Elsevier Editorial System(tm) for Computerized Medical Imaging and Graphics
Manuscript Draft

Manuscript Number: CMIG-D-14-00273R1

Title: Retinal vessel segmentation using multi-scale textons derived from keypoints

Article Type: Original Research Articles

Keywords: Retinal vessel; image segmentation; Texton; keypoints.

Corresponding Author: Dr. Mark Fisher, BSc MSc PhD

Corresponding Author's Institution: University of East Anglia

First Author: Lei Zhang, BSc, MSc

Order of Authors: Lei Zhang, BSc, MSc; Mark Fisher, BSc MSc PhD; Wenjia Wang, BEng, MEng, PhD

Abstract: This paper presents a retinal vessel segmentation algorithm which uses a texton dictionary to classify vessel/non-vessel pixels. However, in contrast to previous work where filter parameters are learnt from manually labelled image pixels our filter parameters are derived from a smaller set of image features that we call keypoints. A Gabor filter bank, parameterised empirically by ROC analysis, is used to extract keypoints representing significant scale specific vessel features using an approach inspired by the SIFT algorithm. We first determine keypoints using a validation set and then derive seeds from these points to initialise a k-means clustering algorithm which builds a texton dictionary from another training set. During testing we use a simple 1-NN classifier to identify vessel/non-vessel pixels and evaluate our system using the DRIVE database. We achieve average values of sensitivity, specificity and accuracy of 78.12%, 96.68% and 95.05% respectively. We find that clusters of filter responses from keypoints are more robust than those derived from hand-labelled pixels. This, in turn yields textons more representative of vessel/non-vessel classes and mitigates problems arising due to intra and inter-observer variability.

Cover Letter

Faculty of Science
School of Computing Sciences

University of East Anglia
Norwich Research Park
Norwich NR4 7TJ
United Kingdom

Email: mark.fisher@uea.ac.uk
Tel: +44 (0) 1603 592671
Fax: +44 (0) 1603 593345

Web:www.uea.ac.uk

The Editor
Computerized medical Imaging and Graphics
Elsevier
1-3 Strand
London
WC2N 5JR

28 May 2015

Your Ref: CMIG-D-14-00273R1

Dear Sir/Madam

Re: Retinal vessel Segmentation using multi-scale textons derived from keypoints

Please find the attached resubmission. The work represents a new approach for unsupervised automated vessel segmentation in fundus images which was developed by Lei Zhang in the context of a PhD in the School of Computing Sciences, UEA. The work is documented in a thesis which was recently accepted and PhD awarded (External Examiner Dr Mike Spann, University of B'ham. UK.

This unsupervised approach builds on a supervised method which was presented at In: Proc. 21st IEEE International Conference on Image Processing, 2014 Oct. 27-30; Paris, France; IEEE: 2014; pp. 252-56.

The co-authors, Fisher and Wang are the student's supervisors.

Yours sincerely

Mark Fisher

Dr Mark Fisher, CEng. MIET, Senior Lecturer; Corresponding Author.

+44 (0) 1603 592671
mark.fisher@uea.ac.uk

We thank the reviewers' for their comments and have now substantially revised the manuscript. In particular we have attempted to improve the description of the technique so it is more accessible. We also renamed the paper as the original title included an explicit reference to SIFT, which is misleading. Our approach is inspired by SIFT but there are important differences.

Reviewer #1: This manuscript described a texton based retinal vessel segmentation approach using Gabor filters and a derivative SIFT. Instead of using difference of Gaussians (DoG) to extract scale dependent features (SIFT), in this manuscript, the scale dependent vessel features are calculated using multi-scale Gabor filters and keypoints calculated through the results (e.g. local maxima) of Gabor filters are used to initialize the k-mean algorithm for the textons extraction.

This is a challenging topic and the idea of using keypoints to initialize the k-mean algorithm for the textons extraction, instead of using labelled ground truth, is quite interesting. The method has been evaluated on 20 images and the presented results are very promising, compared to the other unsupervised approaches.

The general presentation of the paper is clear and the experimental results have been well discussed. However, some parts are lacking details and more explanations are needed. Specific concerns and /or suggestions are as follows:

1) In this paper, the scale dependent features are calculated through multi-scale Gabor filters, why do the authors call it a derivate SIFT?

Our multi-scale approach was inspired by SIFT and shares some (but not all) of the methodology; hence the title. But on reflection we agree that naming it a "derivative of SIFT" is misleading. We have renamed the paper "**Retinal vessel segmentation using multi-scale textons derived from keypoints"**. We've deleted references to DSIFT.

2) Using Keypoints for the initialization of the K-mean is a good point and this is interesting. However, from the k-mean results how do you choose the classes for the final segmentation?

We use k = 5 to give cluster centres that map to 4 vessel scales λ= {6, 8, 9, 13} and the background. We observe that the largest cluster will correspond to the background (as do Oliveira et al. [37]. So our 5 clusters (textons) are labelled implicitly as vessel/non-vessel. As such we describe it as an unsupervised learning approach.

3) Page 20, line 16, it seems there is no "Fig.7(a) circular"?

This is an error; we've corrected it. Note: We split Fig. 1 into two figures. The test phase is now shown as Figure 7 so original Fig. 7 now becomes Fig. 8.

4) Page 23, line 35, there is no Fig.7 a2, b2, c2?

We've corrected this, page 23 line 3 now refers to Fig. 8 (g) (h) and (i).

Reviewer #2:

1. Last sentence of abstract. I think "higher sensitivity" and "better in detecting vessels" are equivalent. I don't think "better…than …manual observer" is a conclusion could be drawn from this manuscript. A more rigorous study and more manual observers are needed to state a method is better than manual observers.

We've reworded the sentence "We achieve average values of sensitivity, specificity and accuracy of 78.12%, 96.68% and 95.05% respectively".

2. Figure 1 appeared before it was mentioned in the text

We repositioned figure 1 on page 10 (cross reference on page 9).

3. Page 10 line 55 ~ page 11 line 12: I don't think it's necessary to explain TP, FP, TN, FN. These should be common sense to the readers of this journal.

We deleted this content, but kept the definitions (Eqns. 3.2 – 3.4). We cite Fawcett [25] for a more detailed description.

4. Page 16 line 5 ~ line 21 and Figure 3, I think these belong to section 3.1.1 as it's still discussing parameter optimization.

We merged section 3.1.1 and 3.1.2 together, deleting the section 3.1.2 and renaming the section 3.1.1 by "Design of Gabor filter and parameter optimization"

5. Page 20 line 46, please explain how "training the dictionary of key point descriptors" was done. Was the manual ground truth used? If yes, I don't think this method can be called unsupervised. Semi-supervised should be more appropriate. Or please explain how this method can be implemented completely independent of manual labelling.

This point illustrates a problem with the original paper (and identified by other reviewers) that we have worked hard to correct in the new submission. We have revised Figure 1 and introduced a new Figure (Figure 7).

The section describing the training of keypoint descriptors has been rewritten (section 3.2). The keypoint descriptors are identified by local maxima of filter outputs over a spatial and scale neighbourhood (see Figure 5). Only the parameters of filter bank are derived with reference to hand labelled ground truth (section 3.1.1). The remaining processing stages (section 3.2, 3.3) are all unsupervised.

6. Page 20 line 57 ~ 60. Should proper image registration be done when matching the key points?

Again, this illustrates a problem with the poor description in the original paper. Hopefully our redrafted paper better describes the 2-stage training process and makes this much clearer. Both training sets yield keypoints, indexed by scale and id (image id + pixel position). Since textons are only characterised by a vector of filter bank outputs they are invariant of position, so image registration is not needed. But we do need to ensure we're matching keypoint descriptors at appropriate scales. This is described on page 20 (first para.).

The need to match keypoints only arises because we choose to use a validation set in the first training stage. The motivation for this is given on page 5 "This two-stage approach allows us to decouple parameter selection from training, so avoiding possible over fitting while also supporting the possibility of using different databases for model selection and training."

7. Page 20 line 16 and line 39. Fig. 7(a) and Fig. 6(b) were referred to help explaining this step. However, I don't think these two figures showed what's described in the text. There should be an additional figure to help the readers understand this paragraph.

