CLASSIFYING CONVEX SETS FOR VESSEL DETECTION IN RETINAL IMAGES

Joes Staal, Stiliyan N. Kalitzin, Michael D. Abrámoff, Tos Berendschot,
Bram van Ginneken and Max A. Viergever

Image Sciences Institute
Heidelberglaan 100
3584 CX Utrecht
The Netherlands
joes@isi.uu.nl

ABSTRACT

We present a method to detect vessels in images of the retina. Instead of relying on pixel classification, as many detection algorithms do, we propose a more natural representation for elongated structures, such as vessels. This new representation consists of primitives called affine convex sets. On these convex sets we apply the classification step.

The reason for choosing this approach is two-fold:
(1) By using a dedicated representation of image structures, one can exploit prior knowledge.
(2) A method based on pixel classification is often computationally unattractive.

The method can also be applied to other image structures, if an appropriate representation for the structures is chosen.

The method was tested on fundus reflection images. We obtained an accuracy of 0.897, a sensitivity of 0.700 and a specificity of 0.923.

1. INTRODUCTION

The appearance of vessels in ocular fundus images plays an important role in the diagnosis of diabetic retinopathy. Based on measurements of features of the vessel characteristics diagnosis, treatment and clinical study can be performed. It is important for eye care specialists to be able to screen patients in a reproducible and reliable manner. Automated systems for delineating the vasculature are therefore of great value.

Most methods previous work is based upon pixel classification [1, 2] or vessel tracking [3, 4]. Our method is based on classification of convex sets, special image primitives for elongated structures like vessels. In section 2 we show how to form convex sets from ridges. Section 3 explains how to use the convex sets for classification. We show the results of our method in section 4 and we conclude in section 5.

2. DETECTION OF ELONGATED STRUCTURES

2.1. Ridge detection

If a two-dimensional gray-value image is regarded as a height map, with the luminance as the height parameter, we can delineate different topological points and point-sets, such as maxima, minima, saddle-points, edges, ridges and valleys. Since we are interested in finding the location of vessels in two-dimensional images, we show how to detect ridges and valleys in that dimension. For topological point-sets in higher dimensional images, see [5].

We define ridges and valleys as those points, where the image has an extremum in the direction of the largest surface curvature. Mathematically, we search for those points in the image $L(\mathbf{x})$, with $\mathbf{x} = (x_1, x_2)^T$, where the first derivative of the luminance in the direction of the largest surface curvature changes sign.

The direction of largest surface curvature is the eigenvector $\hat{\mathbf{v}}$ of the matrix of second order derivatives of the image which has the largest absolute eigenvalue λ. This matrix is often referred to as the Hessian matrix. The first derivative of the image in the direction of $\hat{\mathbf{v}}$ is found by projecting the gradient of the image on it. The sign of λ determines whether we have found a valley ($\lambda > 0$) or a ridge ($\lambda < 0$).

Because taking derivatives of discrete images is an ill-posed operation, we take derivatives at a scale t using the Gaussian scale-space technique (see e.g. [6] and references therein). The main idea is that the image derivatives can be taken by convolving the image with derivatives of a Gaussian:

$$\frac{\partial^i L(\mathbf{x},t)}{\partial {x_j}^i} = \frac{1}{2\pi t}\int_{\mathbf{x}\in\mathbb{R}^2} \frac{\partial^i \mathrm{e}^{-\|\mathbf{x}-\mathbf{x}'\|^2/t}}{\partial {x_j}^i} L(\mathbf{x}')\mathrm{d}\mathbf{x}', \qquad (1)$$

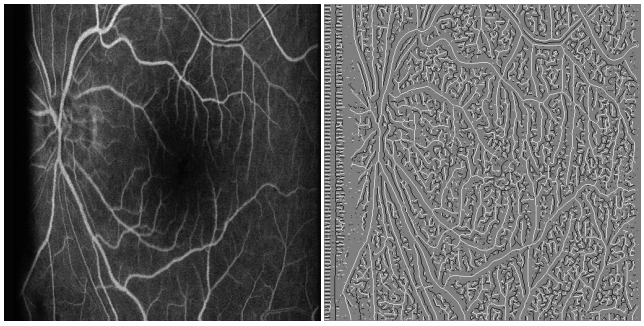
where x_j is the image coordinate w.r.t. which the derivative is taken. Mixed derivatives are computed by taking mixed derivatives of the Gaussian kernel.

