Research Article

An Automatic Cognitive Graph-Based Segmentation for Detection of Blood Vessels in Retinal Images

Rasha Al Shehhi, Prashanth Reddy Marpu, and Wei Lee Woon
Masdar Institute of Science and Technology, Abu Dhabi, UAE

Correspondence should be addressed to Rasha Al Shehhi; ralshehhi@masdar.ac.ae

Received 28 February 2016; Revised 10 April 2016; Accepted 21 April 2016

This paper presents a hierarchical graph-based segmentation for blood vessel detection in digital retinal images. This segmentation employs some of perceptual Gestalt principles: similarity, closure, continuity, and proximity to merge segments into coherent connected vessel-like patterns. The integration of Gestalt principles is based on object-based features (e.g., color and black top-hat (BTH) morphology and context) and graph-analysis algorithms (e.g., Dijkstra path). The segmentation framework consists of two main steps: preprocessing and multiscale graph-based segmentation. Preprocessing is to enhance lighting condition, due to low illumination contrast, and to construct necessary features to enhance vessel structure due to sensitivity of vessel patterns to multiscale/multiorientation structure. Graph-based segmentation is to decrease computational processing required for region of interest into most semantic objects. The segmentation was evaluated on three publicly available datasets. Experimental results show that preprocessing stage achieves better results compared to state-of-the-art enhancement methods. The performance of the proposed graph-based segmentation is found to be consistent and comparable to other existing methods, with improved capability of detecting small/thin vessels.

1. Introduction

Retinal vessel segmentation is a crucial step in analyzing fundus images of the eye for detection and diagnosis of many eye diseases. Some diseases such as glaucoma, diabetic retinopathy, and macular degeneration are very serious and might lead to blindness if they are not detected in time [1, 2]. The information about blood vessels, such as tortuosity and branching patterns, can not only provide information on pathological changes but also help to grade the disease severity and automatically diagnose the disease.

Although retinal vessel segmentation has been widely studied, it is still a challenging problem because of three main reasons. First, the quality of retinal images is highly variable and the segmentation methods face the challenge of low contrast or high homogeneity of illumination conditions [3–6]. Second, the complexity of vascular structures (different scales and orientations) means that most existing methods find it difficult to enhance multiscale vessel-like structures with various linear orientations [5–12]. Third, finding the most optimum model or method which is the most appropriate for variety of data is very difficult [13–15].

The studies of retinal images can be classified into pattern recognition (machine learning/model-based) [6, 16, 17], mathematical morphology [6, 18], kernel-based analysis [10, 11, 17] and tracking-based/path-based (Artificial Intelligent, AI) methods [5, 19, 20]. Here, morphological and AI methods are further discussed because they are more related to presented work in this paper.

Morphological methods examine the geometric vessel-like structure of retinal image by probing it with small patterns called structuring elements (SE) of predefined size and shape. Due to sensitivity of vessel-like patterns to different scales and orientations, most methods use multiscale or/and multiorientation structuring elements [18, 21, 22], such as multisturcure morphological operators [8, 12], and multiscale white top-hat with linear structuring elements [9]. One of the challenges is that there are several structures in retinal images such as optical disk, exudates, microaneurysms, and hemorrhages, which degrade the performance
of vessel detection methods. To overcome this problem, a number of approaches have been proposed in order to decompose components in retinal images. In [6], Morphological Component Analysis (MCA) has been proposed to separate some components such as lesions from vessels.

Tracking-based/path-based methods use regional information (a single vessel rather than the entire vasculature) to find the shortest/cheapest path that matches a vessel profile. The main advantage of this approach is that it provides precise vessel width, unlike other methods. Nowadays, there are many studies which follow this approach, for example, Dijkstra shortest path for vessel patterns [19], graph-cut [5], Bayesian-based tracking [20], and graph-analysis [23].

In this work, we propose a perceptual graph-based segmentation method. The complete framework consists of two stages. The first stage (preprocessing) removes noise as well as unwanted regions such as optical disk and surrounded darker background and produces a higher contrast vessel image. The second stage (segmentation) converts an image into connected graphical layer where each pixel is presented as a node and its spatial/spectral properties are used to merge pixels (nodes) to construct more semantic objects in higher-connected graphical layers. Gestalt perceptual principles, that is, similarity, closure, continuity, and proximity of spatial/spectral properties of nodes, are employed to assemble smaller parts that are most likely to represent a coherent connected vessel-like pattern.

The experimental evaluation is carried out to test the behavior of segmentation algorithm using the standard datasets such as DRIVE, ARIA, and STARE (details of these datasets are given in Section 4.1) using the following major criteria: sensitivity (Se), specificity (Sp), accuracy (Acc), and area under curve (AUC).

In the following, the idea of perceptual Gestalt principles in image segmentation is introduced in Section 2. Section 3 illustrates the proposed hierarchical graph-based segmentation framework in detail. First, it presents the preprocessing part, which includes filtering-based inhomogeneity correction using Gaussian filter, followed by morphology-based illumination enhancement method. Second, it shows the segmentation part, which is based on integrating the idea of perceptual Gestalt principles into object-based features to merge segments. Then, Section 4 presents the datasets, experimental metrics, and results of segmentation. Finally, the conclusions are drawn and some ideas for future work are presented in Section 5.

