

Retinal Vessel Centerline Extraction Using Multiscale Matched Filters, Confidence and Edge Measures

Michal Sofka*, *Student Member, IEEE*, and Charles V. Stewart, *Member, IEEE*

Abstract—Motivated by the goals of improving detection of low-contrast and narrow vessels and eliminating false detections at non-vascular structures, a new technique is presented for extracting vessels in retinal images. The core of the technique is a new likelihood ratio test that combines matched-filter responses, confidence measures and vessel boundary measures. Matched filter responses are derived in scale-space to extract vessels of widely varying widths. A vessel confidence measure is defined as a projection of a vector formed from a normalized pixel neighborhood onto a normalized ideal vessel profile. Vessel boundary measures and associated confidences are computed at potential vessel boundaries. Combined, these responses form a 6-dimensional measurement vector at each pixel. A training technique is used to develop a mapping of this vector to a likelihood ratio that measures the “vesselness” at each pixel. Results comparing this vesselness measure to matched filters alone and to measures based on the Hessian of intensities show substantial improvements both qualitatively and quantitatively. The Hessian can be used in place of the matched filter to obtain similar but less-substantial improvements or to steer the matched filter by preselecting kernel orientations. Finally, the new vesselness likelihood ratio is embedded into a vessel tracing framework, resulting in an efficient and effective vessel centerline extraction algorithm.

Index Terms—vessel extraction, vessel tracing, matched filters, likelihood ratio, retina images

I. INTRODUCTION

THE leading causes of retina-related vision impairment and blindness in the U.S. are diabetic retinopathy, age-related macular degeneration (AMD), and glaucoma. Current estimates indicate that they affect 4.1, 1.8, and 2.2 million adults over age 40, respectively, with projected increases to 7.2, 2.9, and 3.3 million individuals by the year 2020 [1]–[3]. It is believed that half of all blindness can be prevented [4], in part through periodic screening and early diagnosis. Automated image analysis techniques should play a central role because the huge volume of images precludes strictly manual analysis.

Reliable vessel extraction is a prerequisite for subsequent retinal image analysis and processing because vessels are the predominant and most stable structures appearing in the images. Many published algorithms for optic disc detection [5]–

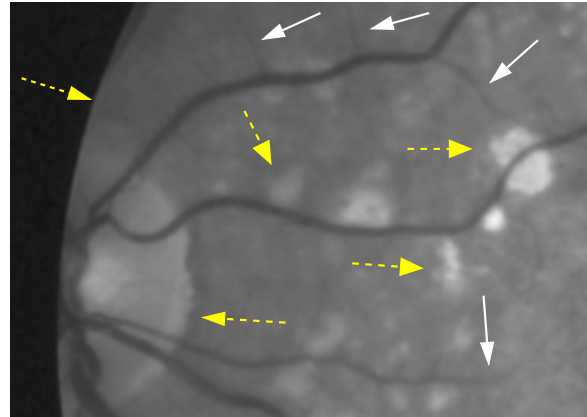


Fig. 1. Illustration of the challenges of retinal vessel extraction. Arrows drawn on the image [in yellow / with dashed lines] indicate pathologies, the boundary of the optic disc and the boundary of the retina, all of which tend to cause false positives in vessel detection. Arrows drawn [in white / with solid lines] highlight narrow or low-contrast vessels which are difficult to detect.

[7], image registration [8]–[14], change detection [15]–[21], pathology detection and quantification [22], tracking in video sequences [23]–[25], and computer-aided screening systems [26]–[29] depend on vessel extraction. The techniques published in the research literature in response to the importance of retinal vessel extraction may be roughly categorized into methods based on matched filters [24], [30], [31], adaptive thresholds [31], [32], intensity edges [33], [34], region growing [35], statistical inferencing [36], mathematical morphology [16], [35], [37], and Hessian measures [22], [38]–[40]. This wide range of techniques closely corresponds to the suite of methods that have been applied throughout the medical image analysis literature [41]. Of particular note, the recent literature has been dominated by Hessian-based methods because of their utility in characterizing the elongated structure of vessels [38], [42]–[45].

Several challenges of vessel extraction in retinal images are illustrated by the images shown in Figures 1 and 2. These challenges may be outlined as follows:

- There is a wide range of vessel widths — from less than a pixel to 12 pixels wide in the example shown.
- Vessels may be low contrast. The central intensity of some vessels differ from the background by as little as 4 grey levels, yet the background noise standard deviation is 2.3 grey levels. Narrow vessels often have the lowest contrast.
- A variety of structures appears in the images, including the retina boundary, the optic disc, and pathologies. The latter are particularly challenging for automatic vessel extraction

Manuscript received March 19, 2006; revised July 31, 2006. This work was supported in part by National Science Foundation Experimental Partnerships under Grant EIA-0000417, in part by the Center for Subsurface Sensing and Imaging Systems under the Engineering Research Centers Program of the National Science Foundation (Award EEC-9986821), and in part by the General Electric Global Research Center. *Asterisk indicates corresponding author.*

M. Sofka is with the Department of Computer Science, Rensselaer Polytechnic Institute, Troy, New York 12180–3590 (email:sofka@cs.rpi.edu).

C.V. Stewart is with the Department of Computer Science, Rensselaer Polytechnic Institute, Troy, New York 12180–3590 (email:stewart@cs.rpi.edu).

Digital Object Identifier 10.1109/TMI.2006.884190

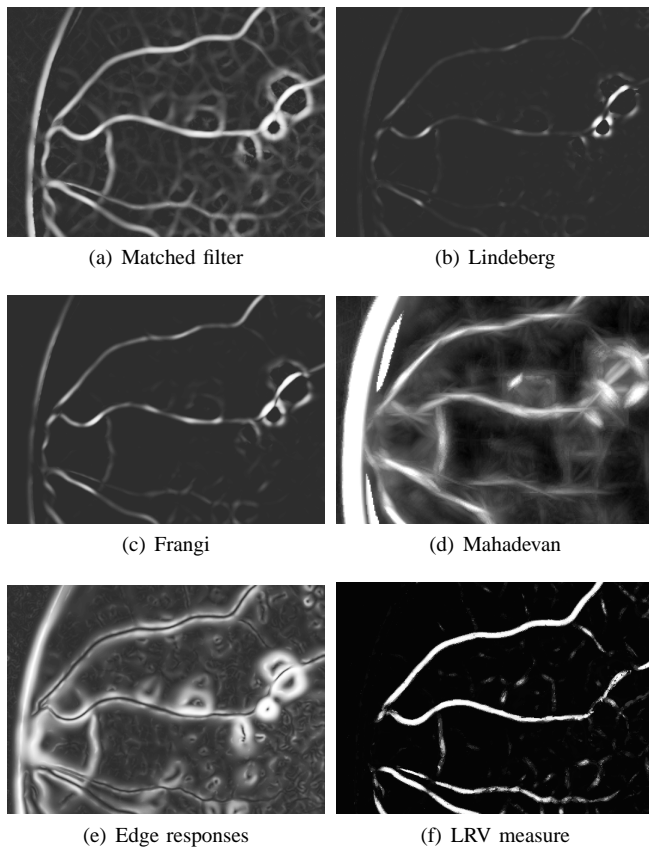


Fig. 2. Illustration of the challenges of retinal vessel extraction on the image shown in Figure 1. **(a)** Matched filter response based on a Gaussian vessel profile; **(b)** Lindeberg's ridge measure [49]; **(c)** Frangi's Hessian-based vesselness measure [42]; and **(d)** Mahadevan's robust model-based measure [36]. All measures show spurious responses to non-vascular structures such as the fundus boundary, the optic disc boundary and pathologies. Edge response image **(e)** shows numerous edges to which the vesselness measures are sensitive. The proposed Likelihood Ratio Vesselness (LRV) measure **(f)** suppresses the response to the bright pathology on the right, highlights thin and low-contrast vessels (e.g. the vessel above the same pathology), and completely removes the boundary of the retina. (Response images in **(a)**-**(f)** were stretched to $[0, 255]$ for display.)

because they may appear as a series of bright spots, sometimes with narrow, darker gaps in between.

- Wider vessels sometimes have a bright strip running down the center (the "central reflex"), causing a complicated intensity cross-section. Locally, this may be hard to distinguish from two side-by-side vessels.

Our focus in this paper is on techniques needed to solve the first three problems — detecting low-contrast vessels and narrow vessels, while avoiding false responses near pathologies and other non-vascular structures.

The limitations of some existing methods are illustrated in Figure 2. All these "vesselness measures" produce substantial responses to non-vascular structures such as the optic disc and the pathologies. All measures produce stronger responses at the boundary of the retina, near the optic disc, and along central pathological structures than for the thin and low-contrast vessels. This problem occurs with measures designed to compensate for responses to edges [42], [43], [46], [47] without directly finding edges, using only the measurement at the hypothetical vessel center. The problem also appears in the results of edge-based

tracing techniques [33], [48] where the vesselness measure is the average of the left and right edge responses.

The main contribution of this paper is the development of an enhanced vesselness measure that addresses the issues illustrated in Figure 1. Motivated by the apparent effectiveness of the matched filter in highlighting low-contrast and narrow vessels and by recent success in using matched filters for retina vessel segmentation [31], we introduce a multi-scale matched filter for vessels, using an appropriate normalizing multiplier to allow the combination of responses across scales. We then augment the matched-filter responses with a new vessel "confidence" measure, analogous to the edge-based measure presented in [50]. It determines how closely an image region follows an ideal vessel profile. Importantly, unlike the matched filter, this measure is independent of amplitude. To these vessel response and confidence values we add edge detection filter responses and confidences taken from the boundary of the purported vessel. This produces a six degree-of-freedom measurement vector at each pixel. Then, we use a training technique to develop a mapping from this vector to a single likelihood ratio that serves as the final "vesselness" measure. This gives a measure at each pixel which may be used either for segmentation of vessel pixels or for identifying the centerline vessel pixels and vessel widths when used in combination with non-maximum suppression. We focus on the latter because the measures are designed to have maximum response along the centerline of the vessel, and because this provides a more compact, geometric description of the vessels than segmentation alone.

We show that the new Likelihood Ratio Vesselness (LRV) measure outperforms the multiscale matched-filter and existing Hessian-based measures using quantitative analysis and using visual inspection of results on retinal images. We also show how the multiscale Hessian may be used to steer the matched filter by selecting its orientation at each pixel and scale. This way, the matched filter is applied only once at each pixel and scale, eliminating much of the excess computation. Alternatively, the Hessian may be used in place of the matched-filter in the LRV measure, producing substantial improvements in Hessian-based vesselness measures. The advantage of this is a lower overall computational cost than the matched-filter-based measure in exchange for a slight decrease in effectiveness.

The new vesselness measure may be used as a low-level operator in many existing vessel extraction techniques for the retina [35], [38]–[40], [51] and in related applications [41], [52]–[55]. By employing a stronger preliminary indication of the locations of vessels, each of these techniques should produce better overall results. As an example of doing this, the last part of this paper shows how to replace the parallel-edge measure with the LRV measure in an efficient, exploratory vessel tracing algorithm [33], [48] which extracts the vessel centerlines and widths without having to analyze each pixel of the image. We compare this new LRV tracing algorithm side-by-side with the result of applying the LRV measure at each pixel and with parallel-edge tracing algorithm [33], [48]. The paper concludes with a summary of remaining issues in retinal vessel extraction.

The paper is organized as follows. In Section II, we give an overview of vessel extraction methods in the research literature with focus on matched filters and the Hessian. We derive the

vessel matched filter in scale space in Section III, and then define the confidence and edge measures in Section IV. We show how these are combined into the single vesselness measure — our LRV measure — using a training technique in Section V. In Section VI, we present a summary of our evaluation technique and extensive experimental results. We propose a tracing version of the LRV measure in Section VII and report experimental results using it in Section VIII. We conclude the paper in Section IX.

II. BACKGROUND

Our background discussion is focused on the core “vesselness” measures that have been proposed in the literature.

