

The following pages of this document constitute a draft version of a chapter and do not represent the final version that will be published. This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder.

© 2007 Taylor and Francis Group, LLC. Personal use of this material is permitted. Permission from Taylor and Francis Group, LLC, must be obtained for all other uses, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

---

## *Contents*

<b>1 Segmentation of Retinal Vasculature Using Wavelets and Supervised Classification: Theory and Implementation</b>	<b>1</b>
<i>João V. B. Soares and Roberto M. Cesar Jr. University of São Paulo, Brazil</i>	
1.1 Introduction . . . . .	1
1.2 Theoretical Background . . . . .	4
1.2.1 The 1-D CWT . . . . .	4
1.2.2 The 2-D CWT . . . . .	5
1.2.3 The 2-D Gabor Wavelet . . . . .	8
1.2.4 Supervised Classification . . . . .	11
1.2.5 Bayesian Decision Theory . . . . .	11
1.2.6 Bayesian Gaussian Mixture Model Classifier . . . . .	12
1.2.7 $k$ -Nearest Neighbor Classifier . . . . .	13
1.2.8 Linear Minimum Squared Error Classifier . . . . .	14
1.3 Segmentation Using the 2-D Gabor Wavelet and Supervised Classification . . . . .	15
1.3.1 Pre-processing . . . . .	17
1.3.2 2-D Gabor Wavelet Features . . . . .	17
1.3.3 Feature Normalization . . . . .	18
1.3.4 Supervised Pixel Classification . . . . .	20
1.3.5 Public Image Databases . . . . .	20
1.3.6 Experiments and Settings . . . . .	21
1.3.7 ROC analysis . . . . .	23
1.4 Implementation and Graphical User Interface . . . . .	25
1.4.1 Overview . . . . .	25
1.4.2 Installation . . . . .	27
1.4.3 Command Line Interface . . . . .	27
1.4.4 Graphical User Interface . . . . .	29
1.5 Experimental Results . . . . .	31
1.6 Conclusion . . . . .	39
1.6.1 Summary . . . . .	39
1.6.2 Future Work . . . . .	39



# 1

---

## *Segmentation of Retinal Vasculature Using Wavelets and Supervised Classification: Theory and Implementation*

**João V. B. Soares and Roberto M. Cesar Jr.**

*University of São Paulo, Brazil*

### CONTENTS

1.1	Introduction .....	1
1.2	Theoretical Background .....	4
1.3	Segmentation Using the 2-D Gabor Wavelet and Supervised Classification .	15
1.4	Implementation and Graphical User Interface .....	25
1.5	Experimental Results .....	31
1.6	Conclusion .....	38
	Acknowledgments .....	41

---

### 1.1 Introduction

Inspection of the optic fundus vasculature can reveal signs of hypertension, diabetes, arteriosclerosis, cardiovascular disease, and stroke [41]. Retinal vessel segmentation is a primary step towards automated analysis of the retina for anomaly detection and also image registration. Automated assessment of the retinal vasculature morphology can be used in a screening tool for early detection of diabetic retinopathy, while retinal image registration is of interest in change detection, mosaic synthesis and real-time tracking and spatial referencing for assistance in laser surgeries.

Different techniques are used for acquiring retinal images. Most common are colored or monochromatic photography and angiography using fluorescent dyes. In monochromatic photography, color filters are used to select light wavelengths that enhance the visibility of various fundus structures. Lighting using wavelengths close to the green region of the spectrum (known as red-free lighting) is frequently employed, as it leaves vessels, hemorrhages and exudates more apparent. Angiographies, on the other hand, require the injection of a small amount of fluorescent dye into the patient, usually sodium fluorescein or indocyanine green. Fluorescence angiography permits recording of blood vessels and flow and also the detection of eventual leakages, of interest for diagnostic purposes. However, it is inadequate

for screening programs, as angiograms can only be obtained by specialists in ophthalmology clinics, and is slightly invasive, presenting a certain risk of side effects to the patient.

In face of the different forms of acquisition, it is desirable that segmentation methods be able to work on images from different modalities. As an additional difficulty, even images of the same modality present large variability depending on the patient, presence of pathologies, camera model, illumination, and focus adjustment. Images with strong illumination variation or deficiency, light reflexes from the cornea and inadequate focus are common and, depending on the application at hand, must be treated appropriately, posing an additional challenge in the development of methods [27].

The retinal vasculature is comprised of two complex networks – one of veins, the other of arteries – that spread out from the optic disk and branch successively to occupy different regions of the fundus. Retinal blood vessels are locally continuous with respect to position, curvature and width, with vessels widths gradually decreasing with the distance to the optic disk [8, 12, 78]. These properties are specially important during the design of tracking algorithms [8, 29, 70]. Additionally, vessels are defined by a pair of parallel borders in which the image derivative presents opposite signs [8, 43, 62], in such a manner that the shapes of vessel cross-sections can be locally approximated by Gaussian functions [77]. These properties provide for vessel models that are usually as valid for retinal photographs as for angiograms, with vessels appearing as lighter than the background in angiograms and darker in colored and red-free photographs.

Diabetic retinopathy screening involves assessment of the optic fundus with attention to a series of indicative features. Of great importance is the detection of changes in blood vessel structure and flow, due to either vessel narrowing, complete occlusions or neovascularization [9, 57, 74]. Neovascularization is a condition associated with proliferative diabetic retinopathy (PDR), an advanced stage of the disease in which new vessels are formed emerging from the area of the optic disk or from peripheral vessels [41]. Prior to neovascularization, it is common to note vessel narrowing/dilation and increase in tortuosity. These changes can be detected by morphological analysis techniques through quantitative measures, such as vessel width and length [61], tortuosity [36], and fractal dimension [39].

Another major application of retinal vessel segmentation is registration of images captured at various instants of time or under different angles. Branching and crossing points or the skeletonized vasculature may be used as spatial landmarks that provide for registration of even images from different modalities [75, 76]. Temporal registration assists in accompanying and detecting diseases and assessing effects of treatment. It may also be used for analysis of blood flow in angiographic video sequences, as the fluorescent dye reaches the optic fundus and retinal vessels. Partial views of the retina taken at different angles can also be combined to synthesize retinal mosaics. The mosaics can then be used for taking accurate morphological measurements or planning and assistance during laser surgeries, through real-time tracking and spatial referencing [27]. Finally, registration of images from different modalities allows visualizing superpositions of clinically important elements that

previously appeared only separately.

The retinal and choroidal vessel structures are used in biometrics for identification/verification of persons in security systems. Among many possible biometric features used, the optic fundus vascular patterns possess possibly the lowest error rate, besides being very stable through time and practically impossible to counterfeit. However, its use has limited reach due to high cost and a slight discomfort to the user subject during image acquisition. These disadvantages could possibly be surpassed, but with the variety of available biometrics, it is hard to predict the future of this particular application [37].

An automated assessment for pathologies of the optic fundus initially requires the precise segmentation of the vessels from the background, so that suitable feature extraction and processing may be performed. Several methods have been developed for vessel segmentation, but visual inspection and evaluation by receiver operating characteristic (ROC) analysis have shown that there is still room for improvement: human observers are significantly more accurate than the methods, which show flaws around the optic disk and in detection of the thinnest vessels [15, 60]. In addition, it is important to have segmentation algorithms that are fast and do not critically depend on configuring several parameters, so that untrained community health workers may utilize this technology. This motivates the use of the supervised classification framework, which only depends on manually segmented images and can be implemented efficiently.

The method and some of the results presented in this chapter were previously published in a journal paper [66]. Here, the method is presented in greater detail, including a denser theoretical review, new tests and analysis of results. An open-source prototype of an interactive software implementation is also described, including details on the most important steps of the segmentation process. The method developed uses the 2-D continuous wavelet transform (CWT) coupled with supervised pixel classification in classes *vessel* and *nonvessel*. Wavelet detection works similarly to matched filters, while at the same time incorporating the scale-angle representation, especially suited for the detection of blood vessels. A significant problem discussed in some works is the choice of an adequate scale for local detection filters. Filters of different scales have been combined empirically [53], while many other algorithms work off only one scale [10, 38, 77]. The approach presented here combines wavelet responses from different scales in feature vectors, allowing for pixel classification based on the statistical training of classifiers.

An essential property of many methods is their foundation on sets of rules that deal with specific situations, leading to complex algorithms that may become parameter-dependent (for example, [38, 40, 58, 77]). Even with a large number of rules, situations such as central vessel reflexes and branching/crossing points usually fail to be dealt with satisfactorily. In turn, the pixel classification approach is conceptually simple and allows the classifier to handle more specific situations, avoiding the formulation of rules and need for parameter adjustment. The approach takes into account only information local to each pixel, leaving space for improvement through the addition of a global inference phase. Even so, it is capable of segmenting the complete vascular networks and does not depend on user interaction, but on avail-

able manual segmentations for training.

Algorithm efficiency must be taken into account in practical and real-time applications, leading to specific considerations during method design [8]. Feature generation and classification of all image pixels may turn out to be a slow process depending on the choice of classifier [68]. In the method described in this chapter, a Bayesian classifier using Gaussian mixture models for class likelihoods was evaluated. The classifier showed good results with respect to ROC analysis, as it allows for complex decision surfaces, while at the same time providing a fast classification phase.

This chapter begins reviewing the theoretical foundations of the segmentation method (Section 1.2), including properties of the 2-D CWT and the 2-D Gabor wavelet, as well as the supervised classifiers evaluated. Section 1.3 describes the approach adopted, that combines wavelet features generation with supervised pixel classification. The same section also lays out the experimental evaluation performed, covering the description of image databases and a brief review of ROC analysis. An overview of the open-source implementation and graphical user interface developed for tests is given in Section 1.4. Experimental results are presented in Section 1.5, followed by the conclusion in Section 1.6, containing a discussion and directions for future work.

---

## 1.2 Theoretical Background

### 1.2.1 The 1-D CWT

Signal analysis may be carried out using transforms capable of emphasizing a signal's relevant properties, allowing processing tasks such as parameter estimation, noise filtering and feature detection to be performed on the transformed signal. The wavelet transform is an example of such a signal processing tool, that originated based on the Fourier transform, and is, as the latter, linear and invertible. Fourier analysis extracts global frequencies from a signal and is therefore appropriate for signals whose properties do not evolve with time (stationary signals). However, it is incapable of locating or analyzing frequencies of limited time duration (present in nonstationary signals). One of the first attempts to overcome this problem was the *short-time Fourier transform* (STFT), also known as Gabor transform [28, 64], in which a sliding window function delimits the signal portion to be analyzed by the Fourier transform. Notwithstanding, the STFT is restricted, as a constant resolution is used for analyzing all frequencies, given by the window's size.

The wavelet transform varies the size of the limiting window when analyzing different frequencies, giving rise to a multiresolution analysis. As observed by Gabor [28], the product between the time and frequency resolutions is lower bounded. This implies that there is a trade-off between time and frequency resolution: increasing time precision leads to decreasing frequency precision and vice-versa. The wavelet transform is defined as to maintain a constant relative bandwidth during anal-

ysis. This means that low frequencies are observed through large windows (i.e. low time precision, high frequency precision), whereas high frequencies are observed through small windows (i.e. high time precision, low frequency precision). This type of frequency varying behavior is encountered in a great variety of signals, thus explaining the success of the wavelet transform. In fact, the development of the CWT theory was motivated by the problem of analyzing microseismic signals derived from oil prospection, which present this type of behavior [31, 33]. Accordingly, the wavelet transform is specially suitable for detecting singularities and analyzing instantaneous frequencies [32, 49].

The CWT is defined as

$$U(b, a) = |a|^{-1/2} \int \psi^*(a^{-1}(x - b)) f(x) dx, \quad (1.1)$$

where  $a \in \mathbb{R}$  is the scale parameter, which relates the analyzing window size to the frequency;  $b \in \mathbb{R}$  is the central time position being analyzed;  $\psi$  is the *analyzing wavelet* (also called *mother wavelet*); and  $\psi^*$  denotes the complex conjugate of  $\psi$ . The transform may be seen as a linear decomposition of the signal as projections onto basis functions  $\psi_{b,a}(x) = \psi(a^{-1}(x - b))$ , each one associated to a given location in time  $b$  and scale  $a$ . The domain of the transform  $U(b, a)$  is referred to as the *time-scale plane*.

There are various analyzing wavelets that may be used depending on the kind of information to be extracted from the signal. The Morlet wavelet is defined as a Gaussian modulated by a complex exponential of frequency  $k_0$ ,

$$\psi(x) = g(x) \exp(-ik_0x) + \text{correction}, \quad (1.2)$$

where  $g(x)$  is a Gaussian function,  $i = \sqrt{-1}$  and the correction term is included to guarantee the wavelet admissibility condition (zero mean). The wavelet's definition shows a strong relation to the STFT, also minimizing the time-frequency joint uncertainty [64].

Wavelet theory represents a unified view of various signal processing techniques developed independently in fields such as pure mathematics, physics and engineering. Being a versatile tool, it is applied in a wide range of applications, such as problems arising in physics [16]; audio, image and video coding [64]; fractal and shape analysis [14]; and biomedical applications [71] (e.g. to detect brain activity in fMRI [72]).

### 1.2.2 The 2-D CWT

The CWT may be extended to higher dimensional spaces, allowing, for instance, the analysis of images (2-D signals) and video sequences (time-varying 2-D signals). This extension of the CWT preserves the fundamental properties of the 1-D case, being founded on coherent states and group theory. The theoretical framework allows the definition of operations over the wavelet space that preserve notions of symmetry. For the 1-D case, the natural operations are translation and dilation, being extended to the 2-D case as translation, dilation and rotation [1, 59].

The CWT is mostly used for signal analysis and feature detection, in contrast to the *discrete wavelet transform* (DWT), preferable for data synthesis and compression, yielding very efficient algorithms. The 2-D DWT provides orthogonal and biorthogonal bases, thus allowing efficient signal representation, while the CWT necessarily leads to a redundant representation. However, the 2-D DWT is formulated as a tensor product of two 1-D schemes. It is therefore restricted to the Cartesian geometry, where directional analysis is much more difficult. In opposition, the introduction of the rotation operation in the 2-D CWT is specially suited for the analysis of oriented image structures by using the so called directional wavelets. This fact explains the adoption of the 2-D CWT with the Gabor wavelet to analyze the blood vessels, as proposed in Section 1.2.3.

The 2-D CWT has been used in a variety of applications in computer vision and image processing, such as analysis of medical and astronomical images [24], texture identification [51], edge detection [49, 50, 52] and fractal analysis [4, 5]. There are also imaging applications in physical problems such as analysis of geological faults and turbulence in fluids [21] or detection of dilation and rotation symmetries [3]. The main 2-D CWT concepts and properties used in this chapter are presented in the following.

The real plane  $\mathbb{R} \times \mathbb{R}$  is denoted as  $\mathbb{R}^2$  and 2-D vectors are represented as bold letters, e.g.  $\mathbf{x}, \mathbf{b}, \mathbf{k} \in \mathbb{R}^2$ . Images are taken as finite energy functions (i.e. square integrable)  $f \in L^2(\mathbb{R}^2)$ . The vector  $\mathbf{x}$  is used to represent a spatial position in the image, while  $\mathbf{k}$  is associated to a given spatial frequency. The spatial frequency domain is defined by the 2-D Fourier transform  $\hat{f}$  of a given image  $f$  as

$$\hat{f}(\mathbf{k}) = (2\pi)^{-1} \int \exp(-i\mathbf{k}\mathbf{x}) f(\mathbf{x}) d^2\mathbf{x}. \quad (1.3)$$

An analyzing wavelet  $\psi \in L^2(\mathbb{R}^2)$  may take on complex values and must satisfy the admissibility condition, described below. A family of wavelets  $\{\psi_{\mathbf{b},a,\theta}\}$  of same shape is defined by translations, dilations and rotations (by  $\mathbf{b}$ ,  $a$  and  $\theta$ , respectively) of the analyzing wavelet:

$$\psi_{\mathbf{b},a,\theta}(\mathbf{x}) = a^{-1} \psi(a^{-1} r_{-\theta}(\mathbf{x} - \mathbf{b})), \quad (1.4)$$

$$\hat{\psi}_{\mathbf{b},a,\theta}(\mathbf{k}) = a \exp(-i\mathbf{b}\mathbf{k}) \hat{\psi}(ar_{-\theta}(\mathbf{k})), \quad (1.5)$$

where  $a > 0$  and  $r_\theta$  denotes the usual 2-D rotation,

$$r_\theta(\mathbf{x}) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta), \quad 0 \leq \theta < 2\pi. \quad (1.6)$$

The translation, dilation and rotation operations generate the 2-D Euclidean group with dilations acting over  $L^2(\mathbb{R}^2)$ , denoted  $G$  [1].

