Exploratory Dijkstra forest based automatic vessel segmentation: applications in video indirect ophthalmoscopy (VIO)

Rolando Estrada,[1,*] Carlo Tomasi,[1] Michelle T. Cabrera,[2] David K. Wallace,[3] Sharon F. Freedman,[3] and Sina Farsiu[3,4]

[1]*Dept. of Computer Science, Duke University, Durham, NC 27708*
[2]*Dept. of Ophthalmology, University of North Carolina, Chapel Hill, NC 27599*
[3]*Dept. of Ophthalmology, Duke University, Durham, NC 27708*
[4]*Dept. of Biomedical Engineering, Duke University, Durham, NC 27708*

[*]*restrada@cs.duke.edu*

Abstract: We present a methodology for extracting the vascular network in the human retina using Dijkstra's shortest-path algorithm. Our method preserves vessel thickness, requires no manual intervention, and follows vessel branching naturally and efficiently. To test our method, we constructed a retinal video indirect ophthalmoscopy (VIO) image database from pediatric patients and compared the segmentations achieved by our method and state-of-the-art approaches to a human-drawn gold standard. Our experimental results show that our algorithm outperforms prior state-of-the-art methods, for both single VIO frames and automatically generated, large field-of-view enhanced mosaics. We have made the corresponding dataset and source code freely available online.

OCIS codes: (100.0100) Image processing; (100.2960) Image analysis; (170.4470) Ophthalmology.

References and links

1. W. Tasman, A. Patz, J. A. McNamara, R. S. Kaiser, M. T. Trese, and B. T. Smith, "Retinopathy of prematurity: The life of a lifetime disease," American Journal of Ophthalmology **141**, 167 – 174 (2006).
2. I. C. for the Classification of Retinopathy of Prematurity, "The international classification of retinopathy of prematurity revisited," Archives of Ophthalmology **123**, 991–999 (2011).
3. D. K. Wallace, G. E. Quinn, S. F. Freedman, and M. F. Chiang, "Agreement among pediatric ophthalmologists in diagnosing plus and pre-plus disease in retinopathy of prematurity," Journal of American Association for Pediatric Ophthalmology and Strabismus **12**, 352 – 356 (2008).
4. S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, "Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation," Opt. Express **18**, 19413–19428 (2010).
5. S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," IEEE Transactions on Medical Imaging **8**, 263–269 (1989).
6. C. Kirbas and F. Quek, "A review of vessel extraction techniques and algorithms," ACM Computing Surveys **36**, 81–121 (2004).
7. Q. Li, J. You, L. Zhang, and P. Bhattacharya, "Automated retinal vessel segmentation using Gabor filters and scale multiplication," SMC **6**, 3521–3527 (2006).
8. E. Ricci and R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," IEEE Transactions on Medical Imaging **26**, 1357–1365 (2007).
9. J. Soares, J. Leandro, R. Cesar Jr, H. Jelinek, and M. Cree, "Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification," IEEE Transactions on Medical Imaging **25**, 1214–1222 (2006).

10. J. Staal, M. Abràmoff, M. Niemeijer, M. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," IEEE Transactions on Medical Imaging **23**, 501–509 (2004).
11. B. Lam, Y. Gao, and A. Liew, "General retinal vessel segmentation using regularization-based multiconcavity modeling," IEEE Transactions on Medical Imaging **29**, 1369–1381 (2010).
12. G. Lathen, J. Jonasson, and M. Borga, "Blood vessel segmentation using multi-scale quadrature filtering," Pattern Recognition Letters **31**, 762–767 (2010).
13. D. Marín, A. Aquino, M. Gegúndez-Arias, and J. Bravo, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," IEEE Transactions on Medical Imaging **30**, 146–158 (2011).
14. A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," IEEE Transactions on Medical Imaging **19**, 203–210 (2002).
15. T. Chanwimaluang and G. Fan, "An efficient blood vessel detection algorithm for retinal images using local entropy thresholding," in *Proceedings of the International Symposium on Circuits and Systems*, (2003), 21–24.
16. M. Martínez-Pérez, A. Hughes, A. Stanton, S. Thom, A. Bharath, and K. Parker, "Retinal blood vessel segmentation by means of scale-space analysis and region growing," in *Proceedings of Medical Image Computing and Computer-Assisted Intervention*, (1999), 90–97.
17. F. Zana and J. Klein, "Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation," IEEE Transactions on Image Processing **10**, 1010–1019 (2002).
18. L. Pedersen, M. Grunkin, B. Ersboll, K. Madsen, M. Larsen, N. Christoffersen and U. Skands, "Quantitative measurement of changes in retinal vessel diameter in ocular fundus images," Pattern Recognition Letters **(21)**, 1215–1223 (Elsevier 2000).
19. M. Cree, D. Cornforth and HF. Jelinek, "Vessel segmentation and tracking using a two-dimensional model," in *Proceedings of Image and Vision Computing New Zealand*, (2005), 345–350.
20. F. Benmansour and L. Cohen, "Tubular structure segmentation based on minimal path method and anisotropic enhancement," International Journal of Computer Vision **92**, 192–210 (2011).
21. H. Li and A. Yezzi, "Vessels as 4-D curves: Global minimal 4-D paths to extract 3-D tubular surfaces and centerlines," IEEE Transactions on Medical Imaging **26**, 1213–1223 (2007).
22. M. Pechaud, R. Keriven, and G. Peyre, "Extraction of tubular structures over an orientation domain," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (IEEE, 2009), 336–342.
23. O. Wink, W. Niessen, and M. Viergever, "Multiscale vessel tracking," IEEE Transactions on Medical Imaging **23**, 130–133 (2004).
24. S. Ahmad, D. Wallace, S. Freedman, and Z. Zhao, "Computer-assisted assessment of plus disease in retinopathy of prematurity using video indirect ophthalmoscopy images," Retina **28**, 1458–1462 (2008).
25. A. Kiely, D. Wallace, S. Freedman, and Z. Zhao, "Computer-Assisted Measurement of Retinal Vascular Width and Tortuosity in Retinopathy of Prematurity," Archives of Ophthalmology **128**, 847–852 (2010).
26. T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (IEEE Computer Society, 1996), 465–470.
27. J. Sethian, *Level set methods and fast marching methods* (Cambridge University Press, 1999).
28. E. Dijkstra, "A note on two problems in connexion with graphs," Numerische Mathematik **1**, 269–271 (1959).
29. R. Estrada, C. Tomasi, M. Cabrera, D. Wallace, S. Freedman, and S. Farsiu, "Enhanced video indirect ophthalmoscopy (VIO) via robust mosaicing," Biomedical Optics Express **2**, 2871–2887 (2011).
30. R. Bellman, *Dynamic programming* (Dover, 2003).
31. T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to algorithms* (MIT Press, 2001).
32. 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 *Proceedings of SPIE*, (2004), 648–656.
33. B. Al-Diri, A. Hunter, D. Steel, M. Habib, T. Hudaib, and S. Berry, "REVIEW-A reference data set for retinal vessel profiles," in *Proceedings of the IEEE Conference on Engineering in Medicine and Biology Society*, (IEEE, 2008), 2262–2265.
34. J. Cohen, "A Coefficient of Agreement for Nominal Scales," Educational and Psychological Measurement **20**, 37–46 (Durham, 1960).
35. J. Gibbons and S. Chakraborti, *Nonparametric statistical inference* (CRC Press, 2003).