Several articles have introduced techniques for vessel or ridge extraction based on the eigen-decomposition of the Hessian computed at each image pixel [38], [42], [49], [56]–[60]. We choose two specific measures here for our analysis, [42] and [49], which have been applied in a number of papers [38]–[40], [44], [45], [55]. They both start from the definition of the *scale-space representation* $L : \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R}$

$$L(x, y; t) = g(x, y; t) * f(x, y), \quad (1)$$

where $g(\cdot; t)$ is a Gaussian function with variance t , f is an image, (x, y) is a pixel location, and $*$ represents the convolution operation [59]. The Hessian of an intensity image in scale space can be obtained at each point by computing

$$H(x, y) = \begin{bmatrix} \frac{\partial^2 L}{\partial x^2} & \frac{\partial^2 L}{\partial x \partial y} \\ \frac{\partial^2 L}{\partial y \partial x} & \frac{\partial^2 L}{\partial y^2} \end{bmatrix} = \begin{bmatrix} L_{xx} & L_{xy} \\ L_{yx} & L_{yy} \end{bmatrix}. \quad (2)$$

The partial derivatives are computed by convolving the image f with a derivative-of-Gaussian kernel. Frangi et al. [42] proposed computing the ratio of the eigenvalues $|\lambda_1| \leq |\lambda_2|$ of the Hessian, as a ridgeness score $\mathcal{R}_B = \lambda_1/\lambda_2$, computing the Frobenius norm, \mathcal{S} , of the Hessian to measure overall strength, and combining these into a vesselness measure. This measure, specialized to 2D images where vessels are darker than the background, is

$$\mathcal{V}_0 = \begin{cases} 0 & \lambda_2 > 0 \\ \exp\left(-\frac{\mathcal{R}_B^2}{2\beta^2}\right) \left(1 - \exp\left(-\frac{\mathcal{S}^2}{2c^2}\right)\right) & \text{otherwise} \end{cases} \quad (3)$$

with parameters $\beta = 0.5$ and c equal to half of the maximum Frobenius norm of the Hessian. An example vesselness image combined across scales is shown in Figure 2(c).

In Lindeberg [49] a ridge point is defined as a location for which the intensity assumes a local maximum (minimum) in the direction of the main principal curvature. Ridge strength is evaluated as

$$\mathcal{A}_{\gamma\text{-norm}} L = t^{2\gamma} \left((L_{xx} - L_{yy})^2 + 4L_{xy}^2 \right), \quad (4)$$

where L_{xx} , L_{yy} and L_{xy} are from the Hessian and $\gamma = 3/4$. An example response image computed across scales is shown in Figure 2(b). In the context of retinal image analysis, Staal et al. [39] described a retinal vessel segmentation algorithm that combines a grouping technique with low-level, Hessian-based classification of pixels.

Chaudhuri [30] et al. used a matched filter to detect retinal ves-

sels modeled as piecewise linear segments with Gaussian cross sections. In Hoover’s algorithm [31] a piece of the blood vessel network is hypothesized by probing an area of the matched-filter response image and, while doing so, the threshold on the matched filter is iteratively decreased. Gang et al. [61] proposes an amplitude-modified second-order derivative of a Gaussian filter to detect vessels at scales that match their widths. Steerable filters are introduced in [62] and extended in [63]. In particular, Freeman [62] et al. shows an example of applying steerable filters to vessel enhancement.

Many techniques in the literature propose to detect vessels by “tracing” along vessel centerlines starting from seed points. Aylward and Bullitt [38] apply Hessian-based vesselness measures in a ridge-traversal method. The exploratory algorithm of Can et al. [33], extended in Fritzsche [48], uses a set of two oriented edge detection kernels to detect parallel vessel boundaries in tracing retinal images.

A variety of other methods has been introduced specifically for retinal vessel extraction. In [36] a model-based technique is proposed for vessel detection in noisy retinal images. While this works well for image regions involving only vessels, the example in Figure 2(d) shows strong response to non-vascular structures. Threshold-based methods find pixel grey levels that separate vessels from non-vessels either in the source or in pre-processed images. This is difficult even with adaptive thresholds [31], [32] due to non-uniform illumination and the presence of pathologies. In [33], [34] vessels are found by finding their intensity edges. Methods based on mathematical morphology [16], [35], [37] use linear structuring elements to enhance vessels. Tolia and Panas [64] use fuzzy clustering of the vessel profiles in a tracing algorithm starting from the optic disc. Tracing can provide substantial speed-up but we have found experimentally that the technique [33], [48] performs most poorly on narrow, low-contrast vessels – exactly what we are concerned with in this paper.

Overall, each of the primary techniques in the literature — Hessian, matched-filter and parallel edge — has limitations when used on images containing low-contrast and narrow vessels as well as significant pathologies. The examples shown in Figures 2(a)–2(d) illustrate this. Recent papers in the literature have built segmentation algorithms on top of the basic Hessian and matched-filter measures, for example using adaptive thresholding [31] and grouping methods [39]. Our approach here is to enhance the primary image-based measures of vesselness. We start with the matched filter because it provides the best overall response to low-contrast and narrow vessels.

III. MULTISCALE MATCHED FILTERS

We start by defining a multiscale matched filter for vessels using a Gaussian vessel profile. The matched filter at one scale is implemented by maximizing the responses over a discrete set of kernel orientations at each pixel. Let $f(u, v)$ be the image intensity at pixel location (u, v) . At each kernel orientation we compute the response to a two-dimensional separable kernel from the convolution of two one-dimensional kernels applied in succession. The kernel in the tangential direction is simply a Gaussian, $g(\cdot)$, with wide support in order to smooth responses along the ridge. The kernel in the normal direction is a second-

order derivative of the Gaussian, $g_{vv}(\cdot)$. In a coordinate system rotated to align u with the tangential direction and v with the normal direction, the matched-filter response is

$$M(u, v; t_u, t_v) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_{vv}(v - v'; t_v) g(u - u'; t_u) f(u', v') du' dv', \quad (5)$$

where t_u and t_v are the variances of the Gaussian in each direction. The scale parameter of the vessel matched filter is the variance t_v .

In order to detect vessels at a variety of widths, we apply the matched filter at multiple scales (i.e. compute (5) for multiple values of t_v) and then combine the responses across scales. Unfortunately, the output amplitudes of spatial operators such as derivatives or matched filters generally *decrease* with increasing scale. To compensate for this effect, Lindeberg [49] introduced γ -normalized derivatives. We use this notion to define a γ -normalized matched filter. The idea is to multiply the matched-filter response at any scale by a normalization factor that takes the form $t^{\gamma/2}$ for scale t . In practice, this normalization is included in the matched-filter kernel. Therefore, kernel coefficients sum to 0 and they are normalized so that filter responses for different kernel scales can be fairly compared to detect vessels of varying widths.

Following [49] we compute γ by proposing an idealized model of the desired feature and its scale, t_0 , and then finding the value of γ that causes the peak response in scale space to occur at $t = t_0$. Our appearance model of an idealized vessel is a Gaussian cross-section with variance t_0 , and a constant intensity in the tangential direction. Because of this, no normalization is needed in the tangential direction and we may consider only the normal direction in deriving the normalization factor. Scale selection must be derived from the entity defining the ridge (vessel) strength – in our case convolution of the vessel profile with the matched-filter kernel of the same orientation. Exploiting the properties of the Gaussian, the γ -normalized matched-filter response to a Gaussian cross-section at location $(0, 0)$ and scale t is

$$\begin{aligned} M_{\gamma\text{-norm}}(0, 0; t) &= - \int_{-\infty}^{\infty} \frac{t^{\frac{\gamma}{2}}}{2\pi\sqrt{t_0}t^{\frac{5}{2}}} (v'^2 - t) \exp\left(-\frac{v'^2}{2} \left(\frac{t+t_0}{tt_0}\right)\right) dv' \\ &= \frac{t^{\frac{\gamma}{2}}}{(2\pi)^{\frac{1}{2}}(t_0+t)^{\frac{3}{2}}}. \end{aligned} \quad (6)$$

To find the peak, we differentiate with respect to the scale parameter and set the derivative to zero. This implies

$$\frac{\frac{\gamma}{2}t^{\frac{\gamma}{2}-1}(t_0+t) - \frac{3}{2}t^{\frac{\gamma}{2}}}{(t_0+t)^3} = 0. \quad (7)$$

Solving for t yields a maximum at $t = t_0\gamma/(3-\gamma)$. Forcing this peak to occur as desired at $t = t_0$, the scale of our model cross-section, we must have $\gamma = 3/2$. Therefore, the normalizing multiplier on the matched-filter response (5) at scale t is $t^{3/4}$. This multiplier has the same value as that derived by

Lindeberg [49]. Formally, a mapping may be derived between the convolution for a matched filter in Equation (6) and Lindeberg's measure derived for second derivative operators applied to a Gaussian ridge profile, but we presented the direct derivation here for clarity and completeness.

A. Parameters and Computation

Using this normalization, the responses may now be fairly compared across scales. Computing the best response, which we will denote by r_v , at each pixel simply requires computing the maximum matched-filter response over all scales and all orientations. We first describe a computationally-expensive approach that uses a fairly small number of orientations and then show how to increase the number of orientations while decreasing the computational time. For the first approach, 9 matched-filter orientations are used in 20° increments, the range of scales, t_v , in the direction normal to the vessel is $0.5 \leq \sqrt{t_v} \leq 3.0$ pixels in 0.5 pixel increments, and $t_u = 9.0$ for all values of t_v . We use a fixed t_u for all t_v because a large t_u is needed for small vessels in order to integrate more vessel pixels, while enlarging t_u further for large t_v we found to be more costly than beneficial, especially due to vessel curvature. Kernel cut-offs are $3\sqrt{t_v}$ and $2\sqrt{t_u}$. We chose these parameters based on our experiments with STARE and DRIVE databases which have images with resolution 700×605 and 565×584 and field of view 35° and 45° , respectively.

Computing a matched-filter response at each pixel location, orientation and scale is expensive. Fortunately, a large fraction of the cost may be eliminated by introducing two simple heuristics. First, at each pixel and scale, t_v , the Hessian is computed, the eigenvector corresponding to its largest eigenvalue is chosen, and its orientation is rounded to the nearest integer degree. A matched filter is then applied with the normal in this direction. In effect, the Hessian is used to steer the matched filter. Because of symmetry, this results in 180 possible discrete directions for which filter kernels are precomputed. Second, at each pixel, scale-space peaks are found using a fine-to-coarse search, stopping at the first peak. In this way, extra computation at coarser scales is avoided. Overall, we have found that these speed-ups eliminate 93% of the matched-filter computations and even provide slightly improved results because the matched-filter directions are chosen in finer increments. Unless otherwise indicated, all results reported here use these speed-up techniques.

Once again, as compared to the Hessian-based operators discussed in Section II, the matched filter provides better responses to thin vessels and to low-contrast vessels. Spurious responses to non-vascular structures such as the fundus boundary, the optic disc boundary and pathologies are still present, however. This, together with a desire to further improve the response for low-contrast vessels, motivates the use of the ‘‘confidence’’ and edge measures we define next.

IV. CONFIDENCE AND EDGE MEASURES

While the multiscale matched filter alone provides better response to low-contrast and narrow vessels than Hessian-based measures (Figure 2(a)), it still provides spurious responses to non-vascular structures. Insight for why this occurs can be obtained by considering what happens when a matched filter

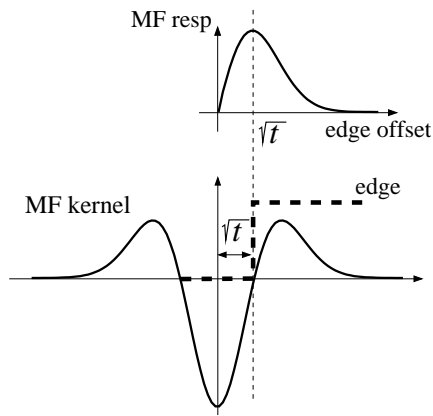


Fig. 3. Vessel matched-filter response to an offset edge. It is highest when the offset is equal to the standard deviation \sqrt{t} of the matched filter.

is applied near a step edge. As in the previous section, the constant intensity along the tangential direction allows us to only consider a 1-D profile. Symmetry arguments show that the matched-filter response is 0 when centered on the edge. On the other hand, as the matched filter is shifted along the normal to the edge, the responses start to increase in magnitude, reaching a peak at an offset of \sqrt{t} (see Figure 3). These responses are often much stronger than the responses to lower contrast vessels, as illustrated around the retina boundary, the pathologies, and the optic disc in Figure 2(a). Similar problems occur with Hessian-based measures. The confidence and edge measures defined in this section are intended to add the information needed to distinguish these spurious responses from responses to true vessels.

