Computerized Medical Imaging and Graphics 52 (2016) 28–43

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

A multi-scale tensor voting approach for small retinal vessel segmentation in high resolution fundus images

Argyrios Christodoulidis[a,*], Thomas Hurtut[a], Houssem Ben Tahar[b], Farida Cheriet[a]

[a] Polytechnique Montréal, Montréal, QC H3C 3A7, Canada
[b] Diagnos Inc., Brossard, QC J4Z 1A7, Canada

ARTICLE INFO

Article history:
Received 24 July 2015
Received in revised form 16 April 2016
Accepted 1 June 2016

Keywords:
Diabetic retinopathy
Fundus imaging
Retinal blood vessel segmentation
Multi-scale line detection
Perceptual organization

ABSTRACT

Segmenting the retinal vessels from fundus images is a prerequisite for many CAD systems for the automatic detection of diabetic retinopathy lesions. So far, research efforts have concentrated mainly on the accurate localization of the large to medium diameter vessels. However, failure to detect the smallest vessels at the segmentation step can lead to false positive lesion detection counts in a subsequent lesion analysis stage. In this study, a new hybrid method for the segmentation of the smallest vessels is proposed. Line detection and perceptual organization techniques are combined in a multi-scale scheme. Small vessels are reconstructed from the perceptual-based approach via tracking and pixel painting.

The segmentation was validated in a high resolution fundus image database including healthy and diabetic subjects using pixel-based as well as perceptual-based measures. The proposed method achieves 85.06% sensitivity rate, while the original multi-scale line detection method achieves 81.06% sensitivity rate for the corresponding images ($p < 0.05$). The improvement in the sensitivity rate for the database is 6.47% when only the smallest vessels are considered ($p < 0.05$). For the perceptual-based measure, the proposed method improves the detection of the vasculature by 7.8% against the original multi-scale line detection method ($p < 0.05$).

1. Introduction

Diabetic retinopathy (DR) is a public health issue with major implications for the eyesight of the affected population. In diabetes, blood glucose levels are higher than in the normal case. Prolonged hyperglycemia has a deteriorating effect upon retinal circulation that triggers many pathophysiological changes, of which the most important are related to the smallest vessels (Archer, 1999). If it is left untreated, this disease can progressively lead to blindness. The prevalence of blindness caused by diabetic retinopathy is 4.8% out of the total of 37 million cases throughout the world (World Health Organisation, 2005). The introduction of screening programs has helped to reduce the incidence of diabetic retinopathy (Scanlon, 2008). However, a major drawback of current screening techniques is their reliance upon visual inspection of fundus images by physicians. Computerized systems assisting retinal specialists' decisions, known as computer-aided diagnosis (CAD) systems, would allow screening programs to cover larger populations and reduce the disease's impact though earlier detection.

Segmentation algorithms can highlight the most important retinal structures and guide the physicians to quickly identify signs of DR. A major category of CAD systems for DR rely on pre-segmenting structures other than dark or bright lesions, i.e. the vessels and the optic disc, before proceeding to the lesion analysis stage (Abràmoff et al., 2010; Antal and Hajdu, 2013; Sopharak et al., 2013). However, the vasculature tends to be coarsely extracted and thus among the candidate lesions, many small and possibly fragmented vessels exist. Consequently, the system can mistake these fragments for dark lesions because they share common features with the latter (Sopharak et al., 2013). To this end, the focus of this work is the segmentation of small retinal vessels that can potentially improve an automatic diagnosis system for retinopathies.

Despite the substantial amount of research on retinal vessel segmentation (Abràmoff et al., 2010; Fraz et al., 2012; Winder et al., 2009), the algorithms that have been proposed focus on the larger vessels while omitting the smaller ones. Since the latter are thinner, with lower contrast than average vessels and possible discontinuities, they need special handling in a separate step. In addition, the performance assessment for these methods is based solely on global binary classification metrics (Fawcett, 2006) that

* Corresponding author.
E-mail addresses: argyrios.christodoulidis@polymtl.ca (A. Christodoulidis), thomas.hurtut@polymtl.ca (T. Hurtut), hbentahar@diagnos.ca (H.B. Tahar), farida.cheriet@polymtl.ca (F. Cheriet).

http://dx.doi.org/10.1016/j.compmedimag.2016.06.001
0895-6111/

do not take into account the fact that the smallest vessels represent approximately 10% of the total surface area of the vessel network (Niemeijer et al., 2004).

To our knowledge, only a limited number of existing studies concentrate on small vessel segmentation in high resolution images, with a reliable evaluation against manual segmentation. The available methods can be grouped in two categories: supervised machine learning and intensity-based approaches. Machine learning approaches (Kovács and Hajdu, 2011; Sofka and Stewart, 2006) require an a priori model trained on a labeled set. This set is usually small because the labeling task is laborious, and not representative of the variability observed in the real world. Even though small and large vessels share common morphological attributes such as elongateness or darker appearance than their background, different kinds of features must be selected in order to perform classification, and the detected features training should be confined to the smallest vessel category. This was partially met in Kovács and Hajdu (2011) where the training phase was divided into large and small vessels; however, during the classification stage both categories were merged and the segmentation performance was assessed globally, without examining the small vessels separately. This method has been used for vasculature subtraction prior to microaneurysm detection in Antal and Hajdu (2013). Sofka and Stewart (2006) proposed a new vesselness measure that is robust to lesions away from the vasculature. The method was validated on small and poorly contrasted vessels, but its output is only the vessel centerlines, which is inadequate if the result must be directly used for lesion detection in a CAD system. In that case, full vasculature segmentation must be performed before any subsequent lesion detection stage.

In the intensity-based methods, a template is applied across the image in the spatial domain in order to segment either full vessels (Zhang et al., 2009; Allen et al., 2011; Oost et al., 2010) or centerlines (You et al., 2011; Mendonca and Campilho, 2006; Xu and Luo, 2010). Although the method in Zhang et al. (2009) proposes an approach to suppress lesion detection away from the vasculature, it is assumed that the small vessels are piecewise straight (Zhang et al., 2009; You et al., 2011; Allen et al., 2011). By contrast, curve fitting was used in Oost et al. (2010) to approximate small vessels' curvatures and to reconnect pre-segmented fragments into larger segments. Another shortcoming of the available intensity-based methods is that the segmentation results are compared quantitatively in all the available categories of vessels without making a distinction between the small and large vasculatures. Moreover, a limited number of methods have been used in the context of CAD systems for lesion detection (Zhang et al., 2010; García et al., 2008). A hybrid method was proposed by Xu and Luo (2010). Small vessel centerlines are first discriminated from noise by classification, then they are merged into larger segments based on local tracking.

Several multi-scale vesselness pipelines have also been proposed to segment the vessels in high resolution images and tested on Erlangen database (Odstrcilik et al., 2013). These works are based on the application of Frangi's model (Frangi et al., 1998) or Gaussian matched filtering (Odstrcilik et al., 2013). These studies modified Frangi's method either by incorporating a pre/post-processing step (Yu et al., 2012; Annunziata et al., 2015; Budai et al., 2013) or by adapting the basic model to handle crossings and bifurcations (Hannink et al., 2014). A Hessian matrix is approximated by partial second-order Gaussian derivatives, but this has the effect of filtering out the small vessels, magnifying the background noise and reconnecting the noise to the detected vessels. Equally, matched filtering is not sensitive enough for the smallest vessels because the template used is too general to represent vessels that demonstrate high curvatures over short distances and limited numbers of cross-sectional vessel entries.

Line detectors are a recent family of methods for identifying linear structures. These were first used to identify linear structures in mammographic data (Zwiggelaar et al., 2004), and recently adapted to the task of detecting retinal vessels (Ricci and Perfetti, 2007). Recently, Nguyen et al. (2013) introduced the multi-scale line detector (MSLD) to accommodate for the various vessel sizes that appear in the retina. This method utilizes a straight sampling segment of variable length that, when rotated around a central pixel, gives the line response. The values of a sliding averaging window are subtracted from the line response to compensate for uneven background illumination. Finally, the line detector responses are thresholded in order to segment the vessels.

The MSLD method efficiently segments large to medium-sized vessels, however several factors hamper the segmentation of the smallest vessels. First, the size of the sliding window is set globally. In the case of a small vessel the window is too large and there are not enough vessel pixels to compensate for the presence of many background values. Second, the small vessels usually appear near the macula and the periphery where the image intensity and contrast are low; as a result, the line response is low and these vessels are fragmented.

In this context, it is useful to consider the tensor voting framework (TVF) developed by Medioni and Kang (2004) and Medioni et al. (2000). This is a bottom-up approach for organizing neighborhood information based on perceptual principles from Gestalt theory used to segment thin fragmented structures. It formalizes the observation that high-level perceptual structures can be formed by grouping individual lower level structures. Tensor voting consists of two components, given as two criteria: (1) the proximity criterion, namely pixels belonging to the same structure are close to each other; and (2) the continuity criterion, namely grouped structures have constant curvature. The process starts by initializing the directions of the structures via a method to estimate orientations, such as ball voting or gradients. Stick voting is then applied to refine the directions, propagate and reconnect the information between the structures, and compute their saliency. Saliency measures the certainty about the existence of a line in a particular location, and it demonstrates a local maximum along the principal direction of the structure. A scale parameter σ_{TVF} controls the range of influence of the voting.

This method has been successfully applied to a range of biomedical imaging applications (Risser et al., 2008; Loss et al., 2011; Yigitsoy and Navab, 2013; Zuluaga et al., 2015; Maggiori et al., 2014). These include filling microvascular discontinuities (Risser et al., 2008) based on interpolating the width between reconnected segments, and detecting important cellular structures (Loss et al., 2011) by iterating a TVF along a range of scales. A variety of algorithmic improvements to the basic framework were proposed in the literature; these were comprehensively reviewed by Maggiori et al. (2014).

In retinal imaging applications, single scale TVF has been applied to either extract full vessels (Park et al., 2007) or segment centerlines (Leng et al., 2011). No particular care was given to the problem of full small vessel segmentation in either of these studies. The work in Park et al. (2007) focused on identifying the optic disc as the converging point of the largest vessels and used a direct thresholding of the saliency map. In this manner, disconnected curvilinear structures belonging to lesions or irrelevant structures can get high saliency values, so thresholding is likely to remove small vessels without suppressing the false positives. The authors in Leng et al. (2011) extracted the centerlines based on steerable tensor voting (Franken et al., 2006). Rather than applying the framework for extracting the saliency of the structures globally, they used the method to link pre-identified edges. Moreover, in a preliminary phase, edge points were discarded based on a global length criterion. But the setting of a non-adaptive global length

parameter is not effective for retaining the smallest vessels and causes a translation in the final centerline location.

This study will present a methodology based on a multi-scale tensor voting framework (MTVF) combined with multi-scale line detection (MSLD) in order to overcome the limitations of the line detection in handling the smallest vessels. Multi-scale tensor voting enables the reconnection of fragmented vessels at variable distances. To a certain degree, it can reconnect non-straight segments and it should also be robust in the presence of background noise. Moreover, non-vascular disconnected responses will be suppressed in the postprocessing phase. Finally, simple inpainting along small vessel centerlines will be applied in order to extract full vessels, thereby allowing TVF to be used for small vessel segmentation.

The rest of the paper is structured as follows: Section 2 describes the proposed methodology for retinal vessel segmentation; in Section 3 the image database as well as the experiments and results are presented; Section 4 discusses the findings; finally, Section 5 concludes this paper.

2. Methodology

The diagram in Fig. 1 demonstrates the proposed pipeline for retinal vessel segmentation. Images are first preprocessed, then vessels are segmented by the application of MSLD. Adaptive thresholding is then applied to extract the large and medium-sized vessels. Next, the MTVF is applied to the segmented image in order to identify and reconnect the smallest vessels. Moreover, starting from the largest vessels, a local tracking procedure merges the information from different scales. In parallel, the smallest vessels are reconstructed to their whole extent. Finally, the images are postprocessed to suppress the background noise.

