Rotation Invariant Eigenvessels and Auto-context for Retinal Vessel Detection

Alessio Montuoro[a], Christian Simader[a] and Georg Langs[a,b] and Ursula Schmidt-Erfurth[a]

[a]Christian Doppler Laboratory for Ophthalmic Image Analysis, Department of Ophthalmology and Optometry, Medical University of Vienna, Austria
[b]Computational Imaging Research Lab, Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, Austria

ABSTRACT

Retinal vessels are one of the few anatomical landmarks that are clearly visible in various imaging modalities of the eye. As they are also relatively invariant to disease progression, retinal vessel segmentation allows cross-modal and temporal registration enabling exact diagnosing for various eye diseases like diabetic retinopathy, hypertensive retinopahty or age-related macular degeneration (AMD).

Due to the clinical significance of retinal vessels many different approaches for segmentation have been published in the literature.[1] In contrast to other segmentation approaches our method is not specifically tailored to the task of retinal vessel segmentation. Instead we utilize a more general image classification approach and show that this can achieve comparable results.

In the proposed method we utilize the concepts of eigenfaces and auto-context. Eigenfaces[2] have been described quite extensively in the literature and their performance is well known. They are however quite sensitive to translation and rotation. The former was addressed by computing the eigenvessels in local image windows of different scales, the latter by estimating and correcting the local orientation. Auto-context[3] aims to incorporate automatically generated context information into the training phase of classification approaches. It has been shown to improve the performance of spinal cord segmentation[4] and 3D brain image segmentation.[5]

The proposed method achieves an area under the receiver operating characteristic (ROC) curve of $A_z = 0.941$ on the DRIVE[6] data set, being comparable to current state-of-the-art approaches.

Keywords: eigenvessels, auto-context, retinal vessel, PCA, random forest, rotation invariant

1. INTRODUCTION

In this paper we propose a supervised method based on the idea of eigenvessels (similar to eigenfaces[2]) and an auto-context loop[3] for result refinement. Rotation invariance is achieved by estimating the local orientation and rotating the eigenvessels correspondingly. The method is then tested on the DRIVE data set.

2. METHODOLOGY

The proposed method uses a supervised learning approach and thus is split into two parts: the training phase (see algorithm 1) and the classification phase (see algorithm 2).

The key points of the algorithm are the orientation estimation (described in section 2.1), the rotated window/feature extraction (described in section 2.2), and the auto-context loop (described in section 2.3).

Each image is converted to gray scale and preprocessed as described by Soares et al.[7] in order to reduce artifacts generated by the circular field of view of fundus cameras (see figure 1).

Further author information: (Send correspondence to Alessio Montuoro)
Alessio Montuoro: E-mail: alessio.montuoro@meduniwien.ac.at

Algorithm 1: training phase

input : set of training images I
and corresponding *labels*
output: set trained classifiers C_n,
eigen_vessels and $pos^{context}$

$\phi \leftarrow orientationEstimation(I)$
$w^{image} \leftarrow rotatedWindows(I, \phi)$
$f^{image}, eigen\_vessels \leftarrow pca(w^{image})$
$w^{labels} \leftarrow rotatedWindows(labels, \phi)$
$pos^{context} \leftarrow peaks(mean(w^{labels}))$
for $n = 0$ **to** $number\_of\_context\_loops$ **do**
  for I_i *in* I **do**
    if $n = 0$ **then**
      $f^{temp} \leftarrow f^{image_{I \setminus I_i}}$
      $f^{I_i} \leftarrow f^{image_{I_i}}$
    else
      $f^{temp} \leftarrow f^{image_{I \setminus I_i}} \cup f_{n-1}^{prediction_{I \setminus I_i}}$
      $f^{I_i} \leftarrow f^{image_{I_i}} \cup f_{n-1}^{prediction_{I_i}}$
    $C_{temp} \leftarrow train(f^{temp}, labels)$
    $P(I_i)_n \leftarrow predict(C_{temp}, f^{I_i})$
  $f_n^{prediction_I} \leftarrow context(P_n, \phi, pos^{context})$
  if $n = 0$ **then**
    $f^{train} \leftarrow f^{image_I}$
  else
    $f^{train} \leftarrow f^{image_I} \cup f_{n-1}^{prediction_I}$
  $C_n \leftarrow train(f_{train}, labels)$
return C_*, *eigen_vessels*, $pos^{context}$

