Automated Measurement of the Arteriolar-to-Venular Width Ratio in Digital Color Fundus Photographs

Meindert Niemeijer*, Xiayu Xu, Alina V. Dumitrescu, Priya Gupta, Bram van Ginneken, *Member, IEEE*, James C. Folk, and Michael D. Abrámoff, *Senior Member, IEEE*

***Abstract*—A decreased ratio of the width of retinal arteries to veins [arteriolar-to-venular diameter ratio (AVR)], is well established as predictive of cerebral atrophy, stroke and other cardiovascular events in adults. Tortuous and dilated arteries and veins, as well as decreased AVR are also markers for plus disease in retinopathy of prematurity. This work presents an automated method to estimate the AVR in retinal color images by detecting the location of the optic disc, determining an appropriate region of interest (ROI), classifying vessels as arteries or veins, estimating vessel widths, and calculating the AVR. After vessel segmentation and vessel width determination, the optic disc is located and the system eliminates all vessels outside the AVR measurement ROI. A skeletonization operation is applied to the remaining vessels after which vessel crossings and bifurcation points are removed, leaving a set of vessel segments consisting of only vessel centerline pixels. Features are extracted from each centerline pixel in order to assign these a soft label indicating the likelihood that the pixel is part of a vein. As all centerline pixels in a connected vessel segment should be the same type, the median soft label is assigned to each centerline pixel in the segment. Next, artery vein pairs are matched using an iterative algorithm, and the widths of the vessels are used to calculate the AVR. We trained and tested the algorithm on a set of 65 high resolution digital color fundus photographs using a reference standard that indicates for each major vessel in the image whether it is an artery or vein. We compared the AVR values produced by our system with those determined by a semi-automated reference system. We obtained a mean unsigned error of 0.06 (SD 0.04) in 40 images with a mean AVR of 0.67. A second observer using the semi-automated system obtained the same mean unsigned error of 0.06 (SD 0.05) on the set of images with a mean AVR of 0.66. The testing data and reference standard used in this study has been made publicly available.**

Manuscript received May 10, 2011; accepted June 01, 2011. Date of publication June 16, 2011; date of current version November 02, 2011. This work was supported in part by the National Eye Institute (R01 EY017066), in part by the Research to Prevent Blindness, and in part by the Marlene S. and Leonard A. Hadley Glaucoma Research Fund. *Asterisk indicates corresponding author*.

*M. Niemeijer is with the University of Iowa Hospitals and Clinics, Iowa City, IA 52242 USA. He is also with the VA Center of Excellence for Prevention and Treatment of Visual Loss, Iowa City, IA 52242 USA (e-mail: meindert-niemeijer@uiowa.edu).

J. C. Folk, A.V. Dumitrescu and P. Gupta are with the University of Iowa Hospitals and Clinics, Department of Opthalmology and Visual Sciences, Iowa City, IA 52242 USA.

X. Xu is with the University of Iowa, Department of Biomedical Engineering, Iowa City, IA 52242 USA.

B. van Ginneken is with the Radboud University Nijmegen Medical Centre, Department of Radiology, Diagnostic Image Analysis Group, 6525 GA Nijmegen, The Netherlands.

M.D. Abrámoff is with the University of Iowa Hospitals and Clinics, Department of Opthalmology and Visual Sciences, Iowa City, IA 52242 USA. He is also with the University of Iowa, Department of Electrical and Computer Engineering, and also the Department of Biomedical Engineering, Iowa City, IA 52242 USA and also with the VA Center of Excellence for Prevention and Treatment of Visual Loss, Iowa City, IA 52242 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2011.2159619

***Index Terms*—Artery, artery vein ratio, artery vein separation, retina, vasculature, vein.**

I. Introduction

THE accurate measurement of retinal vessel parameters is an important problem in eye research, especially in retinal image analysis. Diseases can alter the width of portions, or the entire length, of retinal vessels, increase their curvature or tortuosity, and/or change their reflectance of light. Cardiovascular disease can decrease the width of arteries and increase the widths of veins. Though this fact was appreciated by ophthalmologists for years, only recently, through precise and cumbersome measurements of arterial and venous widths, has it become clear that small changes in the ratio between the widths of arterioles and venules, the arteriolar-to-venular diameter ratio (AVR), are associated with increases in the risk for stroke, cerebral atrophy, cognitive decline, and myocardial infarct. [1]–[3]. Other diseases, including diabetic retinopathy and retinopathy of prematurity are also known to affect the AVR (see [4] for a comprehensive review). Unfortunately, relevant changes in AVR are too subtle to be detected by ophthalmologists during clinical examination and the process of manually or semi-automatically estimating the AVR from digital fundus photographs is too cumbersome and laborious for clinical practice.

Thus, an automated method for determination of the AVR may have substantial impact on clinical practice, and may lead to an improved assessment for patients at risk for cardiovascular and brain disease. This study focuses on a fully automated determination of the AVR from digital color fundus photographs.

The accurate estimation of AVR is challenging and requires optic disc detection, vessel segmentation, accurate vessel width measurement, vessel network analysis, and artery vein classification. Optic disc detection is necessary to determine the location of the region of interest (ROI) where the measurements are obtained. Vessel segmentation must be used to find the vessels themselves and the width of the vessels. Any AVR measurement system must identify which vessels are arteries and which are veins with high accuracy since small classification errors can have a large influence on the final AVR. Please note that while arteriolar/venular are the correct terms, we will use the terms artery/vein interchangably in this work.

Methods for both AV classification as well as AVR determination have been presented previously. A semi-automatic method for the analysis of retinal vascular trees in which the venous and arterial trees were analyzed separately was presented by Martinez-Perez *et al.* [5]. More recent work by Rothaus *et al.* [6]

proposes a method to label all vessels as either artery or vein using an existing vessel segmentation and some manually labeled starting vessel segments. Li *et al.* [7] presented a method for automatically determining the AVR, however, that method still required manual user input to separate arteries from veins. Automated AV classification was first presented by Grisan *et al.* [8]: in this method the vasculature is segmented using a vessel tracking and analysis procedure and the vessel centerlines are detected. After defining an area of interest around the optic disc and dividing this area into four quadrants, color based features are extracted from the vessel segments that are then classified into arteries and veins using an unsupervised clustering method. This method had a total error rate of 12.4% on 24 images. Kondermann *et al.* [9] presented a method for AV classification based on features extracted from vessel profiles as well as features based on the local image intensities around the vessel centerline. The authors obtained an accuracy of 95.32% for assigning manually segmented vessel pixels to the correct class in four images. A combined clustering and classification approach for separating arteries and veins was presented by Vázquez *et al.* [10]. The authors compared different feature sets and classification approaches. They also tested the influence of the distance to the optic disc on their measurements and achieved an accuracy rate of 86.34% for assigning vessels to the artery or vein class. However, not all vessels detected by the system were included in the reference standard.