A. Vessel Confidence Measures

We consider a discrete version of the matched filter at pixel location $\mathbf{x}^T = (u, v)$:

$$M(\mathbf{x}) = - \sum_i \sum_j \mathbf{G}(i, j) \mathbf{F}(u - i, v - j). \quad (8)$$

Here, \mathbf{G} is the matched-filter kernel and \mathbf{F} is the discrete image. We can rearrange \mathbf{G} and the corresponding pixels of \mathbf{F} by putting their columns one-by-one into vectors \mathbf{g} and $\mathbf{f}(\mathbf{x})$, respectively. For example, matrix \mathbf{G} with M columns and N rows will be rearranged into vector \mathbf{g} with $M \cdot N$ elements. Doing this, (8) may be rewritten as

$$M(\mathbf{x}) = -\mathbf{f}(\mathbf{x})^\top \mathbf{g}. \quad (9)$$

In other words, the matched filter at a fixed location is expressed as a dot product of the kernel and image neighborhood both rearranged into vectors. If \mathbf{g} is normalized (without loss of generality) to unit magnitude, this may be considered as a projection of $\mathbf{f}(\mathbf{x})$ onto direction \mathbf{g} . As in any projection, a great deal of information is lost. In particular, the projection $\mathbf{f}(\mathbf{x})^\top \mathbf{g}$ says *nothing* about how well $\mathbf{f}(\mathbf{x})$ fits the profile represented by the kernel \mathbf{g} . This is simply because $\mathbf{f}(\mathbf{x})$ is unnormalized. A vector $\mathbf{f}(\mathbf{x})$ formed from an ideal, low-contrast vessel will be nearly parallel to \mathbf{g} , but may have a much smaller projection than a vector $\mathbf{f}(\mathbf{x})$ formed by an offset step edge (as previously discussed) or even a region made up solely of large magnitude

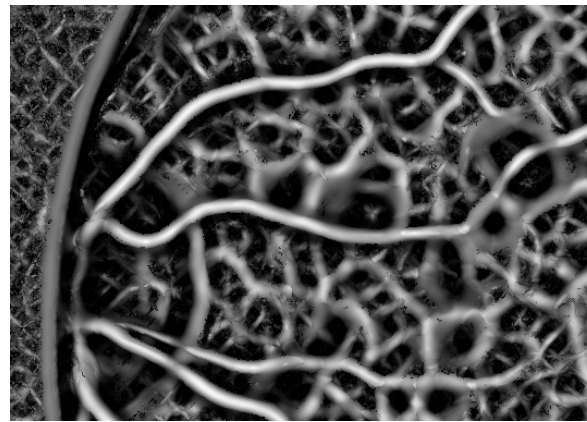


Fig. 4. Vessel confidences for the image in Figure 1. Usually, the confidences are highest at vessel locations but they can still be high around locations offset from edges and other non-vessel structures. Confidence image was stretched linearly to $[0, 255]$ for display.

noise.

We now introduce a vessel “confidence measure” (borrowing terminology from [50]). At each pixel, we determine the peak matched-filter response scale and direction, as described in Section III. We then let \mathbf{g}_m be a vector formed from a segment with the ideal Gaussian profile at the current orientation and scale, t_v — i.e. a Gaussian along the vessel normal and constant along the vessel tangent. (We use the same parameters as in the matched filter.) Let $\mathbf{f}(\mathbf{x})$ be a vector of the image intensity values taken from pixels overlaid by this template. The values in vectors $\mathbf{f}(\mathbf{x})$ and \mathbf{g}_m are independently centered and normalized by first subtracting their mean (so that the sum of the elements in each vector is zero) and then dividing by their magnitude (so that the vectors have magnitude of one). We define the confidence measure at this pixel as

$$\eta_v(\mathbf{x}) = -\mathbf{f}(\mathbf{x})^\top \mathbf{g}_m. \quad (10)$$

This can be viewed as normalized cross-correlation or cosine distance between model and sample vectors. It provides an indication of how closely-aligned the pixels are with the ideal Gaussian vessel profile. In particular, the value is 1 at the center of a Gaussian profile at the exact width, 0 at the center of an ideal step edge, and 0.79 for the peak response when offset by $\sqrt{t_v}$ normal to a step edge. Interestingly, the latter response is still relatively high. An example result is shown in Figure 4.

B. Edge Measures

We have seen that the matched filter, even with an added confidence measure, is not enough to entirely eliminate non-vessel responses. What is needed is a more direct measure of what occurs at the boundary of a purported vessel. To illustrate more concretely why this is important, consider again the difference between the matched-filter response to an offset step edge and the matched-filter response to a true vessel. In the first case, there will be an edge on one side of the region, but not on the other, while in the second case there will be edges on both sides. Stated more strongly, the fact that the second edge is missing is the most significant difference between a vessel and an offset step edge. We capture this information by combining edge strength

and confidence information with the matched-filter strength and confidence.

The first step in doing this is to compute the gradient direction and magnitude at each image pixel in scale space. This is done by applying the algorithm of [49] as implemented in ITK [65]. The maximum normalized response and associated direction are saved at each pixel. The range of scales t_e is $0.5 \leq \sqrt{t_e} \leq 1.75$ in 0.25 increments, and kernel cutoffs are $3\sqrt{t_e}$. Note that this range is much narrower than the range of scales for the matched filter. If the vessels followed an ideal Gaussian profile then the optimum sets of edge and vessel scales would correspond. In practice however the vessel scale depends on the width of the vessel, whereas the edge scale depends on the diffuseness of the vessel boundary. In particular, wider vessels in real images often have edges steeper than the Gaussian-modeled vessels of the same width. Therefore finding edge scale independently of the vessel scale yields better results; we have found the given range sufficient.

The following shows how the edge responses are combined with the independently-computed matched-filter responses (and confidences). At a given pixel location \mathbf{x} , let \mathbf{n} be the (unit) normal to the vessel and $w = 2\sqrt{t_0}$ be the approximate vessel width, both determined by the peak matched-filter response in orientation and in scale (t_0). Given these values, the vessel edges should pass (roughly) through pixel locations $\mathbf{x}_1 = \mathbf{x} + (w/2)\mathbf{n}$ and $\mathbf{x}_2 = \mathbf{x} - (w/2)\mathbf{n}$. Moreover, at these locations, the edge gradient directions should point outward, away from \mathbf{x} (for dark vessels) in direction \mathbf{n} . Therefore, at each of the two potential edge locations \mathbf{x}_1 and \mathbf{x}_2 , we search in a ± 1 pixel interval along \mathbf{n} for the pixel whose gradient vector, when projected onto the outward normal (\mathbf{n} for \mathbf{x}_1 and $-\mathbf{n}$ for \mathbf{x}_2), has the largest positive value. The ± 1 pixel interval – giving 3 possible locations overall – accounts for the modeling error caused by asymmetric vessels. This is only a slight shift in the location and is needed because of discretization effects, because of slight asymmetries in the appearance of the vessel, and because of diffuseness of the edges. We record these largest values along with the matched-filter and confidence responses at \mathbf{x} — in other words at the center of the purported vessel. In doing so, we order the two edge responses by their magnitude, denoting the stronger by $r_{es}(\mathbf{x})$ and the weaker by $r_{ew}(\mathbf{x})$. This ordering makes the measure more discriminating and easier to train. To illustrate, this offset step edge is distinguished by a low value of the weaker response rather than by which side this response occurs on.

Finally, we compute confidences for the strong and weak edge responses. Similar to the matched-filter confidence, the edge confidences measure how well the edge profile fits the expected shape, independent of the magnitude of the edge. To compute confidence, the ideal vessel edge profile is modeled by the cumulative distribution of the Gaussian, with variance t set to the peak scale of the edge. Following the same procedure as for the vessel confidence, each edge confidence measure is computed by projecting the local intensity window onto a segment with the ideal profile oriented along direction $\pm\mathbf{n}$ and centered at the detected edge locations (from the ± 1 pixel interval). The resulting confidences are denoted $\eta_{es}(\mathbf{x})$ and $\eta_{ew}(\mathbf{x})$ and together with the responses recorded at the center pixel, \mathbf{x} (the vessel pixel). Figure 5 shows an example result. Confidences and

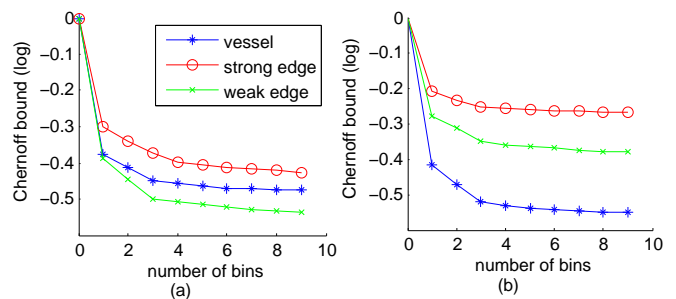


Fig. 6. Logarithm of the Chernoff bound as a function of the number of bins (a) for vessel, strong edge, and weak edge responses, (b) for vessel, strong edge, and weak edge confidences. The discriminative power of the weak edge is higher than that of the strong edge because the existence of the weak edge is the most significant difference between a vessel and an offset step edge.

responses of both edges are high at vessel locations, but the weaker edge response and its confidence are low near edges and other non-vessel structures.

Overall, the combined vessel and edge computations give us three strength measures and three confidence measures at each pixel. We combine these into a six-dimensional vector $\Theta(\mathbf{x}) = (r_v(\mathbf{x}), r_{es}(\mathbf{x}), r_{ew}(\mathbf{x}), \eta_v(\mathbf{x}), \eta_{es}(\mathbf{x}), \eta_{ew}(\mathbf{x}))$.

V. TRAINING AND DECISION CRITERIA

Our next step is to map the 6-D response/confidence measurement vector Θ into a single “vesselness” measure at each pixel. The measure we choose is the optimal test statistic by the Neyman-Pearson Lemma [66], the *likelihood ratio*:

$$\text{LRV}(\Theta) = \frac{p(\Theta|\text{on})P(\text{on})}{p(\Theta|\text{off})P(\text{off})}. \quad (11)$$

Here $P(\text{on})$ is the prior probability of a pixel location being on the centerline of a vessel, $P(\text{off}) = 1 - P(\text{on})$, and $p(\Theta|\text{on})$, $p(\Theta|\text{off})$ represent the corresponding conditional probability density functions (pdfs) for measurement vector Θ . LRV is the desired “Likelihood Ratio Vesselness” measure. Note that when the likelihood ratio is 1.0, it indicates an equal chance that the current location represents vessel or non-vessel. With increasing values of likelihood, we are more sure that a vessel is present.

To train these pdfs we use manual-segmented retinal images, such as in the STARE [31] and DRIVE [51] data sets, as ground truth. Only pixel locations for which the matched-filter response is positive are considered because negative responses will never be a vessel. Non-maximum suppression is applied to the matched filter responses so that pixels near the vessel boundaries are not used in training. The manual segmentations were not thinned for training because even small (1-2) pixel errors in the localization would change decision about whether a response is a vessel or not. Training $P(\text{on})$ is accomplished simply through counting the number of matched filter responses marked as vessels in the manual segments and dividing by the total number of responses. Since we only consider positive matched filter responses, our prior $P(\text{on})$ is the probability of a pixel location with a positive matched filter response being a vessel centerline pixel. $P(\text{off}) = 1 - P(\text{on})$. Training the conditional pdfs uses a histogram-based method adapted from [67]. We chose histograms to estimate the

