Retinal vessel segmentation using multiwavelet kernels and multiscale hierarchical decomposition

Yangfan Wang\textsuperscript{c,a,e}, Guangrong Ji\textsuperscript{d}, Ping Lin\textsuperscript{a,e,*}, Emanuele Trucco\textsuperscript{b}

\textsuperscript{a} Division of Mathematics, University of Dundee, Dundee DD1 4HN, UK
\textsuperscript{b} School of Computing, University of Dundee, Dundee DD1 4HN, UK
\textsuperscript{c} College of Marine Life Science, Ocean University of China, Qingdao 266100, PR China
\textsuperscript{d} College of Information, Ocean University of China, Qingdao 266100, PR China
\textsuperscript{e} Department of Mathematics, University of Science and Technology Beijing, Beijing 100083, PR China

1. Introduction

We propose a novel, general method for segmenting the retinal vasculature in fundus camera images. Locating the retinal vasculature, assessing its morphological properties, and detecting abnormalities play an important role for various purposes. These include, among others, diabetic screening and the detection of lesions associated with diabetes, e.g., diabetic retinopathy, retinopathy of prematurity, and cerebrovascular diseases. For such diagnostic purposes, automatic or semi-automatic image analysis holds important promises. First, rich, quantitative sets of measurements providing clinicians with extensive information extracted from images and supporting accurate diagnosis. Second, repeatable measurements which could contribute to reduce the variability of medical diagnosis. Third, identifying and summarizing key information in the large quantities of data present in retinal exams. For instance, fundus cameras acquire nowadays high-resolution images.

On the non-diagnostic side, much work has been reported on the discovery of biomarkers associated with the retinal vasculature and a variety of conditions, e.g., stroke and hypertension. Typical candidate biomarkers are related to vessel calibre, branching angles and branching coefficients, vessel tortuosity and, less frequently, the fractal dimension of the vasculature network. In cognitive psychology, retinal biomarkers have recently been correlated with cognitive decline.

For image analysis, detecting the retinal vasculature means, in essence, generating a binary mask in which pixels are labeled as vessel or background. The target is to capture as much detail as possible, simultaneously avoiding false positives and, ideally, preserving vessel connectivity. However, it should be noted that many clinical investigations do not use fine vessels, taking measurements only on major ones in a limited region around the optic disc. This may of course depend on the current absence of reliable detectors of fine vessels. The relevant literature is discussed in Section 2.

This paper brings the following contributions:

1. We introduce a novel vessel enhancement technique based on the matched filters with multiwavelet kernels (MFMK). We identify kernels separating vessels from clutter edges and bright, localized features (e.g., lesions).

2. For noise attenuation and vessel localization, we apply a multiscale hierarchical decomposition, which is particularly effective for the normalized enhanced image. This process performs an iterative segmentation at increasing image resolutions, locating smaller and smaller vessels. A single scale
parameter controls the level of detail included in the vessel map. We show a necessary condition to achieve the optimal decomposition, deriving a rule to identify the optimal number of the hierarchical decomposition.

3. Our method does not require preprocessing and training it can therefore be used directly on images with different characteristics. In addition, it relies on adaptive thresholding so that no numerical parameter is tuned manually to obtain a binary mask.

The image segmentation we obtain from the iterative multi-resolution analysis is a gray-scale image. To achieve a binary map, we adopt the spatially adaptive thresholding method [15], which computes a threshold surface over the image. The main idea is to use zero-crossings points as interpolation constraints for the target threshold surface. The interpolating problem is solved by minimizing an energy functional. In this way no threshold value must be tuned by hand.

We report experimental results on two standard retinal data sets, DRIVE and STARE, demonstrating excellent performance in comparison with retinal vessel segmentation methods reported recently.

The remainder of this paper is organized as follows. Section 2 summarizes the state-of-the-art automatic retinal vasculature detection. Section 3 presents our algorithm. Section 4 reports and discusses our experimental results. Section 5 summarizes the paper and offers some conclusive remarks.

Fig. 1. Example of central reflection. (a) Excerpt from fundus image, showing vein (marked 18) and artery (marked 19). (b) Intensity profile across vein, taken along black line is shown. (c) Same for artery, showing central reflection.

2. Related work

Retinal vessel detection systems can be discussed along several dimensions. Here, we consider briefly filters, tracking, supervised learning, and cross-sectional intensity models. This section builds on our discussion of the literature in [16].

Filters. Matched filters for retinal vessel segmentation appear in early works. For instance, in [17] the gray-level profile of the cross-section of a blood vessel is approximated by a Gaussian curve. Vessel segments are searched in all possible directions using a two-dimensional matched filter. Hoover et al. [18] noticed that a single global threshold applied to the filter’s output does not yield a satisfactory classification, and propose a vessel segmentation method that uses local and region-based properties at each pixel. Pixels are classified as vessel or non-vessel by thresholding the image generated by a matched filter using a probing technique. Probing allows a pixel to be tested in multiple region configurations before the final classification. Mendonca and Campilho [19] detected the retinal vascular network automatically by first extracting vessel centerlines using differential filters and then applying morphological operators for filling vessel segments. Mathematical morphology and curvature evaluation are used also by Zana and Klein [20] for the detection of vessel-like patterns. Soares et al. [21] adopt Gabor wavelets as they provide directional selectivity and fine tuning to specific frequencies, enabling noise reduction. In our system, we choose a system of wavelets providing the same properties, and additionally accounting for the different cross-sectional intensity profiles.
encountered in fundus images, and responding selectively to signal (vessels) and clutter.

Vessel tracking methods select a set of reliable seed points on the vessel network, and track the vessels starting from the seeds, similar to classical region growing. For example, Chutatape et al.'s[22] system looks for vessel seed points on the perimeter of the optic disc; once a seed is selected, the vessel is tracked using Kalman filtering. During tracking, the branching points are used as starting points for tracking secondary vessels. Related tracking strategies were adopted in[23–25].

Supervised learning classifies individual pixels and requires hand-labeled gold standard images for training. The system by Sinthanayothin et al.[26] uses a multilayer perceptron, for which the inputs are derived from the principal component analysis (PCA) of the image and edge detection of the first PCA component. Niemeijer et al.[27] extract a feature vector for each pixel from the green channel, then use a K-nearest-neighbor (kNN) algorithm to estimate the probability of the pixel belonging to a vessel. Another supervised method, called primitive-based method, was proposed by Staal et al.[28]. This algorithm is based on the extraction of image ridges (expected to coincide with vessel centerlines) used as primitives for describing linear segments, named line elements. Each pixel is assigned to the nearest line element to form image patches and then classified using a set of features from the corresponding line and image patch. The feature vectors are classified using a kNN classifier and sequential forward feature selection. Soares’ method[21] also adopts supervised classification. Each image pixel is classified as vessel or non-vessel based on the pixel feature vector, which is composed of the pixel intensity and two-dimensional Gabor wavelet transform responses taken at multiple scales. A Gaussian-mixture model classifier is then applied to obtain the final segmentation. Lupascu et al.[16] built a 41-dimensional feature vector for each pixel, encoding information on the local intensity structure, spatial properties, and geometry at multiple scales. The feature vectors are used in an AdaBoost classifier trained on 789,914 gold-standard examples of vessel and non-vessel pixels. Versions of the original image at multiple scales are achieved by building a classic Gaussian scale-space. In our method, instead, pixels are classified according to a multiwavelet scheme.