We revised the figure:

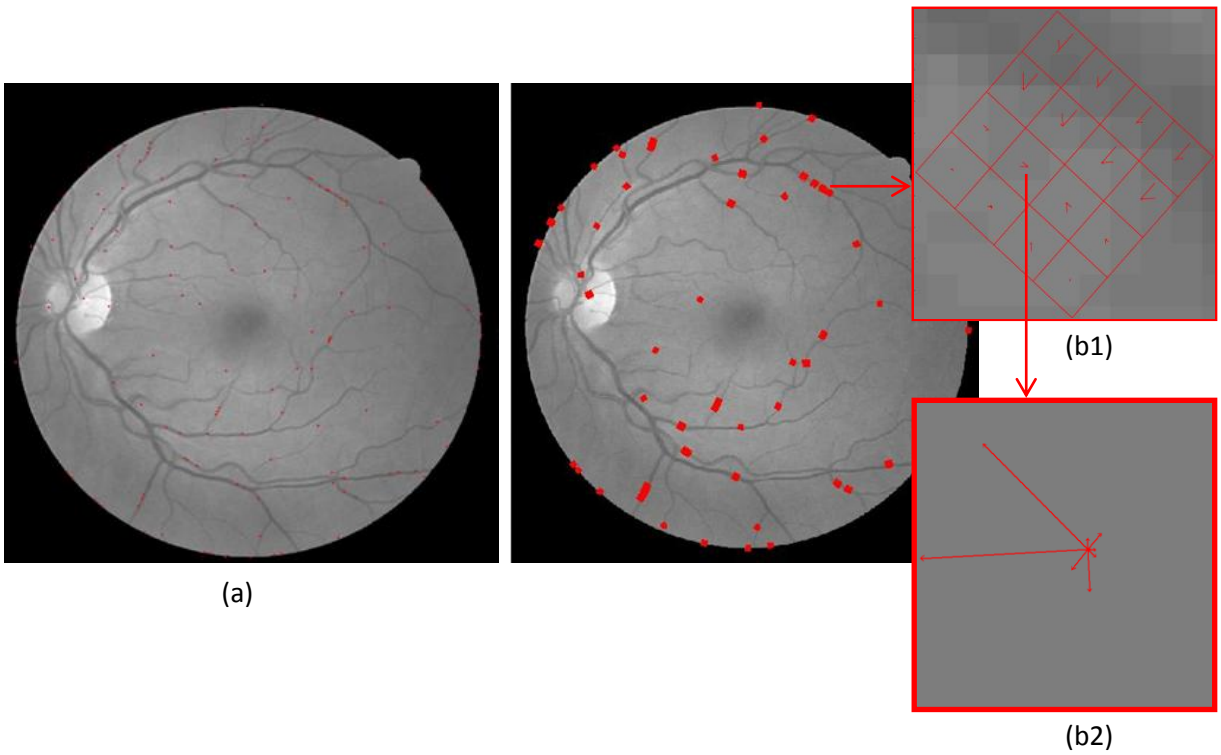


Figure 6: (a) illustrates the key points and (b) illustrates corresponding descriptors (after filtering low-contrast candidates). One of the descriptors is zoomed in and presented in (b1); one of the orientation histograms presented in (b1) is illustrated in (b2).

8. Page 22 lines 32 ~ 42. What does this piece of code do? Is it for constructing a background cluster? In the text, please explain explicitly how the background cluster was constructed. It looks like if the clusters were initialized by the key points, all the generated clusters should be vessel clusters.

We included an improved description (top of page 21) beginning "Our *k*-means clustering algorithm is a hybrid."

We also revised the pseudo-code Algorithm 1.

We also included a paragraph beginning "The motivation for the *k*-means hybrid…" after the pseudo-code description., and cite [39] S Bubeck et al. to further support the approach.

9. Page 24 lines 5 ~ 7. "The comparative results show that our method has much better sensitivity than the best other methods." I don't think this statement is true. Please edit.

Replaced by (page 24)

"The comparative results show that our method has much better sensitivity than the most of the other methods listed in the table 1."

10. Page 25 line 19, researches - reaches?

We redrafted the sentence to read "We report an average sensitivity of 0.7812".

11. Page 26 line 12 ~ line 14. "In terms of the evaluation criterion, a primary goal of vessel segmentation is to detect as many vessel elements as possible" Why? Is there any reference for this argument? Are the small vessels important for diagnostic purpose?

We cite two clinical applications to support this statement (page 26)

"In terms of the evaluation criterion, a primary goal of vessel segmentation is to detect as many vessel elements as possible, since in reality a ratio between diameter of arteries to those of veins also known as A/V ratio is an important clinical biometry to assess the risk of many diseases (e.g. hypertension, brain stroke) **Error! Reference source not found.** However it is also important to detect the tiny vessels as their condition is indicative in diagnosis of Wet Macular Degeneration (WMD). WMD, also known as choroidal neovascularization, is characterised by abnormal growth of the choroidal vascular structure into or around the macula. These new vessels are weak and easily bleed **Error! Reference source not found.**. Achieving good performance in segmenting thin vessels suggests it is better to improve the overall accuracy by pursuing higher sensitivity while maintaining the same or sacrificing only a small fraction of specificity."

Retinal vessel segmentation using multi-scale textons derived from keypoints

Lei Zhang, Mark Fisher, WenJia Wang

Corresponding Author: Dr. Mark Fisher, School of Computing Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK,

mark.fisher@uea.ac.uk, +44(0)1603592671

Retinal vessel segmentation using multi-scale textons derived from keypoints

L. Zhang[a], Mark. Fisher[b], Wenjia. Wang[b]

[a] The Laboratory of Vision Engineering, School of Computer Science, University of Lincoln, Brayford Pool, Lincoln, LN6 7TS, UK

[b] School of Computing Sciences, University of East Anglia, Norwich Research Park, NR4 7TJ, Norwich, UK

Abstract

This paper presents a retinal vessel segmentation algorithm which uses a texton dictionary to classify vessel/non-vessel pixels. However, in contrast to previous work where filter parameters are learnt from manually labelled image pixels our filter parameters are derived from a smaller set of image features that we call keypoints. A Gabor filter bank, parameterised empirically by ROC analysis, is used to extract keypoints representing significant scale specific vessel features using an approach inspired by the SIFT algorithm. We first determine keypoints using a validation set and then derive seeds from these points to initialise a *k*-means clustering algorithm which builds a texton dictionary from another training set. During testing we use a simple 1-NN classifier to identify vessel/non-vessel pixels and evaluate our system using the DRIVE database. We achieve average values of sensitivity, specificity and accuracy of 78.12%, 96.68% and 95.05% respectively. We find that clusters of filter responses from keypoints are more robust than those derived from hand-labelled pixels. This, in turn yields textons more representative of vessel/non-vessel classes and mitigates problems arising due to intra and inter-observer variability.

Keywords: Retinal vessels, image segmentation, Texton, keypoints

1. Introduction

The visual appearance of the retina in digital fundus images is an important indicator of eye disease and can be diagnostic of a range of more general issues relating to health and wellbeing. As such, fundus photography is also used in screening programmes for diseases such as diabetes and hypertension and the morphological properties of retinal vessels are reported in routine clinical and ophthalmological examinations. The projected increase in patients suffering from these disorders together with the tedious nature and high cost of manual examination has motivated research into computer assisted tools that can improve precision, consistency and throughput. Vessel segmentation is a key stage in all published work on interactive and fully-automated computer analysis and diagnosis systems and the focus of our research.

Features such as the displacements of blood vessels within the optic disk (OD) have been used for automated diagnosis of diseases such as Glaucoma, age-related macular degeneration (AMD), diabetic retinopathy (DR) and vascular disorders etc. [1] [2][3][4][5]. In many systems the retinal blood vessels are segmented and used to generate a mask which enables the vessels to be eliminated from subsequent stages of the processing pipeline. Since pathologic changes in the vessel structure manifest the progress of various systemic diseases such as hypertension, diabetes, etc., other tools use features extracted directly from segmented retinal blood vessels to detect or grade the condition. Vessels also play a role as landmarks that describe the locations of other elements in the image [6] or provide features that are used for multimodal image registration [7].

The importance of retinal vessel segmentation has motivated the development of accurate retinal vessel segmentation methods. However although much progress has been made, vessel segmentation is still a challenging problem that remains a focus for ongoing research [14][31][32][33]. The primary factors that challenge the accuracy of the segmentation include:

- Variations in vessel and non-vessel objects, such as the wide range of vessel diameters and the presence of pathological features including hemorrhage, exudates and microaneurysms, etc.
- Variations in local and global image appearance due to illumination that in turn may give rise to various image artefacts. For example, reflection at the vessel surface and refraction within it can produce complex intensity profiles and singularities in the image. Additionally, low and unstable local intensity contrast between the vessels and background may cause miss-segmentation of the small capillary vessels.
- Morphological characteristics of vessels (e.g. how they cross and overlap) and the extremely noisy background can influence the accuracy of segmentation and this is especially challenging for segmentation approaches that attempt to recover vessel structures by region growing and tracking of individual vessel components.

Amongst the methods and algorithms that have been presented for retinal vessel segmentation, those that use prior knowledge (supervised methods) produce correspondingly better segmentation results. Generally, supervised segmentation procedures can be divided into two primary stages. First, small local regions sharing similar visual properties are detected by a low-level algorithm, then those candidate

regions are classified as vessel or non-vessel based on a set of features extracted from the region. Machine learning is often employed to train vessel/non-vessel classifiers based on ground truth labels. In this context, ground truth is derived from reference fundus images that have been manually segmented and precisely labelled by one or more clinical experts. Describing such segmentations using the term *ground truth* can be misleading due to the high intra and inter-observer variability in expert manual segmentations. This is some cause for concern and suggests that supervised segmentation methods may benefit from unsupervised training techniques that compensate for this variability.