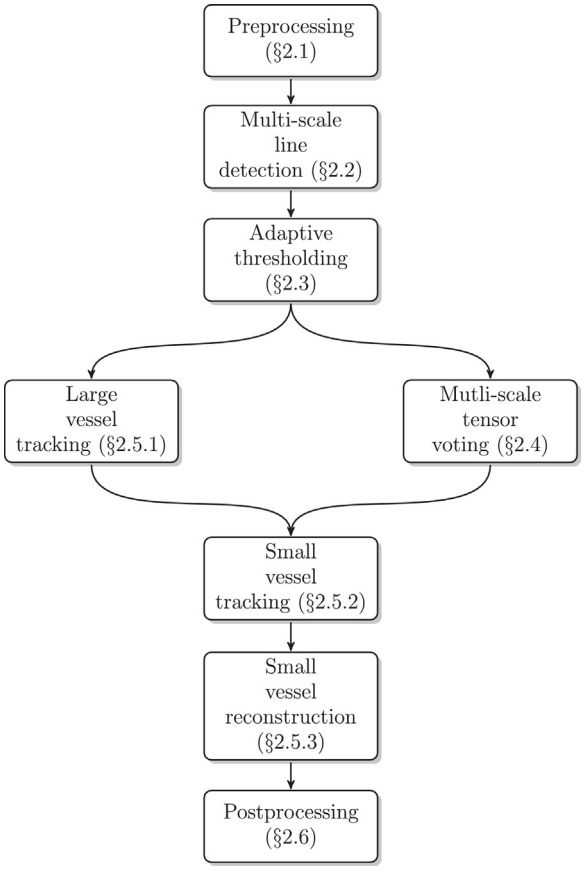


Fig. 1. Block diagram of the proposed pipeline for retinal vessel segmentation.

2.1. Preprocessing

The retinal images are preprocessed to correct for non-uniform lighting, to normalize the image contrast and to suppress the background noise. The green channel of the input image I_{green} is used as it provides the highest contrast between the vessels and the background. The local illumination and contrast correction method proposed by Foracchia et al. (2005) was applied on the images to get $I_{enhanced}$. Moreover, the dual-tree complex wavelet transform (DT-CWT) (Kingsbury, 1998) is applied to $I_{enhanced}$ in order to filter out the noise. Filtering is applied (Donoho, 1995) with the parameter t experimentally selected as described in Section 3.2 to suppress the background noise without distorting the information on the smallest vessels. High values not only can remove the noise, but they can also filter out the smallest vessels. For example, Fig. 2 shows the effect of filtering on the final segmentation result using three different filtering levels $t = 0$, $t = 0.015$ and $t = 2$. In this example, the region is first preprocessed, MSLD is then applied to enhance the vessels and simple thresholding isolates the vessels from the background. The preprocessing step allows us to remove most of the background noise; however, the filtering parameters must be carefully selected to avoid removing the smallest vessels.

2.2. Large to medium caliber vessel segmentation

Following the preprocessing step, MSLD is applied to the images as proposed in Nguyen et al. (2013). The length of the rotating line W, which implements the multi-scale component of the approach, varies from a minimum to a maximum value and the unweighted average response across the scales is computed in order to enhance the vessels. The response is thresholded to distinguish the vessels from the background. In our case, we adjust parameter W according to the vessel resolution (Section 3.2). Fig. 3 shows an example of applying the original MSLD method on an image.

2.3. Adaptive thresholding

The resulting line response is then fed into an adaptive thresholding step to isolate the vessels. In the original MSLD approach, simple thresholding with a fixed value for all the images in the database was proposed (Nguyen et al., 2013). But since there is variability in the acquisition process, the background of the retinal images does not show a uniform distribution across our database. Thus, the choice of a single fixed threshold is inadequate when the whole database is processed. Fig. 4 shows an example where simple thresholding with a fixed value ($T = 120$) is used to segment two healthy images in the database. The given global threshold only partially segments the vessels in the second case (Fig. 4d). Therefore, we apply adaptive thresholding in order to extract the vessels more reliably.

The histogram of the MSLD responses has a specific form, resembling a normal distribution with a distinguishable peak corresponding to the largest anatomical structure, namely the fundus background (Fig. 4a and c). In addition, there is an elevated region or a valley to the right of the peak which corresponds to the foreground. In order to choose an optimal global threshold we can follow the rule of thumb proposed in van Antwerpen et al. (1986). This work suggests finding the central location of the background distribution, i.e. its mean, and adding to that the full width of the background at half height. In order to apply this, we fit a simple Gaussian to the histogram. To obtain the adaptive thresholding value, we modify the rule of thumb slightly by varying the distance from the mean ($\mu_{Gaussian}$). Thus, the threshold value (T) is given by Eq. (1):

$$T = \lfloor \mu_{Gaussian} \rfloor + \alpha \lfloor \sigma_{Gaussian} \rfloor, \quad (1)$$

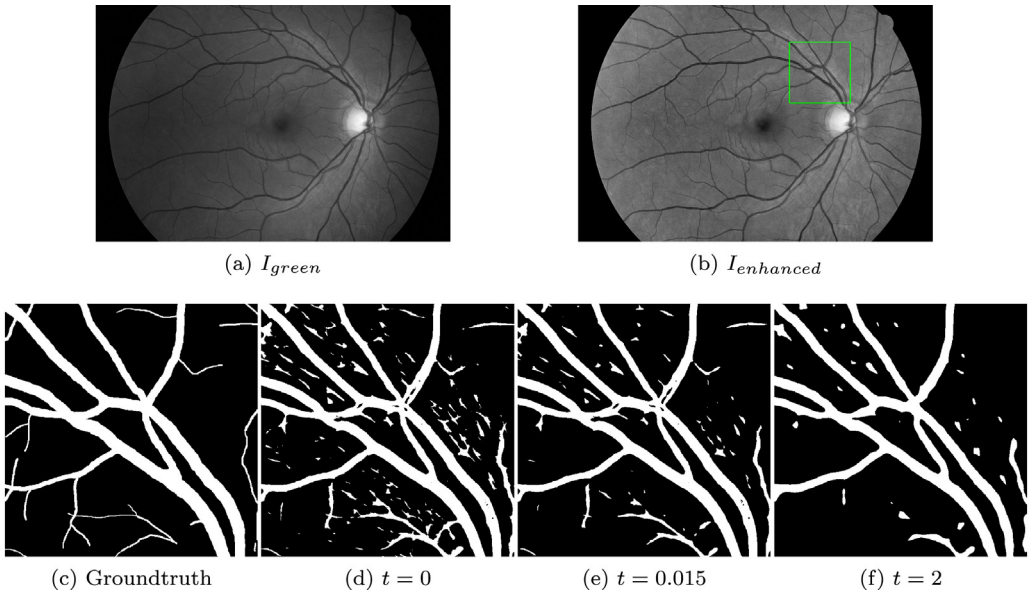


Fig. 2. Effect of the preprocessing steps on the final segmentation result in Image_01_h. (a) input image I_{green}, (b) lighting-corrected and contrast-enhanced image $I_{enhanced}$, (c) groundtruth segmentation corresponding to the green square inset in (b), and (d)–(f) effect of wavelet filtering on the final segmentation result at three different levels: $t = 0$, $t = 0.015$, $t = 2$.

where $\mu_{Gaussian}$ and $\sigma_{Gaussian}$ are the mean and standard deviation of the fitted Gaussian, respectively. The value for weighting factor α is experimentally determined as in Section 3.2. Fig. 4e and f demonstrates the adaptive thresholding result for the two healthy images. For the second image, the method is much more effective at segmenting the vessels than non-adaptive thresholding.

2.4. Small vessel centerline extraction

Following the thresholding step, many smaller vessels remain fragmented, and a MTVF is used to reconnect them. The basic component of single scale tensor voting, namely the second order symmetric tensor, is used to capture the information provided by points as well as curves according to the following expression:

$$\mathbf{T} = \begin{bmatrix} a_{xx} & a_{xy} \\ a_{yx} & a_{yy} \end{bmatrix} = \lambda_1 \widehat{\mathbf{e}}_1 \widehat{\mathbf{e}}_1^T + \lambda_2 \widehat{\mathbf{e}}_2 \widehat{\mathbf{e}}_2^T$$

$$= (\lambda_1 - \lambda_2) \widehat{\mathbf{e}}_1 \widehat{\mathbf{e}}_1^T + \lambda_2 (\widehat{\mathbf{e}}_1 \widehat{\mathbf{e}}_1^T + \widehat{\mathbf{e}}_2 \widehat{\mathbf{e}}_2^T), \quad (2)$$

where entries a represent the a priori information about the orientation of the tensor. In our case, we do not assume any specific direction thus we initialize the tensor with the identity matrix. Additionally, λ_1 (resp. λ_2) is the largest (resp. smallest) eigenvalue, and $\mathbf{e_1}$ (resp. $\mathbf{e_2}$) is the corresponding eigenvector that is normal (resp. parallel) to the local structure. If we set $\lambda_2 = 0$ and $\lambda_1 = \lambda_2$ we get the measures of saliency and ballness, respectively. Ball voting is used to initialize the direction of the structures among the tensors. Since the orientation information is not available for the ball voting step, the points are encoded with the second order identity tensor $\mathbf{T}$. Stick voting follows, where a symmetric fan-shaped voting field centered and aligned with the principal direction of the voter propagates information to every neighboring pixel around the tensor. The voting distance depends on the scale factor σ_{TVF} according to the following equation (Medioni and Kang, 2004, Eq. (5.2)):

$$c = \frac{-16 \log(0.1)(\sigma_{TVF} - 1)}{\pi^2}, \quad (3)$$

Interpretation follows voting; at each pixel, the sum of all votes is analyzed by eigen-decomposition of matrix $\mathbf{T}$. The saliency measure gives a maximum perpendicular to the local stickness direction, so centerlines can be extracted by non-maximum suppression.

We use MTVF to group the fragmented smallest vessels into single longer vessels. This is justifiable if we consider that the gap between the vessels can vary and that higher voting scales can capture neighborhood information at a greater distance. The tensor voting scale σ_{TVF} is varied between a maximum and a minimum value with a predefined step which is determined experimentally (Section 3.2). The same adaptive thresholded image is used as input across all the scales. Here we apply the voting separately for each scale and then we fuse the information by computing the union of the centerlines across all the scales.

The framework is applied without saliency thresholding (Park et al., 2007) or vote weighting (Leng et al., 2011). Indeed, directly

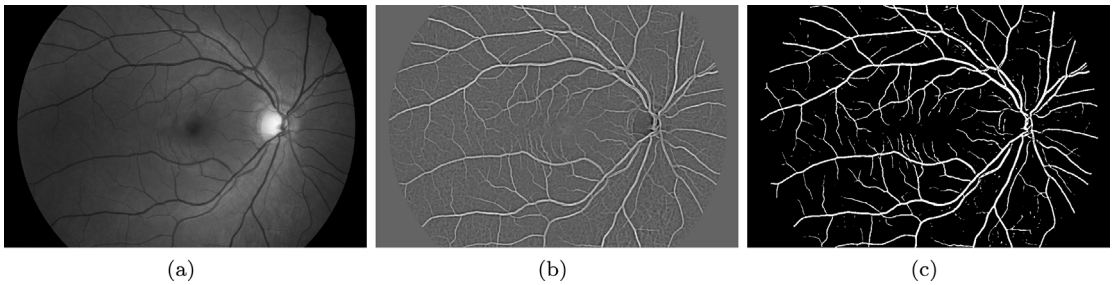


Fig. 3. Example of the line detection approach. (a) Input image, (b) MSLD response, and (c) extracted vessels based on thresholding.