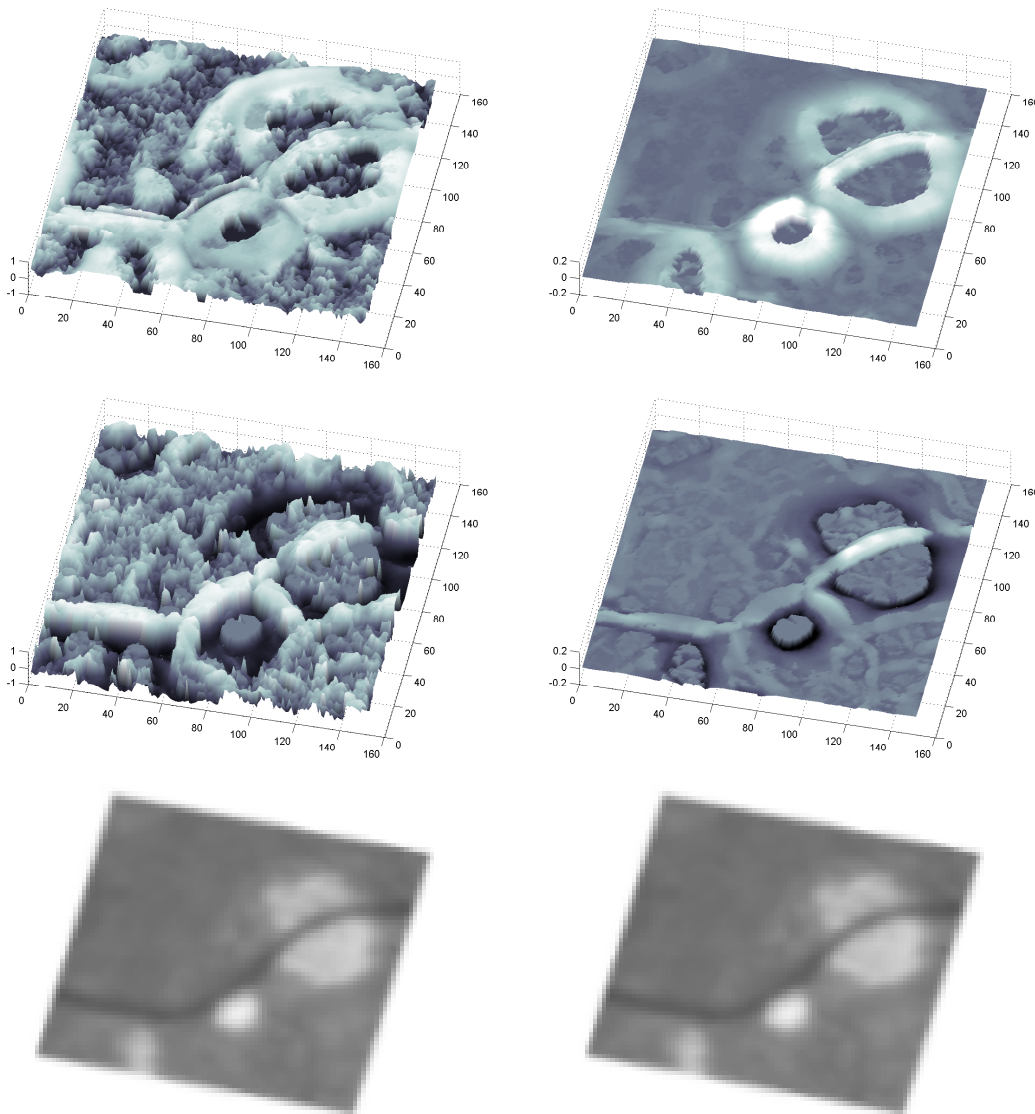


Fig. 5. Confidence (**left column**) and response (**right column**) images of the stronger (**top**) and weaker (**middle**) vessel edge for the source image shown at the bottom. These are displayed as 3D plots. Confidences and responses of both edges are high (and positive) at vessel locations. Around pathologies and other non-vascular structures the weaker edge responses and confidences are close to zero (if no weaker edge is detected) or negative (if it is located on the same diffuse edge as the stronger edge).

distributions because of their simplicity, although clearly other options such as Parzen windows, k -nearest neighbor classifiers, and decision trees could be considered [68]. Two six-dimensional histograms are formed, one to represent $p(\Theta|\text{on})$ for the “on” pixels and one to represent $p(\Theta|\text{off})$ for the “off” pixels. The histograms are normalized to unit magnitude. Because of the high dimensionality, the choice of bin size is crucial: making the bins too small leads to explosion in the number of bins; making them too large leads to a loss in precision. To solve this problem, bin boundaries are chosen adaptively, with the same boundaries used in both histograms.

The choice of boundaries is based on the Chernoff bound on the error for the likelihood ratio [67], [68]:

$$P(\text{error}) \leq P^\xi(\text{on})P^{1-\xi}(\text{off}) \int p^\xi(\Theta|\text{on})p^{1-\xi}(\Theta|\text{off})d\Theta, \quad (12)$$

where $0 \leq \xi \leq 1$. To obtain the bound, the right hand side of the inequality is minimized over ξ . Computationally, the pdf is replaced by a normalized histogram with a discrete set of bins, and the integral in (12) is therefore replaced by a discrete summation.

We now describe how to determine the bin boundaries. Each dimension is handled separately, which means we look for boundaries partitioning 6 1-D histograms. The resulting sequences of boundaries for every dimension are combined by taking their Cartesian product to form the corner points of the axis-aligned bins in 6 dimensions. (Note that the same boundaries are used for both $p(\Theta|\text{on})$ and $p(\Theta|\text{off})$.) In order to determine the boundaries in one dimension, we start by assigning all measurements from the training images to bins in a fine-grained 1D histogram. The Chernoff bound (12) is then minimized over ξ for each boundary point, treating the boundary point as a partition that assigns all

measurements below it to one bin and all measurements above to another bin. The point for which the overall minimum of (12) is achieved is selected as the first partitioning boundary. The second partitioning boundary is determined in a similar manner with the first boundary fixed. Every additional boundary decreases the minimum Chernoff bound, which reaches an asymptotic value when enough boundaries have been added. Figure 6 shows that this happens quickly and only 6 bins per dimension are sufficient to model the distributions accurately. This gives a total of $6^6 \approx 4.7 \times 10^4$ bins, which is a manageable size. Typically, only 15% of the histogram bins are filled. Although this suggests that a more compact representation could be used, the total number of bins is too small for this to be a substantial concern.

The training also gives insight to the discriminating power of each measure we have developed. In particular, the left plot of Figure 6 shows the Chernoff bound as a function of the number of bins for vessel, weak edge, and strong edge responses, while the right shows corresponding plots for the confidences. Overall, the strength measures are more discriminating than the confidence measures, but not dramatically so. This substantiates the motivation for using both types of measures. Below we show experimentally the improvement in the detection rates by augmenting the vessel matched filter with vessel confidences and further improvement when the edge measures are added.

VI. EXPERIMENTAL EVALUATION

The likelihood ratio (11) is our final “vesselness” measure, LRV, intended to be used similarly as the vessel measure of Frangi et al. [42] (3) and the ridge measure of Lindeberg [49] (4). Non-maximum suppression is applied to the LRV values in the direction normal to the vessel orientation at detected points. A threshold is then applied to the surviving LRV values to make the final vessel / non-vessel decisions. A faster but less effective version of the LRV measure is obtained by substituting Frangi vesselness (3) for the matched filter in the measurement vector. All the other steps of the algorithm, including scale selection, remain unchanged.

In our evaluation, we first quantify the improvement gained by using vessel confidences and edge measures together with the matched filter. We then compare detection results using LRV and Frangi’s measure on all retinal vessels in an image. We then present detection experiments focused on thin vessels and on low contrast vessels. Finally, we show qualitative comparison of the LRV measure and the vessel matched filter.

A. Evaluation Technique

We evaluate our technique both quantitatively and qualitatively using two publicly-available data sets with images of diseased retinas, the STARE database [31] and the DRIVE database [51]. There are advantages and disadvantages to each database. STARE includes images with a wider range of appearances and more large-scale pathological structures, but there are fewer images (only 20) and there are substantial interobserver variabilities on the manual segmentations: one observer consistently highlighted more small-scale vessels. More images are available in DRIVE and the interobserver agreement is better, but the pathologies are less prominent. Because of our interest in narrow vessels and avoiding false response to pathological structures, most of the

results we present here are taken from the STARE database with the segmentation produced by the observer who marked more small-scale vessels.

During training we use a leave-one-out strategy for the STARE data set and the dedicated training set for the DRIVE data set. Of the two image databases, DRIVE contains mask images for filtering background and STARE does not. Automatic masking is easy in carefully controlled imaging environments. In practice many images have illumination artifacts, double boundaries caused by acquisition, numbers and tags and different fields of view, which all make masking hard. Moreover, one of the advantages of our technique is that masking is not necessary. Therefore, in this paper, masks were used in experiments only on the DRIVE image set which has the mask images included.

Our analysis is based on vessel centerlines, even though other published results are based on vessel segmentations [31], [32], [39], [51]. There are a number of reasons for this. First, all components of the LRV measure — the matched filter, the two edge responses and all of the confidences — strongly emphasize detecting the center of the vessel. The measures are much weaker near the boundaries of vessel. Second, these centerline points, sometimes together with approximate widths, are what is most needed in algorithms such as registration [12] and optic disc detection [6] that depend on vessel extraction. Third, since the boundaries of the vessels are diffuse and fade subtly into the background, boundary localization is non-exact and highly-subjective, making the labelling of pixels near boundaries less useful in quantitative analysis. Finally, quantitative measures of segmentation are dominated by the performance on wide vessels, simply because wide vessels have more pixels.

Following training we gather performance statistics on test images. For any test image we can determine the number of true positive, false positive, and false negative vessel centerlines extracted using our LRV measure for any choice of final threshold. We define these counts by comparing against a thinned version (using the thinning code in Matlab) of the manual segmentations. A true positive is any detected point within 2 pixels of a point in the thinned manual segmentation. False negatives are counted as the number of points in the thinned manual segmentation not within 2 pixels of a detected centerline point. The tolerance of 2 pixels was chosen to account for localization errors in both the thinned manual segmentation and in the centerline peak location (essentially allowing 1 pixel of error in each).

Finally, we have the question of how to evaluate the quantitative results. ROC curves, which are widely-used, are not very informative when the total number of negatives (in the ground truth) is much larger than the total number of positives; in our case fewer than 3% of the pixels are vessel centerlines. This means a large number of false positives — e.g. equal to the total number of true positives — will be small in terms of percentages of true negatives (see discussion in [69]). This produces an extremely steep, almost step-like, ROC curve. A better measure for evaluating vessel detection is a (1–Precision)-Recall curve. Recall equals the true-positive detection rate just as defined for ROC curves. (1–Precision) is defined as the number of false positives divided by the total number of detections. In another words, it defines the fraction of detections that are wrong.

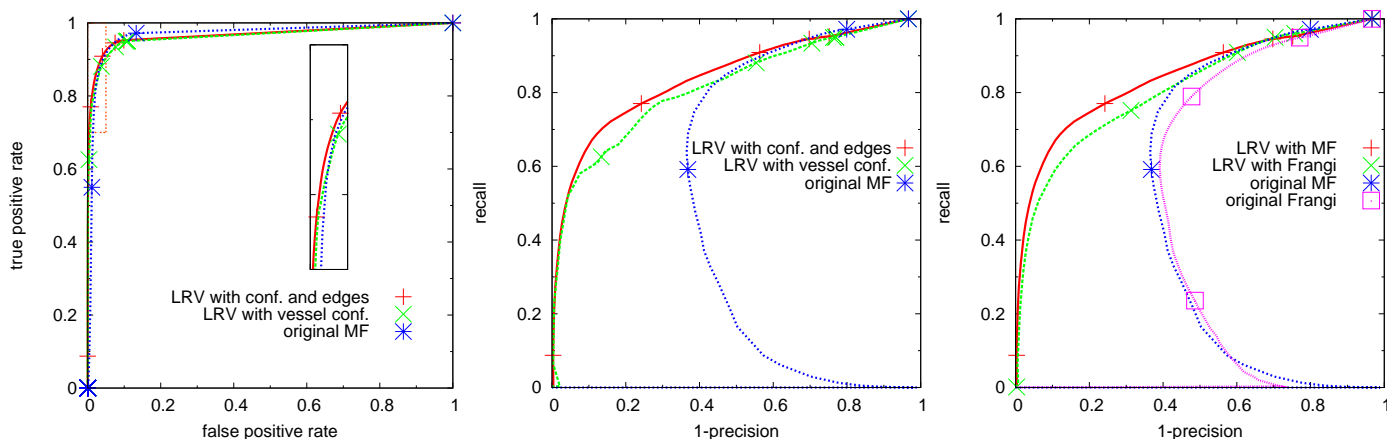


Fig. 7. ROC curves (left) and (1–Precision)–Recall curves (middle) showing the effectiveness of the LRV measure. Comparisons are made between the full measure (LRV), the measure with only vessel confidences, and the multiscale matched filter alone. (1–Precision)–Recall curves are more suitable for comparison because of large number of negatives (non-vessel pixels) in the ground truth. Notice in the middle plot that the matched filter with vessel confidences is as powerful as the full LRV measure until about 50% of all traces are detected. Plot on the right compares the LRV measure (with the matched filter) to the LRV using Frangi’s Hessian-based measure and to the matched filter and Frangi’s measure alone.