2. Contribution of Current Work

The main contribution of this work is to introduce perceptual Gestalt (form, grouping) principles [24–28] of some middle-level image features into graph-based segmentation to discriminate the connected coherent vessel-like patterns from background.

Four Gestalt principles are proposed, inspired from similarity, closure, continuity, and proximity. The theory behind Gestalt grouping is based on how human vision performs perceptual grouping to assemble parts of an image that most likely represent a single object in the scene. Similarity can be used to group segments into one object depending on the number of similar factors such as color, size, and shape. Proximity rule is employed by assembling different parts which are close to each other to present one object. Good continuity is the tendency of elements to be grouped to form smooth contours depending on the number of factors such as orientation of local elements, contour length, and curvature properties. The principle of closure refers to tendency to see an element/object as a complete form or figure, ignoring gaps and incomplete contour lines. It is not necessary to create triangles, circle, and so forth, but it fills the missing information to create familiar shapes [25, 26, 28].

The integration of perceptual Gestalt principles into graph-based segmentation, from computational view, is an important stage to reduce required visual processes to interpret an input image by converting a fully connected layer into locally connected layer [29]. Moreover, such integration helps to cope with undersegmentation/oversegmentation in image layer(s) or between layers.

In this work, the perceptual principles are employed as follows. The first level, defined as color-layer, is built by grouping pixels based on similarity of color between each pixel and its 8-connected neighborhood (Gestalt similarity in illumination characteristics). The second layer, called as black top-hat (BTH) layer, is constructed by grouping adjacent objects, which most likely represent vessel-like shape after applying BTH morphological operator (Gestalt closure in connected components of most common label in BTH-layer). The final Dijkstra-layer is created by conjunction of adjacent objects which have a high probability of constructing connected objects (Gestalt continuity in Dijkstra tracking path). Gestalt proximity is employed by considering 8-connected neighborhood as a Gestalt connectivity patch in all layers.

3. The Proposed Method

The proposed framework (Figure 1) comprises two major stages: preprocessing and segmentation. Preprocessing (green rectangles: A–C) consists of filtering-based inhomogeneity correction and morphology-based illumination enhancement (Section 3.1). Hierarchical graph-based segmentation (red rectangles: 1D–3D) is based on construction of five layers based on the number of objects in retinal images: original RGB image (largest), ROI (region of interest), color, followed by black top-hat (BTH), and finally Dijkstra-layers (smallest) (Section 3.2).

3.1. Preprocessing. Preprocessing involves two main steps to produce more effective feature image which shows a high contrast between vessel and nonvessel objects to facilitate the segmentation. The first step is to remove the effect of background variations by nonuniform illumination and the second step is to eliminate the complexity of vascular structure because of multiple scales and orientations.

In this work, we choose the green channel image for preprocessing, as it exhibits the best vessel/no-vessel contrast in retinal images. The red channel can be saturated and blue
channel has poor dynamic ranges [16]. In addition, using only the green channel decreases the computational time compared to processing all RGB channels. We first extract the region of interest (ROI) representing the fundus, which is a circle-like shape/nonblack region at the center of retinal images without including the surrounding black background. We eliminate the influence of the background by masking it to accelerate further processing stages by focusing only on pixels/objects in ROIs. To find ROI, simple minimum threshold for all red, green, and blue channels is applied to remove unwanted background from RGB channel. However, some missed labeled pixels are created on retinal foreground and background. These noises are eliminated by morphological erosion [30–32], where ROI is shrunk into center of retinal images (Figure 1(A)).

3.1.1. Filtering-Based Inhomogeneity Correction (Figure 1(B)). Due to inhomogeneous light conditions, retinal images may contain background (nonvessels) with high similarity to foreground (vessels), which will degrade the performance of the segmentation method. Therefore, it is important to remove effects of the varying illumination conditions. Zhao et al. [5] applied Retinex theory, adapted from the field of computer vision, to remove unwanted illumination effects by component-wise multiplication of reflected green with green illumination channel \( R_G \times G \). \( R_G \) is component-wise log-subtraction bilateral blurring of green channel \( I(B_G) \) from green channel \( G \). Imani et al. [6] used reflectance component of Retinex theory to reduce illumination differences, but with component-wise subtraction of median blurring of green channel. Some other studies used Contrast-Limited Adaptive Histogram Equalization (CLAHE) method instead of global enhancement methods such as histogram equalization and gamma correction; however, its local enhancement is uniform regardless of whether it is foreground or background [3, 4].

In this work, we use the component-wise subtraction of background from the selected channel to remove the light variations. Let \( x \) be a pixel of a green channel. The retinal image background \( f_B \) is created by applying low-pass Gaussian blurring of green channel \( f_G \) with a reasonable filter size (e.g., \( w = 61 \times 61 \)) to eliminate the effect of the brightest region from retinal images (optical disk). The corrected image \( f_C \) is computed by subtracting \( f_B \) from \( f_G \). This is given as

\[
 f_C(x) = f_G(x) - f_B(x).
\]

Figure 2 illustrates a comparison between previously mentioned correction methods using selected examples from DRIVE, ARIA, and STARE datasets. In general, all methods enhance image contrast, but there are still large areas of homogeneity. Component-wise subtraction of Gaussian blurring is not the best method; however, it has succeeded in eliminating the most noisy part (optical disk) and enhances
Figure 2: Inhomogeneity correction results on selected images. (a) shows green channel of randomly chosen images from the DRIVE training, DRIVE testing, ARIA, and STARE datasets, respectively. (b) shows CLAHE results [3, 4]. (c) shows results of Retinex theory on green channel [5]. (d) presents difference between green channel and median blurring of green channel [6]. (e) presents results of our correction method.

the contrast between vessels and background. This can be especially noticed in ARIA and STARE images after applying morphology enhancement, as depicted in Figure 4. As a consequence, the vessels can be easily identified from the background. Component-wise correction of bilateral filtering has comparable results after applying morphology enhancement, because it is an edge-preserving smoothing filter that maintains edges of vessels [33].

