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

Michal Sofka, and Charles V. Stewart *

August 16, 2005

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 learning technique is applied to map 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 intensity Hessian show substantial improvements both qualitatively and quantitatively. When the Hessian is used in place of the matched filter, similar but less-substantial improvements are obtained. Finally, the new vesselness likelihood ratio is embedded into a vessel tracing framework, resulting in an efficient and effective vessel extraction algorithm.

Index terms: vessel extraction, vessel tracing, matched filters, scale space, retina images, vessel confidence, learning, likelihood ratio

---

*M. Sofka and C.V. Stewart are with the Department of Computer Science, Rensselaer Polytechnic Institute, Troy, New York 12180–3590 (email:{sofka,stewart}@cs.rpi.edu).

1 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 [58, 59, 60]. It is believed that half of all blindness can be prevented [50], 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 [10, 17, 62], image registration [6, 38, 42, 49, 56, 66, 68], change detection [14, 16, 19, 20, 47, 55, 67], pathology detection and quantification [23], tracking in video sequences [30, 40, 52], and computer-aided screening systems [26, 43, 57, 65] 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 [9, 18, 40], adaptive thresholds [18, 22], intensity edges [5, 29], region growing [37], statistical inferencing [35], mathematical morphology [16, 37, 69], and Hessian measures [3, 23, 53, 54]. This wide range of techniques closely corresponds to the suite of methods that have been applied throughout the medical image analysis literature [25]. Of particular note, the recent literature has been dominated by Hessian-based methods because of their utility in characterizing the elongated structure of vessels [3, 11, 33, 46, 64].

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 a particular challenge for automatic vessel extraction 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 primary 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 central-reflex, while important, is less of a concern here because it typically occurs in larger, more-easily-detected vessels. We will return to this issue much later in the paper.

In order to motivate our technical approach, we start by examining the measurements that indicate the potential presence of a vessel in an image neighborhood. These "vesselness measures", which form the basis of many techniques in the literature, primarily include matched filters, Hessian-based measures, and parallel edge models. The limitations of some of these measures are illustrated in Figure 2. All common vesselness measures produce substantial responses to non-vascular structures such as the optic disc and the pathologies. All measures, even those designed to compensate for edge responses [27, 33, 36], 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. Similar problems also appear in the results of edge-based tracing techniques [5, 13].

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 [18], we introduce a multi-scale matched filter for vessels, deriving an appropriate normalizing constant to allow 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 [39], that determines how close an image region

is to 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, producing a six degree-of-freedom measurement vector at each pixel. Then, we use a learning 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.

Intuitively, a strong Likelihood Ratio Vesselness (LRV) measure requires significant matched filter and vessel confidence responses along with responses showing apparent edges at the boundaries of the vessels. We show that the new measure outperforms the multiscale matched filter and existing Hessian-based measures using quantitative analysis of results on retinal images and using visual inspection of results on retinal images and on a 2D angiogram. We also show how the multiscale 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 [3, 37, 41, 53, 54] and in related applications [8, 25, 34, 44, 48]. 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 [5, 13] 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 our earlier parallel-edge tracing algorithm [5, 13]. In addition to improving tracing, this has the added benefit of showing how the LRV measure may be applied efficiently. The paper concludes with a summary of remaining

issues in retinal vessel extraction.

2 Background

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

2.1 Matched Filters

Working from a linear filtering perspective, let the input function (image) be $f(\mathbf{x})$ and let the filter be $s(\mathbf{x})$. Then the output is

$$h(\mathbf{x}) = \int\limits_{-\infty}^{\infty} s(\mathbf{x} - \mathbf{x}')f(\mathbf{x}')d\mathbf{x}' \tag{1}$$

The optimal filter (maximizing the signal-to-noise ratio) is known as the *matched filter* and its shape can be obtained by reversing the shape of the detecting signal [24]. In this way, the filter's shape is *matched* to the signal's shape.

Chaudhuri [9] et al. used a matched filter to detect vessels in retinal images. The filter gives maximum response when its orientation and shape is the same as the intensity profile. Vessels are modelled as piecewise linear segments with Gaussian cross sections. Twelve Gaussian templates at different orientations and a single scale are used. In Hoover's vessel segmentation algorithm [18] 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 [15] proposes amplitude-modified second-order differential of Gaussian filter to detect vessels at scales that match their widths. This is achieved by change of amplitude which is necessary so that responses can be combined over scales.

Steerable filters are introduced in [12] and extended in [51]. In particular, Freeman [12] shows an example of applying steerable filters to vessel enhancement. The second derivative of the Gaussian (G_2) along any direction θ is expressed as a combination of the derivatives at orientations $0, 60$, and 120 degrees. The G_2 filter is steered adaptively along direction of

dominant vessel orientation. The result is less noisy than the output of an isotropic filter of the same frequency passband [12]. It is computed at a single scale.

The main disadvantage of matched filters is their computational cost. They are usually implemented as a convolution of an image with a set of oriented segments, which is especially expensive when computed at multiple scales. We will return to this issue in Sections 6 and 7.

2.2 Hessian-Based Approaches to Ridge Detection

Several papers have introduced techniques for vessel or ridge extraction based on the eigen decomposition of the Hessian computed at each image pixel [3, 11, 31, 32, 45]. Eigenvalues are used in rules to decide if a particular location is a vessel (ridge) pixel. When a pixel passes this test, the eigenvector corresponding to the smaller (in absolute value) eigenvalue points along the ridge. The signs of the eigenvalues determine bright or dark structures. The Hessian is computed at different scales to account for different vessel widths. We choose two specific measures here, [11] and [31], for our analysis. They are used in other papers and many of them have also used slightly-modified, and often simpler versions of these [3, 46, 48, 53, 54, 64].

Frangi et al. [11] proposed using ratios of the eigenvalues $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$ to determine scores $\mathcal{R}_\mathcal{B} = |\lambda_1|/\sqrt{|\lambda_2\lambda_3|}$ and $\mathcal{R}_\mathcal{A} = |\lambda_2|/|\lambda_3|$ and computing the Frobenius norm, $\mathcal{S}$, of the Hessian to measure overall strength. For tube-like structures, including vessels, $\mathcal{R}_\mathcal{B}$ should be low and $\mathcal{R}_\mathcal{A}$ should be high — the latter to distinguish between edges (sheets) and tubes in 3D. Specifically, the vesselness measure for images where vessels are brighter than the background is

$$
\mathcal{V}_0 = \begin{cases} 0 & \lambda_2 > 0 \text{ or } \lambda_3 > 0 \\ \left(1 - \exp\left(-\frac{\mathcal{R}_\mathcal{A}^2}{2\alpha^2}\right)\right) \exp\left(-\frac{\mathcal{R}_\mathcal{B}^2}{2\beta^2}\right) \left(1 - \exp\left(-\frac{\mathcal{S}^2}{2c^2}\right)\right) & \text{otherwise} \end{cases}
\tag{2}
$$

with parameters α, β, and c. The Hessian is computed in scale space and normalized so responses can be combined by taking maximum across scales. In 2D, the first exponential is left out, which explains why this Hessian measure is not particularly effective at distinguishing between edges and vessels. Parameter values suggested in [11] were used to produce the vesselness response image in Figure 2(c).

In Lindeberg [31] 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 in scale space as

$$\mathcal{A}_{\gamma-norm}L = t^{2\gamma}((L_{xx} - L_{yy})^2 + 4L_{xy}^2), \tag{3}$$

where L_{xx}, L_{yy} and L_{xy} are from the Hessian. The factor $t^{2\gamma}$, where t is the variance of the Gaussian smoothing and γ is a parameter, normalizes responses across scales. Responses are combined to yield a single response image (Figure 2(b)) by taking the maximum across scales at each pixel with value of $\gamma = 3/4$.

In the context of retinal image analysis, Staal et al. [53] describes a vessel segmentation algorithm that combines a grouping technique with low-level, Hessian-based classification of pixels into ridge points of local minima, ridge-points of local maxima, and non-ridge points. These are then grouped by region growing into sets that approximate straight line elements. A k-Nearest Neighbor classifier is subsequently applied to segment vessels using an extensive set of geometric and image intensity measures taken from these sets. One area of future work would be to combine our more sophisticated initial processing measures with the grouping and classifier techniques in [53].

Most Hessian-based measures attempt to distinguish between responses to vessel-like structures and responses to other structures through interpretation of the eigenvalues (see [4, 36] in addition). One of the contributions of this paper is that we develop and combine a number of different measures to determine an overall vesselness response, including measures displaced from the center pixel. The advantage of this approach is that each measure provides more specific information — the strength of the response, the confidence of the response, and the exisitence of the nearby edge responses — that will more effectively distinguish vessels from other structures.

2.3 Model-Based Vasculature Extraction

In [35] a model-based technique is proposed for vessel detection in noisy retinal images. The idea is to estimate the parameters of a model in a region surrounding a vessel and compare the estimate against a background noise model in a hypothesis testing framework. Several different

tests are proposed. Although these algorithms are demonstrated to perform well on noisy retina video images, they strongly respond to non-vessel structures and sensitivity to small vessels is lost as shown in the example in Figure 2(d) using the best technique from [35]. These responses should be detected and eliminated through use of a more extensive suite of local, parametric appearance models.

2.4 Vessel Tracing

Many techniques in the literature propose to extract vessels using an exploratory framework that detects vessels by "tracing" along the vessels starting from seed points. Aylward and Bullitt [3] apply Hessian-based vesselness measures in a ridge traversal method. The algorithm distinguishes between plates and tubes by conditions on eigenvalues. Ridge traversal utilizes directions of the eigenvectors. Scale is selected by estimating the object's radius with a convolution kernel. The traversal based on eigen analysis exploits the same properties as Hessian-based measures (Equations 2 and 3).

The exploratory algorithm of Can et al. [5], extended in Fritzsche [13], uses a set of two oriented edge detection kernels to detect parallel vessel boundaries in tracing retinal images. Seeds for tracing are detected at local intensity minima along 1-D horizontal and vertical grid lines. Experimentally, we have found that this technique performs most poorly on narrow, low-contrast vessels – exactly what we are concerned with in this paper.