B. Overall Results

The first part of our quantitative analysis uses overall results combined across all 20 images in the STARE data set (Figure 7). As shown in Figure 7(a), the ROC curves are not very informative, with only very minor differences separating the measures. The (1–Precision)–Recall plots in the rest of the figure are much more informative. Plot (b) shows that use of the matched filter and its associated confidence — a two-component measurement vector at each pixel — in forming the likelihood ratio dramatically improves the performance of the matched filter, while adding the edge responses and confidences needed to form the complete 6-component vector provides significant further improvements. Plot (c) shows that both the Hessian-based (Frangi) and the matched-filter-based LRV measures substantially outperform the original measures, but the matched-filter LRV is clearly superior. (Lindeberg’s ridge measure performs slightly worse than Frangi’s vesselness measure, so it was left out of the the plots.) Finally, the seemingly-strange non-monotonic shape of the matched-filter and Frangi measures alone is easily explained. The highest responses for these measures occur at the retina boundary, the optic disc boundary, and the boundary of pathologies (though perhaps slightly offset from the true boundary, as discussed above), because they are of much higher-contrast than even the more distinct vessels. Hence, for very high thresholds, these are the only responses that survive thresholding.

C. Thin Vessels and Low-Contrast Vessels

In the second part of our analysis we consider the specific problems that motivated our work: detecting thin and low-contrast vessels while avoiding false detection of vessels near pathologies and other non-vascular structures.

We filter the overall quantitative results to develop a fairly rough but indicative analysis of the performance on both thin vessels and low-contrast vessels. For thin vessels, we process the ground truth vessels, eliminating those from the image that are wider than 2 pixels. We do something similar for low-contrast vessels, eliminating those where the difference between the

average foreground intensity and the average nearby background intensity is more than 3 noise standard deviations. (1–Precision)–Recall curves for these two computed using the STARE data set [31] are shown in Figure 8.

We also evaluated the performance using the DRIVE database of images [51] tested for both observers with training using the separate dedicated training set (Figure 9). The improvements are similar but less substantial because the images are of higher quality: their contrast is higher, a brightness does not vary significantly, imaging artifacts such as blur are minimal, the image background is masked out, and the images contain fewer substantial pathologies. The overall higher detection performance of all algorithms can be seen from larger area under the curves.

D. Qualitative Results - Pathologies

In order to give a qualitative feel for the results, several “chips” from the most challenging images are shown in Figure 10. These include a variety of pathologies and thin vessels that are so subtle they sometimes completely fade into the noise. (To obtain the operating point for the qualitative results we set a threshold of $\tau = 2.0$ on the LRV measure, which corresponds to a pixel being twice as likely to be on-vessel as off.) The first thing to observe about the results is that the matched filter (2nd column of Figure 10) highlights many of the vessels, but produces even stronger responses at non-vascular structures, such as the optic disc (4th row) and the pathologies (2nd, 3rd and 5th rows). When the matched filter is combined with the confidence and edge measures in producing the LRV measure, these false responses are dramatically reduced (3rd column). Moreover, the subtle vessels start to stand out (all rows). Occasionally, some of the pathologies still produce strong responses. These cases (row 5, bottom right of the image) usually involve two bright pathologies separated by a narrow strip in which the darker retinal background appears. The LRV measure correctly detects many vessels with central reflex (3rd row, yellow arrow). When the central reflex is severe, more complicated vessel models could be used (both for computing responses and confidences).

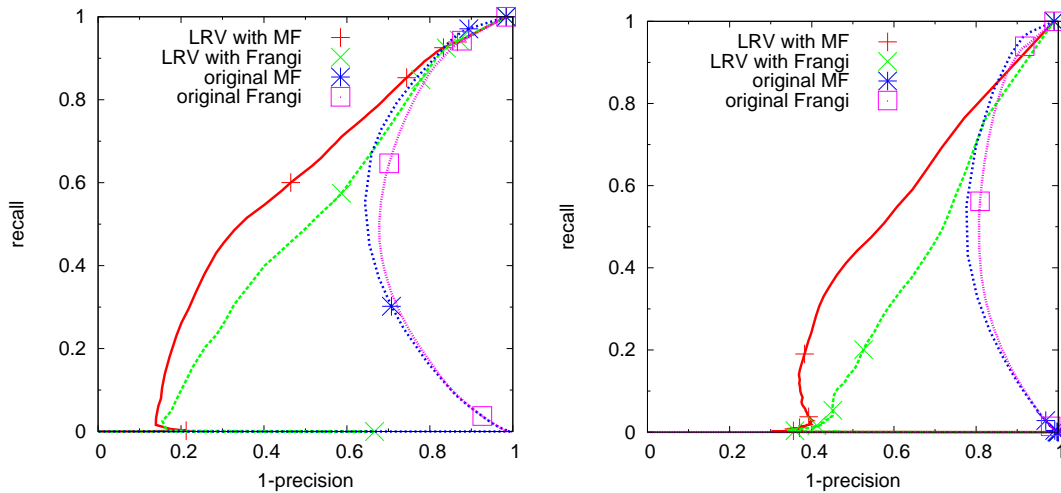


Fig. 8. Using the ground truth segmentation, we generated the (1–Precision)–Recall plots by focusing only on vessels 1–2 pixels wide (left). The plots are computed for the STARE data set [31]. The performance is worse than when evaluating all vessels, but still impressive. Similarly, we used ground truth segmentation to generate the plots studying vessels that have contrast less than 3 noise standard deviations (right).

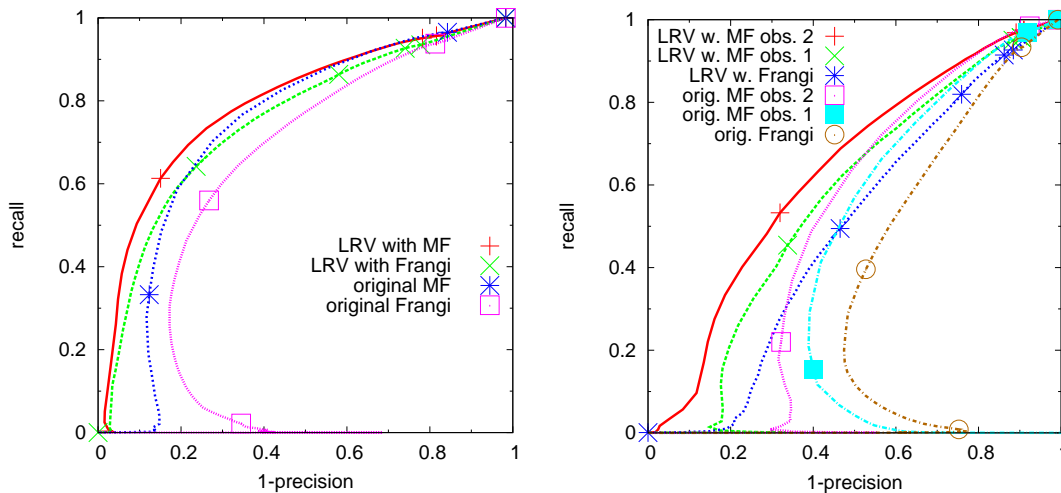


Fig. 9. (1–Precision)–Recall plots by focusing only on vessels 1–2 pixels wide (left) and on vessels that have contrast less than 3 noise standard deviations (right). The plots are computed for the DRIVE database [51]. LRV measure consistently outperforms the original matched-filter and Frangi measures. The difference between observers is larger for the low-contrast analysis so we plotted LRV and original matched-filter results separately for the two observers. Plot for the Frangi’s measure combines results for both of the observers. See the text for discussion.

After applying the threshold most of the narrow vessels still appear, although on occasion these are broken up into short segments. Overall, the results on these challenging images appear qualitatively to be better than other results in the literature.

E. Discussion

One question that arises in studying these results is how to use the performance curves to establish an operating threshold on the LRV measure and any other measure. This is particularly difficult for low-contrast and narrow vessels. For example, when the LRV measure is set to 50% recall of narrow vessels, the 1–precision rate for the narrow vessels (percentage of detections that are false) is 10% for DRIVE and 35% for STARE. For 50% recall of low-contrast vessels, the 1–precision rates are 31% for DRIVE and 57% for STARE. For the matched filter alone, at 50% recall, the 1–precision rates jump to 17% (DRIVE) and

65% (STARE) for narrow vessels and to 43% (DRIVE) and 80% (STARE) for low-contrast vessels.

There are several explanations for these low overall performance numbers. The first can be seen by evaluating the reliability of the manual segmentations used as ground truth. With the DRIVE database, we can use the thinned binary labelling of one observer as the ground truth to test the other. This yields a recall rate of 91% and 1–precision of 9% across all images, but with some images the recall rate drops as low as 76% and with others the 1–precision rate is as high as 25%. With low-contrast and narrow vessels, these numbers are much worse, with 81%/30% for low-contrast vessels and 85%/16% for narrow vessels. For STARE, recall varies between 99% and 56% and 1–precision between 46% and 1% in overall experiments. Clearly, with such high inter-observer variability, using the manual segmentations

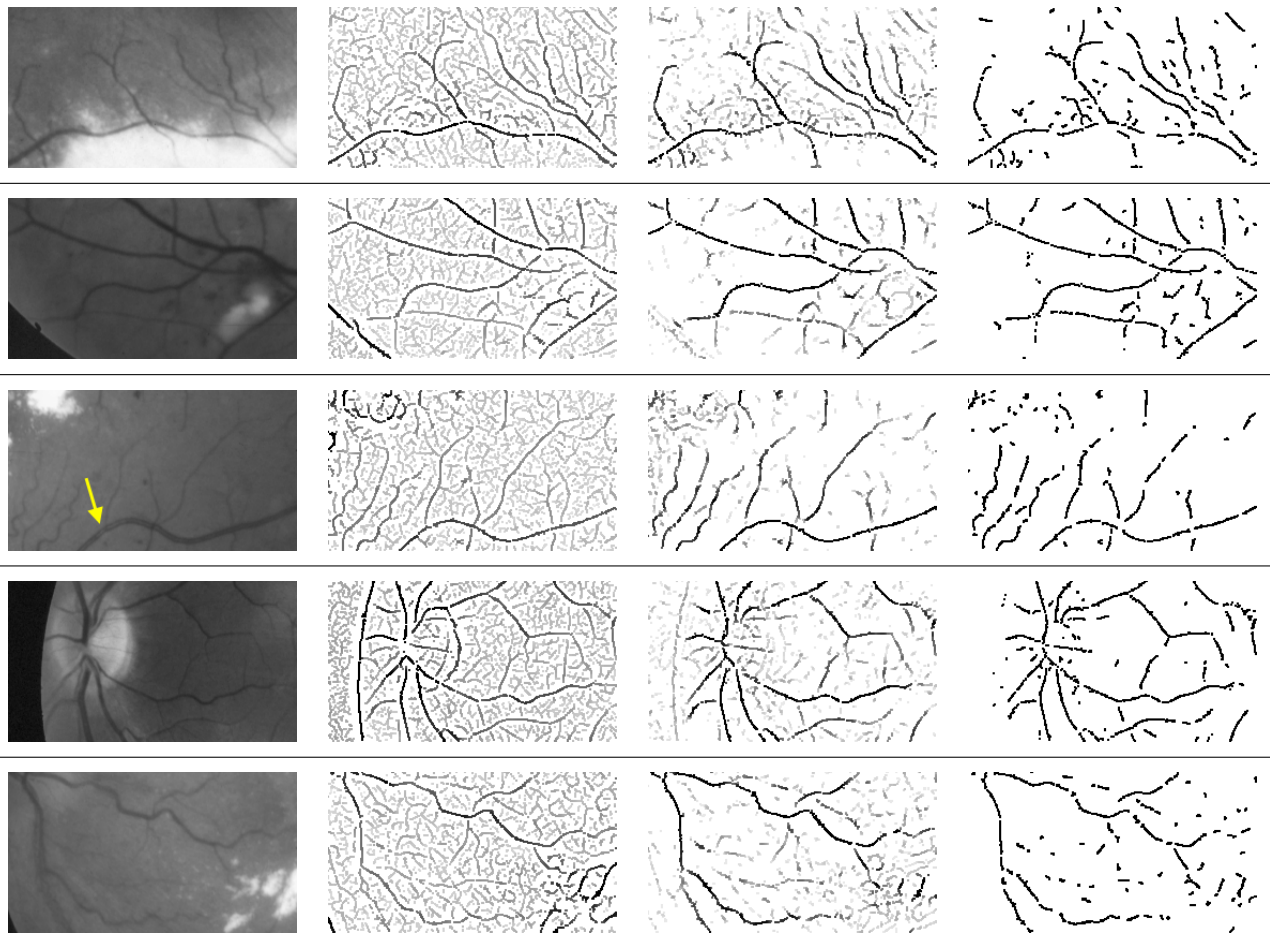


Fig. 10. Results on 5 difficult image chips showing a variety of challenges, including pathologies, other non-vascular structures, and low-contrast, narrow vessels. The figure shows the source images in the **1st column**, the vessel matched-filter response images after non-maximum suppression in the **2nd column**, the Likelihood Ratio Vesselness (LRV) after non-maximum suppression in the **3rd column** and the pixels that remain after thresholding the LRV measure at $\tau = 2.0$ in the **4th column**. The images in the last three columns were stretched to the range $[0, 255]$ for display here, so direct comparison of intensities within and across columns is easier. In generating the images in the fourth column, segments forming connected components with fewer than 4 pixels were removed. See the text for detailed discussion of the results.

to establish a 90% recall rate is not possible.