The CWT  $T_\psi \in L^2(G)$  is defined as the scalar product between  $f$  and each wavelet  $\psi_{\mathbf{b},a,\theta}$ , as a function of  $(\mathbf{b}, a, \theta) \in G$ :

$$T_\psi(\mathbf{b}, a, \theta) = \langle \psi_{\mathbf{b}, a, \theta} | f \rangle \quad (1.7)$$

$$= a^{-1} \int \psi^*(a^{-1}r_{-\theta}(\mathbf{x} - \mathbf{b})) f(\mathbf{x}) d^2\mathbf{x}, \quad (1.8)$$

$$= a \int \exp(i\mathbf{b}\mathbf{k}) \hat{\psi}^*(ar_{-\theta}(\mathbf{k})) \hat{f}(\mathbf{k}) d^2\mathbf{k}. \quad (1.9)$$

The transform is thus a local representation in each of the four dimensions of  $G$ , capable of revealing dilation and rotation symmetries [3, 24]. The wavelet transform may be viewed as a series of correlation operations, presenting stronger responses whenever the transformed wavelet  $\psi_{\mathbf{b}, a, \theta}$  matches the signal portion.

If  $\psi$  and  $\hat{\psi}$  are well localized, the transform acts at constant relative bandwidth, i.e., the bandwidth of the filter is always proportional to the analyzing frequency. Thus, as in the 1-D case, the transform presents high frequency precision in low frequencies and high spatial precision in high frequencies. This explains why the wavelet transform is suitable for detecting image singularities, such as edges in images [49, 50]. The wavelet transform conserves energy and provides a linear decomposition of  $f$  in terms of the wavelet family  $\{\psi_{\mathbf{b}, \theta, a}\}$  with coefficients  $T_\psi(\mathbf{b}, \theta, a)$ . The transform can be inverted, providing a reconstruction of the original signal from the coefficients as

$$f(\mathbf{x}) = c_\psi^{-1} \iiint T(\mathbf{b}, a, \theta) \psi_{\mathbf{b}, a, \theta}(\mathbf{x}) a^{-3} d^2\mathbf{b} da d\theta, \quad (1.10)$$

where  $c_\psi$  is the normalization constant. For the above formula to hold, the wavelet *admissibility condition* must be met,

$$c_\psi = (2\pi)^2 \int |\hat{\psi}(\mathbf{k})|^2 |\mathbf{k}|^{-2} d^2\mathbf{k} < \infty. \quad (1.11)$$

If  $\psi$  satisfies  $\psi \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ , then the admissibility condition reduces to requiring that the analyzing wavelet have zero mean,

$$\hat{\psi}(\mathbf{0}) = 0 \iff \int \psi(\mathbf{x}) d^2\mathbf{x} = 0. \quad (1.12)$$

Besides the admissibility condition, additional desirable properties may be defined, such as that  $\psi$  and  $\hat{\psi}$  be well localized and that  $\psi$  present a certain number of vanishing moments, being thus blind to polynomials up to the corresponding degree.

The CWT can be implemented efficiently employing the Fast Fourier Transform (FFT) [30] and its Fourier domain definition (Equation 1.9). Otherwise, for analyzing only small scales, the analyzing wavelet may be cropped off in the spatial domain so that the standard correlation implementation might actually become faster. The CWT information is highly redundant and, in practice, it is generally calculated over a discrete family of wavelets  $\{\psi_{\mathbf{b}_i, a_j, \theta_k}\}$ . An approximate reconstruction formula can then be given by

$$\tilde{f}(\mathbf{x}) = \sum_{ijk} \psi_{\mathbf{b}_i, a_j, \theta_k} T_\psi(\mathbf{b}_i, a_j, \theta_k), \quad (1.13)$$

where  $\tilde{f}$  is the reconstructed version of  $f$ . The parameters  $\mathbf{b}_i, a_j, \theta_k$  should be chosen as to provide a stable reconstruction, defining the so called *frame*. The precision of the analyzing wavelet is studied in order to provide directives for defining a suitable sampling, so that the frames may satisfactorily cover the spatial and spatial frequency domains [1, 2, 46]. Even so, the resulting representation will necessarily be redundant, in contrast to the DWT, which allows for deriving orthogonal bases.

### 1.2.3 The 2-D Gabor Wavelet

The 2-D Gabor wavelet has been extensively used within the computer vision community, with special attention to its relevance to studies of the human visual system. One of its most important properties is its capability of detecting and analyzing directional structures. It is also optimal in the sense of minimizing the inherent uncertainty in the spatial and spatial frequency domains. The 2-D version of the Gabor filter was probably introduced in [17], which, together with [54], showed that the filter represents a good model for simple cells of the visual cortex. In a subsequent paper [18], Daugman suggested the generalization of the filter to a wavelet, which was then called the Gabor wavelet. On the other hand, Antoine, Murenzi, and collaborators [1, 59] introduced the generalization of the CWT to two or more dimensions based on coherent states and group theory, in which the same wavelet was called Morlet wavelet, being the natural generalization of the 1-D Morlet wavelet. The name most commonly adopted in the computer vision community is Gabor wavelet, which was therefore also used in this work. Applications of the 2-D Gabor wavelet in computer vision include object representation [42, 46] (particularly, faces [23]), texture detection and segmentation [51, 65] and oriented structures analysis [11, 24]. The 2-D Gabor wavelet has also been shown to outperform other linear filters for detection of oriented features [6].

The 2-D Gabor wavelet is defined as

$$\psi_G(\mathbf{x}) = \exp(i\mathbf{k}_0 \mathbf{x}) \exp\left(-\frac{1}{2}(\mathbf{x} A \mathbf{x})\right) + \text{correction}, \quad (1.14)$$

$$\hat{\psi}_G(\mathbf{k}) = (\det B)^{1/2} \exp\left(-\frac{1}{2}((\mathbf{k} - \mathbf{k}_0) B (\mathbf{k} - \mathbf{k}_0))\right) + \text{correction}, \quad (1.15)$$

where  $A$  is a  $2 \times 2$  positive definite matrix which defines the wavelet anisotropy,  $B = A^{-1}$  and  $\mathbf{k}_0 \in \mathbb{R}^2$  defines the complex exponential basic frequency. The anisotropy matrix is taken as

$$A = \begin{bmatrix} \varepsilon^{-1} & 0 \\ 0 & 1 \end{bmatrix}, \quad (1.16)$$

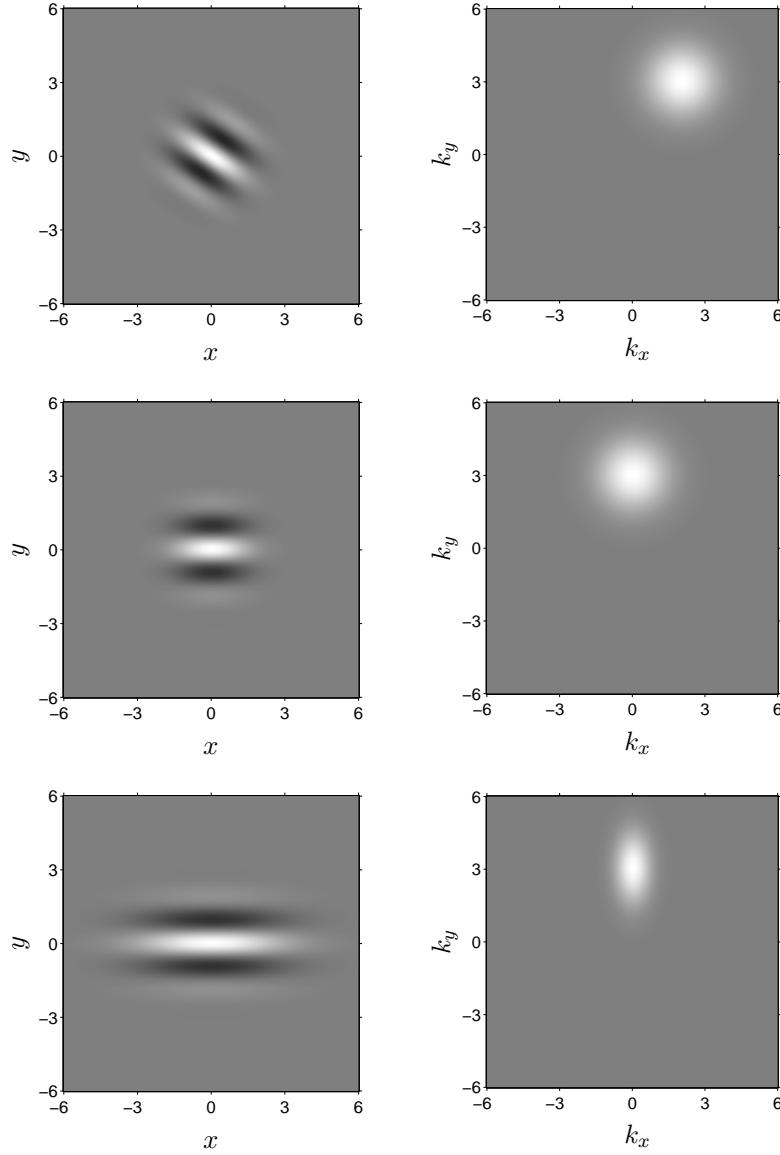
with elongation given by  $\varepsilon \geq 1$ .

The 2-D Gabor wavelet is therefore simply an elongated Gaussian modulated by a complex exponential, which elucidates its relationship to the 1-D Morlet wavelet. The wavelet smoothes the signal in all directions, but detects sharp transitions in the direction of  $\mathbf{k}_0$ . In spatial frequencies, the Gabor wavelet is given by a Gaussian function centered at  $\mathbf{k}_0$  and elongated by  $\varepsilon$  in the  $k_y$  direction. It is thus well localized both in position space, around the origin, and in the spatial frequency space, around  $\mathbf{k}_0$ . Angular selectivity increases with  $|\mathbf{k}_0|$  and  $\varepsilon$ . These effects can be combined choosing  $\mathbf{k}_0 = [0, k_y]$  and  $\varepsilon \gg 1$ . Figure 1.1 shows the effects of changing the wavelet's parameters (both in space and spatial frequencies).

The correction term is necessary to enforce the admissibility condition ( $\hat{\psi}_G(\mathbf{0}) = 0$ ) [1]. However, it is worth noting that the term is numerically negligible for  $|\mathbf{k}_0| \geq 5.6$  and in many situations can be simply dropped. The correction term was not used here for tests, since it implies in loss in spatial localization ( $|\psi_G|^2$  becomes bimodal [35]), though this problem should be addressed in future studies, as discussed in Section 1.6.2.

The definition of a directional wavelet is stated in terms of its spatial frequency domain support [2]. An analyzing wavelet is said to be *directional* if its effective support in the spatial frequency domain is contained within a convex cone with apex at the origin. Wavelets with directional support only in the spatial domain may not be able to detect directional structures, as is the case with the anisotropic Mexican hat [2], which does not satisfy the previously given definition. The Gabor wavelet is thus said to be directional, as its spatial frequency domain support is an ellipse given by the elongated Gaussian of Equation 1.15, as can be seen in the examples of Figure 1.1. If the definition is to be strictly observed, the Gabor wavelet is actually only approximately directional, given the Gaussian tails that are outside the ellipse, but are numerically negligible. The resolving power of the Gabor wavelet has been analyzed in terms of its scale and angular selectivity, helping define wavelet frames that tile the spatial frequency domain [1, 2, 46]. Many other 2-D analyzing wavelets have been proposed in the literature, designed for specific problems, and the reader is referred to [2] for a review.

In the case of vessel segmentation, the Gabor wavelet should be tuned to a suitable frequency, so that vessels may be emphasized while noise and other undesirable structures are filtered out. For the methods presented in this chapter (see Section 1.3.2), the parameters were set to  $\varepsilon = 4$ , resulting in an elongated wavelet, and  $\mathbf{k}_0 = [0, 3]$ , providing a low frequency complex exponential with transitions in the direction of its smaller axis, as shown on the bottom row of Figure 1.1. The Gabor wavelet's capability of detecting oriented features is fundamental in blood vessel detection. It is interesting to note that the wavelet's shape is locally similar that of a blood vessel and is preserved across different orientations and scales. Therefore, stronger wavelet responses are produced when the wavelet is found at the same position, orientation and scale as a vessel in the image.

**FIGURE 1.1**

The Gabor wavelet in the spatial domain (represented by its real part) is shown in the left column, while the frequency domain counterpart is shown in the right column. Different configurations of the parameters are shown to illustrate the corresponding effects. Light and dark gray levels correspond to positive and negative coefficients, respectively. The parameters for the three rows are, from top to bottom:  $\mathbf{k}_0 = [2, 3]$ ,  $\varepsilon = 1$ ;  $\mathbf{k}_0 = [0, 3]$ ,  $\varepsilon = 1$ ; and  $\mathbf{k}_0 = [0, 3]$ ,  $\varepsilon = 4$ .

### 1.2.4 Supervised Classification

In the proposed vessel segmentation approach, image pixels are seen as objects represented by feature vectors, so that statistical classifiers might be applied for segmentation. In this case, each pixel is classified as *vessel* or *nonvessel*, using previously trained classifiers (supervised classification). The training sets for the classifiers are derived from manual segmentations of training images: pixels that were manually segmented out are labeled as *vessel*, while the remaining receive the *nonvessel* label. The approach allows the use of different responses of the wavelet transform to characterize pixels, as well as application to different image modalities, provided a corresponding training set is available. The rest of this section contains descriptions pointing out important qualities of the three classifiers tested: the Bayesian classifier using the Gaussian mixture model (Sections 1.2.5 and 1.2.6), the  $k$ -nearest neighbor classifier (Section 1.2.7) and the linear minimum squared error classifier (Section 1.2.8).

### 1.2.5 Bayesian Decision Theory

Bayesian decision theory is one of the main approaches adopted in pattern recognition problems. It is based on associating costs (or losses) to classification decisions, which are then chosen using the probability distributions of the objects being classified. In most cases, the probability distributions are unknown and must be estimated, which can be done using prior knowledge of the problem structure along with the available training data. *Bayes's decision rule* for an observed feature vector  $\mathbf{v}$  in a two class problem (classes  $C_1$  and  $C_2$ ) can be written in terms of posterior probabilities as:

$$\begin{aligned} &\text{decide } C_1 \text{ if } P(C_1|\mathbf{v}) > P(C_2|\mathbf{v}); \\ &\text{otherwise, decide } C_2. \end{aligned} \quad (1.17)$$

The decision follows natural intuition and is shown to minimize the average probability of error [20].

*Bayes's formula* is written as

$$P(C_i|\mathbf{v}) = \frac{p(\mathbf{v}|C_i)P(C_i)}{p(\mathbf{v})}, \quad (1.18)$$

where  $p(\mathbf{v}|C_i)$  is the class-conditional probability density function, also known as likelihood,  $P(C_i)$  is the prior probability of class  $C_i$  and  $p(\mathbf{v})$  is the probability density function of  $\mathbf{v}$ , also known as evidence. To obtain a decision rule based on estimates from the training data (as will be needed in Section 1.2.6), Bayes's formula is applied to Equation 1.17, resulting in the equivalent decision rule:

$$\begin{aligned} &\text{decide } C_1 \text{ if } p(\mathbf{v}|C_1)P(C_1) > p(\mathbf{v}|C_2)p(C_2); \\ &\text{otherwise, decide } C_2. \end{aligned} \quad (1.19)$$

### 1.2.6 Bayesian Gaussian Mixture Model Classifier

According to Bayesian decision theory, classification decisions can be made using estimates of prior probabilities and likelihood functions. In the experiments performed, prior probabilities  $P(C_i)$  were estimated as  $N_i/N$ , the ratio of class  $C_i$  samples in the training set. The class likelihoods, in turn, were described using the *Gaussian mixture model* (GMM), which consists of a linear combination of Gaussian functions [20, 69]. The Bayesian classifier using the GMM for class likelihoods will be called the GMM classifier. The procedure for estimating the model's parameters was applied separately for each class likelihood. To simplify notation, suppose the class  $C_i$  whose likelihood will be estimated is fixed. The likelihood of the class (earlier denoted  $p(\mathbf{v}|C_i)$ ) is then simply denoted  $p(\mathbf{v}|\phi)$ , being modeled by

$$p(\mathbf{v}|\phi) = \sum_{j=1}^c P_j p(\mathbf{v}|\phi_j), \quad (1.20)$$

where  $c$  Gaussians are used, given by  $p(\mathbf{v}|\phi_j)$  and weights  $P_j$ . Each  $\phi_j$  describes the parameters of Gaussian  $j$ , while the complete set of parameters that describe the model is denoted  $\phi \equiv \{\phi_1, \dots, \phi_c, P_1, \dots, P_c\}$ . The Gaussian functions are

$$p(\mathbf{v}|\phi_j) = \frac{1}{\sqrt{\det(2\pi\Sigma_j)}} \exp\left(-\frac{1}{2}(\mathbf{v} - \mu_j)^T \Sigma_j^{-1} (\mathbf{v} - \mu_j)\right). \quad (1.21)$$

where  $\phi_j = \{\mu_j, \Sigma_j\}$ ,  $\mu_j$  is the mean and  $\Sigma_j$  the covariance matrix that describe the Gaussian. In order for  $p(\mathbf{v}|\phi)$  to be a probability density function, it is necessary that

$$\sum_{j=1}^c P_j = 1 \quad \text{and} \quad \int p(\mathbf{v}|\phi_j) d\mathbf{v} = 1, \quad j = 1, \dots, c. \quad (1.22)$$

GMMs are extensively used for clustering in unsupervised learning problems. In that context, each Gaussian models the likelihood of a class with previously unknown distribution, with each Gaussian weight representing a prior class probability. Here, the GMM is used with a different purpose, in the description of class likelihoods from samples with known labels. The GMM allows likelihoods to be represented with arbitrary precision, being a flexible model applicable to complex and multimodal distributions. In both applications, there is the need to estimate the Gaussians' parameters and weights from the available data.