Algorithm 2: classification phase

input : image to segment I
output: vessel probability map for I

$\phi \leftarrow orientationEstimation(I)$
$f^{image} \leftarrow project(I, \phi, eigen\_vessels)$
for $n = 0$ **to** $number\_of\_context\_loops$ **do**
  if $n = 0$ **then**
    $P_n \leftarrow predict(C_n, f^{image})$
  else
    $P_n \leftarrow predict(C_n, f^{image} \cup f_{n-1}^{prediction})$
  $f_n^{prediction} \leftarrow context(P_n, \phi, pos^{context})$
return $P_{number\_of\_context\_loops}$

Figure 1. preprocessed image (original FOV highlighted)

2.1 Local orientation estimation

We estimate the local image feature orientation $\phi(x, y)$ around the pixel position (x, y) in the input image I using masked central image moments. By varying the size of the window different scales of image features can be considered. Figure 2 shows the result for different scales h and figure 3 shows the circular mask used for $h = 10$.

For each pixel position (x, y) in the input image I we estimate the local image orientation $\phi(x, y)$ by calculating the angle of the eigenvector with the largest eigenvalue as shown in (1).

$$\phi(x,y) = \frac{1}{2} \cdot \arctan\left(\frac{2 \cdot \hat{\mu}^{11}(x,y)}{\hat{\mu}^{20}(x,y) - \hat{\mu}^{02}(x,y)}\right) \quad (1)$$

$$\hat{\mu}^{pq}(x,y) = \frac{\mu^{pq}(x,y)}{s(x,y)} \quad (2)$$

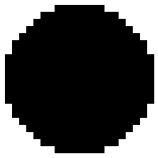
with $\hat{\mu}^{pq}(x, y)$ being the masked second order central image moment around the position (x, y) as defined in (2), $\mu^{pq}(x, y)$ being the masked central image moment around the pixel (x, y) as defined in (3) and $s(x, y)$ the sum of intensities within the mask as defined in (4).

$$\mu^{pq}(x,y) = \sum_{x_i=x-h}^{x+h} \sum_{y_i=y-h}^{y+h} (x_i - \hat{x})^p \cdot (y_i - \hat{y})^q \cdot I[x_i, y_i] \cdot mask[x_i - x, y_i - h] \quad (3)$$

$$s(x,y) = \sum_{x_i=x-h}^{x+h} \sum_{y_i=y-h}^{y+h} I[x_i, y_i] \cdot mask[x_i - x, y_i - h] \quad (4)$$

(a) $h = 2$ (b) $h = 5$ (c) $h = 10$

Figure 2. local orientation estimation for different scales (original FOV highlighted)



(a) $h = 10$

Figure 3. window mask
(size 21 x 21)

with $\hat{x}$ and $\hat{y}$ being the position of the centroid of the window around (x, y) with the size $2 \cdot h + 1$ as defined in (5) and (6).

$$\hat{x} = \frac{1}{s(x,y)} \cdot \sum_{x_i=x-h}^{x+h} \sum_{y_i=y-h}^{y+h} x_i \cdot I[x_i, y_i] \cdot mask[x_i - x, y_i - h] \tag{5}$$

