MULTI-SCALE SEGMENTATION OF THE VASCULAR TREES IN RETINAL IMAGES

Kai Rothaus* and Xiaoji Jiang*

* Department of Computer Science, University of Münster, Germany

{rothaus,xjiang}@uni-muenster.de

Abstract: In this paper, we propose a multi-scale midline extraction method for reconstructing the vascular trees out of retinal images. This approach offers the advantage to handle vessels of different widths simultaneously. On the basis of the used segment representation of the vessels, we compute geometrical features of the vessel segments. Comparing these features with characteristics of vessels the false positive rate is reduced and gaps in the extracted midlines are closed. We give some example results of our approach on different retinal image modes.

Introduction

In this paper we address the problem of segmenting the vascular trees in retinal images. In its general form this problem can be stated as one of extracting a continuous midline of elongated objects of various widths. There are various other applications of midline extraction in the field of image analysis, for instance analysis of heart fibres [1], description of characters, plant root detection [2], analysis of aerial images [3] and finger print analysis.

Considering the graph G of an d dimensional grey-scale image function $L(\xi)$ (with $\xi = (\xi_1, \ldots \xi_d)$) as a hyper-surface (in an (d+1)-dimensional space), one can define ridge and valley points using local characteristics of this hyper-surface. More precisely, we analyse the image gradient ∇L on the boundary of a small neighbourhood in G surrounding the point $(\xi, L(\xi))$. This can be done by choosing the eigenvectors of the Hessian matrix of L as the local coordinate system. Maintz et al. [4] formulate a class of ridge measure based on the isophote curvature, which are the curves on G featuring the same level of grey-scale. Since these measures show a poor performance at saddle zones in discrete domains [5], López et al. have developed a different discretisation, namely $\kappa = -\mathrm{div}(\nabla L/\|\nabla L\|)$. Their measure is the fundament of our method.

Our method is based on a scale-space analysis. Such a multi-scale approach offers the advantage to process images or other objects with various parameter adjustments simultaneously. This leads to different results, which have to be finally combined to a single solution. Therefore, the most challenging task of multi-scale approaches is a consistent combination scheme. Compared to other approaches, the proposed operator minimises the effect of erasing the inferior structure at crossings, bifurcations and neighbouring objects. Furthermore, we directly provide the tangential direction of the midline at the sample points.

We have found other works on multi-scale operators for midline extraction [3, 6, 7]. The commonness of these approaches is the design of the scale-space, which is built directly on the input image L. Typically, these approaches result in fragmentary midlines, due to the fact that we have a loss of spatial precision, when smoothing in the image space. In contrast, we are smoothing the gradient tensor field $\nabla L \cdot \nabla L^{\mathrm{t}}$ component-wise using different Gaussian convolution kernels. This offers the advantage of a more robust localization of the midlines due to the fact that even for thin objects, the gradients at the borders are present in large scales. In the case of applying large Gaussian masks directly on L, thin objects are razed. Furthermore, combining midlines of different image scales could lead to points of discontinuity, when there is an erratic change in the width of the object. With our method this problem is solved, since we firstly combine the vector fields of different scales and then extract the midlines.

This paper is organised as follows. At first we present the scale-space approach for midline extraction. The use of a scale-space offers the advantage to extract thin and wide vessels at the same time. In contrast to earlier approaches the scale-space is not directly computed on the intensity image in our work. Instead, we firstly compute the gradient vector field and use different smoothing technique for vector fields. For that purpose different convolution kernels are used and for each image location exactly one vector is chosen out of the computed scales. The result of this midline extraction process is a set of midline segments, which are organised to build the vascular trees. Finally, we present some results of our approach and conclude with the discussion and give an outlook for future works.

Scale-space approach

In this section we explain the midline extraction procedure, which is the heart of our algorithm. This procedure can be divided into the subtasks: computation of the structural tensor, iterative construction of the scale-space, application of the divergence operator on the combined vector fieldand finally the extraction of the midlines. At

ISSN: 1727-1983

the end of this section we discuss the choice of the parameters.

