.

B. Sheng and S. Mo are with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: shengbin@sjtu.edu.cn).

P. Li is with the Faculty of Information Technology, Macau University of Science and Technology, Macau 999078, China (e-mail: lipingfire@ieee.org).

H. Li, X. Hou, and Q. Wu are with the Sixth People's Hospital, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: wyansh@163.com).

J. Qin is with the Centre for Smart Health, School of Nursing, The Hong Kong Polytechnic University, Hong Kong 999077, China (e-mail: harry.qin@polyu.edu.hk).

R. Fang is with the J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL 32611, USA (e-mail: ruogu.fang@bme.ufl.edu).

D. D. Feng is with the Biomedical and Multimedia Information Technology Research Group, School of Information Technologies, The University of Sydney, Sydney, NSW 2006, Australia (e-mail: dagan.feng@sydney.edu.au).

I. INTRODUCTION

DIABETIC RETINOPATHY (DR), also known as diabetic eye disease, refers to the progressive retinal damage occurring in persons suffering from diabetes. This disease causes narrowing of the small retinal vessels and often show no symptoms in its early stages. However, it can progress rapidly and cause vision loss through several pathways. Ophthalmologists can effectively examine diabetic patients by checking for retinal lesions, microaneurysms, and abnormal/fragile blood vessels. However, owing to the high prevalence of diabetes and shortage of human experts, screening programs are costly and time-consuming for clinics. A reliable automatic retinal image examination method would significantly reduce the workload of ophthalmologists and facilitate a more effective screening process [1]. The motivation of this paper is to develop an automatic and effective retinal vessel extraction from fundus images. We enhance both the low-level feature estimation process and the subsequent vessel extraction process by using an efficient minimum spanning superpixel tree (MSST) detector that can accurately capture both global and local structure features for retinal vessel segmentation. Furthermore, our work could be extended to human identification (ID), as the retinal vessel structure is unique to each person. We believe that in real-life applications, some biometrical features would be more applicable than others [2] in certain cases. For example, in Zhou et al. [3]'s work, human identification was achieved based on a new biometrical method: sclera recognition, which was based on a Gabor wavelet-based sclera vessel patterns detection. And their experimental results have shown comparable recognizing rate to iris recognition in visible wavelengths.

Numerous methods have been proposed for vessel extraction. These methods can be broadly categorized into three types: the machine-learning-based classifications, the tracking-based approaches, and the enhancement/thresholding methods. Among these methods, machine-learning-based classifications usually obtain the best segmentation accuracy, and such strategy can provide satisfactory results for healthy retinal images [4], [5]. However, these methods usually require the ground-truth segmentations for model training, however these ground truth image labels are difficult to be generated in practice. Meanwhile, another problem for these methods is that they often require a time-consuming re-training process when applied to a new image set. On the other hand, most of the tracking-based approaches are interactive, requiring a lot of user interactions [6] which will significantly increase the workload of ophthalmologists. Although [7]'s as-automated-

as-possible algorithm needs limited initial inputs and user interaction, the method requires a given scale tube diameter for the processing. Thus, our technique pursues a fast, robust, automatic and accurate vessel segmentation.

In general, traditional vessel extraction algorithms (such as the thresholding methods) consist of four steps. First, a preprocessing step aims to remove illumination noise and enhance specific features in the original retinal images. Second, the “vesselness” (likelihood of being a vessel) of each image pixel is computed using approaches such as matched filtering [8] and Hessian-based filtering [9]. Third, vessel/non-vessel areas are classified on the basis of the vesselness measurement; in this study, a threshold is used for the classification. Finally, false vessel pixels are removed by examining certain properties, such as connectivity [10]. Kirbas and Quek [11] have categorized vessel segmentations into six major types: (1) pattern recognitions, (2) model-based methods, (3) tracking-based methods, (4) AI-based methods, (5) neural network based methods, and (6) tube-like object detections.

The concept of a matched filter for improve vessel detection in eyes was first adopted in [8]. In their method, 12 different templates are constructed on the basis of a two dimensional design matched filter (MF), which uses intersectional vessel information and minimum changes in distance. Some revised version of matched filter with improved effects were introduced in [10], [12], [13]. [12] presented a method using threshold for area-based vessel classification. Al-Rawi [13] improved the response and performance of a Gaussian MF by proposing a general-purpose optimization technique to determine better filter parameters. [10] introduced a revised matched filter for enhancing the vessel detection processing in the case of non-vessel structures. However, a major drawback of MF-based approaches is the ambiguous assumption that the intensity distribution of a vessel cross section obeys a Gaussian function, which is not always the case.

The line operator in [14] is a useful method for detecting linear structures in mammographic photos. Given a target pixel and multiple lines passing through the pixel, the method selects the largest gray level value as the line strength of the target pixel. Ricci and Perfetti [15] introduced the basic line operator to detect retinal vessels along all directions, as vessels can have different orientations, widths, and intensities. Nguyen et al. [16] generalized the basic line operator and combined various scales to produce a satisfactory segmentation result. Although the basic line operator is effective for vessels with a central reflex (a bright strip that is sometimes present along the center of wider vessels), we find that it has two critical limitations. The first is its ineffectiveness in detecting small vessels owing to illumination noise and extremely small gray level variance in the neighborhood of small vessels. The line operator mainly distinguishes vessels from the background according to their gray levels. However, the gray levels of thin and low-contrast vessels are similar to those of the background. Therefore, it is extremely difficult to set a threshold for discrimination. The second limitation is the spurious response to nonvascular structures such as the optic disc boundary and pathologic area, which commonly have a strong boundary and tend to be falsely detected as vessel boundaries by the line operator.

Supervised methods for retinal segmentation are discussed in the paper written by Singh and Kaur [17]. First of all, the supervised method is trained with the marked dataset. Then, the method is applied to classify retinal pixels. [18], [19] gave the overview of some blood vessel segmentation methods. Supervised classifications are discussed in [20]. Fraz et al. [21] discussed both the unsupervised and supervised methods for blood vessel detections. Fraz et al. [21] introduced the supervised methods in detail. But supervised methods require manually segmented results as a training set. They will be inconvenient without the training set. Recently, there are several new methods to handle retinal vessel extraction. Zhang et al. [22] proposed filtering-based method using 3D rotation. They have applied a left-invariant rotating derivative frame, and a locally adaptive derivative frame. Annunziata et al. [23] presented a vessel extraction using unsupervised learning. An image inpainting method is applied to enhance the vessel detection rate during enhancement processing. Khalaf et al. [24] proposed a new formulation of deep Convolutional Neural Networks that allows simple and accurate segmentation of the retinal vessels using classification. Wang and Jiang [25] investigated various color channels combination to improve vessel segmentation with fine color fusion method.

In addition, there are some novel image segmentation methods. Song [26] applied regularised gradient flux flows for an accurate extraction of object boundary and contour, which not only reduce noise information but also keep nice edge information. Zhang et al. [27] presented a method to estimate bias field and segment images with intensity inhomogeneity at the same time. [28] introduced a level set approach to segment the images with intensity inhomogeneity. Experimental results on different types of images have shown the effectiveness of the method. Funke et al. [29] proposed a learning-based approach for enhancing objects segmentation with the information of boundary cues. The result is favorable when applying to cell segmentation. It has a certain reference to vessel segmentation. These new methods are valuable for vessel segmentation. Many existing methods for vessel segmentation use pixel grid as underlying representation. Felzenszwalb and Huttenlocher [30] used the image grid with local neighborhood pixels, and distinguish value difference in image intensities for neighbors. For vessel segmentation, using pixel grid is not accurate enough. We believe that a pixel form is not good enough for vessel structures, whereas superpixels provide a more natural and efficient way to compute local image features and distinguish vessels from other retinal structures. The superpixel method, which is becoming increasingly common in the field of image processing, can group pixels using the degree of feature similarity, thereby allowing for redundant information about the image to be acquired. This method fully utilizes spatial information and thus provides excellent anti-noise performance. In addition, it can retain the edge information of the original image while enhancing the local consistency. Xu et al. [31] combined superpixel and bottom-hat filtering to segment retinal vessel, but the characteristic of superpixel isn't utilized when using bottom-hat filtering. We need a superpixel-based method instead of pixel-based method to detect vessel superpixels.