1. Introduction

Accurate segmentation and evaluation of the anatomical and pathological features of retinal vessels are critical for the diagnosis and study of many ocular diseases. These include retinopathy of prematurity (ROP). ROP is a disorder of the retinal blood vessels that is a major cause of vision loss in premature neonates [1]. Important features of the disease include increased diameter (dilation) as well as increased tortuosity (wiggliness) of the retinal blood vessels in the portion of the retina centered on the optic nerve (the posterior pole). Increased dilation and

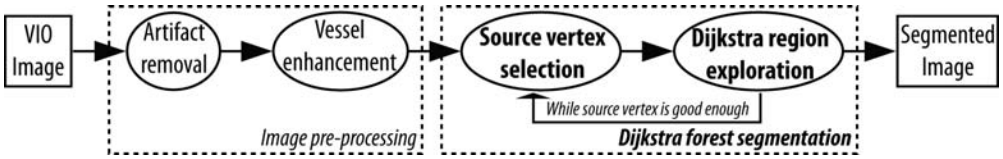

Fig. 1. Proposed VIO vessel segmentation: In the first stage, VIO images are pre-processed with directional local-contrast filters (DLCF) and LoG-Gabor filters to eliminate artifacts and increase contrast. In the second stage, the best, unvisited vessel pixel in the image is repeatedly chosen as a starting point for a dynamic-programming exploration of the unvisited part of the image. The result of each exploration yields a new tree in the growing forest of vessels. Forest growth stops when the best, unvisited vessel pixel is worse than a predefined threshold.

tortuosity of the blood vessels in the posterior pole (called pre-plus in intermediate, and plus in severe circumstances) is an important indicator of ROP severity. [2]. Subjective assessment of plus and pre-plus disease leads to poor agreement between examiners [3]. Manual segmentation of retinal images is not only demanding for experts and excessively time-consuming for clinical use, but is also inherently subjective, and different annotators often yield different results [4]. To address these difficulties, different approaches for automated segmentation of retinal vessels have been tried, with varying levels of success.

Prior methods can be roughly classified into region- and path-based methods. Region-based methods [5–13] classify image pixels directly into vessel and non-vessel pixels. Classification relies on local appearance, as measured by the responses of suitable filter banks at various scales and orientations. In unsupervised region-based approaches, these filter responses are combined into a new image, which is then appropriately thresholded to yield the final classification. Methods in this category employ matched filters [5], piecewise thresholding [14], local entropy [15], and quadrature filters [12]. Supervised region-based methods, on the other hand, assemble the filter responses into feature vectors that are fed to a classifier, which is trained on hand-labeled data. Techniques used within this framework include ridge detection [10], Gabor wavelet filtering [9], line operators [8], and moment invariants [13]. Other region-based approaches have used region growing [16], mathematical morphology [17], and multiconcavity modeling [11].

The goal of path-based methods [18–25], on the other hand, is primarily to trace the centerline of individual vessels, rather than classifying every pixel in the image. Many path-based approaches also estimate vessel thickness as they track each branch, generally by determining the width of the cross-section perpendicular to the current path. Prior work on two-dimensional branch extraction has addressed this topological ambiguity semi-automatically by relying on user-supplied points, requiring either a single seed point [24] or a pair of start- and end-points [22]. User-supplied one-point methods generally employ ridge detection based on differential geometry [26], while two-point methods find a path between the points that minimizes a cost measure designed to penalize paths that stray from the middle of a vessel. Several of these methods rely on front propagation algorithms, such as the fast marching method [27]. In contrast, as described in Section 2, our tracking methodology forgoes the need for external seed points by being robust to a particular tracker's initial position.

Existing methods in both categories have been developed primarily for use on high quality retinal fundus images, such as those obtained with the RetCam imaging system (Clarity Medical Systems, Inc., Pleasanton, CA). However, the usual method for diagnosing ROP is the indirect ophthalmoscope (IO). More recently, Video Indirect Ophthalmoscopy (VIO), in which the physician wears a head-mounted video camera during IO evaluations, has emerged as an economical and convenient method for capturing digital retinal images during ROP examinations. In contrast to RetCam, however, VIO data is often of low quality, fraught with reflections from the IO lens, motion blur, low resolution, and sensor noise. A previous study reported that only 24% of randomly selected video sequences can be utilized for semi-automated evaluation of retinal vessel morphology in ROP [24].