3.1.2. Morphology-Based Illumination Enhancement (Figure 1(C)). Mathematical morphology is a nonlinear method which uses the concepts of set, topology, and geometry to analyze geometrical structures (e.g., shape and form of objects) in images. It examines the geometric structure of an image by probing with structuring elements (B) [30–32].

In this work, black top-hat (BTH) morphological operator is used. This is because the BTH operator is the most
applicable method to extract image structure under lower illumination conditions [30], which is the case in blood vessel detection application. The BTH is obtained by subtracting morphological closing image \( \phi(3(x)) \) of corrected image from corrected image \( f(x) \), as in (2). The morphological closing \( \phi(3(x)) \), which is dilation \( \delta_B(f) \) followed by erosion \( \varepsilon_B(f) \), acts as a shape filter and preserves objects having relevant structure in image (3). This is given as

\[
BTH(f) = \phi(f(x)) - f(x), \quad (2)
\]

\[
\phi(f) = \varepsilon_B(f) \delta_B(f). \quad (3)
\]

All vessel-like patterns are components of linear-shape structures with various horizontal, vertical, and diagonal orientations. To address these differences, we suggest the following approaches.

(i) **Multiscale Multiorientation BTH.** The BTH is adaptively computed by probing the corrected image with multiple linear structuring elements which have a variety of angular orientations. Therefore, a set of linear structuring elements, where each is a matrix representing a line with 3, 5, 9, 15, and 21 pixels of length and rotated in steps of \( \pi/8 \), is used for morphological BTH. As a result, each isolated BTH along each direction will brighten the vessels in that direction, providing that the length of \( B \) is large enough to extract the vessel with largest diameter. Finally, we consider the average of the first three maximum BTH results because the first three BTHs present the highest differences between vessel/nonvessel patterns, as the following:

\[
BTH = \frac{\sum_{3} \text{max} BTH^\theta(f)}{3}, \quad (4)
\]

where \( s \) and \( \theta \) are scale and orientation of linear structuring elements \( B \).

(ii) **Multiscale BTH.** Multiscale BTH is defined as an average of morphological BTHs obtained from probing corrected image with size of 3, 5, 7, 9, and 11 of elliptical-structuring elements \( B \), unlike the previous method which uses linear structuring elements in different directions, given by

\[
BTH = \frac{\sum_{3} BTH^\theta(f)}{5}. \quad (5)
\]

In this paper, two other enhancement methods are selected for comparison: phase-based method [5] and wavelet-based method [10]. In order to reproduce results, the parameters used for these filters are the same as recommended in the corresponding literature [5, 10]. Figure 3 shows the results of applying multiscale BTH, multiscale multiorientation BTH, phase-based method, and wavelet-based method. It can be seen clearly that illumination contrast using wavelet-based method is poor compared to phase-based method and the proposed method, where the vessel-like structure is more distinguishable. However, multiscale BTH method produces more consistent results at optical disk and fovea regions compared to other parts of the vessels. Figure 4 shows the result of applying multiscale BTH and multiscale multiorientation BTH to different correction methods, that is, CLASHA [3, 4], Retinex [5], difference of green and its median blurring [6], and proposed correction method. In summary, it seems that applying morphological operators over all correction images successfully enhances the contrast of vessels. The results of multiscale BTH over proposed correction method seem to be efficient in removing inhomogeneity within the image including optical disk and fovea regions.

### 3.2. Hierarchical Graph-Based Segmentation (Figure 1(D))

In this paper, a cognitive vision approach to graph-based image segmentation is proposed by employing perceptual knowledge of contextual features to provide semantically vessel-like patterns. Moreover, it decreases the required computational cost by processing perceptual features instead of processing the fully connected image in blood vessel applications.

**Graph-based segmentation** represents the input image as a weighted graph \( G = (V, E, W) \), where all vertices/nodes \( v_i \in V \) correspond to a pixel/region of the image \( V = \{v_1, v_2, \ldots \} \) and edges \( e(v_i, v_j) \in E \) correspond to pair of neighboring vertices/nodes. Each edge \( e(v_i, v_j) \in E \) has a corresponding weight \( w(v_i, v_j) \) which represents the dissimilarity between adjacent vertices/nodes \( i \) and \( j \). The dissimilarity between edges is often defined based on selected properties relevant to the application (e.g., similarity of color and shape) [34–36].

In this work, we build an undirected graph with multilayer layers, where the number of vertices/nodes is hierarchically reduced from top-layer to the next bottom-layer because nodes are merged to construct objects with more semantic interpretation from visual perception view. There are two most important questions in graph-based algorithms to be considered in image segmentation. First is the weighting function that presents spectral or/spatial relationship between adjacent nodes. Second is the merging/homogeneity criteria to group adjacent vertices/nodes to have connected components. The graph starts with color-layer and BTH-layer, followed by Dijkstra-layer. The design of each layer is illustrated in the following sections.

