

Segmentation of blood vessels from red-free and fluorescein retinal images

M. Elena Martinez-Perez ^{a,*}, Alun D. Hughes ^b, Simon A. Thom ^b,
Anil A. Bharath ^c, Kim H. Parker ^c

^a Department of Computer Science, Institute of Research in Applied Mathematics and Systems, UNAM, Apdo. Postal 20-726, México, DF 01000, Mexico

^b Department of Clinical Pharmacology, Imperial College London, School of Medicine at St. Mary's Hospital, Paddington, London W2 1NY, UK

^c Department of Bioengineering, Imperial College London, Prince Consort Road, London SW7 2BY, UK

Received 24 January 2006; received in revised form 6 November 2006; accepted 9 November 2006

Available online 3 January 2007

Abstract

The morphology of the retinal blood vessels can be an important indicator for diseases like diabetes, hypertension and retinopathy of prematurity (ROP). Thus, the measurement of changes in morphology of arterioles and venules can be of diagnostic value. Here we present a method to automatically segment retinal blood vessels based upon multiscale feature extraction. This method overcomes the problem of variations in contrast inherent in these images by using the first and second spatial derivatives of the intensity image that gives information about vessel topology. This approach also enables the detection of blood vessels of different widths, lengths and orientations. The local maxima over scales of the magnitude of the gradient and the maximum principal curvature of the Hessian tensor are used in a multiple pass region growing procedure. The growth progressively segments the blood vessels using feature information together with spatial information. The algorithm is tested on red-free and fluorescein retinal images, taken from two local and two public databases. Comparison with first public database yields values of 75.05% true positive rate (TPR) and 4.38% false positive rate (FPR). Second database values are of 72.46% TPR and 3.45% FPR. Our results on both public databases were comparable in performance with other authors. However, we conclude that these values are not sensitive enough so as to evaluate the performance of vessel geometry detection. Therefore we propose a new approach that uses measurements of vessel diameters and branching angles as a validation criterion to compare our segmented images with those hand segmented from public databases. Comparisons made between both hand segmented images from public databases showed a large inter-subject variability on geometric values. A last evaluation was made comparing vessel geometric values obtained from our segmented images between red-free and fluorescein paired images with the latter as the “ground truth”. Our results demonstrated that borders found by our method are less biased and follow more consistently the border of the vessel and therefore they yield more confident geometric values.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Blood vessel segmentation; Multiscale feature extraction; Region growing; Retinal imaging

1. Introduction

The eye is a window to the retinal vascular system which is uniquely accessible for the non-invasive, in vivo study of

a continuous vascular bed in humans. The detection and measurement of blood vessels can be used to quantify the severity of disease, as part of the process of automated diagnosis of disease or in the assessment of the effect of therapy. Retinal blood vessels have been shown to change in diameter, branching angles or tortuosity, as a result of a disease, such as hypertension (Stanton et al., 1995; King et al., 1996; Wong and McIntosh, 2005), diabetes mellitus (Skovborg et al., 1969) or retinopathy of prematurity (ROP) (Gelman et al., 2005). Furthermore, retinal

* Corresponding author. Tel.: +5255 5622 3617; fax: +5255 5622 3620.

E-mail addresses: elena@leibniz.iimas.unam.mx (M.E. Martinez-Perez), a.hughes@imperial.ac.uk (A.D. Hughes), s.thom@imperial.ac.uk (S.A. Thom), a.bharath@imperial.ac.uk (A.A. Bharath), k.parker@imperial.ac.uk (K.H. Parker).

arteriolar or venular changes predict development of hypertension (Wong and McIntosh, 2005; Ikram et al., 2006), new onset diabetes (Wong and McIntosh, 2005), progression to diabetic retinopathy (Klein et al., 2004) and development of diabetic renal disease (Wong et al., 2004). Thus a reliable method of vessel segmentation would be valuable for the early detection and characterisation of morphological changes.

Different techniques are used to acquire images of retinal blood vessels, two of them are of our interest in this work. A non-invasive technique is the retinal fundus photograph taken using a green filter. A variation of the latter, widely used clinically which is consider minimally invasive, implies the use of a topical mydriatic drop to dilate the pupil. These images are generally called *red-free*. A more invasive technique is *fluorescein* angiography which involves an intravenous injection of dye that increases the contrast of the blood vessels against the background. In the present work we will be referring only to monochrome images obtained by extracting the green band from RGB colour images. This band is chosen because it is known to show the improved visibility of the retinal blood vessels. Fig. 1 shows an example of two scanned negatives taken from the same eye before (red-free) and after the injection of fluorescein dye (fluorescein). There have been many studies on the detection of blood vessels in medical images in general but only some of them are related to retinal blood vessels in particular. Most of the work on segmentation of retinal images can be categorised into three approaches: those based on line or edge detectors with boundary tracing (Akita and Kuga, 1982; Wu et al., 1995), those based on matched filters, either 1-D profile matching with vessel tracking and local thresholding (Zhou et al., 1994; Gao et al., 1997, 2000; Tolia and Panas, 1998; Jiang and Mojon, 2003) or 2-D matched filters (Chaudhuri et al., 1989; Zana and Klein, 1997; Hoover et al., 2000), and those supervised methods which require manually labelled images for training (Sinthanayothin et al., 1999; Staal et al., 2004).

We have applied some of these methods but because of the large regional variations in intensity inherent in these retinal images and the very low contrast between vessels and the background, particularly in the red-free photographs, the results were disappointing. Techniques based on line or edge detectors lacked robustness in defining blood vessels without fragmentation and techniques based on matched filters were difficult to adapt to the variations of widths and orientation of blood vessels. Furthermore most of these segmentation methods are developed to work either on red-free or fluorescein images but not on both (Martinez-Perez, 2001).

In this paper we present a method based on multiscale analysis from which we obtain retinal blood vessel width approximation, size and orientation using gradient magnitude and maximum principal curvature of the Hessian tensor, two geometric features based upon the first and the second spatial derivatives of the intensity calculated for each different scale that give information about the topology of the image. We then use a multiple pass region growing procedure which progressively segments the blood vessels using the feature information together with spatial information about the eight-neighbouring pixels, obtaining in this way a segmented binary image. The algorithm works equally well with both red-free fundus images and fluorescein angiograms as will be shown in Sections 3 and 4.

A first approximation to the segmentation of retinal blood vessel using this approach was previously presented (Martinez-Perez et al., 1999), where segmentation method was tested on a small image sample without any validation. An extension of this work is presented here and the method is now tested on two local databases and two public databases of complete manually labelled images (Hoover et al., 2000; Staal et al., 2004). We evaluate our segmentation using the public databases that have been also used by other authors for the same purpose (Jiang and Mojon, 2003; Staal et al., 2004). Validation of segmented vessel diameters and branching angles measurements are also made: between red-free against fluorescein images and between our algorithm and one of the public databases.

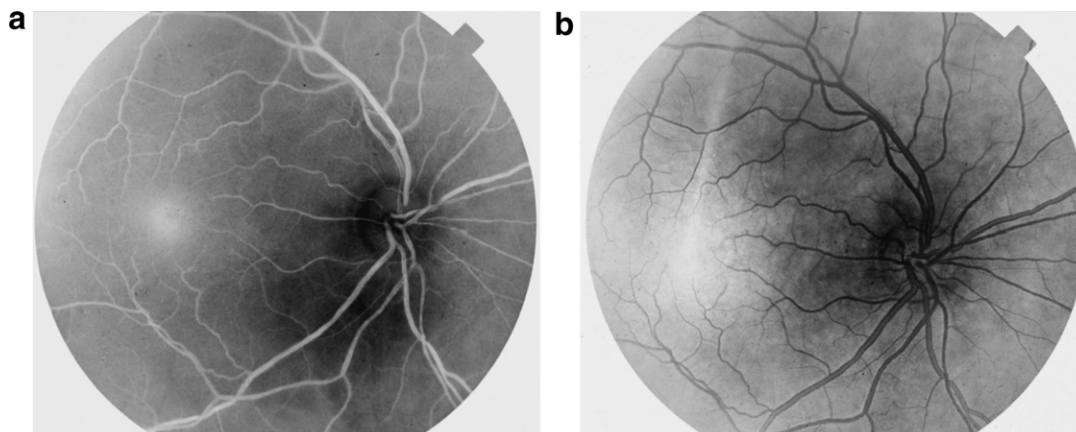


Fig. 1. Negative fundus images from the same eye: (a) red-free and (b) fluorescein. Image size 600×700 pixels.

2. The segmentation method

Since retinal blood vessels have a range of different sizes it is convenient to introduce a measurement that varies within a certain range of scales. Multiscale techniques have been developed to provide a way to isolate information about objects in an image by looking for geometric features at different scales. Under this framework representing information at different scales is defined by convolving the original image $I(x, y)$ with a Gaussian kernel $G(x, y; s)$ of variance s^2 :

$$I_s(x, y; s) = I(x, y) \otimes G(x, y; s) \quad (1)$$

where G is:

$$G(x, y; s) = \frac{1}{2\pi s^2} e^{-\frac{x^2+y^2}{2s^2}} \quad (2)$$

and s is a length scale factor. The effect of convolving an image with a Gaussian kernel is to suppress most of the structures in the image with a characteristic length less than s . Fig. 2 shows different convolution outputs of a portion of a negative red-free retinal image, for $s = 0, 2, 8$ and 14 pixels, showing the progressive blurring of the image as the length scale factor increases.