Cross-sectional intensity models: Several models have been devised to describe the profile of intensity variations across a blood vessel, taken perpendicularly to the vessel’s axis, and to describe the geometry and morphology of the vessels. The simplest and earliest models of the intensity profile are a Gaussian[17] or a top-hat[20]. Matched filters based on these profiles were applied in multiple directions and the maximum response thresholded. This basic algorithm has been shown to generate basic vasculature maps, but significant post-processing is needed to capture small vessels while simultaneously limiting spurious responses. Mixtures of Gaussians[29] and a modified Gaussian[30] have been used to model the dipped (multimodal) intensity profile of some large vessels due to a central light reflection, whereby the edges of the vessel appear darker than the central region (see Fig. 1). Bhalerao et al. in[31] adopt a two-dimensional Hermite function intensity model and a quad-tree structure over a range of spatial resolutions.

Models of shape and morphology have been introduced for accurate estimation of vessel width. An example is the “ribbon of twins” or “tramline” model by Hunter et al.[29,32], whereby two parallel edges at variable distance from each other are expected to detect a vessel locally. This was subsequently used to provide accurate measurements of caliber. The author has provided public data in the REVIEW data set[33], available from http://reviewdb.lincoln.ac.uk/.

Finally, we notice that extensive validation of vasculature detection algorithms is a nontrivial task, generating accurate manual segmentations (binary masks) is an extremely time-consuming task. This point is expanded briefly in Section 4.1.

3. Methods

3.1. Overview and preliminaries

The essential architecture of our algorithm is shown in Fig. 2. We introduce the MFKM vessel enhancement in Section 3.2, the hierarchical optimal decomposition and a stopping criterion in Section 3.3, and adaptive thresholding in Section 3.4. Following common practice, we process the green channel only, in which vessels appear to be well contrasted.

3.2. Enhancement: MF with multiwavelet kernels

In this section we devise a vessel enhancement algorithm based on the multiwavelets kernels. The wavelet method is a powerful and versatile tool that has been applied in many different image processing problems, from image coding to shape analysis. This success is largely due to the fact that wavelet kernels are especially suitable for detecting edges in signals such as blood vessel borders. Classical wavelet theory or its multisresolution analysis[35] is based on one scaling function, and dilates and translates one wavelet function to form a basis. On the other hand, multiwavelets, which allow several (vector) scaling and wavelet functions, will have more freedom in their construction and thus can combine more useful properties than a scalar wavelet. Our objective is twofold: not only to identify blood vessels but also to remove non-vessel edges. It is thus more possible to work under multiwavelet kernels. Multiwavelets (as shown in[36]) are based on a vector scaling function \( \phi(x) = [\phi_1(x), \ldots, \phi_N(x)]^T \), where \( x \) is the one-dimensional space variable. The vector scaling function satisfies

Fig. 2. Essential architecture of our vessel detection algorithm. \( u \) and \( v \) are the signal and noise components, see Section 3.3.
a matrix dilation equation analogous to the scalar case
\[ \phi(x) = \sum_k G[k] \phi(2x-k), \]  
(1)

where the coefficients \( G[k] \) are \( N \) by \( N \) matrices. Associated with these scaling functions are \( N \) wavelets \( \psi(x) = [\psi_1(x), \ldots, \psi_N(x)]^T \), satisfying the matrix wavelet equation
\[ \psi(x) = \sum_k H[k] \phi(2x-k), \]  
(2)

where the coefficients \( H[k] \) are \( N \) by \( N \) matrices.

It turns out that the multiwavelet system proposed by Chui and Lian [37] fulfils our requirement for vessel enhancement, as we shall discuss next. For this system, we have \( N = 2 \) and the non-zero coefficient matrices are
\[
G[0] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \quad G[1] = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix},
\]
\[
G[2] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.
\]
\[
H[0] = \frac{1}{4\sqrt{2}} \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}, \quad H[1] = \frac{1}{4\sqrt{2}} \begin{bmatrix} -4 & 0 \\ 0 & 2\sqrt{2} \end{bmatrix},
\]
\[
H[2] = \frac{1}{4\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.
\]

The scaling functions \( \phi_1, \phi_2 \) and wavelet functions \( \psi_1, \psi_2 \) can be obtained by solving the dilation and wavelet equations (1) and (2) numerically with initial scaling functions
\[
\phi_1(x) = \phi_2(x) = \begin{cases} 1, & 0 \leq x \leq 2, \\ 0, & \text{otherwise}. \end{cases}
\]

For illustration, the scaling functions \( \phi_1, \phi_2 \) and wavelet functions \( \psi_1, \psi_2 \), in the interval \([-6,6]\), are shown in Fig. 3. These scaling and wavelet functions can be used as multiwavelet kernels, or CL-MK in short.

As observed in the figure, the intensity profile of some major vessels may be multimodal due to a central specular reflection [38] (see Fig. 1). The CL-MK \( \phi_1 \) captures the profile generated by central-reflection vessels (see Figs. 1(c) and 3(a)).

We now want to decide which functions can be used as filters to enhance vessels features. To this purpose, we first observe the responses of such filters to ideal Gaussian vessel profiles as well as to step edges (generated by non-vessel clutter) depicted in Fig. 4. Then Fig. 5 shows the responses of CL-MKs \( \phi_1, \phi_2, \psi_1 \), and \( \psi_2 \) to an ideal Gaussian vessel profile (depicted in Fig. 4(b)) and, for comparison, to a general step edge (depicted in Fig. 4(c)). In Fig. 5(a), (c), (e) and (g), the solid lines are the responses of CL-MKs \( \phi_1, \phi_2, \psi_1 \), and \( \psi_2 \) to a Gaussian profile; the dash-dot lines are the responses of CL-MKs \( \phi_1, \phi_2, \psi_1 \), and \( \psi_2 \), respectively, to a general step edge. Both CL-MKs \( \phi_1 \) and \( \phi_2 \) respond strongly to the Gaussian and the step edge, while CL-MKs \( \psi_1 \) or \( \psi_2 \) is relatively weak at positive values. We also compute the local mean of the CL-MK-filtered images, illustrated in Fig. 5(b), (d) and (h). Fig. 5(d) indicates that the local mean of the response of CL-MK \( \phi_2 \) to the step edge is strong but the corresponding response to the vessel profile is very low. This suggests to use \( \phi_2 \) to enhance step edges (non-vessel clutter edges). We now have identified filters to selectively enhance vessels and non-vessel edges.