Since we use second order derivation to compute κ, we need an idealised gradient vector field. The cross-profile of the vessels are frequently bar-like and the gradients could be very small or even vanish near the midline. Therefore, we wish to propagate the meaningful gradients at the edges towards the midline of the vessels. Smoothing the image has known side-effects, which we want to avoid.

The main idea is to construct the scale-space not on the image, but on the image gradient $\nabla L\,(x,y)$. We present a method for smoothing the vector field such that the meaningful gradients at the edges are propagated towards the midlines. The computation of the image gradient in a specific scale is realised by an analysis of a component-wise smoothed structured tensor, which is computed as $\nabla L \cdot \nabla L^{\mathrm{t}}$. For each image location we compute the local structure tensor and convolve the components of this matrices along the image plane. During the iteration we keep up the vector field, which contains the currently best gradient vectors.

Structural tensor

Initially, we compute the image gradient vector field at the pixel grid with the Sobel operator. The used Gaussian mask should be chosen relatively small, since edges introduced by noise are automatically erased in the further process. The edge magnitude is normalised into the interval $[0,1]$. To compensate local differences in contrast we boost the edge magnitude s by the function $b_{t_{\mathrm{infl}}}(s) = 1 - \exp\left(-\frac{s^2}{2 \cdot t_{\mathrm{infl}}^2}\right)$. Let t_{conf} denote the value of $b_{t_{\mathrm{infl}}}$ at its inflection point t_{infl}, which is an external parameter. Gradients holding a magnitude of greater than t_{conf} are considered as confident and will be treated more carefully.

Following the approach of López et al. [5] we apply the component-wise smoothing of the structured tensor and compute the gradient at each image position as the eigenvector to the greatest eigenvector of the smoothed structured tensor field. The motivation of this approach is that the structured tensor is a symmetrical matrix of the form:

$$ST(x,y) = \begin{pmatrix} L_x(x,y)^2 & L_x(x,y) \cdot L_y(x,y) \\ L_x(x,y) \cdot L_y(x,y) & L_y(x,y)^2 \end{pmatrix},$$

with the eigenvector ∇L to the eigenvalue

$$\lambda = s^2 = L_x^2 + L_y^2. \qquad (1)$$

We utilise this property to compute the gradient magnitudes of the different scales.

Forming the scale-space

After computing $ST(x,y)$ for each grid position we construct the scale-space on this structural tensor. For each scale we compute the gradient vector field, which consists of the eigenvector $g_k(x,y)$ to the greatest eigenvector $\lambda_k(x,y)$ of the component-wise smoothed version of $ST(x,y)$. Depending on λ_k we compute the new gradient magnitude. Since the sign of $g_k(x,y)$ is not clear, we have to reconstruct it by some additional considerations.

During the iterative construction of the scale-space we keep $v_{\mathrm{opt}}(x,y)$, the optimal gradient vector field computed so far, with $s_{\mathrm{opt}} := \|v_{\mathrm{opt}}\|$ denoting its magnitude. The scale-space is built by component-wise convolution of the structural tensor repeatedly with the same Gaussian mask G_σ with standard deviation σ. Due to the fact, that the family of Gaussian distributions build a half-group under the convolution operator $\star$, the standard deviation during the k-th iteration is $\sigma_k = \sigma\sqrt{k}$, i.e. $(G_\sigma)^{\star k} = G_{\sigma_k}$.

In the following we explain one iteration step k at a fix grid point (x_0,y_0). Smoothing the structural tensor $ST(x,y)$ results in a matrix of the form

$$\begin{pmatrix} m_{xx} & m_{xy} \\ m_{xy} & m_{yy} \end{pmatrix} := \left((G_\sigma)^{\star k} \star \begin{pmatrix} L_x^2 & L_x L_y \\ L_x L_y & L_y^2 \end{pmatrix} \right)(x_0,y_0)$$