The GMM parameters  $\phi$  are estimated so as to maximize the likelihood of the training samples. The parameters enter the likelihood in a nonlinear fashion, requiring nonlinear optimization techniques. In the experiments performed, the parameters were estimated using the *Expectation-Maximization* (EM) algorithm [19, 20, 69]. EM is an iterative process usually chosen for the estimation of mixture model parameters. It guarantees a local maximum of the function being optimized and can be applied to maximize likelihoods or posterior probabilities.

A difficulty in estimating the GMMs is the choice of the number of Gaussians  $c$  to use. It can be shown that GMMs can approximate with arbitrary precision any

continuous probability density function, given a sufficiently large number of Gaussians. However, large values of  $c$  may cause models to adjust excessively to the data (phenomenon known as over-fitting), while too small a  $c$  might not permit flexibility enough to properly describe the probability function. Experiments were performed varying the values of  $c$  for both classes, leading to different results. Finally, EM guarantees only a local maximum of the likelihood, being dependent on initialization. EM has received a good amount of attention in the last years, being used in a variety of applications. Many methods were developed to overcome the mentioned problems [25], but were not explored in this work.

GMMs represent a halfway between purely nonparametric and parametric approaches, providing a fast classification phase at the cost of a more expensive training algorithm. Nonparametric methods (as  $k$ -nearest neighbor classification presented in Section 1.2.7) are computationally demanding for large numbers of training samples, though they do not impose restrictions on the underlying probability distributions. On the other hand, GMMs guarantee a fast classification phase that depends only on the chosen  $c$  (i.e. independent of the number of training samples), while still allowing for modeling complex distributions.

### 1.2.7 $k$ -Nearest Neighbor Classifier

A simple and popular classification approach is to use the  *$k$ -nearest neighbor* ( *$k$ NN*) rule. Let  $N$  be the total number of labeled training samples. Given an odd number  $k$  of neighbors, a distance measure, and a feature vector  $\mathbf{v}$ , the  $k$ -nearest neighbor classification rule can be summarized in the following steps:

1. out of the  $N$  training samples, identify the  $k$ -nearest neighbors of  $\mathbf{v}$  using the given distance measure;
2. out of the  $k$ -nearest samples identified, count the number of samples  $k_i$  that belong to each class  $C_i$ ;
3. classify  $\mathbf{v}$  as belonging to the class  $C_i$  with the largest number  $k_i$  of samples.

Various distance measures may be used, such as Euclidean or Mahalanobis distances, leading to different outcomes [20]. The Euclidean distance was used in this work for all experiments. The  $k$ NN classifier is nonparametric (independent of a probability distribution model) and capable of producing complex nonlinear decision boundaries. Theoretical superior limits on the classifier's error probability can be established, which decrease with  $N$ , so that when  $k \rightarrow \infty$ , the error tends to the optimal Bayesian error [20].

The classification rule may be interpreted as a decision taken based on estimates of the posterior probabilities  $P(C_i|\mathbf{v})$  – simply estimated as  $k_i/k$  – from the data. From this perspective, as  $k$  increases, so does the confidence on the estimates, but spatial precision is lost, as samples that are far from  $\mathbf{v}$  start being considered. The greater the number of samples  $N$ , the smaller the effect of loss in spatial precision, allowing for larger values of  $k$  to be used. In this study, many training samples are

available, allowing for good estimates, but demanding a large computational effort. A problem with nearest neighbor techniques is the computational complexity of the search for nearest neighbors among the  $N$  training samples. If the dimension of the feature space is fixed, exhaustive search of the nearest neighbors of one sample takes time at least  $O(N)$ , necessary for the calculation of distances. Being so, strategies have been studied to improve performance, including efficient searches and reduction of the number of training samples used [20, 69]. Only the exhaustive search was implemented for the results shown here.

### 1.2.8 Linear Minimum Squared Error Classifier

The *linear minimum squared error classifier* [20, 69], denoted LMSE, was also tested. Linear classifiers are defined by a linear decision function  $g$  in the  $d$ -dimensional feature space:

$$g(\mathbf{v}) = \mathbf{w}^T \mathbf{v} + w_0, \quad (1.23)$$

where  $\mathbf{v}$  is a feature vector,  $\mathbf{w}$  is the weight vector and  $w_0$ , the threshold. The classification rule is to decide  $C_1$  if  $g(\mathbf{v}) > 0$  and  $C_2$  otherwise. To simplify the formulation, the threshold  $w_0$  is accommodated by defining the extended  $(d+1)$ -dimensional vectors  $\mathbf{v}' \equiv [\mathbf{v}^T, 1]^T$  and  $\mathbf{w}' \equiv [\mathbf{w}^T, w_0]^T$ , so that  $g(\mathbf{v}) = \mathbf{w}'^T \mathbf{v}'$ .

The classifier is determined by finding  $\mathbf{w}'$  that minimizes the *sum of squared error* criterion on the training set, defined as

$$J(\mathbf{w}') = \sum_{i=1}^N (\mathbf{v}'_i^T \mathbf{w}' - y_i)^2, \quad (1.24)$$

where  $N$  is the total number of training samples,  $\mathbf{v}'_i$  is the extended  $i^{th}$  training sample, and  $y_i$ , its desired output. The criterion measures the sum of squared errors between the output of the classifier  $\mathbf{v}'_i^T \mathbf{w}'$  and the desired output  $y_i$ . The desired outputs were arbitrarily set to  $y_i = 1$  for  $\mathbf{v}_i \in C_1$  and  $y_i = -1$  for  $\mathbf{v}_i \in C_2$ .

Defining

$$V = \begin{bmatrix} \mathbf{v}'_1^T \\ \mathbf{v}'_2^T \\ \vdots \\ \mathbf{v}'_N^T \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \quad (1.25)$$

the gradient of the criterion function can be written as

$$\nabla J(\mathbf{w}') = \sum_{i=1}^N 2(\mathbf{v}'_i^T \mathbf{w}' - y_i) \mathbf{v}'_i \quad (1.26)$$

$$= 2V^T(V\mathbf{w}' - \mathbf{y}). \quad (1.27)$$

The criterion function is minimized by calculating  $\hat{\mathbf{w}}'$  where the gradient equals zero, resulting in

$$(V^T V) \hat{\mathbf{w}}' = V^T \mathbf{y} \Rightarrow \hat{\mathbf{w}}' = (V^T V)^{-1} V^T \mathbf{y}. \quad (1.28)$$

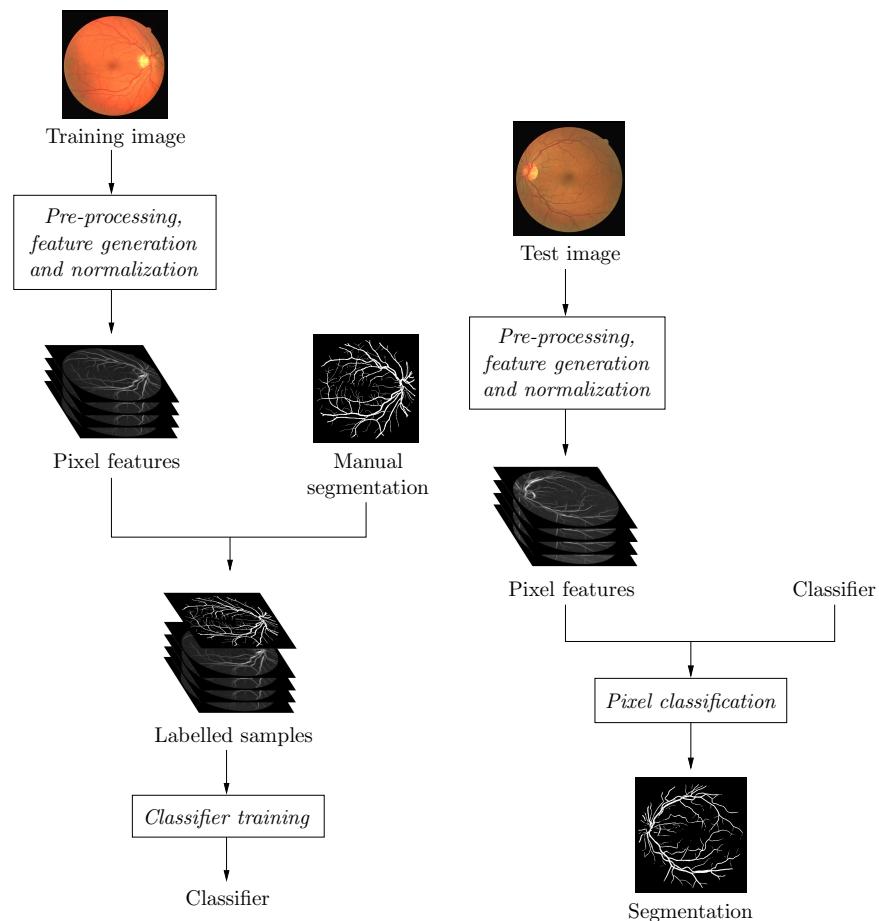
$(V^T V)^{-1} V^T$  is called the *pseudoinverse* of  $V$  and exists only if  $V^T V$  is invertible.  $V^T V$  is the *correlation matrix* of the training samples and will be invertible when  $V$  has full rank  $d + 1$ . Given the nature of the data used in tests and the large number of training samples, this always occurred.

In comparison to the previously described  $k$ NN and GMM classifiers, the LMSE classifier has faster training and test phases, but is restricted in the sense that it is linear, while the others allow for complex decision boundaries. However, as will be shown, the results obtained using LMSE are comparable to those using GMMs (see Section 1.5), representing a reasonable trade-off.

### 1.3 Segmentation Using the 2-D Gabor Wavelet and Supervised Classification

Initially, promising results of vessel detection using the 2-D Gabor wavelet were shown [9, 44]. The segmentation method evolved with the introduction of supervised pixel classification, allowing wavelet responses from different scales to be combined [45]. Finally, the Bayesian Gaussian mixture model classifier showed to be specially appropriate for the segmentation task and was quantitatively evaluated using ROC analysis. The method and some of the results presented in this chapter were previously published in a journal paper [66]. Here, the method is presented in greater detail, including a denser theoretical review, new tests and analysis of results. The method developed uses the 2-D CWT coupled with supervised pixel classification in classes *vessel* and *nonvessel*. Supervised classification requires labeled samples during a training phase, which here are provided from manual vessel segmentations. A diagram illustrating the training of a classifier is presented in Figure 1.2(a). After training, the classifier can be applied to pixels of test images as shown in Figure 1.2(b). The process presented in the diagrams is described in greater detail throughout this section. The supervised learning approach combining different wavelet scale responses simplifies method use, minimizing the need for interaction and parameter configuration.

This section presents the method developed and the experimental evaluation performed. The GMM classifier is shown to provide good results, while at the same time guaranteeing a fast classification phase. For comparison, tests are performed in which the GMM classifier is substituted by the  $k$ NN and LMSE classifiers (see Section 1.2). Performance of the 2-D Gabor wavelet in enhancing blood vessels is demonstrated by comparing results of filtering using a single wavelet scale with results of the 2-D Gaussian matched filter of Chaudhuri *et al.* [10]. Quantitative per-

**FIGURE 1.2**

**Supervised pixel classification approach.** The left diagram illustrates the supervised training of a classifier. The trained classifier can then be applied to the segmentation of test images, as illustrated in the right diagram.

formance evaluation is performed using ROC analysis, which has been previously used for evaluation and comparison of retinal vessel segmentation methods on publicly available image databases [38, 40, 58, 67, 68]. ROC analysis, as well as visual inspection, has shown that there is still room for improvement: human observers are significantly more accurate than the methods, which show flaws around the optic disk and in detection of the thinnest vessels [15, 60].

### 1.3.1 Pre-processing

When the RGB components of colored images are visualized separately, the green channel shows the best vessel to background contrast, whereas the red and blue channels usually present low vessel contrast and are noisy [77]. Hence, for colored images, only the green channel is used for the generation of the wavelet features, as well as to compose the feature vector itself, i.e. the green channel intensity of each pixel is taken as one of its features. The green channel is inverted before the application of the wavelet transform to it, so that the vessels appear brighter than the background. Accordingly, the inversion is also applied to red-free grayscale images and fluorescein angiograms do not require inversion (see Section 1.1).

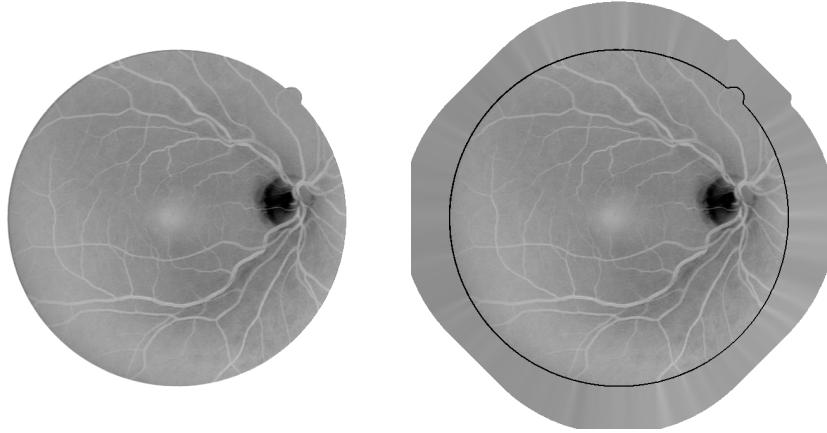
The Gabor wavelet responds strongly to high contrast edges, which may lead to false detection of the borders of the camera's aperture. In order to reduce this effect, an iterative algorithm has been developed. The intent is to remove the strong contrast between the retinal fundus and the region outside the camera's FOV (see Figure 1.3).

The pre-processing algorithm starts with a region of interest (ROI) determined by the camera's aperture and iteratively grows this ROI. Each step of the algorithm consists in the following: first, the set of pixels of the exterior border of the ROI is determined, i.e. pixels that are outside the ROI and are neighbors (using 4-neighborhood) to pixels inside it; then, each pixel value of this set is replaced with the mean value of its neighbors (this time using 8-neighborhood) inside the ROI; finally, the ROI is expanded by inclusion of this set of changed pixels. This process is repeated and may be seen as artificially increasing the ROI, as shown in Figure 1.3(b).

### 1.3.2 2-D Gabor Wavelet Features

Among several available 2-D analyzing wavelets, the Gabor wavelet was adopted for vessel detection here and in previous works [44], based on the following properties. The wavelet is capable of detecting directional structures and of being tuned to specific frequencies (see Section 1.2.3), which is specially important for filtering out the background noise present in retinal images. Furthermore, it has been shown to outperform other oriented feature detectors [6].

The Gabor wavelet parameters must be configured in order to enhance specific structures or features of interest. In the tests performed, the elongation parameter was set to  $\epsilon = 4$ , making the filter elongated and  $\mathbf{k}_0 = [0, 3]$ , i.e. a low frequency complex exponential with few significant oscillations perpendicular to the large axis of the wavelet, as shown in Figure 1.4. These two characteristics are specially suited for the detection of directional features and have been chosen in order to enable

**FIGURE 1.3**

**Image pre-processing for removing undesired border effects.** The inverted green channel of the image presented in Figure 1.6 appears on the left. The inverted green channel after pre-processing, presenting the extended border, appears on the right (the original image limit is presented for illustration).

the transform to present stronger responses for pixels associated with vessels. Note that the wavelet's shape is similar to the vessels', so that the transform yields strong coefficients when at the same position, scale and orientation as a vessel, by means of the scalar product in  $L^2(\mathbb{R}^2)$  (see Section 1.2.2).