Finally, we consider the filters’ response to localized, distracting bright features such as microaneurysms, modeled here as a Gaussian impulse. Fig. 6 shows that CL-MK \( \phi_1 \) and \( \phi_2 \) have a weaker response to such noise than to vessels.

We conclude that \( \phi_1 \) and \( \phi_2 \) are good filter candidates to enhance vessels, separate vessels from clutter, and withstand the distracting effect of localized, bright features. Hence we define the two-dimensional MFMK \( \text{ker}(x,y; a,b) \) as
\[
K_1(x,y; a,b) = \phi_1(a^{-1}(x-b)), \\
K_2(x,y; a,b) = \phi_2(a^{-1}(x-b)), \\
|y| \leq L/2,
\]

(3)

**Fig. 3.** CL multiwavelet kernels: (a) CL scaling function \( \phi_1 \), (b) CL scaling function \( \phi_2 \), (c) CL wavelet function \( \psi_1 \) and (d) CL wavelet function \( \psi_2 \).
where \((x, y), \phi(x), a, b,\) and \(L\) denote a two-dimensional pixel position, one-dimensional CL-MK scaling functions, the dilation parameter (also known as scaling parameter), the translation parameter, and the length of the two-dimensional MFMK in the \(y\) direction, respectively.

For each position and considered scaling value, we are interested in the maximum of the convolution of the two-dimensional image \((\text{Img}(x, y))\) and two-dimensional MFMK \((\ker(x, y; a, b))\) over all possible orientations. The maximum identifies a vessel candidate at that orientation

\[
M_{\ker}(x, y; a, b) = \max_{\theta} (r_{\theta}(\ker(x, y; a, b)) \ast \text{Img}(x, y)),
\]

where \(r_{\theta}\) rotates the kernel function with an angle \(\theta\) and \(\ast\) represents the convolution operation in variables \((x, y)\).

For vessel detection, we take the MFMK \(\ker(x, y; a, b)\) to be CL-MK \(K_{1}(x, y; a, b)\) and compute

\[
M_{K_{1}}(x, y; a, b) = \max_{\theta} (r_{\theta}(K_{1}(x, y; a, b)) \ast \text{Img}(x, y)),
\]

where \(D_{m}\) indicates the local mean of an image. A discrete (pixel) scheme in computing the local mean may be the convolution of the image and a small filter matrix

\[
D_{m}(\text{Img}) = \text{Img} \ast W,
\]

where \(W\) is a \(w \times w\) filter in which all elements are \(1/w^2\). It seems to us that \(w = 5\) or so is a good choice for all the vessel segmentation we have done so far.

Fig. 4. Responses of the matched filter to a Gaussian cross-section and a general step edge: (a) matched filter, (b) Gaussian cross-section, (c) a general step edge and (d) solid line is the matched filter response to a Gaussian cross-section and dash-dot line is the matched filter response to a general step edge.

For non-vessel detection, we take the MFMK \(\ker(x, y; a, b)\) to be CL-MK \(K_{2}(x, y; a, b)\) and compute

\[
M_{K_{2}}(x, y; a, b) = \max_{\theta} (D_{m}(r_{\theta}(K_{2}(x, y; a, b)) \ast \text{Img}(x, y))),
\]

where \(D_{m}\) indicates the local mean of an image. A discrete (pixel) scheme in computing the local mean may be the convolution of the image and a small filter matrix

\[
D_{m}(\text{Img}) = \text{Img} \ast W,
\]
response with $K_2$ (also $a=6$ and $L=9$), i.e., $\mathcal{M}_{K_2}(x,y; a,b)$ in (6), with the same discrete values of $\theta$. Fig. 7(b) shows that vessels and non-vessel edges are highlighted by the CL-MK $K_1$, and noise is attenuated. Fig. 7(c) shows that the local mean of the response of $K_2$ highlights the clutter edges in the central area. Therefore, we can separate vessel targets from noise using CL-MK $K_1$, and vessels from non-vessel clutter by the local mean of CL-MK $K_2$. 

**Fig. 5.** Sub-figures (a), (c), (e), (g): solid line is the MFMK response with CL-MKs $\phi_1, \psi_1$, and $\psi_2$ to the Gaussian cross-section (depicted in Fig. 4(b)), respectively; dash-dot line is the MFMK response with CL-MKs $\phi_1, \psi_1$, and $\psi_2$, respectively, to the step edge (depicted in Fig. 4(c)). Sub-figures (b), (d), (f), (h): the local mean of responses (a), (c), (e) and (g), respectively.
Fig. 6. MFMK responses with CL-MKs $\phi_1$ and $\phi_2$ to a Gaussian section and a noisy ideal Gaussian impulse: (a) a Gaussian section; (b) noisy, ideal Gaussian impulse; (c) solid line is the MFMK response with CL-MK $\phi_1$ to (a) and dash-dot line is the MFMK response with CL-MK $\phi_1$ to (b); and (d) solid line is the MFMK response with CL-MK $\phi_2$ to the Gaussian cross-section and dash-dot line is the MFMK response with CL-MK $\phi_2$ to the Gaussian impulse.

Fig. 7. The maximum modulus $M_{11}(x,y;a,b)$ and the maximum modulus of local mean $M_{11}^{\text{loc}}(x,y;a,b)$ to image Im0001 from the STARE data set: (a) image Im0001 from the STARE data set, (b) $M_{11}(x,y;a,b)$, and (c) $M_{11}^{\text{loc}}(x,y;a,b)$. We used $a=6$, $L=9$ to generate the kernels.
3.3. Multiscale hierarchical decomposition

Multiscale hierarchical decomposition performs an iterative segmentation at varying image resolutions, locating smaller and smaller vessels. The process is controlled by a single scale parameter, which controls the level of detail included in the vessel map. The full algorithm is summarized in pseudocode in Appendix A.

3.3.1. Normalization of the MFMK response image

The normalized image is defined as in [39]

\[ f = \frac{M_{k-1} - \mu}{\sigma} \quad \text{or} \quad \frac{M_{k-2} - \mu}{\sigma}, \]

where \( \mu \) and \( \sigma \) are the mean and standard deviation of the enhanced MFMK images, \( M_{k}(x,y; a;b) \) or \( M_{k-2}(x,y; a;b) \). So the normalized image \( f \), has zero mean and unit standard deviation, which is used for vessel detection or non-vessel detection in the following multiscale hierarchical decomposition.