2.5 Other Techniques in Retina Vessel Extraction

A wide range of techniques has been applied to extracting retinal vasculature. Here is a brief summary. 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 [18, 22] due to illumination changes and pathologies. In [5, 29] vessels are found by finding their intensity edges. These algorithms work well for wide and high-contrast vessels, but not for low-contrast and thin vessels. Methods based on mathematical morphology [16, 37, 69] use linear structuring element to enhance vessels. Usually only large vessels are detected well, and other processing steps are needed to eliminate responses from structures that fit the same morphological

description. Tolias and Panas [61] use fuzzy clustering of the vessel profiles in a tracing algorithm starting from the optic disc.

3   Multiscale Matched Filters

We start by defining a multiscale matched filter for vessels, using a Gaussian vessel profile. This has good response to low-contrast and narrow vessels, as illustrated in Figure 2, but often responses to non-vessel structures as well. Reasons for this, which will become clear as we proceed, motivate the additional measures that we combine with the matched filter.

The matched filter at one scale is implemented by maximizing the responses over a discrete set of kernel orientations. 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}. In a coordinate system rotated to align u with the tangential direction and v with the normal direction, the matched filter response is

$$M(R; u, v; t_u, t_v) = -\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} g_{vv}(v - v'; t_v)g(u - u'; t_u)f(u', v'; t_0)du'dv', \qquad (4)$$

where t_u and t_v are the variances of Gaussian in each direction, and generally $t_u > t_v$.

In order to detect vessels at a variety of widths, we apply the matched filter at multiple scales and then combine the responses across scales. Unfortunately, the output amplitudes of spatial operators such as derivatives or matched filters generally *decrease with scale.* To compensate for this effect, Lindeberg [31] 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. Then, given an ideal profile, which in our case is a Gaussian of variance t_0, the value of γ is chosen for which the peak response in scale space occurs at $t = t_0$. No normalization is needed in the tangential direction because the profile is expected to be constant along this direction. Therefore, we may consider only the normal direction in deriving the normalization factor. Using this together with the properties of

8

the Gaussian, the γ-normalized matched filter response to a Gaussian cross-section at location $(0,0)$ and scale t (in the normal direction) may be calculated as

$$M_{\gamma-norm}(R; 0, 0; t) = -\int\limits_{-\infty}^{\infty} \frac{t^{\frac{\gamma}{2}}}{2\pi\sqrt{t_0 t^5}}(v'^2 - t)e^{-\frac{v'^2}{2}\left(\frac{t+t_0}{tt_0}\right)}dv'$$

$$= \frac{t^{\frac{\gamma}{2}}}{(2\pi)^{\frac{1}{2}}(t_0 + t)^{\frac{3}{2}}}. \tag{5}$$

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

$$\frac{\frac{\gamma}{2}t^{\frac{\gamma}{2}-1}(t_0 + t) - \frac{3}{2}t^{\frac{\gamma}{2}}}{(t_0 + t)^3} = 0. \tag{6}$$

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 (4) at scale t is $t^{3/4}$.

Using this normalization, the responses may now be fairly compared across scales. Computing the best response, r_v, at each pixel simply requires computing the maximum response over all scales and orientations. An example of a resulting "response image" has already been shown in Figure 2(a). Once again, as compared to the Hessian-based operators discussed in Section 2, the matched filter appears to provide 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 measures" we define next.

4   Confidence And Edge Measures

An intuitive insight for why we need to go beyond use of the matched filter alone is to consider what happens when it is applied to an ideal 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 response is 0 for pixels centered on the edge. On the other hand, at pixels offset along the normal to the edge the responses start to increase in magnitude, reaching

a peak at distance $\sqrt{t}$ (see Figure 3). Quite frequently these responses are much stronger than for lower contrast vessels, as illustrated around the retina boundary, the pathologies, and the optic disc in Figure 2. Similar problems occur with Hessian-based measures. By combining these responses with confidences and edge measures, as described in this section, we will be able to distinguish these responses from those of true vessels.

4.1 Vessel Confidence Measure

Consider the matched filter convolution at a single pixel, scale and orientation from the viewpoint of linear algebra.[1] The matched filter kernel weights and the associated pixel intensity values can be "stacked" into two vectors, which we will call $\mathbf{t}_m$ and $\mathbf{a}$, respectively. The convolution is a dot-product $\mathbf{a}^\top \mathbf{t}_m$. If $\mathbf{t}_m$ is normalized (without loss of generality) to unit length, this may be considered as a projection of $\mathbf{a}$ onto direction $\mathbf{t}_m$. As in any projection, a great deal of information is lost. In particular, the projection $\mathbf{a}^\top \mathbf{t}_m$ *says nothing* about how well $\mathbf{a}$ fits the profile represented by the kernel $\mathbf{t}_m$. This is simply because $\mathbf{a}$ is unnormalized. A vector $\mathbf{a}$ formed from an ideal, low-contrast vessel will be nearly parallel to $\mathbf{t}_m$, but may have a much smaller projection than a region vector $\mathbf{a}$ formed by an offset step edge (as previously discussed) or even a region made up of pure but large magnitude noise.

We now introduce a vessel "confidence measure" (borrowing terminology from [39]). At each pixel, let $\mathbf{t}$ be the template formed by the ideal Gaussian profile — i.e. Gaussian along the vessel normal and constant along the vessel tangent — with orientation and scale taken from the peak matched filter response. Let $\mathbf{a}(\mathbf{x})$ be a vector of the image intensity values taken from pixels overlaid by the template. Vectors $\mathbf{a}(\mathbf{x})$ and $\mathbf{t}$ are independently centered (so that the sum of the values in each vector is 0) and then normalized to unit length. We define the confidence measure at this pixel as

$$\eta_v(\mathbf{x}) = \mathbf{a}(\mathbf{x})^\top \mathbf{t}, \tag{7}$$

which can be viewed as normalized cross-correlation 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

---

[1]This discussion is adapted from the work of [39] on edge detection.

of an ideal step edge, and 0.79 for the peak response when offset by $\sqrt{t}$ normal to a step edge. Interestingly, the latter response is still relatively high, motivating the need for further measures. An example result is shown in Figure 4.

4.2 Edge Measures

Our additional measures are developed from the simple observation that vessels have edges on their boundaries. We therefore measure the edge strength and the confidence (similar to the above) on the two sides of a purported vessel and combine these measures with the matched filter and confidence response at the center of the vessel. Intuitively, this should be enough to suppress the response at an offset step edge because only one edge will exist. Stated more strongly, the fact that the second edge is missing is the most significant difference between a vessel and an (offset) step edge.

Edge convolutions are computed in scale space as in [31] and combined by taking the maximum response over scale, producing a gradient vector at each pixel. There is nothing novel about this and we use standard implementations available in ITK [21]. At a given pixel location $\mathbf{x}$ being tested as a vessel, let $\mathbf{n}$ be the (unit) normal to the vessel, and let $w = 2\sqrt{t_0}$ be the vessel width, where t_0 is the variance of the Gaussian vessel profile. The two edges of the vessel should exist at approximately $\mathbf{x} + (w/2)\mathbf{n}$ and $\mathbf{x} - (w/2)\mathbf{n}$, and the normals to the edges should be in the outward directions (for dark vessels), away from the vessel center. Therefore, we look along a small (3 pixel) interval in the normal direction centered at $\mathbf{x} + (w/2)\mathbf{n}$ for the edge pixel with maximum projection of the edge gradient on the outward normal ($\pm\mathbf{n}$). We denote the projected gradient magnitudes of the stronger and weaker of the two maximum responses by $r_{es}(\mathbf{x})$ and $r_{ew}(\mathbf{x})$, respectively. Distinguishing between stronger and weaker vessel edges helps to train the distributions more accurately. Edge confidence measures are computed at the two edge locations giving these maximum responses, as above, using an ideal step edge profile oriented with $\mathbf{n}$ as its outward normal (dark-to-light transition for dark vessels). These 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). An example result is in Figure 5. Confidences of both edges are high at vessel locations. The weaker edge is large only at vessel locations, and is missing around edges and other non-vessel structures.

Only combined measures can separate the image content. Even though the edge confidences are not as powerful as the other measures, they still contribute to better vessel extraction results.

The combined vessel and edge computations give us three strength measures and three confidence measures per pixel. Overall, at each pixel location, we form a six dimensional vector $\boldsymbol{\Theta}(\mathbf{x}) = (r_v(\mathbf{x}), r_{es}(\mathbf{x}), r_{ew}(\mathbf{x}), \eta_v(\mathbf{x}), \eta_{es}(\mathbf{x}), \eta_{ew}(\mathbf{x}))$.

5　Learning and Decision Criteria

Our next step is to map the 6-D response/confidence measurement vector $\boldsymbol{\Theta}$ into a single "vesselness" measure at each pixel. The measure we choose is the optimal test statistic by the Neyman-Pearson Lemma [24]: the *likelihood ratio*,

$$\text{LRV}(\boldsymbol{\Theta}) = \frac{p(\boldsymbol{\Theta}|\text{on})P(\text{on})}{p(\boldsymbol{\Theta}|\text{off})P(\text{off})}. \tag{8}$$

Here $P(\text{on})$ is the prior probability of a pixel being on the centerline of a vessel, $P(\text{off}) = 1 - P(\text{on})$, and $p(\boldsymbol{\Theta}|\text{on})$, $p(\boldsymbol{\Theta}|\text{off})$ represent corresponding conditional probability density functions (pdfs) for measurement vector $\boldsymbol{\Theta}$. LRV is the desired "Likelihood Ratio Vesselness".