with greatest eigenvalue $\lambda_k = 0.5 \cdot \left(s_k^2 + p_k\right)$, where $s_k^2 = m_{xx} + m_{yy}$ is equal to the square of smoothed vector magnitudes and $p_k = \sqrt{(m_{xx} - m_{yy})^2 + 4\,m_{xy}^2}$ can be viewed as a measure of parallelism. If the considered neighbourhood surround (x_0,y_0) consists of parallel gradients, p_k takes the maximum value s_k^2. On the other hand $p_k = 0$ at locations where two orthogonal gradient directions occur (equally weighted). As mentioned above, the gradient vector v_k is computed as the eigenvector to the eigenvalue λ_k. The new magnitude is determined by taking two numbers into account, namely λ_k and the quotient $\frac{p_k}{s_k^2} = \frac{\lambda_k - \tilde{\lambda}_k}{\lambda_k + \tilde{\lambda}_k} \in [0,1]$, with $\tilde{\lambda}_k$ denoting the smaller eigenvalue of $ST(x,y)$. Taking the eigenvalue as magnitude is a consistent extension of the property (1). On the other hand the term $\frac{p_k}{s_k^2}$ is a measure of parallelism and furthermore independent of the edge magnitudes in the neighbourhood. We choose the geometrical mean of these values as the new squared edge magnitude

$$s_{\mathrm{new}}^2 = \sqrt{0.5 \cdot \left(s_k^2 + p_k\right) \cdot \frac{p_k}{s_k^2}}. = \sqrt{\lambda_k \cdot \frac{\lambda_k - \tilde{\lambda}_k}{\lambda_k + \tilde{\lambda}_k}} \qquad (2)$$

In image regions with no preferred gradient orientations the magnitude are pruned by this combination, due to the fact that p_k is low. On the other hand the gradients are boosted in image regions with a clear major orientation even when there is a low base level of gradient magnitudes.

If $s_{\mathrm{new}} > s_{\mathrm{opt}}$ we further process in updating the actual vector v_{opt}, otherwise we leave v_{opt} unchanged. An open problem not discussed yet is the unknown sign of the eigenvector v_k. López et al. [5] propose an alignment into the half-space to which the original gradient points. This could lead to opposed islands of gradients at regions with relative homogeneous grey values. Another side effect is a possible erasement of the gradients introduced by small objects lying in the neighbourhood of dominant

ISSN: 1727-1983

structures. Due to this unwanted behaviour we make a consensus decision based on the image gradient g_k introduced by the smoothed version of the original image $G_{\sigma\sqrt{k}} \star L$ and the actual vector v_{opt}. With respect to the iterative smoothing process and equation (2) we compute two thresholds (in case of $k > 1$):

$$t_{\text{low}} := \sigma_k^{-1} \cdot \sqrt{0.5 \cdot t_{\text{conf}}} \qquad \text{and} \qquad t_{\text{high}} := \sqrt{0.5 \cdot t_{\text{conf}}}$$

Let cos_{grad} denote the cosine of the angle between v_k and g_k, cos_k denote the cosine of the angle between v_k and v_{opt}. The decision rule for the sign of v_{k} is based on

$$sgn := \begin{cases} cos_{\text{grad}} & \text{if } \lambda_k < t_{\text{low}} \vee s_{\text{opt}} < t_{\text{low}} \\ & \vee\, k = 1 \\ cos_k & \text{if } \lambda_k \geq t_{\text{high}} \wedge s_{\text{opt}} \geq t_{\text{low}} \\ (\lambda_k - t_{\text{low}}) \cdot cos_{\text{grad}} & \text{otherwise.} \\ +\left(t_{\text{high}} - \lambda_k\right) \cdot cos_k & \end{cases}$$

If *sgn* is negative the orientation of v_k is reversed by multiplying v_k with -1. This case differentiation is motivated by the consideration that we want to keep informations of consistent gradients. For small values of s_{opt} or λ_k the orientation of g_{opt} is possibly uncertain (first case), so that the decision on the sign should only depend on the smoothed gradient g_k. On the other side, when λ_k and s_{opt} are sufficiently great, we take only the orientation of the currently optimal gradient g_{opt} into account (second case). The third case is a fuzzy-like compromise between the exclusive decisions. Finally, we replace v_{opt} by v_k with the magnitude s_{new}.

