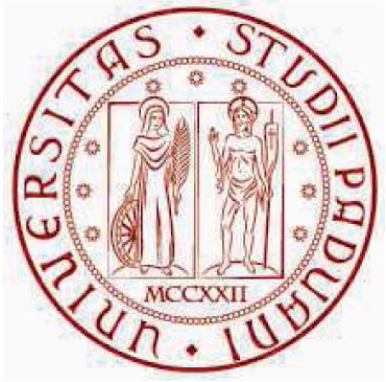
UNIVERSITÀ DEGLI STUDI DI PADOVA



FACOLTÀ DI INGEGNERIA

Corso di laurea Triennale in INGEGNERIA DELL' INFORMAZIONE

SPLINE-BASED REFINEMENT OF CONTOURS IN BINARY MAPS OF RETINAL VESSELS

RELATORE

DEPARTMENT OF INFORMATION ENGINEERING
UNIVERSITY OF PADOVA

Prof. Enrico Grisan
Department
of Information Engineering

CORRELATORE

CVIP DUNDEE

Prof. Manuel Trucco
NRP Professor
of Computational Vision
University of Dundee

LAUREANDO: Alessandro Cavinato

Anno Accademico 2011 - 2012

To my parents

Abstract

Vasculature segmentation and vessel caliber measurements in retinal images can improve early diagnosis of several diseases, such as diabetes, retinopathy of prematurity and hypertension. The aim of this thesis is to present a novel algorithm for improving the vessel contours obtained from binary vessel maps. This is useful for quantitative evaluations like width and tortuosity estimation. Two algorithms are described in this document. Firstly, a simple vessel segmentation strategy filtering the image using a Gaussian kernel and producing a binary vessel mask from the response image by the application of a thresholding step. Secondly, a procedure fitting the two contours of each vessel in the binary map with a cubic spline curve, under a parallelism constraint between the two splines. The second algorithm is the main focus of this work. The performance of the algorithm has been evaluated on the publicly available REVIEW database, which contains a set of images with vasculature showing different characteristics. Images also include several manual measurements made by three independent observers. The method is implemented in MATLAB.

Contents

1 Introduction **5**

2 Methods **7**
2.1 Vessel segmentation algorithm 7
2.2 Mathematical background on spline interpolation 9
2.3 Vessel centreline computation and refinement 11
2.4 Vessel edge extraction 13
2.5 Contours refinement using a coupled spline-based method 15
2.6 Width estimation 19

3 Experimental Results **20**
3.1 REVIEW database 20
3.2 Algorithm performance measures 20
3.3 Processing times 27

4 Discussion and conclusion **28**

1 Introduction

The analysis of retinal images plays a significant role in the clinical diagnosis of several pathologies; for instance, arteriosclerosis, hypertension and diabetes. In fact, Diabetic Retinopathy (DR) is a disease that may cause visual impairments in patients suffering from diabetes mellitus which, after several years, could even lead to blindness. Signs include the appearance of lesions and ischemic regions, areas with poor blood supply, caused by occlusions in retinal vessels. It is then of primary importance to monitor changes in the retinal vascular structure, in order to prevent later complications in patients suffering from DR.

In general, retinal images are clinically relevant, as they permit to have direct access to human vasculature in vivo in a non-invasive manner. Obtaining a binary vessel map from retinal images proves to be useful for many different purposes: among others, the evaluation of vessel tortuosity, often regarded as a symptom of systemic hypertension, and the measurement of vessel calibre as a biomarker for cardiovascular diseases. Hence, the developing of a computerized system for retinal vessel segmentation would be of great benefit to improve the efficacy of ophthalmologists' work, helping them in the diagnosis of some degenerative pathologies. However, devising a reliable automatic retinal vessel detection procedure turns out to be very challenging due, among others, to the wide range of image resolutions and acquisition modes. For this reason, semi-automated methods are often used, despite they may be onerous in terms of time and subject to user errors. In recent years, many vessel segmentation algorithms were presented. These methods can be differentiated in two classes: procedures that use supervised learning and procedures that do not.

The former group includes supervised algorithms that train a classifier, starting from a manually-segmented training set of images in which pixels are labeled as vessel or non-vessel by a human. The classifier is then used to determine whether a pixel belongs to a vessel or not, according to a feature vector which is computed for every pixel in the image. A relevant example is the algorithm by Soares et al.[1], which implements the Gabor wavelet transform to perform a multiscale analysis on the image and then computes pixel feature vectors as the maximum response of the Gabor filter over different orientations and scales; finally, the classification is performed using a Bayesian classifier. In Lupascu et al.[2] the feature vector is composed of 41 elements, according to many local and spatial image properties, like the pixel intensity profile along a segment orthogonal to a vessel, or the output of several filters; afterwards, a trained AdaBoost classifier is used for the classification of each pixel.

The latter group include methods that do not use supervised learning for vessel segmentation. Research results in this field include: exploratory vessel tracing, as shown in Chutatape et al.[3], in which a first set of pixels (seeds), found either automatically or manually, is used as a starting point for the vessel tracking and, as soon as new vessel pixels are found, the set of seed points is updated; deformable models, also called snakes, which are active contour models iteratively adapting their shape to locate vessel edges; thresholding of wavelet coefficients, as the Isotropic Undecimated Wavelet Transform proposed by Bankhead et al.[4]; other Laplacian or 2D Gaussian kernel approaches, as suggested by Vermeer et al.[5].

Comparing the supervised learning methods with the others, the former group leads to better segmentation results on the whole. This is due to the finer

decision techniques adopted by the classifiers: these procedures are suitable for detecting the complex retinal vasculature structure, which sometimes could not be reliably detected by unsupervised procedures. On the other hand, training methods are very expensive in terms of time: in fact, the classifier needs several hand-labeled images to be trained, and each of these takes a long time to be manually segmented. Moreover, human labeling is open to user bias and the same image could be segmented in multiple ways by different observers. On the contrary, unsupervised methods are faster and do not need a training set of images. Nevertheless, these procedures have some drawbacks too: their main limit is that these algorithms are often tailored just for a particular image type or resolution and require adaptations to be applied to other images.

All the methods above aim to obtain a binary image in which the distinction is between pixels belonging or not to the vascular structure. However, the problem of how to reliably estimate the vessel caliber is not addressed by such algorithms: in fact, width estimates taken from raw binary vessel maps, even if taken in locations close to each other, present a relatively high standard deviation, due to the jagged vessel borders occurring in binary maps.

This thesis describes a novel algorithm smoothing raw vessel contours in binary retinal vessels masks. Removing vessel boundaries indentations turns out to be useful to reliably evaluate vascular diameters. For each processed vessel, after identifying its spline-smoothed centreline, the algorithm finds two cubic spline curves fitting the jagged contours. The coefficients of these splines are computed by solving an overconstrained system including both standard spline formulae and a parallel-tangent constraint, which ensures that the vessel profile is as similar as possible to a 2-D curvilinear pipe with parallel borders.

The remainder of this thesis is organized as follows: in Section 2.1 a common vessel segmentation algorithm is described. Section 2.2 provides a mathematical background on spline interpolation. Sections 2.3 to 2.6 deal with the steps of the spline-based procedure described in this document. Algorithm performance evaluated on REVIEW database is discussed in Section 3. Finally, discussion and conclusions are in Section 4.

2 Methods

2.1 Vessel segmentation algorithm

Although this thesis focuses on the vessel width measurement issue, this chapter describes a simple algorithm for the vascular segmentation in retinal images. This generates the binary vessel maps which are regularized by the second algorithm. The vessel segmentation algorithm proposed consists in convolving the image with a Gaussian sliding window in order to enhance the contrast between the vessel and non-vessel pixels, and then apply a thresholding step to the response image.

Such filtering algorithms are widely used in image processing for edge detection and feature extraction. For retinal vessel segmentation, this methodology turns out to be effective and easy to implement because it does not require any previous training step. An approach similar to the one described below was first proposed by Chaudhuri et al. [6].

In image processing, convolving an image with a mask requires moving the kernel through all the pixels of the image. Let I be an image of size $M \times N$ and g the convolution mask of size $p \times p$, where p is an odd integer; then for each pixel (i, j) in I the filter response at that pixel is

$$G(i,j) = \sum_{k=-a}^{a} \sum_{l=-a}^{a} g(k,l) I(i+k, j+l) \tag{2.1.1}$$

where $a = \dfrac{p-1}{2}$ and G is the resulting image.

The goal is to devise a convolution mask that is suitable for vessel detection in retinal images. The key point is to notice that the intensity profiles of the cross-section perpendicular to a vessel at any point can often be modelled as a Gaussian curve. Hence, a set of Gaussian-shaped filters turns out to be appropriate for vessel segmentation (but see [1,2,5] for a discussion of the limitations of this approach). The results presented below were obtained using the filter shown in Figure 1 and described by the function

$$g(x,y) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\dfrac{x^2}{2\sigma^2}}, \qquad -a \le x \le a, \quad -a \le y \le a. \tag{2.1.2}$$

where p is the dimension of the kernel and σ represents the standard deviation of the Gaussian profile.