To train these pdfs we use the manual segmentation of 20 retinal images containing pathologies made publicly available by Hoover's group [18]. A thinning algorithm (Matlab) is applied to obtain centerlines from these segmentations. Training $P(\text{on})$ is accomplished simply through counting the number of resulting centerline pixels and dividing by the total number of pixels. Training the conditional pdfs is based on a method from [28]. Two histograms are formed, one to represent $p(\boldsymbol{\Theta}|\text{on})$ for the "on" pixels and one to represent $p(\boldsymbol{\Theta}|\text{off})$ for the "off" pixels. The histograms are normalized to unit magnitude. They may then be probed for any $\boldsymbol{\Theta}$ to yield approximate values for the two conditional probabilities. Since these histograms are 6-dimensional, however, 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:

$$P(\text{error}) \leq P^{\xi}(\text{on})P^{1-\xi}(\text{off}) \int p^{\xi}(\boldsymbol{\Theta}|\text{on})p^{1-\xi}(\boldsymbol{\Theta}|\text{off})d\boldsymbol{\Theta}, \tag{9}$$

where $0 \leq \xi \leq 1$. The right hand side of the inequality is minimized over ξ. In the discrete domain, the integral is replaced by a sum over histogram bins representing the pdfs. During bin boundary selection every dimension of vector $\boldsymbol{\Theta}$ is processed separately as a one-dimensional histogram. At the beginning, all samples are contained within one bin. Each boundary is then selected such that Chernoff bound (9) for the two category classification is minimized. Every additional boundary decreases the error until an asymptotic limit is reached.

Figure 6 shows that the Chernoff bound quickly reaches its asymptotic value 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, a manageable size. The training also gives insight to the discriminating power of each measure we have developed. In particular, the left plot of the figure shows Chernoff bound as a function of the number of bins for vessel, weak edge, and strong edge ranks, 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. They contribute significantly to making likelihoods more reliable.

6   Experimental Evaluation

6.1   Algorithm Summary

A summary of the overall computation is as follows. At each pixel, the maximum response of the normalized matched filter is found across all orientations and scales. Multiscale edge detection is applied independently to find the strengths of outward edges at the boundaries of vessels. Confidence measures are computed for the vessel, the strong edge, and the weak edge. These form a vector of six measurements, $\boldsymbol{\Theta}$, at each pixel. Finally, this vector is converted to a likelihood ratio (8) using the learned histogram-based pdfs. This likelihood ratio is our final "vesselness" measure, LRV, intended to be used similarly as the vessel measure of Frangi et al.

[11] (2) and the ridge measure of Lindeberg [31] (3). Both of these may be inserted into Θ in place of the matched filter responses. Non-maximum suppression is applied to the LRV values in the direction normal to the vessels. A threshold is then applied to the surviving LRV values to make the final vessel / non-vessel decisions.

Clearly the most expensive part of this algorithm is the multiscale matched filter. After evaluating the overall performance of the algorithm, we will show several ways in which the computational cost may be reduced dramatically. In Section 7, we take this step further to show how the likelihood-ratio vesselness measure may be applied in an exploratory tracing framework.

6.2   Evaluation Technique

We evaluate our technique both quantitatively and qualitatively using the publicly-available data set from Hoover [18], choosing this over other data sets (e.g. [41]) because it includes more images with significant pathologies. Of the two manual segmentations provided, we use the one that contains more small vessels to demonstrate success of our technique.

Our quantitative analysis is based on thinned versions of the manual segmentations, even though other published results are based on binary segmentations of entire vessels. There are several 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 measure is 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 [56] and optic disc detection [17] 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.

In order to use the manual segmentations for training our algorithm, the maxima (over orientation and scale) matched-filter responses are found in the normal directions within a manually-segmented vessel. These locations are deemed "on" vessel and all other pixel locations throughout the images are deemed "off". The counts of these locations provides the prior probabilities, $P(\text{on})$ and $P(\text{off})$. At the "on" locations, the remaining 5 components of the 6D measurement vector

are computed and the conditional probability density $p(\text{on}|\Theta)$ may then be trained. Similarly, the "off-vessel" conditional pdf is computed based on the 6D measurements collected at each off-vessel pixel. This is done using a leave-one-out strategy on the Hoover dataset.

With this training, for any 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. 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 the thinned manual segmentation.

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 false positives (see discussion in [1]). 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 as defined for ROC curves. (1-Precision) is defined as a number of false positives divided by the total number of detections. In another words, it defines the fraction of detections that are wrong.

6.3 Overall Results

We start with an overall, quantitative analysis using results combined across all 20 image in the Hoover data sets (Figure 7). We gathered results using the LRV measure described above and using the LRV measure with Frangi's Hessian vesselness measure used in place of the matched filter. 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 — 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

15

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 raw 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 off 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 much higher-contrast than even the more distinct vessels. Hence, for very high thresholds, these are the only responses that survive thresholding.

These quantitative results show that our likelihood ratio vesselness measure (LRV) is substantially better than the simpler matched-filter and Hessian measures that form the basis for the most successfull vessel extraction and segmentation techniques [3, 18, 22, 41, 53, 54]. Thus, it should be possible to substitute the LRV measure into other existing techniques to improve their performance. One concern in doing this is the computation cost of the matched filter, especially applied at multiple scales. Fortunately, much of this may be eliminated by using the Hessian eigenvectors to choose a single matched-filter direction at each pixel and by searching for scale-space peak in a fine-to-coarse manner. These together provide a 93% speed-up with no loss of performance, yielding an overall cost of under 4 minutes on the images on Hoover's database. Further optimizations are likely possible since this still computes a matched-filter and a confidence at each pixel.

6.4 Thin Vessels, Low-Contrast Vessels and Pathologies

The second part of our analysis considers the specific problems motivating our work — detecting thin and low contrast vessels while avoiding false detection of vessels near pathologies and other non-vascular structures. Already we have implicitly shown a substantial quantitative improvement on non-vascular structures because these form the bulk of the false-positive responses shown in the curves of Figure 7.

We can develop a fairly rough but indicative quantitative analysis of the performance on

16

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 are shown in Figure 8. Note that when the LRV measure is set to detect 50% of all narrow vessels, 35% of the vessels detected are false positives. For the matched-filter alone, this number is 65%. When LRV is set to detect 40% of all low contrast vessels, 55% of the vessels detected are false positives. For the matched filter alone, it is almost 80%. Although this evaluation shows still large number of false positives, we note that these narrow vessels are exactly where different manual segmentations disagree most substantially [18]. Moreover, when we look in detail at the image intensity surrounding low-contrast manual segmentations that the LRV measure misses we do not see the vessels. It seems that higher-level perceptual information is used by the human observer to identify these vessels.

In order to give a qualitative feel for the results, several "chips" from the most challenging images are shown in Figure 9. These include a variety of pathologies and thin vessels which are so subtle they sometimes completely fade into the noise. The first thing to observe about the results are that the matched filter (2nd column of Figure 9) highlights many of the vessels, but produces even stronger responses at non-vascular structures, such as the optic disk (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 receive strong responses. These cases (row 5, bottom right of the image) usually involve two bright pathologies separated by a narrow strip in which the retinal darker retinal background appears. To obtain binary centerlines (4th column) a threshold $\tau = 2.0$ is applied to the LRV measure. (Note that when the vesselness likelihood ratio is 1.0, it indicates an equal chance that the current location represents vessel and a non-vessel. With increasing values of likelihood, we are more sure of a vessel presence. A value of $\tau = 2.0$ was used for all results in this paper.) After applying the threshold most of the narrow vessels still appear, although on occasion these are broken up into short segments. It should be possible to eliminate

this thresholding effect through adaptive thresholding [7] or higher-level grouping techniques. Overall, the results on these challenging images appears qualitatively to be better than other results in the literature, although a direct, quantitative comparison on these challenging image regions has not yet been performed.

As a final evaluation of our LRV measure, Figure 10 shows the result of applying it to an angiogram of the cerebral vasculature. No changes were made to the algorithm to generate this result — the retina image training results were used. Notice that even very thin vessels have high responses and that the angiogram boundary was successfully removed by the algorithm.

7   Vessel Tracing

We now show how to employ the effective new LRV measure in an efficient vessel tracing framework [3], substituting the LRV measures for our earlier parallel-edge vesselness measure [5, 13]. This has the dual-purpose of showing how the LRV measure can be applied efficiently and how vessel tracing results can be substantially improved.

All exploratory tracing algorithms follow the same overall strategy. Starting from a set of oriented seed locations, they recursively trace the vascular (or neuronal [2]) 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 [5, 63].

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 matched filter during tracing and then compute and apply the 6-D measurement vector and resulting likelihood ratio subsequently to the "over-traced" results.

Here are a few important details:

- The seed points we use are the same as the seeds in [13] — local intensity minima along smoothed vertical and horizontal lines through the images. Next, we compute the local, multiscale Hessian around seed locations and determine the vessel tangent direction from the eigenvector corresponding to its 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 highest strength so that when we start tracing, we use the most reliable points first. Only points that have positive matched filter response are retained.

- The iterative tracing procedure is started from each point from a list sorted based on vessel strengths. At step $i \geq 0$ of tracing, 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 change in orientation (in increments of 10 degrees), ± 1 steps in scale space, and ± 2 pixel steps in the normal direction. If 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 post-processing 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 vessel likelihoods. It may be possible to do even better using an adaptive thresholding technique [7].

8   Tracing Results

We evaluate the effectiveness of this likelihood ratio tracing algorithm in side-by-side comparison with two other techniques. One is our likelihood ratio (LRV) applied image-wide with non-maximum suppression. This comparison allows us to see how much is lost by embedding LRV in a tracing framework. The second is our earlier, parallel-edge tracing algorithm [5] as refined

in [13]. The goal of this comparison is to evaluate the improvement over an existing tracing algorithm.

Qualitative results can be seen in Figure 12. The figure shows cropped image regions with parallel and likelihood based tracing as well as likelihoods evaluated at each pixel of the image. The same seed points were used for both algorithms. Thinned ground truth 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 third and fourth example), more accurately determines when to stop tracing in unclear regions (thin vessels near pathologies in the first and third example), correctly ignores pathologies (first and fourth example) and finds even thinnest vessels that usually have very low contrast (fifth example). Notice in a number of cases false positives are marked along false negatives (green line along red line) because of localization errors in ground truth. The tracing results are comparable to the LRV measure evaluated at every point. The LRV measure correctly detects even vessels with central reflex (second example). When the central reflex is severe, more complicated vessel models could be used (both for computing responses and confidences). Alternatively, dark sides of the vessel often detected as two parallel lines could be replaced by a single centerline based on high level reasoning.