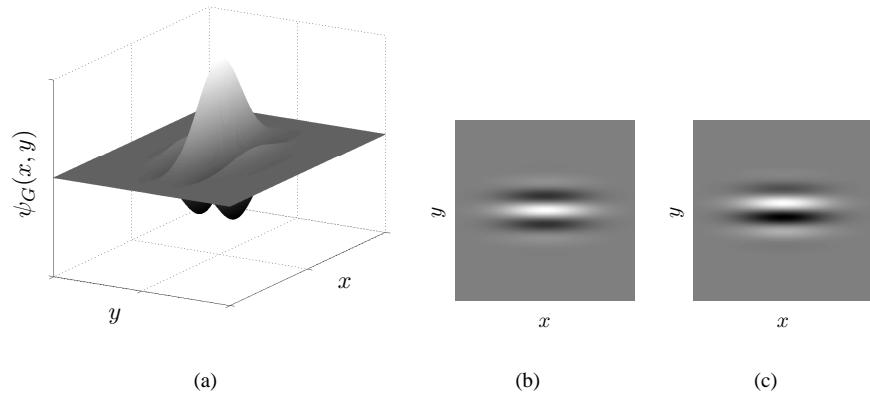
In order to detect vessels in any orientation, for each considered position and scale, the response with maximum modulus over all possible orientations is kept, i.e.

$$M_\psi(\mathbf{b}, a) = \max_{\theta} |T_\psi(\mathbf{b}, a, \theta)|. \quad (1.29)$$

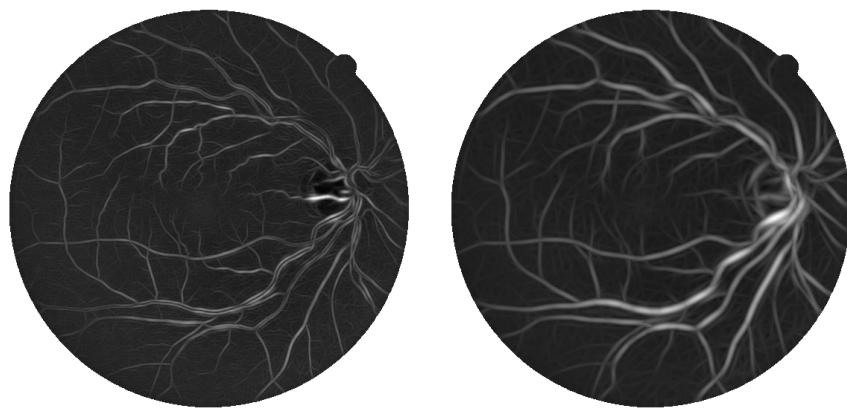
Thus, for each pixel position and chosen scale, the Gabor wavelet transform is computed for  $\theta$  spanning from 0 up to 170 degrees at steps of 10 degrees and the maximum is taken (this is possible because  $|T_\psi(\mathbf{b}, a, \theta)| = |T_\psi(\mathbf{b}, a, \theta + 180)|$ ). The maximum moduli of the wavelet transform over all angles for various scales are then taken as pixel features.  $M_\psi(\mathbf{b}, a)$  is shown in Figure 1.5 for  $a = 2$  and  $a = 5$  pixels.

### 1.3.3 Feature Normalization

The measures used as features may have ranges spanning different orders of magnitude. This can lead to errors in the classification process because of the disparity of each feature's influence in the calculation of feature space distances. A strategy to obtain a new random variable with zero mean and unit standard deviation, compensating for eventual magnitude differences, is to apply the normal transformation to each feature. The normal transformation is defined as [14]

**FIGURE 1.4**

Different representations for the 2-D Gabor wavelet ( $\psi_G$ ) with parameters  $\varepsilon = 4$  and  $k_0 = [0, 3]$ : (a) surface representation of the real part; (b) real part; (c) imaginary part. Darker and lighter shades represent, respectively, positive and negative values.

**FIGURE 1.5**

Maximum modulus of the Gabor wavelet transform over angles,  $M_\psi(b, a)$  (Equation 1.29), for scale values of  $a = 2$  (left) and  $a = 5$  (right) pixels.

$$\hat{v}_i = \frac{v_i - \mu_i}{\sigma_i}, \quad (1.30)$$

where  $v_i$  is the  $i^{th}$  feature of each pixel,  $\mu_i$  is the average value of the feature and  $\sigma_i$ , its standard deviation.

The normal transformation was applied separately to each image, i.e., each image's feature space was normalized using its own means and standard deviations, helping to compensate for intrinsic variation between images, such as global illumination variation.

### 1.3.4 Supervised Pixel Classification

After feature generation and normalization, segmentations are obtained through supervised classification of image pixels into classes  $C_1 = \{\text{vessel pixels}\}$  and  $C_2 = \{\text{non-vessel pixels}\}$ . Supervised classification involves using labeled training samples (samples with known classes) to compose a training set, that then serves as a basis for future classification. Manual image segmentations (see Section 1.3.5) can be used to provide these labels. Many classifiers then go through a training phase, involving, for example, parameter estimation on the training set, preparing them for classification of new, unlabeled, pixel samples (see Figure 1.2).

In the experiments performed, the training sets were composed of labeled pixels from several manually segmented retinal images. Due to the computational cost of training the classifier and the large number of samples, subsets of the available labeled samples were randomly selected to actually be used for training.

Another approach for image segmentation through supervised classification, which showed interesting results, is to form the training set with labeled samples taken from a portion of the image to be segmented [13]. Using this approach, a semi-automated fundus segmentation software could be developed, in which the user only has to draw a small portion of the vessels over the input image or simply click on several pixels associated with vessels. The remaining image would then be automatically segmented based on this specific training set. This approach is interesting since it requires a small effort from the user, which is compensated by the fact that image peculiarities are directly incorporated by the classifier.

### 1.3.5 Public Image Databases

There are different ways of obtaining retinal images, such as with colored digital cameras or through angiography using fluorescent dyes (see Section 1.1). The pixel classification approach can be applied to different image modalities, provided appropriate manual segmentations are available for training. The approach described here has been tested on colored images and grayscale angiograms [44, 45]. In order to facilitate comparisons with other methods, here experiments are presented using two publicly available databases of colored images and corresponding manual segmentations: the DRIVE [68] and STARE [38] databases.

The DRIVE database consists of 40 images (7 of which present pathology), randomly selected from a diabetic retinopathy screening program in the Netherlands, along with manual segmentations of the vessels. They were captured in digital form from a Canon CR5 nonmydriatic 3CCD camera at 45° field of view (FOV). The images are of size  $768 \times 584$  pixels, 8 bits per color channel and have a FOV of approximately 540 pixels in diameter. The images are in compressed JPEG-format, which is unfortunate for image processing, but is commonly used in screening practice.

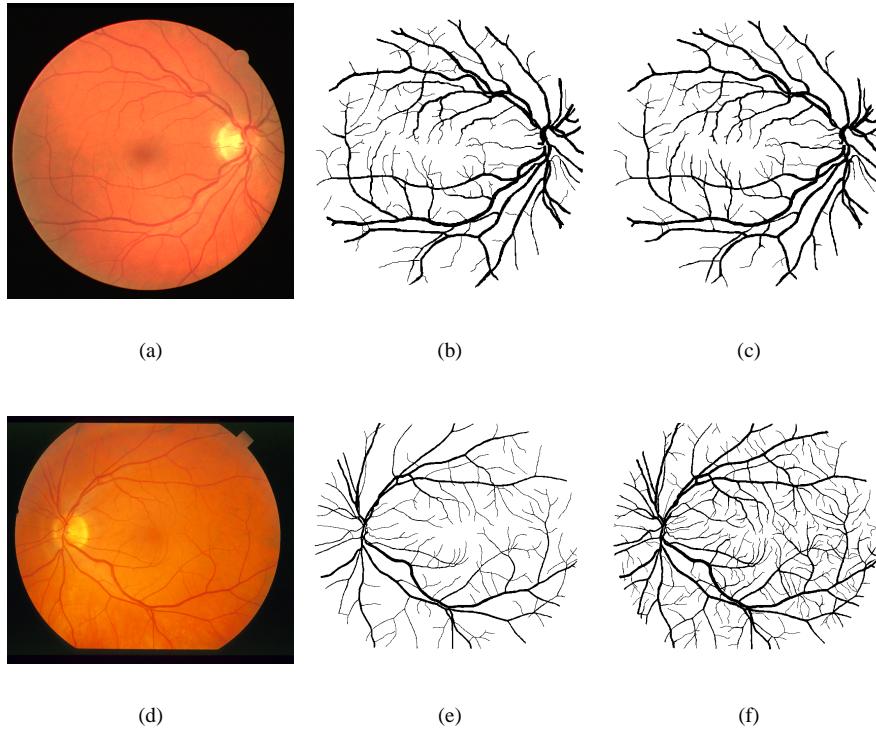
The authors of this database divided the 40 images into fixed training and test sets, each containing 20 images (the training set has 3 images with pathology and the test set, 4) [60, 68]. The images have been manually segmented by three observers trained by an ophthalmologist. Images in the training set were segmented once, while images in the test set were segmented twice, resulting in sets A and B of manual segmentations. The observers of sets A and B produced similar segmentations: in set A, 12.7% of pixels were marked as vessel, against 12.3% vessel for set B. A normal image from the test set and its respective manual segmentations are illustrated in Figure 1.6. Performance is measured on the test set using the segmentations of set A as ground truth. The segmentations of set B are tested against those of A, serving as a human observer reference for performance comparison.

The STARE database consists of 20 digitized slides captured by a TopCon TRV-50 fundus camera at 35° FOV. The slides were digitized to  $700 \times 605$  pixels, 8 bits per color channel. The FOV in the images are approximately  $650 \times 550$  pixels in diameter. Images were selected so that ten of them contained pathologies, which complicate vessel detection. This choice was made so that the performance difference on normal and pathological images could be assessed. Two observers manually segmented all 20 images. The first observer segmented 10.4% of pixels as vessel, against 14.9% vessels for the second observer. The segmentations of the two observers are fairly different in that the second observer consistently segmented much more of the thinner vessels than the first. Figure 1.6 shows an image from the database and its respective manual segmentations. Previous tests using this database [15, 38, 40, 58, 68] calculated performance using segmentations from the first observer as ground truth, with the second observer's segmentations being tested against the first. There is no predefined separation into training and test sets, therefore leave-one-out tests are performed for supervised method performance evaluation.

### 1.3.6 Experiments and Settings

The methods described were tested on the DRIVE and STARE databases with the following settings. The pixel features used for classification were the inverted green channel and its maximum Gabor wavelet transform response over angles  $M_\psi(\mathbf{b}, a)$  (Equation 1.29) for scales  $a = 2, 3, 4, 5$  pixels (see Subsection 1.3.2). These scales were chosen as to span the possible widths of vessels throughout the images, so that all vessels could be detected.

For the DRIVE database, the training set was formed by pixel samples from the 20 labeled training images. For the STARE database, leave-one-out tests where per-

**FIGURE 1.6**

**Images from the DRIVE and STARE databases and their respective manual segmentations:** (a) normal image from the test set of the DRIVE database; (b) manual segmentation from set A; (c) manual segmentation from set B; (d) normal image from the STARE database; (e) manual segmentation by first observer; (f) manual segmentation by second observer.

formed, i.e., every image was segmented using samples from the other 19 images for the training set (see Section 1.3.5). The three classifiers presented in Section 1.2 were evaluated so as to allow their comparison. Due to the large number of pixels, in all experiments, samples were randomly chosen to train the classifiers. The three classifiers were compared using one hundred thousand training samples ( $N = 10^5$ ), because of computational demands of  $k$ NN classification. However, for comparison with other methods from the literature, the GMM classifier was trained using one million samples ( $N = 10^6$ ). The GMM classifier was tested varying the number  $c$  of Gaussians modeling each class likelihood, while experiments with the  $k$ NN classifier used different values for the number  $k$  of neighbors analyzed. Additionally, to verify the dependence of the method on the training set, a test was performed in which the GMM classifier was trained on each of the DRIVE and STARE databases, while

**Table 1.1:** Confusion matrix.

		Ground truth	
		Positive ( $p$ )	Negative ( $n$ )
Method result	Positive	True positive ( $tp$ )	False positive ( $fp$ )
	Negative	False negative ( $fn$ )	True negative ( $tn$ )

being tested on the other.

To demonstrate the performance of the Gabor wavelet in enhancing blood vessels, results of filtering using a single wavelet scale are presented and compared with results of the 2-D Gaussian matched filter of Chaudhuri *et al.* [10]. The parameters of both filters were chosen as to produce the best results:  $a = 4$  pixels for wavelet filtering and  $\sigma = 1$  pixel for the matched filter of Chaudhuri *et al.*

### 1.3.7 ROC analysis

The methods are evaluated using curves in receiver operating characteristic (ROC) graphs. ROC graphs have long been used in signal detection theory and more recently in pattern recognition and medical diagnostic systems [22]. They allow for visualizing the performance of classifiers and diagnostic tests, expressing the trade-off between increased detection and false alarms rates.

ROC curves are formed by ordered pairs of true positive and false positive rates. Different rates are established varying each method's parameters [38, 40, 58] or thresholds on posterior probabilities [68]. For each configuration of parameters or threshold value, a pair formed by a true positive and false positive rate corresponding to the method's outcome is marked on the graph, producing a curve as in Figure 1.7. In the scenario of evaluating vessel segmentations, true positives are pixels marked as vessel (positive) in both the segmentation given by a method and the manual segmentation used as ground truth. False positives are pixels marked as vessel by the method, but that are actually negatives in the ground truth. Classification measures are summarized in the *confusion matrix* (also known as contingency table), illustrated in Table 1.1. The total of positives and negatives in the ground truth are denoted  $p$  and  $n$ ; the total number of true and false positives are  $tp$  and  $fp$ ; and true and false negatives are denoted  $tn$  and  $fn$ , as shown in the confusion matrix. Metrics used in evaluation can then be derived from the matrix elements. True positive and false positive rates, denoted  $tpr$  e  $fpr$ , are given by

$$tpr = \frac{tp}{p}, \quad fpr = \frac{fp}{n}. \quad (1.31)$$

The accuracies (fraction of correctly classified pixels) of methods and human observers tested against the ground truth are also measured during performance evaluation. Accuracy is given by  $\frac{tp+tn}{p+n}$ , simply measuring the fraction of correct classification. Contrasting with ROC measures, accuracy does not express the relation between quantities of true positives and false positives, being sensitive to skews in class distribution.

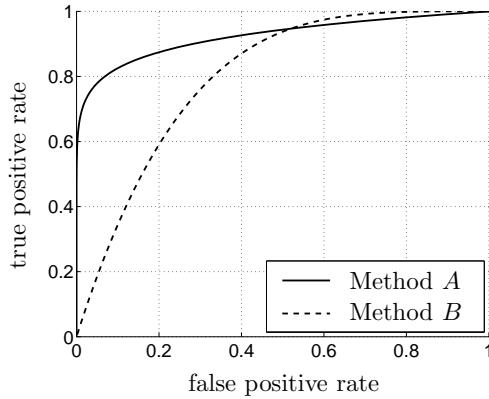
Some points on the ROC graph help give an intuitive idea of the representation (see Figure 1.7). The lower left point  $(0,0)$  represents an extremely conservative method, that assigns only negative to all samples. Thus, it is incapable of producing true positives, though it also does not produce false positives. The opposite strategy, which is to assign positive to all samples, corresponds to the upper right point  $(1,1)$ . The closer an ROC curve is to the upper left corner, the better the method's performance, with the point  $(0,1)$  representing a perfect agreement with the ground truth. Accordingly, an ROC curve is said to dominate another if it is completely above and to the left of it. The diagonal line in which  $tpr = fpr$  represents a random assignment strategy, with different points of the line corresponding to different rates of random assignment of negatives and positives.

A very important property of ROC graphs is their invariance to changes in prior class distributions, represented by the proportion of positive and negative samples. This is possible because the rates analyzed are relative to the total of samples in each class. The same does not apply to other measures, such as accuracy, described previously. In an analogous manner, ROC graphs are also invariant to changes in costs associated with classification decisions [63].

The area under an ROC curve ( $A_z$ ) is used as a single scalar measure of a method's performance.  $A_z$  values, being fractions of a unitary square, necessarily lie between 0 and 1. Nevertheless, as a random classifier would produce a diagonal line connecting  $(0,0)$  and  $(1,1)$  with  $A_z = 0.5$ , in practice, methods should always have  $A_z \geq 0.5$ . A method that agreed completely with the manual segmentations used as ground truth would yield an  $A_z = 1$ . It is possible for a method with larger area under the curve to have a worse performance in given region of the graph than one with smaller area, as illustrated in Figure 1.7. However, the area serves as good estimate of a method's overall performance. The areas under the ROC curves have an important statistical property: they are equivalent to the Wilcoxon statistic [34], the probability that a method assign a larger chance or rank to a randomly chosen positive sample than that assigned to a randomly chosen negative sample.

In order to make conclusions about method superiority, it is not sufficient that a method's ROC curve calculated over a given data set dominate that of another's, as variances must be taken into account. These variances can be estimated using approaches such as cross-validation and bootstrapping [22]. The same principle applies for comparison of accuracy and  $A_z$  measures. In [58, 67, 68], these measures were taken for different methods and compared in paired statistical hypothesis tests, with the accuracy and  $A_z$  of each image segmentation result representing an observation. Here, the evaluation does not include significance tests, though that would provide for a stronger statement in comparisons.