The use of Gaussian kernels to generate multiscale information ensures that the image analysis is invariant with respect to translation, rotation and size (Koenderink, 1984; Witkin, 1984). To extract geometric features from an image, a framework based on differentiation is used. Derivatives of an image can be numerically approximated by a linear convolution of the image with scale-normalised derivatives of the Gaussian kernel.

$$\partial^n I_s(x, y; s) = I(x, y) \otimes s^n \partial^n G(x, y; s) \quad (3)$$

where n indicates the order of the derivative.

The normalisation by scale makes the derivatives dimensionless which means that the derivatives will have the proper behaviour under spatial rescaling of the original image and that structures at different scales will be treated in a similar manner. Differential geometric structures that are used to characterise retinal blood vessels in this work are based on the first and second derivatives which can be extracted by the multiscale approach.

2.1. Feature extraction

Detection of tube-like structures using multiscale approach has been carried out by other researchers (Frangi et al., 1998; Lorenz et al., 1997; Sato et al., 1997). The main purpose of these works was to develop a line-enhancement filter based on the eigenvalue analysis of the Hessian matrix. The filters were applied to 2-D and 3-D medical images such as digital subtraction angiography or magnetic resonance angiography of blood vessels and computer tomography of airways. We use similar information in combination with gradient information to segment blood vessels rather than to enhance them.

Gradient magnitude. The first directional derivatives describe the variation of image intensity in the neighbourhood of a point. Under a multiscale framework, the magnitude of the gradient $|\nabla I_s| = \sqrt{(\partial_x I_s)^2 + (\partial_y I_s)^2}$, represents the slope of the image intensity for a particular value of the scale parameter s . Fig. 3 shows the gradient magnitude at different scales for the subimages shown in Fig. 2, for $s = 2, 8$ and 14 pixels.

Principal curvature. The second directional derivatives describe the variation in the gradient of intensity in the neighbourhood of a point. Since vessels appear as ridge-like structures in the images, we look for pixels where the intensity image has a local maximum in the direction for which the gradient of the image undergoes the largest change (largest concavity) (Eberly, 1996). The second derivative information is derived from the Hessian of the intensity image $I(x, y)$:

$$H = \begin{pmatrix} \partial_{xx} I_s & \partial_{xy} I_s \\ \partial_{yx} I_s & \partial_{yy} I_s \end{pmatrix} \quad (4)$$

Since $\partial_{xy} I_s = \partial_{yx} I_s$ the Hessian matrix is symmetrical with real eigenvalues and orthogonal eigenvectors which are rotation invariant. The eigenvalues, λ_+ and λ_- , where we take $\lambda_+ \geq \lambda_-$, measure convexity and concavity in the corresponding eigendirections. Fig. 4 shows the profile across a negative of a red-free vessel and the corresponding eigenvalues of the Hessian matrix, where $\lambda_+ \approx 0$ and $\lambda_- \ll 0$ for pixels in the vessel. For a negative fluorescein vessel profile, where vessels are darker than the background, the eigenvalues are $\lambda_- \approx 0$ and $\lambda_+ \gg 0$ for vessel pixels.

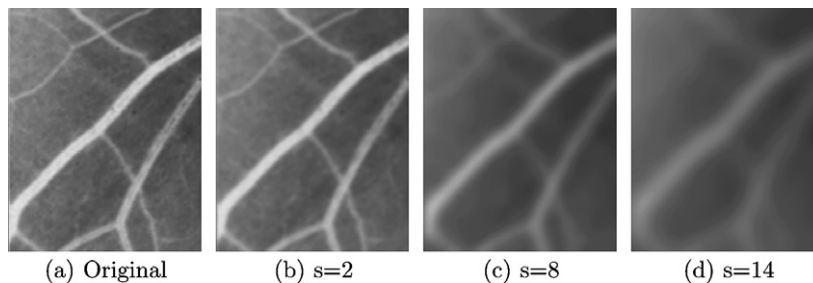


Fig. 2. Multiscale convolution outputs for $s = 0, 2, 8$ and 14 pixels of a portion (720×580 pixels) of a negative of a red-free retinal image (2800×2400 pixels).

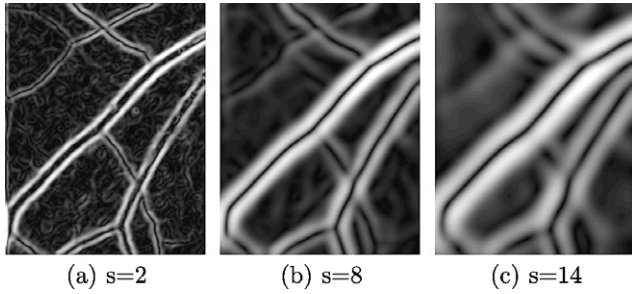


Fig. 3. Magnitude of the gradient at different scales for: $s = 2, 8$, and 14 pixels.

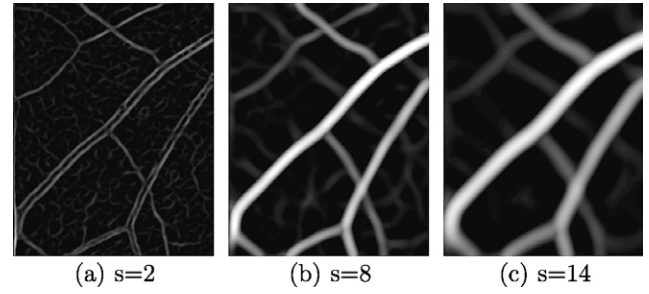


Fig. 5. Maximum principal curvature λ_2 at different scales for: $s = 2, 8$, and 14 pixels.

In order to analyse both red-free and fluorescein images with the same algorithm, we define $\lambda_1 = \min(|\lambda_+|, |\lambda_-|)$ and $\lambda_2 = \max(|\lambda_+|, |\lambda_-|)$. The maximum eigenvalue, λ_2 , corresponds to the maximum principal curvature of the Hessian tensor, which we will refer to as *maximum principal curvature*. Thus, a pixel belonging to a vessel region will be weighted as a vessel pixel if $\lambda_2 \gg 1$, for both red-free and fluorescein images. Fig. 5 shows the maximum principal curvature λ_2 at different scales for the subimages shown in Fig. 2, for $s = 2, 8$ and 14 pixels.

2.2. Multiscale integration based on a diameter-dependent equalisation factor

We calculate the features for all integer values of s , $s_{\min} \leq s \leq s_{\max}$, where s_{\min} and s_{\max} are fixed according to the approximate sizes of the smallest and largest vessel radius to be detected in the image. These two parameters have to be known *a priori* and will depend on the pixel resolution of the original image as well as on the field of view of the fundus camera.

Although the multiscale approach ensures that all spatial points and all scales are treated equally and consistently, it does not address the problem of how to select the appropriate scales for further analysis. An early formulation for feature detection over scales has been proposed by Lindeberg (Lindeberg, 1993). The basic idea of this approach is to apply the feature detector at all scales, and then select levels from the scales at which normalised measures of feature strength assume local maxima with respect to the scale. Some methodologies, particularly developed to enhance tube-like structures, have suggested the maximisation of feature strength over the scales, either maximising normalised measures of filter response (Frangi et al., 1998) or by applying an equalisation of noise level to the normalised derivatives (Sato et al., 1997).

Our approach is similar in that it is based on extracting the information across the scales by finding the local maxima over scales for both measurements of feature strength, but applying an equalisation factor based on the diameter of the vessel we are looking for at a certain scale. If we compute the local maximum along the scales for $|\nabla I_s|$ and

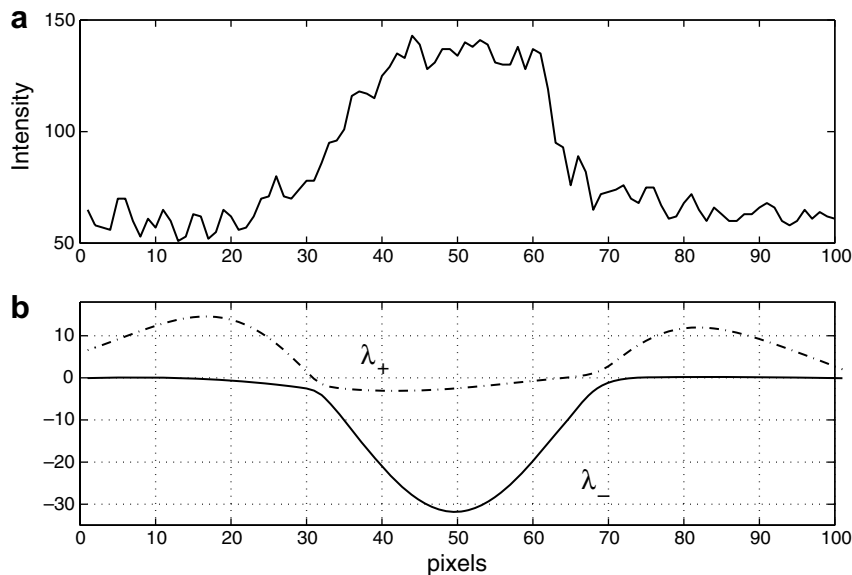


Fig. 4. (a) Intensity profile across a blood vessel from a negative of a red-free image. (b) Eigenvalues, λ_+ (dashed line) and λ_- (solid line). Ridges are regions where $\lambda_+ \approx 0$ and $\lambda_- \ll 0$.

λ_2 from the subimages shown in Figs. 3 and 5, we find that the local maximum response is much higher for large blood vessels than for small ones (Figs. 6a and b).