We also directly compared parallel-edge and LRV-based tracing (Figure 13) using ground truth labelling. Vessel pixels detected by both methods (black), LRV-based tracing only (green) and parallel edge based only (red) are shown. LRV-based tracing outperforms parallel edge tracing in detecting low contrast and thin vessels and produces more complete centerline extraction. The largest improvement is for the thinnest vessels. Occasionally, a vessel is missed by both algorithms due to the lack of a seed point. The image in bottom right shows comparison of LRV-based tracing and LRV evaluated at each pixel followed by non-maximum suppression and thresholding below 2.0. Vessel pixels detected by both (black), only by computing LRV at each pixel (green) and only by LRV-based tracing (red) are shown. Tracing does not detect some vessels (eg. area around the optic disc in this example) due to the lack of an initial point. Non-maximum suppressed likelihoods detect even these vessels and give best results at the expense of exhaustive computation. It takes less than 30 seconds to extract traces from a 700x605 retina

image (unoptimized code on Pentium IV, 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 the eigenvector direction cannot be computed reliably.

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 image. The likelihoods are 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.

9   Discussion and Conclusion

We have introduced the Likelihood Ratio Vesselness (LRV) measure designed to improve the performance of vessel extraction and we 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, adding to the ability of the LRV to detect low-contrast vessels. The 6-component measurement vectors is mapped to the final LRV measure using pdfs learned from training data. Significantly, there are no tuning parameters in the computation of this measure. The computation can be made more efficient by using the Hessian to select matched-filter orientations and by embedding the LRV measure in a tracing framework. Both quantitative and qualitative analysis on challenging retinal images has shown the effectiveness of the technique. The new LRV measure may be used in place of the Hessian and the matched filter in existing vessel detection and segmentation algorithms [18, 53], and based on the experimental evaluations reported here, it should lead to

substantially-improved results.

Problems still remain on the way towards complete retinal vessel extraction algorithm. Most important is extracting the high-level organization of the characteristic venuous 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.

10 Acknowledgements

References

[1] S. Agarwal and D. Roth. Learning a sparse representation for object detection. In *Proceedings of the Seventh European Conference on Computer Vision*, pages 113–130, 2002.

[2] 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 Transactions on Information Technology in Biomedicine*, 6(2):171–187, 2002.

[3] S. Aylward and E. Bullitt. Initialization, noise, singularities, and scale in height-ridge traversal for tubular object centerline extraction. *IEEE Transactions on Medical Imaging*, 21:61–75, 2002.

[4] S. Aylward, J. Jomier, S. Weeks, and E. Bullitt. Registration of vascular images. *International Journal of Computer Vision*, 55(2-3):123–138, 2003.

[5] 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 Transactions on Information Technology in Biomedicine*, 3(2):125–138, 1999.

[6] 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 Transactions on Pattern Analysis and Machine Intelligence*, 24(3):347–364, 2002.

[7] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, Nov. 1986.

[8] R. Chan, W. Karl, and R. Lees. A new model-based technique for enhanced small-vessel measurements in X-ray cine-angiograms. *IEEE Transactions on Medical Imaging*, 19(3):243–255, Mar. 2000.

[9] 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(3):263–269, September 1989.

[10] M. Foracchia, E. Grisan, and A. Ruggeri. Detection of optic disc in retinal images by means of a geometrical model of vessel structure. *IEEE Transactions on Medical Imaging*, 23(10):1189–1195, Oct. 2004.

[11] A. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. In *Proceedings of the 1st International Conference of Medical Image Computing and Computer-Assisted Intervention (MICCAI 1998)*, pages 130–137, 1998.

[12] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(9):891–906, September 1991.

[13] K. Fritzsche, A. Can, H. Shen, C. Tsai, J. Turner, H. Tanenbuam, C. Stewart, and B. Roysam. Automated model based segmentation, tracing and analysis of retinal vasculature from digital fundus images. In J. S. Suri and S. Laxminarayan, editors, *State-of-The-Art Angiography, Applications and Plaque Imaging Using MR, CT, Ultrasound and X-rays*, pages 225–298. Academic Press, 2003.

[14] K. H. Fritzsche. *Computer Vision Algorithms for Retinal Vessel Detection and Width Change Detection*. PhD thesis, Rensselaer Polytechnic Institute, Troy, New York, Dec 2004.

[15] L. Gang, O. Chutatape, and S. 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.

[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. *Medical Image Analysis*, 6(4):407–429, 2002.

[17] A. Hoover and M. Goldbaum. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. *IEEE Transactions on Medical Imaging*, 22(8):951–958, 2003. http://www.parl.clemson.edu/stare/nerve/.

[18] 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(3):203–210, 2000.

[19] A. Houben, M. Canoy, H. Paling, P. Derhaag, and P. de Leeuw. Quantitative analysis of retinal vascular changes in essential and renovascular hypertension. *Journal of Hypertension*, 13:1729–1733, 1995.

[20] 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*, 106(12):2269–2280, 1999.

[21] 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.

[22] 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, Jan. 2003.

[23] J. Jomier, D. K. Wallace, and S. R. Aylward. Quantification of retinopathy of prematurity via vessel segmentation. In *Proceedings of the 6th International Conference of Medical Image Computing and Computer-Assisted Intervention (MICCAI 2003)*, pages 620–626, 2003.

[24] S. M. Kay. *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall PTR, 1998.

[25] C. Kirbas and F. Quek. A review of vessel extraction techniques and algorithms. *ACM Computing Surveys*, 36(2):81–121, 2004.

[26] R. Klein, B. Klein, S. Moss, and et al. Retinal vascular abnormalities in persons with type I diabetes. *Ophthalmology*, 11:2118–25, 2003.

[27] T. M. Koller, G. Gerig, G. Szekely, and D. Dettwiler. Multiscale detection of curvilinear structures in 2-D and 3-D image data. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 864–868, 1995.

[28] S. Konishi, A. L. Yuille, J. M. Coughlan, and S. C. Zhu. Statistical edge detection: Learning and evaluating edge cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(1):57–74, Jan. 2003.

[29] D. Koozekanani, K. Boyer, and C. Roberts. Retinal thickness measurements from optical coherence tomography using a markov boundary model. *IEEE Transactions on Medical Imaging*, 20:900–916, Sept. 2001.

[30] 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 Transactions on Medical Imaging*, 22(12):1519–1536, 2003.

[31] T. Lindeberg. Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30:117–156, Nov. 1998.

[32] T. Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30:79–116, Nov. 1998.

[33] 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 *Proceedings of the First International Conference on Scale-Space Theory in Computer Vision*, pages 152–163, 1997.

[34] 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 Transactions on Medical Imaging*, 23(10):1196–1204, Oct. 2004.

[35] V. Mahadevan, H. Narasimha-Iyer, B. Roysam, and H. Tanenbaum. Robust model-based vasculature detection in noisy biomedical images. *IEEE Transactions on Information Technology in Biomedicine*, 8(3):360–376, 2004.

[36] P. Majer. On the influence of scale selection on feature detection for the case of linelike structures. *International Journal of Computer Vision*, 60(3):191–202, 2004.

[37] 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 *Proceedings*

of the 2nd International Conference of Medical Image Computing and Computer-Assisted Intervention (MICCAI 1999), pages 90–97, 1999.

[38] G. K. Matsopoulos, P. A. Asvestas, N. A. Mouravliansky, , and K. K. Delibasis. Multimodal registration of retinal images using self organizing maps. *IEEE Transactions on Medical Imaging*, 23(12):1557–1563, 2004.

[39] P. Meer and B. Georgescu. Edge detection with embedded confidence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(12):1351–1365, Dec. 2001.

[40] F. Miles and A. Nutall. Matched filter estimation of serial blood vessel diameters from video images. *IEEE Transactions on Medical Imaging*, 12(2):147–152, 1993.

[41] 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 J. M. Fitzpatrick and M. Sonka, editors, *Proceedings SPIE Conference on Medical Imaging*, pages 648–656, 2004.

[42] J. C. Nunes, Y. Bouaoune, E. Delechelle, and P. Bunel. A multiscale elastic registration scheme for retinal angiograms. *Computer Vision and Image Understanding*, 95(2):129–149, 2004.

[43] A. Pinz, S. Bernogger, P. Datlinger, and A. Kruger. Mapping the human retina. *IEEE Transactions on Medical Imaging*, 17(4):606–620, Aug 1998.

[44] F. K. H. Quek and C. Kirbas. Vessel extraction in medical images by wave-propagation and traceback. *IEEE Transactions on Medical Imaging*, 20(2):117–131, Feb. 2001.

[45] 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 *CVRMed-MRCAS'97, First Joint Conference Computer Vision, Virtual Reality and Robotics in Medicine and Medial Robotics and Computer-Assisted Surgery, Proceedings*, pages 213–222, Grenoble, France, 19–22June 1997.

[46] 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 *Proceedings of the 7th International Conference of Medical Image Computing and Computer-Assisted Intervention (MICCAI 2004)*, pages 508–515, Saint-Malo, France, 2004.

[47] 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. *aje*, 150(3):263–270, 1999.

[48] G. Shechter, F. Devernay, E. Coste-Manière, A. Quyyumi, and E. R. McVeigh. Threedimensional motion tracking of coronary arteries in biplane cineangiograms. *IEEE Transactions on Medical Imaging*, 22:493–503, 2003.

[49] 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 Transactions on Pattern Analysis and Machine Intelligence*, 25(3):379–384, March 2003.

[50] J. A. Shoemaker. Vision problems in the U.S. Technical report, U.S. National Institute of Health, 2002.

[51] E. Simoncelli and H. Farid. Steerable wedge filters for local orientation analysis. *IEEE Transactions on Image Processing*, 5(9):1377–1383, 1996.

[52] 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 Transactions on Biomedical Engineering*, 49(9):1059–1067, 2002.

[53] J. Staal, M. Abramoff, M. Niemeijer, M. Viergever, and B. van Ginneken. Ridge based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, Apr 2004.

[54] J. Staal, S. Kalitzin, and M. A. Viergever. A trained spin-glass model for grouping of image primitives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(7):1172–1182, 2005.