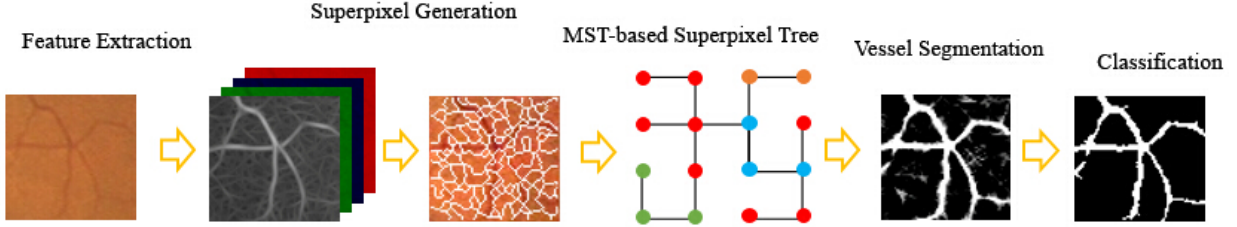


Fig. 1: Overview of the proposed minimum spanning superpixel-based tree detector for retinal vessels.

In this paper, we apply the simple but effective linear iterative clustering (SLIC) algorithm to segment a retinal image into a set of superpixels using a feature combination of spatial distance, color distance, and texture distance. The features are computed for two aims. Firstly, features are computed within a superpixel to determine the probability that the superpixel is part of a vessel. Secondly, the sets divide the input image into patches by distributing each pixel to the nearest primitive. We construct the minimum-spanning tree structure based on these superpixels, called minimum spanning superpixel tree (MSST). Such structure-preserving MSST-based detector is employed to evaluate the vesselness of each superpixel and determine pixels in a patch as either vessel pixels or non-vessel pixels. Finally, segmentation results are refined further by applying a path-opening processing to reduce negative feedback for detection and to avoid non-vascular structures. An overview of the proposed method is shown in Fig. 1. To evaluate our method, we constructed a database of manually labeled images. The dataset is referred as NIVE dataset. The NIVE database can be downloaded at <http://www.mediafire.com/download/t9omausqtf5rfar/NIVE.rar>. It consists of totally 40 fundus images. For comparison, our method is also evaluated using the DRIVE and STARE databases. The results indicate that despite its simplicity, the proposed approach matches or outperforms state-of-the-art methods. The main contribution of this paper is a novel superpixel-based vessel tree detector framework with the following unique characteristics:

- It combines the superpixel and minimum spanning tree to segment retinal vessel.
- It effectively addresses the detection of narrow and low-contrast vessels, preventing influences from other retinal structures.
- It achieves satisfactory segmentation performance in comparison with state-of-the-art methods. (iv) The framework could be extended to human identification (ID), as the retinal vessel structure is unique to each person.

II. METHODS

A. Preprocessing

Fundus images have a special feature of high contrast in the field of view (FOV), false vessel detection may occur around the edge area of the retina image. Therefore, a preprocessing algorithm is executed to remove noise information preparing for further processing the following manipulations as shown in Fig. 2. [32] introduced the key facts and contributions

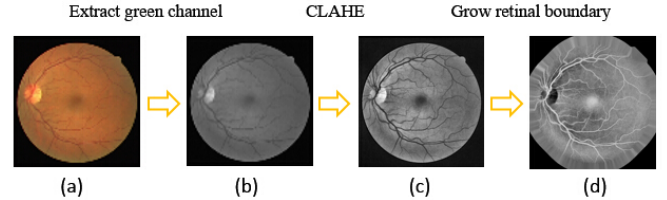


Fig. 2: Step-by-step illustration of preprocessing performed. (a) Input fundus image. (b) Extracted green channel. (c) Local contrast enhancement using CLAHE. (d) Retinal boundary growth using germinating.

for image enhancement using different preprocessing techniques. Our preprocessing step includes contrast enhancement and retinal boundary growth. The CLAHE (contrast-limited adaptive histogram equalization) algorithm could generate image with the effects of local contrast enhancement shown in Fig. 2(c). Therefore, in this study, we adopt CLAHE through dividing our input images with $8 \times 8 = 64$ areas. In addition, to further eliminate the drawbacks generated around camera aperture border through wavelet transformation, we have applied a boundary germinating method using iteration-based computing to expand the concerned area (see Fig. 2(d)). More related information of the processing is discussed in [4].

B. Feature Extraction

(1) Illumination Layer

Inspired by the work in [33], we obtain two layers from an original retinal image I , and one layer is smoother than the other layer. The smoother layer is referred to as the illumination layer L_1 , and the other layer is referred to as the reflectance layer L_2 . The illumination layer is used for subsequent superpixel-based segmentation. The intrinsic image model is formulated as:

$$I = L_1 + L_2 \quad (1)$$

The extraction process utilizes a gradient-based sparsity prior on reconstructed layers and an additional constraint on the smooth layer. To address layer separating, a probability-based method is adopted to obtain most-likely representation of the original image; further details can be found in [33]. The final results are shown in Fig. 3(b) and Fig. 3(c), where the reflectance layer appears to be very bright, whereas some areas are poorly illuminated and unevenly distributed in the FOV. In contrast, the illumination layer is smoother and more consistent, guaranteeing good image quality.

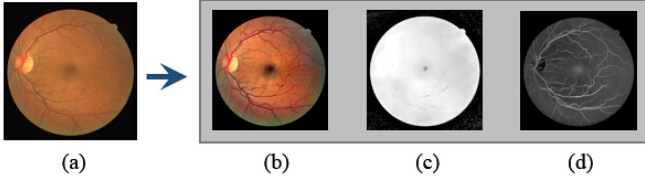


Fig. 3: Feature extraction. (a) Original input image. (b) Illumination layer. (c) Reflectance layer. (d) Texture layer.

(2) Texture Layer

Textures in images play an essential place in many tools like remote sensing, environmental monitoring, and medical image processing. They usually have a natural sequence in the orientation and the multi-narrow-band frequency information. This represents some fundamental features of visual appearances and is crucial in the perception of color. Retinal image textures provide us with spatial color and intensity information, and such information can be applied for superpixel-based vessel segmentation. In this study, Gabor wavelets are employed as texture layer detectors. These wavelets have orientation selectivity, multi-scale properties, a linear stage, and nice localization in spatial and frequency fields, making them proper for texture analyzing. In this section, we utilize the expressions in [34] for the representation, and the Gabor transform is defined in the following as:

$$T_{\psi}(\mathbf{b}, \theta, a) = C_{\psi}^{-1/2} a^{-1} \int \psi^*(a^{-1} r_{-\theta}(\mathbf{x} - \mathbf{b})) I(\mathbf{x}) d^2 \mathbf{x} \quad (2)$$

$$\psi(\mathbf{x}) = e^{j\mathbf{k}_0 \cdot \mathbf{x}} e^{-\mathbf{x} A \mathbf{x} / 2} \quad (3)$$

where, I is the image to be processed, $A = \text{diag}[\eta^{-1/2}, 1]$ denotes the 2 by 2 matrix that the off-diagonal entries are zero, \mathbf{k}_0 denotes a vector for frequency. ψ , ψ^* , a , and \mathbf{b} are the given 2D Gabor variables, more details of the wavelet expressions and processing could be found in [34].

To obtain a good initial response from vessels orientation in different directions, the θ value of the filter is varied from 0 to 170 to generate high-quality feedback. Through all the orientations, the maximum wavelet feedback for every pixel is computed as:

$$M_{\psi}(\mathbf{b}, a) = \max_{\theta} |T_{\psi}(\mathbf{b}, \theta, a)| \quad (4)$$

The thickness of blood vessels in the retina varies from $50\mu m$ to $200\mu m$, with a median of $60\mu m$. Here, a is set to 1.5 experimentally for the high-quality detection of easy-missing vessels, and multiple scales wavelet could be utilized for multi-scale vessel detection. Fig. 4 gives the superpixel blocking result based on illumination layer and on conventional RGB layer. It's obvious that illumination and texture information help generate superpixels adaptively so that they can adhere well to boundaries.