This might be expected, particularly for maximum principal curvature, (λ_2), since the vessels are approximately cylindrical so that the total amount of blood in the light path corresponding to each pixel is larger in large vessels. Thus, there will be more absorption of non-red light in the red-free images and increased fluorescence in fluorescein images in the larger vessels, and therefore higher principal curvatures. In fact this is one qualitative feature that ophthalmologist normally use to differentiate veins from arteries, veins are generally wider and brighter/darker than arteries.

These changes have been reported before by Brinchmann-Hansen and Engvold (1986) who measured a number of parameters from vessel profiles on both arteries and veins, using a scanning micro-densitometric technique. They found greater relative intensity (I_0) of blood column in veins than in arteries, and an increase of I_0 with an increasing diameter in both.

Numerous attempts have been carried out in order to find a model that can describe the very complex effects of multiple light path on retinal fundus images (Smith et al., 2000; Hammer et al., 2001). A Monte Carlo simulation of retinal vessel profiles performed by Hammer et al. (2001) predicts that light that is observed from a retinal vessel at 560 nm wavelength (green band on a RGB colour

image) is predominantly backscattered from the vessel or transmitted once through the vessel.

Unfortunately, *in vivo* measurements the relative magnitude of these components depends on numerous factors including the absorption by the retinal pigment epithelium, absorption by the choroidal blood, the diameter of the pupil, specular and diffuse reflectance properties of the ocular fundus, among others. One of the main effects disturbing the profiles of the small vessels is that the transmittance factor of the red blood cells represents light that has been scattered into angles that are not collected by the instrument. These photons are lost, resulting in an apparent increase in vessel absorption (due to the surrounding area of small vessels) therefore an increase in transmittance. The result is a decrease of contrast between small blood vessels and the retinal background, which make them to look dimmer.

To account for these effects, we introduce a *diameter-dependent* equalisation factor for both strength features at each scale. Since our best approximation of vessel radius at this stage of the algorithm is the scale factor, we used this in our equalisation. Thus, vessels with diameter $d \approx 2s$ are most strongly detected when the scale factor is s , we normalised each feature along scales by d and then kept the local maxima over scales:

$$\gamma = \max_s \left[\frac{|\nabla I_s(s)|}{d} \right]; \quad \kappa = \max_s \left[\frac{\lambda_2(s)}{d} \right] \quad (5)$$

Figs. 6c and d show the local maxima over the scales from the subimage shown in Figs. 3 and 5 after applying the *diameter-dependent* equalisation factor, d , to each scale. We stress that the scale at which we find the local maximum for κ is not used in our algorithm to define the width of the vessel but rather as an approximation that will be refined during the region growing phase of the algorithm.

These values, the local maxima of the feature equalised gradient magnitude, γ , and the local maxima of the feature equalised maximum principal curvature, κ , are the two features used to classify pixels in the image into two region classes, *background* and *vessel*, using a multiple pass region growing procedure.

2.3. Region growing

The region growing algorithm we use is based on an iterative relaxation technique. All the parameters used in the region growing are automatically calculated for each image from the histograms of the extracted features. The classification of pixels as *vessel* or *background* is based primarily upon the maximum principal curvature κ , from which the criteria for determining seeds are defined. Using spatial information from the classification of the eight-neighbouring pixels, classes grow initially in regions with low gradient magnitude, γ , allowing a relatively broad and fast classification while suppressing classification in the edge regions where the gradients are large. In a second stage, the classification constraint is relaxed and classes grow

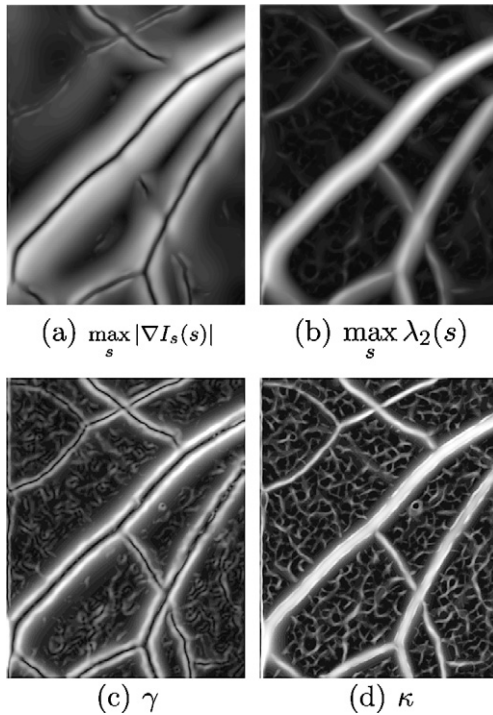


Fig. 6. Local maxima response over the scales for: (a) gradient magnitude $|\nabla I_s|$, (b) maximum principal curvature λ_2 , (c) *diameter-dependent* equalised gradient magnitude γ , and (d) *diameter-dependent* equalised maximum principal curvature κ .

based solely upon κ to allow the definition of borders between regions.

The histograms of both features $h(\gamma)$ and $h(\kappa)$ are calculated and for $h(\gamma)$ only one class is used: *low gradient*, which is defined as $\gamma < \mu_g + \sigma_g$ for the complete histogram (Fig. 7a), where μ_g is the mean and σ_g is the standard deviation. $h(\kappa)$ is automatically divided into two classes using the Otsu threshold algorithm (Otsu, 1979), in which an optimal threshold is selected by discriminant criteria, where the separability between the resulted classes in gray levels is maximised. It uses the zero and first order cumulative moments of the gray level histogram. The means and standard deviations of $h(\kappa)$ for each class are calculated: *background*, for $\kappa/\kappa_{\max} \in [0, t]$ with mean μ_b and standard

deviation σ_b ; and *vessel*, for $\kappa/\kappa_{\max} \in (t, 1]$ with mean μ_v and standard deviation σ_v , where t is the threshold (Fig. 7b). The size of the interval for each class varies depending on the value of a parameter a , which changes during the iteration process. Intervals for each class are defined as: $\mu \pm a\sigma$ (see Fig. 9).

The algorithm begins by planting seeds for each region: *background seeds* are pixels for which $\kappa \leq \mu_b$, whereas *vessel seeds* are defined as $\kappa \geq \mu_v$. Fig. 8a shows an example of the planting stage where vessel seeds are shown in black, background seeds in gray and unknown pixels in white. These conditions ensure that only those pixels in the inner areas of each region, with a very high probability of belonging to the region, are labelled as seeds.

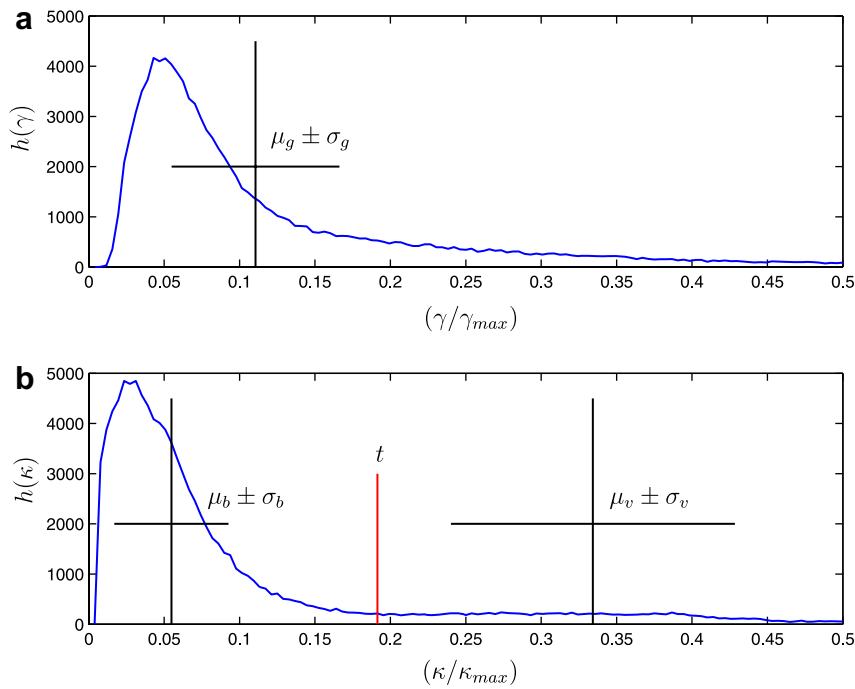


Fig. 7. Parameters used in the region growing algorithm. (a) Histogram of the local maxima of the gradient magnitude strength, γ . One class: *low gradient*, $\gamma < \mu_g + \sigma_g$. (b) Histogram of the local maxima of the maximum principal curvature strength, κ , where t is the threshold. Two classes: *background*, $\kappa/\kappa_{\max} \in [0, t]$ and *vessel*, $\kappa/\kappa_{\max} \in (t, 1]$.

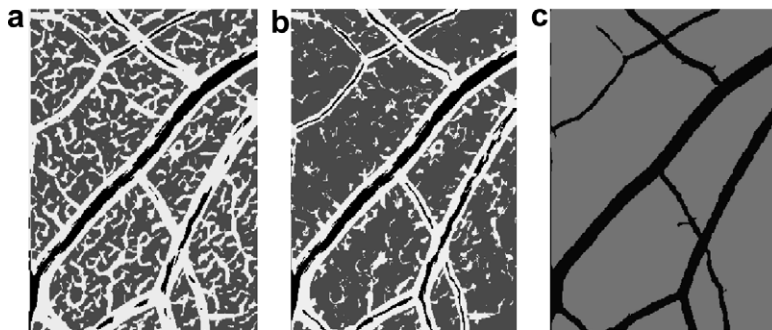


Fig. 8. Region growing algorithm. (a) Planting seeds: black, vessel seeds; gray, background seeds and white, unknown pixels. Two stages of region growing where information of the eight-neighbouring pixels is used. (b) Stage one: growth is restricted to regions with low gradients, allowing vessels to grow where the values of κ lie within a wide interval and (c) stage two: classes grow without the gradient restriction to define borders between classes.