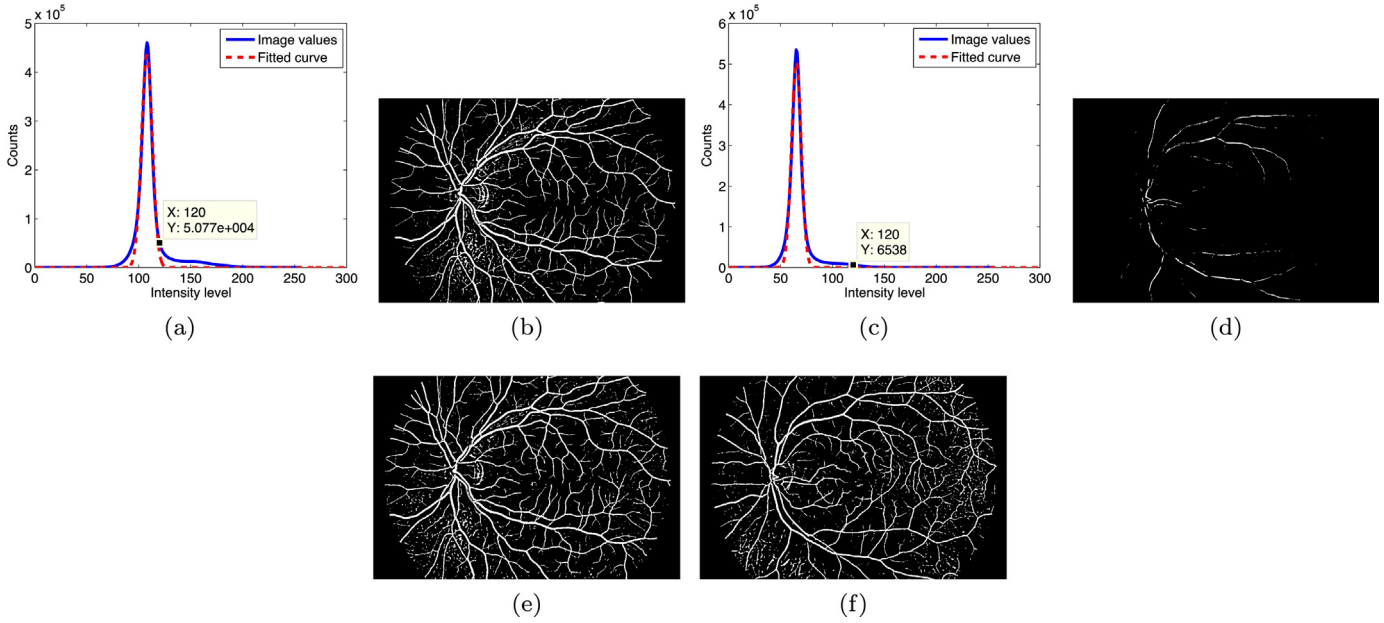


Fig. 4. Example of applying fixed thresholding versus adaptive thresholding in two images of healthy subjects from Erlangen database. (a) histogram of the normalized line response values (X-axis) against the number of counts (Y-axis) of the 2nd healthy subject (blue line) together with the fitted Gaussian (red dashed line), (b) segmentation result of the 2nd healthy subject using a fixed threshold value at $T = 120$, (c) histogram of the normalized line response values of the 10th healthy subject (blue line) together with the fitted Gaussian (red dashed line), (d) segmentation result of the 10th healthy subject using a fixed threshold value at $T = 120$, and (e)–(f) segmentation result of the 2nd and 10th healthy subjects using adaptive thresholding. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

thresholding the saliency map can lead to suppression of important information from the small vessels. Therefore, we instead extract the local maximum from the saliency map without thresholding. Moreover, the grayscale contrast as well as the MSLD response between the smallest vessels and the background is very low. Additionally, there are cases where the background has a higher response than the structure of interest, so weighting the votes can lead to amplifying the saliency on the false positives.

The framework as it was originally proposed by Medioni and Kang (2004) and Medioni et al. (2000) is applied here. In this framework, ball voting takes into account directional information at a greater distance than local gradient (Leng et al., 2011), and the extent of influence is directly determined by the selected scale in the framework. Even though the implementation based on steerable filtering is superior, in terms of computational time, it is an approximation to the original framework (Maggiori et al., 2014). Fig. 5 shows an example of how the method is applied at scales $\sigma_{MTVF} = 5{:}50$ to reconnect small fragmented vessels in a challenging isolated image region, in the periphery. The same adaptive thresholded input image (Fig. 5a) is used for each voting scale. Non-maximum suppression on the saliency map (Fig. 5b–d) extracts the centerlines that, combined into a single map, are used at the reconstruction phase (Fig. 5h). In this example, higher voting scales manage to reconnect the fragmented vessels (Fig. 5e–g).

2.5. Small vessel reconstruction

We choose a tracking approach to combine the previously segmented large vessels with the small vessel centerlines obtained by the MTVF. The advantage of this approach is that we can examine individual vessels as we proceed and obtain useful information that can facilitate their reconstruction. Furthermore, as we track the smallest vessels, we can combine the fused multi-scale centerlines to reconstruct single centerlines. The method uses two centerline maps: (1) the skeletonized image of the large vessels (Section 2.3), and (2) the fused small vessels centerline map (Section 2.4).

2.5.1. Large vessel tracking

The process is initialized by identifying vascular seed points around the optic disc. Vessels are highly interconnected there, forming a tree-like structure that emerges from the optic disc. Therefore, the chance of missing a larger vessel by following all the branches from their root is low. If for example we start the tracking from a distant large vessel that is fragmented, we will stop when we reach a gap, thus stopping the process before reaching any other vessels. On the other hand, by starting at the optic disc we are able to continue tracking by selecting a large vessel that is next to the one that we stopped tracking. Furthermore, initiating tracking from the optic disc allows us to only keep small vessel centerlines and not reconnect background noise. The seeds are the intersection points between a circle with a center at coordinates $C(x, y)$ and radius R with the skeletonized image of the large vessels. The circle parameters are directly dependent on the optic disc size and location; they were manually identified by an expert for all the images in the database.

A depth first based tracking approach is applied on the large vessel skeleton where individual vessels are tracked down to their terminal points. In this way, tracking prioritizes reaching the vascular terminal points, and subsequent initiation of the small vessel tracking, over exploring distant branched vessels. Sequential thinning is applied to obtain the skeleton, thereby suppressing small and noisy branched vessels caused by incomplete and noisy segmentation at the vessel tip.

From there, the next step is the local searching around the end-points reached by tracking, in a 3-by-3 neighborhood A, for a small vessel seed that corresponds to a TVF centerline point. This centerline point will initiate the small vessel tracking process. In that way, we confine our analysis to areas of interest around the large vessel end-points without trying to reconnect background noise. The neighborhood size is crucial for the initialization of the second tracking phase, because it links the large and small vessel segmentation processes. A search neighborhood larger than three pixels risks tracking disconnected noise.

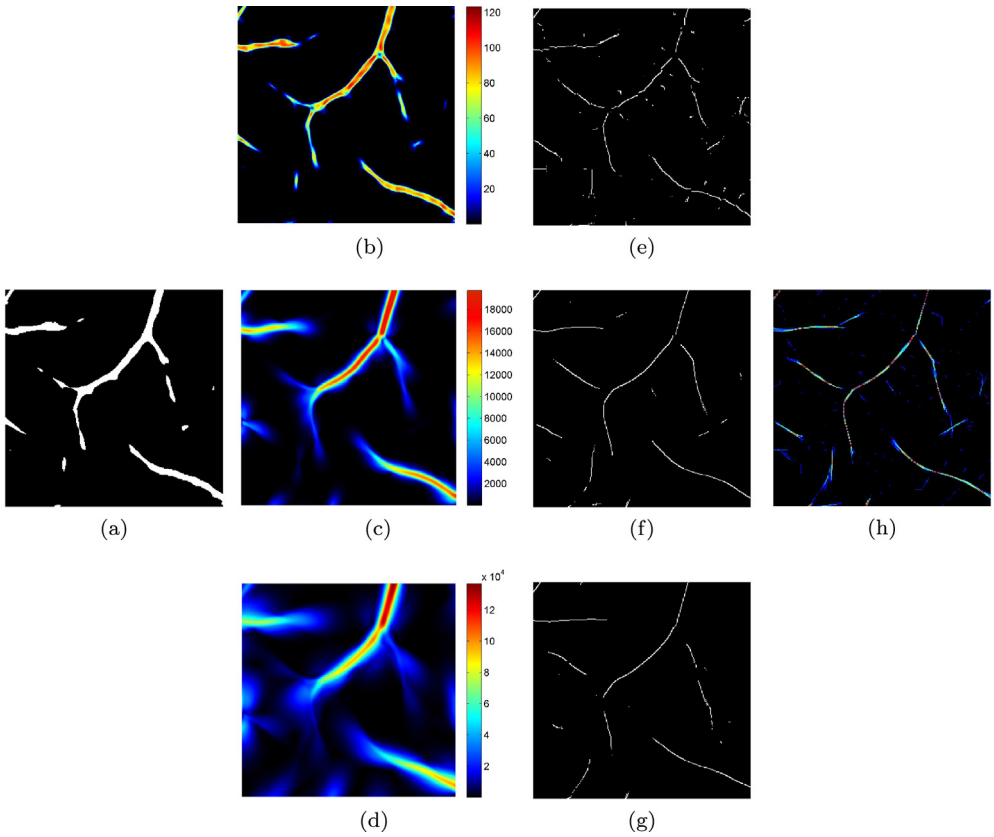


Fig. 5. Small vessel reconnection example. (a) Input image, (b)–(d) saliency maps obtained by small ($\sigma_{TVF} = 5$), medium ($\sigma_{TVF} = 25$) and large ($\sigma_{TVF} = 50$) voting scales, (e)–(g) centerline maps corresponding to each tensor voting scale, and (h) centerlines union across $\sigma_{TVF} = 5{:}50$ where each color represents a different voting scale. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.5.2. Small vessel tracking

We proceed to the small vessel tracking from the combined TVF centerlines across all the available scales. This is an iterative process that loops between three steps: (1) the *searching step*, (2) the *estimation step*, and (3) the *refinement step*. Tracking advances by one pixel at each iteration, and terminates when no new centerline is returned. The search direction $\mathbf{V_{dir}}$ at a given iteration is computed based on the local TVF eigenvector $\mathbf{e_2}$. Fig. 6 shows a small vessel tracking example. More specifically:

1. *Searching step:* Centerlines (blue stars, Fig. 6a) are searched along a line that is placed one step from the previous iteration and oriented perpendicularly to $\mathbf{V_{dir}}$ (red arrow, Fig. 6a).
2. *Estimation step:* Starting from the center of the perpendicular line, the first centerline point that is found is the estimation of the final centerline position (blue star, Fig. 6b).
3. *Refinement step:* We can have multiple centerlines slightly shifted from each other (by 1–2 pixels) that correspond to the same vessel. To merge the multi-scale information into a single centerline point we use the local TVF eigenvector $\mathbf{e_1}$ to obtain perpendicular profile P (yellow line, Fig. 6c). The position that minimizes the sum of squared distances between itself and all the other centerlines along the profile is kept as the refined centerline position (green star, Fig. 6c).

   Direction $\mathbf{V_{dir}}$ (red arrow, Fig. 6a and d) is computed by averaging of the $\mathbf{e_2}$ centerlines along P. Since the directions represent angles, circular averaging is used (Berens, 2009). The resulting directions from eigenvector decomposition can be flipped. If $\mathbf{e_2}$ points in the wrong direction it can guide tracking backwards, thus we compare it with the direction at the previous step and correct for this ambiguity when needed.

2.5.3. Small vessel reconstruction

The previous work in the literature has focused on the extraction of centerlines, while in our context we are interested in obtaining full vessels. In parallel to tracking, full vessel reconstruction is applied by pixel painting along the length of the profile P. The direction as well as the center of P are determined during the refinement stage of the small vessel tracking process (Section 2.5.2). The large vessels' diameters are tracked over the whole skeleton of the larger vessels by minimizing the length of a rotating line, similarly to Nguyen et al. (2013). Then, the reconstructed diameter D_{small} of a small vessel is equal to the tracked diameter $\mu_{diameter}(large\ vessels)$ of the preceding large vessel, averaged over several skeleton points at the end of that vessel. Averaging of the diameter is necessary because the large vessel gradually narrows towards its tip, so we have to normalize $\mu_{diameter}(large\ vessels)$ by considering a number N of preceding points, N being experimentally determined (Section 3.2). The reconstruction process is restricted to the smallest vessels, i.e. those of diameter $D < 8$ pixels (see Section 3.1). As a final step, morphological closing with a small circular element is applied to fill holes in the reconstructed vessel caused by the inpainting of curved vessels along straight lines. Fig. 7 provides an example of small vessel reconstruction.

2.6. Postprocessing

The MSLD response is sensitive to background lesions and artifacts, so non-vascular disconnected components remain after the adaptive thresholding stage. Additionally, branched vessels may remain disconnected after tracking. Junctions are high order morphological features that represent the separation of a main vessel into smaller branches. Both MSLD and TVF are inadequate to

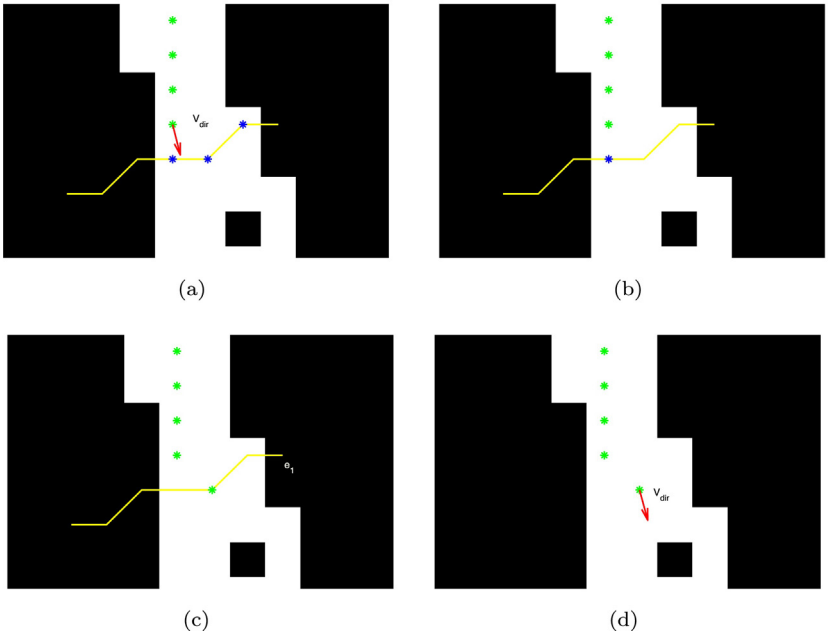


Fig. 6. Small vessel tracking example. (a) Searching of centerlines (blue stars) perpendicularly to $\mathbf{V_{dir}}$ (red arrow), (b) estimation of the final centerline position (blue star), (c) refinement of the final centerline position (green star) by searching along profile P (yellow line), and (d) $\mathbf{V_{dir}}$ search direction for next tracking iteration. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

approximate the bifurcation pattern when it comes to enhancing or segmenting vascular regions at less prominent junctions. MSLD applies a straight line template, while there is a high degree of directional uncertainty that cannot be estimated by second order tensors in the TVF.