Although supervised segmentation methods (reviewed in section 2) are more competitive in terms of performance than unsupervised approaches [36] [37][38], their dependence on ground-truth and the problem of intra- and inter-observer variability limits their robustness [34]. In this and many other fields of medical image understanding, ground truth tends to be sparse as the laborious and skilled nature of the task makes it very expensive to acquire. Even when suitable image databases with corresponding ground truth labels are published, inaccuracies in the segmentation labels will lead to poor performance. Computational complexity is also an issue, especially for supervised vessel segmentation methods that use filter banks to extract vessel and non-vessel features. For instance, matched filters can provide a powerful tool to extract numerous features, however they are computationally expensive since many filter kernels need to be applied to the image. Learning matched filter banks is also expensive due to the high cost of clustering and classifying high dimensional feature vectors. With this in mind, we propose a different approach for retinal vessel segmentation in fundus images which extends previously published work [17][18][19]. Our technique uses the

Gabor filter kernel to extract vessel features in a set of training images. We optimize scale parameters for the filter bank by performing receiver operating curve (ROC) analysis on segmentations produced by a conventional texton classifier trained learning using hand labelled pixels (see section 3.1). We then use the filter bank to identify so-called keypoints in a (validation) training set. Candidate keypoints are found by searching for local maxima in filter outputs, both spatially and across scales (see section 3.2) and keypoint descriptors are stored in a dictionary which is indexed by scale. Once we have learnt a set of keypoint features we build the texton dictionary in a further training cycle. The texton dictionary is built by clustering the multi-scale Gabor filter bank outputs from another set of example images. Vessel clusters are seeded by cluster centres corresponding to keypoint pixels (one at each scale) which are identified by a matching process and a further cluster centre is seeded to form a background (non-vessel) cluster. This two-stage approach allows us to decouple parameter selection from training, so avoiding possible over fitting while also supporting the possibility of using different databases for model selection and training. During testing we assign pixels as vessel or non-vessel by using the texton dictionary as a 1-NN classifier. Our experimental results show that the performance achieved by this filter bank is comparable to that of approaches that use more complex kernels.

The rest of this paper is organized as follows. Section 2 reviews some supervised and unsupervised retinal vessel segmentation methods. In section 3.1, we describe the design of a Gabor filter kernel for feature extraction, and explain how we specify and optimise its parameters. In sections 3.2 and 3.3 we describe how we identify keypoints and generate textons. Section 4 describes how we evaluate of our method and compares

performance with that of other methods. Finally, in section 5, we present conclusions, discuss limitations of the approach and propose further work.

2. Related work

Among researchers adopting supervised segmentation methods, Staal et al. [8] extracted vessel related features in a procedure that employs a pre-processor to highlight vessel ridges. These ridges coincide approximately with vessel centrelines and are used to partition the image into multiple regions comprising convex sets of pixels. In total 27 features are calculated from the regions and these are labelled using a *k*-nearest neighbours (*k*-NN) classifier. Soares et al. [9] proposed a scheme that uses the 2-D Gabor wavelet to reduce noise and extract features, which are then classified using a Gaussian mixture model (GMM) classifier. The GMM is derived from a Bayesian classifier which determines whether a pixel is labelled as belonging to the class vessel or not. Ricci and Perfetti [10], group features in three categories before generating feature vectors. Linear features from vessels are calculated based on the average grey level along lines comprising 15 pixels passing through a target pixel at 12 orientations. A line of three pixels in length, orthogonal to the linear features, is used to identify pixels inside vessels, and its grey level provides a third feature. A classification procedure implemented by a linear support vector machine determines vessel/non-vessel labels. Rezatofighi et al. [11] adopt an approach using Local Binary Patterns (LBP) and a morphological method to extract features. Their classifier is implemented by adaptive Multi-Layer Perceptron (MLP) Artificial Neural Networks and an Adaptive Neuro-Fuzzy Inference System (ANFIS). Moin et al. [12] extract a low dimensional feature vector for vessel and non-vessel classification using Gabor wavelets

and LBP. Supervised classification is based on training a GMM classifier using a similar approach to that of Soares et al. [9]. Lupascu and Tegolo et al. [12] collected various vessel related features based on local, spatial and structural properties. They generated a 41-D feature vector for each pixel in the image. Their classifier uses the AdaBoost iterative boosting algorithm. Marín et al. [14] segment retinal vessels with a Neural Network (NN). They classify pixels based on a 7-D feature vector composed of statistical features derived from grey levels. Fraz et al. [15] also proposed a supervised retinal vessel segmentation method, in which each pixel is represented as a 7-D feature vector and segmented using a GMM classifier. Other notable developments of compound filters for retinal vessel features extraction include work presented in Adjeroh and Kandaswamy, et al. [16] and Zhang et al. [17].

Among many unsupervised segmentation methods, Kande et al. [36] proposed an unsupervised retinal vessel segmentation method using the pixel intensity information from both red and green channels to adjust for non-uniform illumination in the colour fundus images, followed by a vessel enhancement stage using matched filtering. Consequently, enhanced vessels are segmented by adopting spatially weighted Fuzzy C-Means (FCM) cluster based thresholding, which takes the spatial distribution of image pixel intensities into account. The final segmentation is obtained by using label filtering to remove some misclassified pixels. Oliveira et al. [37] proposed an unsupervised retinal vessel segmentation procedure using a combined filter which includes a matched filter, Frangi filter and Gabor filter. The combined filter is used for vessel enhancement and corresponding feature extraction. The extracted features are then clustered by FCM based on their observation that the number of non-vessel elements is larger than the vessel elements and consequently this is used in

classification. Wang et al. [38] proposed an unsupervised retinal vessel segmentation method that does not require a pre-processing stage. The vessels are initially enhanced using a matched filter with a multi-wavelet kernel which is capable of responding to blood vessels and non-vessel edges. The classification task is implemented by using an iterative multi-scale hierarchical algorithm based on a two-class decomposition model which is controlled by one optimal scale parameter. The final binary segmentation is derived by adopting an adaptive threshold.

3. Proposed retinal vessel segmentation scheme

In our work, vessel features are extracted using the Gabor filter (see section 3.1). The filter was originally proposed by Dennis Gabor [20] and subsequently adopted by Daugman [21] to model the response of certain cells in the visual cortex of some mammals. The filter has been widely used within the computer vision community to characterise image texture. In this paper, we present a procedure for parameter selection based on the morphological properties of retinal vessels and use a further parameter λ to control the performance. We adopt a multi-scale framework inspired by the scale invariant feature transform (SIFT) [30] to identify keypoints. Filter responses derived from keypoints initialise a clustering algorithm that subsequently determines the textons. Texton-based approaches have been a significant branch of texture analysis since the term texton was first introduced by Julesz in the 1980s [22]. A practical definition of a texton and an operational framework to enable textons to be automatically generated from an image was provided by Leung and Malik [23] [24]. A texton is an element (e.g. line segment, elongated blobs, cross and terminator) that is defined to represent a basic micro geometric textural structure in an image [22].

Building on this idea, Leung and Malik's operational definition characterises a texton by its responses to a filter bank (F_1,F_2,…,F_n).

$$R = \{F_1 * I(x,y), F_2 * I(x,y), \ldots, F_n * I(x,y)\} \qquad (3.1)$$

If the filter bank convolved with an input image I(x,y) the response at each pixel (Eqn. 3.1), will depend on specific local structures (e.g. lines, blobs) and on the design of the filter bank. Should those structures appear repeatedly, then the feature vectors in {R} may be clustered into corresponding groups.

Our filter bank uses multi-scale Gabor filters characterised by one free parameter that determines the scale. The filter bank is used to detect potentially interesting candidate pixels called keypoints. Filter-response vectors derived from keypoints are used to initialise a clustering algorithm that generates textons that are subsequently used to classify vessel/non-vessel pixels in novel images. We identify keypoints from a small training set (N = 10) of fundus images in a process we call *model selection*. We then perform a further training cycle using another set of images to identify the texton dictionary. In this second training stage, a *k*-NN clustering algorithm is initialised with seeds found by matching keypoints identified by the model selection stage to those in the training set (see section 3.3). An overview of the two-stage training procedure is illustrated in Fig. 1. By ranking the clusters by size they are heuristically mapped to vessel/non-vessel classes (i.e. we label the largest cluster as non-vessel).