[55] 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. *Journal of Hypertension*, 13(12):1724–1728, 1995.

[56] C. Stewart, C.-L. Tsai, and B. Roysam. The dual-bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE Transactions on Medical Imaging*, 22(11):1379–1394, 2003.

[57] 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 Medical & Biological Engineering & Computing*, 40(1):2–13, Jan 2002.

[58] The Eye Diseases Prevalence Research Group. Prevalence of age-related macular degeneration in the united states. *Archives of Ophthalmology*, 122(4):564–572, 2004.

[59] The Eye Diseases Prevalence Research Group. The prevalence of diabetic retinopathy among adults in the united states. *Archives of Ophthalmology*, 122(4):552–563, 2004.

[60] The Eye Diseases Prevalence Research Group. Prevalence of open-angle glaucoma among adults in the united states. *Archives of Ophthalmology*, 122(4):532–538, 2004.

[61] 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(2):263–273, Apr. 1998.

[62] E. Trucco and P. Kamat. Locating the optic disk in retinal images via plausible detection and constraint satisfaction. In *Proceedings IEEE International Conference on Image Processing*, volume 1, pages 155–158, Singapore, Oct.24–27 2004.

[63] C.-L. Tsai, C. Stewart, B. Roysam, and H. Tanenbaum. Repeatable vascular landmark extraction from retinal fundus images using local vascular traces. *IEEE Transactions on Information Technology in Biomedicine*, to appear 2003.

[64] 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 *Proceedings of the 7th International Conference of Medical Image Computing and Computer-Assisted Intervention (MICCAI 2004)*, pages 525–532, Saint-Malo, France, 2004.

[65] 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 Transactions on Medical Imaging*, 21(10), Oct. 2002.

[66] T. Walter, J.-C. Klein, P. Massin, and F. Zana. Automatic segmentation and registration of retinal fluorescein angiographies - application to diabetic retinopathy. In *Workshop on Computer Assisted Fundus Image Analysis*, May 2000.

[67] 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. *Journal of American Medical Association*, 287:2528–33, May 2002.

[68] 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(5):419–428, 1999.

[69] F. Zana and J.-C. Klein. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *IEEE Transactions on Image Processing*, 10(7):1010–1019, 2001.
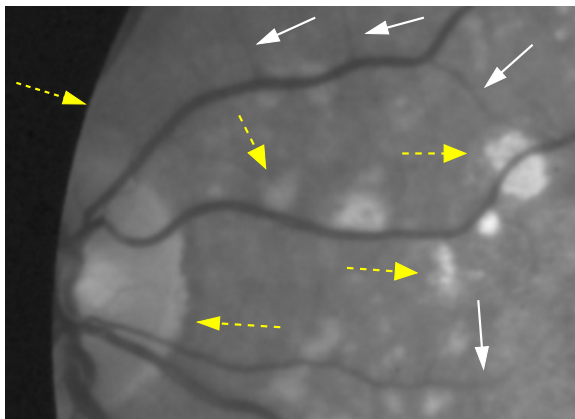
Figure 1: Illustration of the challenges of retinal vessel extraction. Arrows drawn on the image [in yellow / with dashed lines] indicate pathologies, the boundary of the optic disc and the boundary of the retina, all of which tend to cause false positive vessel detection. Arrows drawn [in white / with solid lines] highlight narrow or low-contrast vessels which are difficult to detect.
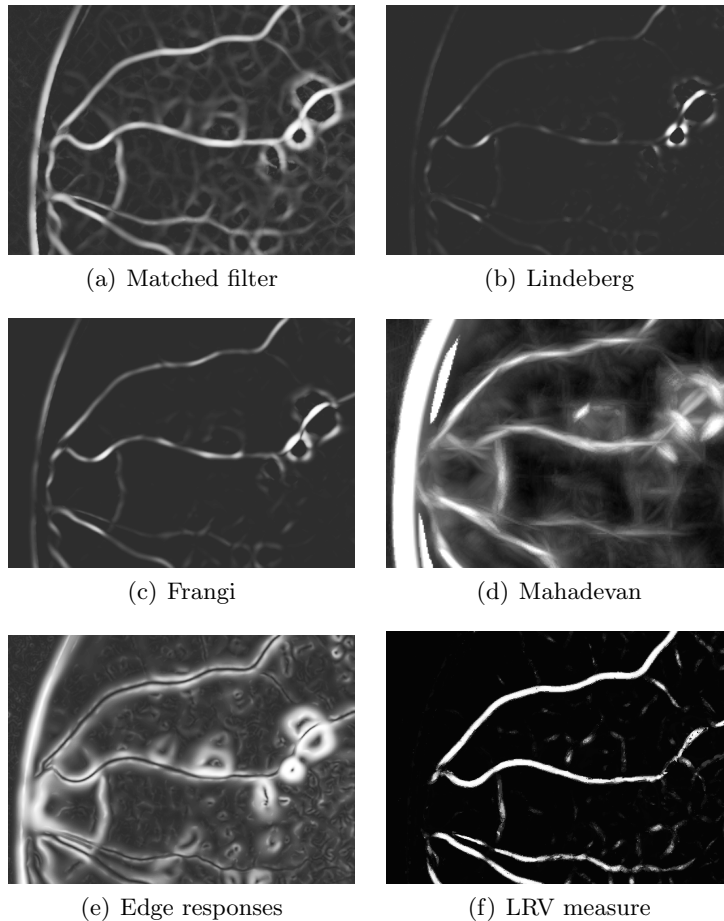
(a) Matched filter