Ruggeri *et al.* [11] was the first group to present a method with an evaluation based on the actual AVR that was measured manually in the image. The authors report a correlation in 14 images varying between 0.73 and 0.83 depending on how the AVR was calculated. Tramontan *et al.* [12] further extended the algorithm with enhanced vessel tracking and structural AV discrimination features obtaining a correlation of 0.88 on 20 images. The system was then implemented as a webservice which allowed graders to check and change the algorithm results [13]. An evaluation by three graders analyzing 30 fundus images showed high reproducibility of the system's measurements between the three graders.

The system we propose and evaluate in this paper combines methods that were previously described by our group as well as newly proposed techniques. The method automatically detects the optic disc, segments the vasculature, determines the vessel width, classifies the detected vessels into arteries and veins and calculates the AVR in the ROI. It is completely automatic and requires no user input. We compare the AVR output of the automatic, proposed algorithm to that of the IVAN semi-automated method developed at the University of Wisconsin. This is the accepted reference standard, and has been used in major studies associating AVR with cardiovascular disease and prognosis [4]. We also validated the classification system that classifies vessels into arteries and veins since it is such an important subcomponent.

This paper is structured as follows. Section II describes the various components of the system. In Section III the data used is presented. Section IV discusses the performed experiments and shows the obtained results. Section V discusses the results of the automated system compared to previous systems as well as its clinical importance.

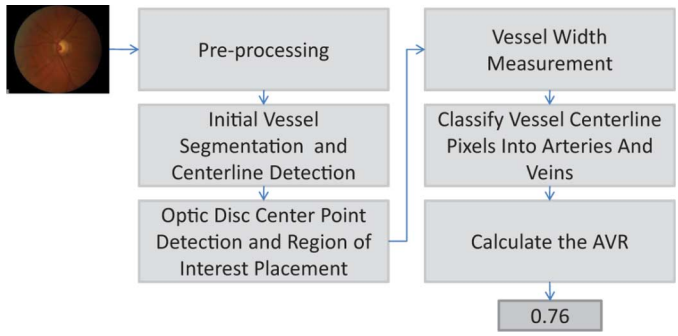

Fig. 1. Overview of all steps in the proposed method.

Fig. 2. Image overlaid with the automatically determined ROI region. The region labeled "B" is where AVR measurements are taken.

II. Methods

Knudtson *et al.* [14], [15] published a protocol for the measurement of the AVR. The protocol precisely defines where and how measurements should be obtained (see Section II-D). The automated method described herein follows this protocol whenever possible. It starts by preprocessing an image to remove the gradient around the border of the field-of-view (FOV) as well as to remove slow intensity variations in the image. Further preprocessing is focused on the detection of the two most important anatomical landmarks. These are the optic disc, the landmark on the retina around which the measurements are obtained and the vasculature, the structure that is actually being measured. After detecting the location of the optic disc, we automatically place the ROI, as defined in the protocol, in the image. Vessel width measurements are obtained and the vessels within the measurement area are classified into arteries and veins. Finally, the AVR is determined. These steps are also shown in Fig. 1.

A. Preprocessing

We have previously described methods for the preprocessing steps of the algorithm and will, therefore, only briefly discuss these here. The interested reader is referred to the references included with the method descriptions.

1) Field of View Mirroring and Background Removal: Digital color fundus photographs have a black border around the FOV (see Fig. 2). The large gradient can disturb feature measurements near the FOV border. It is removed by applying a previously presented mirroring technique [16]. This method mirrors the pixel values from within the circular field of view to outside the FOV. This operation was performed at the original image resolution.

Slow background variations were removed by blurring the image with a Gaussian filter with a large standard deviation and subtracting the blurred image from the original. The value of the standard deviation of the Gaussian filter is not a critical parameter as long as it is large enough to ensure the blurred image contains no visible structures such as vessels. This procedure was performed on both the red as well as the green color planes separately. For all image processing operations the green plane is used, additionally, both the green and red planes are used in the artery vein classification. From here, whenever the green and red plane is mentioned, it refers to the preprocessed versions. The blue color plane was not used.

2) Vessel Segmentation and Preprocessing: To segment the retinal vasculature, a previously presented method [17] based on pixel classification was used. The filter outputs of a Gaussian filter bank were used as features to train a kNN-classifier [18] to detect the vasculature. This method is not scale independent as the Gaussian filterbank features are extracted at particular scales. Additionally, the images with which the vessel segmentation method has been trained (the DRIVE database [19]) have a particular resolution and therefore a particular range of vessel widths (measured in pixels). As the data used in this study had a high resolution, applying the standard method trained with the low-resolution DRIVE data would not produce adequate results. To attain reasonable performance for the vessel segmentation step we have downsampled the images with a factor of 4 before applying the vessel segmentation.

The vessel segmentation method assigns each pixel in the image a likelihood between 0 and 1 that the pixel is within a vessel. This results in a "vesselness map" that can be thresholded to produce a binary vessel segmentation. Before thresholding, we upsampled the vesselness map back to the resolution of the original image, using quintic spline interpolation. To analyze the vessel network we then applied a skeletonization algorithm [20], [21] on the thresholded likelihood map, reducing all vessels to a single centerline one pixel wide. Threshold $T = 0.3$ was used since it gave good results on the training data (the training data is described in Section III). After the skeletonization of the segmented vessels, cross-over and bifurcation points were removed by counting the number of neighbors for all centerline pixels and removing those with more than two neighbors. This is necessary because the vessel width and angle in bifurcations is not well defined and/or difficult to measure in the case of cross-over points. This operation subdivides the vascular network into a collection of vessel segments that are individually analyzed.

3) Optic Disc Detection and ROI Determination: A supervised position regression method [22] was used to detect the centerpoint of the optic disc. This method estimates how far a certain position in the image is from the optic disc center. The estimation is based on measurements obtained in the image and from the vessel segmentation. The target location is found by obtaining estimates in many locations in the image, eliminating those locations that are estimated to be far from the optic disc and searching around the locations estimated to be close to the optic disc center. No prior assumptions about the location of the optic disc in the image are made. The method is first trained using a large set of images (both optic disc and fovea centered) for which the location of the optic disc is known. Our method does not segment the optic disc, therefore we have assumed the optic disc in our testing data to have a constant size. This is a valid assumption due to the uniformity of the data used in this study. All of the images have the same resolution in pixels and cover approximately the same area of the retina. After performing some manual measurements in the training set we assumed a value of 360 pixels for the optic disc diameter (DD), at this diameter, the vast majority of optic discs is completely covered. The optic disc detection proceeds as follows, the automated method detects the center of the optic disc and a circle with diameter of 360 pixels is placed at this location. We assumed this circle corresponds with the optic disk outline in the image.