Since vessels in retinal images have different amplitudes and orientations, the kernel has to be rotated by a finite set of angles $\theta \in [0, 180°]$ and its spread, σ, has to be varied as well, in order to capture as many vessel pixels as possible. The maximum response will be obtained when the filter orientation and amplitude approximate the propagation direction and the diameter of the vessel in the considered pixel. Therefore, for all the pixels (i, j) of the input image I, the maximum absolute value of the filter response over all angles θ and scales σ is kept. In other words, the pixel (i, j) of the output image, after the filtering process, will be

$$G_{max}(i,j) = \arg\max_{\theta,\sigma} \left| \sum_{k=-a}^{a} \sum_{l=-a}^{a} g_{\theta,\sigma}(k,l) I(i+k, j+l) \right| \tag{2.1.3}$$

where $g_{\theta,\sigma}$ is the kernel g rotated by θ and with a spread σ.

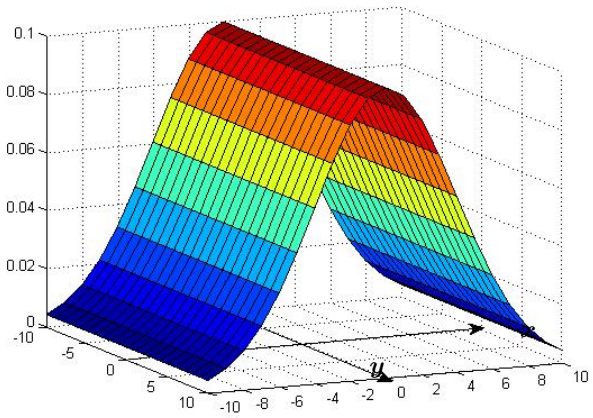


Figure 1: Gaussian kernel with $p = 21$ and $\sigma = 4$.

This procedure will enhance the contrast in the resulting image G_{max} between those pixels belonging to a vessel and those not. Hence, a thresholding step has to be applied to G_{max} in order to obtain a binary vessel mask of the original image I. This means that if the brightness level of a given pixel in G_{max} exceeds a certain threshold t, that point is classified as belonging to a vessel. The threshold can be set either automatically or manually; in this work, the values t used to segment images from REVIEW database are chosen in the latter way.

The main attraction of this unsupervised method is that it is easy to implement and it does not need a classifier to produce binary maps of the vascular structure. Nevertheless, this filter-based algorithm may have strong responses even to non-vessel structures, e.g. the boundaries of the optic disk or some lesions. Moreover, in several retinal images, some wider blood vessels have a bright streak running down the centreline: this *central light reflex* may be enhanced by the filter as well as other edges in the image. Thus, in these cases, a simple thresholding step may produce some false results in the binary image.

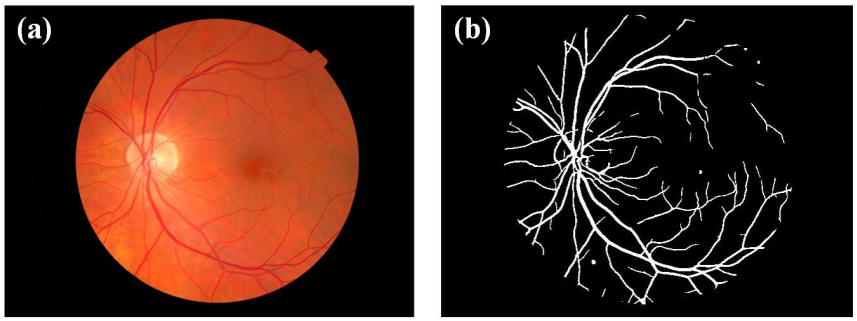


Figure 2: Result of a segmentation procedure. (a) Original image of size 2240×1488. (b) Segmented binary image obtained using a Gaussian filter with σ set to values from 2 to 4. The threshold value used is $t = 48$.

2.2 Mathematical background on spline interpolation

This section provides a theoretical overview on spline data interpolation, a technique which is widely used in several subjects, as approximation theory and numerical analysis. A *spline* is a piecewise polynomial function that is at least of class $\mathcal{C}^2$ at the points where the polynomial pieces join up. These points are called *knots*. Cubic splines are one of the most commonly used splines in interpolating problems; they are called *cubic* because their pieces are third-degree polynomials. This chapter describes how to derive the equations for cubic spline curves that interpolate a finite set of points.

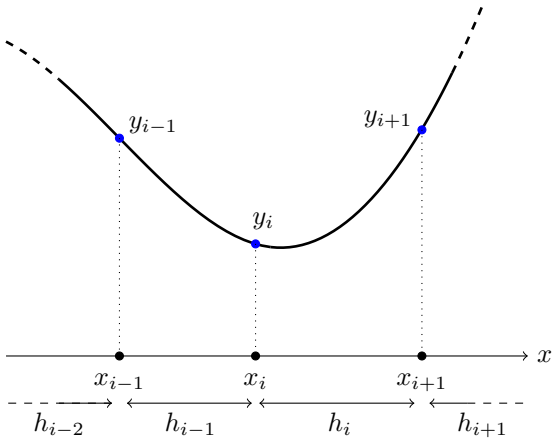


Figure 3: Interpolation with cubic spline between three points.

Given the data points (x_j, y_j) with $j = 0, 1, \dots n$, consider two consecutive knots (x_i, y_i) and (x_{i+1}, y_{i+1}). The general cubic function defined in the interval between these two points is

$$y = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i. \tag{2.2.1}$$

where a_i, b_i, c_i and d_i are the coefficients of the third-degree polynomial in the ith interval. Since the cubic must go through at (x_i, y_i) and (x_{i+1}, y_{i+1}), then

$$y_i = a_i(x_i - x_i)^3 + b_i(x_i - x_i)^2 + c_i(x_i - x_i) + d_i = d_i; \tag{2.2.2}$$

$$\begin{aligned} y_{i+1} &= a_i(x_{i+1} - x_i)^3 + b_i(x_{i+1} - x_i)^2 + c_i(x_{i+1} - x_i) + d_i \\ &= a_i h_i^3 + b_i h_i^2 + c_i h_i + d_i. \end{aligned} \tag{2.2.3}$$

where h_i is the width of the ith interval, i.e. $h_i = x_{i+1} - x_i$.

To ensure a smooth fit across the boundary between two consecutive intervals, the continuity of first and second derivatives is enforced at knots. The first two derivatives of (2.2.1) are:

$$y' = 3a_i(x - x_i)^2 + 2b_i(x - x_i) + c_i, \tag{2.2.4}$$

$$y'' = 6a_i(x - x_i) + 2b_i. \tag{2.2.5}$$

Naming S_i and S_{i+1} the second derivative of the function evaluated at (x_i, y_i) and (x_{i+1}, y_{i+1}) respectively, the following equations are obtained:

$$\begin{aligned} S_i &= 6a_i(x_i - x_i) + 2b_i = 2b_i; \\ S_{i+1} &= 6a_i(x_{i+1} - x_i) + 2b_i = 6a_i h_i + 2b_i. \end{aligned}$$

Thus, the coefficients a_i and b_i can be derived:

$$a_i = \frac{S_{i+1} - S_i}{6h_i} \tag{2.2.6}$$

$$b_i = \frac{S_i}{2}. \tag{2.2.7}$$

Substituting a_i, b_i and d_i, given by (2.2.6), (2.2.7) and (2.2.2), into Equation (2.2.3) and solving for c_i, the last coefficient can be obtained:

$$c_i = \frac{y_{i+1} - y_i}{h_i} - \frac{2h_i S_i + h_i S_{i+1}}{6}. \tag{2.2.8}$$

Other equations are found enforcing the spline first derivative continuity constraint at (x_i, y_i). This means that the curve slope is continuous at the point where the ith and $(i-1)$th intervals join. Equation (2.2.4), computed at (x_i, y_i), gives a new formula for c_i:

$$y_i' = 3a_i(x_i - x_i)^2 + 2b_i(x_i - x_i) + c_i = c_i.$$

In the $(i-1)$th interval, from x_{i-1} to x_i, the first derivative of the function is

$$y' = 3a_{i-1}(x - x_{i-1})^2 + 2b_{i-1}(x - x_{i-1}) + c_{i-1},$$

then the slope at the right end of the interval will be

$$\begin{aligned} y_i' &= 3a_{i-1}(x_i - x_{i-1})^2 + 2b_{i-1}(x_i - x_{i-1}) + c_{i-1} \\ &= 3a_{i-1}h_{i-1}^2 + 2b_{i-1}h_{i-1} + c_{i-1}. \end{aligned} \tag{2.2.9}$$

Since $y_i' = c_i$, equations (2.2.8) and (2.2.9) can be set equal to each other:

$$\frac{y_{i+1} - y_i}{h_i} - \frac{2h_i S_i + h_i S_{i+1}}{6} = 3a_{i-1}h_{i-1}^2 + 2b_{i-1}h_{i-1} + c_{i-1}. \tag{2.2.10}$$

The cubic coefficients in the $(i-1)$th interval are

$$a_{i-1} = \frac{S_i - S_{i-1}}{6h_{i-1}},\ b_{i-1} = \frac{S_{i-1}}{2},\ c_{i-1} = \frac{y_i - y_{i-1}}{h_{i-1}} - \frac{2h_{i-1}S_{i-1} + h_{i-1}S_i}{6} \text{ and } d_{i-1} = y_{i-1},$$

and substituting them in (2.2.10), the following equation is obtained after some simplifications:

$$\begin{aligned} h_{i-1}S_{i-1} + (2h_{i-1} + 2h_i)S_i + h_i S_{i+1} &= 6\left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}}\right) \\ &= 6(f[x_i, x_{i+1}] - f[x_{i-1}, x_i]), \end{aligned} \tag{2.2.11}$$

where $f[x_j, x_{j+1}] = \dfrac{y_{j+1} - y_j}{h_j}$.