(b) Lindeberg

(c) Frangi

(d) Mahadevan

(e) Edge responses

(f) LRV measure

Figure 2: Illustration of the challenges of retinal vessel extraction on a source image from Figure 1. **(a)** Matched filter response based on a Gaussian vessel profile; **(b)** Lindeberg's ridge measure [31]; **(c)** Frangi's Hessian-based vesselness measure [11]; and **(d)** Mahadevan's robust model based measure [35]. 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 **d)** suppresses the response to the bright pathology on the right, highlights thin and low contrast vessels (e.g. 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.)
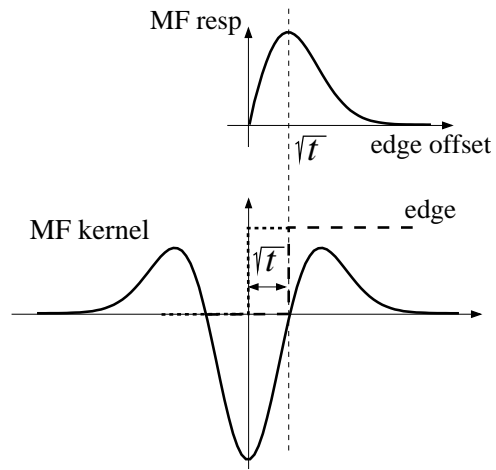
Figure 3: Vessel matched filter response to an offset edge. It is highest when the offset is equal to the standard deviation $\sqrt{t}$ of the matched filter.
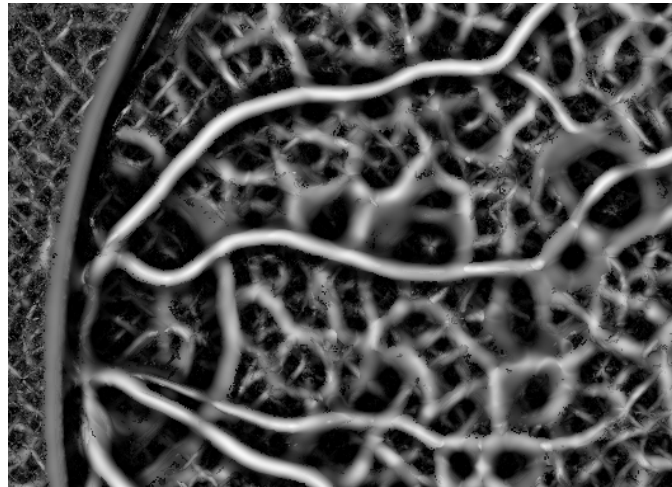
Figure 4: Vessel confidences for the image in Figure 1. They are highest at vessel locations but can still be high around locations offset from edges and other non-vessel structures. Confidence images were stretched to $\langle 0, 255 \rangle$ for display.
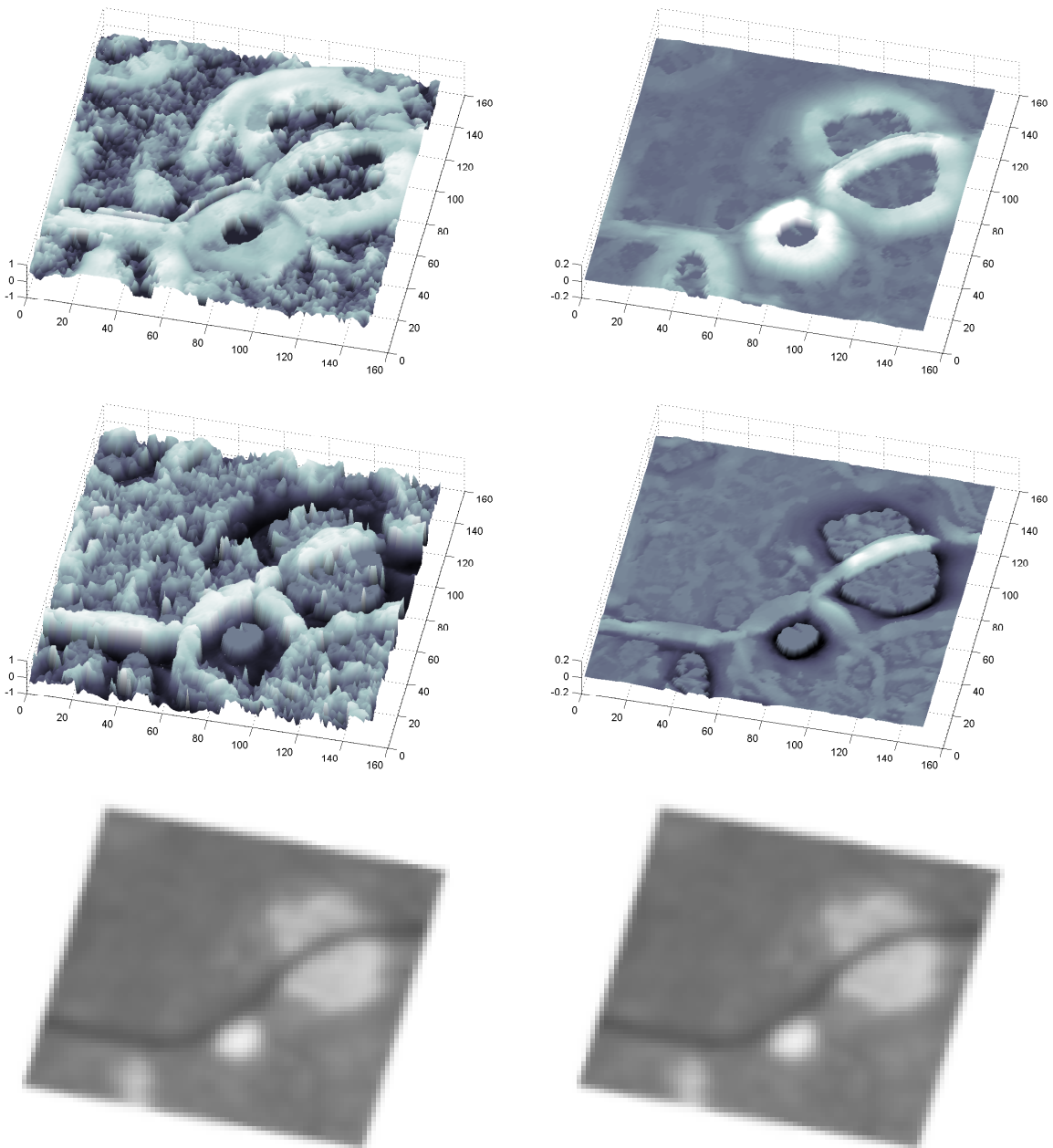
Figure 5: Confidence (left column) and response (right column) images of stronger (top) and weaker (middle) vessel edge for the source image shown at the bottom, displayed as 3D plots. Confidences and responses of both edges are high (and positive) at vessel locations. Around pathologies and other non-vascular 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).
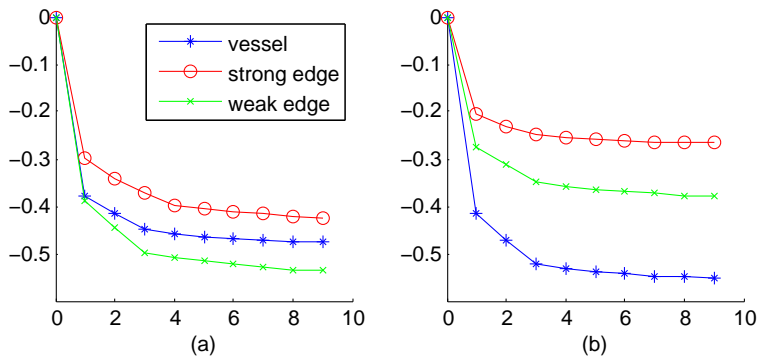
Figure 6: **(a)** Chernoff bound as a function of number of bins for vessel, strong edge, and weak edge responses, **(b)** Chernoff bound for vessel, strong edge, and weak edge confidences. The discriminative power of the weak edge is higher than that of the strong edge because the existence of the weak edge is the most significant difference between a vessel and an (offset) step edge.