C. Minimum Spanning Superpixel Tree Detector

Superpixels are usually formed by over-segmentation. In this work, we take illumination and texture layer information, and employ the SLIC algorithm [35] to partition the original

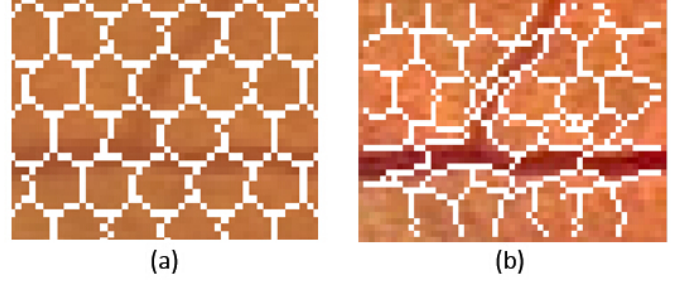


Fig. 4: Comparison of the superpixel blocking result based on conventional RGB layer and illumination layer, respectively. (a) Conventional RGB layer. (b) Illumination layer.

image as a group of superpixels. SLIC begins by sampling starting cluster centers on a regular grid spaced $S = \sqrt{N/K}$ pixels apart. Next, to prevent putting these cluster centers at edges, they are moved to seed locations relate to the lowest gradient position through 3 by 3 neighborhood. Then, the clustering process is iteratively applied to give every pixel to the nearest cluster center according to color and spatial relation. At the end of this process, pixel connectivity is enhanced by relabeling disjoint segments using labels from the largest neighboring cluster. Compared with other superpixel methods, SLIC has low computational complexity and excellent boundary adherence. In addition, SLIC is easy to use because it needs only one parameter to produce the target superpixel number.

Generally, SLIC clusters image pixels using color, texture, and image plane spaces according to a weighted distance metric. In this study, we measure the distance between a pixel i and the SLIC cluster center C_k in the illumination and texture layers, respectively. In addition, we take the spatial distance into consideration, as the pixel position $[x, y]^T$ may vary for different sized images. Just applying the 6D Euclidean distance in the space generated by combining information from all three layers may cause inconsistencies during clustering superpixels of different sizes. For large superpixels, spatial distances outweigh illumination information; more essentialness is given to the spatial information than to color contents. Such operations will generation superpixels which do not conform very nice to edges. While in the case of smaller superpixels, the effects are more accurate. We try to join the two distances as one metric and assign different weights to the three layers. The overall distance measure D is defined as follows:

$$d_{rgb} = \sqrt{(R_{k_1} - R_{k_2})^2 + (G_{k_1} - G_{k_2})^2 + (B_{k_1} - B_{k_2})^2} \quad (5)$$

$$d_t = \sqrt{(g_{k_1} - g_{k_2})^2} \quad (6)$$

$$d_s = \sqrt{(x_{k_1} - x_{k_2})^2 + (y_{k_1} - y_{k_2})^2} \quad (7)$$

$$D = \sqrt{m_{rgb} d_{rgb}^2 + m_t d_t^2 + m_s d_s^2} \quad (8)$$

where $R, G, B, [x_{k_1}, y_{k_1}]^T, [x_{k_2}, y_{k_2}]^T$ denote the red, green, and blue channels of the illumination layer, the pixel position, and the coordinates of the cluster center, respectively. Further, d_{rgb} , d_t , and d_s represent the illumination, texture, and spatial proximities, respectively, and m_{rgb} , m_t , and m_s

are the weights of the illumination distance, texture distance, and spatial distance, respectively. Illumination and texture information both help generate superpixels adaptively so that they can adhere well to boundaries. The information represents some basic characteristics that are crucial in the perception of color. To fully utilize the illumination and texture feature, we set the weight proportional to the difference between pixel and the coordinates of the cluster center.

Fig. 5 shows some superpixel-based fragments of a retinal image (DRIVE, 01). It is clear that superpixel blocks adhere well to vessel boundaries when we combine the three layers of information, especially for some small vessels. Fundus images are known to suffer from uneven illumination noise, which can be eliminated by clustering pixels into superpixels. This in turn would enhance both big and small vessels, because inherent homogeneities are spread across their neighbors in the process of k -means clustering. When a set of superpixels is generated through k -means clustering with the distance measure defined above, the superpixel attributes can be viewed as a set of local nearby homogenous pixels. Based on our observation, vessel superpixels in fundus images are characterized as slender edge areas, which may not be comprehensively preserved by general enhancement methods, such as the well-known bilateral filter, because they deal with slender areas based on a locally averaging operator that tends to smooth out the small homogenous vessel regions. Moreover, they often suffer from severe deviation from the original sharp edges, because the geometric image structures are ignored.

To evaluate the vesselness of superpixels, a new edge-preserving tree detector is employed to determine weighted average. Compared with traditional enhancements, our vessel tree detector distinguishes small connected components (details) from large connected components (major structures) in a non-local manner, achieving impressive results for slender areas such as retinal vessels in fundus images. As shown in Fig. 6, the tree detector utilizes a minimum spanning tree to deal with large connected components in superpixel graph. Each superpixel represents a graph node, and superpixel-based feature differences provide edge weights between the nearest neighboring superpixels. An MSST is formed by repeatedly removing edges with large weights, such that any two close but dissimilar superpixels are automatically dragged apart.

To optimize a graph, there are many frequently used methods like normalized-cut(N-cut), max-flow min-cut, minimum spanning tree and so on. By comparing these methods, we found that MST is the most suitable one in this situation.

- 1) The N-cut method measures both the total dissimilarity between the different groups as well as the total similarity within the groups. It is a NP-complete problem and has high complexity.
- 2) The graph-cut method based on max-flow min-cut theorem utilizes the property that the smallest total weight of the edges which if removed would disconnect the source from the sink. It can get the global optimal solution efficiently, and it has good noise immunity. But we have to choose the pixel inside and outside the target as seed point manually, which limit its application in image segmentation.

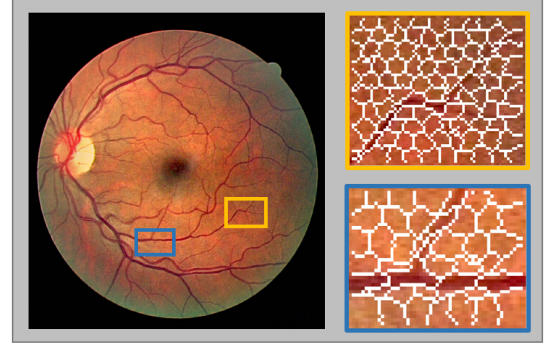


Fig. 5: Patches showing the superpixel region after clustering.

- 3) The MST method can reserve the details inside the boundaries of vessels. Its self-adaptability when searching for minimum weight helps to get the global features, which show the perception of color that is similar to human's eye. Also, it has low complexity. So we choose MST to optimize a graph. Considering the important property that retinal vessels appear to be similar to vascular tree structures, we can easily connect the vessel superpixel sets by constructing the MSST.

Unlike other related algorithms, our approach builds the MSST in the superpixel plane instead of the image pixel grid. Given an image I , let $G = (V, E)$ be an undirected graph in which V (the set of vertices) is the set of superpixels in the retinal image, and every two superpixel nodes p, q are connected by an edge $e_{ij} \in E$, to which a weight is assigned:

$$w((p, q)) = |I(p) - I(q)| \quad (9)$$

where $I(p)$ is the mean value of the image intensity at superpixel p . A tree that spans a component $C \in V$ and has the minimum total weight is the MSST.

With the MST-based superpixel information, we define the similarity between two nodes p and q in the MSST as:

$$S(p, q) = S(q, p) = \exp\left(-\frac{D(p, q)}{\sigma}\right) \quad (10)$$

where σ is a constant to adjust the similarity and $D(p, q)$ denotes the distance between p and q in the MSST. If the two pixels p and q are close, i.e., if $D(p, q)$ has a small value, p and q are similar, and vice versa. Thus, the edge-preserving tree detector can be formulated as:

$$I(p) = \sum_{q \in \Omega} S(p, q) I(q) = \sum_{q \in \Omega} \exp\left(-\frac{D(p, q)}{\sigma}\right) I(q) \quad (11)$$

where Ω is the set of all nodes in the entire superpixel graph. This equation indicates that each pixel receives support weights not within the local user-specified window but from all the other superpixel nodes in the MSST. Therefore, the MSST distinguishes small connected components (details) from large connected components (major structures) in a non-local manner. Thus, the tree filter is able to deal with slender regions such as retinal vessels in fundus images. As shown in Fig. 7, vessel merging and vessel disconnection are found in Marín et al. [36] and Soares et al. [4], while the proposed

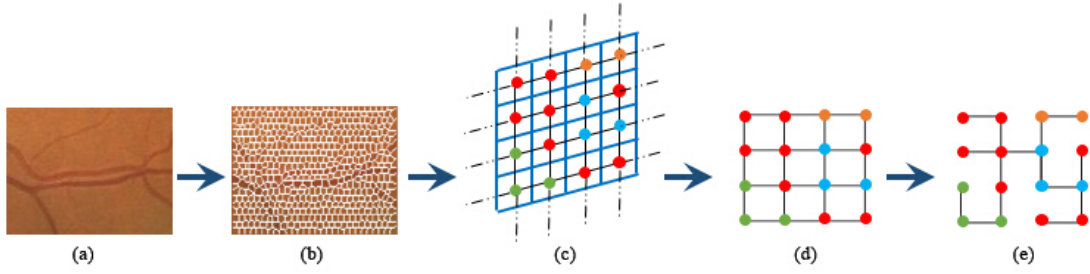


Fig. 6: Illustration of superpixel tree generation. (a) Original image. (b) After applying SLIC. (c) Illustration of evenly located superpixel nodes. (d) Superpixel graph. (e) Minimum spanning tree based on superpixel graph.