In the experiments performed, measures were taken over all test images, considering only pixels inside the FOV defined by the camera aperture. For the GMM and  $k$ NN classifiers, ROC curves are produced by varying the thresholds on posterior pixel probability estimates, while the LMSE ROC curve is produced by varying the threshold  $w_0$  applied to the scalar product between the classifier's weight vector and pixel feature vectors (Equation 1.23). Finally, the ROC curves for filtering using a single wavelet scale and the matched filter of Chaudhuri *et al.* are produced vary-

**FIGURE 1.7**

**Example of an ROC graph with two curves. The area under the curve ( $A_z$ ) of method A is larger than that of method B, though B should be preferred for false positive rates larger than approximately 0.5.**

ing the threshold on the filters' responses. For the DRIVE database, performance is calculated using manual segmentations from set A as ground truth and human observer performance is estimated from manual segmentations from set B, which only provide one true/false positive rate pair, appearing as a point in the ROC graph (Figure 1.13). For the STARE database, the first observer's manual segmentations are used as ground truth and the second observer's true/false positive rate pair is plotted on the graph (Figure 1.14). It is important to point out that the manual segmentations evaluated do not present perfect true/false positive rates, as they disagree on several pixels with the segmentations used as ground truth. Thus, the variance between observers can be estimated, helping set a goal for method performance.

## 1.4 Implementation and Graphical User Interface

### 1.4.1 Overview

The method's implementation for performing experiments originated a package of MATLAB [55] scripts, that now also includes a graphical user interface (GUI) for preliminary testing. The package, named *mlvessel*, is available as open-source code under the GNU General Public License (GPL) [26] at the project's collaborative development website, <http://retina.iv.fapesp.br>. The website also contains some of the results presented here, allowing researchers to evaluate them in new and diverse manners. Experiments can be executed through function calls, with the specification of options and parameters in configuration files. Results are then presented as im-

ages and tables organized in HTML pages, as illustrated in Figure 1.8. Most of the package’s functionality is also accessible through the GUI, outlined in Section 1.4.4.

To facilitate testing, a pipeline architecture was used, in which intermediate results are saved on disk for reutilization. In this manner, pixel features, training sets and classifiers can be reused in different tests, saving up the computational effort necessary for their creation. Components of the pipeline, such as feature generation or classifier training (refer to Figure 1.2), can then be substituted for comparative tests. The software’s implementation is organized in the modules listed below, each corresponding to a directory inside the `src/` directory.

- **tests:** examples of tests, which manipulate other modules. For complete examples, see scripts `testmixed.m` and `testleaveoneout.m`.
- **gui:** the GUI. It is initialized through its main window in `guemain.m`, which provides access to all others.
- **ftrs:** feature generation and manipulation, including creation of training sets. The `createfeatures` functions produce raw pixel features from images. The `createlabelled` functions receive the raw features and also manual segmentations, saving everything as labelled pixel features. Finally, the function `createprocessed` normalizes the labelled features, forming a training set which can then be fed to classifiers.
- **gmm:** the GMM classifier. `gmmcreatemodel` receives a processed training set in order to create and save a classifier. `gmmclassify` receives a saved classifier and raw pixel features, producing each pixel’s class as output.
- **knn:** an exhaustive implementation of the *k*NN classifier.
- **lmse:** the LMSE classifier. Creation and application of the classifier is analogous to that of the **gmm** module.
- **stats:** statistics generation from results. ROC graphs and statistics are organized in HTML pages, as illustrated in Figure 1.8.
- **html:** functions for formatting and saving HTML pages.
- **skel:** post-processing and scale-space skeletonization [14] for application to vasculature segmentations. Part of the module is written in C with a MATLAB programming interface.

MATLAB is a software extensively used for tasks such as image processing and includes an environment that provides for fast, prototypical, software development. However, it is a commercial program and its use requires a license that is usually paid for. A new implementation of the method and interface is being developed in C++, which should be faster, have better usability and be independent of MATLAB, allowing wider reach.

### 1.4.2 Installation

Installation should be simple, consisting in the following steps.

1. Download the `mlvessel` package, version 1.2 or later, from the project's website (<http://retina.iv.fapesp.br>).
2. Unpack the package (`mlvessel.zip` or `mlvessel.tar.gz`), using `unzip` or `gnuzip` and `tar`. Move the unzipped files to the directory where you wish the package to be installed.
3. Run the `mlvesselinstall.m` script from within MATLAB.

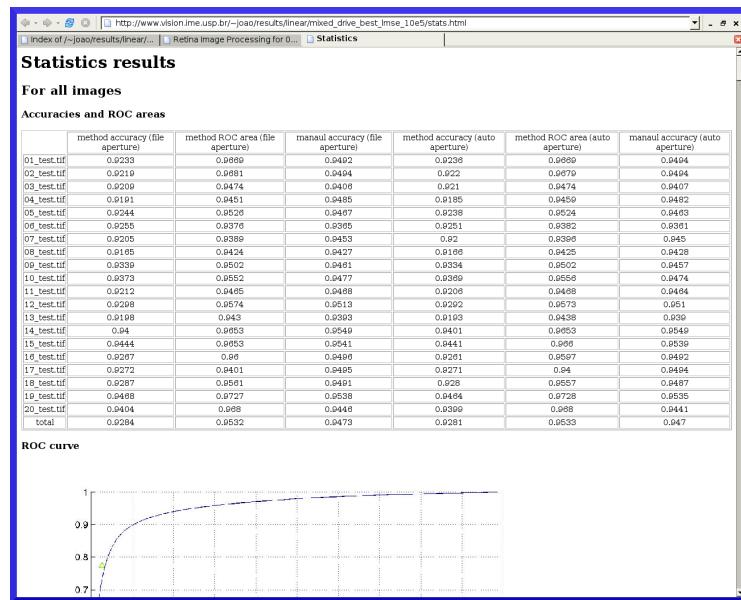
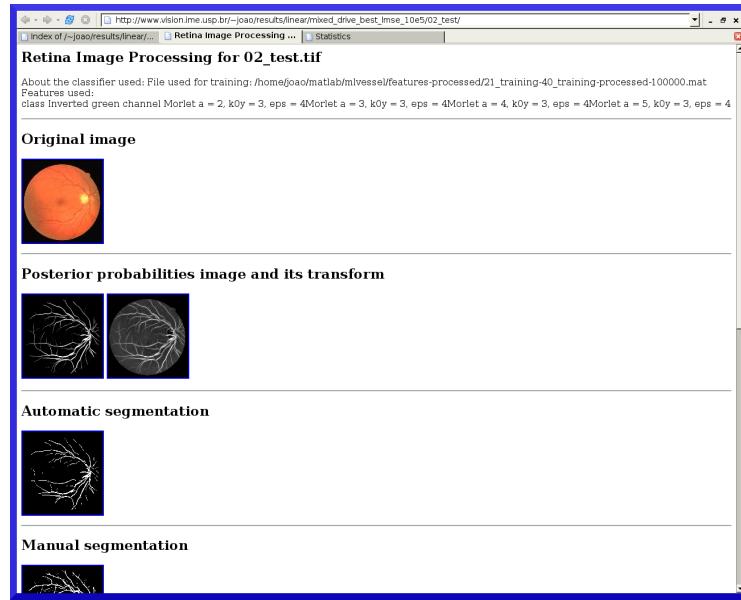
The `mlvesselinstall.m` script compiles the C code in the `skel` module, records the installation directory for future use, adds the `gui` and `tests` modules to the MATLAB path and creates shortcut files for starting the software. You might want to add the `gui` and `tests` modules permanently to the paths in `startup.m`. Alternatively, the software may be started through the shortcuts `mlvessel.bat` and `mlvesselgui.bat` on Windows or `mlvessel.sh` and `mlvesselgui.sh` on Unix platforms, which should add the paths appropriately on startup. It is important to note that the package requires standard MATLAB toolboxes in able to run properly, such as the image processing toolbox.

### 1.4.3 Command Line Interface

All of the package's features are available through functions accessible from the command prompt of the MATLAB environment. Some of the features, such as the `stats` module for generation of statistics from results, are currently only available through function invocation, as the GUI has not yet been developed for them. The tests' outputs are saved as images, `.mat` files, and HTML pages as illustrated in Figure 1.8. The major experiments listed below serve as good starting points and examples of use:

```
>> testmixed(someconfig): experiments for separate train and test sets, as used  
here for the DRIVE database;  
  
>> testleaveoneout(someconfig): leave-one-out experiments, as used here for  
the STARE database.
```

`someconfig` describes a structure containing all experiment configurations, including testing and training image file names, classifier specifications, and pixel features to be used. Examples of configuration structures are in `src/tests/driveconfig.m` and `src/tests/stareconfig.m`. Images for tests were not included in the package, but can be downloaded directly from the DRIVE and STARE databases.

**FIGURE 1.8**

Examples of HTML outputs produced using the **mlvessel** package. Top: segmentation results for an image from the DRIVE database. Bottom: statistics and ROC graphs for results on the DRIVE database.

#### 1.4.4 Graphical User Interface

A GUI has been implemented in MATLAB for preliminary testing by users. It can be initialized through its main window (Figure 1.9(a)) in `guemain.m`, which provides access to all others. The interface currently comprehends:

- visualizing and saving images composed of pixel features;
- specifying, configuring, training, and saving classifiers;
- opening and visualizing classifiers and applying them for image segmentation.

Following is a brief description of GUI use, which illustrates the supervised classification framework, since feature generation and classifier training to image segmentation.

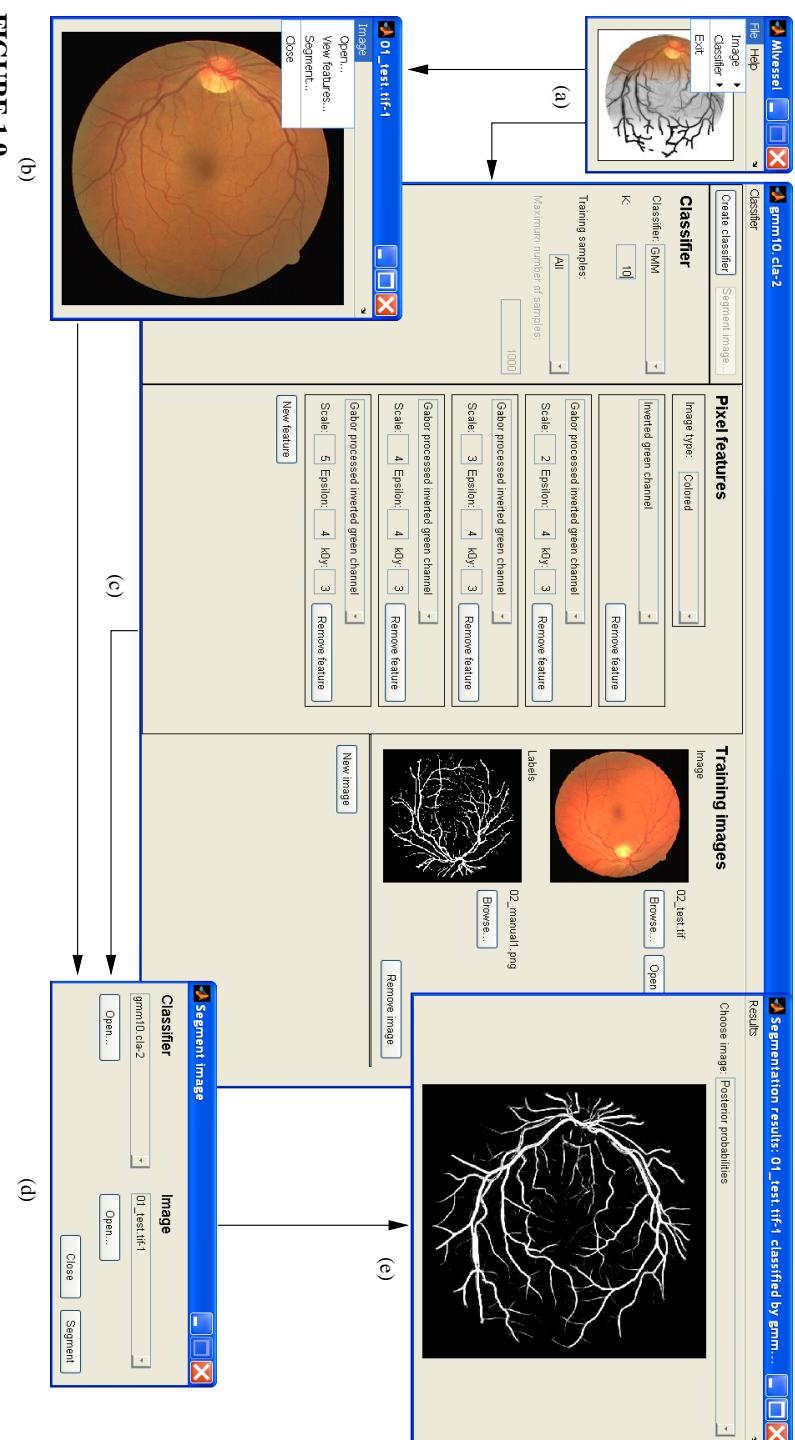
##### 1.4.4.1 Visualizing Pixel Features

Images can be opened through the main window's menu at **File → Image → Open**. After opening an image, pixel features are generated and viewed by accessing **Image → View Features** from the image window menu (Figure 1.9(b)). The **Choose features** dialog will then appear so that features may be specified. First, the image type must be informed. Depending on image type, possible pixel feature choices are the inverted green channel (colored images), inverted grayscale channel (red-free images), or original grayscale channel (fluorescein angiograms) and each of these processed by the Gabor wavelet ( $M_\psi(\mathbf{b}, a)$ , as in Equation 1.29) for different possible choices of the parameters: vertical basic frequency  $k_y$ , elongation  $\epsilon$ , and scale  $a$  (see Equation 1.14). After confirming feature selection, they are generated (feature generation may take a few minutes), after which a window for visualization will appear, providing the option of saving pixel feature images.

##### 1.4.4.2 Creating and Saving a Classifier

Figure 1.9(c) illustrates a classifier window in which a set of classifier options has been chosen. The following steps should be followed in order to create a classifier.

1. In the main program window menu, choose **File → Classifier → New**. This will open a new classifier window.
2. Specify the classifier properties: type of classifier (GMM,  $k$ NN or LMSE), its parameters, and number of training samples.
3. Choose the image type (colored, red-free, or fluorescein angiogram) for the classifier and specify pixel features to be used.
4. Indicate training images and respective manual segmentations.
5. Click on **Create classifier**. Depending on the previous choices, the classifier creation process may take from minutes to hours.



**FIGURE 1.9**  
Windows and dialogs from the GUI illustrating the supervised image segmentation process: (a) main window; (b) image window; (c) classifier window; (d) image segmentation dialog; (e) segmentation result window.

Once the classifier is created, it is ready to be used to segment new images. Classifiers can be saved for future use through their menu at **Classifier** → **Save as** and should be loaded through the main window menu at **File** → **Classifier** → **Open**, which also allows all of the classifier's properties to be viewed.

#### 1.4.4.3 Applying a Classifier for Image Segmentation

In order to segment an image, a suitable classifier must be chosen. The **Segment image** dialog, shown in Figure 1.9(d), is used to match an image with a classifier for segmentation. It can be reached through image window menus at **Image** → **Segment Image** or classifier windows at the **Segment Image** button. The image chosen to be segmented should match the modality and approximate resolution of the images used to train the classifier, as the features generated for segmentation will be the same as specified during classifier creation. In this manner, the segmentation step does not require choosing features or any kind of parameter configuration. The result will then appear as a grayscale image representing posterior probabilities, along with the final segmentation (see Figure 1.9(e)), both of which can then be saved.

---

## 1.5 Experimental Results

Illustrative segmentation results for a pair of images from each of the DRIVE and STARE databases (produced by the GMM classifier with  $c = 20$  and  $N = 10^6$ ), along with corresponding manual segmentations, are shown in Figures 1.10 and 1.11. Figure 1.12 presents results for the same images and settings, but with the GMM classifier being trained on each of the DRIVE and STARE databases, while tested on the other. The results shown are images formed by estimated posterior probabilities of each pixel belonging to class  $C_1$  (vessel), as well as the final segmentation, produced by thresholding the posterior probabilities at  $p(C_1|v) > 0.5$ .