Region growing is an iterative process: an unlabelled pixel is classified as belonging to class j if it has at least one neighbour of class j already classified and if it fulfils a specific condition with initial parameters $a_i = 1$. a_i will specify the size of the class interval for the iteration i . Fig. 9 shows a case where the value of the pixel falls within the current interval for class 2 and will be classified as 2 since it also has nearest neighbours already classified as 2. A grey value which may belong to two different classes is assigned to the one to which most of its neighbours belong. Growing is repeated from left to right and top to bottom until no more pixels are classified. The constraints are relaxed by incrementing the parameters a_i in steps of 0.5 and the growing is repeated.

After the seeds are planted, the conditions for growing are defined as follows: in the first stage, the growing for both classes is restricted to regions with low gradients allowing rapid growth of regions outside of the boundaries, and allowing vessels to grow where the values of κ lie within a wide interval. The condition for class *vessel* is:

$$(\mu_v - a_v \sigma_v) \leq \kappa \text{ AND } \gamma \leq (\mu_g + a_g \sigma_g) \text{ AND } N_v \geq 1 \quad (6)$$

and for class *background*:

$$\kappa \leq (\mu_b + a_b \sigma_b) \text{ AND } \gamma \leq \mu_g \text{ AND } N_b \geq 1 \quad (7)$$

where N_j is the number of neighbours already labelled as class j . Fig. 8b shows the result of the first stage after the restricted growing.

After alternating these two steps until no further classifications are found, the final stage of the algorithm grows vessel and background classes simultaneously without the gradient restriction. Now the condition for class j is:

$$(\mu_j - a \sigma_j) \leq \kappa \leq (\mu_j + a \sigma_j) \text{ AND } N_j \geq 1 \quad (8)$$

and again the condition is relaxed by increasing a from $a = 1$ by steps of 0.5 until all pixels are classified. With this final stage, borders between classes are defined. Fig. 8c shows the result of the growing after this second stage. The result of the whole region growing algorithm just

described, is an adaptive thresholding which depends on both spectral and spatial local information.

3. Results

A total of 114 images from four different sources have been tested. Two sets were provided by St. Mary's Hospital, London. The first one corresponds to 20 images of 10 normotensive and 10 hypertensive subjects. The second one corresponds to 17 paired red-free and fluorescein images (a total of 34). The other two sources tested come from two public databases with 20 and 40 images, respectively (Hoover et al., 2000; Staal et al., 2004).

Both sets provided by St. Mary's Hospital, London are retinal photographs taken using a fundus camera with a 30° field of view (Kowa FX-50R, Kowa, Tokyo, Japan) using Ilford FP4 (125 ASA) photographic film (Ilford Imaging UK Ltd., Knutsford, England). Photographic negatives were digitised using a Nikon 35 mm film scanner (LS-1000, Nikon, Tokyo, Japan). Digitised images were 2400 × 2800 pixels in size. For the analysis of complete fundus images, these were reduced to 600 × 700 pixels by applying a Gaussian pyramid low-pass filter (Burt, 1984).

The two public databases will be described in Section 4. The complete results for all the databases where the present method was tested can be found at: <http://turing.iimas.unam.mx/~elena/Projects/segmenta/Vessel-Segment.html>

Fig. 10 shows the segmentation process described in the previous section: (a) negative red-free gray scale image, (b) γ , (c) κ and (d) the segmented image. For the reduced size images, an interval of $1 \leq s \leq 12$ pixels in steps of 1 pixel was used. It is noticeable that some vessels that are visible in Fig. 10c are missed in Fig. 10d; this is due mainly to the parameters used in the planting seeds stage. We have made a conservative choice by using only those pixels which have a very high probability of belonging to each class. If we use lower levels we include more *vessels* but also more of the undesirable artifacts in the background such as choroidal vessels.

Fig. 11 shows the segmented images corresponding to the original images shown in Fig. 1a and b, red-free and fluorescein, respectively. It can be seen that despite the much poorer contrast, the red-free image segmentation is nearly as good as that of the fluorescein image. Of course, since the fluorescein image has a better contrast, more vessels are segmented. We also note that the segmentation is relatively insensitive to the wide variations in intensity that are inherent in these images. In the fluorescein image, for example, the detail in the segmented image in the much brighter region just beside the darker optic disk in the left centre of the image is nearly the same as in the higher contrast portions of the image (refer to Fig. 1).

Fig. 12 shows another example of paired red-free and fluorescein images in which other components such as exudates, haemorrhages or microaneurysms show up as very dark regions. The blood vessels are detected adequately

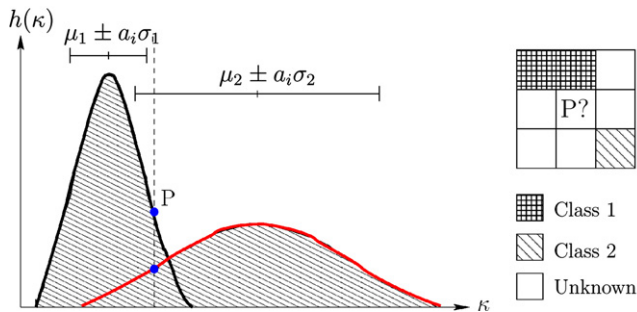


Fig. 9. From the histogram of κ , intervals for each class, j , are defined as: $\mu_j \pm a_i \sigma_j$, where a_i will specify the size of the class interval for the iteration i . An unlabelled pixel P is classified as belonging to class j if it lies inside the interval j and if it has at least one neighbour of class j already classified.

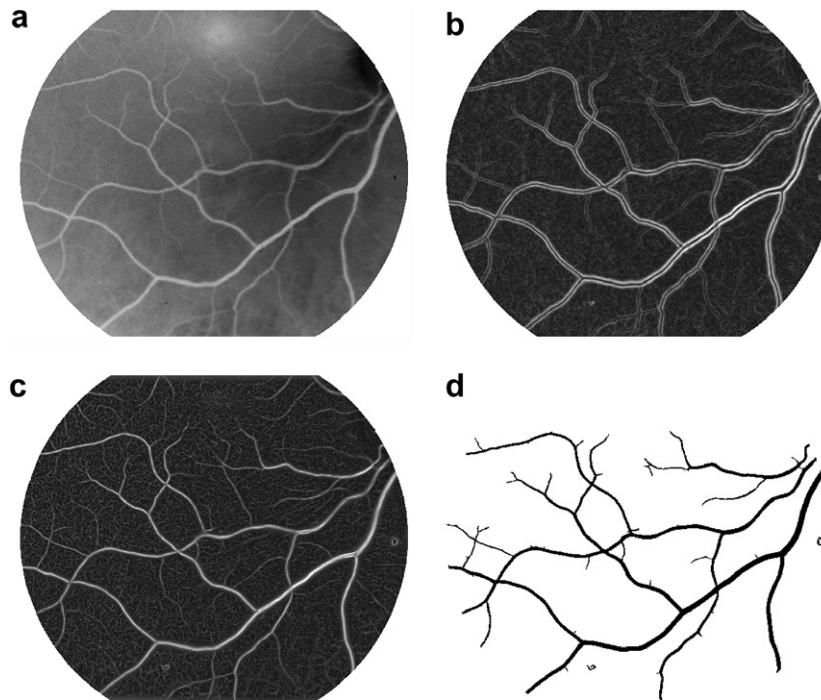


Fig. 10. The segmentation method: (a) original, (b) γ , (c) κ and (d) segmented. Image size 600×700 pixels.

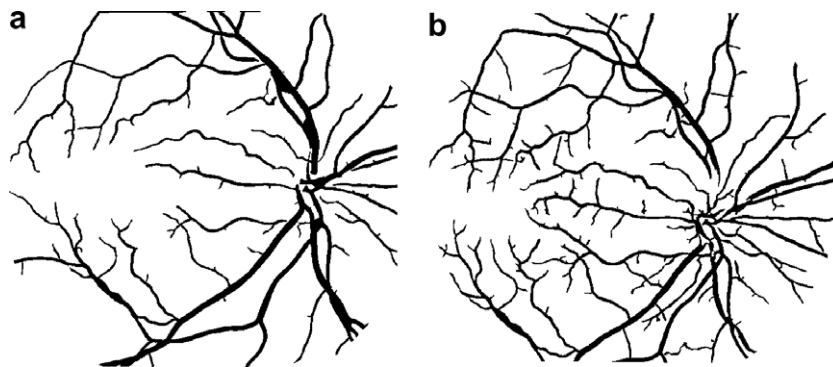


Fig. 11. Segmented images from Fig. 1, respectively: (a) red-free, (b) fluorescein. Image size 600×700 pixels.

despite of the low contrast, but the abnormalities are detected as dark regions attached at the end of some branches. Other detectors, such as those implemented by Sinthanayothin et al. (1999), have to be applied in combination with the blood vessel segmentation process for these kind of images.

4. Validation of segmentation

Before using these results in clinical studies, a validation of the accuracy of the segmentation is necessary. Since direct in vivo measurements of retinal blood vessels in humans are not feasible, three indirect validation studies were undertaken: (1) a comparison of our multiscale (MS) segmentation results with two public databases of hand-segmented images; (2) a comparison of automatic measurements of diameters taken from one MS segmented

image and its corresponding hand-segmented image from one public database and (3) a comparison of automatic measurements of diameters and branching angles made in pairs of red-free and fluorescein images taken in the same eye.