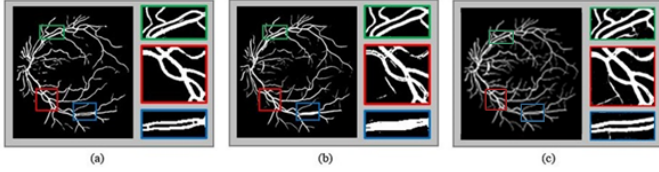


Fig. 7: Comparison of final vessel response and several magnified regions showing improved results. (a) The response image of Marín et al. [36]. (b) The response image of Soares et al. [4]. (c) Result by our tree detector.

vessel tree detector can effectively preserve high-contrast, fine-scale details, thereby achieving impressive results.

A limitation of the proposed vessel tree detector is its high computational complexity because it involves a search procedure as well as computation of tree distances among all pixel pairs. However, by using the proposed MSST non-local aggregation, the vessel tree detector can be recursively implemented and achieve very fast computation. Here, we provide a brief introduction to the proposed MSST non-local aggregation algorithm. A MSST is built using Kruskals algorithm by treating the superpixel graph as a standard 4-connected grid, the nodes of which are the partitioned superpixels. After selecting a node as the root node (v_4), an iterative aggregation algorithm is employed for the root node to receive supports from any other node, as shown in Fig. 8. Let $I_A(v)$ denote the aggregated support values and $P(v_c)$ denote the parent of node v_c ; then, for each node $v \in V$,

$$I_A(v) = I(v) + \sum_{P(v_c=v)} S(v, v_c) \cdot I_A(v_c) \quad (12)$$

Similarly, as shown in Fig. 9, the aggregated support value for each node consists of two parts, the node itself and the supports it receives from its subtrees, given by

$$I_A(v) = I(v) + S(P(v), v) \cdot I_A(P(v)) - S(v, P(v)) \cdot I(v) \\ = S(P(v), v) \cdot I_A(P(v)) + (1 - S^2(v, P(v))) \cdot I_A(v) \quad (13)$$

Thus, the entire non-local aggregation can actually be separated into two steps: (i) aggregation of the original supports from the leaf nodes toward the root node, and (ii) distribution of the supports from the root node toward the leaf nodes. Note that in Eq. (12) and Eq. (13), $S(v, v_c)$, $S(P(v), v)$ and $1 - S^2(v, P(v))$ depend only on the edges of MSST built

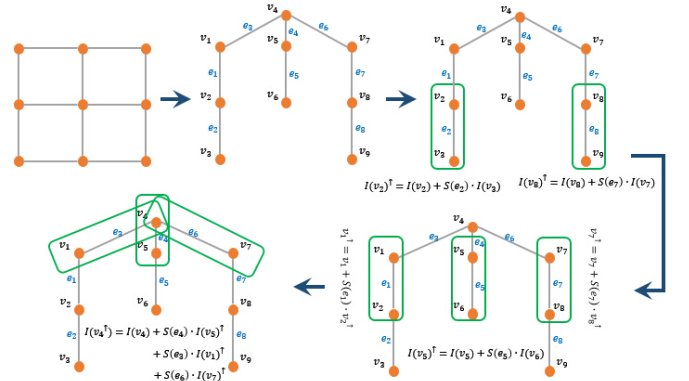


Fig. 8: Step-by-step illustration of upward aggregation process: root node receiving supports from all the other nodes.

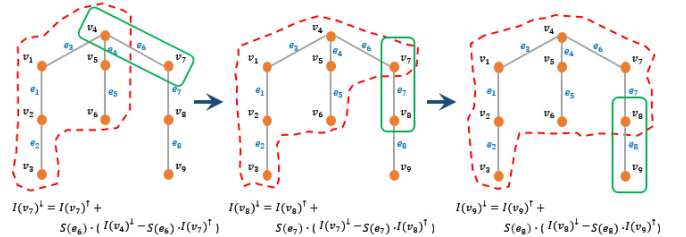


Fig. 9: Aggregated support value for each node includes two parts: the node itself, and supports it receives from its subtrees.

in superpixel graph and can be computed beforehand. Only two addition/subtraction operations and three multiplication operations for one pixel detection are required, therefore the tree detector can achieve real-time running speed for retinal vessel extraction. Fig. 10 compare running times of two implementations. Although they show competitive performance on image size(100 × 100), the implementation with acceleration exhibits significant advantages over that without acceleration when image size is growing to 1000. According to our experiments, the whole superpixel-tree-based filtering with acceleration can process 500 × 500 24-bit RGB-color image in about 4.57 seconds on our CPU, which outperforms state-of-the-art vessel segmentations. We used STARE database to test the runtime of several methods. The average runtime of our method, Zhang et al. [10], Soares et al. [4] and Nguyen et al. [16] are 5s, 10s, about 3m10s, and 3s respectively.

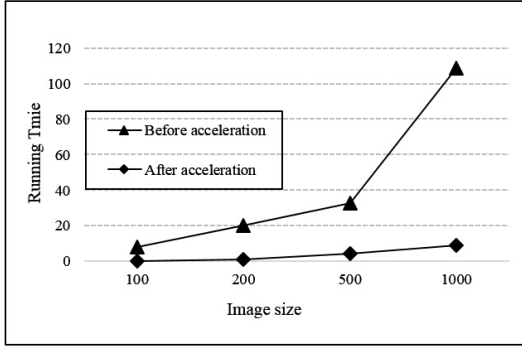


Fig. 10: Comparison of different algorithm implementations in terms of running time. (a) Before acceleration. (b) After acceleration based on upward and downward aggregation.

Combining it with Table I, we can see that our method is efficient and accurate comparing to these methods.

Vessel detection could be further enhanced using techniques of path opening to avoid false extraction. Buckley and Talbot [37] introduced the method of path opening, and it is further enhanced in [38], [39]. Through path opening, we could obtain highly local-based objects' contours with significant contrast difference in the image environments. Here, we use N_{min} to indicate the edge path length, and we use N_k to denote acceptable missing pixels number. When the path meets the requirements of N_{min} and N_k , the algorithm will give largest intensity value to the pixels. For the path filters, the choice of length parameter N_{min} is significant, and it will determine the size of the structure for the elements. Applying the method of path opening processing, undesired information like image noise could be filtered out, while vessels with thin structures could be retained.

III. EXPERIMENTS ON DATASETS

A. DRIVE Dataset

Through a diabetic retinopathy (DR) screening program in Holland, the DRIVE (Digital Retinal Images for Vessel Extraction) dataset is created for doing further researches. The sample size for the DR screening is 453 persons ranging from 31 to 86 years old. There are totally 40 retinal images in the DRIVE dataset, where 7 out of 40 images have different types of pathological changes. The dataset is obtained through using professional retinal image capturing device. The retinal images we downloaded are with the resolution of 565×584 with three channels of color in TIF format, the radius for the FOV is around 270 pixels. Fig. 11(a) and Fig. 11(b) show some example retinal images in this DRIVE dataset. Mainly, the DRIVE dataset is further categorized into two subsets, namely, 20 images to form the training set and 20 images to form the testing set. Under supervision of professional ophthalmologists, three human observers have manually extracted the vessel details from the fundus images for DRIVE dataset, providing a minimum certainty of 70%. Fourteen retinal images in the training set are extracted by the first observer, the other six retinal images are extracted by the second observer. To ensure the accuracy, the retinal images in the testing set are extracted

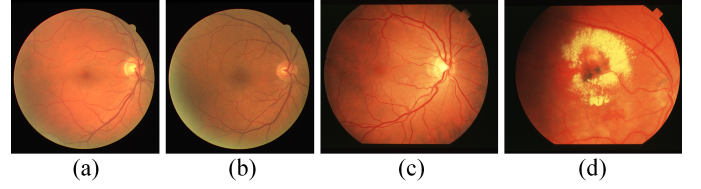


Fig. 11: Sample retinal images in DRIVE dataset and STARE dataset. (a) a fundus image (18_test) in DRIVE. (b) a fundus image (06_test) in DRIVE. (c) a fundus image (im0022) in STARE. (d) a fundus image (im0017) in STARE.