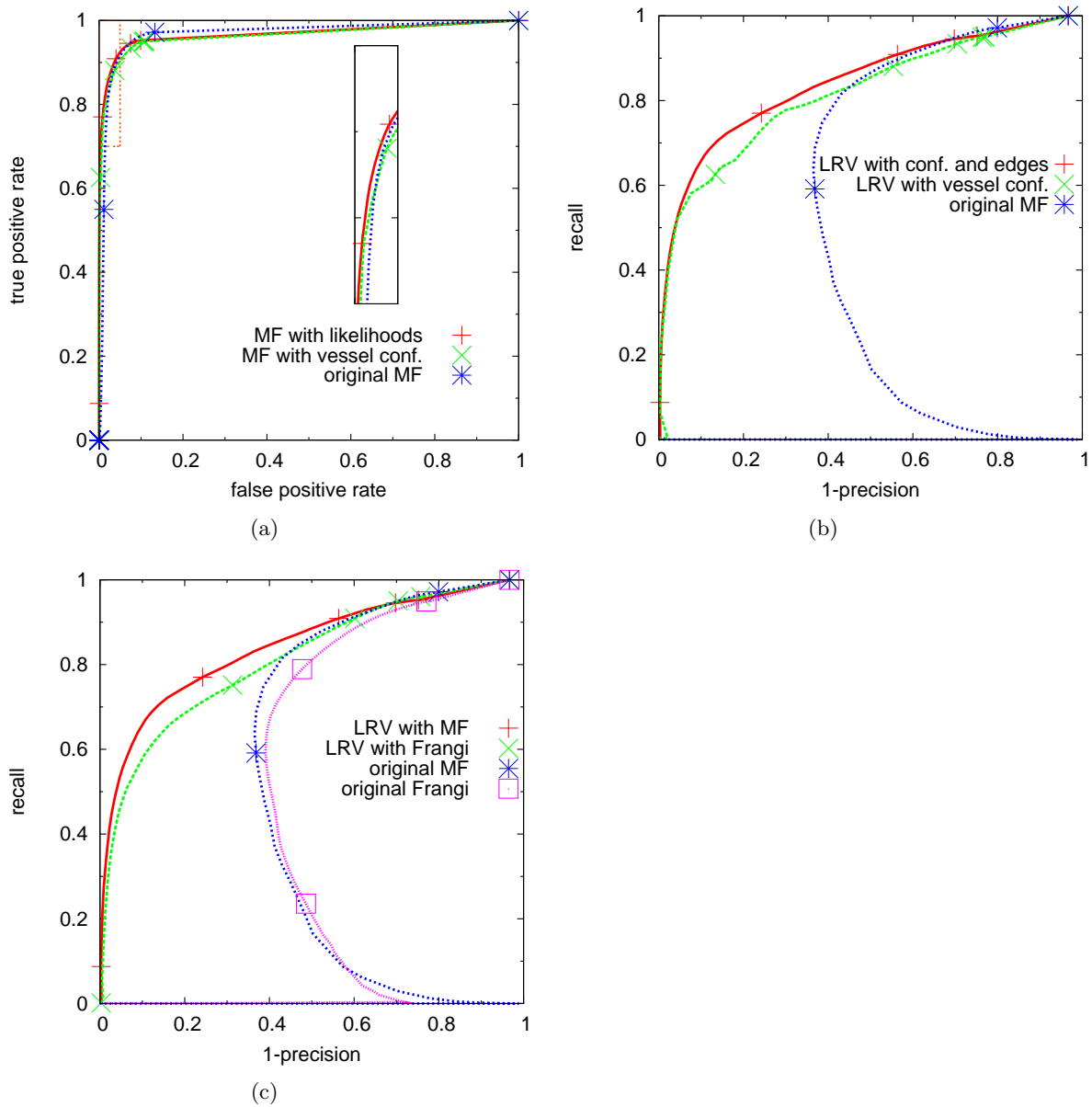
Figure 7: ROC curves **(a)** and (1-Precision)-Recall curves **(b)** showing improvement by vessel confidences and by including edges. (1-Precision)-Recall curves are more suitable for comparison because of large number of negatives (non vessel pixels) in the ground truth. Notice in (b) that matched filter with vessel confidences is as powerful as likelihoods until about 50% of all traces are detected. The matched filter outperforms the Hessian-based measures and the likelihoods improve both **(c)**.
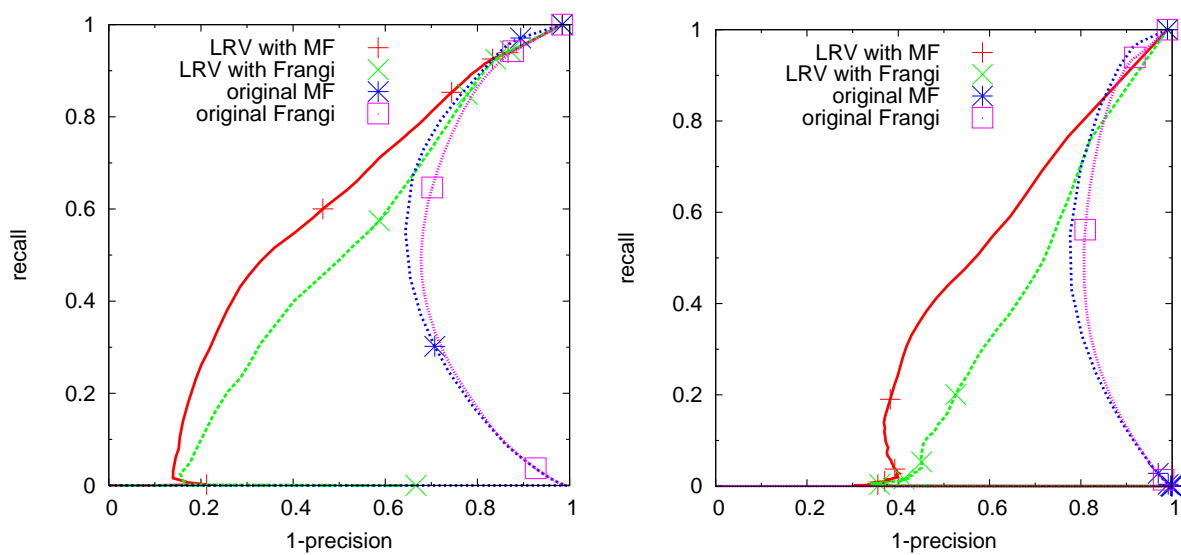
Figure 8: Using ground truth segmentation, we generated the (1-Precision)-Recall plots by focusing only on vessels 1-2 pixels wide **(left)**. The performance is worse than when evaluating all vessels, but still impressive. Similarly, we used ground truth segmentation to generate the plots eliminating vessels that have contrast more than 3 noise standard deviations **(right)**.
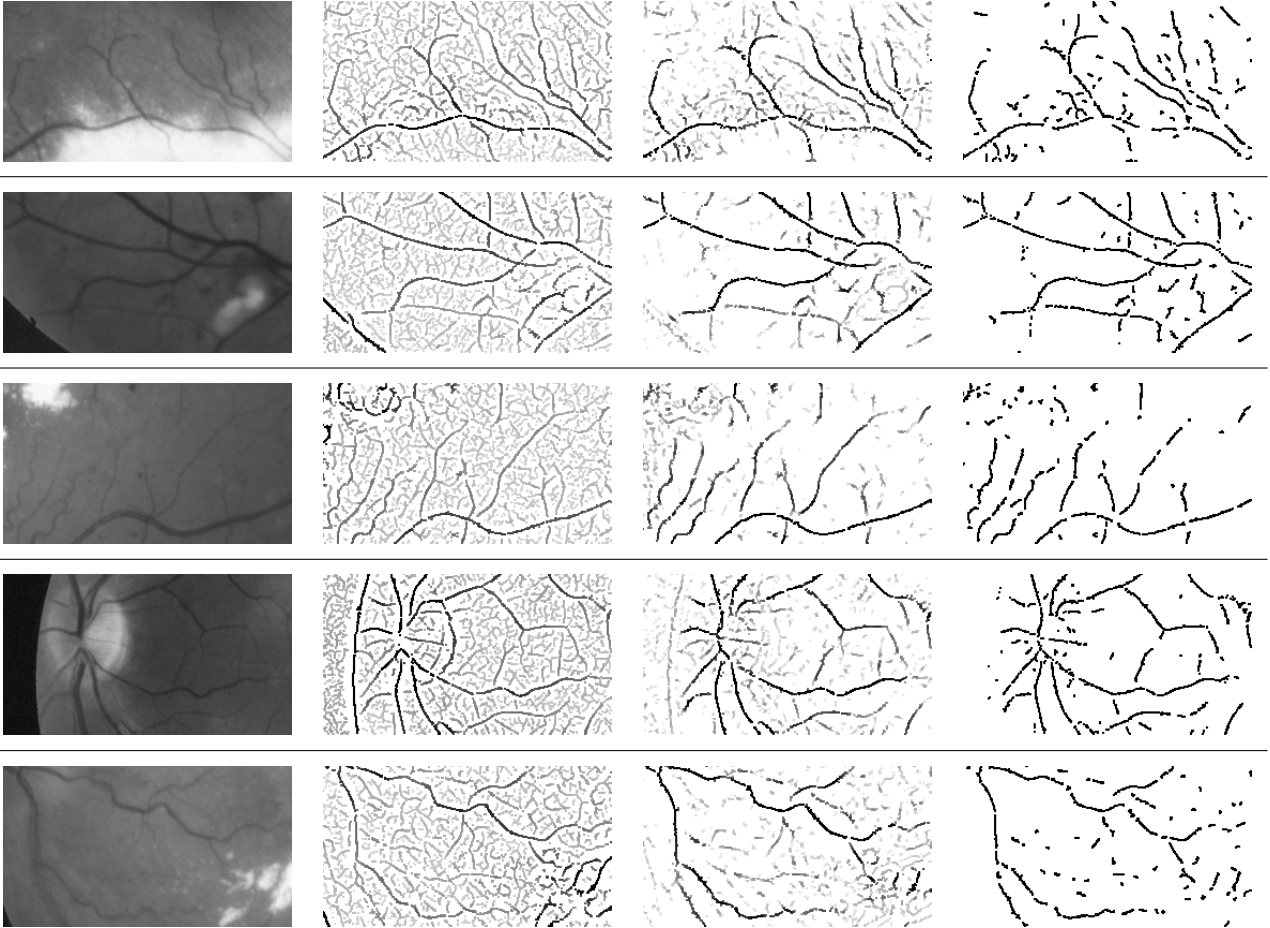
Figure 9: Results on 5 difficult image chips showing a variety of challenges, including pathologies, other non-vascular structures, and low-contrast, narrow vessels. The figure shows the source images in the **1st column**, the vessel matched filter response images after non-maximum suppression in the **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.
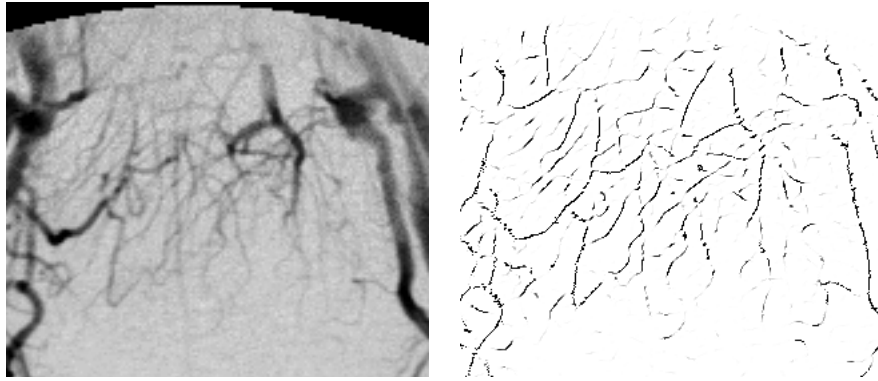
Figure 10: Angiogram of the cerebral vasculature **(left)**, and Likelihood Ratio Vesselness **(right)**. Original retina vessel training data was used.
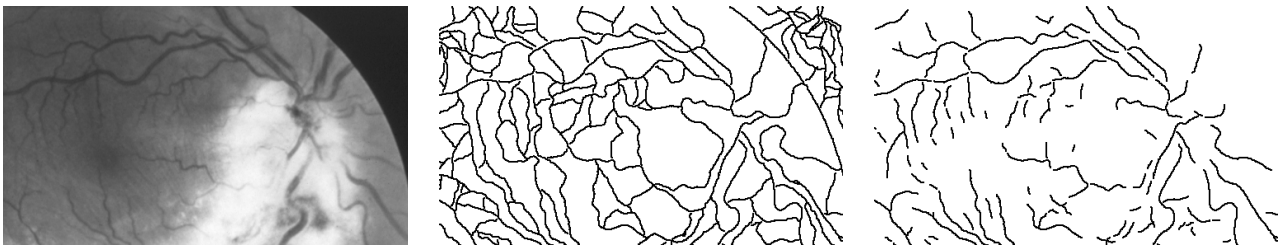