The final step in our pipeline is the morphological cleaning of background artifacts from the image, together with the inclusion of disconnected vessels. All the connected components (CC) that are disconnected from the main vasculature are sequentially analyzed. If they are larger than a specific size ($Area_{CC} > Area_{value}$), elongated enough ($Elongation_{CC} > X$) and not very solid ($Solidity_{CC} < X$), then they belong to vessels and thus they are retained from the thresholded image. Elongation or eccentricity of an ellipse (ε) is defined as:

$$\varepsilon = 2\frac{\sqrt{\left(\frac{Length_{MajorAxis}}{2}\right)^2 - \left(\frac{Length_{MinorAxis}}{2}\right)^2}}{Length_{MinorAxis}}, \quad (4)$$

where $Length_{MajorAxis}$ and $Length_{MinorAxis}$ are the major and minor axis lengths of an ellipse, respectively (Haralick and Shapiro, 1992). Solidity is defined as the ratio of the area over the convex hull area. Conservative values for the parameters used at this step were experimentally determined as described in Section 3.2. This allowed us to clean the images while retaining important vascular structures. Fig. 8 shows an example where the postprocessing step for background noise suppression is applied in a highlighted area of an image.

3. Experiments

3.1. Image database

Our approach was evaluated on high resolution fundus images from the Erlangen database (Odstrcilik et al., 2013). The database contains three categories of images: healthy subjects (H), diabetic retinopathy patients (DR), and glaucomatous patients (G). Each category comprises 15 images acquired using a CANON CF-60UVi camera (Budai et al., 2011) with fixed resolution of 3504×2336 pixels and pixel size of $6.05 \times 6.46\,\mu m$. Manual segmentations of the vessels are provided for the whole database. In this study, all the vessels of small width were considered as small vessels. This definition of size is invariant to location and branching order, however it depends on the acquisition process, the field of view and the image resolution. For this particular high resolution database, we qualitatively selected the vessel diameter value that isolates most of the terminal, non-fragmented vessels, without including vessels that might belong to the larger vascular structures, such as the major arcades. Everything below 8 pixels in diameter was thus considered as a small vessel. This value yielded a good compromise between the aims of identifying complete terminal vessels

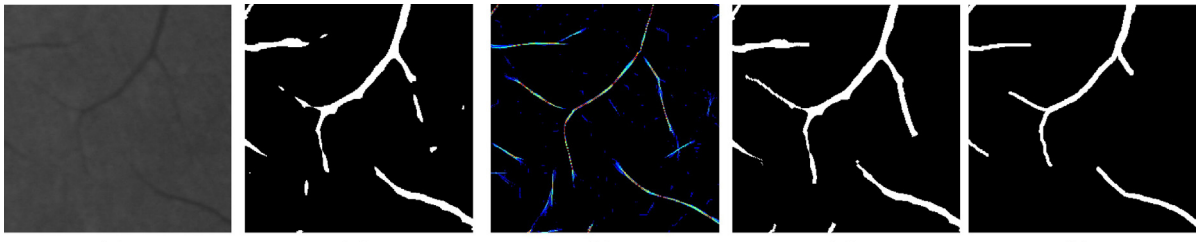


Fig. 7. Small vessel reconstruction example. (a) Input grayscale image, (b) simple thresholded image, (c) centerline union, (d) reconstructed vessels, and (e) reference. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

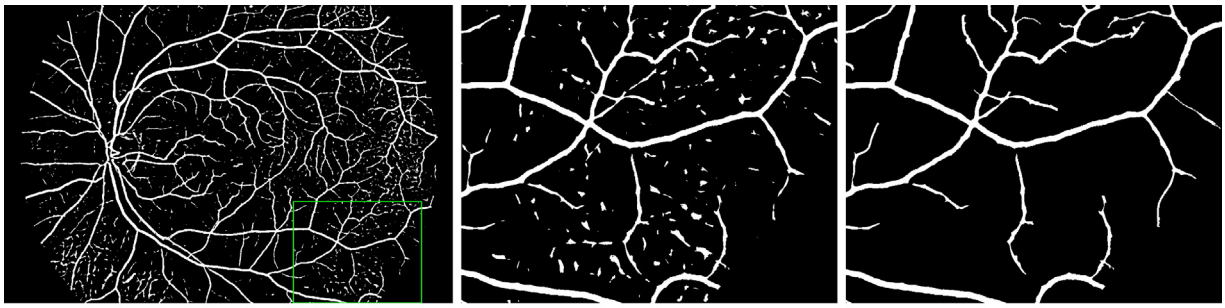


Fig. 8. Postprocessing example. (a) Adaptive thresholded input image, (b) zoomed area corresponding to inset in (a), and (c) same image area after postprocessing (cleaning).

and excluding larger vessels. The rest of the vasculature, with diameter above the 8-pixel threshold, was considered as belonging to the large and medium-sized categories of vessels.

3.2. Parameter setting

The vessel segmentation performance was evaluated at two levels: (1) using the full vasculature, and (2) restricting the analysis to the smallest vessels. The glaucomatous cases in our database were not severe, so they show deterioration on the optic disc but none for the vascular network. Hence, this group of images was used as a training set to tune the method's parameters, while the rest of the database was used for validation purposes. We also compared our vasculature segmentation results to recent methods published in the literature. Before proceeding to the evaluation stage, we examined different parameters values and how their values affected the segmentation performance in the training set. Table 1 lists the method's main parameters.

In order to examine which value t from the wavelet filtering step was most appropriate for our method, we varied this parameter between several representative values: $t = 0$, 0.015, 0.03, 0.06, 0.12, 0.5, and 2. The filtering value that gave the highest improvement in the sensitivity rate against the original MSLD method (Nguyen et al., 2013), at the same noise level, as chosen as the optimum filtering value. The changes in sensitivity rate for the above t values were respectively 1.13%, 1.15%, 0.76%, −0.66%, −2.90%, −7.28% and −8.02%. Therefore, we choose $t = 0.015$ for subsequent experimentation as it gave the highest increase.

For the MSLD step, we considered a maximum available line length of $W = 40$ pixels, which was twice the diameter of an average vessel in the images. Moreover, we determined the adaptive thresholding factor α by testing different values in a range and evaluating the result using the CAL metric (Section 3.3.1). Given the observed mixed distributions of the MSLD responses in Fig. 4, we had to try high sigma values ($\alpha < 3$) to investigate the response of the segmentation at these levels. Even though the maximum was achieved at $\alpha = 2.2$, this value was not optimal because, together with the background noise, small vessels were removed from the result. Retaining the fragmented small vessels, which is necessary for vessel reconnection, was more important than cleaning the background, thus we select a value in the middle of the range, namely $\alpha = 1.55$, as the optimum for segmentation purposes. Fig. 10c shows the effect of varying factor α on the CAL metric.

Next, the effect of using single-scale tensor voting instead of MTVF on the small vessel reconstruction was evaluated. We tried different values for the scale parameter in the range $\sigma_{TVF} = 5{:}50$, and measured the sensitivity (Fig. 10a) and specificity (Fig. 10b) rates across the training set. As we increased the MTVF scale range, more vessels were included in the final result. The inclusion of high voting scales can help in the direction of letting the algorithm to reconnect fragments in a greater distance than if only smaller voting scales are used, however at the expense of a reduction in the specificity rate. Beyond a certain point, though, the inclusion of larger scales not only fails to improve the segmentation performance, but in fact causes more noise to be reconnected to the vessel network. Since we want to segment as many vessels as possible, we used the range of values from the smallest scale to the scale where there is stabilization in the variation of the sensitivity rate ($\sigma_{TVF} = 5{:}50$). Also, it is clear from Fig. 10a and b that MTVF had better performance measures than single scale TVF.

We then evaluated the small vessel reconstruction parameter N to determine the value giving the best improvement in sensitivity rate against the original MSLD method (Nguyen et al., 2013) at the same noise level. The changes in sensitivity rate, as we increased N from 1 to 8 (by steps of 1), were respectively 1.13%, 1.14%, 1.14%, 1.15%, 1.12%, 1.13%, 1.12%, and 1.12%. Therefore, we choose $N = 4$ for

Table 1
Parameters of the proposed approach and the selected values.

Parameter (section)	Function	Selected value
t (§3.1)	Wavelet filtering	$t = 0.015$
W (§3.2)	MSLD (Nguyen et al., 2013)	$W = 40$
α (§3.3)	Adaptive thresholding factor	$\alpha = 1.55$
σ_{TVF} (§3.4)	Multi-scale tensor voting range	$\sigma_{TVF} = 5{:}5{:}50$
N (§3.5)	Large vessel skeleton pixels for reconstruction	$N = 4$
$Area_{value}$, X (§3.6)	Postprocessing	$Area_{value} > 500$, $X = 0.5$

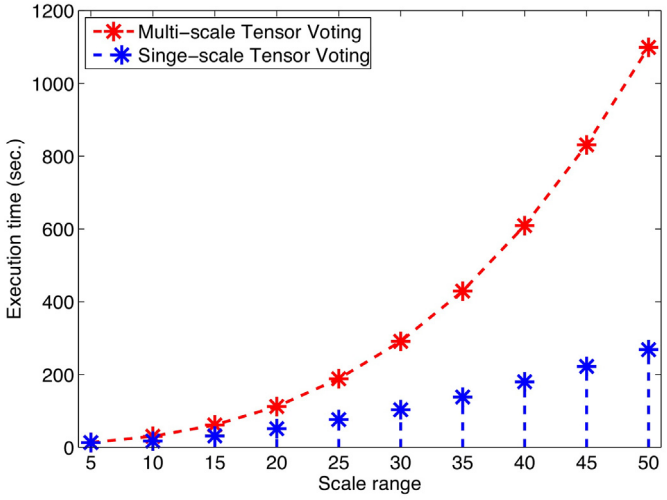


Fig. 9. Single (blue line) and multi (red line) scale tensor voting execution time for a typical retinal image across scales $\sigma_{TVF} = 5{:}5{:}50$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