The AVR calculation protocol [15] precisely defines the region of interest (ROI) in which the AVR should be measured. This ROI is centered on the optic disc (see Fig. 2). The ROI consists of several circular regions whose size is based on the approximate diameter of the optic disc. Region A is between 0.5 and 1 DD from the optic disc center and region B, where vessel measurements are taken, is between 1 DD and 1.5 DD from the optic disc center. All analyses and evaluations performed in this work are based on measurements within region B.

B. Vessel Width Measurement

After vessel segmentation and preprocessing, the vasculature was thinned and subdivided into a set of vessel segments. All vessel segments that were not (partly) inside region B (see Fig. 2) were removed as they were not used in the AVR analysis. Even though the vessel segmentation algorithm we used can successfully localize most vessels, i.e., wide and narrow ones, choosing a single threshold to produce a binary segmentation of the vasculature that can be used to determine the local vessel width is difficult. The values in the likelihood map, as produced by the vessel segmentation algorithm, tend to zero as one moves away from the vessel border into the retinal background. This effect is also dependent on the vessel width with narrower vessels having an overall lower likelihood response than wider vessels. Consequently, relatively small variations in the applied threshold result in substantial vessel width differences and thresholds that give visually pleasing results for wider vessels completely miss smaller vessels. This is further complicated by the fact that the vessel detection is performed at a lower resolution resulting in larger errors after upsampling.

The semi-automated software that was used to establish the reference standard for our AV ratio measurements, IVAN, uses a technique called "tobogganing" [23], [24] in its vessel segmentation algorithm. As we wanted to obtain similar vessel widths to IVAN, we have combined tobogganing with our vessel segmentation. This has the added benefit that we can do the vessel width analysis on the original images in their original resolution instead of on the subsampled images on which the vessels were segmented.

1) Combining Tobogganing and Vessel Pixel Classification: Tobogganing is a segmentation technique that subdivides the image into areas (i.e., "splats") that are homogeneous based on a certain criterion. The technique's results are analogous to the "catchment basins" in watershed segmentation [25]. We used

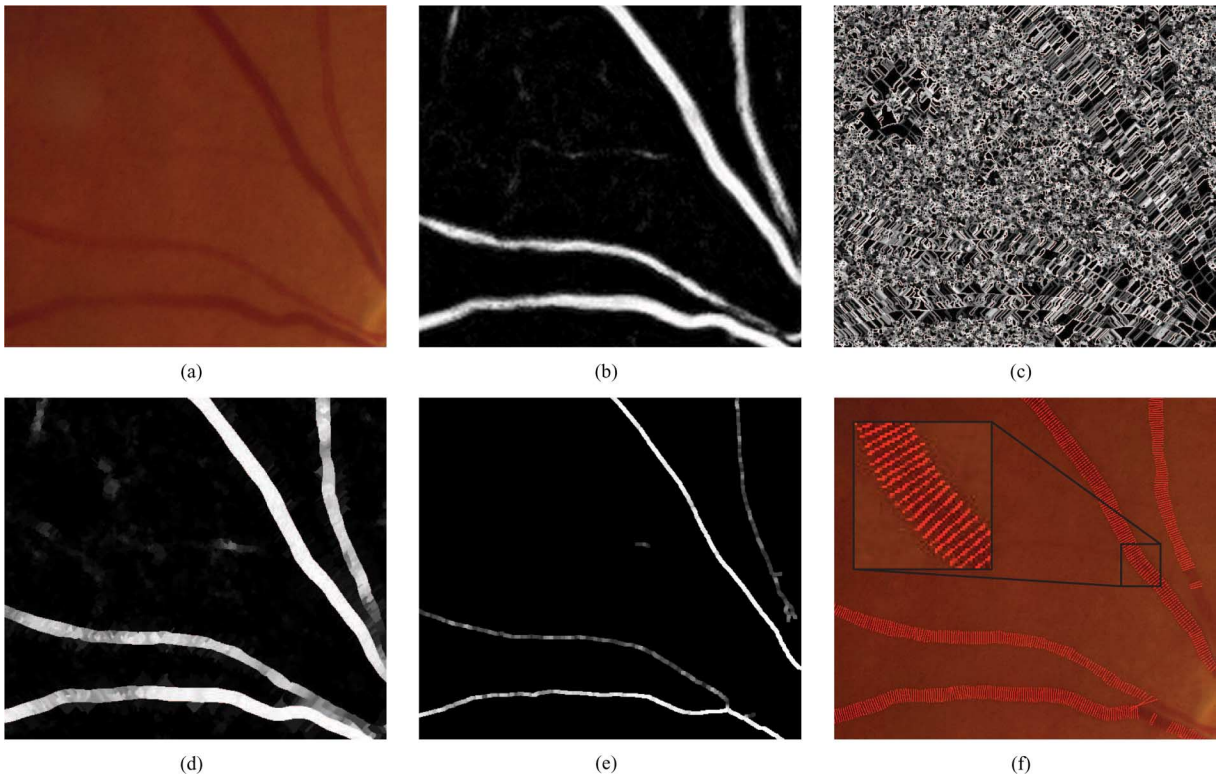

Fig. 3. Image showing the most important processing steps described in Section II-A on a small sub-image. (a) The color fundus image. (b) The vessel likelihood map. A higher pixel value in this image means a higher vessel likelihood. (c) The splat map, the borders of the individual splats are shown. (d) The processed vessel likelihood map. Each splat has been assigned the median vessel likelihood value under the splat. (e) The vessel centerlines where each centerline pixel has been asigned the likelihood it is inside a vein (dilated for display). (f) The final width measures superimposed on the original image. Only every third measurement was plotted.

the multiscale gradient magnitude image [25] to define homogeneity. To calculate the gradient magnitude, the image is convolved with a first derivative of Gaussian filter in both the x and y direction after which the magnitude of the gradient is obtained for each pixel. The gradient magnitude at various scales (i.e., various standard deviations σ) tends to have a maximum at the border of high contrast structures such as the vasculature. A lower scale filter will give more response at the border of small vessels and a higher scale filter will give more response at the border of wider vessels. To obtain the multiscale gradient magnitude image we calculated the scale-normalized gradient magnitude at scales $\sigma = 1$, 2,3,4, 5, 6 and used the maximum value over scale for each of the pixels in the image. After applying the tobogganing method, a splat map with around 150 000 splats [Fig. 3(c)] was obtained. The likelihood map produced by the vessel segmentation algorithm can now be used to determine for each splat the likelihood it is inside a vessel. This is accomplished by assigning to each splat the median likelihood value of all the pixels that are part of the splat. We can assume the splats are either inside or outside the vessel (note that this assumption does not always hold in the case of low contrast, narrow vessels). Given a correct likelihood map this results in the splats inside the vessel being assigned a higher likelihood than those outside the vessel. In the resulting vessel map, the borders of the vessels are better defined and the widths of the vessels become less dependent on the chosen vessel segmentation threshold. This enhanced vessel map was used to determine the vessel width. Fig. 3 shows the process in more detail on a small subimage taken from one of the test images.