4.1. Comparison with public databases

4.1.1. HOOVER database

A public database has been made available by Hoover et al. (2000). This database, which we will call HOOVER, consists of 20 retinal fundus slides captured by a TopCon TRV-50 fundus camera at 35° of field of view. Each slide was digitised to produce a 605×700 pixel image, with 24 bit per pixel (standard RGB). 10 of the images are of subjects with no pathology (normal) and the other 10 with pathology (abnormal). Each of these 20 images were

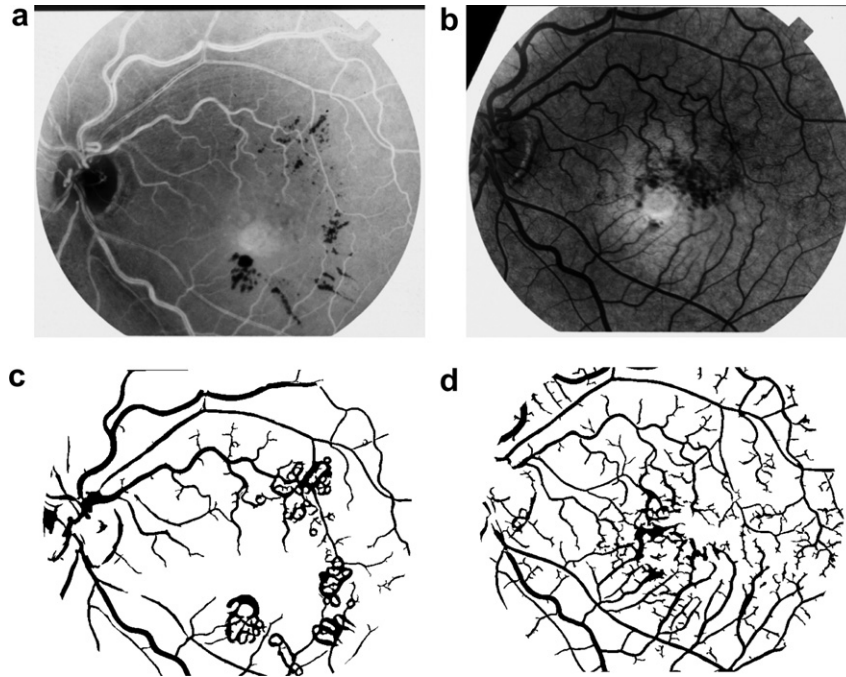


Fig. 12. Example of a pair of fundus negative images which contain other components: (a) red-free image, contains presumable drusen or exudates (showing as dark markings in the inverted image); and (b) fluorescein image, contains a large haemorrhage close to the fovea and several microaneurysms. These components appear in the corresponding segmented images (c) and (d) as dark regions at the end of some branches.

hand-segmented by two different observers, both image processing experts with 2 years of experience working in retinal images. The first observer (AH) labelled fewer vessels than the second (VK). Fig. 13a shows one image from this database segmented by the MS method (using only the green band) and Fig. 13b and c show the two corresponding hand-segmented images, by AH and VK, respectively.

The validations that have been performed by other authors (Jiang and Mojon, 2003; Hoover et al., 2000; Staal et al., 2004) using HOOVER database are based on contin-

gency tables of true classifications. We will follow the same idea in order to compare our results with theirs. We will call our multiscale algorithm of segmentation MS, and we will use, as other authors, AH hand-segmented images as the “ground truth” (GT). Contingency tables are built as follows: any pixel which is marked as *vessel* in both GT and MS is a true positive (TP). Any pixel which is marked as *non-vessel* in both GT and MS is a true negative (TN). Any pixel which is marked as *vessel* in MS but *non-vessel* in GT is a false positive (FP). The true positive rate

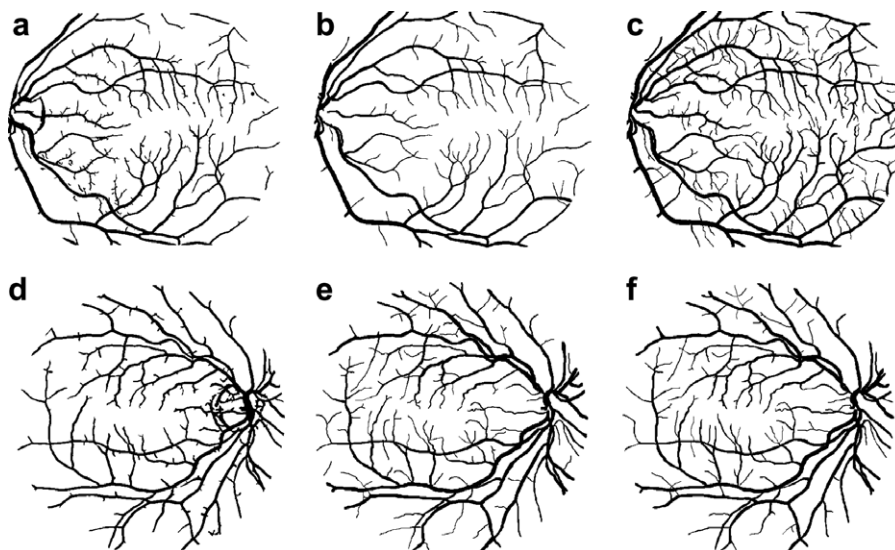


Fig. 13. Top row: image from HOOVER database: (a) MS method; (b) observer AH and (c) observer VK. Bottom row: image from DRIVE database; (d) MS method; (e) observer A and (f) observer B.

(TPR) is established by dividing the number of true positives by the total number of *vessels* in GT. The false positive rate (FPR) is computed by dividing the number of false positives by the total number of *non-vessels* in GT. A measurement of accuracy (Ac) can also be defined by the sum of true positives and true negatives divided by the total number of pixels in the image. For a perfect segmentation, values of true positive rate should be 1, whereas for false positive rate should be 0.

The top of Table 1 shows all these values when comparing MS and VK with AH as the “ground truth”. We can see that the values of both false positive rates are about the same whereas both values of the true positive rate and accuracy are larger for VK than for MS. The latter could suggest room for an improvement of 14.45% in the true positive rate for our method if we use the second observer as the target performance level (TPR = 0.8951). This target was also used by authors Hoover et al. (2000) and Jiang and Mojon (2003) to evaluate the performance of their algorithms with 14.51% and 6.11% of room of improvement, respectively.

The bottom of Table 1 shows the same values but this time taking VK as the “ground truth”. In this case, true positive rate is larger for AH than MS, false positive rate for AH is smaller than MS and therefore accuracy of AH is larger than MS. The latter could suggest a room for an improvement of 3.7% in the true positive rate and of 2.7% in the false positive rate over our method if we use the first observer as the target performance level (TPR = 0.6424).

Table 2 shows the values of true positive and false positive rates reported by other authors using HOOVER database. In the first row of Table 2 we see the values for the second observer against AH which we take as the target performance level the same as Jiang and Mojon (2003) and Hoover et al. (2000). From Table 2 MS method and Hoover et al. (2000) have about the same performance whereas Jiang and Mojon (2003) reported a higher one. Staal et al. (2004) reported the lowest value of true positive rate, but also the lowest value of false positive rate, which resulted in achieving the highest accuracy score among all the other experiments, including those from the hand-

Table 2

Comparisons between different methods as reported by the authors using HOOVER database

Method	TPR	FPR	Ac	<i>n</i>
VK–AH	0.8951	0.0438	0.9522	20
Hoover et al. (2000)	0.7500	0.0438	–	20
Jiang and Mojon (2003)	0.8340	0.0438	–	20
Staal et al. (2004)	0.6970	0.0190	0.9541	19
MS	0.7506	0.0431	0.9410	20

First hand-segmented (AH), second hand-segmented (VK), multiscale method (MS). AH was taken as the “ground truth” in all cases. Average values of: true positive rate (TPR), false positive rate (FPR) and accuracy (Ac). Number of images tested *n*.

segmented VK–AH. However, Staal et al. (2004) values were computed using 19 images and in addition using only the field of view (FOV) area of the fundus camera instead of the complete image to calculate them, therefore we are not able to make a direct comparison as we did with the rest of the authors.

If we compare values of true positive and false positive rates from the hand-segmented sets shown in Table 1 for VK–AH and AH–VK, we notice that while the true positive rate from the first comparison is larger than the second, false positive rates are the other way around, but accuracy of both are exactly the same. In this particular comparisons, we could say by inspection that both hand-segmented sets of images, AH and VK, have a reasonably good segmentation although they have a large variation between them. This variation is not obvious from the accuracy value. (see Fig. 13b, c, e and f). Furthermore, since the number of pixels that are *vessel* occupy a very small part of the image in comparison to the *non-vessel* pixels, the denominator in computing the false positive rate (FPR) will always be large – and thus will be biasing the analysis. Consequently, the value of false positive rate is not a very sensitive indicator for this kind of images.

4.1.2. DRIVE database

A second public database, called DRIVE, has been made available by Staal et al. (2004). This database consists on 40 images which were captured in digital form from a Canon CR5 non-mydratic 3CCD camera at 45° field of view. The images are of size 768 × 584 pixels, 8 bit per colour channel. The images have been divided into a *test* and *train* sets with 20 images each. Both sets were hand-segmented by three observers: one computer science student, one clinical expert and one image processing expert. Images in the *test* set were hand-segmented twice resulting in sets A and B. Images in the *train* set were hand-segmented once. Fig. 13d shows one image from the *test* set segmented by the MS method (using only the green band) and Fig. 13e and f show the two corresponding hand-segmented images, by A and B, respectively.