These are $n-1$ equations in $n+1$ unknowns S_j, with $j = 0, \ldots n$. The most common ways of obtaining two additional relations are either

- set $S_0 = 0$ and $S_n = 0$, getting the so-called *natural cubic spline*, or
- extrapolate the value of S_0 from S_1 and S_2 and the value of S_n from S_{n-1} and S_{n-2}.

The first approach is used hereafter.

Equation (2.2.11) can be written in matrix form for $S_1,\ S_2,\ \dots S_{n-1}$ as $A\boldsymbol{x} = \boldsymbol{b}$, where

$$A = \begin{bmatrix} 2(h_0+h_1) & h_1 & 0 & 0 & 0 & \dots & 0 & 0 \\ h_1 & 2(h_1+h_2) & h_2 & 0 & 0 & \dots & 0 & 0 \\ 0 & h_2 & 2(h_2+h_3) & h_3 & 0 & \dots & 0 & 0 \\ \vdots & & & \ddots & & & \vdots & \vdots \\ 0 & 0 & \dots & & & & h_{n-2} & 2(h_{n-2}+h_{n-1}) \end{bmatrix},$$

$$\boldsymbol{x} = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_{n-1} \end{bmatrix} \text{ and } \boldsymbol{b} = 6 \begin{bmatrix} f[x_1,x_2] - f[x_0,x_1] \\ f[x_2,x_3] - f[x_1,x_2] \\ f[x_3,x_4] - f[x_2,x_3] \\ \vdots \\ f[x_{n-1},x_n] - f[x_{n-2},x_{n-1}] \end{bmatrix}$$

Since A is a tridiagonal matrix, this linear system can be easily solved for $\boldsymbol{x}$. Once S_j values are computed, coefficients a_j, b_j, c_j and d_j can be derived for each interval using equations (2.2.6), (2.2.7), (2.2.8) and (2.2.2).

2.3 Vessel centreline computation and refinement

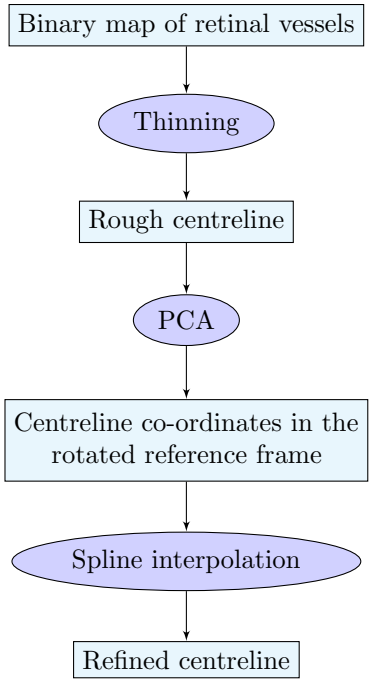


Figure 4: Flowchart representing the main steps of the algorithm for vessel centreline extraction and refinement.

This section presents an algorithm extracting and refining the centreline of each vessel. A flowchart, representing the main steps of this procedure, is shown above for the sake of clarity.

Thinning

A first, rough set of centrelines can be easily obtained from a vessel binary map using a morphological thinning algorithm. This method iteratively erode exterior pixels from the detected vessel structure, until no more erasable pixels exist. The resulting image is a binary mask of the *vessel skeleton*, i.e. a connected chain of pixels 1 pixel thin. It is useful for further processing to separate this binary structure into individual vessel segments by removing branching points from the vessel skeleton. These points are pixels belonging to the thinned centreline that have more than two neighbours. After that, segments that are less than 15 pixels long are removed, as they are considered insignificant for later analysis.

Before proceeding with centreline refinement, a first estimate of vessel width can be computed using the distance transform of the binary vessel mask. The result of this operation is a graylevel image in which pixel intensity values are the Euclidean distances from the considered pixel and the closest background pixel. By doubling the distance values along the thinned centreline, a coarse estimate of vessel diameter in these points is obtained.

PCA

To fit a natural cubic spline to the thinned vessel centreline, it is convenient to start by transforming the reference frame into the principal directions of the vessel points. This guarantees that the centreline is well represented as a function mapping each x value to a single y value. Considering pixel co-ordinates of a given vessel segment as a collection of data $P = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_m\}$, where $\boldsymbol{p}_j = (x_j, y_j)$ are the co-ordinates of the jth segment pixel, the *Principal Component Analysis* (PCA) is used to detect the direction along which data vary the most. These directions are given by the eigenvectors of the data scatter matrix H which is computed as follows:

$$H = \begin{bmatrix} (x_1 - \overline{x}) & (x_2 - \overline{x}) & \ldots & (x_m - \overline{x}) \\ (y_1 - \overline{y}) & (y_2 - \overline{y}) & \ldots & (y_m - \overline{y}) \end{bmatrix} \begin{bmatrix} (x_1 - \overline{x}) & (x_2 - \overline{x}) & \ldots & (x_m - \overline{x}) \\ (y_1 - \overline{y}) & (y_2 - \overline{y}) & \ldots & (y_m - \overline{y}) \end{bmatrix}^T,$$

where $\overline{\boldsymbol{p}} = (\overline{x}, \overline{y})$ is the centre of gravity of the data: $\overline{\boldsymbol{p}} = \dfrac{1}{m} \displaystyle\sum_{i=1}^{m} \boldsymbol{p}_i$.

The scatter matrix H is a 2-by-2 positive-definite and symmetric matrix, hence its eigenvalues are always real valued and positive. The eigenvector $\boldsymbol{e}_1$ with the largest eigenvalue is the direction of greatest data variation; the other eigenvector $\boldsymbol{e}_2$ is orthogonal to $\boldsymbol{e}_1$ given the above properties of H. Thus, the x and y axes are centred in $\overline{\boldsymbol{p}}$ and rotated into $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ directions. The co-ordinates of $\boldsymbol{p}_j \in P$ in the new reference frame are

$$\boldsymbol{p}'_j = (\boldsymbol{p}_j - \overline{\boldsymbol{p}}) \cdot \Lambda,$$

where Λ is the matrix that has $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ as columns.

Spline interpolation

A natural cubic spline is then fitted to $\boldsymbol{p}'_j$ points obtaining a smooth centreline of the selected vessel. Different choices of spline knots among $\boldsymbol{p}'_j$ points lead to slightly different centrelines. Results presented later in this work are obtained using a natural cubic spline with uniform intervals: a reliable interpolation in all tested images has been obtained setting the length of each interval to 10 pixels. Some steps of the centreline refinement process are shown in Figure 5.

Figure 5: Vessel centreline refinement process. (a) Vessels from the binary mask. (b) Thinned centrelines. (c) Refined centrelines plotted on the binary vessel mask.

2.4 Vessel edge extraction

The next step is to extract vessel border points from the binary retinal maps starting from the spline-smoothed centreline and the preliminary vessel widths already computed. The goal is to identify vessel edges using the information given by pixel intensity profile along vessel cross-sections. The refined centreline is smooth enough to compute reliably, most of the times, orthogonal segments that do not intersect each other. Hence, for each centreline pixel C_j, the perpendicular d_j is computed. To ensure that segment d_j will be long enough to pass even through the widest vessels, its length is set to w_j, where w_j is the preliminary vessel width at C_j, estimated from the distance transform described above.

The image pixel intensity profile along d_j is computed using linear interpolation. Since the image is binary, vessel edge points are those pixels where intensity profile changes value. These points are detected finding the two peaks into the first derivative of the binary profile.