3.3.2. Decomposition model

Consider the normalized image \( f(x,y) \) as an intensity function for \( (x,y) \in \Omega \subset \mathbb{R}^2 \) (\( \Omega \) is the domain of the image), which may be written as

\[ f = u_i + v_j, \]

where \( u_i \) is the denoised image and \( v_j \) is the noise part at a noise scale \( \lambda \). Intuitively speaking, the total variation (TV) of a noised image should be larger than that without noise. So minimizing the total variation may result in a denoising process. It is also recognized that the TV minimization is capable of denoising without going too far from the original image. It also relates to the denoising and segmentation approach, finding the minimizer of the total variation may result in a denoising process. It is also related to the denoising and segmentation approach, finding the minimizer of the following functional [see e.g., [41–43]]

\[ J(f, \lambda) = \lambda \|v_j\|_2^2 + \|u_i\|_{BV}, \tag{9} \]

where \( v_j = f - u_i \), the \( L^2(\Omega) \) norm is defined as \( \| \cdot \|_2 = (\int_{\Omega} | \cdot |^2)^{1/2}, \) \( BV \) stands for the homogenous bounded total variation space equipped with the norm of total variation [see [41]] \( \| \cdot \|_{BV} = \| \nabla \cdot \|_2 \) and thus \( \| u_i \|_{BV} = \| \nabla u_i \|_2 = (\int_{\Omega} (u_i)_x^2 + (u_i)_y^2 \, dx \, dy)^{1/2}. \) The first term \( \| v_j \|_2 = \| f - u_i \|_2 \) of (9) is a fidelity term, preventing the process from going too far from the original image. It also relates to the variance of the noise and thus the weight parameter \( \lambda \) in (9) may reflect the scale of noise to be kept in the minimization process. If \( \lambda \) is too small, then only a cartoon representation of \( f \) is kept in \( u_i \). If \( \lambda \) is too large, however, then \( u_i \) remains loaded with too much noise from the image \( f \).

So in this process an image is decomposed into two components: the \( BV \) part, \( u_i \), which captures the main features or structures of interest; and the \( v_j = f - u_i \) part, which captures the remaining elements including noise. What is actually included in each part depends on the scale parameter \( \lambda \) and is therefore scale-dependent. As features interpreted as ‘small’ at a given \( \lambda \) value it may become significant for subsequent values of \( \lambda \), the goal is to identify an optimal sequence \( u_i \) which accumulatively captures as much as possible of the interesting structures (vessels) and simultaneously, as little noise as possible.

A solution to this goal is the multiscale hierarchical decomposition of an image \( f \). It is defined as follows. Given an initial scale parameter \( \lambda_0 \)

\[ f = u_0 + v_0 \quad \text{where} \quad [u_0,v_0] := \arg\min_{u,v} \beta (f, \lambda_0); \]

\[ v_k = u_{k+1} + v_{k+1}, \quad k = 0,1, \ldots, \]

where \( [u_{k+1},v_{k+1}] := \arg\min_{u,v} \beta (f_k, \lambda_0 2^{k+1}). \) \tag{10}

Here we rewrite \( u_{k+1} \) as \( u_k \) and \( v_{k+1} \) as \( v_k \) for convenience. After \( k \) such steps, we end up with a hierarchical decomposition of \( f \)

\[ f = (u_0 + u_1 + \cdots + u_k) + v_k. \tag{11} \]

As \( k \) increases, the ‘interesting’ part \( u = \sum u_k \), where \( u_k \) resolves edges of image \( f \), will contain more and more vessel pixels, meanwhile, at an increased risk of including noise. An optimal stopping criterion will thus be given in Section 3.3.4. Next we show how to minimize the functional \( J(f, \lambda) \), i.e., how to find \( \arg\min_{u,v} \beta (f, \lambda) \) in the process (10).

3.3.3. Functional minimization: Euler–Lagrange equations

The Euler–Lagrange equation of minimizing \( J(f, \lambda) \) defined in (9) is

\[ u_i - \frac{1}{2\lambda} \frac{\partial}{\partial \lambda} \left( \frac{\nabla u_i}{\nabla u_i} \right) = f, \tag{12} \]

usually imposing the Neumann boundary condition

\[ \frac{\partial u_i}{\partial n} = 0 \quad \text{on} \quad \partial \Omega, \tag{13} \]

where \( n \) is the outward normal to the boundary \( \partial \Omega \).

Standard implicit finite difference schemes may be applied for the Euler–Lagrange equation, then the resulting non-linear system can be solved by the fixed-point Jacobi or Gauss–Seidel iterative method with an initial guess, say, \( u_i = 0 \) for the starting point of the decomposition at \( \lambda = \lambda_0 \), or an initial guess \( u_i = u_i^{02} \) for computing the decomposition at \( \lambda = \lambda_0 2^{k+1} \).

We shall use the Gauss–Seidel finite difference scheme suggested in [13,40]

\[ (u_i^{n+1})_{ij} = f_{ij} + \frac{1}{2\lambda h^2} \left[ \frac{(u_i^n)_{ij} - (u_i^n)_{ij+1}}{\sqrt{\epsilon^2 + (D_x(u_i^n))_{ij}^2 + (D_y(u_i^n))_{ij}^2}} - \frac{(u_i^n)_{ij+1} - (u_i^n)_{ij-1}}{\sqrt{\epsilon^2 + (D_x(u_i^n))_{ij+1}^2 + (D_y(u_i^n))_{ij+1}^2}} \right] \]

\[ - \frac{(u_i^n)_{ij} - (u_i^n)_{ij+1}}{\sqrt{\epsilon^2 + (D_x(u_i^n))_{ij}^2 + (D_y(u_i^n))_{ij}^2}} + \frac{(u_i^n)_{ij} - (u_i^n)_{ij-1}}{\sqrt{\epsilon^2 + (D_x(u_i^n))_{ij-1}^2 + (D_y(u_i^n))_{ij-1}^2}} \],

where the domain \( \Omega \) is covered with a discrete grid \( (x_i = ih, y_j = jh), \) \( h \) is the cell size, \( (D_x(u_i^n))_{ij} = (u_i^n(x_{i+1} - u_i^n(x_i)))/h, \) \( (D_y(u_i^n))_{ij} = (u_i^n(y_{j+1} - u_i^n(y_j)))/2h, \) a very small number \( \varepsilon \) is used to avoid singularity when \( \nabla u_i^n = 0 \), and \( D_x(u_i^n)_{ij} \) and \( D_y(u_i^n)_{ij} \) are defined similarly. Letting

\[ C_1 = \frac{1}{\sqrt{\epsilon^2 + (D_x(u_i^n))_{ij}^2 + (D_y(u_i^n))_{ij}^2}}, \]

\[ C_2 = \frac{1}{\sqrt{\epsilon^2 + (D_x(u_i^n))_{ij}^2 + (D_y(u_i^n))_{ij}^2}}, \]

\[ C_3 = \frac{1}{\sqrt{\epsilon^2 + (D_x(u_i^n))_{ij}^2 + (D_y(u_i^n))_{ij}^2}}, \]

and

\[ C_4 = \frac{1}{\sqrt{\epsilon^2 + (D_x(u_i^n))_{ij}^2 + (D_y(u_i^n))_{ij}^2}}. \]
we obtain the computational scheme as
\[
(u_{k+1})_{j}^{n} = \frac{2\lambda k^{2} f_{ij} + C_{1}(u_{k})_{j+1}^{n} + C_{2}(u_{k})_{j}^{n} + C_{3}(u_{k})_{j-1}^{n} + C_{4}(u_{k})_{j-2}^{n}}{2\lambda k^{2} + C_{1} + C_{2} + C_{3} + C_{4}}.
\]  
(15)