Computing the creaseness measure κ

Since the combination of the different scales could introduce boundary effects, we apply the same smoothing procedure on the resulted gradient vector field, but only with a single iteration. The divergence operator computes κ under the constraint that the norm of each vector is 0 or 1. A normalisation of all gradient vectors, which have a positive norm, causes unwanted effects at homogeneous image regions, since the vectors at the propagation hold no confident orientation. To avoid this problem we apply the boosting function $b_{0.33}$ on the vector magnitudes instead of the normalisation.

The divergence operator is based on deviations. Since the vector field has been smoothed already, we implement deviations by the mean of finite forward difference and finite backwards difference with an increment of one pixel. The result is a scalar field, which holds the values $-\kappa(x,y)$.

An additional smoothing of the scalar field κ leads to the extraction of less erratic midlines, but could cause misbehaviour in regions with undefined gradient vectors. In that case we produce vectors, which have no direction, but a positive norm. Due to this problem the smoothing of κ should be used with care. We choose triangular convolution kernels of size three in both image dimensions. This kernel reduces the occurrence of unwanted side-effects, but achieves continuous midlines. The remaining vectors are denoted as the idealised gradient vectors.

Extraction of the midlines

After extracting the pixels holding maximal creaseness using a non maximum suppression strategy, we arrange those which are forming segments. We connect two neighboured pixels, if several conditions are true. Firstly, the difference in creaseness is not allowed to exceed a threshold. The second important constraint is about their gradient direction. These directions correspond to the local orientation of the slope lines, which we consider to be orthogonal on the midline. With this assumption the relative position of two neighboured midline pixel is validated. As a result each pixel has at most two connected neighbours, so that they form chains with exactly two end points. Since on the pixel grid diagonal-wise connected pixel have distance of $\sqrt{2}$ pixels, we resample each segment with an increment of 1 pixel. The resampling of a segment *seg* is done by traversing it from its starting point and keeping the distance, which is covered so far. We compute the tangential direction of *seg* for all locations with an integral distance to the starting point. This can be done by a linear interpolation of the orthogonal vectors on the smoothed gradient direction at the nearest pixel locations on the segment. With this representation of a midline segment we are able to compute various features for each segment (length, local creaseness, average creaseness, turning points, intensity integration on the original image). These features could be used as a basis for decision if a segment should be selected or not.

Parameter adjustment

The presented method can be adjusted by four numerical parameters, namely σ_{Sobel}, t_{infl}, σ_{iter} and n_{iter}. The influence of the standard deviation σ_{Sobel} of the Gaussian mask for the construction of the gradient vector field has been discussed in various works. The parameter t_{infl} offers the possibility to control the inflection point of the boosting function $\text{b}(s)$. A low value results in the consideration of low contrast edges of the input image. The choice of the standard deviation σ_{iter} defines the smallest structures which are preserved during the iteration. Since the Gaussian of the iteration k has a standard deviation $\sigma_{\text{iter}}\sqrt{k}$, the decreasing of σ_{iter} causes more iterations to achieve the same blurring effect. Obviously, this correlation behaves quadratically. The parameter n_{iter} defines the maximal width of the elongated objects, for which the midlines are correctly computed. The increase of n_{iter} has primarily effects on regions with an actual low response, but there could occur unwanted side-effects. It is not possible to avoid these side-effects generally, due to semantical ambiguity. For example, the dominant structure of a horizontal line, which is an arrangement of vertical dark bars, depends on the context. This gap cannot be closed by a multi-scale approach. Experiments have shown, that for retinal images $\sigma_{\text{Sobel}} = 1$, $t_{\text{infl}} = 0.2$, $\sigma_{\text{iter}} = 1$ and $n_{\text{iter}} = 10$ are suitable values. The manual adjustment of these parameter is easy and mainly depends on the image scale.

ISSN: 1727-1983

Extracting the vascular trees

In the last section we have described the approach to extracting midlines of dark elongated objects. Beside the vessel midlines we have a lot of midlines which are produced by non-vessel objects (see Figure 2). In the following we explain how to filter these unwanted detected midlines and keep most of the vessel midlines.

Vessels are objects with are relative straight and connected to each other. For this reason segments with length less than eight pixels are taken out of consideration. Furthermore, a limited difference of the midline directions at two neighboured sample points is allowed. The occurrence of a rapid change of the midline direction is an indicator for a sharp turn. Since vessels normally are straight objects a sharp turn indicates the presence of a non-vessel object or the tracking towards the wrong vessel segment. For this reason a segment is divided in two parts at sample points, where the directions differ more than 13.5°. This corresponds to a restriction of the midline segment creaseness.