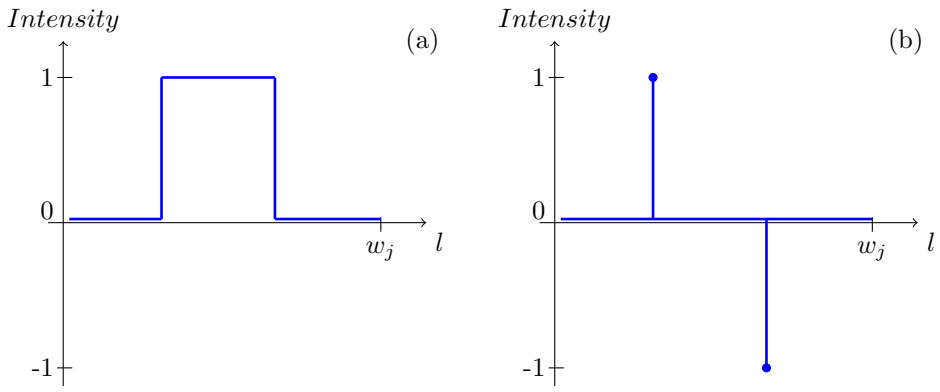


Figure 6: Pixel intensity profile along d_j (a) and its derivative (b). The horizontal axis represent the curvilinear co-ordinate l along the centreline-orthogonal segment.

Thereby, scanning each segment $d_1, d_2, \dots d_m$, where m is the pixel length of the considered centreline, $2m$ coupled contour points, $\boldsymbol{q}_{A,j}$ and $\boldsymbol{q}_{B,j}$, are found; these can be arranged in two lists, $Q_A = \{\boldsymbol{q}_{A,1}, \boldsymbol{q}_{A,2}, \dots, \boldsymbol{q}_{A,m}\}$ and $Q_B = \{\boldsymbol{q}_{B,1}, \boldsymbol{q}_{B,2}, \dots, \boldsymbol{q}_{B,m}\}$, so that in the jth position of the lists are the two vessel edge points detected along d_j. Before proceeding to vessel border points interpolation, Q_A and Q_B must contain points belonging to the same vessel contour. To ensure that, a simple procedure checking whether all the points either in Q_A or in Q_B lie on the same side of the centreline is applied:

Algorithm

input:

Two lists Q_A and Q_B of detected vessel edge points.

steps:

For each pair of consecutive points in Q_A, $\boldsymbol{q}_{A,j} = (x_{A,j}, y_{A,j})$ and $\boldsymbol{q}_{A,j+1} = (x_{A,j+1}, y_{A,j+1})$, compute $\lambda = \dfrac{ax_{A,j} + by_{A,j} + c}{ax_{A,j+1} + bx_{A,j+1} + c}$, where a, b and c are the coefficients of the line $r : ax + by + c = 0$ tangent to the centerline at C_j.

if $\lambda > 0$ *then* $\boldsymbol{q}_{A,j}$ and $\boldsymbol{q}_{A,j+1}$ lie in the same side of the line r, hence they belong to the same vessel contour;

else they are located on the opposite sides of r. Thus, it is necessary to swap $\boldsymbol{q}_{A,j}$ and $\boldsymbol{q}_{B,j}$:

$Temp \leftarrow \boldsymbol{q}_{A,j}$
$\boldsymbol{q}_{A,j} \leftarrow \boldsymbol{q}_{B,j}$
$\boldsymbol{q}_{B,j} \leftarrow Temp$

where $Temp$ is a temporary variable.

output:

Q_A and Q_B lists, both containing points belonging to the same vessel contour.

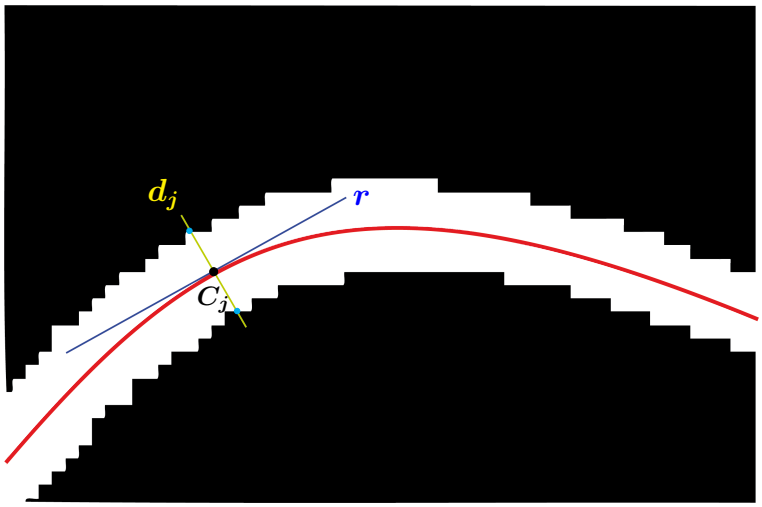


Figure 7: Vessel edge points detection along d_j.

2.5 Contours refinement using a coupled spline-based method

This section deals with the main focus of the thesis: a novel algorithm, based on spline interpolation, fitting points of the two vessel borders detected above. Assuming that retinal vessel boundaries run approximately aligned, a parallel-tangent constraint between the two interpolating splines is enforced at each knot.

The first step is to obtain spline knots as subsampled lists Q_A and Q_B of coupled edge points. The sampling period is chosen as the radius of the smallest vessel bend (fastest turn), expressed in pixel, that we want to detect by interpolating edge points. After subsampling Q_A and Q_B, two spline knots lists L_A and L_B are obtained:

$$L_A = \begin{pmatrix} (x_{A,0}, y_{A,0}) \\ (x_{A,1}, y_{A,1}) \\ \vdots \\ (x_{A,n}, y_{A,n}) \end{pmatrix} \text{ and } L_B = \begin{pmatrix} (x_{B,0}, y_{B,0}) \\ (x_{B,1}, y_{B,1}) \\ \vdots \\ (x_{B,n}, y_{B,n}) \end{pmatrix}$$

Naming y_A the interpolating spline for one vessel border and y_B for the other, the two splines equations for intervals $[x_{A,i}, x_{A,i+1}]$ and $[x_{B,i}, x_{B,i+1}]$ respectively are:

$$y_A = a_i(x - x_{A,i})^3 + b_i(x - x_{A,i})^2 + c_i(x - x_{A,i}) + d_i \tag{2.5.1}$$
$$y_B = \alpha_i(x - x_{B,i})^3 + \beta_i(x - x_{B,i})^2 + \gamma_i(x - x_{B,i}) + \delta_i \tag{2.5.2}$$

The parallelism constraint is enforced at the right end of each pair of coupled intervals $[x_{A,i}, x_{A,i+1}]$ and $[x_{B,i}, x_{B,i+1}]$: the slopes of the tangents to the splines at $y_A(x_{A,i+1})$ and $y_B(x_{B,i+1})$ must be equal. Thereby, grouping Equations (2.5.1) and (2.5.2) with the parallel-tangent constraint, the following overconstrained system is obtained:

$$\begin{cases} y_A = a_i(x - x_{A,i})^3 + b_i(x - x_{A,i})^2 + c_i(x - x_{A,i}) + d_i \\ y_B = \alpha_i(x - x_{B,i})^3 + \beta_i(x - x_{B,i})^2 + \gamma_i(x - x_{B,i}) + \delta_i \\ y'_A(x_{A,i+1}) = y'_B(x_{B,i+1}) \end{cases}$$

The last equation in the system leads to

$$3a_i h_{A,i}^2 + 2b_i h_{A,i} + c_i = 3\alpha_i h_{B,i}^2 + 2\beta_i h_{B,i} + \gamma_i, \tag{2.5.3}$$

where $h_{A,i} = x_{A,i+1} - x_{A,i}$, $h_{B,i} = x_{B,i+1} - x_{B,i}$ and coefficients $a_i, b_i, c_i, \alpha_i, \beta_i, \gamma_i$ can be substituted using standard formulae for spline coefficients (2.2.6),(2.2.7) and (2.2.8). Hence Equation (2.5.3) becomes

$$3\frac{S_{A,i+1} - S_{A,i}}{6h_i}h_{A,i}^2 + 2h_{A,i}\frac{S_{A,i}}{2} + \frac{y_{A,i+1} - y_{A,i}}{h_{A,i}} - \frac{2h_{A,i}S_{A,i} + h_{A,i}S_{A,i+1}}{6} =$$
$$= 3\frac{S_{B,i+1} - S_{B,i}}{6h_i}h_{B,i}^2 + 2h_{B,i}\frac{S_{B,i}}{2} + \frac{y_{B,i+1} - y_{B,i}}{h_{B,i}} - \frac{2h_{B,i}S_{B,i} + h_{B,i}S_{B,i+1}}{6},$$

where $S_{A,i} = y''_A(x_{A,i})$ and $S_{B,i} = y''_B(x_{B,i})$ according to notation used in section 2.2. After some simplifications, previous equation becomes:

$$\frac{1}{3}S_{A,i+1}h_{A,i} + \frac{1}{6}S_{A,i}h_{A,i} - \frac{1}{3}S_{B,i+1}h_{B,i} - \frac{1}{6}S_{B,i}h_{B,i} = f_B[x_{B,i}, x_{B,i+1}] - f_A[x_{A,i}, x_{A,i+1}], \tag{2.5.4}$$

where $f_A[x_{A,i}, x_{A,i+1}] = \dfrac{y_{A,i+1} - y_{A,i}}{h_{A,i}}$, $f_B[x_{B,i}, x_{B,i+1}] = \dfrac{y_{B,i+1} - y_{B,i}}{h_{B,i}}$ and $i = 0, \ldots, n-1$.