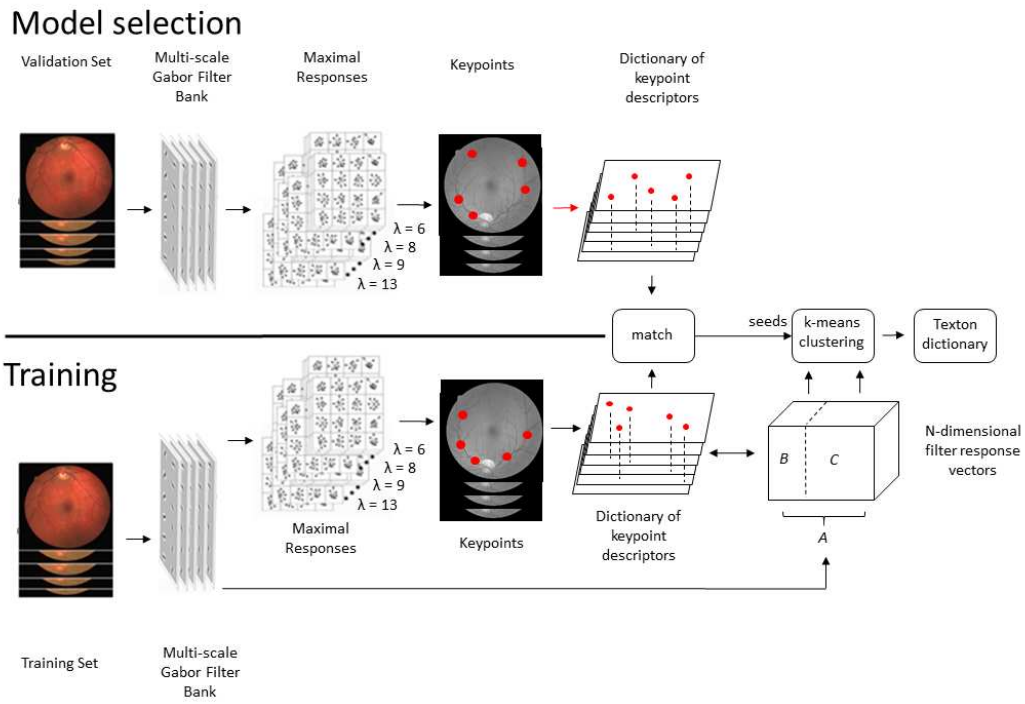


Figure 1: Using keypoints to generate the texton dictionary

We test the system using a 1-NN classifier used to label pixels as vessel/non-vessel in novel fundus images drawn from the publicly available DRIVE database. The DRIVE database, collected by Staal et al. [8] contains 40 images that were captured by a Canon CR5 fundus camera at 45^{o} field of view. Each image is digitized at a resolution of 565×584 pixels captured at 8 bits per colour plane and stored as TIFF format. The data is divided into training and test sets, each set comprising 20 images. Manual segmentations are available for both sets. For the test set, two sets of manual segmentation are provided by two clinical experts. In practice, the first set is used as ground truth whilst the other can be used to provide a measure of inter-observer variation. These measurements are useful for comparing and qualifying results of automatic computer generated segmentation algorithms. For the training set, only a single set of manual vessel tree segmentation is available. Additionally, a mask for each

image is available for both sets that can be used to remove the rim border surrounding the field of interest.

The performance of our proposed method is measured in terms of accuracy, sensitivity and specificity,

$$Sensitivity = \frac{TP}{TP+FN} \tag{3.2}$$

$$Specificity = \frac{TN}{TN+FP} \tag{3.3}$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{3.4}$$

Where *TP, TN, FP, FN* represent True Positive, True Negative, False Positive and False Negative fractions respectively [25].

3.1. Feature extraction using optimized Gabor filters

Psychophysical studies have indicated that the human visual system processes images by analyzing the frequency and orientation of localised components within the visual field [26]. However, while the Fourier transform is able to perform global frequency analysis it is unable to localize signals in the spatial domain. The classical way to combine spatial and frequency information is to introduce spatial dependency into Fourier analysis. For example, the Windowed Fourier Transform (WFT) defines a so-called window function $W(x)$ which allows us to measure the spectrum of the one dimensional function $f(x)$ in the spatial domain of $W(x)$.

$$\mathrm{F}_w(\omega, \varepsilon) = \int_{-\infty}^{\infty} f(x) W(x-\varepsilon)\, e^{-i\omega x}\, dx \tag{3.5}$$

If the window function is Gaussian, the WFT becomes the Gabor transform, named after Dennis Gabor [20] who defined one dimensional function as

$$\mathrm{G}(x) = \mathrm{G}_{re}(x) + i\mathrm{G}_{im}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} e^{i2\pi fx} \qquad (3.6)$$

where $\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{x^2}{2\sigma^2}}$ represents a Gaussian envelope, f is centre frequency and $e^{i2\pi fx}$ is a complex sinusoid which consists of real component $\mathrm{G}_{re}(x)$ and imaginary component $\mathrm{G}_{im}(x)$. The real part also can be denoted as:

$$\mathrm{G}_{re}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \cos(2\pi fx) \qquad (3.7)$$

and the imaginary part as

$$\mathrm{G}_{im}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \sin(2\pi fx) \qquad (3.8)$$

Daugman [21] extended the Gabor concept to two dimensions and adopted it to model the response of simple cells in the receptive field of the visual cortex of some mammals. The two dimensional Gabor function can be defined as

$$\mathrm{G}(x, y) = e^{-\frac{1}{2}(\frac{x^2}{\sigma_x} + \frac{y^2}{\sigma_y})} \cos(2\pi fx + \varphi) \qquad (3.9)$$

where σ_x and σ_y determine the spread of the Gaussian envelope and φ is the phase of the sinusoidal wave. When φ equals $\frac{\pi}{2}$ or $-\frac{\pi}{2}$, the equation becomes an odd-symmetric function. Note, in practice, only the real part of the Gabor filter is convolved with the image weighted by the Gaussian envelope. Since the Gabor filter is localised spatially different image textures can be extracted depending on values of the filter parameters; in our work, we choose the parameters of the Gabor filter kernel by considering the morphological characteristics of retinal vessels. The most important vascular properties are vessel width and their direction (i.e. angle). In practice, the width of the vessel covers a wide range of diameters (from 2-12 pixels), and their length can

extend at any orientation. Hence, the geometric properties of vessels represent linear or bar structures and this information can be used to design appropriate Gabor filters for retinal vessel segmentation.

3.1.1 Design of Gabor filter and parameter optimization

Normally, the Gabor function uses multiple parameters which control the performance of the filter. In Eqn. 3.9, these parameters include f, φ, σ. The phase of the sinusoidal wave φ controls the symmetry of filter kernel. In order to determine the orientation of the filter kernel, we introduce the orientation parameter θ. Hence Eqn. 3.9 becomes:

$$\begin{aligned} \mathrm{G}_\theta(x,y) &= e^{-\frac{1}{2}\left(\frac{x'^2}{\sigma_x}+\frac{y'^2}{\sigma_y}\right)}\cos(2\pi f x'+\varphi) \\ x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \end{aligned} \qquad (3.10)$$

Assuming the spread of the Gaussian envelope in *x* and *y* directions are the same, we set the terms $\sigma_x = \sigma_y$. Because the linear structures (vessel segments) have an approximate rectangular shape, the kernel should be anisotropic, so we insert a parameter γ into Eqn. 3.10, which gives the spatial aspect ratio and determines the ellipticity of the Gabor kernel. In Eqn. 3.10, the spatial frequency term f also can be expressed as $1/\lambda$, where λ is the wave length parameter. Consequently, Eqn. 3.10 can be written as

$$\begin{aligned} G_{\lambda,\theta,\varphi,\sigma,\gamma}(x,y) &= e^{\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)}\cos\left(2\pi\frac{x'}{\lambda}+\varphi\right) \\ x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \end{aligned} \qquad (3.11)$$

If γ=1, the kernel is circular; the term σ is the standard deviation of Gaussian envelope. Because the vessel boundaries in fundus images are presented in a plane approximately normal to the sensor plane their edges are assumed to be parallel, and we model the vessel as an even symmetric function with corresponding *centre-on* and *centre-off* responses given by $\varphi = 0$ and $\varphi = \pi$.

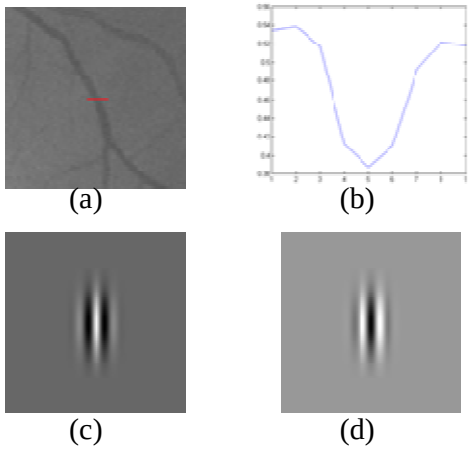


Figure 2: The characteristics of vessel boundaries. (a) shows a panel cropped from a grey-level retinal image; (b) is a grey level profile transept (marked in red) crossing the vessel; (c) and (d) show the responses of the symmetric Gabor kernel with centre-on (c) $\boldsymbol{\varphi = 0}$ and centre-off (d) $\boldsymbol{\varphi = \pi}$.