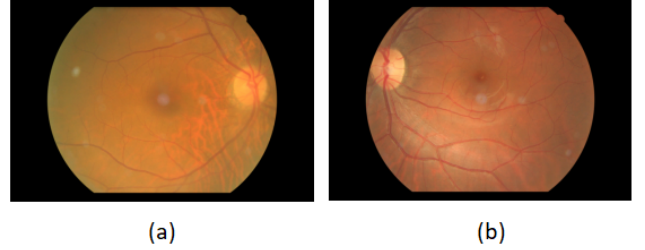


Fig. 12: Sample fundus images from NIVE. (a) a fundus image (im21) in NIVE. (b) a fundus image (im23) in NIVE.

by twice giving testing set A and testing set B. Thirteen retinal images in the testing set A are extracted by the first observer, the other seven retinal images are extracted by the second observer; the third observer works independently to extract the testing set B. Generally, the testing set A and testing set B agree with each other in vessel pixel extraction. In this paper, without losing generality, we apply the testing set A as the benchmark for ground-truth vessel extraction.

B. STARE Dataset

Hoover et al. [12] created the STARE (STRUCTURED Analysis of the RETINA) dataset for research and analysis. There are totally 20 images in this dataset, and half of them have pathological changes. The dataset is also obtained via professional retinal image capturing device, where the FOV is set at 35. The fundus images downloaded are with the resolution of 700×605 with three channels of color in PPM format, the radius for the FOV is around 325 pixels. Fig. 11(c) and (d) show some sample retinal images in this STARE dataset. Two human observers have extracted the vessel details from the fundus images for STARE dataset. The two human observers have slight differences in indication of vessel information, namely, 10.4% of total pixels for observer one and 14.9% of total pixels for observer two. In this paper, we assume to utilize the extraction results from the first human observer as the benchmark for ground-truth vessel extraction for STARE dataset. The whole images are applied in [12] to perform the vessel segmentation, which may cause some problems in extraction. Here, we only apply the area of field of view (FOV) for the processing using our approach.

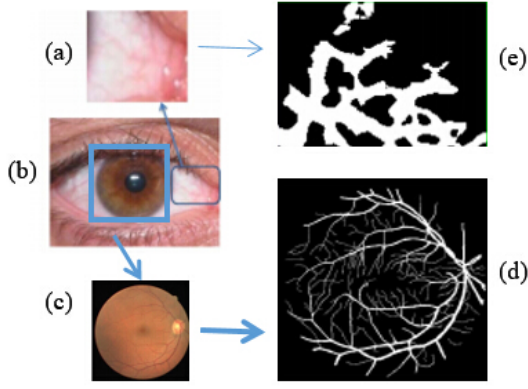


Fig. 13: Comparison of our approach with [3] in terms of human identification. (a) Segmented sclera region. (b) Original sclera color image. (c) Relative retinal image. (d) Segmented retinal vessel. (e) Segmented vessel of sclera region.

C. NIVE Dataset

Through a diabetic retinopathy screening project in Shanghai, the NIVE dataset is created for research study. There are totally 40 retinal images in the NIVE dataset. The Canon CR-DGi is applied to as the high-quality professional fundus image capturing device in the project for taking the retinal images setting at a field of view of 45. The retinal image quality in NIVE is with sufficient resolution allowing analysis of small and thin vessels. The retinal images downloaded in NIVE are with the resolution of 600×400 with three channels of color in TIF format, the radius for the FOV is around 225 pixels. Fig. 12 shows some examples from the NIVE fundus images.

IV. EXPERIMENTAL RESULTS AND EVALUATIONS

We have applied our approach in the above mentioned three datasets for the evaluation of our method, namely, the DRIVE, STARE, and NIVE dataset. Comparisons have been made with the state-of-the-art methods for all the three datasets, also with the ground-truth vessel extraction defined in the last section for the DRIVE and STARE datasets. We have compared our approach with Zhou et al. [3] in terms of human identification. As shown in Fig. 13, only the sclera vessel structure is extracted from original eye image in [3]. In contrast, our method acquires the whole retinal vessel patterns, utilizing which the human identification process can achieve higher accuracy. Moreover, with superpixel-based graph structure employed, our method enables combination with other biometrics, such as iris recognition to perform multi-modal biometrics. Overall, the retinal vessel segmentation exhibits promising prospects for positive human identification.

A. Quality Assessment

Qualitative assessment for vessel extraction quality performs an essential role in retinal image segmentation. Usually, the overall vessel segmentation quality for each algorithm is evaluated according to conventional evaluation metrics mainly related to different kinds of detection rates [40]. However, the conventional evaluation metric suffers from a drawback [40],

TABLE I: Performance for Vessel Extraction (STARE)

Approaches	C	A	L	CAL Score
Manual	0.9998	0.8476	0.8600	0.7288
Hoover et al. [12]	0.9998	0.6866	0.7299	0.5010
Soares et al. [4]	0.9942	0.7494	0.7690	0.5729
Nguyen et al. [16]	0.9914	0.7903	0.8082	0.6332
Zhang et al. [10]	0.9986	0.7828	0.8007	0.6259
The Proposed	0.9996	0.8242	0.8382	0.6906

TABLE III: Performance for Vessel Extraction (DRIVE)

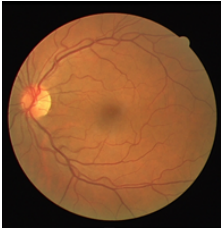





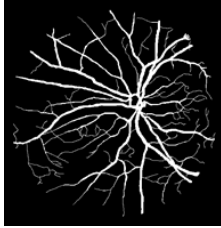

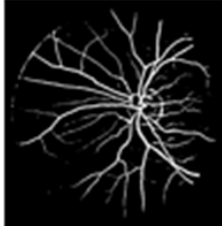



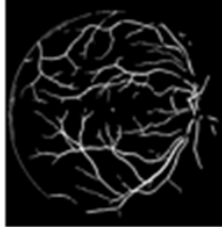
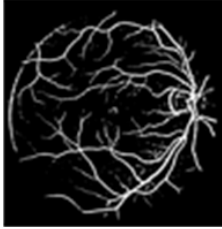
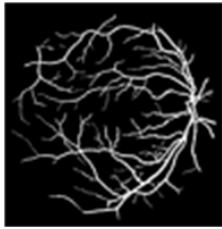
Approaches	C	A	L	CAL Score
Manual	1	0.9398	0.9347	0.8784
Marín et al. [36]	0.9990	0.8327	0.8314	0.6916
Soares et al. [4]	0.9952	0.8920	0.8889	0.7891
Nguyen et al. [16]	0.9895	0.8727	0.8687	0.7502
Zhang et al. [10]	0.9988	0.8097	0.8108	0.6557
The Proposed	0.9998	0.9006	0.8988	0.8092

where the medical scientists usually do not describe the retinal images using exact math as the exact positions for the retinal vessels are hard to be addressed by medical scientists. Manual extraction of blood vessels in images are not fully objective and easy to make mistakes. The retinal vessel segmentations from different experts for even the same fundus image may not be the same. Especially for the thin vessels, the segmentation results from different human experts may look roughly the same but do not match exactly. Thus, if two experts jointly segment the same fundus image, such facts will downgrade the segmentation accuracy to a greater degree than the case in which one of the observers does not label any vessel. Consequently, attempts to produce reference-standard images with high accuracy are worth discussion. In this paper, we hence apply a new method for qualitative evaluation of vessel extraction quality for retinal images. Generally, we have to compute the CAL score [40] to evaluation the retina segmentation quality as: $C \times A \times L$, where, the component C is the vessel connectivity, the component A is the area, and the component L is the length for vessel extraction. The ranges for all the three valuables C , A , and L are between 0 and 1, with 1 as the highest score. The final CAL score is the product of C , A , and L . Gegúndez-Arias et al. [40] has justified that the CAL score offers better quality evaluation considering the facts of human perception in retinal image segmentation.