#### 3.2.1. Gestalt Similarity and Proximity in Contextual Color Features (Figure 1(ID))

In order to build spectral level, the graph-based framework proposed in [35] is applied. It translates Gaussian smoothed input image into a graph, where each pixel \( v_i \in V \) is mapped to vertex/node and each edge \( e_k = (v_i, v_j) \in E \) reflects spectral relationships between adjacent pixels. We consider 8-connected neighborhood as a Gestalt connectivity patch. The initial weighting function is Euclidean distance between red, green, and blue components between two adjacent vertices/nodes \( v_i \) and \( v_j \). This is given as

\[
w(v_i, v_j) = \sqrt{(R_{v_i} - R_{v_j})^2 + (G_{v_i} - G_{v_j})^2 + (B_{v_i} - B_{v_j})^2}. \quad (6)
\]

The spectral vertices/nodes are hierarchically merged based on degree of spectral similarity (for more details of the
algorithm, read from [35]). The output of spectral-grouping is not robust enough to perceptually interpret resulting segmented regions as semantic components, as depicted in Figure 5(c). High-level criteria may efficiently enhance results by grouping/splitting components into more meaningful spatial structures starting from spectral components instead of pixels. Therefore, grouping in higher layers is similar to the spectral-layer, but with other weighting and merging criteria for different measures of similarity. At each stage, components are merged iteratively from small to large until convergence (no more merging). The stop condition is based on prior knowledge and it is used to prevent oversegmentation/undersegmentation.

3.2.2. Gestalt Closure and Proximity Based on Contextual BTH Features (Figure 1(2D)). A sequential connected component labeling approach [37, 38] is employed to separate
vessel-like components from background components by scanning all pixels in vessel-layer and consequently labeling each connected component. For convenience, we consider the case of 8-connected neighborhood (Gestalt proximity patch). The suggested labeling threshold is based on ratio of mean average over standard deviation average of vessel-layer with 10 homogeneous pixels as the minimum number of pixels in one component.

We consider the labeled-component with highest number of pixels as background (label 0) and other components as foreground components (labels 1, 2, 3, 4, 5, etc.). On the other hand, non-masked area is typically left unchanged by the labeling process. As a result, background label \( L_B \) is used to cut edges of spectral components or vertices from spectral components (see (7)) to build BTH-layer where the number of nodes in this layer is lower than the number of nodes in spectral-layer, as depicted in Figure 5(d):

\[
BTH(v) = \begin{cases} 
vessel, & v \in L_F, 

\text{no-vessel}, & v \in L_B. 
\end{cases} 
\]  

(7)

### 3.2.3. Gestalt Continuity and Proximity Based on Contextual Features (Figure 1(3D)).

The BTH-layer is updated by assigning a new weight for edges between BTH-spectral connected components. The weight is based on the smallest Euclidean distance between vertices of all adjacent components.

The development of our continuity feature is motivated by the need for a representation of long contour representation suitable to visually perceive vessel-like patterns as continuous irregular lines. In order to apply continuity principle, Dijkstra algorithm [19, 39] is employed within window size \( w = 50 \times 50 \) by considering each first component in BTH-layer block as a source point of the graph-path and each furthest component in window \( w = 50 \times 50 \) as target point of graph-path. All vertices in Dijkstra path from source to target are iteratively merged to be one component, until convergence of vessel and nonvessel regions (Figure 5(e)).

Figure 5 shows some examples from standard datasets. The spectral segmentation is not sufficient to obtain meaningful structures, because lighting contrast of vessel-like/non-vessel-like patterns is low. Therefore, the results are with undersegmented/oversegmented regions, so higher-level criteria should be used efficiently to group or split components into more meaningful spatial structures (BTH). This shows the importance of using shape features to eliminate wrongly labeled vessel-like patterns. On the other hand, other properties such as asymmetry and length/width [40] help to decrease the number of nodes in BTH-layer; however, they are not enough to find connectedness of nonadjacent nodes in BTH-layer.

As a result, Graph-based characteristics (Dijkstra algorithm) are introduced in image segmentation to determine connectedness of non-neighbor nodes, which can be integrated to build a complete contour. Table 1 presents the number of components in vessel multilayer graph, starting from initial green channel, ROI image, and BTH and Dijkstra-layers.
Figure 5: Demonstrative results of each stage of applying graph-based segmentation. (a) shows some selected images from the DRIVE training, DRIVE testing, ARIA, and STARE datasets. Gold standards are shown in (b). (c) presents spectral-layer. The BTH-layer after applying graph-cut to vessel-layer (multiscale BTH) is illustrated in (d). (e) shows the final-layer after applying Dijkstra path.
4. Experimental Evaluation

We have employed three public retinal image datasets to evaluate the proposed segmentation framework. In this section, a brief introduction to these datasets is provided in Section 4.1; evaluation criteria including preprocessing and segmentation criteria are defined in Section 4.2, followed by experimental results in Section 4.3.

4.1. Data. We obtained human retinal images from publicly available datasets: DRIVE, ARIA, and STARE. All datasets consist of RGB components of retinal images with their corresponding ground truth images where blood vessel-like structures are segmented. These datasets are selected because of availability of gold standard from manual annotations of retinal vessels by experts.