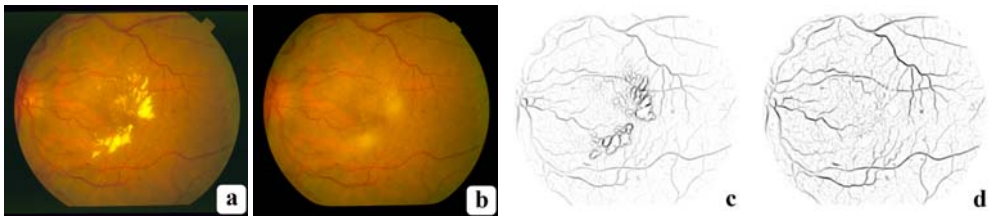

Fig. 2. DLCF exudate removal: **(a)** An image from the STARE dataset [14]. **(b)** The image after DLCF. **(c)** Matched filtering [5] applied to (a). **(d)** Matched filtering applied to (b). The non-vascular filter responses around the exudates have been eliminated in (d) without affecting the true vessel responses.

In this paper, we propose a hybrid method that extends the path-based methodology into a region-based segmentation scheme for detecting retinal vessels. Our complete approach works in two stages, as illustrated in Fig. 1. The first stage pre-processes the input image to remove both lens and motion artifacts, and to construct a high-contrast vessel map. The second stage builds a forest of tree-like vessel regions through a sequence of exploration waves on the vessel map: the most vessel-like pixel $\mathbf{s}_0$ in the image is used as the starting point for an exploration wave that searches for the best tree-like vessel region in the image around $\mathbf{s}_0$ by means of the single-source, multi-destination version of Dijkstra's shortest path algorithm [28]. This exploration returns an entire *tree region* for part of the vessel system, that is, it handles branching naturally and efficiently, and preserves vessel thickness. When this exploration ends, a new exploration begins at the best remaining starting point $\mathbf{s}_1$ in the unexplored part of the image, which yields a new vessel tree region. Our method stops constructing new regions when the best unexplored starting point is no longer likely to be part of the vessel system. Unlike existing single-source, single-destination vessel analysis methods [20–23], our single-source, multiple-destinations approach automatically explores the complete vasculature in a retinal image, and requires no user intervention whatsoever.

Furthermore, the initial single-frame image enhancement step can be optionally replaced by a multi-frame image mosaicing technique. We have recently developed such a technique to combine several low-quality VIO frames into a high-quality, large field-of-view (FOV) composite [29]. As our results in Section 3 show, our approach obtains superior segmentation results on both types of –raw and composite– VIO images compared to current state-of-the-art segmentation methods.

The rest of this paper is organized as follows: we first detail our automated dynamic-programming segmentation method in Section 2 and then describe our experiments in Section 3. We present the experimental results in Section 4 and discuss their significance and explore future directions in Section 5.

2. Exploratory Dijkstra forest based vessel segmentation method

We represent each VIO image (or composite) as a graph of nodes, $G = (V,E)$, where each node corresponds to a pixel and the links connecting the nodes are called arcs. In this formulation, the ordered pair of node and arc sets are represented by V and E, respectively. Path-based methods for vessel extraction define the *cost* of traversing the arc that connects any two neighboring pixels in the image in such a way that arcs between vessel pixels are more likely to have lower cost. Vessel extraction then looks for paths that traverse the image from neighbor to neighbor and have minimum aggregate cost, and are thereby likely to follow vessels. If the cost aggregation rule is associative, minimum-cost paths can be found efficiently [30].

We depart from previous work within this framework in two major ways. First, we find vessel *regions*, rather than simply vessel paths. In other words, we preserve vessel thickness, rather than merely finding the skeleton, or centerline, of each vessel. This is important, because eye disease diagnosis often requires consideration of vessel thickness. Second, we employ a se-

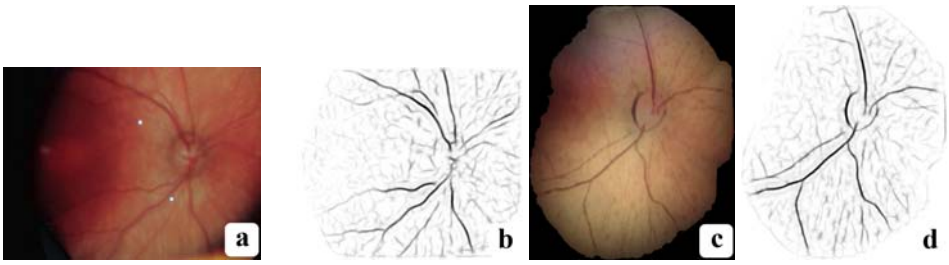

Fig 3. LoG-Gabor filtering: **(a)** A sample VIO frame. **(b)** Frame after LoG-Gabor filtering. **(c)** A sample mosaic. **(d)** Mosaic after LoG-Gabor filtering. The isotropic LoG filtering enhances vessel contrast, while the anisotropic Gabor wavelets selectively enhance elongated structures.

quence of searches for vessel regions that start at *source points* $\mathbf{s}_0, \mathbf{s}_1, \ldots$ automatically selected in decreasing order of their likelihood to be part of a vessel, as detailed in Subsection 2.4. This novelty eliminates the need for a user to select vessel starting points by hand.

Thus, we use the single-source, multiple-destination version of Dijkstra's shortest path algorithm [28], rather than the single-source, single-destination version used in prior work. In other words, rather than *connecting* a start point with a destination point, our method *explores* the image outward from an (automatically selected) source point. This exploratory strategy has two advantages: it eliminates the need for selecting a destination point manually, and it finds vessels as tree-like image regions, thereby accounting for vessel branching naturally and efficiently.

The computational cost of this important change of perspective is trivial, as the only difference between the single-destination and multi-destination algorithms is when they stop: the single-destination algorithm stops when it reaches the designated vertex, while the multi-destination algorithm stops when a target threshold on the path cost has been reached. Both versions of Dijkstra's algorithm have the same computational complexity of $O(|E|+|V|\log|V|)$, where $|\cdot|$ indicates the cardinality or size of a set. This complexity is achievable with a heap-based priority queue implementation [31].