Using Equation (2.2.11) for both contours and the parallel-tangent constraint given by (2.5.4), the system can be written in matrix form $A\boldsymbol{x} = \boldsymbol{b}$:

$$A = \left[\begin{array}{c|c} M_1 & 0 \\ \hline 0 & M_2 \\ \hline M_3 & M_4 \end{array}\right]$$

$$\boldsymbol{x} = \begin{bmatrix} S_{A,1} \\ S_{A,2} \\ S_{A,3} \\ \vdots \\ \hline S_{A,n-1} \\ \hline S_{B,1} \\ S_{B,2} \\ S_{B,3} \\ \vdots \\ S_{B,n-1} \end{bmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{bmatrix} 6(f_A[x_1, x_2] - f_A[x_0, x_1]) \\ 6(f_A[x_2, x_3] - f_A[x_1, x_2]) \\ 6(f_A[x_3, x_4] - f_A[x_2, x_3]) \\ \vdots \\ \hline 6(f_A[x_{n-1}, x_n] - f_A[x_{n-2}, x_{n-1}]) \\ \hline 6(f_B[x_1, x_2] - f_B[x_0, x_1]) \\ 6(f_B[x_2, x_3] - f_B[x_1, x_2]) \\ 6(f_B[x_3, x_4] - f_B[x_2, x_3]) \\ \vdots \\ \hline 6(f_B[x_{n-1}, x_n] - f_B[x_{n-2}, x_{n-1}]) \\ \hline f_B[x_0, x_1] - f_A[x_0, x_1] \\ f_B[x_1, x_2] - f_A[x_1, x_2] \\ \vdots \\ f_B[x_{n-1}, x_n] - f_A[x_{n-1}, x_n] \end{bmatrix}$$

where

$$M_1 = \begin{bmatrix} 2(h_{A,0}+h_{A,1}) & h_{A,1} & 0 & 0 & 0 & \dots & 0 \\ h_{A,1} & 2(h_{A,1}+h_{A,2}) & h_{A,2} & 0 & 0 & \dots & 0 \\ 0 & h_{A,2} & 2(h_{A,2}+h_{A,3}) & h_{A,3} & 0 & \dots & 0 \\ \vdots & & & \ddots & & & \\ 0 & 0 & \dots & & & h_{A,n-2} & 2(h_{A,n-2}+h_{A,n-1}) \end{bmatrix},$$

$$M_2 = \begin{bmatrix} 2(h_{B,0}+h_{B,1}) & h_{B,1} & 0 & 0 & 0 & \dots & 0 \\ h_{B,1} & 2(h_{B,1}+h_{B,2}) & h_{B,2} & 0 & 0 & \dots & 0 \\ 0 & h_{B,2} & 2(h_{B,2}+h_{B,3}) & h_{B,3} & 0 & \dots & 0 \\ \vdots & & & \ddots & & & \\ 0 & 0 & \dots & & & h_{A,n-2} & 2(h_{B,n-2}+h_{B,n-1}) \end{bmatrix},$$

$$M_3 = \begin{bmatrix} \frac{1}{3}h_{A,0} & 0 & 0 & 0 & \dots & 0 \\ \frac{1}{6}h_{A,1} & \frac{1}{3}h_{A,1} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{6}h_{A,2} & \frac{1}{3}h_{A,2} & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ & & & & & \frac{1}{3}h_{A,n-2} \\ 0 & 0 & 0 & \dots & 0 & \frac{1}{6}h_{A,n-1} \end{bmatrix},$$

$$M_4 = \begin{bmatrix} -\frac{1}{3}h_{B,0} & 0 & 0 & 0 & \dots & 0 \\ -\frac{1}{6}h_{B,1} & -\frac{1}{3}h_{B,1} & 0 & 0 & \dots & 0 \\ 0 & -\frac{1}{6}h_{B,2} & -\frac{1}{3}h_{B,2} & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ & & & & & -\frac{1}{3}h_{B,n-2} \\ 0 & 0 & 0 & \dots & 0 & -\frac{1}{6}h_{B,n-1} \end{bmatrix}$$

Since interpolating functions are natural cubic splines, coefficients $S_{A,0}$, $S_{A,n}$, $S_{B,0}$ and $S_{B,n}$ are null.

Least-squares solution

If a system $A\boldsymbol{x} = \boldsymbol{b}$ is overdetermined, i.e. it has more equations than unknowns, it is generally impossible to find an exact solution $\boldsymbol{x}$. Vector $A\boldsymbol{x}$ belongs to the column space of A, also called range of A, which is the space spanned by A columns. If $\boldsymbol{b} \notin range(A)$, the right hand of equation $A\boldsymbol{x} = \boldsymbol{b}$ can be replaced with the orthogonal projection $\boldsymbol{\pi}$ of $\boldsymbol{b}$ onto the column space of A. This leads to equation $A\hat{\boldsymbol{x}} = \boldsymbol{\pi}$, where $\hat{\boldsymbol{x}}$ is the least squares system solution i.e. the vector $\hat{\boldsymbol{x}}$ that minimizes the norm of the residual $\boldsymbol{\omega} = A\boldsymbol{x} - \boldsymbol{b}$:

$$||A\hat{\boldsymbol{x}} - \boldsymbol{b}|| \leq ||A\boldsymbol{x} - \boldsymbol{b}||, \text{ for all } \boldsymbol{x}.$$

Since the error vector $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{\pi}$ is orthogonal to $range(A)$, it follows that

$$A^T\boldsymbol{e} = A^T(\boldsymbol{b} - A\hat{\boldsymbol{x}}) = \boldsymbol{0}.$$

Solving this equation for $\hat{\boldsymbol{x}}$, the estimated system solution is obtained:

$$\hat{\boldsymbol{x}} = (A^TA)^{-1}A^T\boldsymbol{b}.$$

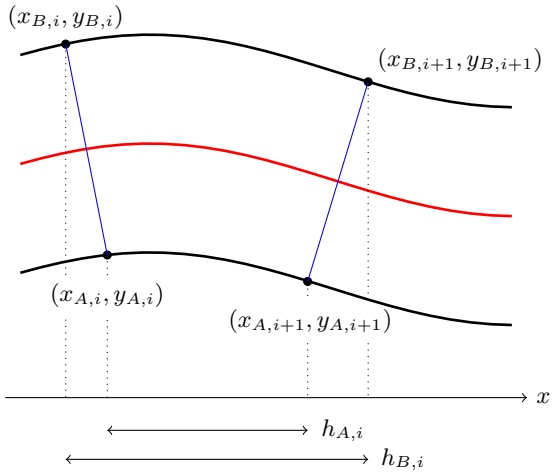


Figure 8: Spline interpolation of two pair of points enforcing the parallel-tangent constraint. The red curve is the centreline. The curve in the upper part of the image is the spline fitting vessel contour B. The one in the lower part is fitting vessel contour A.

Having values $S_{A,i}$ and $S_{B,i}$ from $\hat{\boldsymbol{x}}$, the coefficients $a_i, b_i, c_i, d_i, \alpha_i, \beta_i, \gamma_i$ and δ_i are given by standard spline formulae, reported below for the sake of clarity:

$$
\begin{aligned}
a_i &= \frac{S_{A,i+1} - S_{A,i}}{6h_{A,i}} \\
b_i &= \frac{S_{A,i}}{2} \\
c_i &= \frac{y_{A,i+1} - y_{A,i}}{h_{A,i}} - \frac{2h_{A,i}S_{A,i} + h_{A,i}S_{A,i+1}}{6} \\
d_i &= y_{A,i} \\
\alpha_i &= \frac{S_{B,i+1} - S_{B,i}}{6h_{B,i}} \\
\beta_i &= \frac{S_{B,i}}{2} \\
\gamma_i &= \frac{y_{B,i+1} - y_{B,i}}{h_{B,i}} - \frac{2h_{B,i}S_{B,i} + h_{B,i}S_{B,i+1}}{6} \\
\delta_i &= y_{B,i}.
\end{aligned}
$$

2.6 Width estimation

The contours refinement method proposed improves vessel width estimations in binary images since vascular boundaries are smoothed and the typical indentation of binary edges is removed. Thus, multiple diameter measurements along the same vessel will not present a high standard deviation any longer. The vessel width at point C_j lying on the spline-smoothed centreline is estimated computing the Euclidean distance between points D_j and E_j: these are the points belonging to the two refined contours and lying on segment d_j, orthogonal to centreline at C_j.