We are now able to define a scalar field $\rho(\mathbf{x},t)$ over our image that takes value -1 for valleys, 1 for ridges and 0 elsewhere as follows:

$$\rho(\mathbf{x},t) = -\tfrac{1}{2}\operatorname{sign}(\lambda(\mathbf{x},t))\times \\ |\operatorname{sign}(\mathbf{g}(\mathbf{x}+\epsilon\hat{\mathbf{v}},t)\cdot\hat{\mathbf{v}}) - \operatorname{sign}(\mathbf{g}(\mathbf{x}-\epsilon\hat{\mathbf{v}},t)\cdot\hat{\mathbf{v}})|, \qquad (2)$$

where the gradient vector $\mathbf{g}(\mathbf{x},t)$ is defined as $\boldsymbol{\nabla} L(\mathbf{x},t)$, $\lambda(\mathbf{x},t)$ is the largest eigenvalue by absolute value of the Hessian matrix $\mathbf{H}(\mathbf{x},t) = \boldsymbol{\nabla}\boldsymbol{\nabla}^T L(\mathbf{x},t)$ and $\hat{\mathbf{v}}(\mathbf{x},t)$ is the unit-length normalized eigenvector belonging to that eigenvalue. In (2) $\hat{\mathbf{v}}$ is evaluated at $(\mathbf{x},t)$. The parameter ϵ is the spatial accuracy with which the point-sets are detected. In the continuous case the limit $\epsilon \to 0$ is taken, but in the discrete pixel case $\epsilon = 1.0$ pixel is a natural choice.

This work is carried out in the framework of the NWO research project STW-UGN/4496.



(a) (b)

Fig. 1. (a) An optical fluorescence image of the fundus of the human retina obtained from a scanning laser ophthalmoscope. Image size is 512×512 pixels. The artefacts on the boundaries of the image are due to signal underflow. (b) The ridges (white) and the valleys (black) are obtained at scale $t = 8.0$ pixel^2. Note the large response of the ridge detector w.r.t. to noise in the background.

In figure 1 we show an example of the ridge detection at a fundus image. We see that the structures we are interested in are a sub-set of the ridge pixels.

2.2. Affine convex sets: grouping ridge pixels

The problem of detecting the vessels of images like the one in figure 1 (a) is already reduced to classifying those ridge pixels of figure 1 (b) which delineate vessels. We can discard the valley pixels, because the vessels are the bright structures. Still, the ridge pixels are not an optimal representation of the vessels. To obtain a better representation we are going to group ridge pixels which belong to the same ridge. This leads to a representation of pixel-sets which are approximately straight line-elements.

The grouping method is a simple region growing algorithm which compares an already grouped ridge pixel with ungrouped pixels in a neighborhood of radius ϵ_{c}, where the subscript c stands for connectivity. The comparison between the grouped and a candidate pixel within the neighborhood is based on two conditions: (1) Is the direction of the ridges on which the pixels were found similar? (2) If so, are the pixels on the same ridge or are they on parallel ridges? The first condition can be checked by taking the scalar product of the principal eigenvectors of the Hessian matrix at the location of the pixels. As noted before, these vectors denote the direction of principal surface curvature and were used to detect the ridges, to which they are perpendicular. If the pixels have similar orientation the scalar product will be close to 1. The second condition can be checked by computing the unit-length normalized vector $\hat{\mathbf{r}}$ between the locations of the two pixels under consideration and taking the vector product between this vector and the principal direction of the grouped pixel. If the pixels are on the same segment, the vector product will be close to 1.

Mathematically we check the following inequalities:

$$\|\mathbf{x}_{\mathrm{g}} - \mathbf{x}_{\mathrm{u}}\| \leq \epsilon_{\mathrm{c}}, \quad (3)$$

$$|\hat{\mathbf{v}}(\mathbf{x}_{\mathrm{g}}, t) \cdot \hat{\mathbf{v}}(\mathbf{x}_{\mathrm{u}}, t)| \geq \epsilon_{\mathrm{o}}, \quad (4)$$

$$|\hat{\mathbf{v}}(\mathbf{x}_{\mathrm{g}}, t) \wedge \hat{\mathbf{r}}| \geq \epsilon_{\mathrm{p}}, \quad (5)$$

where the subscript g stands for grouped, u for ungrouped, o for orientation and p for parallelism. The ϵ's determine the measure for similarity.