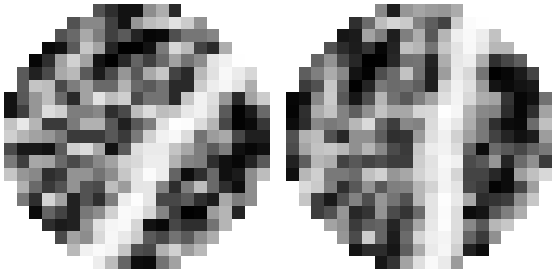
$$\hat{y} = \frac{1}{s(x,y)} \cdot \sum_{x_i=x-h}^{x+h} \sum_{y_i=y-h}^{y+h} y_i \cdot I[x_i, y_i] \cdot mask[x_i - x, y_i - h] \tag{6}$$

2.2 Rotation invariant feature extraction

2.2.1 Training

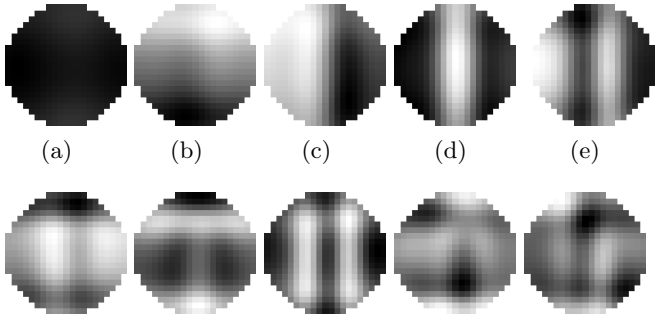
During training a small window around each pixel within the field of view of the fundus camera is extracted (as shown in Fig. 4(a)). This window is then rotated according to the local image orientation ϕ (as shown in Fig. 4(b)).

Similar to eigenfaces[2] we then perform a principal component analysis (PCA) on the serialized window, the resulting eigenvectors can be seen in figure 5.



(a) raw masked window (b) rotated window

Figure 4. image windows, $h = 10$
(histogram equalized and inverted for visualization)



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

Figure 5. first 10 eigenvessels for $h = 10$
(histogram equalized and inverted for visualization)

Figure 6. eigenvessel shown in Fig. 5(d) pre-rotated in steps of 18° for scales $h = [2, 5, 10]$

2.2.2 Feature generation

The features for each pixel position in an image could now be generated by extracting a window around the position, rotating it according to the local image orientation ϕ and projecting it into the feature space using the eigenvessels computed during training. For computational efficiency the eigenvessels can be rotated in a preprocessing step and the projection can be made using the eigenvessel with the same rotation as the local image orientation. An example of such rotated eigenvessels can be seen in figure 6.

In order to gain a degree of scale invariance the above steps are performed for multiple scales and the final feature vector f_{image} is formed by concatenating the feature vector of different scales.

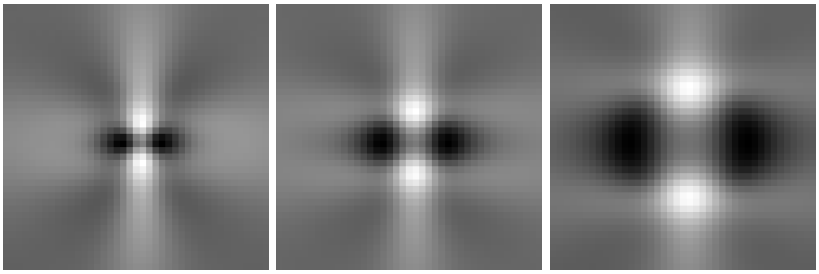
2.3 Auto-context

The idea behind auto-context[3] is to use the probability maps given by a first classification step as additional context information in a second classification step. For this Tu et al.[3] extract the class probabilities around each pixel at around 4.000 locations (both single pixel probabilities and means of small patches) and append these to the feature vector. We tried to reduce this number by learning the position of relevant context locations from the training labels.

2.3.1 Context position estimation

A good feature should be able to discriminate between different classes, meaning in the case of vessel segmentation that a context position relative to the current pixel that has a equal probability of being a vessel or a background pixel is a bad feature. On the other hand a position that has a very high probability of being either one of the classes has a high discriminative value.