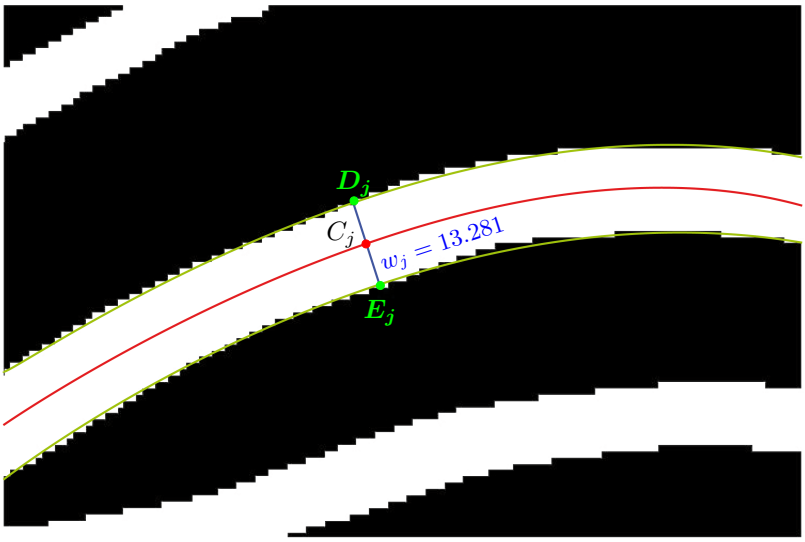


Figure 9: The vessel width at C_j is estimated as the Euclidean distance w_j between D_j and E_j.

3 Experimental Results

3.1 REVIEW database

The diameter measurement performance is evaluated using the REVIEW database [8], which is publicly available at *http://ReviewDB.lincoln.ac.uk*. This database comprises four image sets presenting a good overview of different possible kind of retinal images: high-resolution (HRIS dataset), central light reflex (CLRIS dataset), vascular diseases (VDIS dataset) and kickpoints (KPIS dataset). All the images were assessed by three different experts that manually marked vessel edge points: the co-ordinates of these hand-selected points are also available in the database. Therefore, each pair of edge points presents three width estimates, one for each observer; the average of these three diameter measurements is considered as the ground truth vessel width. REVIEW database contains a total of 5066 manually marked profiles.

A brief overview about the four image sets follows:

- HRIS: It is composed by 4 high-resolution images (3584×2438 pixels) representing different grades of Diabetic Retinopathy. It contains 90 hand-selected segments for a total of 2368 profiles. Measurements reported in REVIEW database refer to HRIS images that are downsampled by a factor of four, so that estimated widths are known to ± 0.25 pixels, leaving out human error.
- CLRIS: Two images representing early artheriosclerotic changes, often with the presence of a light streak, known as central light reflex, running down the central length of the blood vessel. Image resolution is 1440×2169 and 285 profiles from 21 segments are reported.
- VDIS: It consist of 8 images (1360×1024 pixels), with the presence of different type of Diabetic Retinopathy in 6 of them. Manually selected profiles are 2249. Since images in this dataset are considerably noisy, vessel measurement algorithms generally perform worse here.
- KPIS: In contrast with previous three datasets, which contain full fundus images, KPIS consists of two images representing several large and non-tortuous vessel segments. Images resolution is 760×570 and the dataset comprises 164 profiles.

The authors of the database proposed the Extraction of Segment Profiles (ESP) procedure [9] and they compared its performance with Gregson's [10], Half Height Full Width (HHFW) [11], 1D Gaussian [12] and 2D Gaussian [13] algorithms on the REVIEW database.

3.2 Algorithm performance measures

Naming O_1, O_2 and O_3 the three observers that manually marked vessel profiles in REVIEW database, the reference standard vessel diameter at ith location is the mean ψ_i of width measurement from O_1, O_2 and O_3. For comparison of different algorithms efficiency, the error χ_i is defined as

$$\chi_i = w_i - \psi_i$$

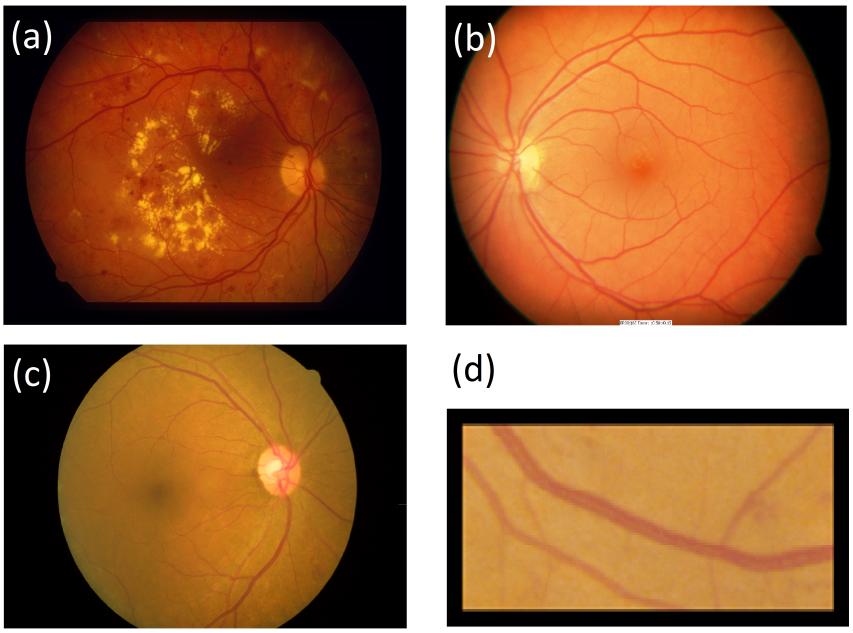


Figure 10: Example images from REVIEW database. (a) HRIS. (b) VDIS. (c) CLRIS. (d) KPIS.

where w_i is the width at ith location estimated by the algorithm under examination. The standard deviation σ_χ of the error is used to evaluate algorithm performance. The error mean μ_χ and the sandard deviation σ and μ of the measurements w are also reported for the sake of completeness.

A further useful parameter for performance evaluation is the success rate (SR): it is considered as the number of meaningful measurements returned by the algorithm over the total number of profile reported in the database. In fact, vessel segmentation procedures sometimes produce inaccurate binary masks, with either the lack of some vessel parts or too noisy vessel profiles. Since the algorithm proposed in this thesis works on a binary image, resulting from a previous vessel segmentation step, its performance strongly depends on the quality of the input binary mask. Hence, in the cases mentioned above, diameter measurements are not reported for algorithm performance evaluation, because width estimations are meaningless: this affects the algorithm success rate.

In order to compare ψ_i with the diameter w_i, estimated with the spline based method, the following procedure has to be applied: for each pair of ground truth vessel edge co-ordinates $(x_{A,O}, y_{A,O})$ and $(x_{B,O}, y_{B,O})$ in REVIEW, the reference standard centreline point $C_O(x_C, y_C)$ is computed, where

$$x_C = \frac{x_{A,O} + x_{B,O}}{2} \text{ and } y_C = \frac{y_{A,O} + y_{B,O}}{2}.$$

After that, the method described in section 2.6 to find w_i is applied: in this case, C_j is the point lying on the spline-smoothed centreline which is the closest to C_O.

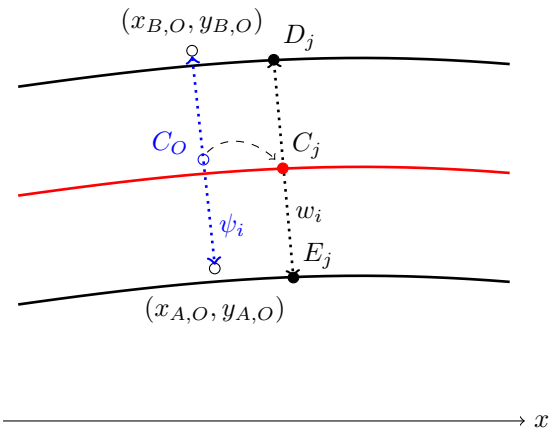


Figure 11: Ground truth centreline point C_O and the closest point C_j lying on the spline-smoothed centreline.

Tables 1 to 4 reports the performance of algorithms [9], [10], [11], [12], [13] and of Xu's graph-based method [14], [15]. The accuracy of the spline-based algorithm presented in this document is shown in Tables 6 to 8. These tables show results of the algorithm applied to different binary images obtained with several segmentation procedures.