Validations using DRIVE database have been made by their authors only, as far as we are aware. They follow the same idea of contingency tables described before and they

Table 1
HOOVER database

Comparison	TPR	FPR	Ac	<i>n</i>
MS–AH	0.7506	0.0431	0.9410	20
VK–AH	0.8951	0.0438	0.9522	20
MS–VK	0.6064	0.0354	0.9232	20
AH–VK	0.6424	0.0082	0.9522	20

Top: comparisons of multiscale (MS) segmentation and the second hand-segmented (VK) set against the first hand-segmented (AH) set as the “ground truth”. Bottom: comparisons of multiscale (MS) segmentation and the first hand-segmented (AH) set against second hand-segmented (VK) set as the “ground truth”. Average values of: true positive rate (TPR), false positive rate (FPR) and accuracy (Ac). Number of images tested *n*.

Table 3
DRIVE database

Method	TPR	FPR	Ac	<i>n</i>
B–A	0.7760	0.0275	0.9473	20
Staal et al. (2004)	0.6780	0.0170	0.9441	20
MS	0.7246	0.0345	0.9344	20
MS–GT	0.7187	0.0422	0.9270	20
MS–GT	0.7172	0.0337	0.9343	18

Top: *Test* set. Comparisons between the second hand-segmented (B) set, Staal et. al. (2004) and multiscale (MS) methods against the first hand-segmented (A) set as the “ground truth”. Bottom: *Train* set. Comparisons of multiscale (MS) segmentation method against the hand-segmented (GT) set as the “ground truth”. Average values of: true positive rate (TPR), false positive rate (FPR) and accuracy (Ac). Number of images tested *n*.

present their results using only the *test* set with the A hand-segmented set as a reference. We tested the MS method on both hand-segmented sets. The top of Table 3 shows the evaluation made with the *test* set using the hand-segmented set A as the “ground truth”. We can see a larger value of true positive rate for MS than Staal et al. (2004) and also a larger value of false positive rate for MS than Staal et al. (2004). The latter makes the accuracy of Staal et al. (2004) larger.

The bottom of Table 3 shows our evaluation made with the *train* set using the only one hand-segmented set available as the “ground truth” (GT). Here we can see that if we take out only two of the worst images from the set

the true positive rate of MS keeps almost the same value than that of the complete set but the false positive rate decreases and therefore the accuracy increases. This means that the detection of vessels is not improved, we are only eliminating those errors due to the background artifacts.

After validating our results with both public databases using contingency tables and reminding that the higher the value of true positive rate means the more right blood vessels detected and the lower the value of false positive rate means the lower the error of segmentation. We conclude that with this evaluation it is not possible to know if the false positive rate value is due to an error in the detection of the borders of the vessels (under or over estimation of widths) or due to the detection of blobs due to some pathology present on the retina layer (background). Therefore we believe that this methodology to evaluate the accuracy of segmentation does not really measure the accuracy of detecting blood vessels but how many errors the method has on both sides of the classes against a “ground truth”, which could be itself difficult to build.

The main objective of the current work is to present a methodology for segmenting blood vessels in retinal images. Since one of the important goals of any segmentation method that detects blood vessels in these images is to measure vessel morphological properties that can be useful in clinical research to characterise changes due to diseases, we consider that this kind of evaluation technique does not give us the information that we really need. Because of this in the next sections we will present other approach to

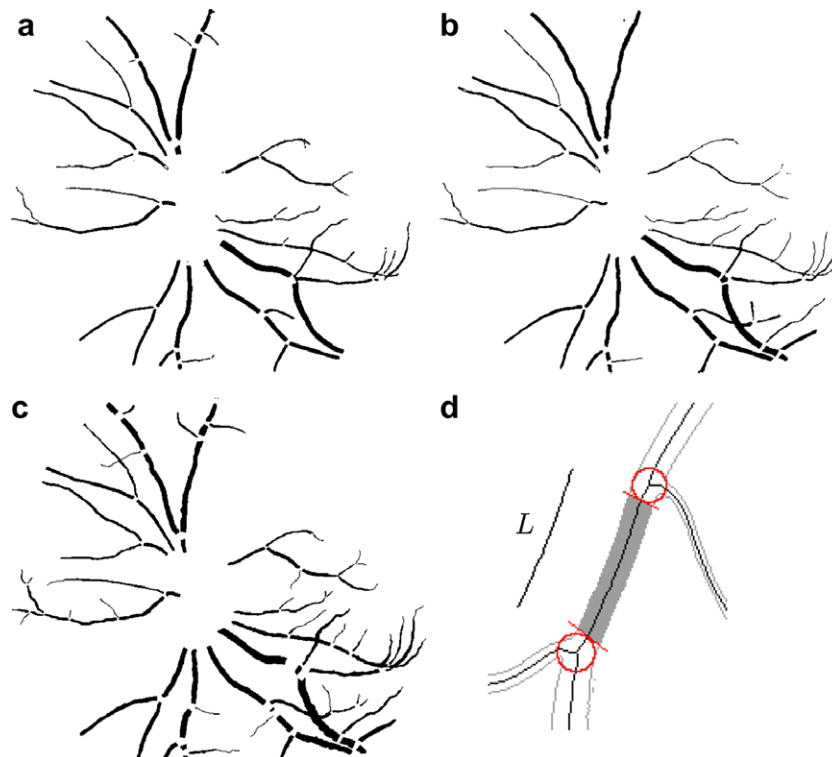


Fig. 14. Vessel diameters measured: (a) MS segmented image, (b) AH observer, (c) VK observer and (d) the average diameter is calculated from the shaded area as: $d = \text{area}/\text{length}$ in pixels, where *length* is the distance along the skeleton line.

evaluate the accuracy of the segmentation by performing some geometric measurements from the segmented images and comparing them with those taken from the “ground truth” images.

4.2. Comparison of measurements taken from a pair of segmented MS and HOOVER database images

We selected one of the images from HOOVER database (im0163) and measured the average diameters of a number of trees using a previously reported method for the measurement of geometric properties in binary retinal images (Martinez-Perez et al., 2002). Fig. 14a–c show the vessel segments that were measured for MS, AH and VK, respectively.

Table 4
Comparisons of vessel diameter measurements in pixels between multi-scale (MS) segmentation, first (AH) and second (VK) hand-segmented set

Comparison	$X_a - X_b$	$\Delta X / \langle X \rangle$	n	p
MS–AH	0.318 ± 0.684	0.171 ± 0.269	49	≤ 0.003
MS–VK	-0.949 ± 0.967	-0.224 ± 0.231	55	< 0.001
AH–VK	-1.134 ± 0.903	-0.328 ± 0.339	57	< 0.001

The average of differences $X_a - X_b$, where subscripts refer to first set (a) and reference set (b) are expressed in pixels. $\Delta X / \langle X \rangle$ is the average of the normalised differences and n , is the number of paired vessel segments measured. The p values are calculated using the two sided Wilcoxon signed-rank test. $p < 0.05$ was considered statistically significant.

Fig. 14d shows in gray the vessel area which is measured. For each vessel an average segment diameter is calculated as: $d = \text{area}/\text{length}$ in pixels, where length is the distance along the skeleton line.

In the absence of a true measure, we defined a normalised difference between measures X_a and X_b as:

$$\frac{\Delta X}{\langle X \rangle} = \frac{X_a - X_b}{\langle X \rangle} \quad \text{where } \langle X \rangle = \frac{1}{2}(X_a + X_b) \quad (9)$$

X_a corresponds to the first measurement and X_b to Bland and Altman (1986). Since samples of vessel diameters do not follow a normal distribution we use a two-sided Wilcoxon signed-rank test to look for significance at $p < 0.05$. Table 4 shows the result of comparisons of average vessel diameter measurements.

Table 4 shows that measurements of average diameters in MS were larger than those of AH and smaller than those of VK. Similarly hand-segmented diameters in AH were smaller than those of VK. All values of p showed that measurements are significantly different. These significant differences cannot be taken as error but only as indications of the variability between all diameter measurements.

Fig. 15 shows a very clear example of this variation. In the top row of Fig. 15, a branch extracted from bottom-centre-left (about 6:35) of Fig. 14a–c is shown, for MS, AH and VK, respectively. The bottom row of Fig. 15 shows the parent vessel of that branch zoomed, where borders are superimposed on the original image with bright dots, for MS, AH and VK, respectively.