To find such positions we calculated the relative prior probability during the training phase using the training labels. Figure 7 shows the result of this computation. The center of the image corresponds to the mean class label (and thus the prior class probability) the rest is the prior class probability relative to the center after being rotated by the local image orientation computed at different scales. As one can see the highest probability for a pixel of the class vessel is when looking in the direction of the local image orientation (i.e. the top in Fig. 7) in front and behind the current pixel. This can be explained intuitively by the fact that if the current pixel is a background pixel there is a high and almost uniform probability of the surrounding pixels also being a background pixel. On the other hand if the current pixel is a vessel pixel there is a high probability of finding



(a) $h = 2$ (b) $h = 5$ (c) $h = 10$

Figure 7. multiscale relative prior vessel probability

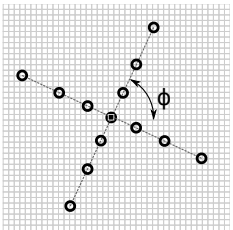
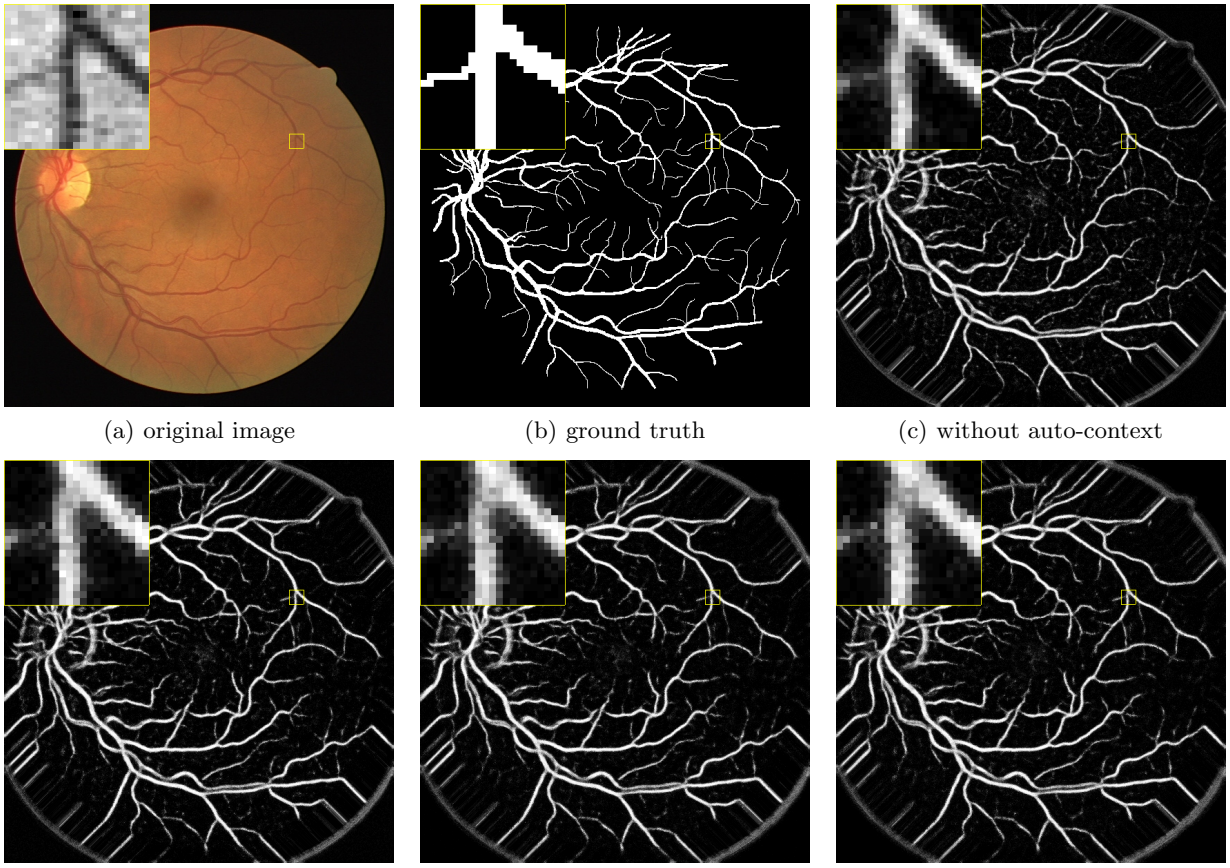