Since retinal vessels appear relatively darker compared with their background, we choose $\varphi = \pi$ for our kernel model. Neurophysiological research shows that the parameter λ and σ are not independent [27], Petkov and Kruizinga [28] reported that the ratio σ/λ is related to the half-response spatial frequency bandwidth *b* and can be set as follows.

$$\frac{\sigma}{\lambda} = \frac{1}{\pi}\sqrt{\frac{\ln 2}{2}}.\frac{2^b+1}{2^b-1} \qquad (3.12)$$

In practice, the bandwidth b controls the number of excitatory and inhibitory zones in the model. Three zones, one inhibitory and two excitatory are visible in the retinal

vessel structure (Fig. 2(b)), so we choose b=3. We set the spatial aspect ratio γ as 0.85 as our previous work [18] suggests this ratio to be optimal. Since σ and λ are correlated, only one of them (λ) is considered a free parameter. Hence, we rewrite Eqn. 3.11 as

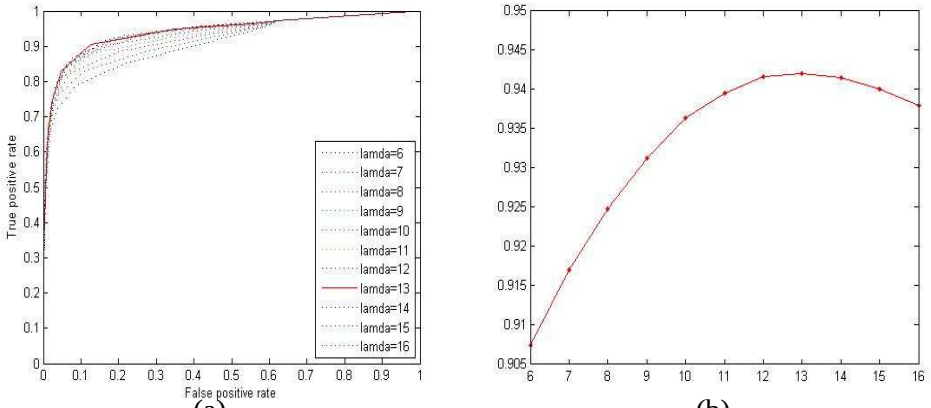
$$G_{\lambda,\theta}(x,y) = e^{\left(-\frac{x'^2+0.72y'^2}{0.12\cdot\lambda^2}\right)}\cos\left(2\pi\frac{x'}{\lambda}+\pi\right)$$
$$x' = x\cos\theta + y\sin\theta \qquad (3.13)$$
$$y' = -x\sin\theta + y\cos\theta$$

The orientation of the retinal vessel is another significant structural characteristic. Consequently, the Gabor filter kernels are designed to cover 12 different orientations in 15^o increments. Hence, our filter bank comprises a set of Gabor kernels parameterised by λ.

The choice of λ, and hence using the relationship in Eqn. 3.12, the kernel size σ, is of primary importance with regard to the Gabor filter's ability to extract vessels and the computational cost. This can be illustrated by plotting a family of ROC curves based on ground truth for a range of filter responses obtained from a training set of images. Typical results are plotted in Fig. 3(a) and corresponding AUCs (area under the curves) are shown in Fig. 3(b). These results suggest a value of $\lambda = 13$ is optimal (with the largest AUC) for a filter bank detecting vessels at a single scale. Using Eqn. 3.12 we find σ is 3.12 and hence a suitable kernel size is 9×9.

To assess the performance of this Gabor filter bank, we use ROC curves to compare its performance in a vessel segmentation test to that of an improved matched filter proposed by Al-Rawi [35] and also to a second order derivate of Gaussian [17]. The matched filters are optimized based on classic matched filter proposed by Chaudhuri et

al. [29]. The results presented in Fig. 4 (a) indicate that the Gabor filter outperforms the matched filter and second order derivative of Gaussian, as we can see the Gabor curve is closer to the top left corner. A typical example image from the training set and the output of the Gabor filter bank ($\lambda = 13$) is illustrated in Fig. 4 (b) and (c) respectively.



(a) (b)

Figure 3: (a) the ROC curves obtained by different lambda values; (b) the AUCs of ROCs with different lambdas.

Although the single scale exhibits good performance and Fig. 4 (c) confirms the main vessels are extracted, many tiny vessels may be missed. To address this problem we adopt a multi-scale approach. A scale range λ=4-15 is chosen empirically to extract key vessel features. By tracking key vessel features across multiple scales, the detection of the keypoints is more robust to noise and this is particularly useful when detecting tiny vessels.

3.2. Identifying keypoints

Our approach to identifying keypoints is inspired by the scale invariant feature transform (SIFT), originally proposed by Lowe [30] to extract distinctive local image features. SIFT identifies so-called keypoint candidates based on responses of a set of difference of Gaussian (DoG) filters. Points with low contrast are discarded and the

location and scale of the remaining keypoints are assigned one or more orientations based on the surrounding local image gradients.

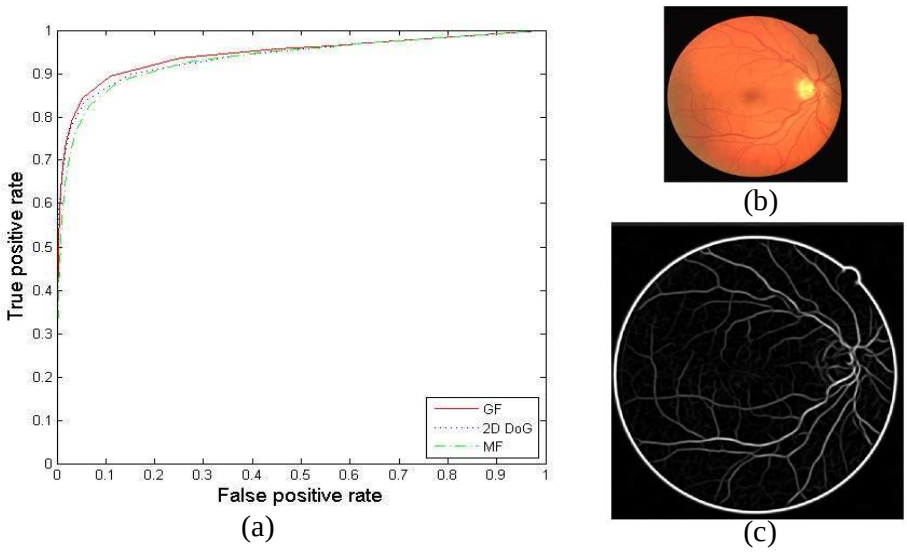


Figure 4: (a) Comparative ROCs between Gabor filter and Matched filter; (c) An example of the Gabor filter response ($\boldsymbol{\lambda = 13}$) for the image shown in (b).

Our approach is somewhat similar, but uses a Gabor filter bank to identify vessel keypoint candidates rather than a DoG filter. We apply this to find points of interest corresponding to scale specific vessel features. These pixels provide initial clusters for a 2-stage *k*-means hybrid clustering process used for subsequently forming the final texton dictionary. The filter bank uses Gabor filters (see section 3.1) at multiple scales $\lambda \in \{4, 6, 8, 9, 13, 15\}$. Although the feature vectors can be calculated for each pixel in the image, to reduce the computational cost only the most representative pixels are used. These are identified by comparing each pixel (marked as cross in Fig. 5) to its 8 neighbours (marked as circles) at the same scale whilst comparing the pixel to its 18 neighbours at adjacent scales (i.e. similar to SIFT). In this way vessel keypoints correspond to local (spatial and scale) maximal filter responses.

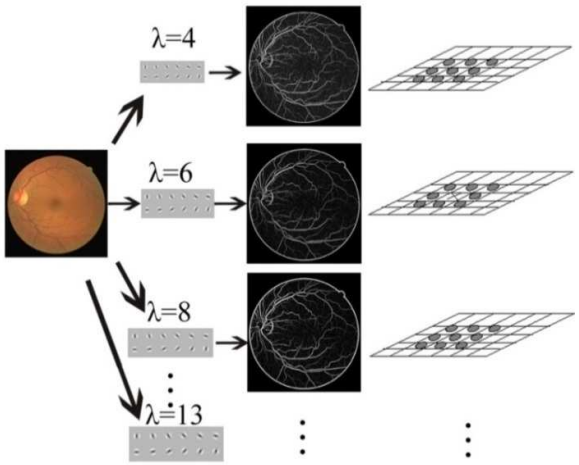


Figure 5: Identifying keypoints

Fig. 6 (a) illustrates how potential keypoints are found using our filters. Not only the main vessels produce keypoints, but also a few keypoints are located in the areas around the end of some capillaries. This is because the multiple scales of Gabor filters match a range of vessel structures in the image. For instance, assuming λ equals 4, 6, and 8, the filter (λ=4) extracts thin vessels thus the responses of capillaries are evident. The filter (λ=8) is tuned to detect wider vessels however the responses of tiny vessels may not be evident. The filter (λ=6) responds to intermediate vessels between λ=4 and λ=8. Comparing filter responses for each pixel to its 26 neighbours and searching for local maxima allows us to extract potential keypoints at four specific scales $\lambda \in \{6, 8, 9, 13\}$.