DRIVE (Digital Retinal Images for Vessel Extraction). It consists of training and testing image with 565 × 584 pixels, which are obtained from a diabetic retinopathy screening program in the Netherlands. The set of 40 photographic images have been randomly selected, 33 do not show any sign of diabetic retinopathy, and 7 show signs of mild early diabetic retinopathy. The manual segmentation of set A is used as ground truth. The DRIVE dataset is available at http://www.isi.uu.nl/Research/Databases/DRIVE/.

ARIA (Automated Retinal Image Analysis). It consists of three groups: 92 images of age-related muscular degeneration, 59 images of patients with diabetes, and 61 images of healthy eyes, collected by St. Paul’s Eye Unit and University of Liverpool. Each image was captured at a resolution of 768 × 576 pixels. The manual segmentation from observer DGP is used as ground truth. The ARIA dataset is available at http://aria.cvs.rochester.edu/#&panel1-1.

STARE (Structural Analysis for the Retina). It consists of 20 images with 10 images showing evidence of pathology, obtained from University of California. Each image was captured by 8 bits per plane at 605 × 700 pixels. Hand-labeled vessel networks, provided by Adam Hoover (1st manual) and Valentina Kouznetsova (2nd manual), are used as ground truth. The STARE datasets are available at http://www.ces.clemson.edu/~ahoover/stare/probing/index.html.

4.2. Evaluation Measures

4.2.1. Preprocessing Measures. In order to evaluate contrast enhancement, several objective methods are proposed: Contrast Improvement Index (CII), Peak Signal-to-Noise Ratio (PSNR), and Mean Structural Similarity (MSSIM) [8, 45, 46]. Th CII is ratio of contrast of the enhanced and original images. CII is computed as follows:

\[
\text{CII} = \frac{C_{\text{enh}}}{C_{\text{original}}},
\]

where \( C \) presents the contrast between foreground (vessel) and background (retinal regions except vessels). \( fg \) and \( bg \) are mean gray-level value of foreground and background, respectively. The larger \( C \) and consequently the larger CII, the more obvious the difference between foreground and background (higher contrast and better enhancement). The PSNR measures intensity changes of original and enhanced images based on Mean Square Error (MSE), as follows:

\[
\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{max}_I^2}{\text{MSE}} \right).
\]

To compute \( M_{\text{SSIM}} \), the image is decomposed into \( k \) blocks of size \( n \times m \) and, for each block, SSIM is computed as follows:

\[
\text{SSIM} = \frac{(2\bar{x}\bar{y} + C_1) \left( 2\sigma_{xy} + C_2 \right)}{\left( \sigma_x^2 + \sigma_y^2 + C_2 \right) \left( \bar{x}^2 + \bar{y}^2 + C_1 \right)},
\]

where \( \bar{x}, \bar{y}, \sigma_x, \) and \( \sigma_y \) are means and standard deviations in original and enhanced blocks. \( \sigma_xy \) corresponds to covariance measure, \( C_1 = (0.01 \times 255)^2 \), and \( C_2 = (0.03 \times 255)^2 \).

4.2.2. Segmentation Measures. Four common metrics are employed to measure the performance of the proposed segmentation: sensitivity (Se) or True Positive Rate (TPR), specificity (Sp) or False Positive Rate (FPR), accuracy (Acc), and area under curve (AUC). Sensitivity indicates the proportion of vessel pixels identified by the proposed method that coincide with vessel pixels in ground truth images (11), while specificity is the proportion of detected pixels by the proposed method that are not vessels in ground truth images (12). Accuracy indicates the proportion of vessel/nonvessel patterns that are correctly identified in terms of the total number of pixels in retinal images (13). ROC is obtained by plotting the fraction of Se/TPR versus Sp/FPR (14). The closer

<table>
<thead>
<tr>
<th>Datasets</th>
<th>DRIVE training</th>
<th>DRIVE testing</th>
<th>ARIA</th>
<th>STARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial layer</td>
<td>329960</td>
<td>329960</td>
<td>442368</td>
<td>423500</td>
</tr>
<tr>
<td>Initial layer with ROI</td>
<td>210702</td>
<td>208174</td>
<td>350013</td>
<td>288761</td>
</tr>
<tr>
<td>Spectral-layer</td>
<td>9874</td>
<td>9816</td>
<td>15094</td>
<td>12917</td>
</tr>
<tr>
<td>BTH-layer</td>
<td>3285</td>
<td>3719</td>
<td>3130</td>
<td>4968</td>
</tr>
<tr>
<td>Dijkstra-layer</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 2: Graph-based segmentation performance without and with correction methods on DRIVE, ARIA, and STARE datasets, respectively. 2nd row: CLAHE \((w = 3 \times 3)\) [3, 4, 41]. 3rd row: Retinex \((w = 3 \times 3\), spatial spread based on low-pass filter \(\sigma_d = 0.3\), and geometric spread of the image intensity \(\sigma_r = 0.3\) [5]. 4th row: subtraction median blurring image of green \((w = 31 \times 31)\) from green image [6]. 5th row: subtraction low-pass Gaussian blurred image of green \((w = 61 \times 61)\) from green image. Se: sensitivity, Sp: specificity, Acc: accuracy, and AUC: area under curve.