2) Measuring the Vessel Width: Measurement of the local vessel width must be performed perpendicular to the local vessel angle in order to minimize errors during the measurement process. The local vessel angle is determined for all centerline pixels in every vessel segment. We have defined the local vessel angle as the direction of the largest eigenvector of the covariance matrix of the coordinate of the centerline pixel along with the coordinates of its seven connected neighbors to both sides (i.e., 15 coordinates in total). As it is unknown where the vessel begins or ends, the range of the angles is $[0 \ldots \pi]$. Near the end of the vessel segment only centerline coordinates inside the vessel are used, 8 for the end pixel.

For each centerline pixel the local vessel width was measured by finding the left and right vessel edges in the enhanced vessel map and calculating the distance between them. To determine the locations of the edges of the vessel, the likelihood was measured along a line through the centerline pixel and perpendicular to the local vessel angle. Starting from the centerline pixel the vessel border was found in both the left and right directions. The likelihood threshold at which a splat is no longer part of a vessel is a critical parameter in this algorithm. As the likelihood assigned to vessel splats varies over the image and is dependent on local vessel contrast and vessel width, a local vessel threshold was determined for every centerline pixel. The vessel likelihood under the centerline pixel was multiplied by a ratio

TABLE I
COMPLETE SET OF FEATURES EXTRACTED FOR EACH CENTERLINE PIXEL

Nr.	Feature description
1-3	Normalized Mean Hue, Saturation and Intensity across the vessel.
4-5	Normalized Mean Red and Green plane intensities across the vessel.
6-8	Standard deviation of Hue, Saturation and Intensity across the vessel.
9-10	Standard deviation of Red and Green plane intensities across the vessel.
11-13	Normalized Hue, Saturation and Intensity under the centerline pixel.
14-15	Normalized Red and Green plane intensity under the centerline pixel.
16-19	Normalized highest and lowest intensity in the Red and Green plane across the vessel.
20-27	Intensity under the centerline pixel in a Gaussian blurred ($\sigma = 2, 4, 8, 16$) version of the Red and Green plane.

to determine the appropriate value for the vessel threshold. A ration of 0.7 was found to give good results on the training set.

After all vessel widths for a vessel segment were determined, error correction was performed by finding sudden, local changes in the vessel width. Based on measurements in the training set, we defined sudden changes as more than 3 pixels from one centerline pixel to the other. These happen when a vessel splat is not included or a background splat is included in the vessel width measurement. The ratio threshold was varied locally until the width measurement was similar (<3 pixels difference) to the average width at the preceding 8 vessel centerline pixels. If the vessel width could not be adjusted to match the mean width, the originally detected vessel width was kept. Even though this would leave possibly erroneous measurements in the analysis, failed measurements mostly occurred in narrow, low contrast vessels that had little influence on the AVR. In addition to measuring the vessel width we also stored the location of the left and right vessel boundary for each centerline pixel. Using these two points, a profile across the vessel was defined, this profile is used to extract feature data from across the vessel.

C. Classification Into Arteries and Veins

To enable separate analysis of the arteries and the veins in the image, the previously detected vessel segments need to be assigned to one of these two classes. We used a supervised system, i.e., trained with examples. The system is an adapted and enhanced version of the classification system described in [26]. After a one time training procedure the method can be used to classify previously unseen centerline pixels into either artery or vein (AV classification). The pre-processing procedure and vessel width measurements as detailed in Sections II-A and II-B are applied to all images in the training set. An expert indicated whether a given major vessel was an artery or vein for each of the training images.

1) Training Phase: In the training phase a classifier is trained using the expert labeled vessels in the training set, in order to distinguish between both classes of centerline pixels. As not all vessels in the training set were marked as artery or vein, centerline pixels from unmarked vessels were not included in the training dataset. For all remaining centerline pixels in the training images, a set of 27 local features was extracted. Table I shows the list of extracted features, some of the features in this list were used previously in [8] and [12], in general we chose features that characterize the color as well as the color variation in the vessel. All features measured across the vessel are based on the profiles as determined in the previous section.

The absolute color of the blood in the vessels varies between images and across subjects. This variation has several causes. Primarily, the amount of hemoglobin oxygen saturation influences the reflectivity of the blood column, and this difference allows the difference between higher saturation arterial from lower saturation venous blood to be visualized. Next, lens absorption for different wavelengths is influenced by aging and the development of cataract, causing shifts in the spectral distribution of light reflected by blood. Individual difference in pigmentation of the retinal pigment epithelium below the blood vessels also influence the spectrum of reflected light. Finally, across examinations, even from the same subject, differences in flash intensity, flash spectrum, nonlinear optical distortions of the camera, flash artifacts, and focus also cause considerable variability. These factors complicate classification substantially, and normalization to zero mean and unit standard deviation of the vessel color features for every individual image is therefore important for successful classification. After sampling the features for each centerline pixel, the appropriate labels were assigned based on the reference standard and all training samples were stored in a training dataset. This sampling process was repeated for all images in the training set.

The training set was then split into a separate classifier selection training and test set. An extensive comparison of several different classifiers was performed using these two sets: k-nearest neighbor classifier, support vector machine classifier, linear discriminant classifier, and a quadratic discriminant classifier. The classifier that maximized the area under the receiver operator characteristic (ROC) curve [27] was selected. For each experiment the features were normalized to zero mean and unit standard deviation. Note that this is a normalization across all training samples (i.e., images) and is different from the normalization of the color features that is performed for each training image individually. In this preliminary experiment the linear discriminant classifier showed the best results. Feature selection decreased performance on the training set, so the complete set of features was used for the experiment on the test data.

2) Applying the AV Classification to Unseen Data: After the one time training phase was finished, the trained classifier was applied to the images in the test set. All the test images were first preprocessed similarly to the training images. For each centerline pixel the complete set of 27 features was extracted. The trained classifier was then used to assign a soft label $l = [0 \ldots 1]$. Here, a label close to 0 meant the pixel was likely in an artery and a label close to 1 meant a pixel was likely in a vein (see Fig. 4 for an example). We assumed that all pixels in a vessel segment are either in an artery or a vein. Each soft

label assigned to a centerline pixel can be regarded as a vote for the label of the complete segment. Combining these votes can be done in many different ways but we have found, using preliminary experiments on the training set, that taking the median label for the entire vessel segment works well.

Due to variation in the local image characteristics, the soft labels assigned to each of the segments can vary over the image. A global threshold will not always successfully separate the arteries from the veins within a single subject and will vary between subjects. To perform the final classification we used the prior knowledge that arteries and veins usually come in pairs. This means that, when going around the optic disc in a circle in region B of the ROI, one will generally encounter an artery after first encountering a vein and vice versa. Since this rule does not always hold and finding the matching vessel segment for a particular different vessel segment is nontrivial, we propose to use a voting procedure.