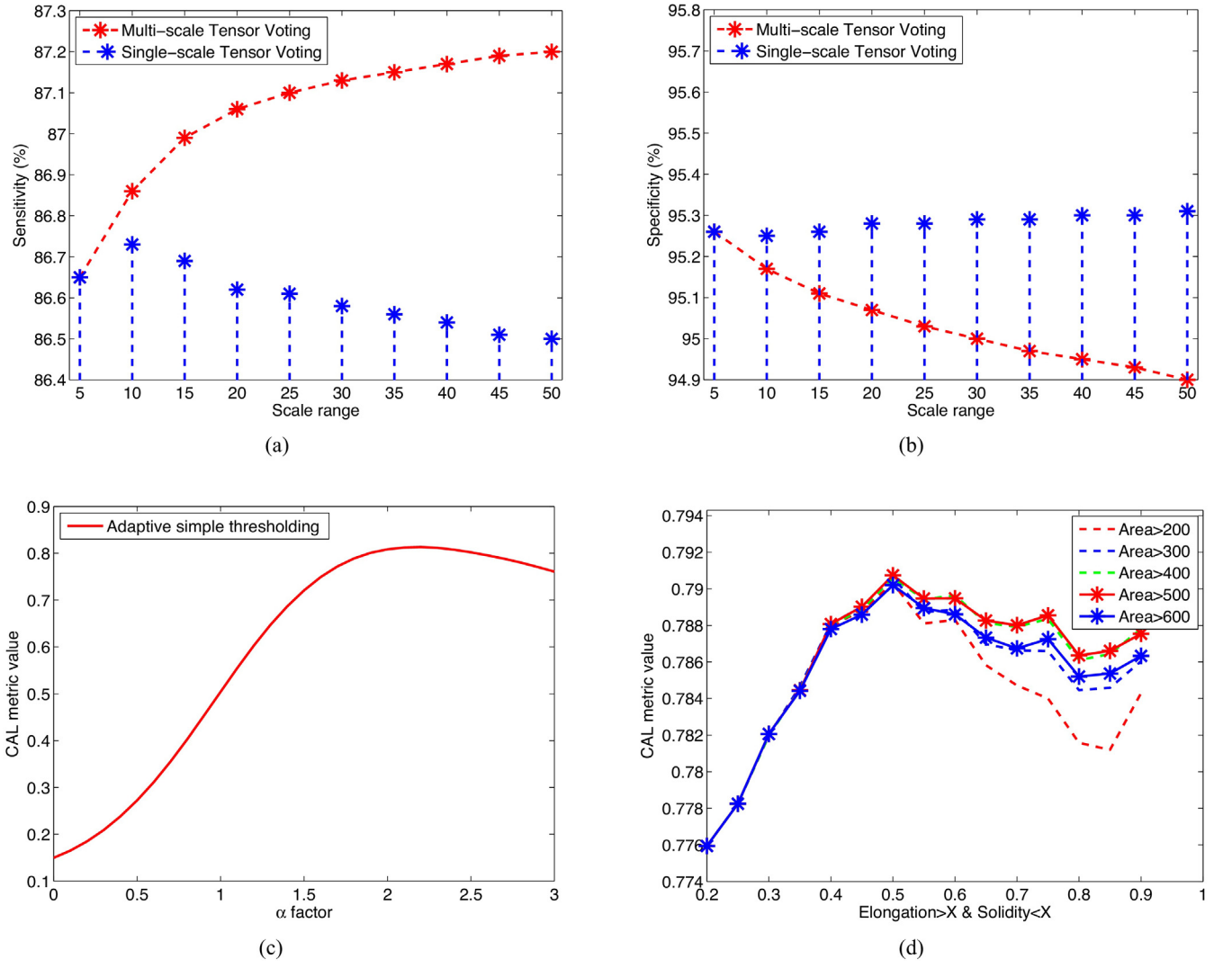


Fig. 10. Analysis of parameter settings. (a) Sensitivity rates for multi- and single-scale TVF vs. scale range, (b) specificity rates of multi- and single-scale TVF vs. scale range, (c) CAL metric vs. adaptive thresholding factor, and (d) CAL metric vs. postprocessing parameters. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the rest of the experiments. Finally, the postprocessing parameter values were selected by examining the CAL performance metric. As described in Section 2.6, these parameters establish the morphological limits between real vessel segments (to be reconnected) and artifacts (to be removed). We varied $Area_{value}$ from 200 to 600 and X from 0.2 to 0.9. Values of $Area_{value} = 500$, or 10 times the largest vessel diameter, and $X = 0.5$, gave the maximum perceptual response (Fig. 10d). Table 1 summarizes the results of the parameter setting process. The computational time across the different scales in MTVF for a typical image is given in Fig. 9.

3.3. Vessel segmentation performance

In our experiments we compared two segmentation methods: (1) the MSLD approach (Nguyen et al., 2013), and (2) our proposed approach. This was done by applying each of the methods on the test set composed of the healthy (H) and diabetic retinopathy (DR) groups of images from the Erlangen database.

3.3.1. Full vasculature analysis

The standard discrepancy metrics of sensitivity, specificity and accuracy were employed. The methods were also assessed using a perceptual evaluation metric proposed by Gegúndez-Arias et al. (2012) that is tailored to retinal applications.

For the MSLD approach (Nguyen et al., 2013), we first determined that a thresholding value ($T > 0.671$) gave the same specificity level as our approach; then we compared their sensitivities. We were interested to see if we had an improvement with respect to the rate of successfully identified vessels, or true positives (TPs), at the same background noise level (specificity). Given the same level of background noise, we can easily examine which method identifies more vessel pixels. Our method can be represented as a single operating point on the ROC curve and not as a complete curve because the final segmentation depends on many parameters that must be optimized at the same time.

Table 2 shows the results across the test set. In the following discussion, the mean and standard deviation of a given performance measure are abbreviated as M and SD. Also, we tested the statistical significance of the results using T-test analysis; when relevant, this is presented as t-value (degrees of freedom), p-value. For the sensitivity, our approach ($M = 85.06$, $SD = 4.03$) segmented significantly more vessels than the original MSLD method (Nguyen et al., 2013) ($M = 81.06$, $SD = 3.65$), $t(58) = 4.0344$, $p < 0.05$. The specificity was equivalent in both methods, while the accuracy was slightly better with our method. However, the metric used here is global in

Table 2
Full vasculature performance evaluation on test set in terms of sensitivity specificity accuracy.

Method	Sensitivity ± SD	Specificity ± SD	Accuracy ± SD
MSLD (Nguyen et al., 2013)	81.06% ± 3.65%	95.82% ± 2.14%	94.35% ± 1.72%
Proposed approach	85.06% ± 4.03%	95.82% ± 1.15%	94.79% ± 1.06%

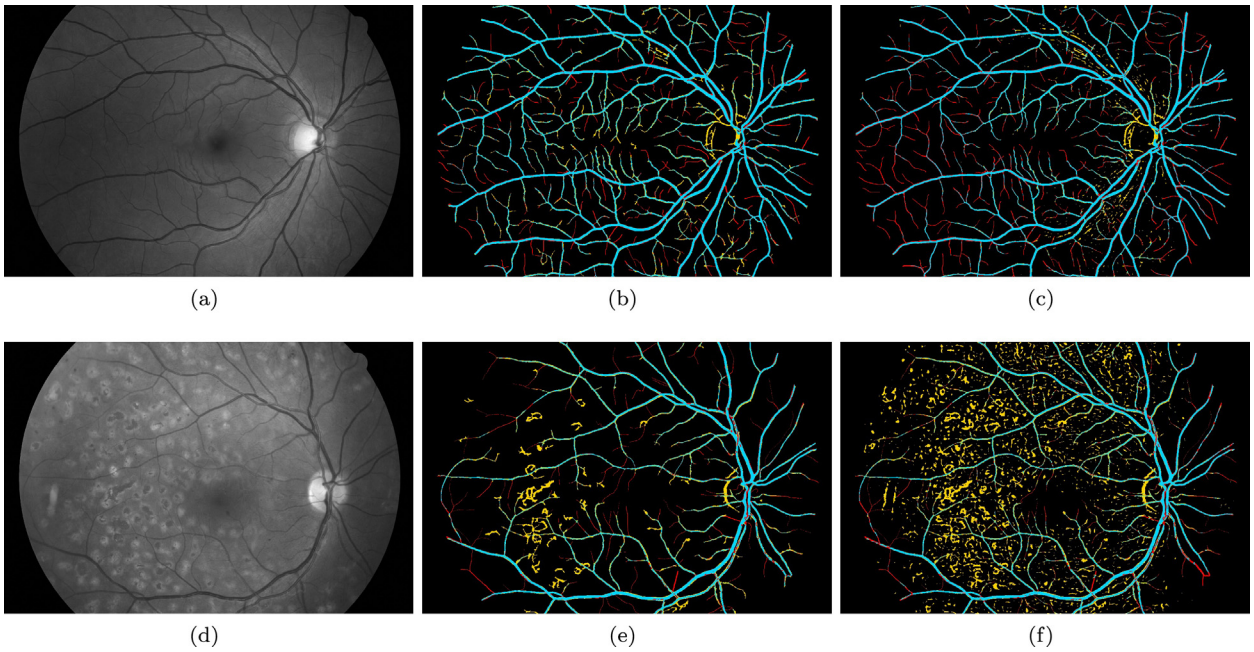


Fig. 11. Full vessel segmentation examples in Image_01_healthy (first row) and Image_04_diabetic (second row). (a) and (d) original image, (b) and (e) segmentation result of our approach, (c) and (f) segmentation result of MSLD method. In the segmentation maps (b, c and e, f), blue pixels represent the true positives, red pixels are the false negatives, yellow pixels are the false positives, and black pixels are the true negatives. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the sense that it cannot distinguish between the smaller and larger vessels in the segmentation results.

Fig. 11 shows examples of the two segmentation methods for two images, one from the (H) group and another from the (DR) group. For the top image, more small peripheral vessels were identified by our approach (Sensitivity: 82.07% vs. 73.12%). Additionally, less noise originating from the nerve fiber layer, which is adjacent to the major arcades, was segmented. The major arcades are defined as the largest vessel branches, i.e. the first order vessel branches that are superior and inferior to the optic disc. In both our approach and MSLD (Nguyen et al., 2013), the border of the optic disc was wrongly segmented. For the bottom image, the vessel segmentation performance is similar in both methods (sensitivity: 80.82% vs. 80.89%). However, our approach identifies considerably fewer background lesions than MSLD, which is reflected in the specificity rates (96.27% vs. 92.13%).

Generally, our approach relies on the connectivity criterion in the postprocessing step to exclude background noise that is not connected to the main vasculature. In DR images, it succeeds in eliminating much of the background noise, including lesions, and this contributes to lowering the false positive rate. Consequently, a higher T value must be used to threshold the MSLD response to achieve the same specificity rate as our approach for pathological cases. For healthy cases, the difference in sensitivity between the two methods is even greater, as exemplified in Fig. 11.

Table 3 shows the results in terms of the perceptual performance metric (CAL, or Connectivity Area Length). This measure is the product of three individual metrics: connectivity, area and length. The connectivity criterion penalizes any deviation from the number of connected components in the reference image. Meanwhile, the length (resp. area) criterion measures the degree of coincidence, in terms of total skeleton length (resp. vessel area), between the groundtruth and tested segmentations (Gegúndez-Arias et al., 2012). Length is an important individual metric because it shows to what extend each method can identify vessels. This criterion evaluates the skeletonized vasculature, thereby disregarding the influence of the pixel areas of the different categories of vessels.

Our approach ($M = 0.779$, $SD = 0.060$) achieves significantly higher CAL performance than the original MSLD method (Nguyen et al., 2013) ($M = 0.701$, $SD = 0.127$), $t(58) = 3.0203$, $p < 0.05$. The proposed method seeks to maximize resulting vascular connectivity, thus that particular criterion is very close to 1 (or perfectly preserved connectivity compared to the reference image) in our approach. However, the CAL metric penalizes false positives, which explains the lower results for the area and length criteria for both

Table 3
Full vasculature performance evaluation on test set in terms of CAL metric.

Method	Connectivity ± SD	Area ± SD	Length ± SD	CAL ± SD
MSLD (Nguyen et al., 2013)	0.997 ± 0.002	0.867 ± 0.074	0.804 ± 0.084	0.701 ± 0.127
Proposed approach	0.999 ± 0.000	0.902 ± 0.036	0.862 ± 0.034	0.779 ± 0.060

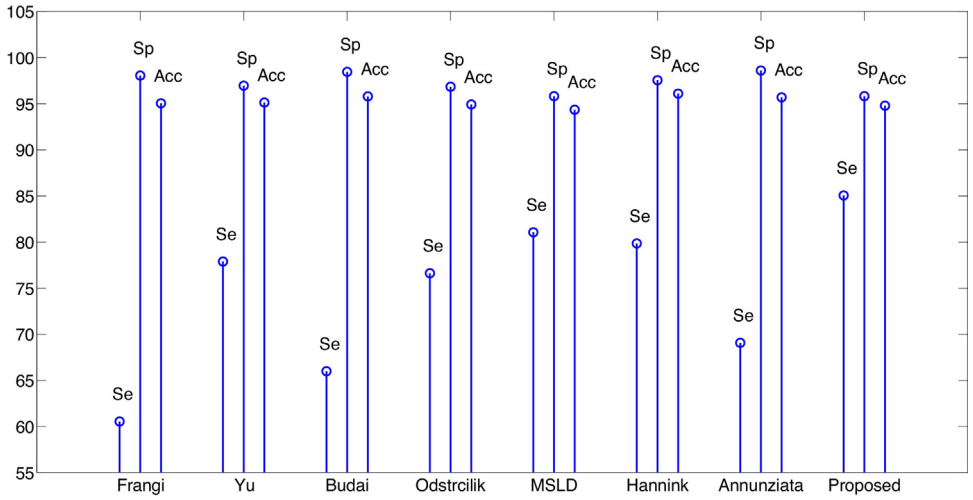


Fig. 12. Differences in the discrepancy metrics across the available segmentation methods.

Table 4
Performance comparison of different methods on test set in terms of sensitivity specificity accuracy.

Method	Sensitivity	Specificity	Accuracy
Frangi et al. (1998)	60.55%	98.05%	95.05%
Yu et al. (2012)	77.89%	96.96%	95.13%
Budai et al. (2013)	66.00%	98.45%	95.80%
Odstrcilik et al. (2013)	76.62%	96.84%	94.92%
MSLD (Nguyen et al., 2013)	81.06%	95.82%	94.35%
Hannink et al. (2014)	79.85%	97.55%	96.10%
Annunziata et al. (2015)	69.08%	98.61%	95.70%
Proposed approach	85.06%	95.82%	94.79%

segmentation methods. Nonetheless, our approach achieves area and length results that are more than 3% higher than MSLD.