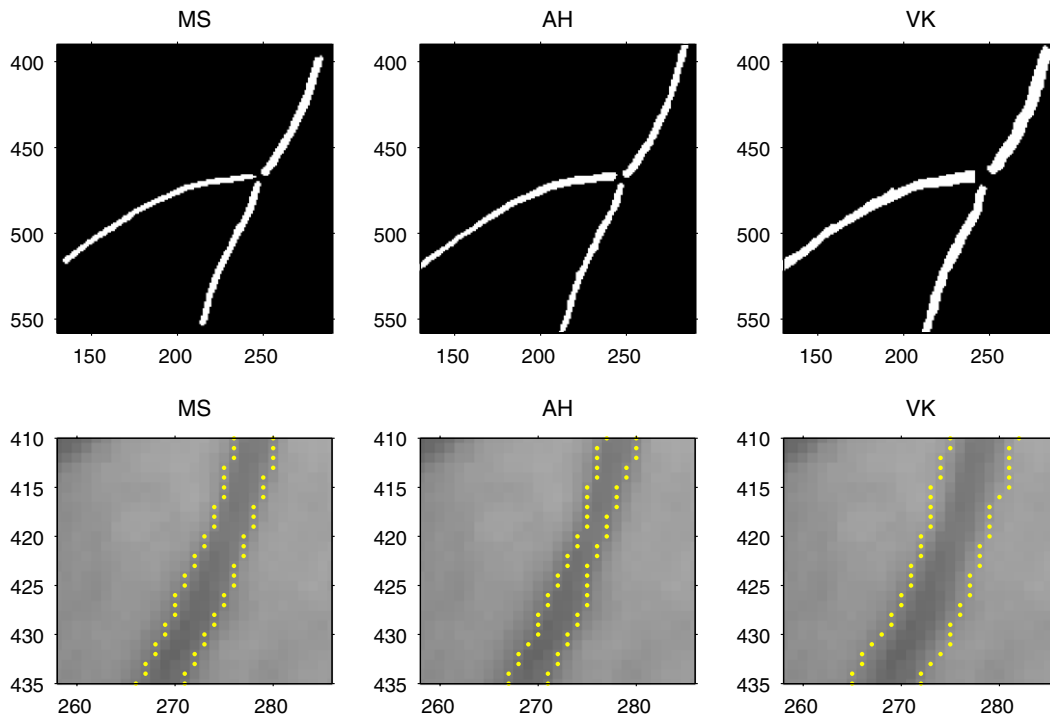


Fig. 15. Top row: branch extracted from bottom-centre-left (about 6:35) of Fig. 14 (a–c), for MS, AH and VK, respectively. Bottom row: the parent vessel zoomed and borders superimposed on the original image with yellow dots, for MS, AH and VK, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Average vessel diameter for each of the vessel segments shown in the bottom row of Fig. 15 are: 5.13, 4.08 and 7.34 pixels, respectively. The large variability on measurements from the same vessel segment can be seen, although borders found by the MS method appear to be less biased and to follow more consistently the border of the vessel.

4.3. Comparison of measurements taken from pairs of red-free and fluorescein images

Since the algorithm presented in this paper is able to segment both red-free and fluorescein images, a third validation was undertaken to assess the precision of the automatic measurements. Measurements from red-free images, that are most commonly used clinically, are compared against the measurements from better quality fluorescein images, which are normally considered to be more accurate. The latter are used by some authors as the “ground truth” when comparing diameter measurements (Gao et al., 1997).

Original images, from St. Mary’s, London database, were collected with the same setting as described in Section 3. Images were digitised with 2800×2400 pixels in size, a region of interest (bifurcation) was selected and segmented with an integer interval of scales $1 \leq s \leq 25$ pixels. Sixteen randomly taken individual bifurcations were automatically measured in red-free and fluorescein paired images. Fig. 16d–f show the same bifurcations of Fig. 16a–c after

the fluorescein dye injection. Borders of the segmentation MS method are superimposed with white dots on the original images.

The average diameter, d , is calculated as before (see Fig. 14d). Bifurcation angles, ω , were measured along the skeleton lines within a distance from the bifurcation centre to a fixed circle of $5R$ where R is the radius of the maximum circle centred on the bifurcation point that fits inside the boundary of the bifurcation, as those circles shown in Fig. 14d between the two bifurcation points (refer to Martinez-Perez et al. (2002) for more details).

For this comparison, X_R corresponds to red-free measurements and X_F corresponds to fluorescein. Table 5 summarises the results.

It can be seen that automatic measurements of both diameter and branching angle in fluorescein and red-free images are not significantly different, like those shown in the example of Fig. 15.

Table 5
Fluorescein against red-free comparisons

Feature	$X_R - X_F$	$\Delta X / \langle X \rangle$	n	p
d	0.66 ± 2.58	0.04 ± 0.16	48	0.15
ω	0.90 ± 6.90	0.01 ± 0.09	16	0.67

The differences $X_R - X_F$, where subscripts refer to red-free (R) and fluorescein (F), are expressed in pixels for diameter d and degrees for branching angle ω . $\Delta X / \langle X \rangle$ is the average of the normalised differences and n , is the number of paired vessel segments measured. The p values are calculated using the two sided Wilcoxon signed-rank test. $p < 0.05$ was consider statistically significant.

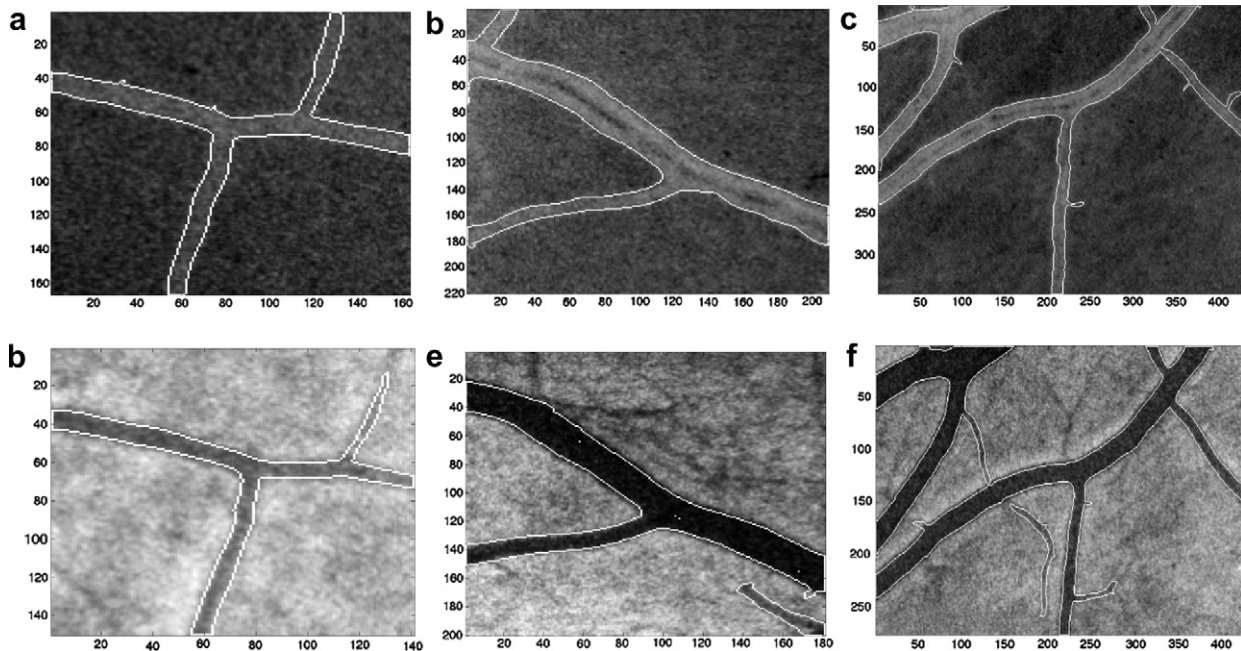


Fig. 16. Some examples of individual bifurcations in gray scale negatives of red-free and fluorescein images. (a–c) are the original subimages of red-free, (d–f) are the respective pairs of fluorescein bifurcation. Borders of the segmentation with the MS method are superimposed with white dots on the original images. Subimages are taken from gray scale images of size 2400×2800 pixels.

5. Conclusions

We have presented an algorithm to segment retinal blood vessels from both red-free and fluorescein images which combines: (1) the multiscale feature extraction that gives information about the width property of blood vessels that is independent of their orientation in the image, (2) two geometric properties of tube-like structures based on the first and second derivatives of intensity that give weights to pixels with a high probability of belonging to vessels, and (3) a multiple pass region growing procedure. The region growing algorithm is relatively fast because in the initial stage growth is restricted to regions with low gradients, allowing vessels to grow where the values of the maximum principal curvature lie within a wide interval. This allows rapid growth outside the boundary regions. In the final stage when the borders between classes are defined, the algorithm grows vessel and background classes simultaneously without the gradient restriction.

It should be pointed out that although the idea of the *diameter-dependent* equalisation factor applied to the multiscale information is empirical, it can be justified by a simple model in which the intensity of the image is proportional to the amount of blood in the light path corresponding to that pixel. However, from Fig. 10b and c it can be noticed that the intensity of small vessels is still lower than large vessels; the assumption of a linear dependency on diameter of this factor may be too simplistic and further research is needed to find the appropriate factor (Styles et al., 2005). Results shown in Fig. 12c and d and in some images from the public databases also suggest that further processing techniques are required to extract objects such as haemorrhages, drusen, exudates or microaneurysms in combination with the segmentation of blood vessels (Sinthanayothin et al., 1999). In any case, the algorithm presented seems to work acceptably in detecting retinal blood vessels over a large range of widths in the majority of the images tested so far.

From the first validation made with the public databases, comparisons of our results with other authors that used the HOOVER database showed that they are apparently the same as Hoover et al. (2000) and they have a lower performance than Jiang and Mojon (2003). In the case of the DRIVE database MS performs better on detecting correct vessels than Staal et al. (2004) but MS detects more blobs from background particularly in those images with strong pathologies. The accuracy value is anyhow very similar. However, we concluded that an analysis of true and false positive rate values does not give us the information needed about the accuracy of segmenting vessels when geometric measurements such as vessel widths are of interest. In order to demonstrate this, we performed a second validation with an image of HOOVER database where we found a large variability in diameter measurements and therefore we emphasise the difficulty of building a reliable “ground truth” image for blood vessel segmentation validation on complete retinal images.