On all remaining segments we apply the connection procedure The open problem of handling bifurcations is discussed afterwards. Finally, we describe the influence of the thresholds and parameter used in the following procedure.

Connecting midline segments

The presence of noise could separate a vessel in disconnected segments. Therefore, a reconnecting algorithm is essential for extracting vascular trees. Small gaps primarily appear at vessel crossings, near pathological objects, in low-contrast regions and in noisy regions, respectively. To close these gaps, we define a distance function on the segment endpoints.

The Euclidian distance is insufficient for this purpose, since it does not take the course of the midlines into account. To overcome this problem we firstly compute the direction of the segments at both of their endpoints. This can be done by taken the mean direction accumulated at a fixed number of the last sample points of the segment. We use 10 samples to compute the ending directions and run the connection process. Thereby, some segments are merged and so the total number of segments decreases. Then we repeat the same operation on the remaining segments using 15, 20 and finally 5 samples for direction computation. The segments have typically highly curved short end pieces, independent of the true course of the segments. This may bias the direction estimation. For this reason we use 5 samples in the last processing cycle.

In the following we define the distance function on an endpoint pair p_1 and p_2 (see Figure 1). Let seg_1 and seg_2 denote the corresponding segments and α_i is the angle between the extended segment (towards the computed ending direction) and the line between p_1 and p_2. If both angles α_1 and α_2 are less than 40° and the L_1-distance between p_1 and p_2 does not exceed 30 pixel we compute the B-Spline S_{p_1,p_2} connecting the segments. Thereby, the

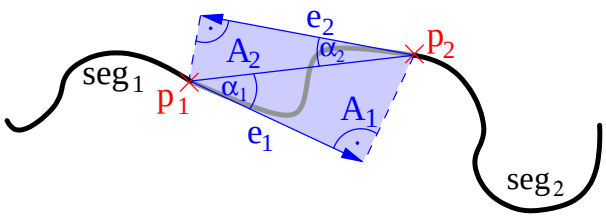

Figure 1: Connection of two segments

control points are chosen carefully on the elongation lines e_1 and e_2. The distance between the two endpoints is a combination of

(1) the length l_{spline} of S_{p_1,p_2}
(2) the lengths l_1 and l_2 of seg_1 and seg_2 and
(3) the mean area A of the triangles A_1, A_2 (see Figure 1).

With these values we experimentally derive the distance function:

$$D_{\text{ep}}(p_1,p_2) = \sqrt{l_1^{-1}+l_2^{-1}} \cdot \left(0.75 \cdot \sqrt{A} + 0.25 \cdot l_{\text{spline}}\right).$$

Let EP denote the set of all unconnected endpoints. In the following the connection process based on EP and the D_{ep} is depicted. At first the distance matrix on EP is computed. We connect two endpoints p_1 and p_2 if they are pair wise their nearest neighbours subject to some conditions discussed below. Thereby, the connection segment is chosen as the already computed B-Spline S_{p_1,p_2} and afterwards p_1 and p_2 are eliminated from EP. The two segments belonging to p_1 and p_2 and the connection segment together form a new segment. This procedure is repeated until EP becomes stable, meaning that there is no pair of unconnected endpoints which is a nearest neighbour pair and satisfies all the conditions.

To avoid a connection of relative short segments with relative great distance, some constraints are checked.

Firstly, the length of the connecting segments each segment should consist of at most $1/3$ of connection segments. Secondly, two segments are connected only if the length of the connection segment between them is shorter than twice the minimum of the single segment lengths. Finally, all remaining segments with length less than 16 pixels are deleted.

Handling of bifurcations

During the connection process the presence of vessel bifurcations implies a special handling. At bifurcations there exists normally one dominated vessel and a smaller one. As a result the midline segment s_1 corresponding to the smaller vessel has one endpoint p near the midline segment s_2 of the dominated vessel. Since there are no correct counterparts for p, the segment s_1 should not be extended at p. To detect such situation we are still investigating a bifurcation filter, marking possible bifurcation points and allow connections between them and near endpoints.