The second explanation for the low overall performance numbers comes from the fact that the LRV measure is computed in just a small image region. When we look in small image regions surrounding locations of false negatives in the LRV measure (for a given threshold) we often do not see vessels. Only stepping back to larger regions do these vessels start to “appear”. This suggests that higher-level information must be used. The simplest of these could be adaptive thresholding [32], [70], but more sophisticated grouping techniques are likely to be necessary and several steps in this direction have been taken in the literature [39], [40]. Still, our quantitative and qualitative results show that the low-level LRV measure substantially outperforms existing measures in a relative comparison and therefore should be inserted into existing algorithms in place of other measures such as the Hessian.

We computed paired t -tests on the areas enclosed by the (1–Precision)–Recall curves. Area is computed for each curve by analyzing every test image separately. Statistics on the areas (mean and variance) are calculated for each measure and the final critical value t is compared against the tabulated distribution. Doing so, we found out that LRV measure is significantly better

($p < 10^{-14}$ for STARE and $p < 0.0005$ for DRIVE) than the matched-filter and Frangi measures alone for overall, thin vessel and low-contrast vessel experiments. In addition, matched-filter-based LRV is significantly better ($p < 0.0001$ for STARE and $p < 0.005$ for DRIVE) than the Frangi-based LRV.

The final question in using our LRV measure is computational efficiency. We showed in Section III-A how applying the Hessian can be used to steer the matched filter, avoiding the cost of applying it at all orientations. This reduces the overall cost of our implementation to under 4 minutes on the images in the STARE database. A number of further optimizations is likely possible, for example avoiding the application of the matched filter altogether when the Hessian response is too low. Rather than investigating these possibilities here, we turn now to the question of embedding the LRV measure in a tracing framework.

VII. VESSEL TRACING

We now show how to substitute the LRV measure for our earlier parallel-edge vesselness measure [33], [48]. This has the dual-purpose of improving the efficiency of the LRV measure and improving the effectiveness of tracing.

All exploratory tracing algorithms follow the same overall

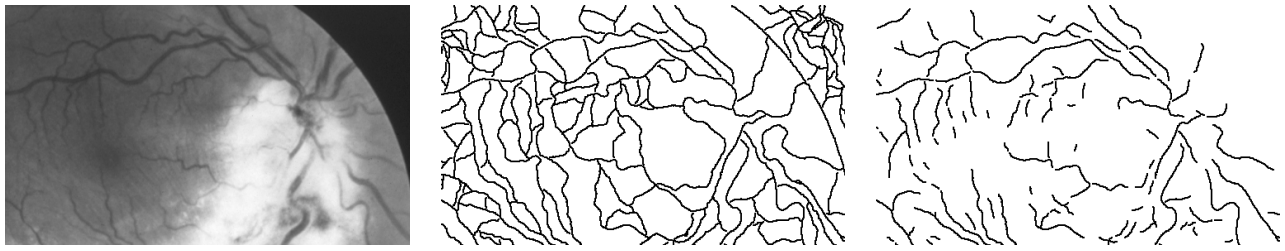


Fig. 11. Vessel traces of a source image (left) extracted using the matched filter to extract candidate traces (middle) and after applying full 6-dimensional LRV measure (right). Most of non-vessel structures are successfully removed while traces of even thin and barely visible vessels are kept.

strategy. Starting from a set of oriented seed locations, they recursively trace the vascular (or neuronal [71]) structure by stepping along the vessel tangent direction and locally detecting the next point on the vessel. Usually, the “centerline” location, vessel width, and tangent direction are estimated simultaneously. Tracing stops for a given seed point when it intersects a vessel already traced or when the vessel is not reliably detected. Tracing stops overall when all seeds have been explored. Usually, vessel branching and cross-over points are detected and localized in a post-processing step [33], [72].

Our approach follows this strategy as well, with a small but important variation. The natural approach would be to use the 6-D measurement vector of matched-filter, confidence and edge responses, computed over a small neighborhood in location, orientation and scale to compute the likelihood ratio and then apply a threshold to decide the presence of a vessel. We have found, however, that this can lead to missed traces for the most subtle vessels, prematurely halting tracing. Instead we apply only the vessel matched filter during tracing. This is very sensitive and detects both subtle vessels as well as results in many traces at non-vessel locations (Figure 11(middle)). We then compute the 6-D measurement and apply the likelihood ratio subsequently to these “over-traced” results. In effect we are using the vessel matched filter in tracing to generate candidate locations at which to evaluate the LRV measure. This is less expensive than starting from large number of seed points and applying LRV measure during tracing. Earlier in the paper (Section VI) we evaluated the LRV measure at each pixel.

Here are a few important details:

- The seed points we use are the same as the seeds in [48] — local intensity minima along smoothed vertical and horizontal lines through the images. Next, the multiscale Hessian is computed around each seed location and the vessel tangent direction is determined as the eigenvector corresponding to the smallest eigenvalue. Vessel width is found from the maximum matched-filter response in scale space applied in this direction. The seed points are then sorted based on strength so that when tracing starts, the most reliable points are used first. Only points that have positive matched-filter response are retained.
- The iterative tracing procedure is started from each seed point in the sorted list. At step $i \geq 0$ of tracing from a seed point, let \mathbf{x}_i be the current vessel location, θ_i be the current tangent orientation, and σ_i be the current vessel scale. The position, orientation and scale are refined by searching for the maximum scale-space response over a ± 20

degree change in orientation (in increments of 10 degrees), ± 1 steps in scale space, and ± 2 pixel steps in the normal direction. If the largest response is non-positive, tracing is terminated. Otherwise, the best parameters are recorded, a step of two pixels is taken along direction θ_i , and the process is repeated.

- The verification step computes the confidences and edge responses at each trace point using the location, orientation, and scale parameters to determine where to compute these values. The likelihood ratio is computed for the resulting measurement vector. The trace point is discarded if the likelihood ratio is below threshold $\tau = 2.0$. An example in Figure 11 shows retina vessel traces before and after applying the LRV measure and τ threshold. It may be possible to do even better using adaptive thresholding [32], [70].

VIII. TRACING RESULTS

We evaluate the effectiveness of this likelihood ratio tracing algorithm in side-by-side comparison with two other techniques.¹ One is our likelihood ratio (LRV) applied image-wide with non-maximum suppression. This allows us to compare the performance when embedding the LRV measure in a tracing framework. The second is the parallel-edge tracing algorithm [33] as refined in [48]. The goal of this comparison is to evaluate the improvement over an existing tracing algorithm. Training the conditional pdfs is done as in Section V. Estimating $P(\text{on})$ is accomplished by counting the number of traced pixels using the matched filter (in the “over-traced” image) manually marked as vessels in the manual segments and dividing by the total number of pixels. As before, we use a leave-one-out strategy in training from the STARE data set and the dedicated training set when training for the DRIVE data set.

Qualitative results can be seen in Figure 12. The figure shows cropped image regions with parallel-edge and LRV tracing as well as with the LRV measure evaluated at each pixel of the image. The same seed points were used for both algorithms. The thinned manual segmentation was used to mark true positives (blue), false negatives (red), and false positives (green). The new tracing algorithm successfully handles difficult vessel branching and crossover points (two overlapping vessels in the 3rd and 4th example), more accurately determines when to stop tracing in unclear regions (thin vessels near pathologies in the 1st and

¹The tracing executable can be found at <http://www.vision.cs.rpi.edu/vessels/>.

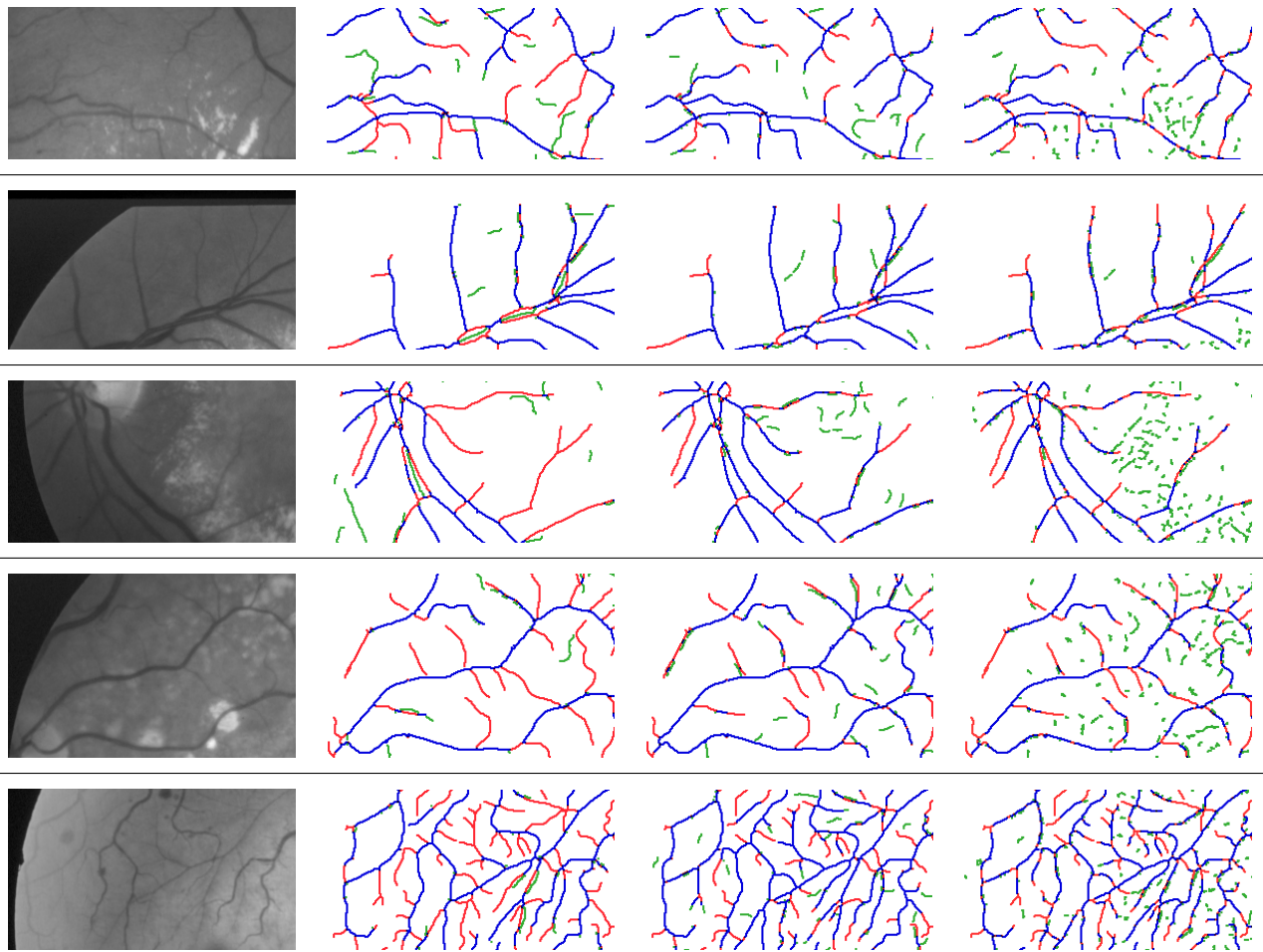


Fig. 12. Source image (**1st column**), tracing results using parallel edge algorithm [33] (**2nd column**), Likelihood Ratio Vesselness (LRV) based tracing (**3rd column**) and the LRV after non-maximum suppression and thresholding below 2.0 (**4th column**). True positives are in blue, false negatives in red, and false positives in green. The new tracing algorithm successfully handles difficult vessel branching and crossover points (two overlapping vessels in the 3rd and 4th example), more accurately determines when to stop tracing in unclear regions (thin vessels near pathologies in the 1st and 3rd example), correctly ignores pathologies (1st and 4th example) and finds many of the thinnest vessels which usually have very low contrast (5th example). Notice in a number of cases false positives are marked alongside false negatives (green line along red line) because of localization differences with the ground truth. The tracing results are comparable to the LRV measure evaluated at every point. Vessels with fewer than 4 connected pixels were removed.