We can use the same numerical process to obtain the solutions \(u_{k}\) of the hierarchical multiscale decomposition (10) from the associated Euler–Lagrange equations
\[
(u_{k+1})_{j}^{n} - \frac{1}{2\lambda k^{2}} \text{div} \left( \nabla u_{j+1}^{n} \right) = v_{k}.
\]  
(16)

where \(\lambda_{k+1} = \lambda_{0} 2^{-k+1}\) and \(u_{k+1} = u_{k+1}\) satisfy the boundary condition given in (12). The multiscale decomposition is obtained by iterating the above process with \(u_{0}\) being obtained from (12) (taking \(\lambda = \lambda_{0}\) and \(v_{0} = f - u_{0}\)).

3.3.4. Identifying the optimal decomposition
As mentioned earlier, increasing the number of hierarchical decompositions increases the risk of including noise in the segmented image. In this section, we will present a necessary condition to achieve the optimal number of decompositions.

After the MFMK enhancement, the normalized image \(f\) from Eq. (8) contains enhanced target vessels, enhanced edges from distracting clutter, and noise surviving the attenuation performed by the enhancement. We can think of \(f\) as a sum of two components, \(s\), containing only the interesting elements, and \(n\), containing the uninteresting rest of the image
\[f = s + n.
\]  
(17)

As explained in Section 3.3.3, our goal is to find a hierarchical decomposition \(f = u_{0} + u_{1} + \cdots + u_{k} + n_{k}\) such that \(u = \sum_{k} u_{k}\) is the best possible approximation of \(s\). Then we can also write
\[f = u + n,
\]  
(18)

where \(n = n_{k}\) is associated with \(\lambda = \lambda_{0} 2^{-k}\). From the normalization, we also have \(f = 0\).

Let us first identify that there exist a finite number of decompositions achieving an optimal segmentation in the sense of the signal-to-noise ratio (SNR). To do this, we consider the signal-to-noise ratio of the recovered (denoised) image \(u\)

\[\text{SNR}(u) = 10 \log \frac{V(s)}{V(u - s)} = 10 \log \frac{V(s)}{V(u - n)}
\]  
(19)

where \(\log = \log_{10}\) and \(V(q)\) is the variance of signal \(q\)
\[V(q) = \frac{1}{|\Omega|} \int_{\Omega} (q - \overline{q})^{2}.
\]  
(20)

Here, \(|\Omega|\) is the area of the domain \(\Omega\) of the image \(f\) and \(\overline{q}\) is the mean value of \(q\), i.e., \(\overline{q} = (1/|\Omega|) \int_{\Omega} q\). The initial SNR of the input image, \(\text{SNR}_0 (u = 0, v = f)\), is therefore

\[\text{SNR}_0 = \text{SNR}(0) = 10 \log \frac{V(s)}{V(u - s)} = 0.
\]

As the iteration index \(k\) increases, and \(u\) in (18) is close to the ideal \(s\) in (17), then

\[\text{SNR}(u) = \text{SNR}(s) \rightarrow +\infty.
\]  
(21)

If \(k \rightarrow +\infty\), then

\[\text{SNR}(u) = \text{SNR}(f) = 10 \log \frac{V(s)}{V(n)} \ll +\infty.
\]

So there exists a value \(k = k_{opt}\) such that \(\text{SNR}(u)\) attains its maximum value in the process of the multiscale hierarchical decomposition. Associated with \(k_{opt}\) is the optimal signal-to-noise value, \(\text{SNR}_{opt}\)

\[\text{SNR}_{opt} = \max \text{SNR}(u_{k}).
\]  
(22)

and the decomposition pair, \((u_{opt}, v_{opt})\) that gives \(\text{SNR}_{opt}\). Based on the decomposition process in Eq. (10), we know \(V(v) = 0\) if the image \(f\) contained no noise at all, and \(V(v) = V(f)\) if \(f\) were a purely noisy image.

Next we need a way to compute the optimal value, \(k_{opt}\), for the number of decompositions. We start by recalling the definition of the covariance of two signals
\[\text{cov}(q, r) = \frac{1}{|\Omega|} \int_{\Omega} (q - \overline{q})(r - \overline{r}) \, dx \, dy.
\]  
(23)

and state the following theorem.

**Theorem 1.** For any \((s, n)\) pair consisting of two uncorrelated signals \(s\) and \(n\), a necessary condition for the maximum \(\text{SNR}(u)\) in the range \(V(v) \in (0, V(f))\) using the multiscale hierarchical decomposition model is
\[\frac{\partial \text{SNR}}{\partial V(v)} = 0
\]  
(24)
or
\[\frac{\partial \text{cov}(n, v)}{\partial V(v)} = \frac{1}{2}.
\]  
(25)

**Proof.** The proof of Theorem 1 follows the one in [14] using the definition (19) and by noticing that \(n\) and \(s\) are fixed quantities in a given image \(f\) and \(V(n-v) = V(v) + V(n) - 2 \text{cov}(n, v)\). \(\square\)

So, intuitively speaking, as long as the noise remains limited \((\text{cov}(n, v)/\partial V(v) > \frac{1}{2})\), the hierarchical decomposition is iterated. The process continues until the following condition holds:
\[\frac{\partial \text{cov}(n, v)}{\partial V(v)} \leq \frac{1}{2}.
\]  
(26)

To use the stopping criterion suggested by Theorem 1, we must compute an approximation of the variation of \(\text{cov}(n, v)\) with respect to \(V(v)\). This can be estimated by noticing that \(\text{cov}(n, v)\) and \(V(v)\) both vary with respect to the scaling parameter \(\lambda\). Assume we can devise a treatable model \(\hat{n}\) for the noise \(n\), such that \(\hat{n}\) has similar statistics to \(n\). Further, as indicated in [14], it appears that \(\text{cov}(n, v)\) as a function of \(\lambda\), is almost independent from the underlying image \(f\). We can then use the chain rule to get an approximate formula
\[\frac{\partial \text{cov}(n, v)}{\partial V(v)} = \frac{\partial \text{cov}(n, v)}{\partial \lambda} \frac{\partial \lambda}{\partial V(v)} = \frac{\partial \lambda}{\partial \lambda} \right|_{u = u_{k} + v_{k}} \right|_{u = u_{k} + v_{k}}.
\]  
(27)

The first term on the right-hand side may be precomputed to estimate the behavior of \(\text{cov}(n, v)\) with respect to the scaling parameter \(\lambda\), where the underline image may simply be taken to be the purely estimated noise \(\hat{n}\). The second term can be numerically calculated easily under the framework of the multiscale hierarchical decomposition since, through the decomposition process, we have computed \(v_{k}\) for a series of values of \(\lambda\).