Figure 8. relative context positions rotated by local image orientation ϕ



(a) original image (b) ground truth (c) without auto-context

(d) with 1 auto-context loop (e) with 2 auto-context loops (f) with 3 auto-context loops

Figure 9. results on test image 1 of the DRIVE data set

other vessel pixels along the local image orientation (i.e. the vessel) and a high probability of finding background pixels perpendicular to the vessel.

We now use the local maxima in this relative prior probability maps (i.e. highest vessel probability) and the local minima (i.e. highest background probability maps) as positions for our context features. The resulting context feature vector is simply computed by finding 13 probability values (four for the minima / maxima at each scale and one in the center) relative to the current pixel rotated by the local image orientation as show in figure 8.

2.3.2 Auto-context training loop

The first step is to train a classifier C_0 (in our case a random forest[8]) on the set of all feature vectors f^{image} and their corresponding *labels* ("vessel" and "background") of all available training images. In order to be able to use auto-context we need a realistic probability prediction for each image I_{train} in our training set. If we would simply use the trained classifier C_0 to make this prediction we would get unrealistic results since I_{train} would be part of the data on which C_0 was trained. In order to avoid this we trained a temporary classifier C_{temp} on all images in the training set, except I_{train} and used C_{temp} to get a prediction P_0 for I_{train}. We repeated this step for each image I_{train} in our training set and thus got a realistic prediction for each image.

This prediction maps P_0 where then used to extract the context features $f_0^{prediction}$. By concatenating f^{image} and $f_0^{prediction}$ a new feature vector is formed and used to train the next classifier C_1 in the context loop. As seen in algorithm 1 this process is repeated multiple times.

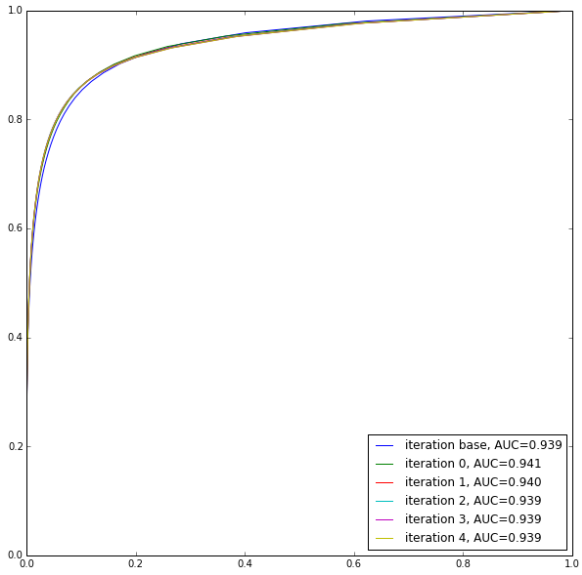


Figure 10. receiver operating characteristic (ROC) curve and area under the curve (A_z) for the proposed method on the DRIVE data set