2.1. Arcs and Arc Costs

We view each VIO color image as an $X \times Y \times 3$ matrix I. Prior to processing, we first remove the image's artifacts by using directional local contrast filtering (DLCF) as defined in [29]. Figure 2 illustrates the effect of this image enhancement step. A pixel position in I is given by a two-dimensional vector of integers, $\mathbf{p} = [x,y]^{\mathrm{T}}$. The value at each pixel position is given by a three-dimensional vector $I(\mathbf{p})$ of red, green, blue values normalized between 0 and 1.

We define two features that determine the arc costs at each pixel $\mathbf{p}$: the green channel intensity $I_g(\mathbf{p})$ and the inverted response $F(\mathbf{p})$ to a Laplacian-of-Gaussian filter followed by a Gabor filter bank, or Laplace-Gabor filtering, as detailed in [29]. The vessel map F maximizes the discriminability of vessels, as illustrated in Fig. 3.

To apply Dijkstra's algorithm to I, we define a weighted lattice graph on the set $V = \{\mathbf{p}\}$ of all pixel locations in the image. There is an arc $e = (\mathbf{v}, \mathbf{v}')$ in the arc set E for this directed graph $G = (V, E)$ for any ordered pair of 8-neighbors, that is, whenever

$$\max(|x - x'|, |y - y'|) = 1\,. \tag{1}$$

A non-negative cost is defined on each arc, with the intent that arcs inside and along vessels cost less than arcs that have one or both endpoints outside any vessel. Specifically, we define the cost of arc e as the following convex linear combination:

$$c(e) = \sum_{m=1}^{4} w_m e^{\alpha z_m(e)} \qquad \text{where} \qquad \sum_{m=1}^{4} w_m = 1 \tag{2}$$

and $z_m(e)$ indicates the m-th element in the following four-dimensional feature vector:

$$\mathbf{z}(e) = \mathbf{z}(\mathbf{v},\mathbf{v}') = \Big[I_g(\mathbf{v}'),\, |I_g(\mathbf{v}) - I_g(\mathbf{v}')|,\, F(\mathbf{v}'),\, |F(\mathbf{v}) - F(\mathbf{v}')| \Big] \ . \tag{3}$$

Therefore, a low-cost arc is an arc whose destination point $\mathbf{v}'$ is dark ($I_g(\mathbf{v}') << 1$) and has a low inverted Laplace-Gabor response ($F(\mathbf{v}') << 1$), and such that the two arc endpoints are similar in both brightness ($|I_g(\mathbf{v}) - I_g(\mathbf{v}')| << 1$) and Laplace-Gabor response ($|F(\mathbf{v}) - F(\mathbf{v}')|$).

The exponential in Eq. 2 provides a non-linear scaling of the arc's features that emphasizes the divide between vascular and non-vascular feature values, and the scalar α controls the growth rate of the exponential term. In our experiments, we set the values of both α and the coefficients w_m based on training images from our dataset, as explained in Section 3.

Input: Graph G, source vertex $\mathbf{s}$, threshold τ.
Output: Dijkstra region R.

```
Q = initialize_priority_queue();
push(Q, s, 0);
while not_empty(Q) do
    [v_c, γ̃_{0,c}] = pop(Q);
    if not_visited(v_c) then
        set_visited(G, v_c);
        V_c = neighbors(G, v_c);
        foreach v ∈ V_c do
            h = c(γ̃(s, v_c)) + c(v_c, v);
            if h < τ then
                push(Q, v, h);
            end
        end
    end
end
R = visited(G);
```

Algorithm 1. Exploratory Dijkstra vessel segmentation: starting from a single pixel, the algorithm progressively explores the rest of the image such that every unvisited pixel has a higher minimum path cost than every visited pixel. The algorithm keeps adding pixels until a cost boundary is reached.

2.2. Path Costs

A path γ between any two nodes $\mathbf{v}$, $\mathbf{v}'$ in V is composed of a sequence of neighboring lattice locations:

$$\gamma(\mathbf{v},\mathbf{v}') = (\mathbf{v} = \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k = \mathbf{v}'), \tag{4}$$

subject to the constraint that $(\mathbf{v}_i, \mathbf{v}_{i+1}) \in E$, for $i \in [1, k-1]$. In short, γ is a curve discretized as a sequence of neighboring pixels. The cost of γ is defined as the sum of the costs of its arcs:

$$c(\gamma) = \sum_{i=1}^{k-1} c(\mathbf{v}_i, \mathbf{v}_{i+1}) \ . \tag{5}$$

The associative nature of this definition allows splitting a path's total cost into disjoint sub-path costs at any point along γ:

$$c(\gamma(\mathbf{v},\mathbf{v}')) = c(\gamma(\mathbf{v},\mathbf{v}_i)) + c(\gamma(\mathbf{v}_i,\mathbf{v}')) \qquad \text{for any} \qquad i \in [2, k-1] \,, \tag{6}$$

with which we can efficiently determine the minimum cost path between any two nodes $\mathbf{v}$ and $\mathbf{v}'$. That is, we use Dijkstra's algorithm to compute:

$$\tilde{\gamma}(\mathbf{v},\mathbf{v}') = \underset{\gamma \in \Gamma(\mathbf{v},\mathbf{v}')}{\operatorname{argmin}}\, c(\gamma), \tag{7}$$

where $\Gamma(\mathbf{v},\mathbf{v}')$ is the set of all possible paths between the two nodes.