To obtain an approximate sample \(\hat{n}\) of the image noise \(n\), we need to select a patch of the normalized enhanced image containing neither vessels nor clutter. Then we can compute the first term of the right-hand side of (27) and thus the stopping condition (26) so that the multiscale decomposition can stop at the optimal point. In Section 3.5 an example will show that a patch which is not carefully selected may result in a bad result. Based on the experiments which we have carried on a large number of patches, we find that the patch with size 60 × 80 and...
with the lowest average value of all these patches can do a good job in stopping the decomposition process properly (i.e., closer to the optimal decomposition level). The patch with the lowest average value indicates that it contains the fewest vessels and clutter compared to the other patches with the same sizes. Thus such a patch with a fixed size may be automatically selected under this lowest average criterion. The example in Section 3.5 also shows that this automatic selection criterion results in a good vessel segmentation.

3.4. Adaptive thresholding

The final step is to produce a binary mask of the vessels. This requires to threshold the optimal-SNR image produced by the multiscale hierarchical decomposition. A fixed threshold does not provide adequate results. In order to extract an accurate vessel mask, we adopt an adaptive thresholding method based on sequential relaxation [15]. This method makes the threshold vary across the image, resulting in a threshold surface. The main idea is to extract edge points corresponding to zero-crossings [44] and use them as interpolation points of the threshold surface. Using this sequential relaxation algorithm, we can obtain the threshold surface \( \pi(x,y) \) to be used with \( u \).

From the analysis of the normalized MFMK response, we know that the response magnitude at vessel pixels is larger near the centerline than near the vessel edges. Therefore binarization is performed as follows:

\[
\text{Out}(x,y) = \begin{cases} 
1, & \pi(x,y) \leq u(x,y), \\
0, & \text{otherwise}, 
\end{cases}
\]

where Out stands for the finally segmented binary mask of the vessel image. For the purpose of completeness, the algorithm given in [15] is summarized in Appendix B.

3.5. An example

Illustrative segmentation results are shown in Fig. 8 for image Im0162 from the STARE data set. We have used \( a = 6, L = 9 \) in the

![Fig. 8. Segmentation results produced by multiscale hierarchical decomposition with \( \lambda_0 = 0.01 \) and \( \lambda = 2^l \lambda_0 \): (a) Im0162 image from the STARE data set, (b) Segmentation image \( u \) at scaling parameter \( \lambda_0 \), (c) Segmentation image \( u \) at scaling parameter \( \lambda_1 \), (d) Segmentation image \( u \) at scaling parameter \( \lambda_2 \), (e) Segmentation image \( u \) at scaling parameter \( \lambda_3 \), (f) Segmentation image \( u \) at scaling parameter \( \lambda_4 \), (g) Segmentation image \( u \) at scaling parameter \( \lambda_5 \), (h) Segmentation image \( u \) at scaling parameter \( \lambda_6 \), (i) Segmentation image \( u \) at scaling parameter \( \lambda_7 \), (j) Segmentation image \( u \) at scaling parameter \( \lambda_8 \), (k) Segmentation image \( u \) at scaling parameter \( \lambda_9 \).]
CL-MK $K_1$ as before and the initial scaling parameter $l_0 = 0.01$ in the multiscale hierarchical decomposition. For this image, the noise patch with size 60 x 80 and with the lowest average value is selected to compute the stopping condition using (27) and (26). The optimal segmentation (maximal SNR) is achieved for $l_6$ (see Fig. 8(h)), and $\hat{\text{cov}}(n, v)/\hat{\Sigma}V(v) < 1/2$ when $l = l_7$. Meanwhile, if we choose a noise patch with size 60 x 80 and with the highest average value, the decomposition is stopped at $l_8$, which contains a lot of noise (see Fig. 8(j)).

3.6. Post-processing: clutter reduction

To reduce the interference of large bright areas like the optic disc (OD) or bright lesions, which can generate false positives, we need a rough sketch of the bright area applying binarization (using a threshold of 4/5 of the highest brightness value, determined by observations over the whole image set) and a convex-hull algorithm [44] easily implemented in MATLAB. We note that this algorithm is not designed as a lesion detector, but only to limit the effect region of clutter. The OD region could of course be discarded by running a specialized OD detector [45,46]. When the effect region of the clutter is identified, non-vessel edges in the region are then detected using the CL-MK $K_2$ or $\phi_2$ and finally removed from the segmented retinal image.

We use image Im0001 in the STARE data set as an example. Fig. 9(a) and (b) shows the rough regions of the optic disc, lesions or exudates by global binarization and convex hull methods, respectively. The latter outlines the region of optic disc, lesions or exudates. Fig. 9(c)–(e) shows three segmentation results taking $a = 6$, $L = 9$ CL-MK in $K_2$ and $l_0 = 0.01$ in the multiscale hierarchical decomposition. Fig. 9(f) shows the final non-vessel edges with the help of the region outlined in Fig. 9(b). These edges will be excluded in the final segmented blood vessel image (see results in Section 4.2).

4. Results

4.1. Materials and procedures

The standard paradigm to validate medical image processing algorithms is to compare their outputs with ground truth, or gold standard, generated by several human experts. In case of vasculature detection, this requires experts to generate full binary maps of the vasculature in each image. It is therefore extremely time-consuming to generate accurate and extensive sets of ground-truth masks. Only a few sets are available publicly, limiting the power of comparative performance assessment. In contrast, other image processing areas have developed extensive, entirely public data sets, e.g., PASCAL for object recognition and classification (http://pascalvoc.dsi.unimi.it/).

To enable comparative assessment, we use images and associated manual segmentations from two public data sets available on the web, DRIVE [28] and STARE [18]. Both DRIVE and STARE include ground truth segmentations for their images. The field of view (FOV) is approximately 50° for images of both sets. The DRIVE data set consists of 40 images (768 x 584, eight bits per channel). The images are in compressed JPEG format and have been divided into a training and a test set, each containing 20 images. They were manually segmented by three observers trained by an ophthalmologist.
From STARE we used 20 images (700 × 605 pixels, eight bits per channel). The FOV in the images is approximately 650 × 550 pixels in diameter. Ten of the images contain pathology. Two observers segmented all images. The segmentations of the two observers are fairly different, in that the second observer segmented much more of the thinner vessels than the first one; we evaluate performance against the more demanding segmentations by the first observer. As stated before, we work in the green channel.

The algorithm has been implemented in MATLAB on an IBM running WinXP on an Intel(R) Core(TM)2 Duo CPU (3.16 GHz) with 2 GB memory. On this platform, it took about 3.5 min to process one image with seven decompositions. Considering that these results were obtained with interpreted MATLAB on a standard PC, the method is clearly an interesting candidate for incorporation into assisted-diagnosis systems expected to supply an answer with a limited time.