Table 1.2 presents comparative results for the three different classifiers tested. Areas under the ROC curves ( $A_z$ ) and accuracies are presented for the GMM classifier with different values of  $c$ ; kNN classifier with different values of  $k$  and the LMSE classifier. The results for this table were produced with classifiers trained with one hundred thousand samples ( $N = 10^5$ ). Diverse segmentation methods are compared in Table 1.3, that presents areas under the ROC curves and accuracies for the following methods: GMM classifier with  $c = 20$  and  $N = 10^6$ ; the same classifier being trained on each of the DRIVE and STARE databases and tested on the other (crossed test); filtering using a single Gabor wavelet scale; the matched filter of Chaudhuri *et al.* [10]; and the methods of Jiang *et al.* [40] and Staal *et al.* [68], as published in [68]. Accuracies are also presented for manual segmentations from human observers. ROC curves for the DRIVE and STARE databases produced using the GMM classifier with  $c = 20$  and  $N = 10^6$ , filtering using a single Gabor wavelet scale, the matched filter of Chaudhuri *et al.*, as well as points representing human observer

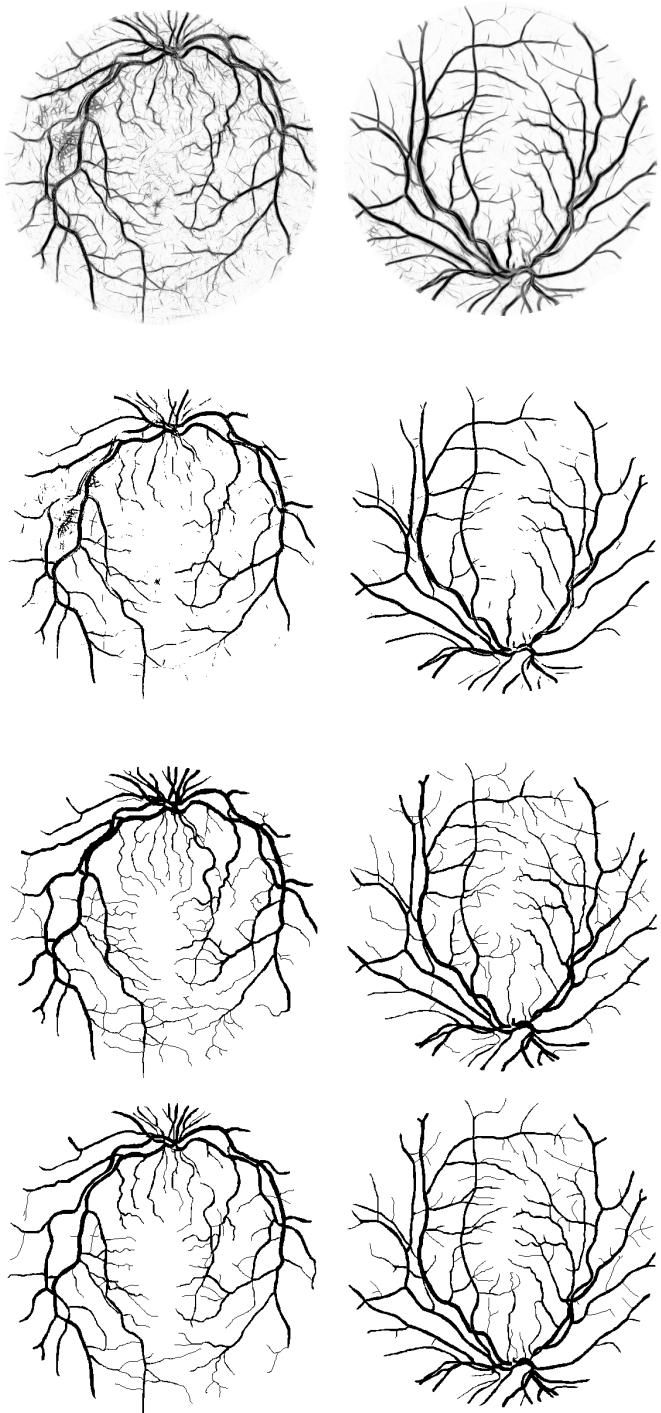


FIGURE 1.10

Results produced by the GMM classifier with  $c = 20$  and  $N = 10^6$ , along with respective manual segmentations, for two images from the DRIVE database. First column images are the estimated posterior probabilities of each pixel belonging to class  $C_1$  (vessel), while the second column images are the final segmentations, produced by thresholding posterior probabilities at  $p(C_1|v) > 0.5$ . The third and fourth columns are manual segmentations from sets A and B, respectively. The posterior probability images had their histograms modified for better visualization.

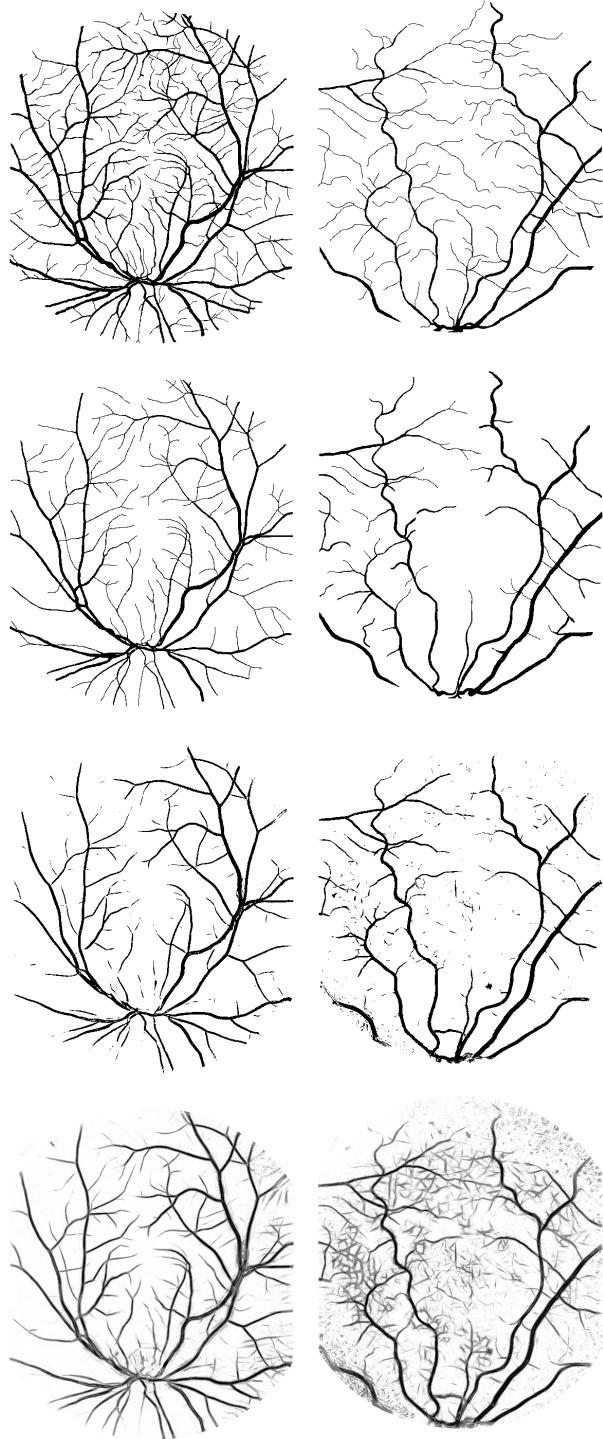


FIGURE 1.11

Results produced by the GMM classifier with  $c = 20$  and  $N = 10^6$ , along with respective manual segmentations, for two images from the STARE database. First column images are the estimated posterior probabilities of each pixel belonging to class  $C_1$  (vessel), while the second column images are the final segmentations, produced by thresholding posterior probabilities at  $p(C_1|v) > 0.5$ . The third and fourth columns are manual segmentations from the first and second observer, respectively. The posterior probability images had their histograms modified for better visualization. Top row images originate from a normal case, while the bottom ones originate from a pathological case.

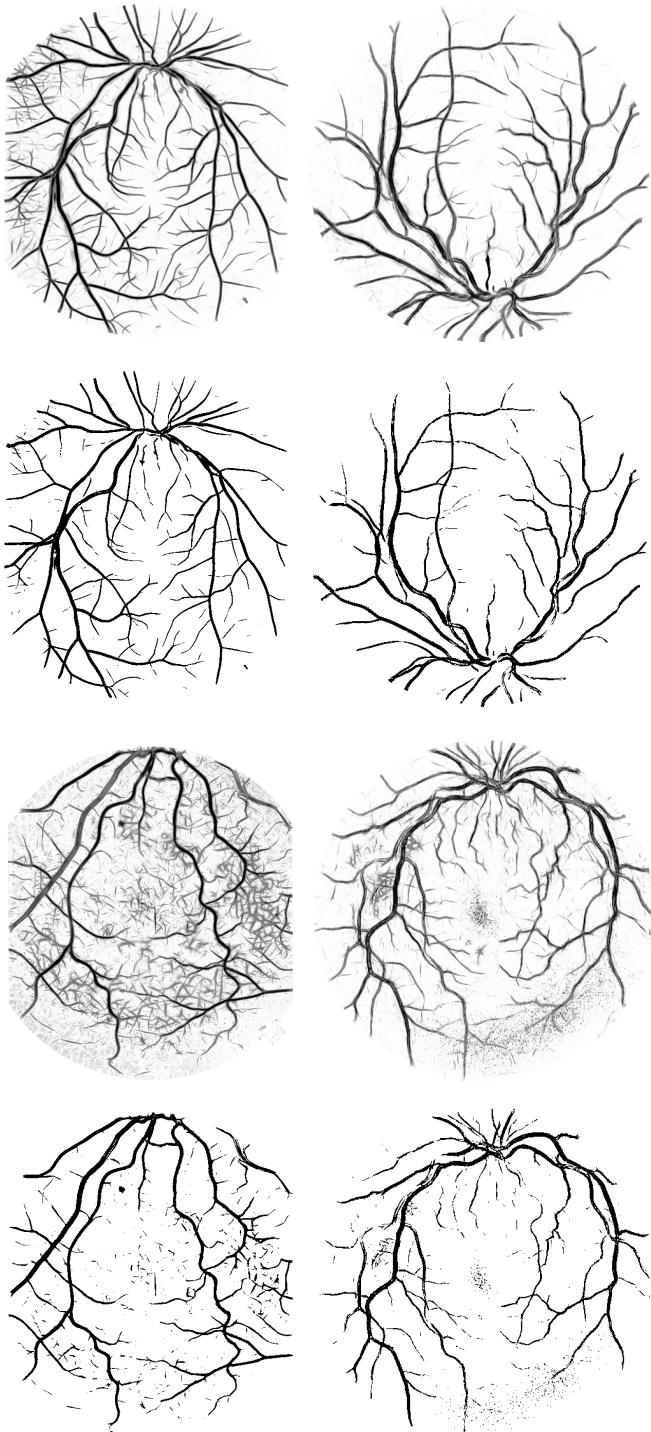


FIGURE 1.12

Results produced by training the GMM classifier with  $c = 20$  and  $N = 10^6$  on each of the DRIVE and STARE databases, while testing it on the other, for images presented in Figures 1.10 e 1.11. First and third column images are the estimated posterior probabilities of each pixel belonging to class  $C_1$  (vessel), while the second and fourth column images are the final segmentations, produced by thresholding posterior probabilities at  $p(C_1|v) > 0.5$ . The posterior probability images had their histograms modified for better visualization.

**Table 1.2:**  $A_z$  and accuracies for using different classifiers.  $A_z$  indicates the area under the ROC curve, while the accuracy is the fraction of correctly classified pixels. All classifiers were trained using one hundred thousand samples ( $N = 10^5$ ).

Segmentation method	Database			
	DRIVE		STARE	
	$A_z$	Accuracy	$A_z$	Accuracy
LMSE	0.9532	0.9284	0.9602	0.9365
$k$ NN, $k = 1$	0.8220	0.9201	0.8166	0.9273
$k$ NN, $k = 8$	0.9339	0.9446	0.9336	0.9460
$k$ NN, $k = 32$	0.9529	0.9473	0.9558	0.9480
$k$ NN, $k = 64$	0.9568	0.9475	0.9612	0.9482
$k$ NN, $k = 128$	0.9591	0.9479	0.9636	0.9480
$k$ NN, $k = 256$	0.9605	0.9478	0.9653	0.9478
$k$ NN, $k = 512$	0.9609	0.9476	0.9658	0.9472
GMM, $c = 1$	0.9287	0.9227	0.9409	0.9244
GMM, $c = 5$	0.9549	0.9419	0.9616	0.9437
GMM, $c = 10$	0.9582	0.9446	0.9657	0.9474
GMM, $c = 15$	0.9592	0.9454	0.9657	0.9469
GMM, $c = 20$	0.9600	0.9468	0.9666	0.9478
GMM, $c = 30$	0.9609	0.9468	0.9661	0.9476
GMM, $c = 40$	0.9610	0.9473	0.9665	0.9479

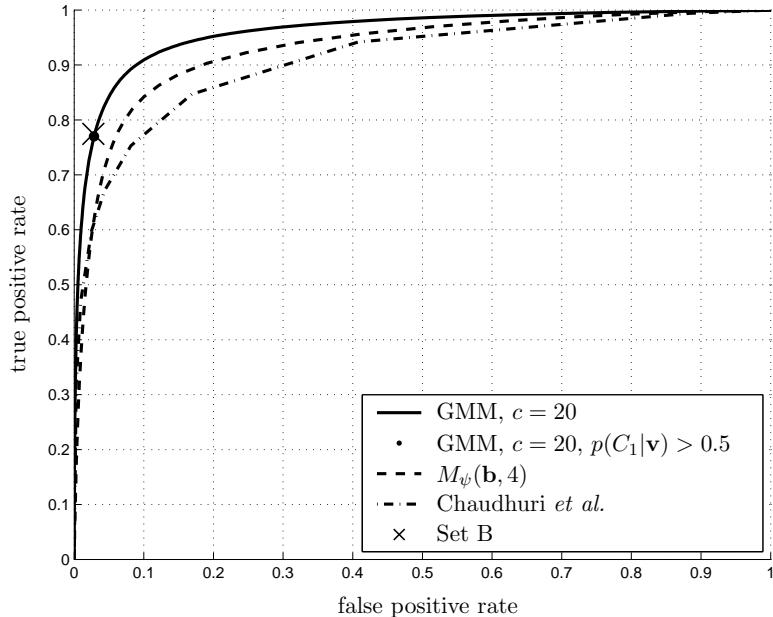
performance, are shown in Figures 1.13 and 1.14.

The Expectation-Maximization training process for the GMMs is computationally more expensive as  $c$  increases, but can be done off-line, while the classification phase is fast. The  $k$ NN classifier, implemented in a straightforward exhaustive manner, doesn't go through training, but has a very demanding classification phase. In turn, the LMSE classifier is both fast in training and classification, but provides poorer results, as shown in Table 1.2. Fixing the dimension of the feature space, classification of an image's pixel feature vectors using the GMM classifier is fast, taking time  $O(cP)$ , where  $P$  is the total of pixels in the image. Comparatively, the  $k$ NN classifier takes time  $O(NP)$  solely for the calculation of the distances used, where  $N$  is the number of training samples. The process of feature generation is basically the calculation of the wavelet coefficients, which is done by a series of correlations. By using the Fast Fourier Transform and the Fourier definition of the wavelet transform (Equation 1.15), these are done in  $O(P \log_2 P)$  [30]. A non-optimized MATLAB implementation was used for tests. On an AMD Athlon XP 2700+ PC (2167 MHz clock) with 1 GB memory, feature generation for typical images from the DRIVE and STARE databases takes about 3 minutes. Estimation of the GMM parameters with  $c = 20$  and  $N = 10^5$  takes up to 2 hours (though this would speed up considerably with an optimized implementation), while the classification of an image's pixels using this classifier takes less than 10 seconds. On the other hand, classification of an image's pixels with the  $k$ NN classifier with the same number of training samples takes about 4 hours.

Note that the  $A_z$  and accuracy for the GMM classifier increased with  $c$  (Table 1.2).

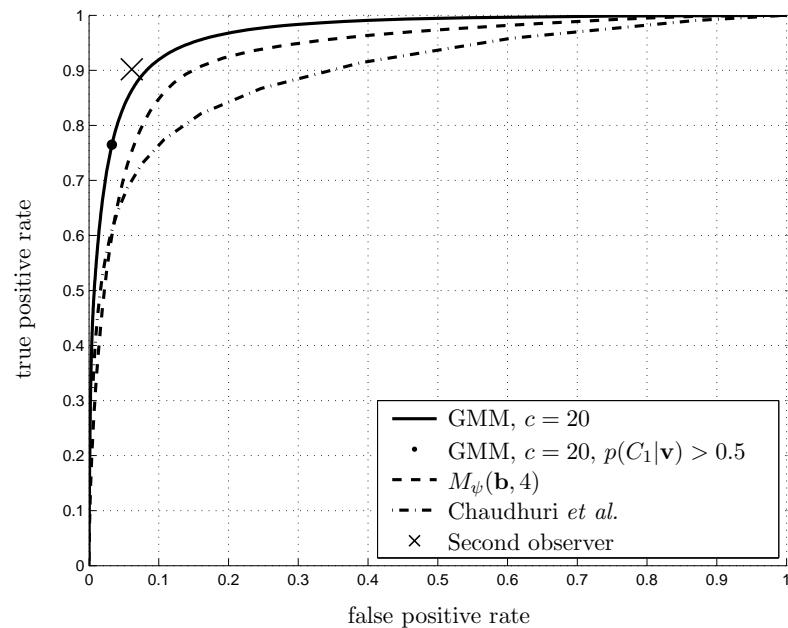
**Table 1.3:**  $A_z$  and accuracies for different segmentation methods and also a second human observer.  $A_z$  indicates the area under the ROC curve, while the accuracy is the fraction of correctly classified pixels.