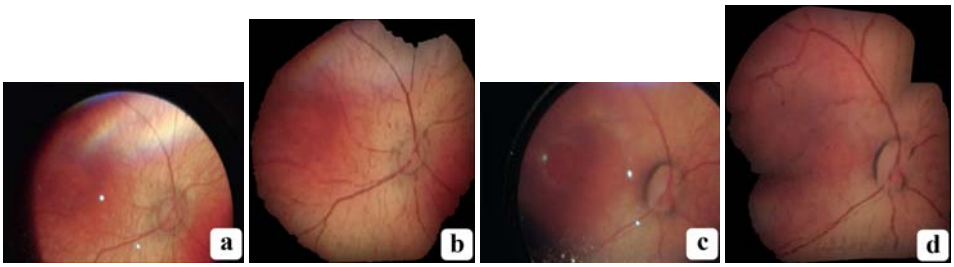

Fig. 4. VEVIO images: Two pairs of manually selected frames (a), (c) and two automatically generated mosaics (b), (d). Although each pair was obtained from the same video, the manual frames were not used to generate the mosaics. The mosaics were constructed using selected source frames as described in [29].

2.3. *Exploratory Dijkstra Segmentation*

Dijkstra' minimum cost algorithm solves Eq. 7 for any graph with non-negative arc costs [28]. More generally, it finds a minimum cost path $\tilde{\gamma}(\mathbf{s},\mathbf{v})$ between a single source vertex $\mathbf{s}$ and (potentially) every other vertex $\mathbf{v}$ in the graph.

As discussed earlier, instead of simply connecting user-defined points, we employ an *exploratory* strategy by using the single-source, many-destinations version of Dijkstra's method. Starting from a single position $\mathbf{s}$ on a major vessel, this strategy enables us to segment this major vessel and all the less prominent vessels that branch out of it, without any need for setting any destination point. Instead, we set an exploration threshold τ on the cost of any path, and find all the minimum-cost paths $\tilde{\gamma}$ from $\mathbf{s}$ in G such that $c(\tilde{\gamma}) < \tau$. Algorithm 1 outlines our exploratory Dijkstra vessel segmentation method.

With the lattice arc costs defined in Eq. 2, the exploratory Dijkstra algorithm will preferentially visit vascular pixels before exploring non-vascular ones, since the cost to reach the latter is generally much higher. When it stops, it will have visited the *Dijkstra region*:

$$R_\tau(\mathbf{s}) = \{\mathbf{v} \,|\, \tilde{\gamma}(\mathbf{s},\mathbf{v}) \leq \tau\} \,. \tag{8}$$

The segmentation's accuracy is thus dependent on the value of τ. However, our choice of τ is made less sensitive by the exponential in Eq. 2, which increases the separation between the vascular and non-vascular pixel classes. This lower sensitivity reduces both the problem of "leakage", in which a segmentation goes beyond the correct vessel boundary and the problem of stopping too soon. For our experiments, we set τ based on the training set images, as explained in Section 3.

Input: Graph G over image domain V, inverted Laplace-Gabor responses F, exploratory threshold τ, filtering threshold ψ.
Output: Dijkstra forest $\mathbf{R}$.
$\mathbf{R} = \emptyset$;
while $\mathbf{s} < \psi$ **do**
  $\mathbf{s}$ = `argmin`$_V(F)$;
  R = `exploratory_dijkstra(`$G, \mathbf{s}, \tau$`)`;
  $\mathbf{R} = \mathbf{R} \cup R$;
  $V = V \setminus \mathbf{R}$;
end

Algorithm 2. Dijkstra forest vessel segmentation: The algorithm adds disjoint Dijkstra regions until the minimum inverted Laplace-Gabor response at the source pixel exceeds ψ. The operation $V \setminus \mathbf{R}$ represents $\{x \in V \mid x \notin \mathbf{R}\}$.

2.4. *Dijkstra Forest*

The exploratory Dijkstra method outlined in Subsection 2.3 efficiently segments a Dijkstra region $R_\tau(\mathbf{s})$ given a single source vertex $\mathbf{s}$. As Fig. 3 (d) exemplifies, however, the vasculature

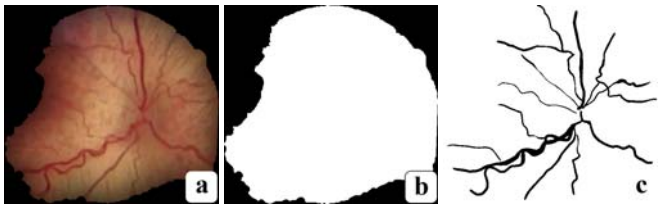

Fig. 5. VEVIO ROI: **(a)** The original mosaic. **(b)** The binary mask outlining the ROI for the mosaic in (a). **(c)** The corresponding manual gold standard. Only pixels that appear white in (b) are taken into account for the metrics tallied in our results.

in the retina extends from more than one primary vessel. Furthermore, the low quality and blur of VIO frames can obscure large sections of the vascular network, and break up the vasculature into several disconnected regions. Therefore, in order to segment all visible vessels better, we extend the single source method to multiple sources.

To this end, we first generate the initial region $R_0 = R_\tau(\mathbf{s}_0)$ from a first source point $\mathbf{s}_0$ as described above. We then select a new source vertex $\mathbf{s}_1$ from those vertices in V that are not part of R_0, and generate a new region R_1 from it, such that $R_0 \cap R_1 = \emptyset$. By repeating, we thus form a *Dijkstra forest*:

$$\mathbf{R} = \{R_0, R_1, \cdots, R_K\}, \qquad \text{where} \qquad F(\mathbf{s}_0) \leq \ldots \leq F(\mathbf{s}_K) \leq \psi. \tag{9}$$

Here, ψ is a threshold on the highest allowable inverted Laplace-Gabor response. We stop adding new regions to the forest when the highest response outside $\mathbf{R}$ is higher than ψ. Algorithm 2 outlines the complete Dijkstra forest computation. As with τ, we determine ψ in our experiments using the training set of images in our database (Section 3). In our experiments, each image requires around 10 source vertices.

3. Experiments

To validate the effectiveness of our proposed segmentation method, we collected a new VIO retinal vessel dataset from pediatric patients and manually segmented the corresponding vascular system to produce the associated ground truth. In this section, we outline the dataset construction process and our methodology for comparing the various segmentation methods to the ground truth.