**Fig. 10.** The segmentation process of image Im0001 in the STARE data set using our proposed method: (a) Im0001 in the STARE data set; (b) the maximum modulus $M(a, b)$; (c) the maximum modulus of local mean $\overline{M}(a, b)$; (d) the segmentation results of Fig. 10(b) using our proposed method; and (e) the segmentation results of Fig. 10(c) using our proposed method; (f) the final blood vessel segmentation result after we remove the non-vessel edge from Fig. 10(d) according to non-vessel information in Fig. 10(e).
In the remainder of this section we report first qualitative results aimed to give a visual feeling of the quality of the vessel maps generated by our method. We then report comparative, quantitative results using the standard ROC method. We compare our results with those reported by Soares et al. [21], Niemeijer et al. [27], Staal et al. [28], and Jiang and Mojon [34].

4.2. Qualitative results

Fig. 10(a) shows image Im0001 in the STARE data set. This image presents bright lesions or exudates, large and small, cluttering a wide area. Fig. 10(b) shows the MFMK response with CL-MK \( K_1 \), i.e., \( M_{\text{CL-MK}_1}(a,b) \), where several non-vessel responses are visible. With the help of local mean of the MFMK response with CL-MK \( K_2 \) (i.e., \( M_{\text{CL-MK}_2}(a,b) \)) to the image, we can obtain a good localization of the non-vessel responses, shown in Fig. 10(c). Our segmentation result is shown in Fig. 10(d) \( (\lambda_0 = 0.01) \). Fig. 10(e) displays the segmentation results of lesions or exudates in the image (same as Fig. 9(f)). Subtracting Fig. 10(e) from Fig. 10(d), we obtain Fig. 10(f) which gives the final blood vessel segmentation after removing non-vessel responses.

Fig. 11 shows another example using image 01 from the DRIVE data set.

4.3. Quantitative results

Figs. 12 and 13 show segmentation of our method on STARE images and compare them with those obtained by Hoover et al. [18] and ground truth segmentations.

The performance of the binary classifier described earlier is measured using ROC curves (see [47]) of the segmentation in comparison with the first-observer-based standard from DRIVE and STARE databases. ROC curves are represented by plotting true positive fractions versus false positive fractions as the discriminating parameter related to \( \pi(x,y) \) defined in (28) is varied. As usual, the true positive fraction (TPF), also called sensitivity, is determined by dividing the number of pixels correctly classified as vessel pixels \( (TP) \) by the total number of vessel pixels in the gold standard segmentation

\[
\text{TPF} = \frac{TP}{P} = \frac{TP}{TP + FN},
\]

where \( FN \) is the number of pixels incorrectly classified as non-vessel pixels. The false positive fraction (FPF) is the number of pixels incorrectly classified as vessel pixels \( (FP) \) divided by the total...
number of non-vessel pixels in the gold standard

\[
FPF = \frac{FP}{N} = \frac{FP}{TP + TN}
\]

Here, \(TN\) is the number of pixels correctly classified as non-vessel pixels. The axes of the plot are rescaled, so the true positives and false positives vary between 0 and 1. The area under the ROC curve (\(Az\)) measures discrimination, i.e., the ability of the classifier to correctly distinguish between vessel and non-vessel pixels. An area of 1 indicates a perfect classification. The accuracy (\(ACC\)) for one image is the fraction of pixels correctly classified at a specified threshold

\[
ACC = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + FN + FP + TN}
\]

Fig. 14 shows the ROC curves for the DRIVE and STARE databases. We compute the \(ACC\) for each parameter value and the optimal value is the one which gives the maximal \(ACC\) and is marked in ROC curves.

Table 1 shows the performance of our method on both DRIVE and STARE data set using different methods, which have been ordered by types of methods in the introduction chapter for segmenting the retinal vasculature. On the 10 STARE images with lesions, our proposed method is the best in terms of \(Az\) and \(ACC\). Results with DRIVE images indicate that our method performs very well, but its \(Az\) and \(ACC\) fall slightly behind that of Soares et al. [21]. The performances shown here are those reported on www.isi.uu.nl/Research/Databases/DRIVE/.

4.4. Discussion

We offer here a few comments on our experimental results, including directions for future developments.

First, our algorithm is efficient. With interpreted in MATLAB, the average processing time with STARE and DRIVE images was 3.5 min to process one image with about seven decomposition levels.

Second, results could be arguably improved using a more accurate noise estimation procedure (Section 3.3) or a better size of the patch. For current testing databases of images our algorithm leads to very good results.

Third, like every vessel segmentation scheme based on pixel classification, the detected vessels can be fragmented. In our results, fragmentation is limited and large vessels are the least affected. Hence the method seems suitable, among others, for correlational studies like the ones reported in [12,11], considering morphometric measurements on major vessels only, or local width measurement.

5. Conclusion

We have presented a novel, general method for segmenting the retinal vasculature in fundus camera images. Our method does not require training and generates very competitive results, both visually and in terms of ROC performance, compared with quite a few algorithms reported in the literature. On the basis of our experimental assessment, we believe this paper brings three main contributions.

First, a novel vessel enhancement technique based on the matched filters with multiwavelet kernels, which copes well with substantially different cross-sectional intensity profiles (with and without central reflections). The wavelet functions are shown to respond to vessels and clutter respectively, thus allowing us to identify and eliminate a good percentage of erroneous candidates.

Second, an iterative multiscale hierarchical decomposition, based on a two-class model of the normalized enhanced image. The process is controlled by a scale parameter; we showed a necessary condition to achieve the optimal decomposition and derived the associated value of the scale parameter.
Third, the system relies on adaptive thresholding, so that no thresholds are tuned manually. It may be difficult to avoid the manual intervention for the selection of a patch for noise identification (in order to estimate the stopping time of the multiscale hierarchical decomposition). As a first attempt in this regard, we present an automatic selection method by looking at the lowest average of patches as long as the size of the patch is provided.

The vessel maps generated by the system seem perfectly usable in any application not requiring perfect vessel connectivity, for instance studies investigating the correlation of biomarkers associated with the retinal vasculature and various systemic conditions or cognitive performance. The retinal measurements used in these studies are, typically, caliber of major vessels within a limited distance from the optic disc, branching angles at junctions in selected areas, tortuosity, and less frequent fractal dimension. All of these can be estimated from maps generated by our system.

The proposed method falls in the category of pixel classifiers (vessel or non-vessels). An accurate connectivity map can not of these can be estimated from maps generated by our system.

The vessel maps generated by the system seem perfectly usable in any application not requiring perfect vessel connectivity, for instance studies investigating the correlation of biomarkers associated with the retinal vasculature and various systemic conditions or cognitive performance. The retinal measurements used in these studies are, typically, caliber of major vessels within a limited distance from the optic disc, branching angles at junctions in selected areas, tortuosity, and less frequent fractal dimension. All of these can be estimated from maps generated by our system.