Potential keypoints are filtered by removing those of low contrast (Fig. 6 (b)). Image gradient and orientation is computed within a local window centred on the remaining keypoints. Image gradient magnitude $m(x, y)$ and corresponding orientation $\theta(x, y)$ are calculated using pixels differences at each scale, expressed as follows:

$$m(x,y) = \sqrt{(R(x+1,y) - R(x-1,y))^2 + (R(x,y+1) - R(x,y-1))^2} \quad (3.14)$$

$$\theta(x,y) = tan^{-1}\left(\frac{R(x+1,y)-R(x-1,y)}{R(x,y+1)-R(x,y-1)}\right) \quad (3.15)$$

where R represents the image smoothed by a Gaussian filter chosen by inheriting the parameters of the Gaussian envelope from those used by the scale specific Gabor function used to identify the keypoints. For example, for $\lambda \in \{4, 6, 8, 9, 13, 15\}$, Eqn. 3.12 gives a corresponding range $\sigma \in \{0.96, 1.45, 1.93, 2.17, 3.13, 3.61\}$. Gradients and orientations are accumulated into the keypoint descriptor by quantising and summing the gradient magnitudes and orientations across scales within a 4 x 4 window. Each descriptor is formed as a 4x4 grid, of which each sub-grid contains an orientation histogram, binned in 8 directions. An example is shown in Fig. 6 (b1) and (b2).

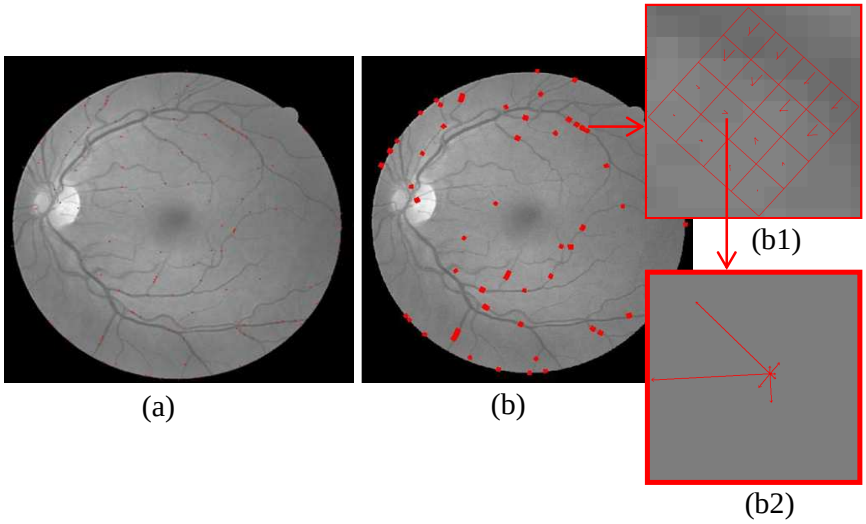


Figure 6: (a) illustrates the potential keypoints and (b) illustrates corresponding descriptors (after filtering low-contrast candidates). One of the descriptors is zoomed in and presented in (b1); one of the orientation histograms presented in (b1) is illustrated in (b2).

To evaluate our approach we extract keypoints from a training set of 20 images drawn from the DRIVE database. We further split these images into two subsets. A validation subset is used for model selection (i.e. identify a set of keypoint descriptors) and the other is used to train textons. For images in the training subset we detect keypoints in the same way we did for model selection, and apply the same scheme to generate their descriptors. Up to this point, each keypoint is represented by its descriptor which is formed as a 4x4x8=128 dimensional vector. We the match the best keypoint candidates found from the training set to keypoint descriptors found from the validation set (using a Euclidean distance metric). Again, this approach is similar to that used to identify natural objects using SIFT [30], but in this case using a matched Gabor filter bank. This ensures that the majority of keypoints map to vessel structures and so there is no need for ground truth labels. In an extreme condition, we can even use all of detected key points derived from a novel image without matching process. The matched key points are used in the next stage to identify seed candidates which are subsequently used to initialise the clustering algorithm when generating textons (Fig. 1). The next section describes how we form the texton dictionary and use it for vessel segmentation.

3.3. Generation of textons and segmentation

Textons are filter responses identified as an *n*-dimensional vector R (see Eqn.3.1) at each pixel position in the image. Filter responses from pixels in *m* training images are clustered into *k* groups using a *k*-means clustering algorithm. The *k*-means algorithm is initialised by filter responses from keypoints and runs iteratively until either it converges or a maximum number of iterations is reached. We choose $k = 5$ to reflect significant classes of objects that are visible in the images, i.e. vessels (at 4 scales) and

background in our experiment. Our *k*-means clustering algorithm is a hybrid. We control the way in which clusters are formed in the feature space by carefully initialising the 5 cluster centres selected by randomly choosing one keypoint descriptor at each scale $\lambda \in \{6, 8, 9, 13\}$ and one "wildcard" descriptor from the validation set dictionary. The 5 keypoint descriptors in the training set dictionary that best match these are then used to index into the filter response vectors to give initial cluster centres Y_i. We denote the set of filter response vectors $A = B \cup C$ where B denotes a set of responses that map to keypoint pixels and C denotes those that arise from other pixels. We then cluster the filter response vectors in two stages. In the first stage we cluster B, then merge the remaining response vectors C and complete the process. The pseudo code for the *k*-means hybrid algorithm is as follows:

Algorithm 1: *k*-means (Hybrid)

Input: *A*, a set of filter response vectors; *B*, a set of keypoint response vectors; *C*, a set of other response vectors (Note: $A = B \cup C$); k, number of clusters (Note: k = 5); Y_i, cluster centres (Note: i = 1..k);

Output: clusters (labelled filter response vectors)

1: Initialise cluster centres Y_i
2: **while** termination condition is not satisfied **do**
3: Assign *B* to closest cluster centre Y_i
4: Update Y_i based on assignment
5: **end while**
6: **while** termination condition is not satisfied **do**
7: Assign *A* to closest cluster centre Y_i
8: Update Y_i based on assignment
9: **end while**

The motivation for the *k*-means hybrid is based on a conclusion reported by Bubeck et al. [39]. Here, they conjecture that "when there is at least one initial centre in each of the underlying clusters, then the initial centres tend to stay in the clusters they had been

placed in". Our approach encourages clusters to form around filter responses that map to four scales of vessel structures. The 'wildcard' clustering centre (k=5) will tend form the background (non-vessel) cluster centre in the second stage of the clustering process and this will attract most membership labels. This *k*-means hybrid not only improves the stability of k-means algorithm, but also ensures the clustering procedure starts from centres that map to a range vessel widths. Since the labelling of clusters is implicit training is unsupervised and can be constructed without assistance of ground truth.

To identify textons, we first rank the clusters based on their size. The largest cluster in the list maps onto the background texton class, since the background comprises the largest number of pixels in an image. The remaining clusters are considered as vessel related textons. We store these textons in a texton dictionary which is used in test stage.

In the test phase, we filter each novel image with the multi-scale Gabor filter bank to generate corresponding responses at each pixel. Pixels are then labelled as vessel/non-vessel by a 1-NN classifier which assigns them to the nearest texton cluster in the dictionary (Fig. 7).

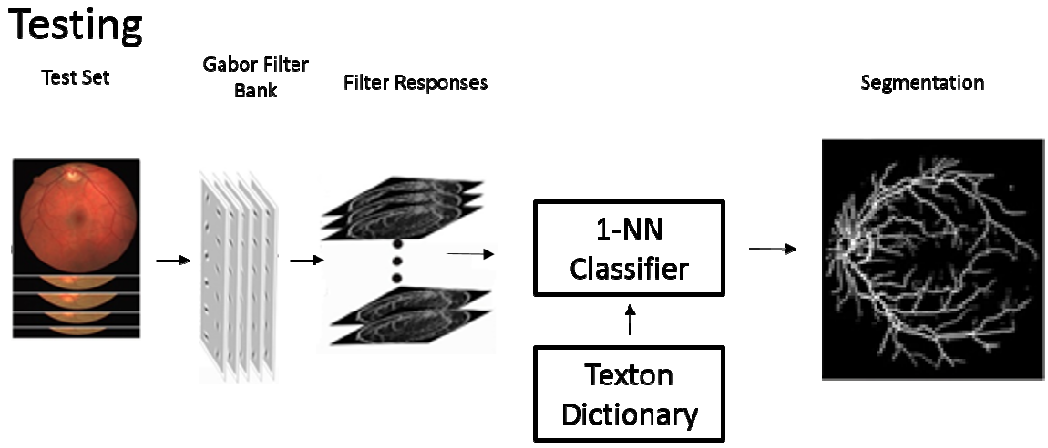


Figure 7: Classifying Vessel/Non-Vessel Pixels

4. Experimental results and evaluation

The proposed method was tested and evaluated on the DRIVE data sets. Each image in test set of the DRIVE database was segmented. Example segmentation results shown in Fig. 8 (g) (h) and (i) illustrate that as well as extracting veins and arties (i.e. vessels with wide diameter) our method also performs well in segmenting capillaries (tiny vessels).