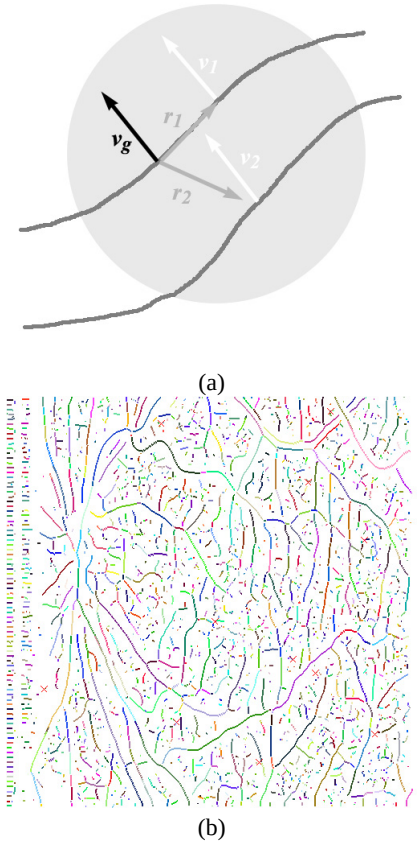


Fig. 2. (a) The dark curved lines are two ridges. The disk represents the neighborhood within which pixels are compared. The dark arrow is the direction belonging to a grouped pixel. The light arrows are the directions of still ungrouped pixels. The pixel that belongs to the same ridge will be grouped to the pixel with the dark arrow, because it satisfies conditions in (3)-(5). The pixel on the parallel ridge does not satisfy condition (5) and will not be grouped. (b) The convex sets of the ridges of figure 1 (b). Every grouped set has its own color. The settings used are $\epsilon_{\mathrm{c}} = 3.0$ pixel, $\epsilon_{\mathrm{o}} = 0.95$ and $\epsilon_{\mathrm{p}} = 0.95$.

We refer to these sets as (affine) convex sets. Convex because they approximate straight line-elements in two dimensions (in higher dimensions they approximate convex sets) and affine because of the parallelism condition. Figure 2 (a) illustrates the construction of the convex sets. In figure 2 (b) we show the convex sets of the ridges of figure 1 (b).

3. CLASSIFICATION OF CONVEX SETS

For classifying the convex sets certain features are computed which are attributed to each set. The classification is mainly based on the properties of the mean profile of each set. As explained before, to each ridge pixel corresponds a vector $\hat{\mathbf{v}}$, pointing in the direction perpendicular to the ridge. For all pixels $\mathbf{x}$ in each set the profile is sampled along a line in the direction of $\hat{\mathbf{v}}(\mathbf{x})$. The mean pro-

file is computed by averaging all profiles in the set. The following characteristics are computed:

1. The width of the profile, defined as the distance between the first strongest edge in the direction of $\hat{\mathbf{v}}$ and $-\hat{\mathbf{v}}$, starting at the midpoint of the profile.
2. The mean of the strength of the edges, where the strength is defined as the absolute value of the slope at the location of the edge.
3. The height of the profile (its value at the midpoint location) divided by its width.
4. The edge-strength divided by the width of the profile.
5. A curvature measure of the convex set. It is based on the absolute value of the scalar product of the direction vector $\hat{\mathbf{v}}$ at consecutive locations in the set. These values are added and divided by the number of pixels in the set. This measure is one for straight segments.

Note that we have not taken into account the length of the convex set or the average pixel value. The reason is that the smaller vessels cannot be distinguished from background based on these features.

The classifier we have chosen is a kNN-classifier for which an optimized implementation can be found at [7].

Using k neighbors of which n are labeled as vessel the a posteriori probability for being vessel is given as

$$p_i = n_i/k, \tag{6}$$

where the subscript i refers the convex set with index i. As [8] shows, (6) converges to the real probability in the limit of an infinite number of examples.