3.1. Benchmark dataset

Existing benchmark retinal vessel segmentation datasets such as the DRIVE [32], STARE [14] and REVIEW [33] databases do not include VIO images. The relatively lower quality and artifacts in VIO images present a number of unique challenges for automated analysis methods. Thus, there is a need for a benchmark VIO dataset. To address this issue, we constructed a thirty-two image database of VIO images, the Vessel Extraction in Video Indirect Ophthalmoscopy (VEVIO) dataset. VEVIO consists of sixteen manually selected frames and sixteen corresponding enhanced large FOV mosaics from sixteen different premature infants imaged. All images are of each patient's right eye. Figure 4 showcases some of the frames and mosaics in the dataset. Four steps were needed to construct the VEVIO dataset: video recording, manual frame selection, automatic mosaicing and manual vessel segmentation.

3.1.1. VIO recording

This study was approved by the Duke University Institutional Review Board. Informed consent was obtained from parents or legal guardians of all participating infants. All VIO videos were acquired during ROP clinical bedside examinations at the Duke Medical Center, Durham, NC, USA. Each examination was carried out between August and October 2010. The videos were

Table 1. Segmentation results on the test set: The results include the twenty-two test images: eleven manual frames and eleven automatic mosaics. Existing methods were applied to both the raw frames and the frames pre-processed with DLCF. Our proposed method outperforms existing approaches, even with pre-processing, regardless of metric.

Method	F-measure	Kappa	Accuracy	A_z
The proposed method	**0.5228** ($\pm$ 0.07)	**0.4987** ($\pm$ 0.07)	0.9337 ($\pm$ 0.05)	**0.8647** ($\pm$ 0.06)
Matched filters[b]	0.489 ($\pm$ 0.09)	0.4646 ($\pm$ 0.09)	0.9322 ($\pm$ 0.05)	0.7977 ($\pm$ 0.08)
Local entropy[b,*]	0.4504 ($\pm$ 0.11)	0.4049 ($\pm$ 0.16)	0.8839 ($\pm$ 0.19)	0.7104 ($\pm$ 0.1)
GMM Gabor[b,*]	0.3234 ($\pm$ 0.19)	0.3046 ($\pm$ 0.19)	**0.9341** ($\pm$ 0.04)	0.7921 ($\pm$ 0.17)
Matched filters[a,*]	0.3847 ($\pm$ 0.17)	0.3313 ($\pm$ 0.2)	0.7481 ($\pm$ 0.34)	0.7682 ($\pm$ 0.1)
Local entropy[a,*]	0.2808 ($\pm$ 0.23)	0.2545 ($\pm$ 0.22)	0.892 ($\pm$ 0.17)	0.7106 ($\pm$ 0.1)
GMM Gabor[a,*]	0.2861 ($\pm$ 0.2)	0.2652 ($\pm$ 0.19)	0.9304 ($\pm$ 0.04)	0.7716 ($\pm$ 0.18)
K-means Gabor[b,*]	0.1777 ($\pm$ 0.12)	0.1667 ($\pm$ 0.12)	0.9328 ($\pm$ 0.04)	0.7727 ($\pm$ 0.16)
K-means Gabor[a,*]	0.1536 ($\pm$ 0.12)	0.1411 ($\pm$ 0.12)	0.9308 ($\pm$ 0.04)	0.7599 ($\pm$ 0.17)

[a] Raw frames. [b] Pre-processed frames. [*] F-measure: $p < 0.05$.

recorded using a Keeler Wireless Digital Indirect Ophthalmoscope (Keeler Instruments Inc, Broomall, PA, USA). An assistant operated the video recording software provided by Keeler on a computer at the bedside. Each video was recorded at a resolution of 720×576 pixels in 24-bit color and saved as an interlaced, compressed Audio Video Interleaved (AVI) file.

3.1.2. Manually selected frames

During the recording of the bedside examination, the assistant viewed a real-time feed of the video being recorded and manually screen-captured a number of frames, using Keeler's recording software, when she considered that the video feed was well-centered and in focus. One of the authors (MTC) later examined each set of manually captured frames and selected the highest quality image of each right eye.

3.1.3. Automatic mosaics

To generate the corresponding ten mosaics, we applied our automatic mosaicing pipeline [29] to each video. The set of frames suitable for mosaicing into a single image were automatically selected by our method and did not rely on the manually captured frames. From the thousands of frames in each video, our method retained the twenty frames with the highest frame-quality scores. Each mosaic was constructed from five of those twenty frames. While it is possible to construct the mosaic from the highest five scoring frames directly, to ensure that the mosaics had the widest possible field of view we manually selected the final five frames.

3.1.4. Manual vessel segmentation

In order to provide a quantitative assessment of the various automated methods' performance, we produced a gold standard segmentation by manually tracing all the visible retinal vessels in each of the twenty VIO images. MTC, a practicing ophthalmologist, traced each image in Adobe Photoshop CS3 (Adobe Systems Inc., San Jose, CA) using a Wacom Intuous3 graphics tablet (Wacom Co. Ltd, Kazo-shi, Saitama, Japan). This tablet uses a pressure-sensitive pen that mimics a real brush, thus allowing the user to dynamically alter the thickness of a pen stroke. The set of vessel tracings for each VIO image were then saved as a separate binary image mask.

3.2. Comparison to other methods

We divided the VEVIO dataset into a training set of ten images and a test set of twenty-two images. Each set included frame/mosaic pairs taken from the same videos, so that there was no overlap between the training and test patients. In order to compare our method to existing methods fairly, we contacted a large number of research groups who had developed methods

Table 2. Segmentation results on the single (not mosaiced) test frames: Each method was trained or optimized using the frames in the training set and the parameters where then kept fixed for the testing stage. Existing methods were applied to both the raw frames and the frames pre-processed with DLCF.