Segmentation method	Database			
	DRIVE		STARE	
	$A_z$	Accuracy	$A_z$	Accuracy
GMM, $c = 20, N = 10^6$	0.9614	0.9466	0.9671	0.9480
Crossed test, GMM, $c = 20, N = 10^6$	0.9522	0.9404	0.9601	0.9328
$M_\psi(\mathbf{b}, 4)$	0.9312		0.9351	
Chaudhuri <i>et al.</i> [10]	0.9103		0.8987	
Jiang <i>et al.</i> [40]	0.9327	0.8911	0.9298	0.9009
Staal <i>et al.</i> [68]	0.9520	0.9441	0.9614	0.9516
Second observer		0.9473		0.9349



**FIGURE 1.13**

ROC curves for results from the DRIVE database produced using the GMM classifier with  $c = 20$  and  $N = 10^6$ , filtering using a single Gabor wavelet scale ( $M_\psi(\mathbf{b}, 4)$ ) and the matched filter of Chaudhuri *et al.*. Point marked as  $\bullet$  corresponds to classifications produced by applying the threshold  $p(C_1|\mathbf{v}) > 0.5$  and the point marked as  $\times$  corresponds to set B, the second set of manual segmentations. The GMM classifier has  $A_z = 0.9614$ .

**FIGURE 1.14**

ROC curves for results from the STARE database produced using the GMM classifier with  $c = 20$  and  $N = 10^6$ , filtering using a single Gabor wavelet scale ( $M_\psi(\mathbf{b}, 4)$ ) and the matched filter of Chaudhuri *et al.* Point marked as  $\bullet$  corresponds to classifications produced by applying the threshold  $p(C_1|\mathbf{v}) > 0.5$  and the point marked as  $\times$  corresponds to the second observer's manual segmentations. The GMM classifier has  $A_z = 0.9671$ .

After some value of  $c$ , though, it is expected that the performance decline, since the model will probably be excessively adjusted (over-fit) to the training data (see Section 1.2.6). In a similar fashion, the  $A_z$  of the  $k$ NN classifier increased with  $k$ , but after a certain value a loss in the spatial precision of estimates is expected, which would also lead to performance decline (see Section 1.2.7). This kind of behavior is observed by the accuracy decrease of the  $k$ NN classifier for  $k \geq 128$ . The decrease is caused by the small number of training samples coupled with high values of  $k$ , leading to an excessive amount of test samples being attributed to the nonvessel class. As ROC analysis is invariant to skews in class distribution, the  $A_z$  values were not affected, though it is still expected they decrease for larger values of  $k$ .

The manual segmentations help give an estimate of the variance between human observers. On the DRIVE database, the GMM classifier ROC curve is very close to the point representing the second set of manual segmentations (Figure 1.13). The methods presents very good results on normal, well-behaved images, but pathological or poorly illuminated images (as discussed below) lower the overall performance. It is curious to note that, on the STARE database, the accuracy of some methods is higher than that of the second observer (Table 1.3). The second observer's manual segmentations contain much more of the thinnest vessels than the first observer (lowering their accuracy), while the method, trained by the first observer, is able to segment the vessels at a similar rate. Nevertheless, the ROC graph (Figure 1.14) still reflects the higher precision of the second observer, due to some difficulties found by the method, as discussed below.

Visual inspection reveals typical problems that must be solved by future work. The major errors are in false detection of noise and other artifacts. False detection occurs in some images for the border of the optic disc, hemorrhages, microaneurysms, exudates and other types of pathologies that present strong contrast. Another difficulty is the detection of the thinnest vessels that are barely perceived by human observers. Also, the method did not perform well for very large variations in lighting throughout an image, but this occurred for only one image out of the 40 tested from both databases.

The crossed test, in which images from each database were segmented by a classifier trained on the other, presented slightly worse results than the standard experiments (Table 1.3 and Figure 1.12). Though the databases are similar, there is a difference in the typical vessel widths found in each database's images, which contributed significantly to the performance loss. In the results from the DRIVE database, the thinner vessels are poorly detected and there is an increase in noise appearing as false positives. Results from the STARE database present much more noise and also pathologies detected as false positives. While the performance difference is not large, this shows that even for the simple vessel structures there is a certain dependence of the method on the training set. The classifier is capable of adjusting to specific acquisition conditions, but not able to generalize perfectly to others.

## 1.6 Conclusion

### 1.6.1 Summary

The Gabor wavelet shows itself efficient in enhancing vessel contrast while filtering out noise, showing better performance than the matched filter of Chaudhuri *et al.* [10]. Information from wavelet responses at different scales is combined through the supervised classification framework, allowing proper segmentation of vessels of various widths. Of the three classifiers tested, the LMSE presented the worst performance, but with fast training and classification phases. The  $k$ NN classifier showed good performance, but with a slow classification phase, complicating its use in interactive applications. Finally, the GMM classifier has a demanding training process, but guarantees a fast classification phase and good performance, similar to that of the  $k$ NN. Feature generation consists of computing a series of wavelet coefficients and can be implemented efficiently, resulting in an approach that can be included in an interactive tool. An open-source prototype of an interactive software for vessel segmentation – based on a graphical user interface that assists the process of classifier training and vessel segmentation – was described, including details on the most important steps.

The supervised classification framework demands use of manual labelings, but allows classifiers to be trained for different image modalities, possibly adjusted to specific camera or lighting conditions and are otherwise automatic, i.e., adjustment of parameters or user interaction is not necessary. The experiments in which the GMM classifier was trained on each of the DRIVE and STARE databases and tested on the other showed the dependence of the method on the training set. As an alternative training process, labeled samples from the image to be segmented provided by a user may be used, allowing image peculiarities to be incorporated [13]. Recent supervised methods have shown good results with respect to ROC analysis, even though they are restricted to local vessel detection [60, 68]. Supervised classification avoids formulation of complex rules that deal with specific situations, as the complexity can be directly incorporated by the classifier. Given the conceptual simplicity of the approach presented here, it could easily be applied to the segmentation of other oriented structures, such as neurons or roads in aerial photography, in spite of no experiment having been performed in this sense.

### 1.6.2 Future Work

Analyzing the method and the experimental evaluation performed, various points can be thought of to be better studied and worked upon. In particular, visual inspection of results reveals difficulties that suggest ideas for the future evolution of methods.

The method did not respond well to large lighting variations within images. These variations could be compensated in a pre-processing step or otherwise somehow incorporated into pixel feature vectors. The Gabor wavelet presented in Equation 1.14,

when adjusted to low frequencies, does not present zero mean. Thus, in the implementation evaluated, it was not totally insensitive to background variations. This could be corrected with the introduction of the wavelet's correction term [1]. The term, however, when added to low frequency wavelets, leads to loss in spatial precision, being prejudicial when analyzing the modulus. To overcome this, it is possible to separately analyze the real and imaginary parts of the coefficients (or equivalently, modulus and phase) or modify the wavelet appropriately, as was done in [35].

In order to distinguish borders and other artifacts from vessels, information about the presence of edges could be explicitly incorporated, possibly through the inclusion of new wavelet coefficients in the feature vectors. The wavelet transform is capable of providing a local representation of the image, revealing directional structures, as the vessels in this work. Representation using the Gabor wavelet is done by deriving frames, that provide complete but redundant representations [1, 46]. It is worth noting that there are fast algorithms for calculation of 2-D CWT frames, which could speed up considerably the application of the method [73].

The training approach that uses samples from the image being segmented provided by an operator may present better results, since image peculiarities are incorporated by the classifier. It can be difficult or tiresome for a user to interactively supply a representative sampling. Being so, it would be interesting to study a strategy to take advantage of both the available samples from other images as well as those from the image to be segmented. In this manner, with minimal user effort, information about specific image conditions and the large amount of examples from other images would be taken into account.

A drawback of the approach presented here is that it only takes into account information local to each pixel through the wavelet transform, ignoring useful information from image shapes and structures. The segmentation results can be slightly improved through a post-processing of the segmentations for removal of noise and inclusion of missing vessel pixels as in [45]. An intermediate result of the method is the intensity image of posterior probabilities, to which global vessel segmentation algorithms could be applied, providing more precise results. Many approaches have been studied that take into account shape and structure for vessel segmentation, such as tracking [8, 29, 70], threshold probing [38, 40], region growing [53] and deformable models [48, 56].

Depending on the application, different evaluation methods become more appropriate [7]. For example, the evaluation of the vasculatures' skeletons or tracings would not take into account vessel widths, but could measure other qualities such as the presence of gaps and detection of branching and crossing points. Vessel widths are fundamental in change and pathology detection, but can be discarded in applications like image registration or fractal analysis of the vasculature. In [47], true and false positives are defined from skeletons in the following manner: if a pixel from some method's skeleton is at a distance inferior to 3 pixels to a pixel from a manual segmentation, it is counted as true positive; otherwise, it is a false positive. In [67], a similar approach is presented, in which thinned versions of the manual segmentations are used and, additionally, true and false negatives are defined. ROC analysis using complete vessel segmentations, on the other hand, takes vessel widths into account,

in such a manner that wide vessels have a much larger influence than thin ones. Another interesting form of evaluation would be directly through an application, such as in detection of neovascularization by means of analysis and classification of the vessel structure [9]. Qualitative visual assessment is still very important for segmentation evaluation, allowing the identification of strengths and weaknesses of each method and localizing specific regions in which a given method performs better or worse. Still, it is a laborious and subjective analysis. Though very good ROC results are presented, visual inspection (as discussed in Section 1.5) shows typical difficulties of the method to be worked on.

A major difficulty in evaluating the results is the establishment of a reliable ground truth [27]. Human observers are subjective and prone to errors, resulting in large variability between observations. Thus, it is desirable that multiple human-generated segmentations be combined to establish a ground truth, which was not the case in the analysis presented. Moreover, in an ideal scenario, manual segmentations would be provided by ophthalmologists, so that the classifier be calibrated accordingly. As an alternative to manually segmenting several images, synthetic images might be used. The images would then have segmentations known beforehand and problems of human observers such as subjectivity, errors and workload would be avoided. The problem with synthetic images is that it is practically impossible to faithfully reproduce the complexity of retinal image formation, yielding this approach inadequate for a complete evaluation. Nonetheless, it could be applied in partial experiments, such as in measuring the adequacy of vessel models or robustness under different kinds of noise. For future work, a more complete evaluation is desirable: it would be interesting to evaluate performance on different image modalities; hypothesis tests could be performed for method comparison (as done in [58, 67, 68]); and confidence bands for ROC curves could also be presented [22].

Properties of the retinal vasculature can be quantified through morphological analysis techniques, assisting in the diagnosis of diabetic retinopathy. Fractal [9] and multifractal [57] analysis provide numeric indicators of the extent of neovascularization, while vessel lengths, widths [61] and curvature [36] have been identified as important characteristics for diagnosing diabetic retinopathy. The segmentation method presented shows good results and large potential for improvement, being theoretically founded on the 2-D CWT and statistical classification. It is capable of producing results in reasonable time and lessens the need for user interaction, representing an effort towards automated retinal assessment. A new implementation of the method is being developed, that will be faster, have better usability and independent of the MATLAB platform, avoiding the need for licenses. It is expected that the softwares and graphical user interfaces developed evolve while being tested by final users, including non-specialized health workers, ophthalmologists and researchers.

---

### Acknowledgments

The authors are extremely grateful to H. F. Jelinek and M. J. Cree for their long-lasting collaboration. The authors would also like to thank J. J. Staal and colleagues and A. Hoover for making their databases publicly available, Dr. A. Luckie and C. McQuillin from the Albury Eye Clinic for providing fluorescein images used during research, and P. Mani for running the tests of the Chaudhuri *et al.* method as part of her B. E. degree final year project. JVBS is grateful to CNPq (131403/2004-4) and FAPESP (2006/56128-1). RMC is also grateful to CNPq (300722/98-2, 474596/2004-4 and 491323/2005-0) and FAPESP (2005/00587-5).

---

## **References**

- [1] J.-P. Antoine, P. Carette, R. Murenzi, and B. Piette. Image analysis with two-dimensional continuous wavelet transform. *Signal Processing*, 31:241–272, 1993.
- [2] J.-P. Antoine and R. Murenzi. Two-dimensional directional wavelets and the scale-angle representation. *Signal Processing*, 52:259–281, 1996.
- [3] J.-P. Antoine, R. Murenzi, and P. Vandergheynst. Directional wavelets revisited: Cauchy wavelets and symmetry detection in patterns. *Applied and Computational Harmonic Analysis*, 6:314–345, 1999.
- [4] F. Argoul, A. Arnéodo, J. Elezgaray, G. Grasseau, and R. Murenzi. Wavelet analysis of the self-similarity of diffusion-limited aggregates and electrodeposition clusters. *Physical Review A*, 41(10):5537–5560, 1990.
- [5] A. Arnéodo, N. Decoster, and S. G. Roux. A wavelet-based method for multifractal image analysis. I. Methodology and test applications on isotropic and anisotropic random rough surfaces. *The European Physical Journal B*, 15:567–600, 2000.
- [6] F. J. Ayres and R. M. Rangayyan. Performance analysis of oriented feature detectors. In *Proc. of the 18th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI)*, pages 147–154. IEEE Computer Society, 9–12 October 2005.
- [7] K. W. Bowyer and P. J. Phillips, editors. *Empirical Evaluation Techniques in Computer Vision*. IEEE Computer Society, 1998.
- [8] A. Can, H. Shen, J. N. Turner, H. L. Tanenbaum, and B. Roysam. Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms. *IEEE Transactions on Information Technology in Biomedicine*, 3(2):125–138, 1999.
- [9] R. M. Cesar-Jr. and H. F. Jelinek. Segmentation of retinal fundus vasculature in non-mydriatic camera images using wavelets. In J. Suri and T. Laxminarayan, editors, *Angiography and Plaque Imaging: Advanced Segmentation Techniques*, pages 193–224. CRC Press, 2003.
- [10] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Transactions on Medical Imaging*, 8:263–269, 1989.

- [11] J. Chen, Y. Sato, and S. Tamura. Orientation space filtering for multiple orientation line segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:417–429, 2000.
- [12] O. Chutatape, L. Zheng, and S. M. Krishnan. Retinal blood vessel detection and tracking by matched Gaussian and Kalman filters. In *Proc. of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, volume 20, pages 3144–3149, 1998.
- [13] D. J. Cornforth, H. F. Jelinek, J. J. G. Leandro, J. V. B Soares, R. M. Cesar-Jr., M. J. Cree, P. Mitchell, and T. R. J. Bossomaier. Development of retinal blood vessel segmentation methodology using wavelet transforms for assessment of diabetic retinopathy. *Complexity International*, 11, 2005. Available from: <http://www.complexity.org.au/ci/vol11/>. Cited May 2006.
- [14] L. F. Costa and R. M. Cesar-Jr. *Shape analysis and classification: theory and practice*. CRC Press, 2001.
- [15] M. J. Cree, J. J. G. Leandro, J. V. B. Soares, R. M. Cesar-Jr., G. Tang, H. F. Jelinek, and D. J. Cornforth. Comparison of various methods to delineate blood vessels in retinal images. In *Proc. of the 16th National Congress of the Australian Institute of Physics*, Canberra, Australia, 2005. Available from: <http://aipcongress2005.anu.edu.au/index.php?req=CongressProceedings>. Cited May 2006.
- [16] I. Daubechies. *Ten lectures on wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.
- [17] J. G. Daugman. Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, 20:847–856, 1980.
- [18] J. G. Daugman. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(7):1169–1179, 1988.
- [19] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1):1–38, 1977.
- [20] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley and Sons, 2001.
- [21] M. Farge. Wavelet transforms and their applications to turbulence. *Annual Review of Fluid Mechanics*, 24:395–457, 1992.
- [22] T. Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27:861–874, 2006.