4. RESULTS

Four images like the one in figure 1 (a) were available, for which we detected the ridges at scale $t = 8.0$ pixels^2. The convex sets were extracted using $\epsilon_{\mathrm{c}} = 3.0$ pixel, $\epsilon_{\mathrm{o}} = 0.95$ and $\epsilon_{\mathrm{p}} = 0.95$. This left us with $11,051$ convex sets. These sets were labeled by hand as vessel and background, resulting in $1,276$ vessel sets and $9,775$ background sets. To be able to evaluate our method, we performed leave-one-out experiments on the data set, i.e. we used three of the data sets to build the kNN-classifier and classified the fourth with it. We did this for all four data sets. A set was classified as vessel if $p_i > 0.5$. In the kNN-classification we used a k-value of 20. The results are shown in table 1.

For the accuracy, sensitivity and specificity we show the results in table 2. The accuracy is defined as the total number of correctly classified sets divided by the total number of sets. The sensitivity is the fraction of correctly classified vessels w.r.t. the total number of manually labeled vessels and the specificity is the fraction of correctly classified background w.r.t. the manually labeled background. As an example of a classified image, we show in figure 3 (a) the input image and in 3 (b) the sets that were manually labeled as vessel. In figure 3 (c) the classified vessels are shown, where $p_i > 0.5$ is taken as threshold. Figure 3 (d) shows the a posteriori probabilities for all sets in 3 (a), see also equation (6). The gray-values for the sets range from white (definitely vessel) to black (definitely background).

5. DISCUSSION AND CONCLUSIONS

In this paper we have investigated the local classification of image structures. For this purpose we have introduced the notion of affine convex sets. These sets are a good representation of image structures with an intrinsic lower dimension than that of the embedding image. The advantage of such a representation is that use is made of the prior knowledge that the structures are elongated. As a result, instead of classifying $512 \times 512 \approx 2.6 \cdot 10^5$ pixels, we classified approximately $2.7 \cdot 10^3$ sets for each image.

We have attributed image features to the convex sets and classified the sets based on those features. However, to check that our method performs better than human observers, we should compare the algorithm performance with intra- and inter-observer variability, which was not done in this study.

Another point of investigation is that not all vessels respond to the ridge detector. This is due to the various widths of the vessels and the fixed scale that is used. A multi-scale approach could solve this problem.

In stead of taking a kNN-classifier we could also have opted for a feed-forward neural network, but the advantages of the kNN-classifiers are that they are powerful non-parameterized classifiers and very simple to train, as opposed to neural networks. We did not investigate the use of support vector machines [9], because we wanted to approximate the a posteriori probabilities for the classification as well, which are obtained almost for free with neural networks [10] and kNN-classifiers [8], whereas it is not clear at the present moment how to obtain them from support vector machines.

It should also be noted that no feature selection was performed. We have chosen certain features, without testing the results by adding or removing features. Furthermore, it would be interesting to investigate more features such as histogram information around each convex set.

	vessel	background	total
correct	893	9,020	9,913
incorrect	755	383	1,138
total	1,276	9,775	11,051

Table 1. Results of classification of the convex sets in leave-one-out experiments. The lower line shows the values of the manual labeling.

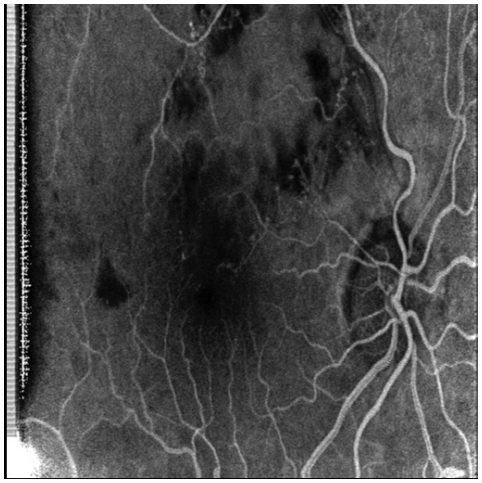
	acc.	sens.	spec.
class. result	0.897	0.700	0.923
all negative	0.867	0	1

Table 2. The accuracy, sensitivity and specificity of our method is shown in the first line. The second line shows what happens if all the sets are being classified as background.