Table 4 and Fig. 12 gives the performances of different methods proposed in the literature in terms of the standard discrepancy metrics. The segmentation methods of Annunziata et al. (2015), Hannink et al. (2014), Odstrcilik et al. (2013), Budai et al. (2013), Yu et al. (2012) and Frangi et al. (1998) were considered. The methods are sorted chronologically. The proposed approach achieves considerably higher sensitivity than all the other methods for all categories of images, which translates to the segmentation of more vessels. Specificity-wise, our method obtains 95.82% which, aside from the MSLD (Nguyen et al., 2013), is the lowest result, with the next closest being for Odstrcilik et al. (2013), at 96.84%. When the accuracy metric is considered, our approach ranks in the lower half of the methods, just ahead of the MSLD (Nguyen et al., 2013), and very close to Odstrcilik et al. (2013).

3.3.2. Small vessels analysis

The standard discrepancy measures cannot account for the heterogeneity in the pixel areas of vessel segments of different diameters (Niemeijer et al., 2004). Thus, we confined the following part of our analysis to the smallest vessels. The smallest vessels of the groundtruth images were first isolated from the rest of the vessel network. To achieve this, the minimization of a rotating line on the pixels belonging to the skeletonized reference image was applied as in Section 2.5.3. Next, the diameter value at each skeleton point is assigned to all vessel pixels along the normal direction at that point. Then, the pixels belonging to vessels of diameter less than 8 pixels were isolated, and the final result of this operation was subtracted from the rest of the reference image. The segmentation algorithms were assessed by computing the same discrepancy measures as those used for the full vasculature analysis (Section 3.3.1), and by assigning the large vessel pixels from the reference to the true negative class. The segmentation maps of Frangi et al. (1998) and Hannink et al. (2014) were also available, so our analysis included these methods.

Table 5 shows the performance of the four methods on the smallest vessels when the pixel-based discrepancy measures were considered. Our approach ($M = 68.64$, $SD = 6.78$) achieved significantly higher sensitivity than MSLD (Nguyen et al., 2013) ($M = 62.17$, $SD = 8.51$, $t(58) = 3.225$, $p < 0.05$), Hannink et al. (2014) ($M = 54.08$, $SD = 12.36$, $t(58) = 5.6580$, $p < 0.05$), and Frangi et al. (1998) ($M = 49.01$, $SD = 11.11$, $t(58) = 8.2576$, $p < 0.05$) methods. In fact, the difference between our method and the available ones is greater than when all the vessels were considered. Our approach detected approximately 6% more small vessels than the MSLD (Nguyen et al., 2013). Moreover, this increase doubles and triples when our approach is compared with the Hannink et al. (2014) and A.F. Frangi methods Frangi et al. (1998), respectively. In our analysis we remove the detected medium to large-sized vessels, however the oversegmentation of the vessels boundary is not taken into account which explains the differences in the specificity and accuracy metrics between the MSLD and the proposed approach.

Table 6 shows the performance of the four methods on the smallest vessels when the perceptual-based discrepancy measure is considered. Even though our approach ($M = 0.363$, $SD = 0.064$) achieved a higher overall CAL performance than the original MSLD (Nguyen et al., 2013) ($M = 0.334$, $SD = 0.085$, $t(58) = 1.4487$, $p > 0.05$) and Hannink et al. (2014) ($M = 0.313$, $SD = 0.089$, $t(58) = 2.4551$, $p > 0.05$) methods, the difference is not statistically significant. The

Table 5
Smallest vessel ($D < 8$) segmentation evaluation on test set in terms of sensitivity specificity and accuracy.

Method	Sensitivity ± SD	Specificity ± SD	Accuracy ± SD
Frangi et al. (1998)	49.01% ± 11.11%	95.82% ± 2.29%	94.58% ± 2.09%
MSLD (Nguyen et al., 2013)	62.17% ± 8.51%	96.84% ± 2.44%	95.91% ± 2.24%
Hannink et al. (2014)	54.08% ± 12.36%	97.99% ± 2.05%	96.82% ± 1.80%
Proposed approach	68.64% ± 6.78%	96.54% ± 1.33%	95.80% ± 1.32%

Table 6
Smallest vessel ($D < 8$) segmentation evaluation on test set in terms of CAL metric.

Method	Connectivity ± SD	Area ± SD	Length ± SD	CAL ± SD
Frangi et al. (1998)	0.741 ± 0.148	0.548 ± 0.087	0.402 ± 0.077	0.177 ± 0.079
MSLD (Nguyen et al., 2013)	0.980 ± 0.006	0.594 ± 0.091	0.563 ± 0.070	0.334 ± 0.085
Hannink et al. (2014)	0.991 ± 0.005	0.569 ± 0.090	0.544 ± 0.075	0.313 ± 0.089
Proposed approach	0.991 ± 0.004	0.624 ± 0.060	0.582 ± 0.053	0.363 ± 0.064

A.F. Frangi method (Frangi et al., 1998) obtained the lowest performance of all; this poor performance can be explained by the excessive amount of disconnected background noise produced by that method. The individual criteria of the CAL metric heavily penalize false positives, so the net effect is that when all three criteria are multiplied, the final CAL metric ranges from around 1% to a maximum of 36%.

Figs. 13–15 provide examples where the analysis focuses on the smallest vessels. Each figure illustrates the segmentation results of the four methods under comparison (Frangi et al. (1998), MSLD (Nguyen et al., 2013), Hannink et al. (2014), and Proposed) for the square region corresponding to the green inset superimposed on the small vessel groundtruth image. The segmentation results of each algorithm are colorized to reflect the confusion matrix: blue pixels represent the true positives, red pixels are the false negatives, yellow pixels are the false positives, and black pixels are the true negatives.

Fig. 13 (Image_15_healthy), presents the best performance of the proposed approach for both the large and small vessel categories. Qualitatively, our approach segments most of the vessels without retaining any disconnected background noise, however the terminal vessels and the boundary of the large vessels are overestimated in terms of length and width. Also, some false positive vessels are reconnected to the main vasculature. Quantitatively, our approach achieves the highest sensitivity but the lowest specificity.

Fig. 14 (Image_10_healthy), presents the worst performance of the proposed approach for the small vessel category. Our method has the same qualitative behavior as in the best case; it segments most of the vessels without retaining any disconnected background noise. But this is achieved at the expense of overestimating the vessel diameters/endpoints and reconnecting vessel-like false positives components. Even though this image represents the worst-case performance, our method still reaches the highest sensitivity rate among the considered methods.

Finally, Fig. 15 (Image_06_diabetic) presents the worst performance of the proposed approach in terms of full vascular network segmentation. An interesting feature of this image is that it includes many convoluted vessels that could be neovessels. Our approach managed to segment most of these vessels. The majority of them were segmented earlier in our pipeline (at the adaptive thresholding stage), while the MTVF partially completed the network (Fig. 16). We note as well that the vessel widths were overestimated. At the same time, the other three methods failed to segment many vessels in this region of interest. This is reflected in their sensitivity rates, which are lower than for our approach.

3.3.3. *Group differences between healthy and diabetic performance*

We divided the entire test set into the healthy (H) and diabetic (DR) images in order to assess how the proposed method performed on the separate groups as compared to the overall case. For the healthy group, the improvement in sensitivity over the original MSLD method (Nguyen et al., 2013) widened: our method detected 11% more small vessels than MSLD. However, this was accomplished at the expense of more false positives, probably due to overestimation of the vessel endpoints and widths or from

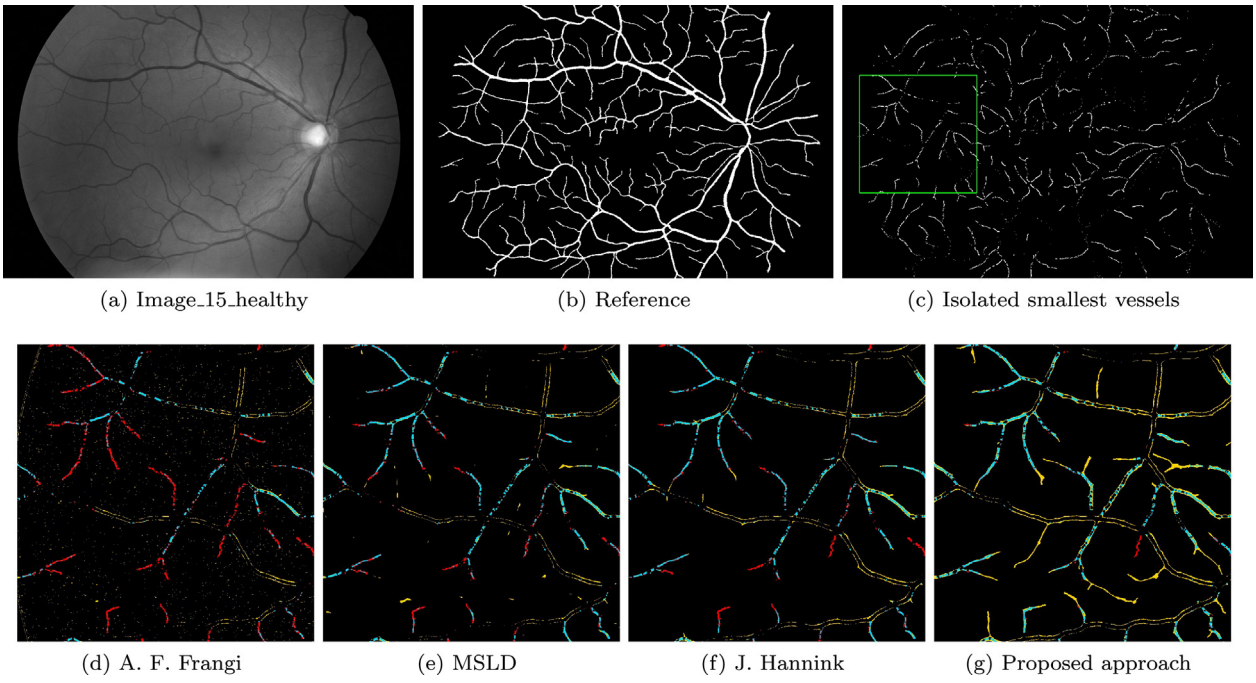


Fig. 13. Small vessel segmentation example in Image_15_healthy. (a) Grayscale input image, (b) manual segmentation of the whole vasculature, (c) isolated smallest vessels, (d) Frangi et al. (1998) segmentation result: ($Sens._{full}$ = 70.97% , $Specif._{full}$ = 98.44% , $Sens._{smallest}$ = 40.36% , $Specif._{smallest}$ = 99.34%), (e) MSLD (Nguyen et al., 2013) segmentation result: ($Sens._{full}$ = 87.28% , $Specif._{full}$ = 97.14% , $Sens._{smallest}$ = 72.41% , $Specif._{smallest}$ = 97.85%), (f) Hannink et al. (2014) segmentation result: ($Sens._{full}$ = 86.27% , $Specif._{full}$ = 97.80% , $Sens._{smallest}$ = 66.30% , $Specif._{smallest}$ = 98.47%), and (g) Proposed approach segmentation result: ($Sens._{full}$ = 92.91% , $Specif._{full}$ = 95.99% , $Sens._{smallest}$ = 83.83% , $Specif._{smallest}$ = 96.36%). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