ISSN: 1727-1983

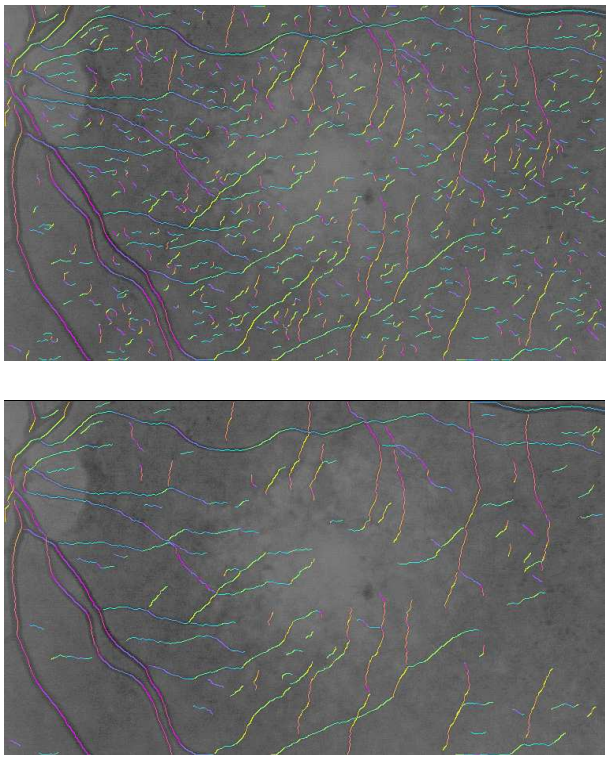

Figure 2: Result on retinal angiography without (top) and with (bottom) connecting segments

Parameter Adjustment

We use several experiment based thresholds and weights for the filtering and the connection process. Some of them depend on the image scale since for instance the minimal length of segments or the maximal creaseness of the midlines. These scale-dependent parameters could be easily modulated to other image material. Other parameters should be varied carefully or even kept fixed. These are the weights for the combination rule defining the distance function D_{EP} and the pruning thresholds to avoid segment chains, which contain a high fraction of B-Spline segments. We are investigating an optimisation of the parameters on objective criteria.

Results

We have tested our method on angiographies (Figure 2) and retinal images (Figure 3). Thereby, we use adequate pre-processing techniques for enhancing the contrast of the images. Since there are retinal images of various quality, we have not investigated a unique pre-processing method, which are able to handle all those different retinal images. In Figures 2 and 3 the midlines are superimposed on a grey-scale version of the original. The colour of the midline pixels specifies the direction of the midline at this position. Thereby, turquoise correspondes to vertical and red correspondes to horizontal directions.

It should be possible to enhance our results by an improved pre-processing method. The results depends somewhat on the parameters for the multi-scale midline extraction,