During this procedure, all vessel segments intersecting with a circle of a certain diameter around the optic disc and within region B of the ROI were eligible for matching. Finding the nearest neighbor vessel segment on a circle is straightforward, this can be done by finding the nearest intersection point on the circle. The soft AV labels of both points were compared and the vessel segment with the highest soft label received a vote for "vein" and the other received an "artery" vote. Then the next nearest unpaired vessel was selected and the procedure was repeated. It is obvious that the outcome of this procedure is dependent on the starting vessel segment. By picking a different vessel segment as the starting segment, the distribution of the AV votes amongst the vessel segments will vary. All vessel segments eligble for matching were therefore selected once as the starting vessel and the matching procedure was repeated. Finally, the votes were counted and each of the vessel segments was assigned a hard label (i.e., either artery or vein based on the received votes). Vessel segments with a equal number of artery and vein votes were excluded from the analysis.

D. Calculating the AVR

The arteriolar to venular ratio is defined as $\mathrm{AVR} = \mathrm{CRAE}/\mathrm{CRVE}$ where CRAE is the Central Retinal Artery Equivalent and CRVE is the Central Retinal Vein Equivalent. To calculate these numbers, Knudtson *et al.* [15] describe an iterative process for matching up vessels and calculating the CRAE and CRVE. The widest six veins and arteries (these do not have to be paired) are entered although fewer total number of widths can be used in case not enough measurement points are available. Algorithm 1 shows how we implemented this procedure.

The proposed automated method supplements the manual procedure, outlined in the previous paragraph, with several important steps. This is needed to deal with the fact that the final voting procedure as described in Section II-C2 is based on measurements obtained on a circle with a certain diameter. The chance that all vessel segments in the ROI will intersect with a single circle of any particular diameter is small. Therefore, the voting procedure and AVR calculation should be repeated at various diameters within the ROI (see Fig. 5). The diameters we chose were from 1 DD to 1.5 DD in steps of 0.1 DD where DD was 360 pixels, so the voting and AVR calculation procedure was repeated six times. Note that this samples the ROI equidistantly. For each circle diameter, the AV voting procedure is performed, the local vessel width is measured (see Section II-B) and stored in two vectors, A for arteries and V for veins. Next, Algorithm 1 is used to calculate the AVR. The resulting six AVR values are averaged to arrive at the final AVR estimate for the complete image.

Algorithm 1 The algorithm used to calculate the AVR

Input: Vector A of length $|A|$ containing the widths of the found arteries and vector V of length $|V|$ containing the widths of the found veins.
Output: The AVR.

Sort A and V in decreasing order and set the length of both vectors to $\min(|A|, |V|, 6)$.

while $|A| > 1$ **do**
  Select and remove the first element f and last element l from A
  Store $\sqrt{(f^2 + l^2) * 0.88}$ in vector C
  if $|A| <= 1$ **then**
    $A = A \cup C$
    Sort A in decreasing order
  end if
end while
Clear C
while $|V| > 1$ **do**
  Select and remove the first element f and last element l from V
  Store $\sqrt{(f^2 + l^2) * 0.95}$ in vector C
  if $|V| <= 1$ **then**
    $V = V \cup C$
    Sort V in decreasing order
  end if
end while
Divide the remaining element in A by the remaining element in V to obtain the final AVR.

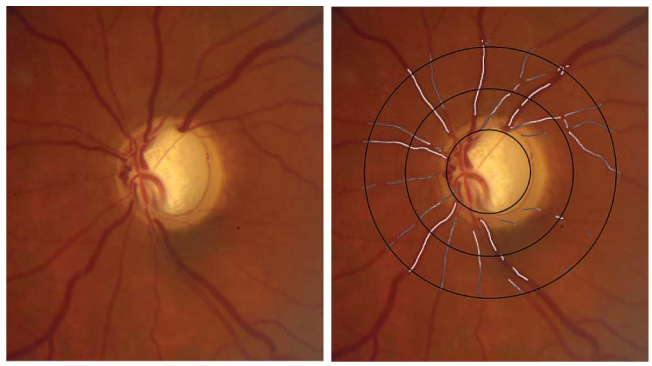

Fig. 4. AV centerline classification result overlaid on a retinal image. A centerline pixel with a higher pixel intensity has a higher likelihood to be a vein according to the algorithm.

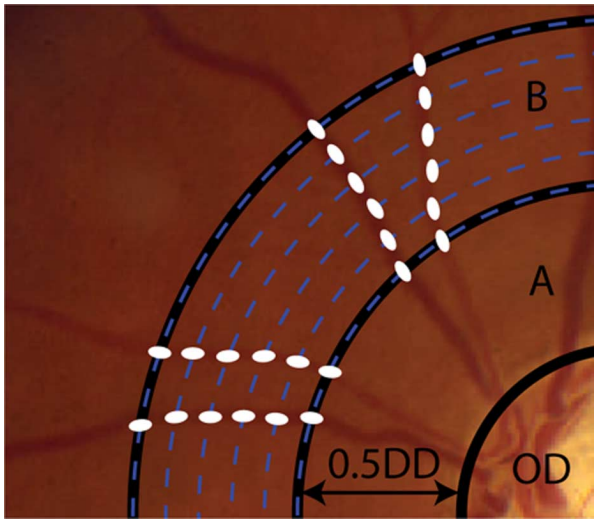

Fig. 5. Detail of a fundus photograph showing the measurement diameters and the vessel crossings where width measurements are obtained. The white dots indicate the points on the vessels where measurements are obtained.

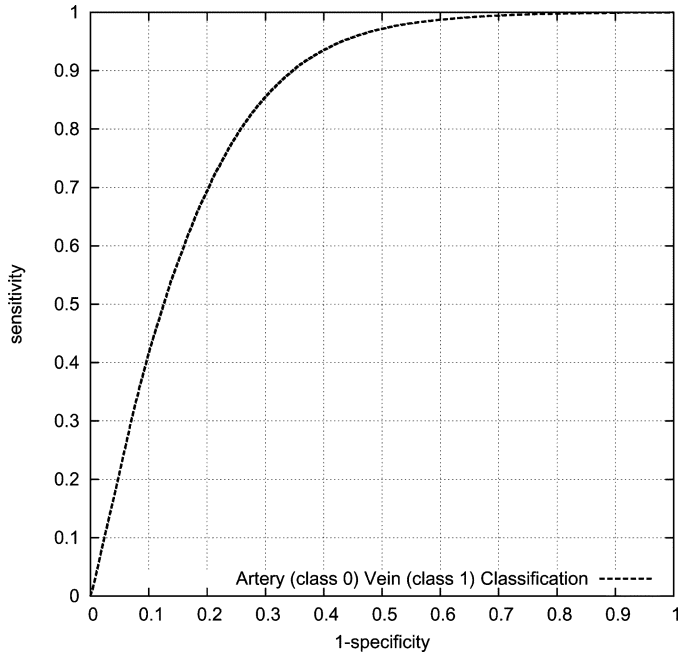

Fig. 6. ROC curve of the proposed system for assigning centerline pixels to either artery or vein class. Area under the curve is 0.84.