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Correction methods</th>
<th>Se</th>
<th>Sp</th>
<th>Acc</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRIVE</td>
<td>Without</td>
<td>0.896</td>
<td>0.779</td>
<td>0.789</td>
<td>0.837</td>
</tr>
<tr>
<td></td>
<td>CLAHE</td>
<td>0.841</td>
<td>0.867</td>
<td>0.865</td>
<td>0.806</td>
</tr>
<tr>
<td></td>
<td>Retinex</td>
<td>0.804</td>
<td>0.934</td>
<td>0.923</td>
<td>0.854</td>
</tr>
<tr>
<td></td>
<td>Green-MF (Green)</td>
<td>0.911</td>
<td>0.619</td>
<td>0.645</td>
<td>0.764</td>
</tr>
<tr>
<td></td>
<td>Green-GF (Green)</td>
<td>0.921</td>
<td>0.702</td>
<td>0.931</td>
<td>0.869</td>
</tr>
<tr>
<td>ARIA</td>
<td>Without</td>
<td>0.743</td>
<td>0.913</td>
<td>0.898</td>
<td>0.828</td>
</tr>
<tr>
<td></td>
<td>CLAHE</td>
<td>0.781</td>
<td>0.823</td>
<td>0.819</td>
<td>0.802</td>
</tr>
<tr>
<td></td>
<td>Retinex</td>
<td>0.768</td>
<td>0.701</td>
<td>0.707</td>
<td>0.734</td>
</tr>
<tr>
<td></td>
<td>Green-MF (Green)</td>
<td>0.794</td>
<td>0.667</td>
<td>0.679</td>
<td>0.731</td>
</tr>
<tr>
<td></td>
<td>Green-GF (Green)</td>
<td>0.721</td>
<td>0.944</td>
<td>0.924</td>
<td>0.832</td>
</tr>
<tr>
<td>STARE</td>
<td>Without</td>
<td>0.629</td>
<td>0.924</td>
<td>0.850</td>
<td>0.774</td>
</tr>
<tr>
<td></td>
<td>CLAHE</td>
<td>0.643</td>
<td>0.902</td>
<td>0.838</td>
<td>0.770</td>
</tr>
<tr>
<td></td>
<td>Retinex</td>
<td>0.650</td>
<td>0.894</td>
<td>0.839</td>
<td>0.769</td>
</tr>
<tr>
<td></td>
<td>Green-MF (Green)</td>
<td>0.670</td>
<td>0.866</td>
<td>0.821</td>
<td>0.769</td>
</tr>
<tr>
<td></td>
<td>Green-GF (Green)</td>
<td>0.563</td>
<td>0.975</td>
<td>0.879</td>
<td>0.774</td>
</tr>
</tbody>
</table>

the curve approaches the top left corner with AUC close to 1, the better the performance of the proposed method is [47]:

\[
Se = \frac{TP}{TP + FN},
\]

\[
Sp = \frac{TN}{TN + FP},
\]

\[
Acc = \frac{TP + TN}{TP + FP + TN + FN},
\]

\[
AUC = \frac{Se + Sp}{2},
\]

where TP is defined as the number of vessel pixels correctly detected in the retinal images, FP is defined as the number of nonvessel pixels detected as vessels, TN is the number of nonvessel pixels correctly detected, and FN is the number of vessel pixels detected as nonvessel pixels.

4.3. Results. In order to evaluate the efficiency of the proposed segmentation method, we use two approaches. First, each individual stage of the proposed method is evaluated by its comparable stages of other methods considered in this study (i.e., inhomogeneity correction and illumination enhancement) across all three datasets. Second, the performance of the proposed segmentation is compared with other works across DRIVE and STARE datasets.

4.3.1. Inhomogeneity Correction Assessment. The optical disk and fovea area contribute to most of the false detection of vessel pixels in most blood vessel detection frameworks [9, 10, 48]. Subtracting green channel from low-pass Gaussian blurring of green channel increases the contrast between optical disk/foveal area and vessel pixels. Therefore, the noisy area will not be enhanced after morphology enhancement, especially after multiscale BTH. Figure 6 presents an example of applying different correction methods, mentioned in Section 3.1.1, to one of the images from DRIVE testing dataset. Our correction method (Figure 6(e)) shows a superior performance for the vessel subtraction and is more similar to ground truth image, compared to other methods. Table 2 shows that the accuracy of our correction method is the highest in all three datasets.

4.3.2. Illumination Enhancement Assessment. Figure 7 shows an example of applying enhancement methods to an image selected randomly from DRIVE testing dataset. The results of multiscale BTH outperform other enhancement methods. The false detection of segmentation method after applying multiscale multiorientation BTH is because it fails to remove optical disk and some exudates. Optic disk area, which can be seen as brightest, round, vertically slightly oval disk, can be easily preserved by multiscale multiorientation BTH, which identifies line-shape structures in vertical, horizontal, or diagonal angularity with varying sizes. Moreover, exudates can be identified as areas with varying size and shapes, where it is difficult to be eliminated by multiscale multiorientation BTH. In contrast, multiscale BTH maintains just all ellipse-shape structures regardless of their orientations and then finds magnitude of directional blurring of BTH. As depicted in Table 3, the accuracy of multiscale BTH is above 90% in all
Figure 6: Demonstrative comparison of correction methods on one of the DRIVE testing images. The first row shows the results of inhomogeneity correction method and its consequence final segmentation in the second row. (a) RGB image and 1st manual. (b) Green channel without correction. (c) CLAHE correction. (d) Retinex correction. (e) Difference between green and median blurring of green channels. (f) Difference between green and Gaussian blurring of green channel.