Finally, we plan to incorporate the vessel detection reported here into VAMPIRE (Vasculature Assessment and Measurement Platform for Images of the Retina) [48], our software application designed to collect efficiently large numbers of measurements of the retinal vasculature for correlational studies.

Conflict of interest statement

None declared.

Acknowledgments

YW was funded by the China Scholarship Council (CSC) and was partially supported by the National High-Tech R&D Program (863 Program 2012AA10A412). PL is partially supported by the Leverhulme research fellowship No. RF/9/REG/2009/0507 and University of Science and Technology Beijing research grant (No. 06108038). ET is partially supported by the Scottish Northern Research Partnership. We thank Adrià Perez Rovira and the VAMPIRE group for their suggestions and help.

Appendix A. Multi-wavelet multiscale segmentation algorithm

1. Parameters: \( \lambda_0 \), \( N_p \) (the max loop number), \( sp \) (size of the normalized MFMK response image), and \( sn \) (size of the similar noise patch).

2. Initialize \( k = 0 \), \( \nu_{k-1} = f \leftarrow \text{normalized MFMK response image} \), \( u_{k-1} = 0 \), \( \lambda_{k-1} = 0 \), \( u = 0 \) (where size of \( u, u_k, \nu_k \) are \( sp \))

   (a) \( \lambda_k \leftarrow \lambda_0 2^{k} \).
   (b) Compute \( u_k \) by (15) with \( u_{k-1}, \lambda_{k-1} \).
   (c) \( \nu_k = f - u_{k-1} + u_{k} \).
   (d) \( \text{DEcov}^{\nu_k} \leftarrow (\lambda_k - \lambda_{k-1})/(V(\nu_k) - V(\nu_{k-1})) \) (where \( \text{DEcov}^{\nu_k} \) is the recomputing a discrete estimation of \( \partial \text{cov}(\hat{h}, \nu)/\partial \lambda \) in vector \( \text{DEcov}^{\nu_k} \).
   (e) Exit loop if \( \text{DEcov}^{\nu_k} < \frac{1}{2} \) or \( k = N_p \).

3. Return \( u \)

Precomputing a discrete estimation of \( \partial \text{cov}(\hat{h}, \nu)/\partial \lambda \).

1. Initialize \( f \leftarrow \text{noise patch with } sn \) (where we let \( sn \) be \( 80 \times 60 \)).

   \( k = 0 \), \( \nu_{k-1} = f, u_{k-1} = 0, \lambda_{k-1} = 0, u = 0 \) (where size of \( u, u_k, \nu_k \) are \( sn \)), \( \text{Ecov}_{k-1} = 0 \).

2. Loop \( (k = 0; k++; k < N_p) \)

   (a) \( \lambda_k \leftarrow \lambda_0 2^{k} \).
   (b) Compute \( u_k \) by (11) with \( u_{k-1}, \lambda_{k-1} \).
   (c) \( \nu_k = f - u_{k-1} + u_k \).
   (d) \( \text{Ecov}_k \leftarrow < \nu_k, \hat{n} > \) by (23).
   (e) \( \text{DEcov}^{\nu_k} \leftarrow (\text{Ecov}_k - \text{Ecov}_{k-1})/(\lambda_k - \lambda_{k-1}) \).
   (f) \( f \leftarrow \nu_k \).

3. Return vector

\( \text{DEcov}^{\nu_k} = [\text{DEcov}^{\nu_1}, \ldots, \text{DEcov}^{\nu_k}] \).

Appendix B. Sequential relaxation algorithm (SRA)

The key idea of the algorithm is to minimize the following energy functional:

\[
\min \frac{\partial u}{\partial t} = \alpha \int_{\Omega} |(\Delta \phi \ast \nu)|^2 \, dx \, dy + \beta \int_{\Omega} \mathcal{F}(\nabla (\nu)) \, dx \, dy \tag{29}
\]

where \( \phi \) is a symmetric smooth kernel to reduce noise and avoid breaking the boundaries of objects; \( \Delta \phi \) (the Laplacian of \( \phi \)) is referred to as the LOG filter in [44], \( \mathcal{F}(\nabla \nu) = (\nabla \nu)^T(1 + |\nabla \nu|^2) \), which is an increasing positive function, \( \alpha \) and \( \beta \) are positive parameters to balance the two terms on the right. The associated PDE is

\[
\frac{\partial u}{\partial t} = -\alpha \Delta \phi (\Delta \phi \ast \nu) + \beta \text{div} \left( \mathcal{F}(\nabla \nu) \frac{\nabla \nu}{|\nabla \nu|} \right). \tag{30}
\]

Numerical procedure: Let \( \tau = 0, 1, 2, \ldots \), be non-negative integers, and \( \Delta t \) the step size.

1. Set \( \tau = 0 \) and assign the initial values for \( \hat{u} \) and \( \hat{v} \):

\[
\hat{u}^0 = \hat{v}^0 = u. \tag{31}
\]

2. For \( k = 0, 1, 2, \ldots \), complete the evolution of \( \hat{v} \) as follows:

\[
\hat{v}^{k+1} = \hat{v}^k + \Delta t \left( \beta \frac{\Delta \phi^k}{(1 + |\nabla \hat{v}^k|^2)^{3/2}} - 2 \cdot \Delta \phi (\phi \ast \hat{v}^k) \right). \tag{32}
\]

3. Set \( \tau = \tau + 1 \), complete the evolution of \( \hat{u} \) by \( \hat{u}^\tau = \lim_{k \to N_k} \hat{u}^k \) (where \( N_k \) is a big positive integer), and re-initialize \( \hat{v} \).

4. Go back to step 2 until the value of \( |\hat{u}^{\tau + 1} - \hat{u}^\tau|^2 / |\hat{u}^\tau|^2 \) less than a given value.

References


Emanuele Trucco received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Genoa, Genoa, Italy, in 1984 and 1990, respectively. He is currently the Northern Partnership Professor of Computational Vision in the School of Computing, University of Dundee, Dundee, UK, and a Honorary Clinical Researcher at the Ninewells Clinical Hospital, Dundee. He was with the Joint Research Centre of the European Union, Ispra, and the University of Edinburgh and Heriot Watt University, Edinburgh, UK. He was an Editor of Pattern Analysis and Applications. He has authored or coauthored more than 170 refereed papers on computer vision and its applications in manufacturing, robotics, communications and bioengineering, and two books, one of which (with A. Verri) has been a standard for computer vision (1457 citations Google Scholar). His current research interests include articulated human motion tracking and medical (especially retinal) image analysis. Dr. Trucco was an Editor-in-Chief of the IEEE Proceedings on Signal, Speech and Image Processing, and an Editor of IEEE Transactions on Systems, Man and Cybernetics C. He is a Regular Invited Member of the scientific committees of all the major conferences on computer vision. He has been a Co-Organizer of the British Machine Vision Conference in 1996, 2006, and 2011.