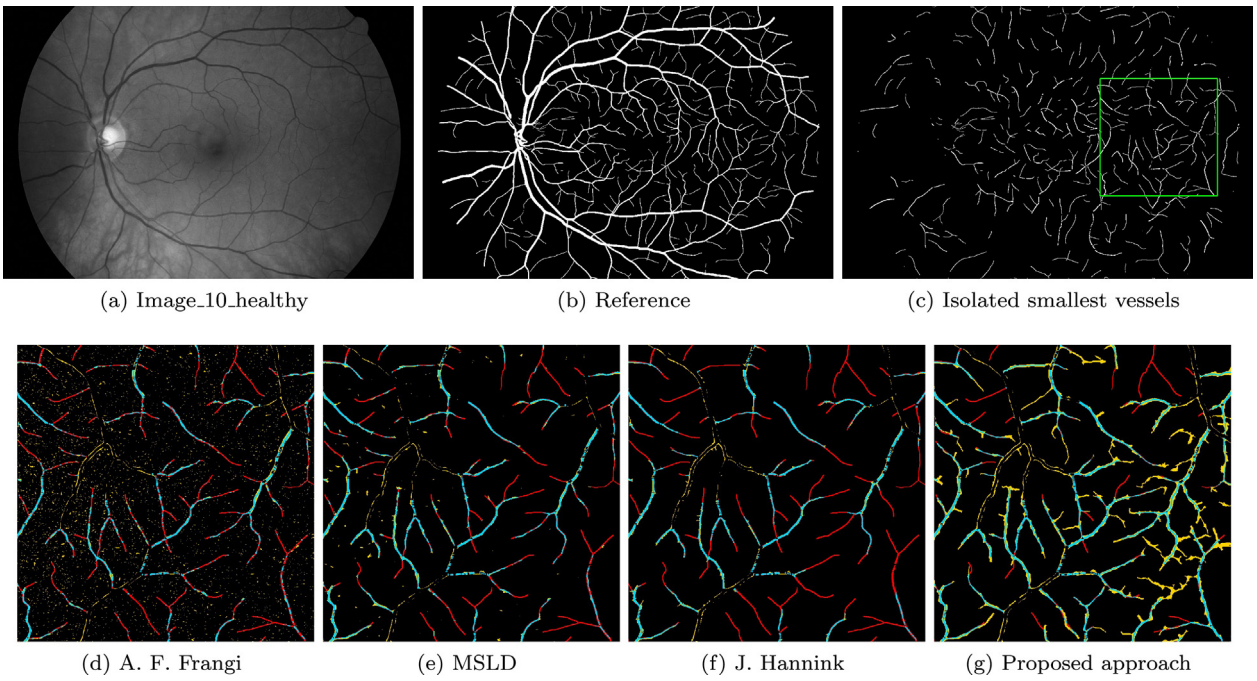


Fig. 14. Small vessel segmentation example in Image_10_healthy. (a) Grayscale input image, (b) manual segmentation of the whole vasculature, (c) isolated smallest vessels ($D < 8$), (d) Frangi et al. (1998) segmentation result: ($Sens._{full} = 70.97\%$, $Specif._{full} = 98.44\%$, $Sens._{smallest} = 39.47\%$, $Specif._{smallest} = 97.58\%$), (e) MSLD (Nguyen et al., 2013) segmentation result: ($Sens._{full} = 79.37\%$, $Specif._{full} = 97.58\%$, $Sens._{smallest} = 53.15\%$, $Specif._{smallest} = 98.45\%$), (f) Hannink et al. (2014) segmentation result: ($Sens._{full} = 74.94\%$, $Specif._{full} = 98.53\%$, $Sens._{smallest} = 35.01\%$, $Specif._{smallest} = 99.32\%$), and (g) Proposed approach segmentation result: ($Sens._{full} = 82.54\%$, $Specif._{full} = 96.96\%$, $Sens._{smallest} = 57.08\%$, $Specif._{smallest} = 97.59\%$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

false positive reconnections. The opposite trend was observed for the DR group. The gain in sensitivity over MSLD (Nguyen et al., 2013) dropped, but in terms of specificity the proposed method segmented less noise. This can be accounted for by the inherent connectivity criterion that removes isolated background noise and artifacts. When the two image groups are merged, the performances are averaged so the specificity is the same across the database and the sensitivity is reduced compared to the healthy case. Tables 7 and 8 give the results for the standard discrepancy metrics separately in each image group.

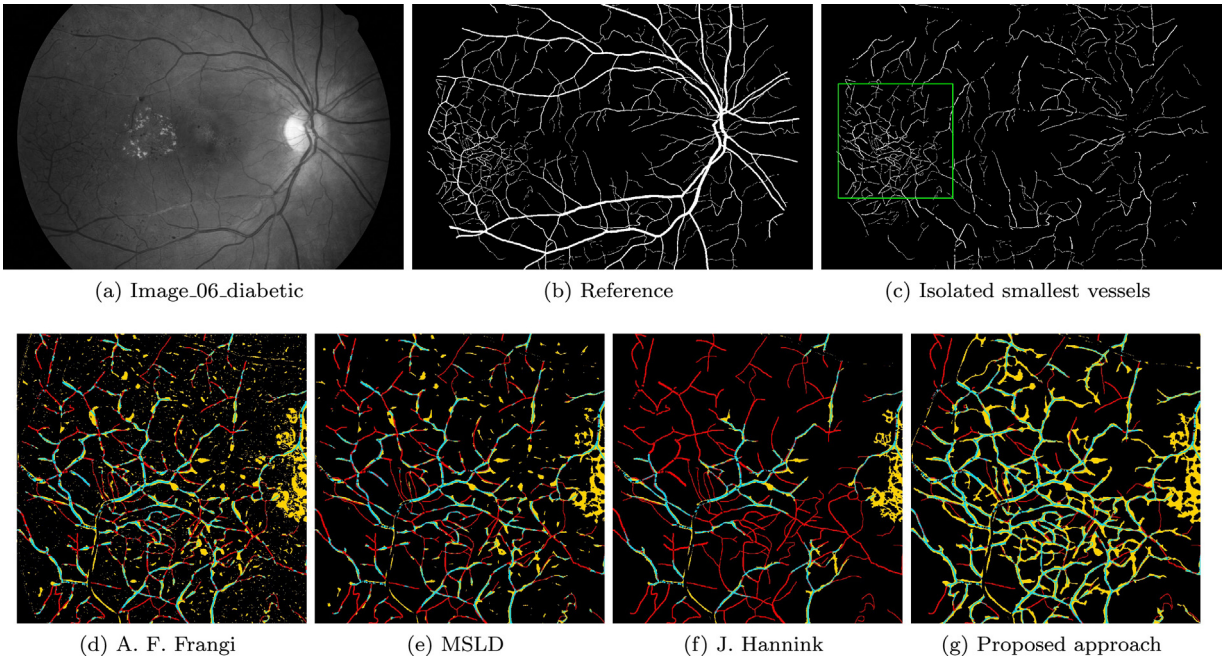
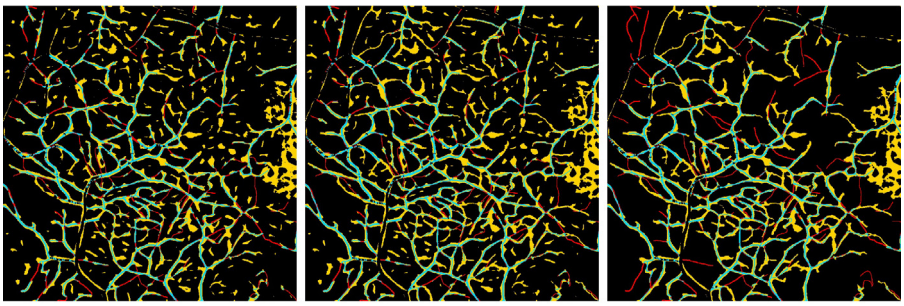


Fig. 15. Small vessel segmentation example in Image_06_diabetic. (a) Grayscale input image, (b) manual segmentation of the whole vasculature, (c) isolated smallest vessels ($D < 8$), (d) Frangi et al. (1998) segmentation result: ($Sens._{full} = 71.26\%$, $Specif._{full} = 92.94\%$, $Sens._{smallest} = 54.19\%$, $Specif._{smallest} = 92.44\%$), (e) MSLD (Nguyen et al., 2013) segmentation result: ($Sens._{full} = 72.59\%$, $Specif._{full} = 93.90\%$, $Sens._{smallest} = 55.35\%$, $Specif._{smallest} = 94.79\%$), (f) Hannink et al. (2014) segmentation result: ($Sens._{full} = 68.25\%$, $Specif._{full} = 95.85\%$, $Sens._{smallest} = 42.56\%$, $Specif._{smallest} = 96.66\%$), and (g) Proposed approach segmentation result: ($Sens._{full} = 74.43\%$, $Specif._{full} = 93.62\%$, $Sens._{smallest} = 58.62\%$, $Specif._{smallest} = 94.47\%$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) (b) (c)

Fig. 16. Small vessel segmentation at different stages in proposed pipeline, for Image_06_diabetic. (a) Adaptive thresholding result, (b) small vessel reconnection result, and (c) postprocessed result. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 7
Full vasculature performance evaluation on test set in terms of sensitivity specificity accuracy.

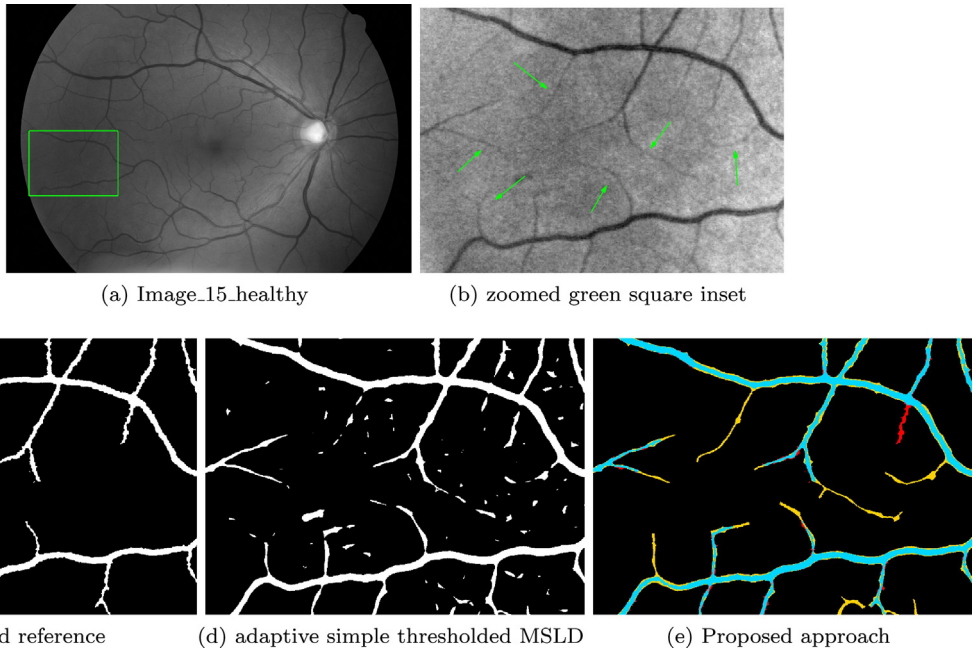
Category	Method	Sensitivity ± SD	Specificity ± SD	Accuracy ± SD
H	MSLD (Nguyen et al., 2013)	80.50% ± 4.04%	97.66% ± 0.55%	95.74% ± 0.63%
	Proposed approach	86.60% ± 3.40%	96.67% ± 0.43%	95.54% ± 0.43%
DR	MSLD (Nguyen et al., 2013)	81.62% ± 3.26%	93.99% ± 1.41%	92.95% ± 1.23%
	Proposed approach	83.53% ± 4.13%	94.97% ± 1.02%	94.05% ± 0.98%

Table 8
Smallest vessel ($D < 8$) segmentation evaluation on test set in terms of sensitivity specificity accuracy.

Category	Method	Sensitivity ± SD	Specificity ± SD	Accuracy ± SD
H	MSLD (Nguyen et al., 2013)	58.10% ± 8.64%	97.86% ± 0.51%	96.90% ± 0.46%
	Proposed approach	69.90% ± 7.15%	96.96% ± 0.42%	96.30% ± 0.39%
DR	MSLD (Nguyen et al., 2013)	66.25% ± 6.31%	94.33% ± 1.36%	93.61% ± 1.21%
	Proposed approach	67.38% ± 6.39%	95.55% ± 1.06%	94.81% ± 1.06%

4. Discussion

Segmenting the smallest retinal vessels is still a challenging problem. The MSLD method (Nguyen et al., 2013) is based on straight line segments, which is inadequate in this context. Using tensor voting can help by retaining only the information that is perceptually significant to extending the segmentation to the smallest vessels. In this paper, we have proposed a novel hybrid approach



(a) Image_15_healthy (b) zoomed green square inset

(c) zoomed reference (d) adaptive simple thresholded MSLD (e) Proposed approach

Fig. 17. Missing small vessels in the reference segmentation of Image_15_healthy. (a) Grayscale input image, (b) zoomed region corresponding to inset in (a); green arrows indicate faint vessels, (c) corresponding reference region, (d) adaptive thresholded MSLD, (e) final segmentation result of the proposed method. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

that combines the two methods. Multi-scale information from the TVF is combined in a new way that allows us to reconnect vessels at greater distances than single scale tensor voting. Besides, our approach can fully reconstruct the vessels, which enables the use of the framework for lesion detection in CAD systems.

Even though the proposed method achieved higher performance than MSLD (Nguyen et al., 2013) in the majority of full and small vessel experiments, several challenges remain to be addressed. In the final result, the segmented small vessels can be slightly translated from their groundtruth counterparts. Similarly, their reconstructed widths do not exactly coincide in some locations with the corresponding vessels in the manually segmented image. Moreover, small vessel terminal points are over-estimated. Additionally, the TVF can reconnect vessel-like false positive structures to the main vasculature, while conversely it can miss real vessels lying after junctions.