Table 3: Graph-based segmentation performance of four enhancement methods: local phase-based [5], wavelet-based [10], multiscale multiorientation BTH, and multiscale BTH methods on DRIVE, ARIA, and STARE datasets, respectively. Se: sensitivity, Sp: specificity, Acc: accuracy, and AUC: area under curve.

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Enhancement methods</th>
<th>Se</th>
<th>Sp</th>
<th>Acc</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Phase-based</td>
<td>0.820</td>
<td>0.827</td>
<td>0.826</td>
<td>0.823</td>
</tr>
<tr>
<td></td>
<td>Wavelet-based</td>
<td>0.827</td>
<td>0.768</td>
<td>0.773</td>
<td>0.798</td>
</tr>
<tr>
<td></td>
<td>Multiscale multiorientation BTH</td>
<td>0.635</td>
<td>0.924</td>
<td>0.945</td>
<td>0.803</td>
</tr>
<tr>
<td></td>
<td>Multiscale BTH</td>
<td>0.863</td>
<td>0.930</td>
<td>0.924</td>
<td>0.893</td>
</tr>
<tr>
<td>DRIVE</td>
<td>Phase-based</td>
<td>0.469</td>
<td>0.818</td>
<td>0.785</td>
<td>0.644</td>
</tr>
<tr>
<td></td>
<td>Wavelet-based</td>
<td>0.725</td>
<td>0.683</td>
<td>0.687</td>
<td>0.704</td>
</tr>
<tr>
<td></td>
<td>Multiscale multiorientation BTH</td>
<td>0.813</td>
<td>0.893</td>
<td>0.871</td>
<td>0.828</td>
</tr>
<tr>
<td></td>
<td>Multiscale BTH</td>
<td>0.746</td>
<td>0.947</td>
<td>0.920</td>
<td>0.841</td>
</tr>
<tr>
<td>ARIA</td>
<td>Phase-based</td>
<td>0.432</td>
<td>0.776</td>
<td>0.855</td>
<td>0.699</td>
</tr>
<tr>
<td></td>
<td>Wavelet-based</td>
<td>0.469</td>
<td>0.773</td>
<td>0.711</td>
<td>0.620</td>
</tr>
<tr>
<td></td>
<td>Multiscale multiorientation BTH</td>
<td>0.660</td>
<td>0.944</td>
<td>0.888</td>
<td>0.803</td>
</tr>
<tr>
<td></td>
<td>Multiscale BTH</td>
<td>0.606</td>
<td>0.977</td>
<td>0.910</td>
<td>0.800</td>
</tr>
<tr>
<td>STARE</td>
<td>Phase-based</td>
<td>0.432</td>
<td>0.776</td>
<td>0.855</td>
<td>0.699</td>
</tr>
<tr>
<td></td>
<td>Wavelet-based</td>
<td>0.469</td>
<td>0.773</td>
<td>0.711</td>
<td>0.620</td>
</tr>
<tr>
<td></td>
<td>Multiscale multiorientation BTH</td>
<td>0.660</td>
<td>0.944</td>
<td>0.888</td>
<td>0.803</td>
</tr>
<tr>
<td></td>
<td>Multiscale BTH</td>
<td>0.606</td>
<td>0.977</td>
<td>0.910</td>
<td>0.800</td>
</tr>
</tbody>
</table>

datasets and consequently their AUC values are most closer to 1.

The suggested preprocessing enhancement method is tested by combining it with correction methods and comparing it with other enhancement methods, that is, local-phased and wavelet-based methods, in terms of the following criteria: CII, PSNR, and $M_{SSIM}$ (Table 4). The CII and PSNR of multiscale BTH with low-pass Gaussian blurring correction are relatively largest. On the other hand, $M_{SSIM}$ of both median and Gaussian corrections are comparable. The reason is that both filtering methods succeed in efficiently removing noisy areas by filtering operators of large size.

4.3.3. Comparison between Proposed Segmentation and State-of-the-Art Methods. The segmentation performance of the proposed method in terms of sensitivity, specificity, accuracy, and area under curve is compared with other state-of-the-art methods (matched filtering, supervised, unsupervised, and artificial methods) in the most public datasets: DRIVE
and STARE, as depicted in Table 5. ARIVE dataset is not used because only a few studies used ARIVE dataset. The comparison shows that the proposed method is comparable to most of blood vessel detection frameworks. The slightly lower value is due to the fact that the proposed segmentation shows smaller and thinner vascular structures which are not generally annotated in the ground truth images. This can be clearly seen in Figure 8 which shows examples from STARE and ARIA datasets where the proposed segmentation method succeeds in detecting small and thin vessel-like patterns. It is also due to overestimated vessel width, which is a result of initial spectral segmentation (Section 3.2.1), ignoring all spatial properties to build initial regions.

5. Discussion and Conclusion

In this paper, we proposed a new method to detect blood vessels in fundus images, which is based on three main steps: filtering-based correction, morphological-illumination enhancement, and graph-based segmentation.

Low-pass Gaussian blurring (filtering-based correction) is suggested to remove most noisy areas in retinal images, that
Table 5: Performance of segmentation methods based on sensitivity (Se), specificity (Sp), accuracy (Acc), and area under curve (AUC) on DRIVE and STARE datasets. The hand segmented images from first-manual observers are used as benchmarks (1st STARE manual is selected because all works used it).