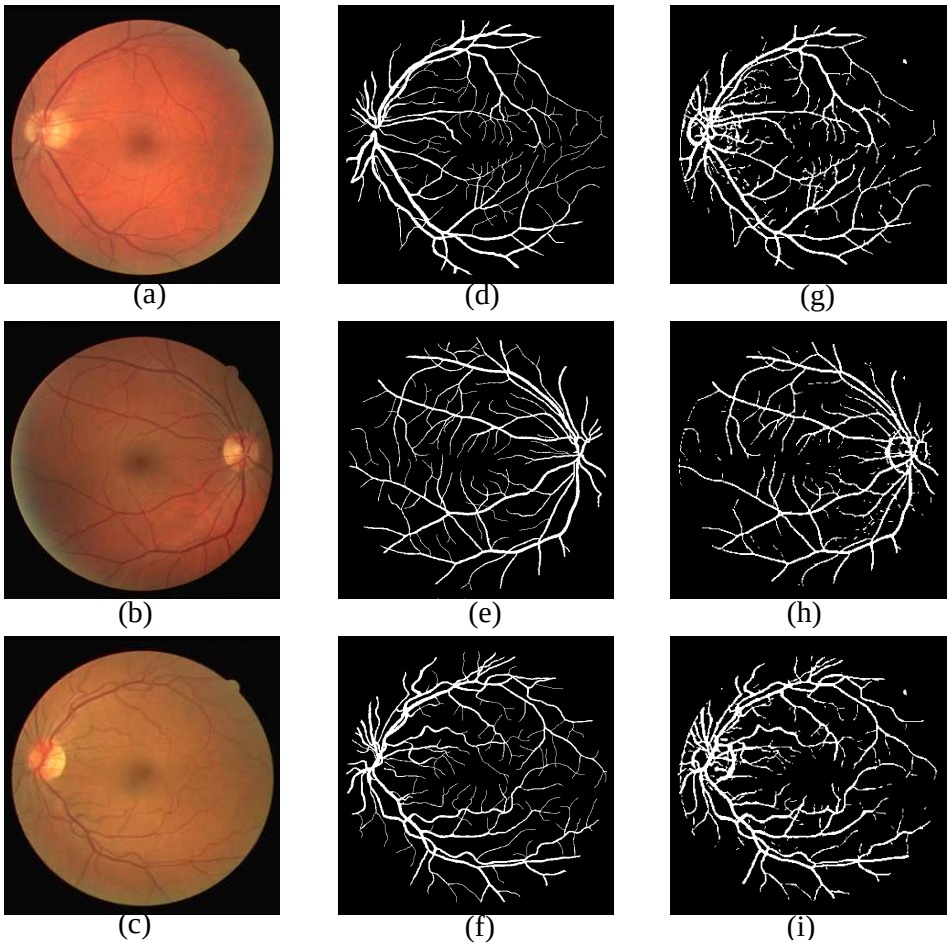


Figure 8: (a)-(c) Images from the DRIVE database; (d)-(f) the corresponding ground truths; (g)-(i) segmentation results achieved by textons using multi-scale Gabor filters and keypoints.

Even the most tiny vessels at the end of vessel network that exhibit a diameter close to the limit of that identified in ground truth are detected. In order to qualify the

performance of the proposed method, segmentation results from each test image were compared to its ground truth. The average specificity, sensitivity and accuracy for our method are 0.9668, 0.7812 and 0.9504 respectively.

Table 1 compares our method with other approaches published between 2004 and 2013. Most of methods presented in table 1 are supervised classifier based methods but a method employing vessel tracing [33] and one based on graph cuts [31] are also included, as both have been proposed very recently and their performance is outstanding. The comparative results show that our method has much better sensitivity than the most of the other methods listed in the table 1.

Table 1: Comparative results between our retinal vessel segmentation method and other state-of-the-art methods on DRIVE database.

Method	Year	Sensitivity	Specificity	Accuracy
2nd observer	-	0.7761	0.9725	0.9473
Staal [8]	2004	0.7194	0.9773	0.9441
Soares [9]	2006	0.7283	0.9788	0.9466
Ricci & Perfetti [10]	2007	-	-	**0.9595**
Rezatofighi [11]	2008	0.7308	0.9723	0.9410
Fraz [15]	2011	0.7525	0.9722	0.9476
Marín et al. [14]	2011	0.7067	0.9801	0.9454
Condurache [32]	2012	0.9094	0.9591	**0.9516**
Ocbagabir et al.* [33]	2013	0.7131	0.9824	**0.9583**
Salazar-Gonzalez* [31]	2010	0.7197	0.9665	0.9479
Our method*	2014	0.7812	0.9668	0.9504

Note: *denote unsupervised approaches; bold values represent accuracy values that outperform our method.

Normally, it is difficult to balance the sensitivity and specificity. Increasing sensitivity tends to reduce specificity and this in turn may change the overall accuracy. In reality, the tiny vessels have extremely low contrast comparing with background, thus if the algorithm is particularly designed for tiny vessel elements extraction in order to increase the sensitivity of segmentation, more non-vessel elements from background may be detected as vessels. This will lead to a reduction in specificity and accuracy. We report an average sensitivity of 0.7812. This figure suggests that using keypoint descriptors and multi-scale Gabor filters may offer advantages in terms of better vessel segmentation performance. The figure for the average specificity of 0.9668 also illustrates that the algorithm maintains good performance and is robust to artefacts present in the background that might be mistaken for vessel structures.

To the best of our knowledge, the maximum accuracy achieved by previous methods is 0.9595. This value is achieved by a supervised approach proposed by Ricci and Perfetti [10]. However they didn't report values of sensitivity and specificity, which are more useful in evaluating the performance of segmentation and classification methods particularly when the testing data is unbalanced (e.g. more background pixels than vessel pixels). The results published by Condurache [32] show the best overall sensitivity (0.9094). However, the sensitivity of their supervised approach exceeds that of the 2nd observer by quite a margin and suggests that there may be a tendency for the uncertainty in ground truth labels and test methodolgy adopted by the DRIVE database to favour supervised techniques. With this in mind, unsupervised approaches such as ours and the vessel tracing method recently proposed by Ocbagabir et al. [33] perform

comparatively well. Compared to our method theirs delivers comparable accuracy (0.9583), but the sensitivity of our method is approximately 7% better while the specificity just 1.5% poorer. Although the lower specificity indicates that more non-vessel elements were segmented as vessels, the better sensitivity demonstrates that our method has performed 7% better with respect to extracting vessels from the background by sacrificing only just 1.5% on specificity. In terms of the evaluation criterion, a primary goal of vessel segmentation is to detect as many vessel elements as possible, since in reality a ratio between diameter of arteries to those of veins also known as A/V ratio is an important clinical biometry to assess the risk of many diseases (e.g. hypertension, brain stroke) [40] However it is also important to detect the tiny vessels as their condition is indicative in diagnosis of Wet Macular Degeneration (WMD). WMD, also known as choroidal neovascularization, is characterised by abnormal growth of the choroidal vascular structure into or around the macula. These new vessels are weak and easily bleed [41]. Achieving good performance in segmenting thin vessels suggests it is better to improve the overall accuracy by pursuing higher sensitivity while maintaining the same or sacrificing only a small fraction of specificity.

We wish to emphasize that the primary goal of this experiment was to develop an unsupervised method that achieves relatively good performance compared with supervised methods. Since ground truth is only used to parameterise the multi-scale Gabor filter bank we believe we have succeeded in this respect. Our experimental results show that using keypoint descriptors and multi-scale Gabor filters delivers competitive performance compared with other supervised methods.

5. Conclusions and summary

The performance of segmentation algorithms that learn from ground truth labels supplied by human observers can suffer due to inconsistencies in the labelling process. Addressing this problem we proposed a novel unsupervised retinal vessel segmentation method that uses keypoint descriptors to generate the texton dictionary used to discriminate between vessel and non-vessel pixels in novel images. Textons are assigned vessel and non-vessel labels implicitly in this approach and as such, our clustering algorithm is not constrained by ground truth. Our multi-scale framework allows us to extract features relating to different vessel widths and these are used to automatically generate textons at salient scales. The evaluation results show that the performance of our proposed unsupervised retinal vessel segmentation is competitive compared to other supervised methods. Additionally, we believe that the scheme we propose for selecting potential vessel candidates based on keypoints can also be used as a general tool for selecting the most appropriate initial seeds in segmentation approaches that trace vessels.

Although the performance of the proposed method (sensitivity, specificity, and accuracy) is good, we can identify some limitations and weaknesses. Visually, some false positive pixels are evident in our segmentation results, especially in an area around the optic disc (OD) and in the left part of the peripheral area. This is because the area surrounding the OD and the outer circle exhibit strong contrast and so there are significant gradient changes on their boundaries. We believe that the overall accuracy of our method can be improved by removing these false positive pixels. A direct way to

handle this limitation is adding a pre-processing stage before implementing segmentation. We intend to address these limitations in our further work.

Conflict of interest statement

The authors declare they have no conflict of interest with regard to the work presented.

References

[1] Nayak, J., Acharya, U. R., Subbanna, P. S, Shetty, N., and Lim, T. C., Automated Diagnosis of Glaucoma Using Digital Fundus Images, J. Med. Syst., 2009; 33:337-46.