Method	F-measure	Kappa	Accuracy	A_z
The proposed method	**0.5403** ($\pm$ 0.06)	**0.5127** ($\pm$ 0.06)	0.9101 ($\pm$ 0.06)	**0.8773** ($\pm$ 0.06)
Matched filters[b]	0.5025 ($\pm$ 0.07)	0.4745 ($\pm$ 0.07)	0.9086 ($\pm$ 0.06)	0.7735 ($\pm$ 0.1)
Local entropy[b,*]	0.4347 ($\pm$ 0.1)	0.3646 ($\pm$ 0.2)	0.8123 ($\pm$ 0.25)	0.7092 ($\pm$ 0.05)
GMM Gabor[b,*]	0.2297 ($\pm$ 0.19)	0.2128 ($\pm$ 0.18)	0.9153 ($\pm$ 0.05)	0.7182 ($\pm$ 0.18)
Matched filters[a,*]	0.2938 ($\pm$ 0.18)	0.2078 ($\pm$ 0.19)	0.5402 ($\pm$ 0.4)	0.7144 ($\pm$ 0.11)
Local entropy[a,*]	0.0955 ($\pm$ 0.15)	0.0638 ($\pm$ 0.15)	0.8284 ($\pm$ 0.22)	0.7097 ($\pm$ 0.05)
GMM Gabor[a,*]	0.1549 ($\pm$ 0.15)	0.134 ($\pm$ 0.13)	0.908 ($\pm$ 0.05)	0.6771 ($\pm$ 0.17)
K-means Gabor[b,*]	0.1348 ($\pm$ 0.12)	0.1258 ($\pm$ 0.12)	**0.916** ($\pm$ 0.05)	0.7085 ($\pm$ 0.17)
K-means Gabor[a,*]	0.0866 ($\pm$ 0.1)	0.0746 ($\pm$ 0.09)	0.9121 ($\pm$ 0.05)	0.6761 ($\pm$ 0.17)

[a] Raw frames. [b] Pre-processed frames. [*] F-measure: $p < 0.05$.

for retinal vessel segmentation. The results presented here were all obtained using the source code of the groups that kindly made their methods available to us.

In this work, we were able to test both supervised and unsupervised state-of-the-art approaches. We obtained source code for the unsupervised methods of Chaudhuri et al. (matched filters) [5] and Chanwimaluang and Fang (local entropy) [15]. We also obtained code for the supervised classification based on Gabor responses of Soares et al. [9]. For the latter, we tested two types of classifiers: Gaussian mixture models (GMM) and K-nearest neighbors (KNN).

For the supervised methods, we trained the different classifiers on the training data using the learning code made available by Soares et al. [9]. For the unsupervised methods, we optimized their parameters by exhaustively determining the values which resulted in the best possible F-measure for the training set. We then kept the parameters fixed for the testing stage. The optimal thresholds for each method are summarized in Table 4. We tested each existing method on the manually selected frames in two ways: (1) using the raw frames directly captured from the video and (2) using the frames after DLCF pre-processing. The raw frames capture how existing methods fare on VIO data as is, while the pre-processed images allowed us to gauge how our Dijkstra forest segmentation itself compared to other methods on the same source data.

4. Results

Our experimental results are summarized in Tables 1, 2 and 3. Each table is ranked according to the F-measure (Appendix A) in the first column. Table 1 includes all twenty-two testing set images (eleven frames and eleven mosaics). As noted above, the testing data includes only data from new patients that were not part of the training data. This table illustrates how a method generalizes to novel data, regardless of image type. The subscripts next to each state-of-the-art method indicate whether we used the raw frames or the frames after DLCF filtering. For all methods, the filtered frames allowed significantly better results. Table 2 includes the segmentation results for the test frames, while Table 3 tallies the results for the test mosaics.

Each table includes the mean F-measure, Cohen's Kappa [34], accuracy (Appendix A), and area under the ROC curve (A_z) for each method with the corresponding standard deviation in parentheses. For each image, each metric was calculated inside a region-of-interest (ROI) that only includes the image's retinal pixels. We obtained each image's ROI by applying our hue masking method outlined in [29]. In short, hue masking retains only those pixels that match the color profile of the current retina. Figure 5 illustrates the ROI mask for a particular mosaic.

Each metric was determined on a pixel-by-pixel basis. For a given automatic segmentation, a pixel is considered a *true positive* if both it and the matching pixel in the ground truth image are ones. If both are zero, it corresponds to a *true negative*. A mismatch in which the automatic

Table 3. Segmentation results on the test mosaics: Each method was trained or optimized using the mosaics in the training set and the parameters were then kept fixed for the testing stage.

Method	F-measure	Kappa	Accuracy	A_z
The proposed method	**0.5053** ($\pm$ 0.08)	**0.4847** ($\pm$ 0.08)	**0.9573** ($\pm$ 0.01)	**0.8522** ($\pm$ 0.05)
Matched filters	0.4755 ($\pm$ 0.1)	0.4547 ($\pm$ 0.1)	0.9559 ($\pm$ 0.01)	0.8219 ($\pm$ 0.04)
Local entropy	0.466 ($\pm$ 0.1)	0.4453 ($\pm$ 0.1)	0.9556 ($\pm$ 0.01)	0.7115 ($\pm$ 0.14)
GMM Gabor	0.4172 ($\pm$ 0.15)	0.3964 ($\pm$ 0.15)	0.9529 ($\pm$ 0.01)	0.8461 ($\pm$ 0.12)
K-means Gabor*	0.2205 ($\pm$ 0.11)	0.2076 ($\pm$ 0.1)	0.9496 ($\pm$ 0.01)	0.8368 ($\pm$ 0.12)

* F-measure: $p < 0.05$.

segmentation produced a one and the ground truth had a zero is a *false positive*. The converse mismatch is a *false negative*.