<table>
<thead>
<tr>
<th>Methods</th>
<th>DRIVE</th>
<th>STARE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>Bankhead et al. [10]</td>
<td>0.703</td>
<td>0.028</td>
</tr>
<tr>
<td>Estrada et al. [19]</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Wang et al. [11]</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Nguyen et al. [42]</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Orlando and Blaschko [43]</td>
<td>0.785</td>
<td>0.967</td>
</tr>
<tr>
<td>Salazar-Gonzalez et al. [44]</td>
<td>0.7512</td>
<td>0.0316</td>
</tr>
<tr>
<td>Mithun et al. [18]</td>
<td>0.471</td>
<td>0.970</td>
</tr>
<tr>
<td>Azzopardi et al. [4]</td>
<td>0.744</td>
<td>0.978</td>
</tr>
<tr>
<td>Zhao et al. [5]</td>
<td>0.744</td>
<td>0.978</td>
</tr>
<tr>
<td>Imani et al. [6]</td>
<td>0.752</td>
<td>0.975</td>
</tr>
<tr>
<td>Proposed method</td>
<td>0.850</td>
<td>0.944</td>
</tr>
</tbody>
</table>

Figure 8: Illustrative comparison between results obtained from manual observers and the proposed segmentation on selected sample from STARE and ARIA datasets. (a) Selected sample. (b) 2nd manual sample (selected because it presents most small/thin blood vessels). (c) Segmented sample.

is, optical disk and fovea. This correction method may not be the best method; however, it has succeeded in detecting the most noisy areas from retinal images.

Due to high sensitivity of multistructure elements to edges in all directions, multiscale/multiorientation and multiscale morphological BTHs (morphological-illumination enhancement) are proposed which are capable of detecting most of the blood vessel edges which are thin and small. The deficiency of missing some thin vessels is due to the minimum threshold of connected component labeling of BTH-layer. Therefore, the need for appropriate thresholding method to find thin vessel and avoid false-edge pixel is important. Hence, one of our future works is to find more suitable connected component thresholding method to increase the accuracy of the proposed method.

Hierarchical graph-based segmentation is based on applying perceptual Gestalt principles (e.g., similarity, closure, proximity, and continuity) of spectral/spatial features between nodes in graphical multilayers. For instance, similarity of spectral characteristics between adjacent nodes is informative; however, semantic ambiguity still exists because of the appearance, shape, or other higher-level features’ similarity. Moreover, closure principle is applied by combining nodes if their connected component labels are not part of the common connected component label of BTH-layer. From authors’ viewpoint, there are other laws behind this closure principle which are similar in size and similar in orientation because BTH operator preserves multistructure which has variety in size and orientation.
This work solved the problem of the priority of the highest importance principle, unlike other works [28]. The major law in this work is proximity (connectedness). This is because the adjacent nodes in each layer are taken into account to make vessel-like pattern stand out from its background as a separate object by grouping small subobjects together with their surroundings based on other principles.

The quantitative performance of enhancement as well as segmentation shows that the proposed method works well on healthy retinal image with accuracy 93.4%. However, a main drawback is that it tends to produce false detection in overlapping areas. This has lowered the overall accuracy of the proposed method, especially on STARE images. To solve this problem, two-label graph-cut step will help to reduce incorrect detection and improve performance of the proposed segmentation.

In this paper, the proposed method is objectively evaluated. From the view of cognitive psychology, subjective evaluation is an efficient method to test the performance of perceptual segmentation. Therefore, we aim to investigate human visual perception for segmenting and consequently detecting blood vessels (new ground truth) and comparing with segmentation results.

Algorithm run-time is an issue of concern to assess the algorithm performance. In this paper, the computation time of the proposed method is 10 minutes and 3 seconds. The preprocessing part takes around 3 s and segmentation part takes about 10 m to analyze graph algorithms. The proposed algorithm is quite slow because of sequential connected component labeling and Dijkstra algorithms in graph-based segmentation. Therefore, in the future, we aim to parallelize connected component labeling and Dijkstra algorithms by translating our graph-based segmentation into parallel segmentation using massively parallel GPU [49].

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References


[20] Y. Yin, M. Adel, and S. Bourennane, "Retinal vessel segmenta-
segmentation using color image morphology and local binary
automatic graph-based approach for artery/vein classification in
NY, USA, 1935.
psychology in visual perception: I. Perceptual grouping and
[27] A. Richtsfeld, M. Zillich, and M. Vincze, "Implementation of
scenes based on visual attention and gestalt grouping laws," in
Proceedings of the IEEE International Conference on Systems,
[29] P. Soille, Morphological Image Analysis: Principles and Applica-
some morphological hierarchies on edge-weighted graphs," in
Mathematical Morphology and Its Applications to Signal and
Image Processing: 11th International Symposium, ISMM 2013,
[32] M. Elad, "Retinex by two bilateral filters," in Proceedings of the
5th International Conference on Scale Space and PDE Methods in
[33] M. R. Khokher, A. Ghafoor, and A. M. Siddiqui, "Image
segmentation using fuzzy rule based system and graph cuts," in
Proceedings of the 12th International Conference on Control,
[34] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based
Submit your manuscripts at http://www.hindawi.com