III. Materials

Sixty-five digital color fundus photographs were acquired for training and testing of the system. All images were obtained from patients with primary open angle glaucoma at the University of Iowa Hospitals and Clinics using a 30° Zeiss fundus camera (Carl Zeiss Meditec, Dublin, CA), with digital back (OIS systems, Sacramento, CA). The images were centered on the disc. The dimensions of the images are 2392×2048 pixels with 8-bits per pixel per color plane, and stored in JPEG format for export. To train the AV classification component and determine the parameters for the algorithm, 25 digital color fundus photographs were randomly selected from the set of 65. The remaining 40 images were assigned to the test set and were only used to evaluate the complete system.

An ophthalmologist (AVD) labeled the major vessels in the images of the training set as either artery or vein to train the artery vein classification method. As only the vessel centerlines needed to be labeled, precise vessel segmentation was not needed. Labeling was done by manually drawing a line over the major vessels using a standard painting program. The colors blue and red were used for veins and arteries respectively.

Two components of the presented method were evaluated on the test images: the AV classification and the AVR determination. To evaluate the AV classification, the vessels in each of the images in the test set were manually labeled as either artery or vein by an ophthalmologist (AVD). In contrast to the way the major arteries and veins were labeled in the training set, only those parts of all vessels (i.e., including the small vessels) that were inside the ROI were labeled in the test set. We asked the expert to label all vessels in the ROI as either an artery or a vein.

To set the AVR reference standard, a semi-automated computer program developed by the University of Wisconsin, Madison, was used (IVAN). We asked two ophthalmologists to process the images in the test set using this software. Both ophthalmologists were instructed in the use of the software using the protocols defined by the software developers. IVAN finds the optic disc and places the AVR measurement ROI on the image. It then finds vessels in the image and labels each of these as artery or vein. It also attempts to measure the vessel width. IVAN is semi-automated; the ROI localization, the vessel width measurements and the artery vein classification need manual adjustment. On average a human observer takes around 10 minutes per image to perform the analysis. The software is capable of producing several different AVR measures. We selected "Big 6" as this one corresponds to the AVR described in [15]. The ratios obtained by the first ophthalmologist (MDA) were used as the reference standard and the second ophthalmologist's (AVD) ratios were used to determine the variability between experts.

The 40 images in our test set have been made publicly available on the internet in the Iowa Normative Set for Processing Images of the REtina (INSPIRE-AVR) [28]. Included with INSPIRE-AVR are the AVR values we measured using IVAN, our reference standard.

IV. Experiments and Results

The proposed system was applied to all 40 test images. The system was able to find the optic disc and successfully placed the AVR in all 40 images. This was verified by visual inspection and the ROI was centered on the optic disc in all 40 images.

To compare the artery/vein classification with the labeling by the human expert we performed an ROC analysis with class 0 (the "negative" class) being artery and class 1 (the "positive" class) being vein. Note that this analysis was performed on the vessel centerline pixels only. The proposed classification system assigned each centerline pixel a value indicating the likelihood that it was inside of a vein. That is, a value of 0 indicates a centerline pixel likely inside an artery and a value of 1 indicates a centerline pixel likely in a vein. The ROC curve is shown in Fig. 6. The system attained an area under the ROC curve of 0.84.

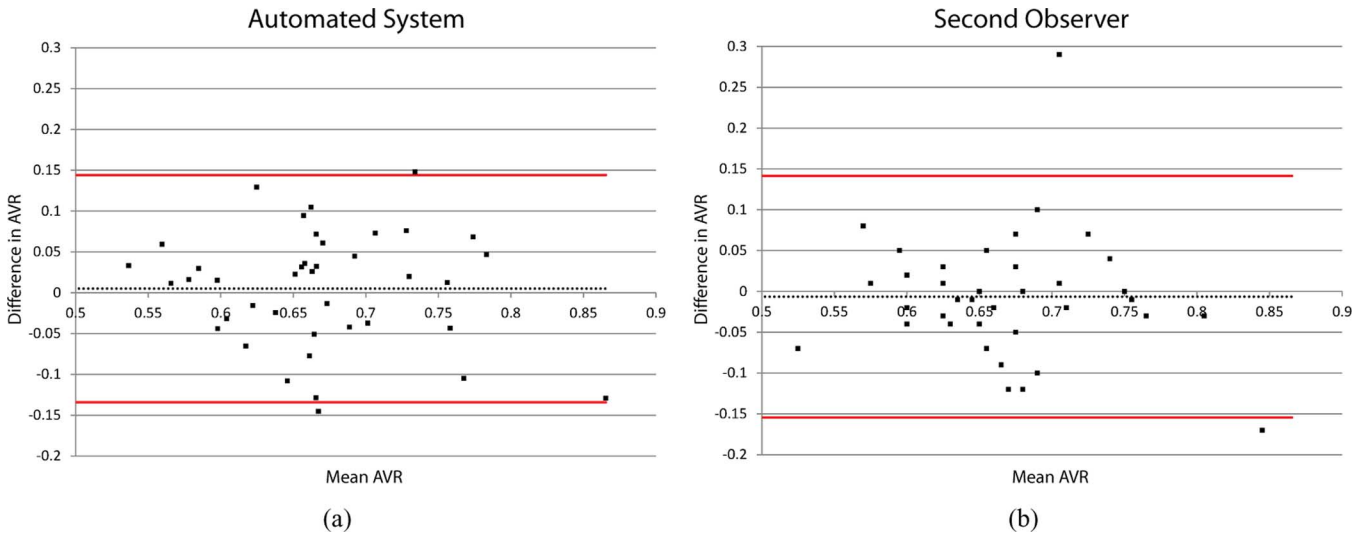

Fig. 7. Bland–Altman plots of the agreement between the automatic system and the reference standard (a) and between the second observer and the reference standard (b). The red lines represent the 95% limits of agreement. The dotted line represents the mean difference between AVR measurements.

This means that, given a randomly selected negative (artery centerline pixel) example and a randomly selected positive (vein centerline pixel) example, in 84% of cases the automatic system will correctly assign the vein centerline pixel a higher likelihood value than the artery centerline pixel.

To evaluate the ability of the system to assign an AVR value to an image we directly compared the AVRs as produced by the system and the second observer with the reference standard. The Student's paired t-test showed that there was no significant difference between the reference standard and the system's measurements ($p = 0.66$). The same holds for the second observer's measurements ($p = 0.59$). Table II shows the results for the individual images. To visually assess the agreement between both the automatic system and the second observer and the reference standard we have plotted the results in Bland–Altman plots [29] in Fig. 7. This graph plots the mean of two AVR measurements against the difference between them and allows a visual assessment of the distribution of errors and the agreement between the two methods.