6. REFERENCES

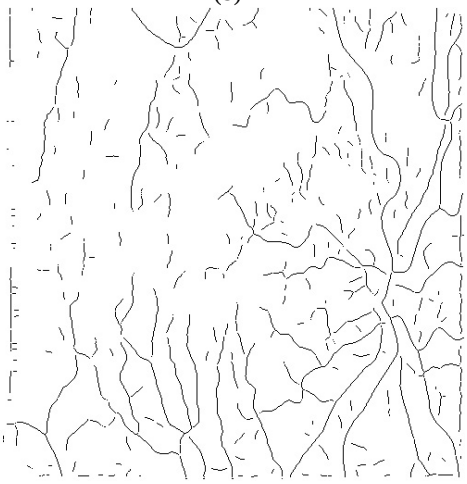
[1] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. on Med. Imag.*, vol. 19, no. 3, pp. 203–210, 2000.

[2] M. E. Martínez-Pérez, A. D. Hughes, A. V. Stanton, S. A. Thom, A. A. Bharath, and K. H. Parker, "Retinal blood vessel segmentation by means of scale-space analysis and region growing," in *Proc. of the 2nd Int. Conf. on Med. Image Comp. and Comp. Ass. Interv.*, G. Goos, J. Harmanis, and

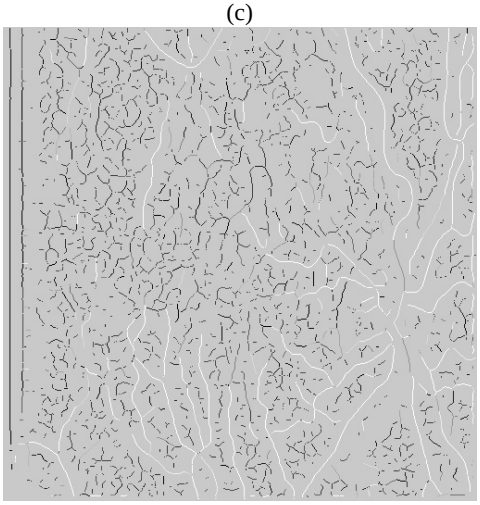


(a)

(b)



(c)



(d)

Fig. 3. (a) Input image. The artefacts around the boundaries are due to signal overflow. (b) The manual labeled vessel sets. (c) The sets that are classified as vessel. (d) The a posteriori probabilities of the convex sets. The classification result ranges from white ($p_i = 1$) to black ($p_i = 0$), see also equation (6).

J. van Leeuwen, Eds., Berlin, 1999, number 1679 in Lecture Notes in Computer Science, pp. 90–97, Springer.

[3] Y. Tolias and S. Panas, "A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering," *IEEE Trans. Med. Imag.*, vol. 17, no. 2, pp. 263–273, 1998.

[4] B. Kochner, D. Schuhmann, M. Michaelis, G. Mann, and K.-H. Englmeier, "Course tracking and contour extraction of retinal vessels from color fundus photographs: most efficient use of steerable filters for model-based image analysis," in *Proc. SPIE on Medical Imaging*, K. M. Hanson, Ed., 1998, number 3338, pp. 755–761.

[5] S. N. Kalitzin, J. J. Staal, B. M. ter Haar Romeny, and M. A. Viergever, "A computational method for segmenting topological point sets and application to image analysis," *IEEE Trans. on Patt. Anal. and Mach. Intell.*, vol. 23, no. 5, pp. 447–459, 2001.

[6] L. M. J. Florack, *Image structure*, Kluwer Academic Publishers, Dordrecht, 1997.

[7] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching," *Journal of the ACM*, vol. 45, pp. 891–923, 1998, for an implementation see: `http://www.cs.umd.edu/~mount/ANN/`.

[8] R. O. Duda and P. E. Hart, *Pattern classification and scene analysis*, Wiley, New York, 1973.

[9] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines (and other kernel-based learning methods)*, Cambridge University Press, Cambridge, 2000.

[10] C. M. Bishop, *Neural networks for pattern recognition*, Clarendon Press, Oxford, 1996.