Figure 11: Vessel traces of a source image **(left)** before **(middle)** and after applying vessel likelihoods **(right)**. Non vessel structures are successfully removed while traces of even thinnest and barely visible vessels are kept.
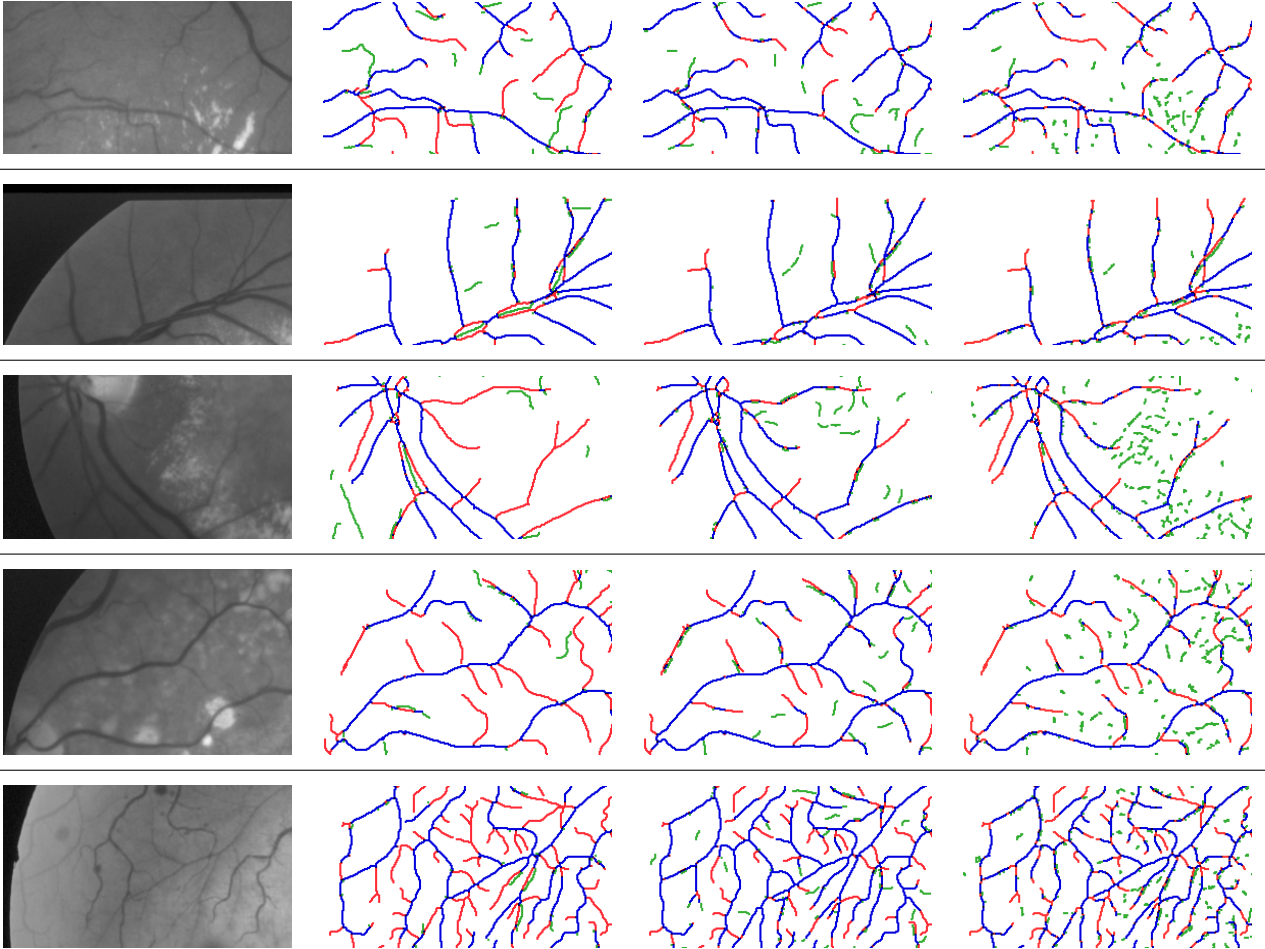
Figure 12: Source image **(1st column)**, tracing results using parallel edge algorithm [5] **(2nd column)**, Likelihood Ratio Vesselness (LRV) based tracing **(3rd column)** and the LRV after non-max suppression and thresholding below 2.0 **(4th column)**. True positives are in blue, false negatives in red, and false positives in green. The new tracing algorithm successfully handles difficult vessel branching and crossover points (two overlapping vessels in the third and fourth example), more accurately determines when to stop tracing in unclear regions (thin vessels near pathologies in the first and third example), correctly ignores pathologies (first and fourth example) and finds even the thinnest vessels that usually have very low contrast (fifth example). Notice in a number of cases false positives are marked along false negatives (green line along red line) because of localization errors in ground truth. The tracing results are comparable to the LRV measure evaluated at every point. Traces with less than 4 connected components in the last column were removed.
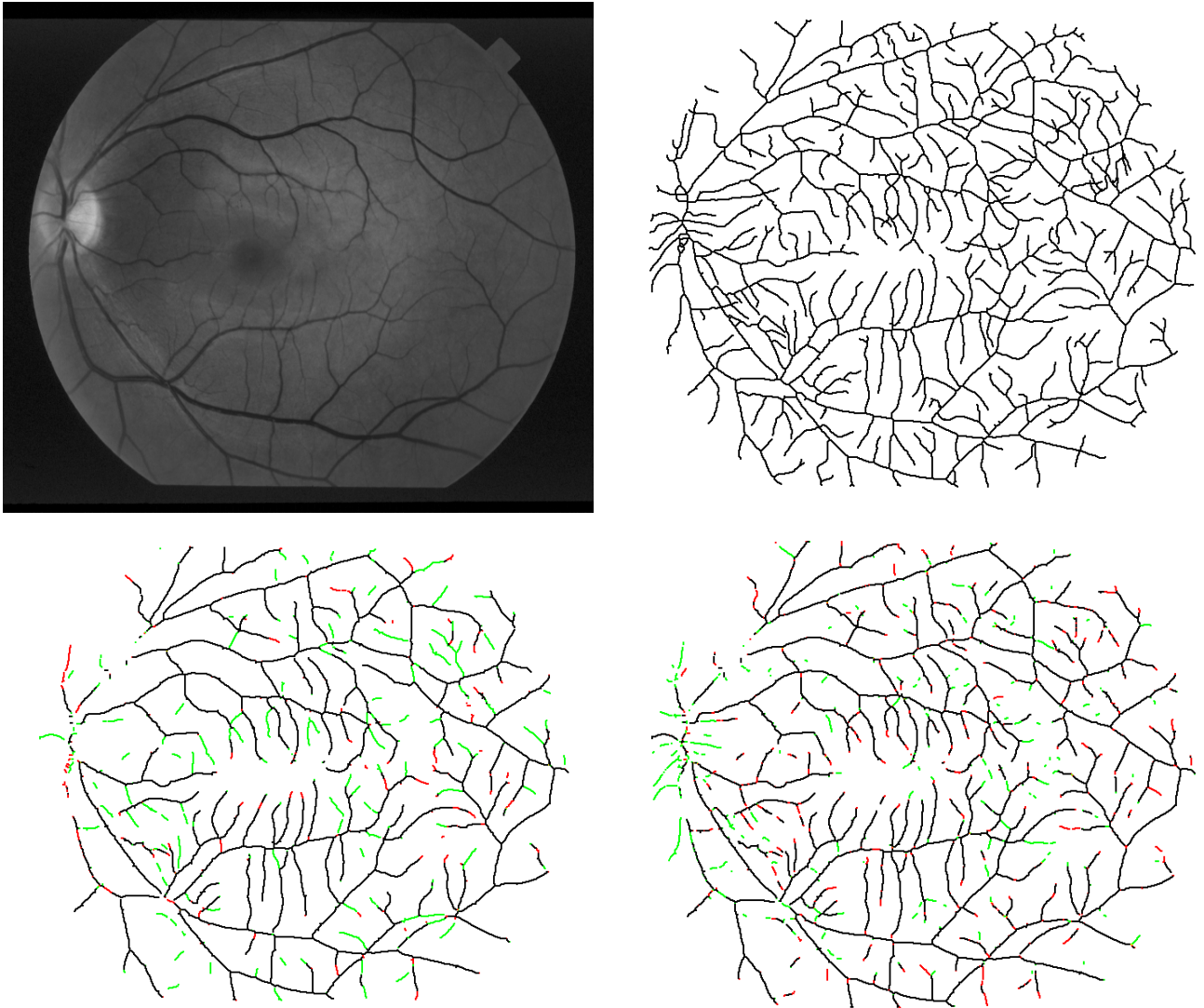
Figure 13: Source image **(top left)** and ground truth thinned labelling **(top right)** for comparison of results for parallel edge and LRV-based tracing **(bottom left)**. Vessel pixels detected by both methods (black), LRV-based tracing only (green) and parallel edge based only (red) are shown. LRV-based tracing outperforms parallel edge tracing in detecting low contrast and thin vessels and produces more complete centerline extraction. Image in **(bottom right)** shows comparison of LRV-based tracing and LRV evaluated at each pixel followed by non-maximum suppression and thresholding below 2.0. Vessel pixels detected by both (black), only by computing LRV at each pixel (green) and only by LRV-based tracing (red) are shown. Occasionally, tracing does not detect some vessels (eg. area around the optic disc in this example) due to the lack of an initial point. LRV computed at each pixel correctly highlights even these vessels.
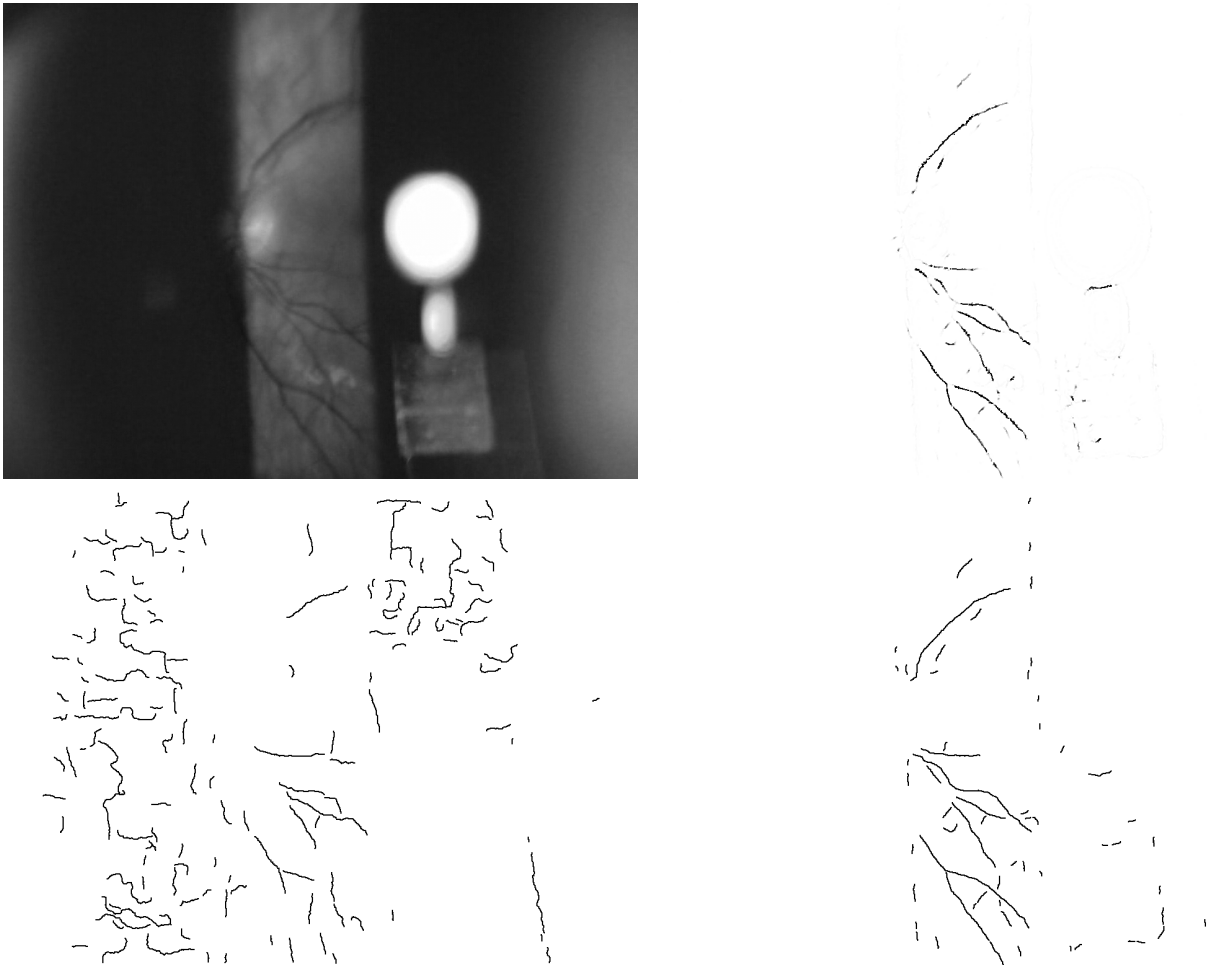
Figure 14: Example of a slit lamp image **(top left)**, vessel likelihoods **(top right)**, parallel edge based tracing **(bottom left)**, and likelihood based tracing **(bottom right)**. The likelihoods are powerful to highlight vessels but ignore the background noise which is the main reason for spurious traces of the parallel edge based algorithm. The new tracing algorithm correctly finds vessels that are missed by the parallel edge based tracing.