B. Assessment Using DRIVE Dataset

Some retinal vessel extractions of images in DRIVE dataset are shown in Table II. The input retinal images are shown in the first column, the manual vessel extractions from human expert are shown in the second column. The third column shows the fundus image extraction effects from Zhang et al. [10], and the fourth column shows the fundus image extraction effects using Nguyen et al. [16]. Our retinal image extraction effects are shown in the fifth column. It can be seen clearly from Table II that the results using [10] are easily disconnected, while the extraction results from [16] are

TABLE II: Vessel Extraction Comparison on the DRIVE Dataset

I.D.	Input	Manual	Extraction Effects		
			Zhang et al. [10]	Nguyen et al. [16]	Our Proposed
#1					
#2					
#3					

not sensitive to thin and light blood vessels. However, our approach is able to perform well on the DRIVE dataset and provide good solutions to thin vessels compared to the state-of-the-art methods. Our method could extract more essential thin vessels compared with existing approaches. Table III shows the quality evaluation of vessel extraction on the DRIVE dataset. It can be seen clearly from Table III that our proposed approach offers better CAL score compared with the state-of-the-art methods. For each of the three components, namely, connectivity, area, and length, our approach also provides better results.

C. Assessment Using STARE Dataset

Some retinal vessel extractions of images in STARE dataset are shown in Table IV. The three input retinal images are shown in the first column, the three manual vessel extractions from human expert are shown in the second column. The third column shows the fundus image vessel extraction effects from Hoover et al. [12], and the fourth column shows the fundus image extraction effects using Nguyen et al. [16]. Our retinal image extraction effects are shown in the fifth column. It can be seen clearly from Table IV that the results using Hoover et al. [12] include vessels that are not extracted correctly, while the extraction results from [16] have blood vessels that are connected not very well. Compared to the state-of-the-art methods, our approach is able to perform well on the STARE dataset and provide good solutions to especially thin vessels in addition to thick vessels. Table I shows the quality evaluation

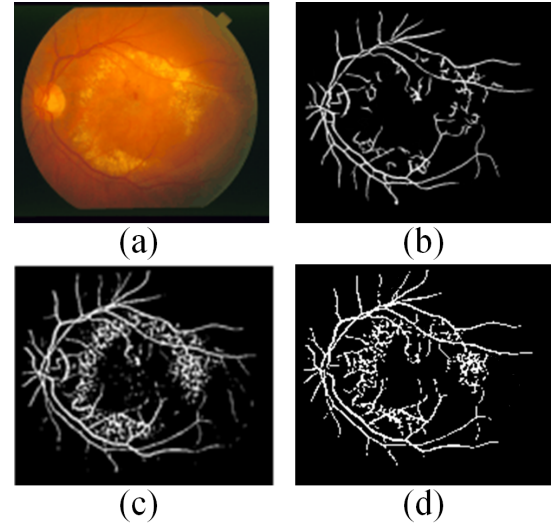


Fig. 14: Vessel extractions for a pathological retinal image in the STARE dataset. (a) Input retinal image. (b) Manual ground-truth segmentation. (c) Extraction using Nguyen et al. [13]. (d) Segmentation using the proposed approach.

of vessel extraction on the STARE dataset. It can be seen clearly from Table I that our approach offers better CAL score compared with the state-of-the-art approaches [4], [10], [12], [16]. For the area and length, our approach provides better results; for the connectivity, the results using our approach are

TABLE IV: Vessel Extraction Comparison on the STARE Dataset

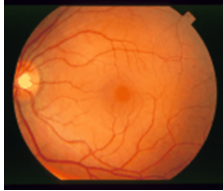
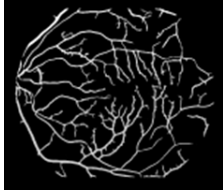
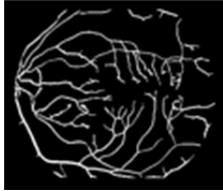

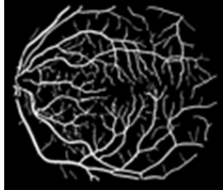
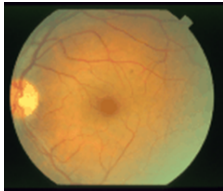
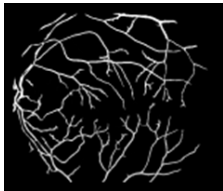
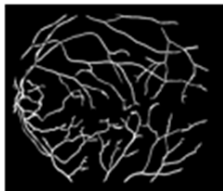
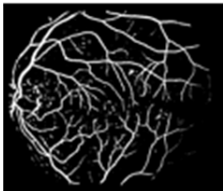
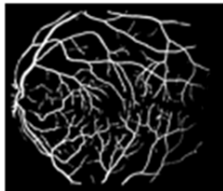
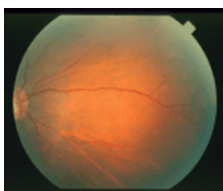
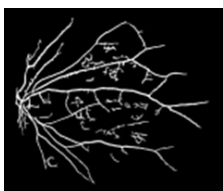
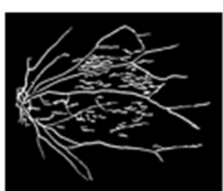
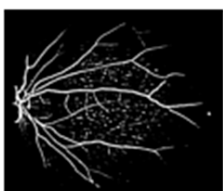
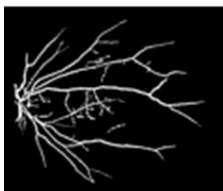
I.D.	Input	Manual	Extraction Effects		
			Hoover et al. [12]	Nguyen et al. [16]	Our Proposed
#1					
					
					

TABLE VI: Performance for Vessel Extraction (NIVE)

Approaches	<i>C</i>	<i>A</i>	<i>L</i>	<i>CAL Score</i>
Zhang et al. [10]	0.9982	0.2882	0.3881	0.1117
Nguyen et al. [16]	0.9989	0.3265	0.4196	0.1369
The Proposed	0.9992	0.3966	0.4679	0.1854

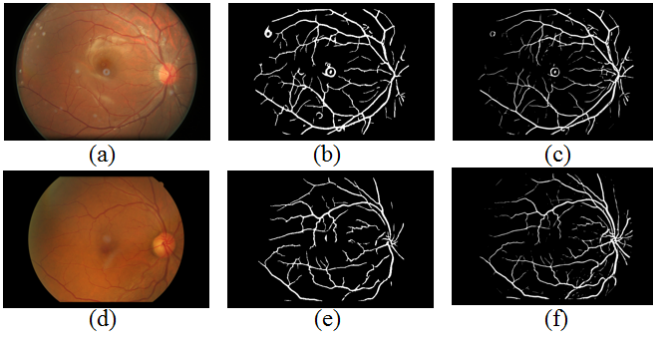


Fig. 15: Comparison of our method with Zhang et al. [10] on segmentation of retinal images from NIVE. (a), (d) show the input retinal images. (b), (e) show the extraction results using [10]. (c), (f) show extraction results using our approach.

comparable to the state-of-the-art method. Another advantage of our method is the performance in the case of lesions. As shown in Fig. 14, there are some hard exudates in the original image, which are highly unfavorable for vessel extraction. The region of exudates not only covers the blood vessels originally exist in image but also degrades the entire image quality. The results from [16] are evidently prone to exudates, while our method filters them out with better quality.