[2] Fuente-Arriaga, J. A. D., Felipe-Riverónb, E. M. and Garduño-Calderónc, E., Application of vascular bundle displacement in the optic disc for glaucoma detection using fundus images. Computers in Biology and Medicine. 2014; 47:27-35.

[3] Bhuiyan, A., Karmakar, C., Xiao, D., Ramamohanarao, K. and Kanagasingam, Y., Drusen quantification for early identification of age related macular degeneration (AMD) using color fundus imaging. In: 35rd Annual International Conference of the IEEE, Engineering in Medicine and Biology Society (EMBC), 2013 July 3-7; Osaka, Japan: IEEE; 2013. pp. 7392-5.

[4] Narasimhan, K., Neha, V.C. and Vijayarekha, K., Hypertensive retinopathy diagnosis from fundus images by estimation of AVR. In: Rajesh, R., Ganesh, K. and Koh, S.C.L. (Eds), Procedia Engineering. ICMO2012: international conference on modelling optimization and computing; 2012. 38:980-93.

[5] Usher, D., Dumskyj, M., Himaga, M., Williamson, T. H., Nussey, S. and Boyce, J., Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening. Diabetic Medicine. 2003; 219(1):84-90.

[6] Yun, W. L. at al., Identification of different stages of diabetic retinopathy using retinal optical images. Information Sciences. 2008; 178(1):106-121.

[7] Zana, F. and Klein, J. C., A Multimodal Registration Algorithm of Eye Fundus Images Using Vessels Detection and Hough Transform. IEEE Transactions on Medical Imaging.1999; 18(5):419-428.

[8] Staal, J., Abramoff, M.D., Niemeijer, M., Viergever, M.A. and van Ginneken, B., Ridge based vessel segmentation in color images of the retina. IEEE Transactions on Medical Imaging. 2004; 23(4):501-509.

[9] Soares, J. V. et al., Retinal vessel segmentation using the 2D Gabor wavelet and supervised classification. IEEE Transactions on Medical Imaging. 2006; 25(9):1214-1222.

[10] Ricci, E. and Perfetti, R., Retinal blood vessel segmentation using line operators and support vector classification. IEEE Transactions on Medical Imaging. 2007; 26(10):1357-65.

[11] Rezatofighi, S. H., Roodaki, A. and Noubari, H. A., An enhanced segmentation of blood vessels in retinal images using contourlet. In: 30th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society (EMBC); 2008 Aug 20-24; Vancouver, Canada: IEEE; 2008. pp. 3530-3.

[12] Moin, M. S., Tavakoli, H. R. and Broumandnia, A., A new retinal vessel segmentation method using preprocessed Gabor and local binary patterns, in: Machine Vision and Image Processing (MVIP); 2010 Oct. 27-28; Isfahan, Iran: IEEE; 2010. pp.1-6.

[13] Lupascu, C. A., Tegolo, D. and Trucco, E., FABC: Retinal vessel segmentation using AdaBoost., IEEE Transactions on Information Technology in Biomedicine. 2010; 14(5):1267-74.

[14] Marín, D. and Aquino, A., Gegúndez-Arias, M.E. and Bravo, J. M., A new supervised method for blood vessel segmentation in retinal images by using gray-level and

moment invariants-based features. IEEE Transactions on Medical Imaging. 2011; 30:146-58.

[15] Fraz, M. M. at al., A supervised method for retinal blood vessel segmentation using line strength, multiscale Gabor and morphological features. In: 2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA); 2011 Nov. 16-18; Kuala Lumpur, Malaysian: IEEE; 2011. p. 410-5.

[16] Adjeroh, D. A., Kandaswamy, U. and Odom, J. V., Texton-based segmentation of retinal vessels, Journal of the Optical Society of America. 2007; 24(5): 1384-93.

[17] Zhang, L., Fisher, M. and Wang, W. J., Locating blood vessels in retinal images using unified Textons. In: 17th conference on Medical image understanding and analysis. 2013 July 17-19; Birmingham, UK; 2013. pp. 213-8.

[18] Zhang, L., Fisher, M. and Wang, W. J., Retinal vessel segmentation using Gabor filter and textons, In: 18th conference on Medical image understanding and analysis. 2014 July 9-11; London, UK; 2014. pp. 155-60.

[19] Zhang, L., Fisher, M. and Wang, W., Comparative performance of texton based vascular tree segmentation in retinal images, In: Proc. 21st IEEE International Conference on Image Processing, 2014 Oct. 27-30; Paris, France; IEEE: 2014; pp. 252-56.

[20] Gabor, D., Theory of communication. Journal of the Institute of Electrical Engineers. 1946; 93:429-57.

[21] Daugman, J. G., Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. Journal of the Optical Society of America. 1985; 2(7):1160-9,

[22] Julesz, B., Textons, the elements of texture perception, and their interactions. Nature. 1981; 290(5802):91-7.

[23] Leung, T. and Malik, J., Representing and recognizing the visual appearance of materials using three dimensional textons. Int. J. Comput. 2001; 43:29-44.

[24] Varma, M. and Zisserman, A., A Statistical Approach to Texture Classification from Single Images. International Journal of Computer Vision. 2005; 62:61-81.

[25] Fawcett, T., An introduction to ROC analysis, Pattern Recognition Letters, 2006, 27:861-74.

[26] Campbell, F. W. and Robson, J. G., Application of Fourier Analysis to the Visibility of Gratings. Journal of Physiology. 1968; 197:551-66.

[27] Petkov, N., Biologically motivated computationally intensive approaches to image pattern recognition. Future Generation Computer System. 1995; 11:451-65.

[28] Petkov, N. and Kruizinga, P., Computational models of visual neurons specialized in the detection of periodic and aperiodic oriented visual stimuli: Bar and grating cells. *Biol. Cybern.* 1997; 76:83-96.

[29] Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M. and Goldbaum, M., Detection of blood vessels in retinal images using two-dimensional matched filters. IEEE Transactions on Medical Imaging. 1989; 8(3):263-9.

[30] Lowe, D. G., Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision, 2004; 60(2): 91-110.

[31] Salazar-Gonzalez, A.G., Li, Y. M. and Liu, X. H., Retinal blood vessel segmentation via graph cut. In: Proc. 11th International Conference on Control Automation Robotics & Vision; 2010 Dec. 7-10; Singapore: IEEE; 2010. p. 225 – 30.

[32] Condurache, A. P. and Mertins, A., Segmentation of retinal vessels with a hysteresis binary-classification paradigm. Computerized Medical Imaging and Graphics. 2012; 36(4):325-35.

[33] Ocbagabir, H., Hameed, I., Abdulmalik, S. and Barkana Buket, D., A novel vessel segmentation algorithm in color images of the retina. In: Proc. IEEE Long Island Systems on Applications and Technology Conference. (LISAT); May 3-3, Farmingdale, NY: IEEE, 2013. pp. 1-6.

[34] Li, X., Aldridge, B., Fisher, R. and Rees, J., Estimating the ground truth from multiple individual segmentations incorporating prior pattern analysis with application to skin lesion segmentation. In: Proc. 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Micro, March 20-April 2, Chicago, USA: IEEE: 2011. pp.1438-41.

[35] Al-Rawi, M., Qutaishat, M. and Arrar, M. An improvement matched filter for blood vessel detection of digital retinal images. Computers in Biology and Medicine. 2007, 37(2):262-7.

[36] Kande, G. B., Subbaiah, P. V. and Savithri, T. S., Unsupervised fuzzy based vessel segmentation in pathological digital fundus images. Journal of Medical Systems. 2010, 34(5):849-58.

[37] Oliveira, W. S., Ren, T. I. and Cavalcanti, G. D. C., An Unsupervised Segmentation Method for Retinal Vessel Using Combined Filters. In: Proc. IEEE 24th International Conference on Tools with Artificial Intelligence, Nov. 7-9, Athens, Greece, IEEE: 2012, pp.750-6.

[38] Wang, Y. F., Ji,, G. R., and Lin, P., Retinal vessel segmentation using multi-wavelet kernels and multi-scale hierarchical decomposition. Pattern Recognition. 2013, 46(8):2117-33.

[39] Bubeck, S, Meila, M. and Luxburg, U. V. How the Initialization Affects The Stability Of The K-Means Algorithm, ESAIM: PS, 2012, 16:436-52.

[40] Niemeijer, M. et al., Automated Measurement of the Arteriolar-to-Venular Width Ratio in Digital Color Fundus Photographs, IEEE Transactions on Medical Imaging, 2011, 30(11):1941-50.

[41] Priya, R. and Aruna, P., Automated Diagnosis Of Age-Related Macular Degeneration From Color Retinal Fundus Images, In: Proc. 3rd International Conference on Electronics Computer Technology (ICECT), 2011, 2:227-30.

Highlights

- We present a retinal vessel segmentation approach that uses textons.
- Vessel textons are derived from responses of a multi-scale Gabor filter bank.
- We train on keypoint descriptors instead of labelled ground truth.
- We show our unsupervised approach performs well compared to previous work.
- Our method outperforms other unsupervised approaches on the Drive data set.