HRIS Dataset:

Algorithm name	Measurement		Error		SR %
	μ	σ	μ_χ	σ_χ	
First observer: O_1	4.12	1.25	-0.23	0.288	100
Second observer: O_2	4.35	1.35	0.002	0.256	100
Third observer: O_3	4.58	1.26	0.23	0.285	100
Ground truth: O	4.35	1.26	-	-	100
Gregson's algorithm	7.64	-	3.29	2.841	100
HHFW	4.97	-	0.62	0.926	88.3
1D Gaussian	3.81	-	-0.54	4.137	99.6
2D Gaussian	4.18	-	-0.17	6.019	98.9
ESP	4.63	-	0.28	0.42	99.7
Xu's graph	4.56	1.30	0.21	0.567	100

Table 1: Performance of main width measurement methods in HRIS dataset

As it is shown on Table 1, the accuracy σ_χ of Gregson's algorithm and of 1D and 2D Gaussian methods is particularly poor on the High Resolution Image Set: even though they all score an high SR, the error standard deviation is inferior compared to other algorithms. On the other hand, ESP and graph-based method performances are comparable to the observers: 1.52 times the mean accuracy of the observers in the first case, 2.05 times in the second.

CLRIS Dataset:

Algorithm name	Measurement		Error		SR %
	μ	σ	μ_χ	σ_χ	
First observer: O_1	13.19	4.01	-0.61	0.567	100
Second observer: O_2	13.69	4.22	-0.11	0.698	100
Third observer: O_3	14.52	4.26	0.72	0.566	100
Ground truth: O	13.80	4.12	-	-	100
Gregson's algorithm	12.8	-	-1.0	2.841	100
HHFW	-	-	-	-	0
1D Gaussian	6.3	-	-7.5	4.137	98.6
2D Gaussian	7.0	-	-6.8	6.019	26.7
ESP	15.7	-	-1.90	1.469	93.0
Xu's graph	14.05	4.47	0.08	1.78	94.1

Table 2: Performance of main width measurement methods in CLRIS dataset

Due to the bright strip running down some vessels centreline, the detection of vessel borders for diameter measurements turns out to be tricky in this dataset. In fact, the central light reflex causes a complicated intensity cross-section and some border detection methods can interpret one vessel as two different vessels running parallel each other. For this reason HHFW, 1D and 2D Gaussian algorithms are unstable in this dataset. Even if ESP and Xu's procedures accuracy is not close to the observers, these methods still outperform the other edge location methods in this dataset.

VDIS Dataset:

Algorithm name	Measurement		Error		SR %
	μ	σ	μ_χ	σ_χ	
First observer: O_1	8.50	2.54	-0.35	0.543	100
Second observer: O_2	8.91	2.69	0.06	0.621	100
Third observer: O_3	9.15	2.67	0.30	0.669	100
Ground truth: O	8.85	2.57	-	-	100
Gregson's algorithm	10.07	-	1.22	1.494	100
HHFW	7.94	-	-0.91	0.879	78.4
1D Gaussian	5.78	-	-3.07	2.110	99.9
2D Gaussian	6.59	-	-2.26	1.328	77.2
ESP	8.80	-	-0.05	0.766	99.6
Xu's graph	8.35	3.00	-0.53	1.43	96.0

Table 3: Performance of main width measurement methods in VDIS dataset

Since VDIS dataset is sensibly noisy, algorithms accuracy is comparatively weaker here. Again ESP have the best accuracy $\sigma_\chi = 0.766$ pixels (1.26 times the mean accuracy of the observers). Also the error standard deviation reported

by Half Height Full Width is close to the observers, but its SR is quite low, so HHFW results are not comparable with the others.

KPIS Dataset:

Algorithm name	Measurement		Error		SR %
	μ	σ	μ_χ	σ_χ	
First observer: O_1	7.97	0.47	0.45	0.233	100
Second observer: O_2	7.60	0.42	0.08	0.213	100
Third observer: O_3	7.00	0.52	-0.53	0.234	100
Ground truth: O	7.52	0.42	-	-	100
Gregson's algorithm	7.29	-	-0.23	0.602	100
HHFW	6.47	-	-1.05	0.389	96.3
1D Gaussian	4.95	-	-2.57	0.399	100
2D Gaussian	5.87	-	-1.65	0.337	100
ESP	6.56	-	-0.96	0.328	100
Xu's graph	6.38	0.59	-1.14	0.67	99.4

Table 4: Performance of main width measurement methods in KPIS dataset

In this low resolution dataset, all the algorithms score high SR and exhibit good performance. The worse accuracy is reported by the graph-based method: σ_χ is 2.96 times the mean standard deviation of the observers and vessel withs are underestimated ($\mu_\chi = -1.14$ pixels).

Segmentation using a Gaussian filter

This paragraph shows the results of the spline-based contour refinement method on retinal binary images obtained as described in Section 2.1. Since the size of the vessels, in terms of pixels, strictly depends on image resolution, the standard deviation of the Gaussian filter should be finely tuned to each particular image, in order to detect both narrow and wide vessels. For this reason, the matched filter scales are adjusted for each dataset and same scales are used for images with similar resolutions.

HRIS dataset has the highest resolution, but all its images are downsampled by a factor of 4 before being input into the filtering procedure. It turns out that the best choice of σ values in Equation (2.1.1) are 2, 3 and 4; in CLRIS σ is set to values from 4 to 6; in VDIS the filter scales are 3 and 4; in the low resolution KPIS, σ is set to values 2 and 3. Results reported below are achieved using a Gaussian shape sliding-window of size 33×33 pixels. Filtered images are then thresholded to obtain binary vessel maps, which are input into the spline-based algorithm for width measurements. The threshold values used in this work are reported in table 5 and the results, referring to the ground truth observer O measurements, are shown in Table 6.

Dataset	Image 1	Image 2	Image 3	Image 4	Image 5	Image 6	Image 7	Image 8
HRIS	97	68	60	100	-	-	-	-
CLRIS	54	53	-	-	-	-	-	-
VDIS	15	23	29	37	50	41	27	31
KPIS	50	56	-	-	-	-	-	-

Table 5: Threshold values used to obtain binary vessel maps from filtered images.

Dataset	Measurement		Error		SR %
	μ	σ	μ_χ	σ_χ	
HRIS	3.935	1.398	-0.425	0.760	95.65
CLRIS	13.808	3.681	-0.159	1.229	90.18
VDIS	8.174	2.819	-0.792	1.381	92.08
KPIS	6.061	0.285	-1.325	0.319	93.90

Table 6: Performance of spline-based method for width measurement, applied to binary images that obtained after filtering them with Gaussian sliding-window.

The method proposed in this thesis has a performance which is comparable to the observers in HRIS dataset: $\sigma_\chi = 0.760$ pixels (2.75 times the mean of observers' σ_χ). Nevertheless, Xu's graph-based method and ESP algorithm perform slightly better. On the contrary, in CLRIS dataset, the spline-based method reports the best accuracy, despite the central light reflex. It's success rate is yet lower than ESP and Xu's algorithms. On the noisy VDIS dataset the width measurement by spline-based contours refinement reports the second best performance after ESP. Finally, in KPIS dataset the spline-based method reports again the best performance, even though it scores a SR slightly lower that other algorithms.

Segmentation using 2-D Gabor wavelet transforms

This chapter deals with the results of the spline-based algorithm applied to retinal binary images obtained with another vessel segmentation procedure. The software used to produce binary masks is VAMPIRE [16]: this is an application detecting the vessel structure and quantifying vessel features in retinal fundus images. It implements a version of Soares's algorithm [1] for the vessel segmentation step. This method is a supervised procedure using a Bayesian classifier, with class-conditional probability density functions modelled as Gaussian mixtures, to label each pixel as vessel or non-vessel. The pixel feature vector includes multi-scale measurements taken from the 2-D Gabor wavelet transform, enhancing vessels appearance.

The main drawback of this segmentation algorithm is that it requires to train the Bayesian classifier. Since the Soares's algorithm version implemented in VAMPIRE is not trained on images from REVIEW database, the resulting binary masks are very noisy and inaccurate for the most part. This affects the width measurement accuracy of the spline-based method, as it can be inferred by Table 7.

Dataset	Measurement		Error		SR %
	μ	σ	μ_χ	σ_χ	
HRIS	5.304	1.229	1.034	0.666	86.91
CLRIS	13.792	2.527	1.195	1.354	70.18
VDIS	7.980	3.010	-0.901	2.283	74.30
KPIS	-	-	-	-	-

Table 7: Performance of spline-based method for width measurement applied to binary images that are obtained using Soares's algorithm.

The great deterioration of the success rate is due to large number of undetected vessel pixels in binary images obtained with VAMPIRE. In fact, many locations in which REVIEW reports width measurements, are not even classified as vessel pixels by Soares's algorithm. Nevertheless, at those locations where vessels are detected properly, the spline-based algorithm reports width measurements that are comparable with the ground truth: σ_χ is 2.41 times the mean accuracy of the observers in HRIS and 2.22 times in CLRIS. Results in VDIS database are sensibly poorer compared with those reported in Table 3: this is due to the fact that images in VDIS present an high level of noise that makes the vessel segmentation harder. Particularly, the first and the last image of this dataset are such noisy that Soares's algorithm is not able to detect most of the locations reported in REVIEW for width measurements. The next paragraph proposes the results of spline-based algorithm on binary images obtained after a noise-reduction procedure. Performances in KPIS dataset are not reported, as VAMPIRE can produce meaningful results just processing fundus retinal images.