3rd example), correctly ignores pathologies (1st and 4th example) and even finds many of the thinnest vessels (5th example). Notice in a number of cases that false positives are marked alongside false negatives (green line along red line) because of localization differences with the manual segmentations. The tracing results are comparable to the LRV measure evaluated at every point. Occasionally, a vessel is missed by both tracing algorithms (often around the optic disc) due to the lack of a seed point. Another reason for a vessel to remain undetected (3rd example on the left, marked as red in all results) is when the local structure is very poor and it does not follow the vessel model well. In this particular case, the LRV responses are between 1.0 and 2.0 and are therefore removed by thresholding. It is possible that also this vessel would get detected with better thresholding or a grouping technique. False detections that are inside manually-segmented vessels but not within the error tolerance of a vessel centerline are exhibited by a red trace (false negative) very close to a green trace (false positive) as in the 4th column of the 2nd example. Because it is a richer description, the LRV measure more effectively detects two side-by-side vessels while parallel-

edge tracing tends to detect them as one vessel (2nd and 3rd example in Figure 12). Tracing introduces a hysteresis that leads to less fragmentation than the raw LRV measure.

Quantitative performance is summarized in Figure 13. The new algorithm outperforms parallel edge tracing and brings improvement by more than 10 percentage points in the number of detected traces overall (measured for same false positive fractions and same likelihood threshold $\tau = 2.0$). For a small number of false detections, LRV tracing outperforms use of the image-wide LRV because vessels are traced starting from the strongest seed points. The situation changes when there are not enough seed points to detect as many vessels as with the exhaustively-computed LRV. The tracing curves do not reach point (1, 1) in the plots because tracing starts from seed points, meaning that vessels having no seeds are never traced. Similarly, point (0, 0) is not reached because whole vessel segments are traced at a time in the order of decreasing seed strength and for small number of false detections there will already be several correct vessel segments detected. It takes less than 30 seconds to extract traces from a 700×605 retina image (unoptimized code on Pentium IV,

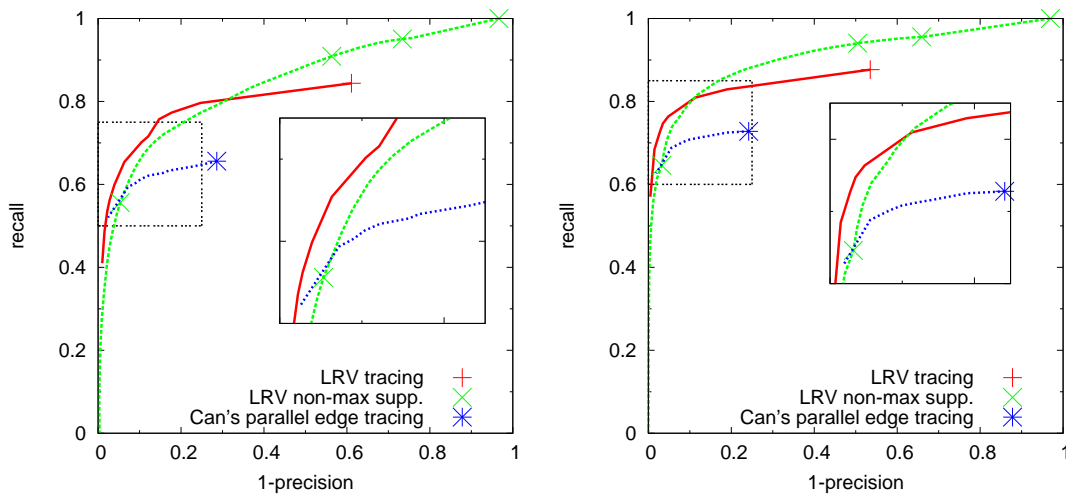


Fig. 13. (1–Precision)-Recall curves for LRV-based tracing results compared to parallel edge tracing algorithm [33] and LRV measure evaluated at each pixel followed by non-maximum suppression. Computed for the STARE data set [31] (**left**) and DRIVE database [51] (**right**). The new algorithm outperforms parallel edge tracing and brings improvement by more than 10 percentage points of the number of detected traces overall. For small number of false detections, tracing outperforms image wide LRV because vessels are traced starting from the strongest seed points. The situation changes when there are not enough seed points to detect as many vessels as with exhaustively computed LRV.

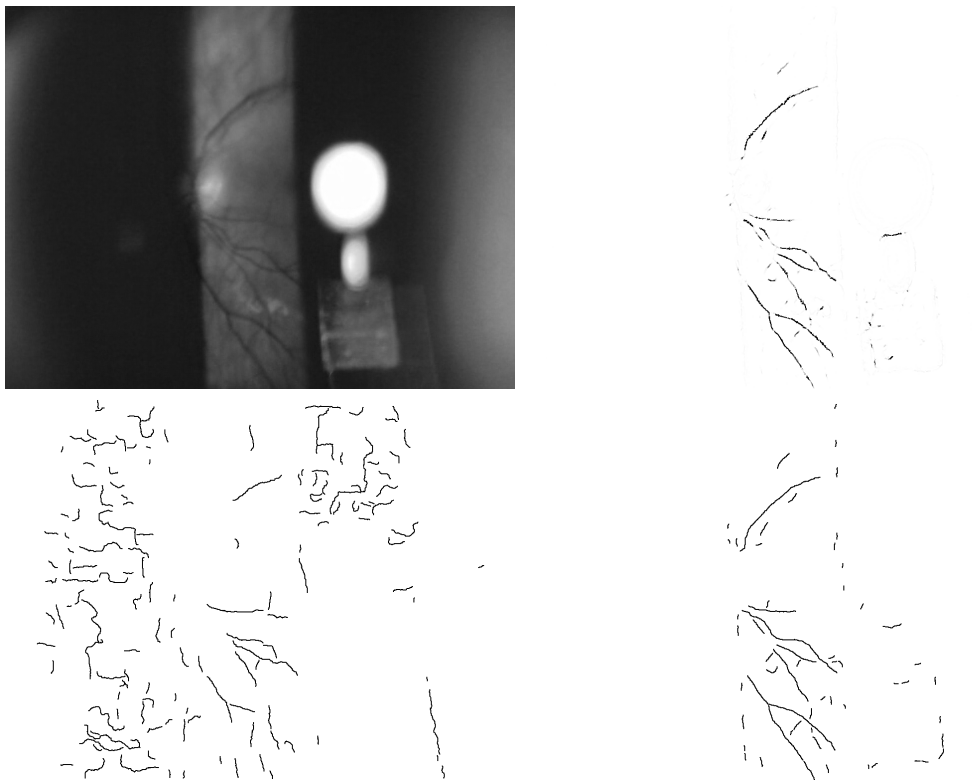


Fig. 14. Example of a slit lamp image (**top left**), LRV measure computed image wide (**top right**), parallel-edge based tracing (**bottom left**), and LRV-based tracing (**bottom right**). The LRV measure is powerful enough to highlight vessels but ignore the background noise which is the main reason for spurious traces of the parallel-edge based algorithm. The new tracing algorithm correctly finds vessels that are missed by the parallel-edge based tracing.

3.2 GHz). The processing time decreases to less than 10 seconds if we compute the vessel tangent directions using the eigenvectors of the Hessian as in pixel-by-pixel processing. The results can be slightly worse at locations where the tracing cannot continue due to poor local structure and inaccuracies in the eigenvector direction.

One final example that reinforces the effectiveness of LRV

tracing is shown in Figure 14 where both tracing algorithms are applied to an image from a slit-lamp biomicroscope. The likelihoods are powerful enough to highlight vessels but ignore most of the background noise. This noise causes detection of spurious traces by the parallel-edge based algorithm. The new tracing algorithm correctly finds vessels that are missed by the parallel-edge based tracing.

IX. DISCUSSION AND CONCLUSION

We have introduced the Likelihood Ratio Vesselness (LRV) measure designed to improve the performance of vessel extraction and we have shown how it can be embedded in an efficient tracing framework. The LRV is formed from a 6-component vector of matched-filter and edge responses and their associated confidences. Each component is designed to help address the problem of detecting narrow and low-contrast vessels, while avoiding responses to other retinal image structures. In particular, the elongated template of the multiscale matched filter tends to preserve vessels that are only a pixel wide (and typically low-contrast), whereas isotropic measures such as the Hessian tend to substantially blur these structures. The edge responses are useful in distinguishing between offset edges near pathologies and true vessels. The confidence measures emphasize the shape of the intensity surface rather than the magnitude of the responses, enhancing the ability of the LRV to detect low-contrast vessels. The 6-component measurement vectors are mapped to the final LRV measure using pdfs learned from training data. Significantly, after designing multiscale filtering and training procedures there are no run time tuning parameters in the computation of this measure. The computation is made more efficient by using the Hessian to select matched-filter orientations and by embedding the LRV measure in a tracing framework. The latter substantially outperforms our existing retinal vessel tracing algorithm. Both quantitative and qualitative analysis on challenging retinal images have shown the effectiveness of the LRV measure. The new measure may be used in place of the Hessian and the matched filter in existing vessel detection and segmentation algorithms [31], [39]. Based on the experimental evaluations reported here, this should lead to substantially-improved results. The fact that the Hessian may be used to steer the application of the matched-filter and confidence measures makes the computation tractable.

Problems still remain on the way towards complete retinal vessel extraction algorithm. Most important is extracting the high-level organization of the characteristic venous and arterial trees of the retina. Doing so will help to distinguish veins and arteries, to avoid incorrect placement of vessels between two pathological structures, to separate parallel vessels from a single vessel with a central reflex, and will further improve the ability to detect low-contrast vessels that are not apparent from the local intensity structure alone. Together with the generalization of our confidence-based tracing algorithm to three dimensions, this forms the most important future direction for our research.

ACKNOWLEDGEMENTS

The authors thank Tomasz Malisiewicz for use of his generalized tracing implementation that made testing LRV tracing substantially easier. The authors thank Amitha Perera and Matt Turek for discussions that led to some of the ideas in this paper. The authors also thank the reviewers for their valuable comments which helped improve the manuscript in many ways.