V. Conclusion and Discussion

This study showed that a completely automatic method can estimate the AVR in retinal color images with a mean error similar to that of a human expert who was using the reference standard system IVAN. The automatic method also successfully classified retinal vessel pixels into being part of an artery or vein. The total running time of the algorithm implemented in C++ from initial retinal image to measured AVR value was 9 min on average, running on a single core of an Intel 2.65 GHz Core 2 Duo. The code was not optimized for speed.

Compared to previously presented methods, the area under the ROC curve for the AV classification may not seem to be an improvement. However, it is important to note that, in contrast with previously presented methods, we classified all detected individual vessel centerline pixels inside region B of the ROI. This includes vessels for which the observers were not able to see whether they were an artery or a vein without tracing the vessel back to its source. A major issue in artery/vein classification is the variability in the vessel appearance between subjects and even within the same image. Many of the previously presented methods have used clustering instead of classification to overcome the challenges presented by the high variability between subjects. We believe the use of color features that are normalized for each individual image, combined with a supervised classification approach, can prevent some of the issues with inter-subject variability a supervised method may have.

Table II shows the error with respect to the reference standard of the automated system for each image. The mean AVR values and their standard deviations of the reference standard are very close between the automated system and observer 2; a statistical test showed there was no significant difference between the means. Nevertheless, there are eight AVR measurements by the automated system that have an error above 0.10 when compared with the reference standard. However, of these, five show a relatively good agreement with the measurement done by the second observer. The Bland–Altman plots (see Fig. 7) show that both the automated system and the second observer have no substantial bias as the mean difference between the AVR measurements is close to 0. The 95% limits of agreement for both the automated and second observer are also almost the same. However, the second observer has two outliers without which the 95% limits of agreement would have been "tighter" than those of the automated system showing there is still room for improvement.

There are several limitations of the presented method and our study. The fact that we chose to use IVAN based AVR measurements to compare against our automated system limits our ability to draw conclusions about the "true" performance of the method. However, as IVAN is the de facto standard approach for measuring the AVR, used in many clinical trials and studies, we think that comparing the proposed method with (semi-automated) IVAN is a valid choice. Potentially, a better evaluation of fully automated AVR estimation would be comparing it to clinical outcome parameters. We would like to do this in the near

TABLE II
INDIVIDUAL RESULTS FOR EACH IMAGE. COLUMN "REFERENCE" CONTAINS THE REFERENCE STANDARD READING, "SYSTEM" IS THE OUTPUT FROM THE AUTOMATIC SYSTEM AND "OBS. 2" CONTAINS THE READING FROM THE SECOND OBSERVER. ALL NUMBERS IN ITALICS REPRESENT DIFFERENCES BETWEEN THE FIRST AND SECOND

Image	Reference	System	*Error*	Obs. 2	*Error*
1	0.70	0.62	*0.08*	0.71	*0.01*
2	0.76	0.81	*0.05*	0.75	*0.01*
3	0.66	0.81	*0.15*	0.69	*0.03*
4	0.75	0.76	*0.01*	0.75	*0.00*
5	0.53	0.59	*0.06*	0.61	*0.08*
6	0.93	0.80	*0.13*	0.76	*0.17*
7	0.63	0.70	*0.07*	0.68	*0.05*
8	0.70	0.59	*0.11*	0.65	*0.05*
9	0.65	0.68	*0.03*	0.64	*0.01*
10	0.78	0.74	*0.04*	0.75	*0.03*
11	0.65	0.58	*0.07*	0.65	*0.00*
12	0.67	0.74	*0.07*	0.65	*0.02*
13	0.64	0.68	*0.04*	0.71	*0.07*
14	0.69	0.77	*0.08*	0.76	*0.07*
15	0.56	0.57	*0.01*	0.85	*0.29*
16	0.64	0.70	*0.06*	0.74	*0.10*
17	0.57	0.60	*0.03*	0.62	*0.05*
18	0.62	0.59	*0.03*	0.58	*0.04*
19	0.64	0.67	*0.03*	0.61	*0.03*
20	0.68	0.67	*0.01*	0.68	*0.00*
21	0.52	0.55	*0.03*	0.45	*0.07*
22	0.62	0.58	*0.04*	0.63	*0.01*
23	0.67	0.71	*0.04*	0.63	*0.04*
24	0.71	0.67	*0.04*	0.62	*0.09*
25	0.57	0.59	*0.02*	0.58	*0.01*
26	0.72	0.74	*0.02*	0.76	*0.04*
27	0.65	0.63	*0.02*	0.64	*0.01*
28	0.56	0.69	*0.13*	0.49	*0.07*
29	0.73	0.60	*0.13*	0.61	*0.12*
30	0.64	0.66	*0.02*	0.63	*0.01*
31	0.63	0.61	*0.02*	0.68	*0.05*
32	0.72	0.68	*0.04*	0.70	*0.02*
33	0.59	0.61	*0.02*	0.61	*0.02*
34	0.61	0.71	*0.10*	0.59	*0.02*
35	0.65	0.68	*0.03*	0.61	*0.04*
36	0.74	0.59	*0.15*	0.64	*0.10*
37	0.69	0.64	*0.05*	0.62	*0.07*
38	0.82	0.72	*0.10*	0.79	*0.03*
39	0.61	0.70	*0.09*	0.64	*0.03*
40	0.74	0.81	*0.07*	0.62	*0.12*
mean	0.67	0.67	*0.06*	0.66	*0.05*
SD	0.08	0.07	*0.04*	0.08	*0.05*
min	0.52	0.55	*0.01*	0.45	*0.00*
max	0.93	0.81	*0.15*	0.85	*0.29*

future. A second limitation is that our approach was developed for, and was evaluated on, high resolution retinal fundus images which are typically obtained in a clinical setting. We would like to further improve the system so it can be used on images obtained by nonmydriatic fundus cameras, which typically have lower resolution and contrast. A final limitation is the fact we assumed a fixed size for the optic disc. We expect the use of different data may require retraining of the artery vein classification method, also, an optic disc segmentation procedure would have to be added to deal with varying optic disc sizes but the basic procedure will remain the same.

The AVR is a known, independent, risk factor for many systemic diseases, including cerebral atrophy, cognitive decline, stroke and cardiovascular disease, as well as a metric for retinal diseases, including diabetic retinopathy, hypertensive retinopathy and retinopathy of prematurity. However, the AVR is not in clinical use because it is so cumbersome to obtain. Our approach obtains the AVR from a high quality retinal image centered on the disc, and is more cost-effective than manual estimation, with the only human intervention consisting of taking the retinal photographs. This approach has the potential to have a major impact on the early detection and treatment of common neurological, cardiovascular and retinal diseases.