Segmentation with 2-D Gabor wavelet transforms after median filtering

In order to reduce impulsive noise and "salt and pepper" effect present in several REVIEW images, a median filtering is applied to each image before input it into Soares's algorithm. *Median filtering* is a non-linear smoothing operation, often used in image processing, as it turns out to be effective when the goal is to reduce noise and preserve edges at the same time. In this work, median filtering is performed moving a 3×3 sliding-window through all the pixels of the image: each pixel in the centre of the window is replaced with the median of the brightnesses in its 3×3 neighbourhood. Since the median of brightness values in the neighbourhood is not affected by individual noise spikes, the resulting image is free from impulsive noise. Filtered images are then processed using VAMPIRE software. The resulting binary images are input into the spline-based algorithm for diameter estimates and the measurements performance against REVIEW is reported in Table 8.
The removal of impulsive noise leads to an improved success rate in each dataset. The accuracy σ_χ in HRIS is even better than those reported in Table 6 and 7: $\sigma_\chi = 0.639$ pixels (2.31 times the observers' mean accuracy). On the contrary, in CLRIS and VDIS the standard deviations of the error are slightly worse compared to those obtained without previously applying a median filtering.

Dataset	Measurement		Error		SR %
	μ	σ	μ_χ	σ_χ	
HRIS	5.18	1.262	0.947	0.639	91.05
CLRIS	13.638	3.425	1.453	1.739	77.89
VDIS	9.211	3.655	0.126	2.323	80.35
KPIS	-	-	-	-	-

Table 8: Performance of spline-based method for width measurement applied to binary images that obtained using Soares's algorithm on median filtered REVIEW images.

KPIS images are not processed for the same reasons mentioned above.

3.3 Processing times

The algorithm described in this document takes few seconds to find the two splines refining the boundaries of one single vessel. Timings are related with the length of the processed segment: the following table gives computation times required to fit contours of vessels with different lengths.

Vessel centreline length (pixels)	Time (seconds)
57	2.07
97	2.78
209	5.60
362	7.15

Table 9: Mean computation times of the spline-based algorithm applied to four different vessels.

These processing times are obtained using a MATLAB implementation of the spline-based algorithm. Tests has been made on an Intel(R) Core(TM)2 Duo CPU (2.26 GHz) with 3GB RAM memory.

4 Discussion and conclusion

This thesis has proposed a novel algorithm refining vessel boundaries in retinal binary vessel maps. Since these black-and-white vessel maps often present jagged edges that are not suitable for width estimations, the spline-based algorithm provides an improved version of the input binary image, in which vessel contours are smoothed and refined. However, this procedure does not perform vessel extraction; rather, it refines the boundaries of vessels that have been already located by previous vascular segmentation methods.

The algorithm is general enough to be applied to any kind of retinal binary mask, even to black-and-white images not representing a full eye fundus. As it can be inferred from results reported in Section 3.2, one important drawback of this approach is that its diameter evaluation performance strictly relies on the accuracy of the input binary image. In fact, if vessels are not reliably segmented or not detected at all by the segmentation step, the diameter measurement taken with the spline-based algorithm will be meaningless. Although this limitation is significant, the problem may be solved integrating the method with a reliable segmentation procedure. The method proposed in Section 2.1, convolving the image with a Gaussian shape filter, and Soares's algorithm [1] are just two of several vascular segmentation methodologies. The integration of the spline-based algorithm with more accurate vessel segmentation procedures, may lead to an improved performance in width measurements.

Future work Results reported in this thesis are encouraging and suggest that the spline-based algorithm has a good potential for improving vessel contours obtained from binary vessel masks. However, some issues must still be solved by future work.

- Firstly, the choice of spline knots. The set of contour points that are interpolated by the two spline curves depends on how edge points lists Q_A and Q_B of section 2.5 are subsampled. The choice made in this work is straightforward: the sampling period is set as the smallest vessel bend radius that has to be detected by the spline interpolation. A more thoughtful solution, taking into account some vessel parameters, such as the curvature and length of the centreline, may lead to an improved contours refinement.

- Secondly, a procedure generating more reliable binary vessel masks. Even if image segmentation is one of the most studied topics in medical image processing and many innovative solutions are often suggested by scientists, further research is still needed to find a segmentation procedure suitable to different kind of retinal images.

References

[1] Soares J.V.B, Leandro J.J.G, Caesar Jr.R.M, Jelinek H.F, Cree M.J, "Retinal Vessel Segmentation Using the 2-D Gabor Wavelet and Supervised Classification", *IEEE Trans on Med Imaging*, Vol 25, No. 9, September 2006, pp 1214-1222.

[2] Lupaşcu C.A, Tegolo D, Trucco E, "FABC: Retinal Vessel Segmentation Using AdaBoost", *IEEE Trans on Info Tech in Biomedicine*, Vol 14, No. 5, September 2010, pp 1267-1274.

[3] Chutatape O, Zheng L, Krishnan S, "Retinal blood vessel detection and tracking by matched Gaussian and Kalman filters", *Proc IEEE Int Conf Emg Bio Soc*, Vol 20, No.6, 1998, pp 3144-3149.

[4] Bankhead P, Scholfield C.N, McGeown J.G, Curtis T.M, "Fast Retinal Vessel Detection and Measurement Using Wavelets and Edge Location Refinement", *PLoS ONE*, Vol 7, Issue 3, March 2012, e32435, pp 1-12.

[5] Vermeer K.A, Vos F.M, Lemij H.G, Vossepoel A.M, "A model based method for retinal blood vessel detection", *Comput Biol Med*, Vol 34, 2004, pp 209-219.

[6] Chaudhuri S, Chatterjee S, Katz N, Nelson M, Goldbaum M, "Detection of blood vessels in retinal images using two-dimensional matched filters", *IEEE Trans on Med Imaging*, Vol 8, No. 3, September 1989, pp 263-269.

[7] Curtis F.G, Wheatley P.O (1970), *Applied numerical analysis*, Addison-Wesley, Reading, 1989.

[8] Al-Diri B, Hunter A, Steel D, Habib M, Hudaib T, Berry S, "REVIEW - A Reference Data Set for Retinal Vessel Profiles", *30th Annual International Conference of the IEEE Engineering in Medicine and Biology*, August 2008, pp 2262-2265.

[9] Al-Diri B, Hunter A, Steel D, "An active contour model for segmenting and measuring retinal vessels", *IEEE Trans on Med Imaging*, Vol 28, No. 9, March 2009, pp 1488-1497.

[10] Gregson P.H, Shen Z, Scott R.C, Kozousek V, "Automated grading of venous beading", *Computers and Biomedical Research*, Vol 28, No. 4, 1995, pp 291-304.

[11] Brinchmann-Hansen O, Heier H, "Theoretical relations between light streak characteristics and optical properties on retinal vessels", *Acta Ophthalmologica supplement*, Vol 179, 1986, pp 33-37.

[12] Zhou L, Rzeszotarski M, Singerman L, Chokreff J, "The detection and quantification of retinopathy using digital angiograms", *IEEE Trans on Med Imaging*, Vol 13, No. 4, December 1994, pp 619-626.

[13] Lowell J, Hunter A, Steel D, Basu A, Ryder R, Kennedy R.L, "Measurement of retinal vessel widths from fundus images based on 2-D modeling", *IEEE Trans on Med Imaging*, Vol 23, No. 10, October 2004, pp 1196-1204.

[14] Xu X, Niemeijer M, Song Q, Sonka M, Garvin M.K, "Vessel boundary delineation on fundus images using graph-based approach", *IEEE Trans on Med Imaging*, Vol 30, No. 6, June 2011, pp 1184-1191.

[15] Xu X, Niemeijer M, Song Q, Garvin M.K, Reinhardt J, Abràmoff "Retinal vessel width measurements based on a graph-theoretic method", *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2011, pp 614-644.

[16] Perez-Rovira A, MacGillivray T, Trucco E, Chin K.S, Zutis K, Lupaşcu C.A, Tegolo D, Giachetti A, Wilson P.J, Doney A, Dhillon B, "VAMPIRE: Vessel Assessment and Measurement Platform for Images of the REtina", *33th International IEEE EMBS Conference*, Boston (USA), 2011.

[17] Sonka M, Hlavac V, Boyle R (1993), *Image Processing, Analysis, and Machine Vision, Second edition*, PWS an Imprint of Brooks and Cole Publishing, 1998.

[18] Gonzales R.C, Woods R.E, Eddins S.L (2003), *Digital Image Processing using MATLAB, Second edition*, Gatesmark Publishing, 2009.