REFERENCES

- [1] The Eye Diseases Prevalence Research Group, "Prevalence of age-related macular degeneration in the united states," *Arch. Ophthalm.*, vol. 122, no. 4, pp. 564–572, 2004.
- [2] —, "The prevalence of diabetic retinopathy among adults in the united states," *Arch. Ophthalm.*, vol. 122, no. 4, pp. 552–563, 2004.
- [3] —, "Prevalence of open-angle glaucoma among adults in the united states," *Arch. Ophthalm.*, vol. 122, no. 4, pp. 532–538, 2004.
- [4] J. A. Shoemaker, "Vision problems in the U.S." U.S. National Institute of Health, Tech. Rep., 2002.
- [5] M. Foracchia, E. Grisan, and A. Ruggeri, "Detection of optic disc in retinal images by means of a geometrical model of vessel structure," *IEEE Trans. Med. Imag.*, vol. 23, no. 10, pp. 1189–1195, Oct. 2004.
- [6] A. Hoover and M. Goldbaum, "Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels," *IEEE Trans. Med. Imag.*, vol. 22, no. 8, pp. 951–958, 2003, <http://www.parl.clemson.edu/stare/nerve/>.
- [7] E. Trucco and P. Kamat, "Locating the optic disk in retinal images via plausible detection and constraint satisfaction," in *Proc. IEEE Int. Conf. on Image Proc.*, vol. 1, Singapore, Oct.24–27 2004, pp. 155–158.
- [8] A. Can, C. Stewart, B. Roysam, and H. Tanenbaum, "A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, no. 3, pp. 347–364, 2002.
- [9] G. K. Matsopoulos, P. A. Asvestas, N. A. Mouravliansky, , and K. K. Delibasis, "Multimodal registration of retinal images using self organizing maps," *IEEE Trans. Med. Imag.*, vol. 23, no. 12, pp. 1557–1563, 2004.
- [10] J. C. Nunes, Y. Bouaoune, E. Delecquelle, and P. Bunel, "A multiscale elastic registration scheme for retinal angiograms," *Comput. Vis. Image Und.*, vol. 95, no. 2, pp. 129–149, 2004.
- [11] H. Shen, C. Stewart, B. Roysam, G. Lin, and H. Tanenbaum, "Frame-rate spatial referencing based on invariant indexing and alignment with application to laser retinal surgery," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 3, pp. 379–384, March 2003.
- [12] C. Stewart, C.-L. Tsai, and B. Roysam, "The dual-bootstrap iterative closest point algorithm with application to retinal image registration," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1379–1394, 2003.
- [13] T. Walter, J.-C. Klein, P. Massin, and F. Zana, "Automatic segmentation and registration of retinal fluorescein angiographies - application to diabetic retinopathy," in *Work. Comp. Assisted Fundus Img. Analysis*, May 2000.
- [14] F. Zana and J. C. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform," *IEEE Trans. Med. Imag.*, vol. 18, no. 5, pp. 419–428, 1999.
- [15] K. H. Fritzsche, "Computer vision algorithms for retinal vessel detection and width change detection," Ph.D. dissertation, Rensselaer Polytechnic Institute, Troy, New York, Dec 2004.
- [16] C. Heneghan, J. Flynn, M. O'Keefe, and M. Cahill, "Characterization of changes in blood vessel width and tortuosity in retinopathy of prematurity using image analysis," *Med. Image Anal.*, vol. 6, no. 4, pp. 407–429, 2002.
- [17] A. Houben, M. Canoy, H. Paling, P. Derhaag, and P. de Leeuw, "Quantitative analysis of retinal vascular changes in essential and renovascular hypertension," *J. of Hypertension*, vol. 13, pp. 1729–1733, 1995.
- [18] L. Hubbard, R. Brothers, W. King, L. Clegg, R. Klein, L. Cooper, R. Sharrett, M. Davis, and J. Cai, "Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the atherosclerosis risk in communities study," *Ophthalmology*, vol. 106, no. 12, pp. 2269–2280, 1999.
- [19] A. Sharrett, L. Hubbard, L. Cooper, P. Sorlie, R. Brothers, F. Nieto, J. Pinsky, and R. Klein, "Retinal arteriolar diameters and elevated blood pressure the atherosclerosis risk in communities study," *Am. J. of Epidemiology*, vol. 150, no. 3, pp. 263–270, 1999.
- [20] A. Stanton, B. Wasan, A. Cerutti, S. Ford, R. Marsh, P. Sever, S. Thom, and A. Hughes, "Vascular network changes in the retina with age and hypertension," *J. of Hypertension*, vol. 13, no. 12, pp. 1724–1728, 1995.
- [21] T. Wong, R. Klein, A. Sharrett, M. Schmidt, J. Pankow, D. Couper, B. Klein, L. Hubbard, and B. Duncan, "Retinal arteriolar narrowing and risk of diabetes mellitus in middle-aged persons," *J. of Am. Med. Assoc.*, vol. 287, pp. 2528–33, May 2002.
- [22] J. Jomier, D. K. Wallace, and S. R. Aylward, "Quantification of retinopathy of prematurity via vessel segmentation," in *Proc. 6th MICCAI*, 2003, pp. 620–626.
- [23] D. Koozekanani, K. L. Boyer, and C. Roberts, "Tracking the optic nervehead in oct video using dual eigenspaces and an adaptive vascular distribution model," *IEEE Trans. Med. Imag.*, vol. 22, no. 12, pp. 1519–1536, 2003.
- [24] F. Miles and A. Nutall, "Matched filter estimation of serial blood vessel diameters from video images," *IEEE Trans. Med. Imag.*, vol. 12, no. 2, pp. 147–152, 1993.
- [25] N. H. Solouma, A.-B. M. Youssef, Y. A. Badr, , and Y. M. Kadah, "A new real-time retinal tracking system for image-guided laser treatment," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 9, pp. 1059–1067, 2002.

- [26] R. Klein, B. Klein, S. Moss, T. Wong, L. Hubbard, K. Cruickshanks, and M. Palta, "Retinal vascular abnormalities in persons with type I diabetes," *Ophthalmology*, vol. 11, pp. 2118–25, 2003.
- [27] A. Pinz, S. Bernogger, P. Datlinger, and A. Kruger, "Mapping the human retina," *IEEE Trans. Med. Imag.*, vol. 17, no. 4, pp. 606–620, Aug 1998.
- [28] T. Teng, M. Lefley, and D. Claremont, "Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy," *IEE Med. & Bio. Eng. & Comp.*, vol. 40, no. 1, pp. 2–13, Jan 2002.
- [29] T. Walter, J.-C. Klein, P. Massin, and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy - detection of exudates in color fundus images of the human retina," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, Oct. 2002.
- [30] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Trans. Med. Imag.*, vol. 8, no. 3, pp. 263–269, September 1989.
- [31] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 203–210, 2000.
- [32] X. Jiang and D. Mojon, "Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 1, pp. 131–137, Jan. 2003.
- [33] A. Can, H. Shen, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from live high-resolution retinal fundus images using direct exploratory algorithms," *IEEE Trans. Inform. Technol. Biomed.*, vol. 3, no. 2, pp. 125–138, 1999.
- [34] D. Koozekanani, K. Boyer, and C. Roberts, "Retinal thickness measurements from optical coherence tomography using a markov boundary model," *IEEE Trans. Med. Imag.*, vol. 20, pp. 900–916, Sept. 2001.
- [35] 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 *Proc. 2nd MICCAI*, 1999, pp. 90–97.
- [36] V. Mahadevan, H. Narasimha-Iyer, B. Roysam, and H. Tanenbaum, "Robust model-based vasculature detection in noisy biomedical images," *IEEE Trans. Inform. Technol. Biomed.*, vol. 8, no. 3, pp. 360–376, 2004.
- [37] F. Zana and J.-C. Klein, "Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation," *IEEE Trans. Image Processing*, vol. 10, no. 7, pp. 1010–1019, 2001.
- [38] S. Aylward and E. Bullitt, "Initialization, noise, singularities, and scale in height-ridge traversal for tubular object centerline extraction," *IEEE Trans. Med. Imag.*, vol. 21, pp. 61–75, 2002.
- [39] J. Staal, M. Abramoff, M. Niemeijer, M. Viergever, and B. van Ginneken, "Ridge based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr 2004. [Online]. Available: <http://www.isi.uu.nl/Research/Databases/DRIVE/>
- [40] J. Staal, S. Kalitzin, and M. A. Viergever, "A trained spin-glass model for grouping of image primitives," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 27, no. 7, pp. 1172–1182, 2005.
- [41] C. Kirbas and F. Quek, "A review of vessel extraction techniques and algorithms," *ACM Computing Surveys*, vol. 36, no. 2, pp. 81–121, 2004.
- [42] A. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Proc. 1st MICCAI*, 1998, pp. 130–137.
- [43] C. Lorenz, I.-C. Carlsen, T. M. Buzug, C. Fassnacht, and J. Weese, "A multi-scale line filter with automatic scale selection based on the hessian matrix for medical image segmentation," in *Proc. 1st Int. Conf. Scale-Space Theory in Comp. Vis.*, 1997, pp. 152–163.
- [44] Y. Sato, S. Yamamoto, and S. Tamura, "Accurate quantification of small-diameter tubular structures in isotropic CT volume data based on multiscale line filter responses," in *Proc. 7th MICCAI*, Saint-Malo, France, 2004, pp. 508–515.
- [45] H. G. van Andel, E. Meijering, A. van der Lugt, H. Vrooman, and R. Stokking, "Vampire: Improved method for automated center lumen line definition in atherosclerotic carotid arteries in CTA data," in *Proc. 7th MICCAI*, Saint-Malo, France, 2004, pp. 525–532.
- [46] T. M. Koller, G. Gerig, G. Szekely, and D. Dettwiler, "Multiscale detection of curvilinear structures in 2-D and 3-D image data," in *Proc. ICCV*, 1995, pp. 864–868.
- [47] P. Majer, "On the influence of scale selection on feature detection for the case of linelike structures," *Int. J. Comp. Vis.*, vol. 60, no. 3, pp. 191–202, 2004.
- [48] K. Fritzsche, A. Can, H. Shen, C. Tsai, J. Turner, H. Tanenbaum, C. Stewart, and B. Roysam, "Automated model based segmentation, tracing and analysis of retinal vasculature from digital fundus images," in *State-of-The-Art Angiography, Applications and Plaque Imaging Using MR, CT, Ultrasound and X-rays*, J. S. Suri and S. Laxminarayan, Eds. Academic Press, 2003, pp. 225–298.
- [49] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *Int. J. Comp. Vis.*, vol. 30, pp. 117–156, Nov. 1998.
- [50] P. Meer and B. Georgescu, "Edge detection with embedded confidence," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, no. 12, pp. 1351–1365, Dec. 2001. [Online]. Available: citeseer.nj.nec.com/456269.html
- [51] M. Niemeijer, J. Staal, B. van Ginneken, M. Loog, and M. Abramoff, "Comparative study of retinal vessel segmentation methods on a new publicly available database," in *Proc. SPIE Conference Med. Imaging*, J. M. Fitzpatrick and M. Sonka, Eds., 2004, pp. 648–656.
- [52] R. Chan, W. Karl, and R. Lees, "A new model-based technique for enhanced small-vessel measurements in X-ray cine-angiograms," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 243–255, Mar. 2000.
- [53] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, and R. Kennedy, "Measurement of retinal vessel widths from fundus images based on 2-D modeling," *IEEE Trans. Med. Imag.*, vol. 23, no. 10, pp. 1196–1204, Oct. 2004.
- [54] F. K. H. Quek and C. Kirbas, "Vessel extraction in medical images by wave-propagation and traceback," *IEEE Trans. Med. Imag.*, vol. 20, no. 2, pp. 117–131, Feb. 2001.
- [55] G. Shechter, F. Devernay, E. Coste-Manière, A. Quyyumi, and E. R. McVeigh, "Three-dimensional motion tracking of coronary arteries in biplane cineangiograms," *IEEE Trans. Med. Imag.*, vol. 22, pp. 493–503, 2003.
- [56] J. Damon, "Properties of ridges and cores for two-dimensional images," *J. Math. Imag. Vis.*, vol. 10, no. 2, pp. 163–174, Mar. 1999.
- [57] D. Eberly, *Ridges in image and data analysis*. Kluwer Academic, 1996.
- [58] S. N. Kalitzin, J. Staal, B. M. ter Haar Romeny, and M. A. Viergever, "A computational method for segmenting topological point-sets and application to image analysis," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, no. 5, pp. 447–459, 2001.
- [59] T. Lindeberg, "Feature detection with automatic scale selection," *Int. J. Comp. Vis.*, vol. 30, pp. 79–116, Nov. 1998.
- [60] Y. Sato, S. Nakajima, H. Atsumi, T. Koller, G. Gerig, S. Yoshida, and R. Kikinis, "3D multi-scale line filter for segmentation and visualization of curvilinear structures in medical images," in *Proc. 1st Joint Conf. Comp. Vis., Virt. Real. and Robot. in Med. and Med. Robot. and Comp. Assist. Surg.*, Grenoble, France, 19–22 June 1997, pp. 213–222.
- [61] L. Gang, O. Chutatape, and S. Krishnan, "Detection and measurement of retinal vessels in fundus images using amplitude modified second-order gaussian filter," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 2, pp. 168–172, 2002.
- [62] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 9, pp. 891–906, September 1991.
- [63] E. Simoncelli and H. Farid, "Steerable wedge filters for local orientation analysis," *IEEE Trans. Image Processing*, vol. 5, no. 9, pp. 1377–1383, 1996.
- [64] Y. A. Tolias and S. M. Panas, "A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering," *IEEE Trans. Med. Imag.*, vol. 17, no. 2, pp. 263–273, Apr. 1998.
- [65] L. Ibáñez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide: The Insight Segmentation and Registration Toolkit (version 1.4)*. Kitware Inc., 2003.
- [66] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall PTR, 1998.
- [67] S. Konishi, A. L. Yuille, J. M. Coughlan, and S. C. Zhu, "Statistical edge detection: Learning and evaluating edge cues," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 1, pp. 57–74, Jan. 2003.
- [68] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. John Wiley and Sons, 2001.
- [69] S. Agarwal and D. Roth, "Learning a sparse representation for object detection," in *Proc. Seventh ECCV*, 2002, pp. 113–130.
- [70] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.
- [71] K. Al-Kofahi, S. Lasek, S. D., C. Pace, G. Nagy, J. N. Turner, and B. Roysam, "Rapid automated three-dimensional tracing of neurons from confocal image stacks," *IEEE Trans. Inform. Technol. Biomed.*, vol. 6, no. 2, pp. 171–187, 2002.
- [72] C.-L. Tsai, C. Stewart, B. Roysam, and H. Tanenbaum, "Model-based method for improving the accuracy and repeatability of estimating vascular bifurcations and crossovers from retinal fundus images," *IEEE Trans. Inform. Technol. Biomed.*, vol. 8, pp. 122–130, 2004.