For the third validation, where measurements of diameters and branching angles were made automatically between paired red-free and fluorescein images, we did not find any significant difference for either diameters or branching angles. Because of the better quality of fluorescein images, measurements taken from this technique are normally considered to be more accurate than those from red-free, and are taken as a “ground truth” measure by some authors (Gao et al., 1997, 2000). In addition, we did not observe any bias in the automatic measurements as that found in the example shown in Fig. 15. We conclude from our comparisons that the approach presented gives comparable estimates of diameter and branching angles in both red-free and fluorescein image. MS method performs satisfactorily in having a relatively high true positive rate value with confident vessel widths, however a method to eliminate the abnormalities detected as blobs have to be implemented as we have aforementioned.

The appearance of the retinal blood vessels can be an important diagnostic indicator of various clinical disorders of the eye and the body. Quantitative analysis of retinal blood vessels from fundus images has usually been studied in terms of individual bifurcations, measuring a few of the most clearly visible bifurcations in an image. Accurate detection of vessels widths and continuous blood vessel networks, based on a robust segmentation, will allow the study of morphological changes resulting from disease: geometric factors such as diameters, branching angles, lengths, tortuosity, length to diameter ratios, etc. (Gelman et al., 2005) and network properties such as connectivity, branch ordering, purely topological indices which describe the shape of branching trees (Martinez-Perez et al., 2002; Hughes et al., 2006), and if appropriate, fractal properties.

Acknowledgement

M. Elena Martinez-Perez would like to acknowledge to the Mexican National Council for Science and Technology (CONACYT) for financial support.

References

- Akita, K., Kuga, H., 1982. A computer method of understanding ocular fundus images. *Pattern Recogn.* 15, 431–443.
- Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1 (8476), 307–310.
- Brinchmann-Hansen, O., Engvold, O., 1986. Microphotometry of the blood column and the light streak on retinal vessels in fundus photographs. *Acta Ophthalmol. (Suppl.)* 179, 9–19.
- Burt, P., 1984. The pyramid as a structure for efficient computation. In: Rosenfeld, A. (Ed.), *Multiresolution Image Processing and Analysis*, Information Sciences. Springer-Verlag, Berlin, pp. 6–35 (Chapter 2).
- Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M., Goldbaum, M., 1989. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Trans. Med. Imag.* 8, 263–269.
- Eberly, D., 1996. *Ridges in Image and Data Analysis*. Computational Imaging and Vision. Kluwer Academic Publishers, Netherlands.
- Frangi, A., Niessen, W., Vincken, K., Viergever, M., 1998. Multiscale vessel enhancement filtering. In: Wells, W., Colchester,

- A., Delp, S. (Eds.), 1994, Medical Image Computing and Computer-Assisted Intervention – MICCAI'98, Lecture Notes in Computer Science, vol. 1496. Springer, Berlin, pp. 130–137.
- Gao, X., Bharath, A., Hughes, A., Stanton, A., Chapman, N., Thom, S., 1997. Towards retinal vessel parametrisation. In: SPIE Conference on Medical Imaging, pp. 734–744.
- Gao, X., Bharath, A., Stanton, A., Hughes, A., Chapman, N., Thom, S., 2000. Quantification and characterisation of arteries in retinal images. *Comp. Meth. Program. Biom.* 63, 133–146.
- Gelman, R., Martinez-Perez, M.E., Vanderveen, D.K., Moskowitz, A., Fulton, A.B., 2005. Diagnosis of plus disease in retinopathy of prematurity using retinal image multiscale analysis. *Invest. Ophthalm. Vis. Sci.* 46 (12), 4734–4738.
- Hammer, M., Leistriz, S., Leistriz, L., Schweitzer, D., 2001. Light paths in retinal vessel oximetry. *IEEE Trans. Biomed. Eng.* 48, 592–598.
- Hoover, A., Kouznetsova, V., Goldbaum, M., 2000. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Trans. Med. Imag.* 19, 203–210.
- Hughes, A.D., Martinez-Perez, M.E., Jabbar, A.S., Hassan, A., Witt, N.W., Mistry, P.D., Chapman, N., Stanton, A.V., Beevers, G., Pedrinelli, R., Parker, K.H., Thom, S.A., 2006. Quantification of topological changes in retinal vascular architecture in essential and malignant hypertension. *J. Hypertens.* 24 (5), 889–894.
- Ikram, M., Witteman, J., Vingerling, J., Breteler, M., Hofman, A., de Jong, P., 2006. Retinal vessel diameters and risk of hypertension: the rotterdam study. *Hypertension* 47, 189–194.
- Jiang, X., Mojon, D., 2003. Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images. *IEEE Trans. Pattern Recogn. Anal. Mach. Intell.* 25, 131–137.
- King, L., Stanton, A., Sever, P., Thom, S., Hughes, A., 1996. Arteriolar length–diameter ($l:d$) ratio: a geometric parameter of the retinal vasculature diagnostic of hypertension. *J. Hum. Hypertens.* 10, 417–418.
- Klein, R., Klein, B., Moss, S., Wong, T., Hubbard, L., Cruickshanks, K., Palta, M., 2004. The relation of retinal vessel caliber to the incidence and progression of diabetic retinopathy – xix: the Wisconsin epidemiologic study of diabetic retinopathy. *Arch. Ophthalmol.* 122, 76–83.
- Koenderink, J., 1984. The structure of images. *Biol. Cybern.* 50, 363–370.
- Lindeberg, T., 1993. On scale selection for differential operators. In: Heia, K., Hogdra, K., Braathen, B. (Eds.), *Proceedings of the 8th Scandinavian Conference on Image Analysis*, Tromsø, Norway, pp. 857–866.
- Lorenz, C., Carlsen, I., Buzug, T., Fassnacht, C., Weese, J., 1997. Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical images. In: Troccaz, J., Grimson, E., Mösges, R. (Eds.), *Proceedings of the CVTMed-MRCAS'97, Lecture Notes in Computer Science*, pp. 233–242.
- Martinez-Perez, M.E., 2001. Computer analysis of the geometry of the retinal vasculature. Ph.D. Thesis, Imperial College of Science, Technology and Medicine, University of London.
- Martinez-Perez, M.E., Hughes, A.D., Stanton, A.V., Thom, S.A., Bharath, A.A., Parker, K.H., 1999. Retinal blood vessel segmentation by means of scale-space analysis and region growing. In: Taylor, C., Colchester, A. (Eds.), *MICCAI-99, Lecture Notes in Computer Science*, vol. 1679. Springer-Verlag, pp. 90–97.
- Martinez-Perez, M.E., Hughes, A.D., Stanton, A.V., Thom, S.A., Chapman, N., Bharath, A.A., Parker, K.H., 2002. Retinal vascular tree morphology: a semi-automatic quantification. *IEEE Trans. Biomed. Eng.* 49, 912–917.
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* SMC-9, 62–66.
- Sato, Y., Nakajima, S., Atsumi, H., Koller, T., Gerig, G., Yoshida, S., Kikinis, R., 1997. 3D multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. In: Troccaz, J., Grimson, E., Mösges, R. (Eds.), *Proceedings of the CVTMed-MRCAS'97, Lecture Notes in Computer Science*, pp. 213–222.
- Sinthanayothin, C., Boyce, J., Cook, H., Williamson, T., 1999. Automated localisation of the optic disc, fovea and retinal blood vessels from digital colour fundus images. *Brit. J. Ophthalmol.* 83, 902–910.
- Skovborg, F., Nielsen, A., Lauritzen, E., Hartkopp, O., 1969. Diameters of the retinal vessels in diabetic and normal subjects. *Diabetes* 18, 292–298.
- Smith, M., Denninghoff, K.A.L., Hillman, L., 2000. Effect of multiple light paths on retinal vessel oximetry. *Appl. Opt.* 39, 1183–1193.
- Staal, J., Abramoff, M., Niemeijer, M., Viergever, M., van Ginneken, B., 2004. Ridge-based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imag.* 23, 501–509.
- Stanton, A., Wasan, B., Cerutti, A., Ford, S., Marsh, R., Sever, P., Thom, S., Hughes, A., 1995. Vascular network changes in the retina with age and hypertension. *J. Hypertens.* 13, 1724–1728.
- Styles, I., Claridge, E., Orihuela-Espina, F., Calcagni, A., Gibson, J., 2005. Quantitative interpretation of multi-spectral fundus images. In: Amini, A., Manduca, A. (Eds.), *Proceedings of SPIE, Medical Imaging 2005: Physiology, Function and Structure from Medical Images*, vol. 5746, pp. 267–278.
- Tolias, Y., Panas, S., 1998. A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering. *IEEE Trans. Med. Imag.* 17, 263–273.
- Witkin, A., 1984. Scale-space filtering: a new approach to multi-scale description. In: Ullman, S., Richards, W. (Eds.), *Image Understanding*. Ablex, Norwood, NJ, pp. 79–95.
- Wong, T., McIntosh, R., 2005. Hypertensive retinopathy signs as risk indicators of cardiovascular morbidity and mortality. *Brit. Med. Bull.* 73–74, 57–70.
- Wong, T., Shankar, A., Klein, R., Klein, B., 2004. Retinal vessel diameters and the incidence of gross proteinuria and renal insufficiency in people with type 1 diabetes. *Diabetes* 53, 179–184.
- Wu, D.-C., Schwartz, B., Schwoerer, J., Banwatt, R., 1995. Retinal blood vessel width measured on color fundus photographs by image analysis. *Acta Ophthalmol. Scand. Suppl.* 215, 33–40.
- Zana, F., Klein, J., 1997. Robust segmentation of vessels from retinal angiography. In: *Proceedings of the International Conference on Digital Signal Processing*, Santorini, Greece, pp. 1087–1090.
- Zhou, L., Rzeszotarski, M., Singerman, L., Chokreff, J., 1994. The detection and quantification of retinopathy using digital angiograms. *IEEE Trans. Med. Imag.* 13, 619–626.