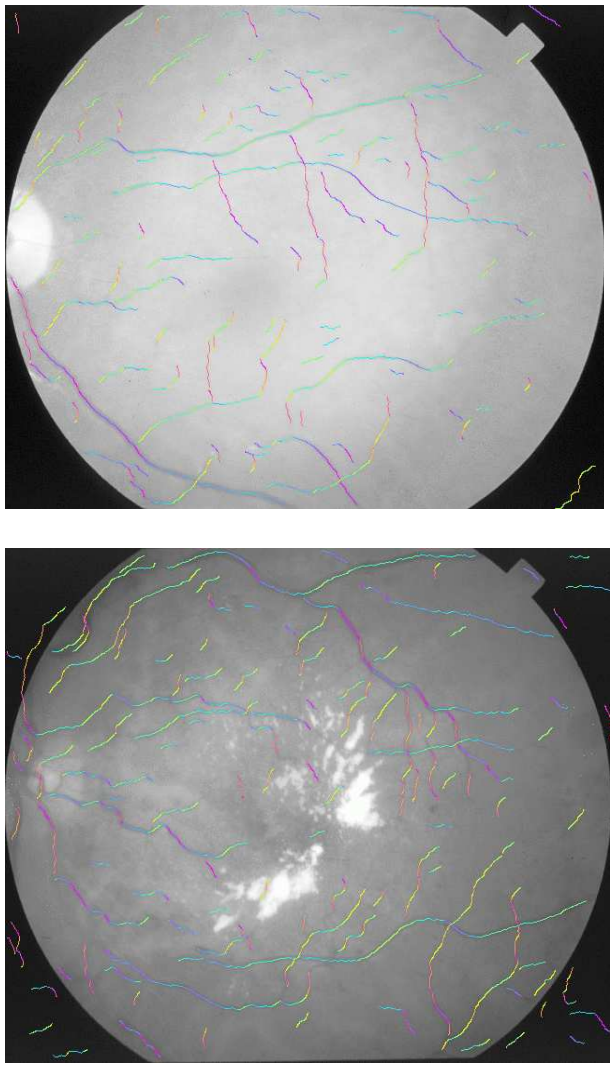

Figure 3: Result on retinal images [8]: normal (top) and pathological eye (bottom)

By comparing the midlines which are extracted using the scale-space approach (Figure 2 top row) with the remaining midlines after applying the vessel combination scheme (Figure 2 bottom row), it is evident that the false positive rate is reduced.

We have not tested the accuracy of our method on the ground truth data base of Hoover [8]. A visual inspection of the results affirmed that we have reached our purpose to extract vessels of various widths simultaneously. To quantify these observation we are working on an extension of our algorithm, which computes the vessel borders. The result of this vessel segmentation could be compared with ground truth data [8].

Discussions and conclusion

In this paper, we have proposed a multi-scale approach for vessel extraction. Our method enables the pos-

ISSN: 1727-1983

sibility to extract the structure of the vascular tree, with only little adjustments of parameter concerning the general image quality. We are able to reconstruct parts of the vessel trees, even though a special handling of vessel bifurcation is not yet taken into consideration. By using characteristics of vessels, the false positive rate could be visibly decreased. Furthermore, the extracted midline segments can be used for the computation of vessel features like the mean curvature the length and the angles at crossings, respectively. An extension of the perceptual grouping of the vessel segments for handling bifurcations should be possible. After this extension the vascular trees could be extracted completely.

Acknowledgment

The authors would like to thank A. Hoover and B. Török for the image materials.

References

[1] LUNKENHEIMER P.P., REDMANN K., KLING N., ROTHAUS K., JIANG X., CRYER C.W., WÜBBELING F., NIEDERER P., HO S.Y., and ANDERSON R.H. The three-dimensional architecture of the left ventricular myocardium. *submitted for publication*, 2005.

[2] ERZ G. and POSCH S. A region based seed detection for root detection in minirhizotron images. In Bernd Michaelis and Gerald Krell, editors, *Pattern Recognition, 25th DAGM Symposium*, volume 2781 of *Lecture Notes in Computer Science*, pages 482–489. Springer, 2003.

[3] C. STEGER. An unbiased detector of curvilinear structures. *IEEE Trans. on PAMI*, 20(2):113–125, February 1998.

[4] J.B.A MAINTZ, P.A. VAN DEN ELSEN, and M.A. VIERGEVER. Evaluation of ridge seeking operators for multimodality medical image matching. *IEEE Trans. on PAMI*, 18(4):353–365, April 1996.

[5] A.M. LÓPEZ, F. LUMBRERAS, and J. SERRAT. Creaseness from level set extrinsic curvature. In H. Burkhardt and B. Neumann, editors, *Proc. of the 5th ECCV*, pages 156–169, 1998.

[6] J.M. GAUCH and S.M. PIZER. Multiresolution analysis of ridges and valleys in grey-scale images. *IEEE Trans. on PAMI*, 15(6):636–646, June 1993.

[7] T.M. KOLLER, G. GERIG, G. SZEKELY, and D. DETTWILER. Multiscale detection of curvilinear structures in 2-d and 3-d image data. In *Proc. of the 5th ICCV*, pages 864–869, June 1995.

[8] A. HOOVER, V. KOUZNETSOVA, and M. GOLDBAUM. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Trans. on Medical Imaging*, 19(4):203–210, March 2000.

ISSN: 1727-1983