In summary, we have successfully developed a fully automated method that is capable of estimating the AVR in retinal color images with a small error that is similar to the error of a human expert.

ACKNOWLEDGMENT

The authors would like to acknowledge the University of Wisconsin–Madison Fundus Photograph Reading Center, Department of Ophthalmology and Visual Sciences, for providing the IVAN software package used in this research.

REFERENCES

[1] T. Y. Wong, A. Shankar, R. Klein, B. E. K. Klein, and L. D. Hubbard, "Prospective cohort study of retinal vessel diameters and risk of hypertension," *Br. Med. J.*, vol. 329, pp. 799–800, 2004.

[2] L. D. Hubbard, R. J. Brothers, W. N. King, L. X. Clegg, R. Klein, L. S. Cooper, A. R. Sharrett, M. D. Davis, and J. Cai, "Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the atherosclerosis risk in communities study," *Ophthalmology*, vol. 106, pp. 2269–2280, 1999.

[3] S. R. Lesage, T. H. Mosley, T. Y. Wong, M. Szklo, D. Knopman, D. J. Catellier, S. R. Cole, R. Klein, J. Coresh, L. H. Coker, and A. R. Sharrett, "Retinal microvascular abnormalities and cognitive decline: The ARIC 14-year follow-up study," *Neurology*, vol. 73, no. 11, pp. 862–868, 2009.

[4] C. Sun, J. J. Wang, D. A. Mackey, and T. Y. Wong, "Retinal vascular caliber: Systemic, environmental, and genetic associations," *Survey Ophthalmol.*, vol. 54, no. 1, pp. 74–95, 2009.

[5] M. E. Martínez-Pérez, A. D. Hughes, A. V. Stanton, S. A. Thom, N. Chapman, A. A. Bharath, and K. H. Parker, "Retinal vascular tree morphology: A semi-automatic quantification," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 8, pp. 912–917, Aug. 2002.

[6] K. Rothaus, P. Rhiem, and X. Jiang, "Separation of the retinal vascular graph in arteries and veins," in *LNCS*, 2007, vol. 4538, pp. 251–262.

[7] H. Li, W. Hsu, M. L. Lee, and T. Y. Wong, "Automatic grading of retinal vessel caliber," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 7, pp. 1352–1355, Jul. 2005.

[8] E. Grisan and A. Ruggeri, "A divide et impera strategy for automatic classification of retinal vessels into arteries and veins," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2003, pp. 890–893.

[9] C. Kondermann, D. Kondermann, and M. Yan, "Blood vessel classification into arteries and veins in retinal images," in *Med. Imag. 2007: Image Process.*, San Diego, CA, 2007, vol. 6512, pp. 651247–651249.

[10] S. G. Vázquez, N. Barreira, M. G. Penedo, M. Ortega, and A. Pose-Reino, "Improvements in retinal vessel clustering techniques: Towards the automatic computation of the arterio venous ratio," *Computing*, vol. 90, no. 3–4, pp. 197–217, 2010.

[11] A. Ruggeri, E. Grisan, and M. De Luca, "An automatic system for the estimation of generalized arteriolar narrowing in retinal images," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2007, pp. 6463–6466.

[12] L. Tramontan, E. Grisan, and A. Ruggeri, "An improved system for the automatic estimation of the arteriolar-to-Venular diameter ratio (AVR) in retinal images," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2008, pp. 3550–3553.

[13] L. Tramontan, E. Poletti, D. Fiorin, and A. Ruggeri, "A web-based system for the quantitative and reproducible assessment of clinical indexes from the retinal vasculature," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 818–821, Mar. 2011.

[14] L. D. Hubbard, R. J. Brothers, W. N. King, L. X. Clegg, R. Klein, L. S. Cooper, A. R. Sharrett, M. D. Davis, and J. Cai, "Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the atherosclerosis risk in communities study," *Ophthalmology*, vol. 106, no. 12, pp. 2269–2280, 1999.

[15] M. D. Knudtson, K. E. Lee, L. D. Hubbard, T. Y. Wong, R. Klein, and B. E. K. Klein, "Revised formulas for summarizing retinal vessel diameters," *Curr. Eye Res*, vol. 27, no. 3, pp. 143–149, 2003.

[16] M. Niemeijer, M. D. Abràmoff, and B. van Ginneken, "Information fusion for diabetic retinopathy cad in digital color fundus photographs," *IEEE Trans. Med. Imag.*, vol. 28, no. 5, pp. 775–785, May 2009.

[17] M. Niemeijer, J. J. Staal, B. van Ginneken, M. Loog, and M. D. Abràmoff, "Comparative study of retinal vessel segmentation methods on a new publicly available database," in *Proc. SPIE: Med. Imag.*, 2004, vol. 5370, pp. 648–656.

[18] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching in fixed dimensions," *J. ACM*, vol. 45, no. 6, pp. 891–923, 1998.

[19] J. Staal, M. Abràmoff, M. Niemeijer, M. Viergever, and B. van Ginneken, "Ridge based vessel segmentation in color image of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.

[20] M. Ahmed and R. Ward, "A rotation invariant rule-based thinning algorithm for character recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 12, pp. 1672–1678, Dec. 2002.

[21] P. I. Rockett, "An improved rotation-invariant thinning algorithm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1671–1674, Oct. 2005.

[22] M. Niemeijer, M. D. Abràmoff, and B. van Ginneken, "Fast detection of the optic disc and fovea in color fundus photographs," *Med. Image Anal.*, vol. 13, no. 6, pp. 859–870, 2009.

[23] J. Fairfield, "Toboggan contrast enhancement for contrast segmentation," in *Proc. Int. Conf. Pattern Recognit.*, 1990, pp. 712–716.

[24] X. Yao and Y. Hung, "Fast image segmentation by sliding in the derivative terrain," in *Proc. SPIE Conf. Intell. Robots Computer Vision X: Algorithms Techn.*, 1991, pp. 369–379.

[25] E. Mortensen, "Simultaneous multi-frame subpixel boundary definition using Toboggan-based intelligent scissors for image and movie editing," Ph.D. dissertation, Brigham Young Univ., Provo, UT, 2000.

[26] M. Niemeijer, B. van Ginneken, and M. Abràmoff, "Automatic classification of retinal vessels into arteries and veins," in *Medical Imaging 2009: Computer-Aided Diagnosis*, 2009, vol. 7260, pp. 72601F–72601F.

[27] C. E. Metz, "ROC methodology in radiologic imaging," *Investigative Radiol.*, vol. 21, no. 9, pp. 720–733, 1986.

[28] INSPIRE-AVR: Iowa normative set for processing images of the retina-artery vein ratio [Online]. Available: http://webeye.ophth.uiowa.edu/component/k2/item/270

[29] J. M. Bland and D. G. Altman, "Measuring agreement in method comparison studies," *Stat. Methods Med. Res.*, vol. 8, no. 2, pp. 135–160, 1999.