Each of the four metrics captures some form of similarity between a method's output and the corresponding ground truth. The retinal vessel segmentation literature has traditionally favored accuracy and area under the ROC curve, A_z, as the primary metrics [5–11, 13, 14, 32]. While A_z is an adequate measure of classifier robustness, as we argue in Appendix A, we believe the F-measure is a much more appropriate measure than accuracy for analyzing segmentation results in this type of data. Due to the very low prior probability of a pixel being part of a vessel, methods that only segment a small fraction of each image will still obtain competitive accuracy scores. The F-measure, on the other hand, provides a ratio-independent summary of the overlap between two segmentation's pixel labels. Therefore, the approach of labeling very few pixels as vascular will yield a very low F-measure score due to the large number of false negatives.

We applied a Wilcox signed-rank test between our proposed method's F-measure distribution and the F-measures of every other method [35]. Methods for which the difference was statistically significant $(p < 0.05)$ are marked with an $*$ in each table.

5. Discussion

As Tables 1, 2 and 3 show, our proposed method compares favorably to existing supervised and unsupervised methods. Regardless of metric, our method consistently outperformed existing state-of-the-art approaches in our experiments by better balancing the likelihood of false positives and negatives. In contrast, Fig. 6 illustrates how a method such as the GMM classifier has good recall, but poor precision, while a more conservative method such as the KNN classifier has better precision, but worse recall. In the first case, the segmentation has too many non-vascular pixels, while the latter segmentation misses a significant portion of the vasculature. Our method's connectivity contraints allow us to strike a good balance between these two objectives by better disambiguating between similarly valued pixels. In other words, our method is more likely to label a pixel as vascular if it can be directly connected to a large vascular region than if it is isolated, since the latter case is more indicative of noise rather than an actual vessel.

Finally, it is also worth noting how DLCF pre-processing has a sizable impact on the segmentation results of existing methods. All state-of-the-art methods performed significantly better on pre-processed frames than raw frames. As an extreme example, note in Table 2 the four-fold improvement in the F-measure of the matched filters method when using pre-processed frames.

In the future, we wish to expand our VEVIO database with more images from more patients. The presented experimental results highlight the challenges that VIO data present for vessel segmentation methods. The F-measures reported in this paper indicate significant room for further improvement. To encourage further research in this area, we have made the VEVIO dataset and the MATLAB code that we have developed for this project publically available at `http://www.duke.edu/~sf59/Estrada_BOE_2012.htm`

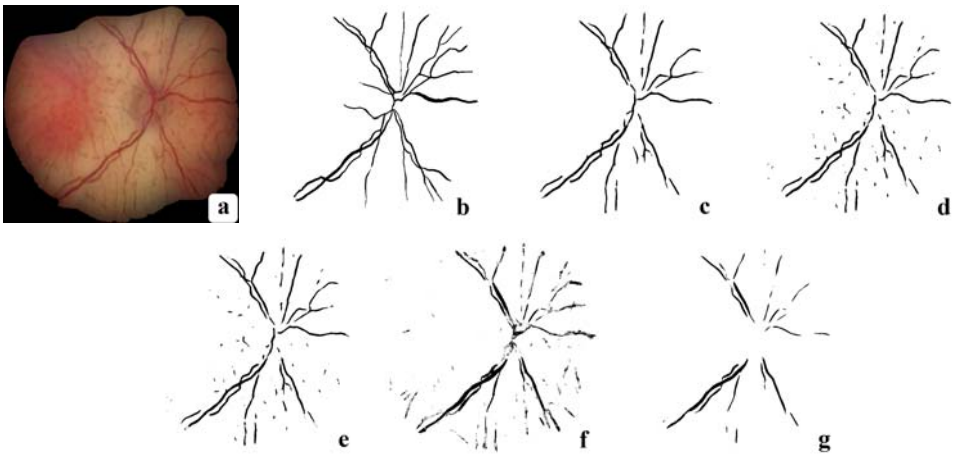

Fig. 6. Vessel segmentation on a mosaic: **(a)** Original image **(b)** Manual segmentation **(c)** Dijkstra forest **(d)** Matched filters **(e)** Local entropy **(f)** GMM classifier **(g)** KNN classifier

Acknowledgements

This work was supported in part by the Knights Templar Eye Foundation, Inc. pediatric ophthalmology research grant, Research to Prevent Blindness 2011 Dukes Unrestricted Grant award, and a fellowship from the Duke University Center for Theoretical & Mathematical Sciences. We would like to thanks Prof. Yuliya Lokhnygina for her invaluable help with the biostatistical analysis. We would also like to thank the research groups that kindly provided their source code to us, especially Dr. João V. B. Soares.

A. F-measure vs. accuracy

Traditionally, accuracy has been one of the key metrics for evaluating vessel segmentation results. For a binary classification, this metric is defined thus:

$$\text{accuracy} = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \tag{10}$$

where t_p and f_p indicate true and false positives respectively, while t_n and f_n tally true and false negatives. The unbiased F-measure, on the other hand, is given by:

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{11}$$

where precision and recall are defined as:

$$\text{precision} = \frac{t_p}{t_p + f_p}, \quad \text{recall} = \frac{t_p}{t_p + f_n}. \tag{12}$$

Accuracy becomes less informative when one of the two classes if far more likely than the other, as is the case for vascular vs. non-vascular pixels. On average, vascular pixels only comprise about 5-10% of an image. This means that a classifier that labels all pixels as non-vascular can already boast a 90-95% accuracy. The F-measure, on the other hand, provides a better balance between labeling pixels correctly or incorrectly, since it is not affected by class sizes.

B. Parameter values

Table 4. Parameter values

The proposed method	Exploratory threshold: 5×10^{-5}	Filtering threshold: 0.7
Matched filters	Raw threshold: 0.5	Pre-processed threshold: 0.2254
Local entropy	Raw threshold: 0.5253	Pre-processed threshold: 0.7677
GMM Gabor	Raw threshold: 0.5	Pre-processed threshold: 0.5
K-means Gabor	Raw threshold: 0.77	Pre-processed threshold: 0.82