The errors in the reconstructed width of a small vessel can be attributed to the pixel painting process. The diameter obtained from the larger vessel is used without taking into account local information such as the local saliency from TVF. This gives rise to false positives and to an equivalent reduction in the true positive counts. The issue of vessel over-extension was also raised by Leng et al. (2011), however their solution was based on simple postprocessing, taking into account the information from an initial centerline map extracted by a Canny edge detector. The vessel over-extension problem, together with the false positive reconnection problem, is related to the quality of the available reference. Precise manual segmentation of the smallest vessels is not obvious in high resolution images, even for an expert. Therefore, the use of multiple experts who manually segment the same images would allow a consensus groundtruth to be constructed. With such a reference, there would be higher confidence about the smallest vessels' positions and extents. In our case, there were instances where the expert did not identify small vessels or did not extend them until their endpoints. Fig. 17 highlights examples where small vessels of weaker contrast are missing or they were not completed in the manual segmentation. Our approach, on the other hand, managed to identify and/or complete these vessels.

Furthermore, there is significant difficulty in reconnecting vessel segments that are located beyond junctions. As was highlighted in Section 2.6, the high level of orientation uncertainty causes the TVF to fragment the vessels at junctions, even if these structures are well defined. In the proposed method, we try to retain branched vessels via morphological processing, but alternatives could be considered. The ballness measure (λ_2) reaches local maxima at junctions. Using it directly via concurrent tracking and searching for branched vessels to reconnect could be advantageous. Alternatively, the use of higher than second-order tensors (Maggiori et al., 2014) could resolve the inability of the TVF to infer the information at junctions. Finally, new metrics should be proposed that will be tailored to the analysis of the smallest vessels by removing the influence of the larger vessels.

5. Conclusion

A novel approach was proposed to segment small vessels in high resolution retinal images. Simple thresholding the MSLD (Nguyen et al., 2013) response leads to the fragmentation of the low contrasted smallest vessels. Instead, this work proposes the combination of an adaptively thresholded MSLD with a perceptual organization approach (TVF) to reconnect the small fragmented vessels. Multiple scales for the TVF are considered which promote the reconnection in a greater distance than single scale. The method was applied in high resolution fundus images. According to the available data, our approach can segment more vessels than MSLD (Nguyen et al., 2013), both when all the vessel diameters are considered and when the analysis is confined to the smallest vessels. In addition, our method can further extend the detected vasculature and produce a more connected result than previous works; both these advantages translate to better perceptual performance. Compared to other published methods, our approach segments more vessels at the expense of slightly higher background noise. However, the proposed algorithm should be validated in datasets containing images with different pixel resolutions.

Methods relying on a connectivity criterion for segmenting the vessels are well suited for preprocessing the retinal vasculature in CAD systems. On the other hand, simple thresholding does not set any restriction on the background responses, which can result in false negatives at a subsequent lesion detection step, i.e. failure to detect potentially vision-threatening lesions.

Future work will focus on utilizing the ballness measure (λ_2) computed in the tensor voting framework for neovascularization detection. This category of vessels is even more challenging to detect because neovessels are very fine and tortuous structures with low contrast that form chaotic networks around the optic disc or in other areas of the retina. Their presence in the fundus image signals the transition to the final stage of diabetic retinopathy and thus their accurate detection is essential.

Acknowledgements

This study was funded by the grant RDCPJ419502-11 from National Science and Engineering Research Council of Canada (NSERC). The authors would like to thank the members of the LIV4D group (Polytechnique Montréal) and Diagnos Inc. for their support as well as Philippe Debanné for revising this manuscript. They are also grateful to J. Hannink for granting access to the segmentation maps of his method.

References

Abràmoff, M.D., Garvin, M.K., Sonka, M., 2010. Retinal imaging and image analysis. IEEE Rev. Biomed. Eng. 3, 169–208.

Allen, K., Joshi, N., Noble, J.A.,2011. Tramline and NP windows estimation for enhanced unsupervised retinal vessel segmentation. In: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, pp. 1387–1390.

Annunziata, R., Garzelli, A., Ballerini, L., Mecocci, A., Trucco, E., 2015. Leveraging multiscale Hessian-based enhancement with a novel exudate inpainting technique for retinal vessel segmentation. IEEE J. Biomed. Health Inf. (99), 1.

Antal, B., Hajdu, A., 2013. Improving microaneurysm detection in color fundus images by using context-aware approaches. Comput. Med. Imaging Graph. 37 (5), 403–408.

Archer, D.B., 1999. Diabetic retinopathy: some cellular, molecular and therapeutic considerations. Eye 13 (4), 497–523.

Berens, P., 2009. Circstat: a matlab toolbox for circular statistics. J. Stat. Softw. 31 (10), 1–21, URL http://www.jstatsoft.org/v31/i10.

Budai, A., Odstrcilik, J., Kolar, R., Hornegger, J., Jan, J., Kubena, T., Michelson, G., 2011. A public database for the evaluation of fundus image segmentation algorithms. In: Proc. of ARVO, FL: USA: Fort Lauderdale.

Budai, A., Bock, R., Maier, A., Hornegger, J., Michelson, G., 2013. Robust vessel segmentation in fundus images. Int. J. Biomed. Imaging.

Donoho, D.L., 1995. De-noising by soft-thresholding. IEEE Trans. Inf. Theory 41 (3), 613–627.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognit. Lett. 27 (8), 861–874.

Foracchia, M., Grisan, E., Ruggeri, A., 2005. Luminosity and contrast normalization in retinal images. Med. Image Anal. 9 (3), 179–190.

Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.,1998. Multiscale vessel enhancement filtering. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI'98. Springer, pp. 130–137.

Franken, E., van Almsick, M., Rongen, P., Florack, L., ter Haar Romeny, B.,2006. An efficient method for tensor voting using steerable filters. In: Computer Vision – ECCV 2006. Springer, pp. 228–240.

Fraz, M.M., Remagnino, P., Hoppe, A., Uyyanonvara, B., Rudnicka, A.R., Owen, C.G., Barman, S.A., 2012. Blood vessel segmentation methodologies in retinal images – a survey. Comput. Methods Programs Biomed. 108 (1), 407–433.

García, M., Sánchez, C.I., Lopez, M.I., Diez, A., Hornero, R.,2008. Automatic detection of red lesions in retinal images using a multilayer perceptron neural network.

In: Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE. IEEE, pp. 5425–5428.
Gegúndez-Arias, M.E., Aquino, A., Bravo, J.M., Marin, D., 2012. A function for quality evaluation of retinal vessel segmentations. IEEE Trans. Med. Imaging 31 (2), 231–239.
Hannink, J., Duits, R., Bekkers, E.,2014. Crossing-preserving multi-scale vesselness. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014. Springer, pp. 603–610.
Haralick, R.M., Shapiro, L.G., 1992. Computer and Robot Vision, 1st ed. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
Kingsbury, N.G., 1998. The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters. In: Proc. 8th IEEE DSP Workshop, Utah, vol. 8, Citeseer, p. 86.
Kovács, G., Hajdu, A.,2011. Extraction of vascular system in retina images using averaged one-dependence estimators and orientation estimation in hidden markov random fields. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2011. IEEE, pp. 693–696.
Leng, Z., Korenberg, J.R., Roysam, B., Tasdizen, T.,2011. A rapid 2-d centerline extraction method based on tensor voting. In: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, pp. 1000–1003.
Loss, L.A., Bebis, G., Parvin, B., 2011. Iterative tensor voting for perceptual grouping of ill-defined curvilinear structures. IEEE Trans. Med. Imaging 30 (8), 1503–1513.
Maggiori, E., Manterola, H.L., del Fresno, M., 2014. Perceptual grouping by tensor voting: a comparative survey of recent approaches. IET Comput. Vis. 9 (2), 259–277.
Medioni, G., Kang, S.B., 2004. Emerging Topics in Computer Vision. Prentice Hall PTR.
Medioni, G., Lee, M.-S., Tang, C.-K., 2000. A Computational Framework for Segmentation and Grouping. Elsevier.
Mendonca, A.M., Campilho, A., 2006. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. IEEE Trans. Med. Imaging 25 (9), 1200–1213.
Nguyen, U., Bhuiyan, A., Park, L.A., Ramamohanarao, K., 2013. An effective retinal blood vessel segmentation method using multi-scale line detection. Pattern Recognit. 46 (3), 703–715.
Niemeijer, M., Staal, J., van Ginneken, B., Loog, M., Abramoff, M.D.,2004. Comparative study of retinal vessel segmentation methods on a new publicly available database. In: Medical Imaging 2004. International Society for Optics and Photonics, pp. 648–656.
Odstrcilik, J., Kolar, R., Budai, A., Hornegger, J., Jan, J., Gazarek, J., Kubena, T., Cernosek, P., Svoboda, O., Angelopoulou, E., 2013. Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database. IET Image Process. 7 (4), 373–383.
Oost, E., Akatsuka, Y., Shimizu, A., Kobatake, H., Furukawa, D., Katayama, A.,2010. Vessel segmentation in eye fundus images using ensemble learning and curve fitting. In: 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, pp. 676–679.
Park, J., Kien, N.T., Lee, G.,2007. Optic disc detection in retinal images using tensor voting and adaptive mean-shift. In: 2007 IEEE International Conference on Intelligent Computer Communication and Processing. IEEE, pp. 237–241.
Ricci, E., Perfetti, R., 2007. Retinal blood vessel segmentation using line operators and support vector classification. IEEE Trans. Med. Imaging 26 (10), 1357–1365.
Risser, L., Plouraboué, F., Descombes, X., 2008. Gap filling of 3-d microvascular networks by tensor voting. IEEE Trans. Med. Imaging 27 (5), 674–687.
Scanlon, P., 2008. The English national screening programme for sight-threatening diabetic retinopathy. J. Med. Screen. 15 (1), 1–4.
Sofka, M., Stewart, C.V., 2006. Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures. IEEE Trans. Med. Imaging 25 (12), 1531–1546.
Sopharak, A., Uyyanonvara, B., Barman, S., 2013. Simple hybrid method for fine microaneurysm detection from non-dilated diabetic retinopathy retinal images. Comput. Med. Imaging Graph. 37 (5), 394–402.
van Antwerpen, G., Verbeek, P., Groen, F.,1986. Automatic counting of asbestos fibres. In: 1986 Proceedings of the 3rd European Signal Processing Conference (EUSIPCO). Elsevier, pp. 891–896.
Winder, R.J., Morrow, P.J., McRitchie, I.N., Bailie, J., Hart, P.M., 2009. Algorithms for digital image processing in diabetic retinopathy. Comput. Med. Imaging Graph. 33 (8), 608–622.
World Health Organisation, 2005. Prevention of Blindness from Diabetes Mellitus: Report of WHO Consultation in Gevena, Switzerland, 9–11 November 2005.
Xu, L., Luo, S., 2010. A novel method for blood vessel detection from retinal images. Biomed. Eng. Online 9 (1), 14.
Yigitsoy, M., Navab, N., 2013. Structure propagation for image registration. IEEE Trans. Med. Imaging 32 (9), 1657–1670.
You, X., Peng, Q., Yuan, Y., Cheung, Y.-m., Lei, J., 2011. Segmentation of retinal blood vessels using the radial projection and semi-supervised approach. Pattern Recognit. 44 (10), 2314–2324.
Yu, H., Barriga, S., Agurto, C., Zamora, G., Bauman, W., Soliz, P.,2012. Fast vessel segmentation in retinal images using multiscale enhancement and second-order local entropy. In: SPIE Medical Imaging. International Society for Optics and Photonics, p. 83151B.
Zhang, D., Li, Q., You, J., 2009. A modified matched filter with double-sided thresholding for screening proliferative diabetic retinopathy. IEEE Trans. Inf. Technol. Biomed. 13 (4), 528–534.
Zhang, B., Wu, X., You, J., Li, Q., Karray, F., 2010. Detection of microaneurysms using multi-scale correlation coefficients. Pattern Recognit. 43 (6), 2237–2248.
Zuluaga, M.A., Rodionov, R., Nowell, M., Achhala, S., Zombori, G., Mendelson, A.F., Cardoso, M.J., Miserocchi, A., McEvoy, A.W., Duncan, J.S., et al., 2015. Stability, structure and scale: improvements in multi-modal vessel extraction for SEEG trajectory planning. Int. J. Comput. Assist. Radiol. Surg., 1–11.
Zwiggelaar, R., Astley, S.M., Boggis, C.R., Taylor, C.J., 2004. Linear structures in mammographic images: detection and classification. IEEE Trans. Med. Imaging 23 (9), 1077–1086.