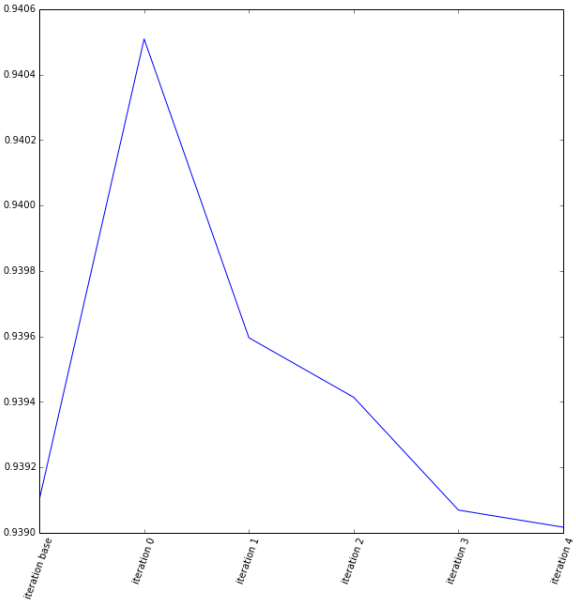


Figure 11. A_z after 0 to 5 auto-context loops

3. RESULTS & VALIDATION

The training and testing was performed on the DRIVE[6] data set. The data set consists of 40 color fundus images of which 33 show no sign of diabetic retinopathy and 7 show signs of mild early diabetic retinopathy. The set is split into 20 training and 20 test images, each having at least one manual segmentation. In figure 9 the ground truth and the classification result of test image 1 of the DRIVE data set can be seen. A quantitative evaluation was performed by calculating the ROC curve and the area under the curve (A_z) (see image 10). The segmentation results compare very favorably to other proposed methods[9] outperforming all but the method proposed by Staal et al.[6]

4. CONCLUSIONS AND FURTHER WORK

Interestingly the auto-context loop enhances the result after the first iteration (as can be seen in figure 11), however after multiple iterations it degrades the segmentation of vessel endings and vessel bifurcations which reduces the overall A_z. This is due to the fact that the auto-context learns the average appearance of vessels, which includes a order of magnitude more samples of straight vessel sections than vessel bifurcations and vessel endings. Results could be further improved by treating bifurcations and vessel endings as own classes and extending the proposed method to a multiclass method.

This method does not make any assumptions about the geometry or any other feature of retinal vessels but learns the feature representation from the training data and thus could be applied to classification tasks similar to vessel segmentation. Furthermore an extension to three dimensional data sets (e.g. retinal optical coherence tomography[10]) is possible and will be investigated further.

ACKNOWLEDGMENT

The financial support of the Austrian Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged.

REFERENCES

[1] Khan, M. I., Shaikh, H., Mansuri, A. M., and Soni, P., "A review of retinal vessel segmentation techniques and algorithms," *International Journal of Computer Technology and Applications* **2**, 1140–1144 (September 2011).

[2] Turk, M. A. and Pentland, A. P., "Face recognition using eigenfaces," in [*Computer Vision and Pattern Recognition, 1991. CVPR '91. IEEE Conference on*], 586–591, IEEE (1991).

[3] Tu, Z., "Auto-context and its application to high-level vision tasks," in [*Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*], 1–8 (June 2008).

[4] Kawahara, J., McIntosh, C., Tam, R., and Hamarneh, G., "Augmenting auto-context with global geometric features for spinal cord segmentation," in [*Machine Learning in Medical Imaging*], Wu, G., Zhang, D., Shen, D., Yan, P., Suzuki, K., and Wang, F., eds., *Lecture Notes in Computer Science* **8184**, 211–218, Springer International Publishing (2013).

[5] Tu, Z. and Bai, X., "Auto-context and its application to high-level vision tasks and 3d brain image segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **32**, 1744–1757 (Oct 2010).

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

[7] Soares, J. V. B., Le, J. J. G., Cesar, R. M., Jelinek, H. F., Cree, M. J., and Member, S., "Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification," *Medical Imaging, IEEE Transactions on* **25**, 1214–1222 (2006).

[8] Breiman, L., "Random forests," *Machine Learning* **45**, 5–32 (Oct. 2001).

[9] Niemeijer, M., Staal, J., van Ginneken, B., Loog, M., and Abramoff, M. D., "Comparative study of retinal vessel segmentation methods on a new publicly available database," *Proc. SPIE* **5370**, 648–656 (2004).

[10] Hitzenberger, C. K., Trost, P., Lo, P.-W., and Zhou, Q., "Three-dimensional imaging of the human retina by high-speed optical coherence tomography," *Optics Express* **11**, 2753–2761 (2003).