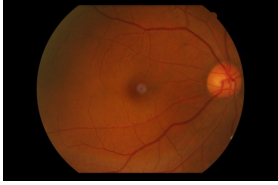
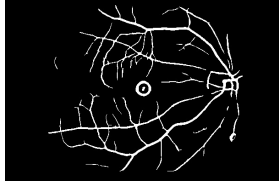
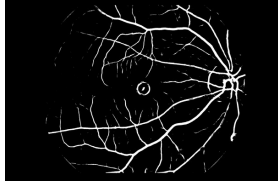
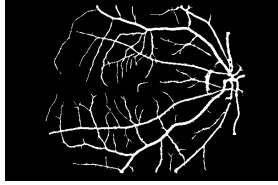
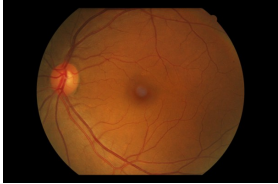
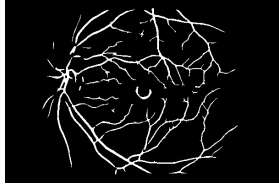
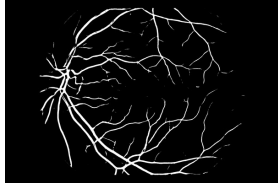
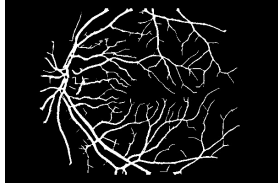
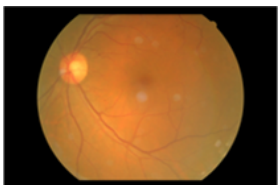
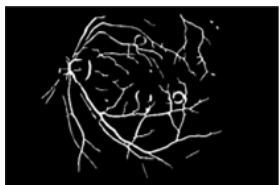
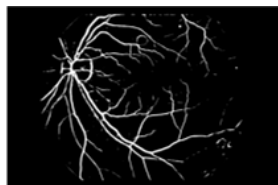

D. Assessment Using NIVE Dataset

To illustrate the effectiveness of our approach more clearly, we have assessed its performance on the NIVE dataset and compared it with other methods. Fig. 15 shows two original NIVE images. In addition, it shows the vessel extraction effects generated by the methods of Zhang et al. [10] as well as those of our proposed method. The main problem with the results of [10] includes the limited ability to handle vessel extraction in the condition of noise information. Moreover, we can see that the segmentation effects of [10] tend to overlook narrow and low-contrast retinal vessels that are actually essential in diabetic retinopathy analytics. In the mean time, our approach is able segment out most of the key vessels based on the input retinal images. Some retinal vessel extractions of images in NIVE dataset are shown in Table V. The three input retinal images are shown in the first column. The second column shows the fundus image vessel extraction effects from Zhang et al. [10], and the third column shows the fundus image extraction effects using Nguyen et al. [16]. Our retinal image extraction effects are shown in the fourth column. It can be seen clearly from Table V that our approach gets favorable results comparing to the other methods. Table VI shows the quality evaluation of vessel extraction on the NIVE dataset. For retinal vessel images which are hard to segment, our method still performs better than these methods.

V. CONCLUSION AND FUTURE WORK

Effective vessel segmentation for retinal images is a critical research domain in medical imaging. Owing to inherent non-uniform illumination artifacts and complicated neighboring pathologies, existing methods have achieved reliable results for wide vessels, but they have failed to adequately segment thin

TABLE V: Vessel Extraction Comparison on the NIVE Dataset

I.D.	Input	Extraction Effects		
		Zhang et al. [10]	Nguyen et al. [16]	Our Proposed
#1				
#2				
#3				

and low-contrast vessels. We introduced a new identification scheme for retinal vessel segmentation using superpixel-based tree structure. Instead of operating on the pixel grid, we regularized the proposed scheme by combining global shape, texture, color, and space information in the superpixel graph. We refined our results further by employing an MSST-based vessel tree detector to evaluate the vesselness for each superpixel. Furthermore, path-opening filters are employed to enhance the effectiveness for more accurate extraction of retinal blood vessels avoiding unnecessary non-vascular influences. We further will work to 3D images, and investigate techniques for classifying different categories of retina-related disease and avoiding inaccurate placement of vessel structures for different pathologies. The current SLIC method has limitations like time-consuming and unsmooth boundary. As essential parts of our future work, we will study the latest efficient segmentations for much smoother boundary, including accurate semantic segmentation [41], improved fast SLIC [42] and segmentation using deep learning [43], [44]. GPU parallelism will also be incorporated for efficient processing.

REFERENCES

- [1] B. Yin, H. Li, B. Sheng, X. Hou, Y. Chen, W. Wu, P. Li, R. Shen, Y. Bao, and W. Jia, "Vessel extraction from non-fluorescein fundus images using orientation-aware detector," *Medical Image Analysis*, vol. 26, no. 1, pp. 232–242, 2015.
- [2] J. Wang, C. Lu, M. Wang, P. Li, S. Yan, and X. Hu, "Robust face recognition via adaptive sparse representation," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2368–2378, 2014.
- [3] Z. Zhou, E. Y. Du, N. L. Thomas, and E. J. Delp, "A new human identification method: Sclera recognition," *IEEE Trans. Syst., Man, Cybern. A*, vol. 42, no. 3, pp. 571–583, 2012.
- [4] J. V. Soares, J. J. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree, "Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification," *IEEE Trans. Med. Imag.*, vol. 25, no. 9, pp. 1214–1222, 2006.
- [5] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, 2004.
- [6] L. Wang, V. Kallem, M. Bansal, J. Eledath, H. Sawhney, K. Karp, D. J. Pearson, M. D. Mills, G. E. Quinn, and R. A. Stone, "Interactive retinal vessel extraction by integrating vessel tracing and graph search," in *MICCAI*, 2013, pp. 567–574.
- [7] T. Behrens, K. Rohr, and H. S. Stiehl, "Robust segmentation of tubular structures in 3-D medical images by parametric object detection and tracking," *IEEE Trans. Syst., Man, Cybern. B*, vol. 33, no. 4, 2003.
- [8] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Trans. Med. Imag.*, vol. 8, no. 3, 1989.
- [9] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *MICCAI*, 1998.
- [10] B. Zhang, L. Zhang, L. Zhang, and F. Karray, "Retinal vessel extraction by matched filter with first-order derivative of Gaussian," *Computers in Biology and Medicine*, vol. 40, no. 4, pp. 438–445, 2010.
- [11] C. Kirbas and F. Quek, "A review of vessel extraction techniques and algorithms," *ACM Computing Surveys*, vol. 36, no. 2, pp. 81–121, 2004.
- [12] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 203–210, 2000.
- [13] M. Al-Rawi, M. Qutaishat, and M. Arrar, "An improved matched filter for blood vessel detection of digital retinal images," *Computers in Biology and Medicine*, vol. 37, no. 2, pp. 262–267, 2007.
- [14] R. Zwigglelaar, T. C. Parr, and C. J. Taylor, "Finding orientated line patterns in digital mammographic images," in *BMVC*, 1999, pp. 1–10.
- [15] E. Ricci and R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," *IEEE Trans. Med. Imag.*, vol. 26, no. 10, pp. 1357–1365, 2007.
- [16] U. T. Nguyen, A. Bhuiyan, L. A. Park, and K. Ramamohanarao, "An effective retinal blood vessel segmentation method using multi-scale line detection," *Pattern Recognition*, vol. 46, no. 3, pp. 703–715, 2013.
- [17] N. Singh and L. Kaur, "A survey on blood vessel segmentation methods in retinal images," in *EDCAV*, 2015, pp. 23–28.
- [18] M. I. Khan, H. Shaikh, and A. M. Mansuri, "A review of retinal vessel segmentation techniques and algorithms," *International Journal of Computer Technology & Applications*, vol. 02, no. 05, 2011.
- [19] H. Ocbagabir, I. Hameed, S. Abdulmalik, and D. Barkana Buket, "A novel vessel segmentation algorithm in color images of the retina," in *LISAT*, 2013, pp. 1–6.