- [23] R. S. Feris, V. Krueger, and R. M. Cesar-Jr. A wavelet subspace method for real-time face tracking. *Real-Time Imaging*, 10:339–350, 2004.
- [24] R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, and A. F. Frere. Analysis of asymmetry in mammograms via directional filtering with Gabor wavelets. *IEEE Transactions on Medical Imaging*, 20:953–964, 2001.
- [25] M. A. T. Figueiredo and A. K. Jain. Unsupervised learning of finite mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:381–396, 2002.
- [26] Free Software Foundation. Gnu general public license, 1991. Available from: <http://www.gnu.org/copyleft/gpl.html>. Cited July 2006.
- [27] K. H. Fritzzsche, A. Can, H. Shen, C.-L. Tsai, J. N. Turner, H. L. Tanenbaum, C. V. Stewart, and B. Roysam. Automated model-based segmentation, tracking, and analysis of retinal vasculature from digital fundus images. In J. Suri and S. Laxminarayan, editors, *Angiography and Plaque Imaging: Advanced Segmentation Techniques*, pages 225–297. CRC Press, 2003.
- [28] D. Gabor. Theory of communication. *Journal of the IEE*, 93:429–457, 1946.
- [29] L. Gang, O. Chutatape, and S. M. Krishnan. Detection and measurement of retinal vessels in fundus images using amplitude modified second-order Gaussian filter. *IEEE Transactions on Biomedical Engineering*, 49(2):168–172, 2002.
- [30] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison-Wesley, second edition, 2002.
- [31] P. Goupillaud, A. Grossmann, and J. Morlet. Cycle-Octave and related transforms in seismic signal analysis. *Geoexploration*, 23:85–102, 1984.
- [32] A. Grossmann. Wavelet transforms and edge detection. In S. Albeverio, P. Blanchard, M. Hazewinkel, and L. Streit, editors, *Stochastic Processes in Physics and Engineering*, pages 149–157. D. Reidel Publishing Company, 1988.
- [33] A. Grossmann and J. Morlet. Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM Journal on mathematical analysis*, 15:723–736, 1984.
- [34] J. A. Hanley and B. J. McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143:29–36, 1982.
- [35] J. D. Harrop, S. N. Taraskin, and S. R. Elliott. Instantaneous frequency and amplitude identification using wavelets: Application to glass structure. *Physical Review E*, 66(2):026703, August 2002.

- [36] W. E. Hart, M. Goldbaum, B. Côté, P. Kube, and M. R. Nelson. Measurement and classification of retinal vascular tortuosity. *International Journal of Medical Informatics*, 53:239–252, 1999.
- [37] R. Hill. Retina identification. In A. K. Jain, R. Bolle, and S. Pankanti, editors, *Biometrics: Personal Identification in Networked Society*, pages 123–141. Kluwer Academic Publishers, 1998.
- [38] A. Hoover, V. Kouznetsova, and M. Goldbaum. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 19:203–210, March 2000.
- [39] H. F. Jelinek, J. J. G. Leandro, R. M. Cesar-Jr., and M. J. Cree. Classification of pathology in diabetic eye disease. In *WDIC2005 ARPS Workshop on Digital Image Computing*, pages 9–13, Brisbane, Australia, 2005.
- [40] X. Jiang and D. Mojon. Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(1):131–137, 2003.
- [41] J. J. Kanski. *Clinical Ophthalmology: A systematic approach*. Butterworth-Heinemann, London, 1989.
- [42] V. Krueger and G. Sommer. Gabor wavelet networks for object representation. *Journal of the Optical Society of America*, 19:1112–1119, 2002.
- [43] M. Lalonde, L. Gagnon, and M.-C Boucher. Non-recursive paired tracking for vessel extraction from retinal images. In *Proc. of the Conference Vision Interface 2000*, pages 61–68, 2000.
- [44] J. J. G. Leandro, R. M. Cesar-Jr., and H. F. Jelinek. Blood vessels segmentation in retina: Preliminary assessment of the mathematical morphology and of the wavelet transform techniques. In *Proc. of the 14th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI)*, pages 84–90. IEEE Computer Society, 2001.
- [45] J. J. G. Leandro, J. V. B. Soares, R. M. Cesar-Jr., and H. F. Jelinek. Blood vessels segmentation in non-mydriatic images using wavelets and statistical classifiers. In *Proc. of the 16th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI)*, pages 262–269. IEEE Computer Society Press, 2003.
- [46] T. S. Lee. Image representation using 2D Gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):959–971, 1996.
- [47] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, and R. L. Kennedy. Measurement of retinal vessel widths from fundus images based on 2-D modeling. *IEEE Transactions on Medical Imaging*, 23(10):1196–1204, October 2004.

- [48] R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):158–175, 1995.
- [49] S. Mallat and W. L. Hwang. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, 38(2):617–643, 1992.
- [50] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(7):710–732, 1992.
- [51] B. S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):837–842, 1996.
- [52] D. Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 207:187–217, 1980.
- [53] M. E. Martínez-Pérez, A. D. Hughes, A. V. Stanton, S. A. Thom, A. A. Bharath, and K. H. Parker. Retinal blood vessel segmentation by means of scale-space analysis and region growing. In *Medical Image Computing and Computer-assisted Intervention (MICCAI)*, pages 90–97, 1999.
- [54] S. Marčelja. Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America*, 70(11):1297–1300, 1980.
- [55] MATLAB website at MathWorks, 2006. Available from: <http://www.mathworks.com/products/matlab>. Cited May 2006.
- [56] T. McInerney and D. Terzopoulos. T-snakes: Topology adaptive snakes. *Medical Image Analysis*, 4:73–91, 2000.
- [57] C. P. McQuillin, H. F. Jelinek, and G. Joss. Characterisation of fluorescein angiograms of retinal fundus using mathematical morphology: a pilot study. In *5th International Conference on Ophthalmic Photography*, page 152, Adelaide, 2002.
- [58] A. M. Mendona and A. Campilho. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE Transactions on Medical Imaging*, 25:1200–1213, 2006.
- [59] R. Murenzi. *Ondelettes multidimensionnelles et application à l'analyse d'images*. PhD thesis, Université Catholique de Louvain, 1990.
- [60] M. Niemeijer, J. J. Staal, B. van Ginneken, M. Loog, and M. D. Abràmoff. Comparative study of retinal vessel segmentation methods on a new publicly available database. In J. M. Fitzpatrick and M. Sonka, editors, *SPIE Medical Imaging*, volume 5370, pages 648–656, 2004.

- [61] L. Pedersen, B. Ersbøll, K. Madsen, M. Grunkin, U. Skands, M. Larsen, and N. Christoffersen. Quantitative measurement of changes in retinal vessel diameter in ocular fundus images. *Pattern Recognition Letters*, 21(13-14):1215–1223, 2000.
- [62] A. Pinz, S. Bernogger, P. Datlinger, and A. Kruger. Mapping the human retina. *IEEE Transactions on Medical Imaging*, 17:606–619, 1998.
- [63] F. Provost and T. Fawcett. Robust classification for imprecise environments. *Machine Learning*, 42(3):203–231, 2001.
- [64] O. Rioul and M. Vetterli. Wavelets and signal processing. *IEEE Signal Processing Magazine*, 8:14–38, 1991.
- [65] C. Sagiv, N. A. Sochen, and Y. Y. Zeevi. Gabor features diffusion via the minimal weighted area method. In *Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR)*, 2002.
- [66] J. V. B. Soares, J. J. G. Leandro, R. M. Cesar-Jr., H. F. Jelinek, and M. J. Cree. Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. *IEEE Transactions on Medical Imaging*, 25:1214–1222, 2006.
- [67] M. Sofka and C. V. Stewart. Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures. *IEEE Transactions on Medical Imaging*, 25:1531–1546, 2006.
- [68] J. J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken. Ridge based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, 2004.
- [69] S. Theodoridis and K. Koutroumbas. *Pattern Recognition*. Academic Press, USA, first edition, 1999.
- [70] Y. A. Tolias and S. M. Panas. A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering. *IEEE Transactions on Medical Imaging*, 17:263–273, April 1998.
- [71] M. Unser and A. Aldroubi. A review of wavelets in biomedical applications. *Proceedings of the IEEE*, 84:626–638, 1996.
- [72] D. Van De Ville, T. Blu, and M. Unser. Integrated wavelet processing and spatial statistical testing of fMRI data. *NeuroImage*, 23(4):1472–1485, December 2004.
- [73] P. Vandergheynst and J.-F. Gobbers. Directional dyadic wavelet transforms: design and algorithms. *IEEE Transactions on Image Processing*, 11(4):363–372, 2002.

- [74] T. Y. Wong, W. Rosamond, P. P. Chang, D. J. Couper, A. R. Sharrett, L. D. Hubbard, A. R. Folsom, and R. Klein. Retinopathy and risk of congestive heart failure. *Journal of the American Medical Association*, 293(1):63–69, 2005.
- [75] G. Yang and C. V. Stewart. Covariance-driven mosaic formation from sparsely-overlapping image sets with application to retinal image mosaicing. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 804–810, 2004.
- [76] F. Zana and J.-C. Klein. A multimodal registration algorithm of eye fundus images using vessels detection and hough transform. *IEEE Transactions on Medical Imaging*, 18:419–428, 1999.
- [77] F. Zana and J.-C. Klein. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *IEEE Transactions on Image Processing*, 10:1010–1019, 2001.
- [78] L. Zhou, M. S. Rzeszotarski, L. J. Singerman, and J. M. Chokreff. The detection and quantification of retinopathy using digital angiograms. *IEEE Transactions on Medical Imaging*, 13(4):619–626, 1994.

---

## **Index**

- Accuracy, 23–24, 35, 38  
Analyzing wavelet, 5, 6, 9  
Anisotropy, 8, 9, *see also* Elongation  
Antoine and Murenzi studies, 5, 8  
Area under the curve ( $A_z$ ), 24, 31, 35, 38  
 $A_z$ , *see* Area under the curve ( $A_z$ )  
Bayes's formula, 11  
Bayesian decision theory, 11  
Biometric identification and verification, 3  
Branching vessel points, 2, 3, 40  
Cesar, Jr. studies, 1–41  
Change detection, 1, 2, 40  
Chaudhuri *et al.* filter, 15, 23, 24, 31, 39  
Chaudhuri studies, 3, 15, 23, 31, 39  
Class-conditional probability density function, *see* Likelihood  
Classifier, GMM, 4, 12–13, 15, 22, 24, 26, 31, 35, 38, 39  
Classifier, *k*NN, 13–15, 22, 24, 26, 31, 35, 38, 39  
Classifier, LMSE, 14–15, 24, 26, 31, 35, 39  
Colored retinal photography, 1, 2, 17, 20, 21, 29  
Confusion matrix, 23  
Contingency table, *see* Confusion matrix  
Correction term, wavelet, 5, 9, 40  
Correlation filter, 3, 7, 35, *see also* Chaudhuri *et al.* filter  
Correlation matrix, 15  
Crossing vessel points, 2, 3, 40  
Daugman studies, 8  
Diabetic retinopathy, 1, 2, 21, 41, *see also* Proliferative diabetic retinopathy (PDR)  
Directional wavelets, 6, 8–9  
Distance measures, 13, 18  
Distance, Euclidean, 13  
Distance, Mahalanobis, 13  
DRIVE database, 20–22, 38  
Edge detection, 6, 7  
Elongation, 8, 9, 17, 29  
EM, *see* Expectation-Maximization (EM) algorithm  
Euclidean distance, 13  
Euclidean group, 6, 7  
Evaluation of vessel segmentation, 21–25, 40–41  
Expectation-Maximization (EM) algorithm, 12, 13  
False positive rate, 23, 25, 40  
Fast Fourier transform, 7, 35  
Feature normalization, 18–20  
FFT, *see* Fast Fourier transform  
Fluorescein angiography, 1, 2, 17, 20, 29  
Fourier analysis, 4  
Fractal analysis, 2, 5, 6, 40, 41  
Gabor studies, 4, 8  
Gabor transform, *see* Short-time Fourier transform (STFT)  
Gabor wavelet, 8–9  
Gaussian cross-sectional profiles, 2, *see also* Chaudhuri studies  
Gaussian functions, 2, 5, 9, 12

- Gaussian mixture model (GMM), 4, 12–13, 35, 39  
Global vessel segmentation, 3, 40  
GMM, *see* Gaussian mixture model (GMM)  
GMM classifier, 4, 12–13, 15, 22, 24, 26, 31, 35, 38, 39  
GNU General Public License (GPL), 25  
Graphical user interface (GUI), 25, 26, 29–31  
Grossmann and Morlet studies, 5  
Ground truth, 21, 23–25, 41  
Group, Euclidean, 6, 7  
GUI, *see* Graphical user interface (GUI)
- Hat wavelet, Mexican, 9  
HTML results pages, 26, 27  
Hypothesis testing, 24, 41
- Inner product, *see* Scalar product  
Interaction, user, 3, 20, 39–41, *see also* Graphical user interface (GUI)
- Jelinek studies, 2, 3, 15, 20, 41
- k*NN classifier, 13–15, 22, 24, 26, 31, 35, 38, 39
- Leandro studies, 3, 15, 17, 20, 40  
Length, vessel, 2, 41  
Likelihood, 11–13  
LMSE classifier, 14–15, 24, 26, 31, 35, 39
- Mahalanobis distance, 13  
Manual vessel segmentation, 3, 4, 11, 15, 20–21, 23–26, 29, 31, 38–41  
Matched filter, *see* Correlation filter  
MATLAB, 25–27, 29, 35, 41  
Mexican hat wavelet, 9  
*mlvessel* package, 25–31  
Monochromatic retinal photography, 1, *see also* Red-free retinal photography
- Morlet wavelet, 5, 8  
Morlet, Grossmann and, studies, 5  
Morphology of the retinal vasculature, 1–2, 41  
Mother wavelet, *see* Analyzing wavelet  
Multiresolution, 4  
Multiscale vessel detection, 3, 9, 15, 18, 21, 39  
Murenzi, Antoine and, studies, 5, 8
- Neovascularization, 2, 41  
Nonparametric classification, 13  
Nonstationary signals, 4–5  
Normalization, feature, 18–20
- Over-fitting, 13, 38
- Package, *mlvessel*, 25–31  
Parametric classification, 13, 14  
PDR, *see* Proliferative diabetic retinopathy (PDR)  
Pipeline architecture, 26  
Posterior probability, 11–13, 23, 24, 31, 40  
Prior probability, 11, 12, 24  
Proliferative diabetic retinopathy (PDR), 2  
Pseudoinverse, 15  
Public image databases, 20–21
- Receiver operating characteristic (ROC) analysis, 17, 23–25, 40, 41  
Red-free retinal photography, 1, 2, 17, 29  
Registration, 1–3, 40  
Retinal image acquisition, 1–2  
Retinal image databases, 20–21  
Retinal image variability, 2, 20, 38, 39  
ROC analysis, *see* Receiver operating characteristic (ROC) analysis
- Scalar product, 5, 6, 14, 18, 24  
Scale, wavelet, 3, 5, 9, 15, 18, 21  
Scale-space skeletonization, 26

- Short-time Fourier transform (STFT), 4  
Skeletonization, scale-space, 26  
Skeletonized vasculature, 2, 26, 40  
Soares studies, 1–41  
Software, *mlvessel*, *see mlvessel package*  
Staal studies, 4, 17, 20, 21, 24, 31, 39, 41  
STARE database, 20–22, 38  
Stationary signals, 4  
Statistical hypothesis testing, 24, 41  
STFT, *see Short-time Fourier transform (STFT)*  
Sum of squared error criterion, 14  
Supervised pixel classification, 3, 11, 15, 20, 29, 39  
Time-frequency uncertainty, 4–5, 8  
Tortuosity, vessel, 2  
Tracking, vessel, 2, 40  
True positive rate, 23, 25, 40  
User interaction, 3, 20, 39–41, *see also Graphical user interface (GUI)*  
Vasculature morphology, 1–2, 41  
Vasculature skeleton, 2, 26, 40  
Vessel branching points, 2, 3, 40  
Vessel crossing points, 2, 3, 40  
Vessel detection using wavelets, 6, 9, 15, 17–18  
Vessel length, 2, 41  
Vessel segmentation evaluation, 21–25, 40–41  
Vessel segmentation, manual, 3, 4, 11, 15, 20–21, 23–26, 29, 31, 38–41  
Vessel tortuosity, 2  
Vessel tracking, 2, 40  
Vessel width, 2, 21, 38–41  
Wavelets and wavelet transforms, 4–9  
analyzing wavelet, 5, 6, 9  
anisotropy, 8, 9, *see also Elongation*
- basis, 5, 6  
correction term, 5, 9, 40  
directional wavelets, 6, 8–9  
elongation, 8, 9, 17, 29  
Euclidean group, 6, 7  
frame, 8, 9  
Gabor wavelet, 8–9  
group, Euclidean, 6, 7  
Mexican hat wavelet, 9  
Morlet wavelet, 5, 8  
nonstationary signals, 4–5  
one-dimensional continuous transform, 4–5  
scale, 3, 5, 9, 15, 18, 21  
stationary signals, 4  
time-frequency uncertainty, 4–5, 8  
time-scale plane, 5  
two-dimensional continuous transform, 5–8  
two-dimensional discrete transform, 6  
vessel detection, 6, 9, 15, 17–18  
Width, vessel, 2, 21, 38–41  
Wilcoxon statistic, 24