- [20] E. James and A. Francisco, "On supervised methods for segmentation of blood vessels in ocular fundus images," *Advances in Image and Video Processing*, vol. 1, no. 1, pp. 21–37, 2013.
- [21] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "Blood vessel segmentation methodologies in retinal images – A survey," *Comput. Meth. Programs Biomed.*, vol. 108, no. 1, pp. 407–433, 2012.
- [22] J. Zhang, B. Dashtbozorg, E. Bekkers, J. P. W. Pluim, R. Duits, and B. M. ter Haar Romeny, "Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores," *IEEE Trans. Med. Imag.*, vol. 35, no. 12, pp. 2631–2644, 2016.
- [23] R. Annunziata, A. Garzelli, L. Ballerini, A. Mecocci, and E. Trucco, "Leveraging multiscale hessian-based enhancement with a novel exudate inpainting technique for retinal vessel segmentation," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 4, pp. 1129–1138, 2016.
- [24] A. F. Khalaf, I. A. Yassine, and A. S. Fahmy, "Convolutional neural networks for deep feature learning in retinal vessel segmentation," in *IEEE ICIP*, 2016, pp. 385–388.
- [25] X. Wang and X. Jiang, "Enhancing retinal vessel segmentation by color fusion," in *IEEE ICASSP*, 2017, pp. 891–895.
- [26] H. Song, "Active contours driven by regularised gradient flux flows for image segmentation," *Electronics Letters*, vol. 50, no. 14, 2014.
- [27] K. Zhang, Q. Liu, H. Song, and X. Li, "A variational approach to simultaneous image segmentation and bias correction," *IEEE Trans. Cybern.*, vol. 45, no. 8, pp. 1426–1437, 2015.
- [28] K. Zhang, L. Zhang, K. M. Lam, and D. Zhang, "A level set approach to image segmentation with intensity inhomogeneity," *IEEE Trans. Cybern.*, vol. 46, no. 2, pp. 546–557, 2015.
- [29] J. Funke, F. A. Hamprecht, and C. Zhang, "Learning to segment: Training hierarchical segmentation under a topological loss," in *MICCAI*, 2015, pp. 268–275.
- [30] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *IJCV*, vol. 59, no. 2, 2004.
- [31] Y. Xu, J. Liu, S. Lin, D. Xu, C. Y. Cheung, T. Aung, and T. Y. Wong, "Efficient optic cup detection from intra-image learning with retinal structure priors," in *MICCAI*, 2012, pp. 58–65.
- [32] T. Walter and J.-C. Klein, "Segmentation of color fundus images of the human retina: Detection of the optic disc and the vascular tree using morphological techniques," in *Medical Data Analysis*, Heidelberg, Germany: Springer, 2001.
- [33] Y. Li and M. S. Brown, "Single image layer separation using relative smoothness," in *IEEE CVPR*, 2014, pp. 2752–2759.
- [34] A. Arnéodo, N. Decoster, and S. Roux, "A wavelet-based method for multifractal image analysis. I. Methodology and test applications on isotropic and anisotropic random rough surfaces," *The European Physical Journal B - Condensed Matter and Complex Systems*, vol. 15, no. 3, pp. 567–600, 2000.
- [35] R. Achanta, A. Shaji, and S. Smith, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, 2012.
- [36] D. Marín, A. Aquino, M. E. Gegúndez-Arias, and J. M. Bravo, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," *IEEE Trans. Med. Imag.*, vol. 30, no. 1, pp. 146–158, 2011.
- [37] M. Buckley and H. Talbot, *Flexible Linear Openings and Closings*. Boston, MA, USA: Springer, 2000, pp. 109–118.
- [38] H. Heijmans, M. Buckley, and H. Talbot, "Path-based morphological openings," in *IEEE ICIP*, vol. 5, 2004, pp. 3085–3088.
- [39] H. Heijmans, M. Buckley, and H. Talbot, "Path openings and closings," *Journal of Mathematical Imaging and Vision*, vol. 22, no. 2-3, 2005.
- [40] M. E. Gegúndez-Arias, A. Aquino, J. M. Bravo, and D. Marín, "A function for quality evaluation of retinal vessel segmentations," *IEEE Trans. Med. Imag.*, vol. 31, no. 2, pp. 231–239, 2012.
- [41] B. Wang, X. Yuan, X. Gao, X. Li, and D. Tao, "A hybrid level set with semantic shape constraint for object segmentation," *IEEE Trans. Cybern.*, vol. 49, no. 5, pp. 1558–1569, 2019.
- [42] X. Li, K. Liu, and Y. Dong, "Superpixel-based foreground extraction with fast adaptive trimaps," *IEEE Trans. Cybern.*, vol. 48, no. 9, pp. 2609–2619, 2018.
- [43] D. Nie, L. Wang, E. Adeli, C. Lao, W. Lin, and D. Shen, "3-D fully convolutional networks for multimodal isointense infant brain image segmentation," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 1123–1136, 2019.
- [44] F. Liao, X. Chen, X. Hu, and S. Song, "Estimation of the volume of the left ventricle from MRI images using deep neural networks," *IEEE Trans. Cybern.*, vol. 49, no. 2, pp. 495–504, 2019.



Bin Sheng received the B.A. degree in English and the B.E. degree in computer science from the Huazhong University of Science and Technology, Wuhan, China, and the M.S. degree in software engineering from the University of Macau, Macau, China, and the Ph.D. degree in computer science from The Chinese University of Hong Kong, Hong Kong, China.

He is currently an Associate Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. His current research interests include virtual reality and computer graphics.



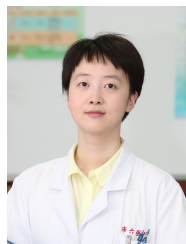
Ping Li received the Ph.D. degree from The Chinese University of Hong Kong, Hong Kong, China.

He is currently an Assistant Professor with the Macau University of Science and Technology, Macau, China. His current research interests include image/video stylization, big data visualization, GPU acceleration, and creative media. He has one image/video processing national invention patent, and has excellent research project reported worldwide by *ACM TechNews*.



Shuangjia Mo is currently pursuing the graduation degree with the Department of Computer Science and Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China.

He is currently with the Visual Media and Data Management Laboratory, Department of Computer Science and Engineering, Shanghai Jiao Tong University. His current research interests include medical image analysis and computer vision.



Huating Li received the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, and Pennington Biomedical Research Center, Baton Rouge, LA, USA.

She is currently an Associate Professor with Shanghai Jiao Tong University affiliated with Sixth People's Hospital and Shanghai Diabetes Institute. Her current research interests include the role of cytokines in the development of fatty liver disease, diabetes, and other obesity-related diseases.

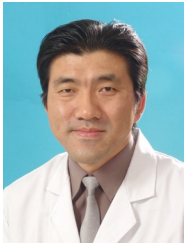
Dr. Li was a recipient of the Shanghai Science and Technology Progress Award.



Xuhong Hou received the Ph.D. degree in medicine from Shanghai Jiao Tong University, Shanghai, China.

She is currently an Associate Chief Technician with Shanghai Jiao Tong University affiliated with Sixth People's Hospital. Her current research interest includes epidemiology of metabolic diseases.

Dr. Hou is a member of the Clinical and Preventive Group of the Chinese Medical Association and a member of the Epidemiology and Preventive Medicine Group of the Chinese Diabetes Society.



Qiang Wu received the M.D. degree in ophthalmology from Shanghai Jiao Tong University, Shanghai, China.

He is currently a Full Professor and the Director with the Department of Ophthalmology, Shanghai Jiao Tong University, Shanghai, China, affiliated with Sixth People's Hospital. His current research interest includes treatment of cataract and diabetic retinopathy diagnosis.



Ruogu Fang received the Ph.D. degree in electrical and computer engineering from Cornell University, Ithaca, NY, USA.

She is currently an Assistant Professor with the J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL, USA. Her current research interests include data, brain, and health.

Dr. Fang was a recipient of the single-PI NSF CRII Award, the Ralph Lowe Junior Faculty Enhancement Award from ORAU, and the Robin Sidhu Memorial Young Scientist Award from the Society of Brain Mapping and Therapeutics.



Jing Qin received the Ph.D. degree in computer science and engineering from The Chinese University of Hong Kong, Hong Kong, China, in 2009.

He is currently an Assistant Professor with the School of Nursing, The Hong Kong Polytechnic University, Hong Kong, where he is also a Key Member with the Centre for Smart Health. His current research interests include virtual/augmented reality for healthcare and medicine training, medical image processing, deep learning, and health informatics.



David Dagan Feng (F'03) received the Ph.D. degree in computer science from the University of California at Los Angeles, Los Angeles, CA, USA, in 1988.

He is currently the Head with the School of Information Technologies, the Director with Biomedical and Multimedia Information Technology Research Group, and the Research Director with the Institute of Biomedical Engineering and Technology, The University of Sydney, Sydney, NSW, Australia. He has published over 700 research papers, pioneered

several new research directions, and made a number of landmark contributions in his field.

Dr. Feng was a recipient of the Crump Prize for Excellence in Medical Engineering in the University of California at Los Angeles, Los Angeles, CA, USA. He has served as the Chair of the International Federation of Automatic Control Technical Committee on Biological and Medical Systems, has organized/chaired more than 100 major international conferences/symposia/workshops, and has been invited to give over 100 keynote presentations in 23 countries/regions. He is a fellow of the Australian Academy of Technological Sciences and Engineering.