

Vessel Segmentation in 2D-Projection Images Using a Supervised Linear Hysteresis Classifier

Alexandru Paul Condurache
Institute for Signal Processing
University of Luebeck
Ratzeburger Allee 160, D-23538 Luebeck

Til Aach
Institute of Imaging and Computer Vision
RWTH Aachen University
D-52056 Aachen, Germany

Abstract

2D projection imaging is a widely used procedure for vessel visualization. For the subsequent analysis of the vasculature, precise measurements of e.g. vessel area, vessel length or the number of vessel segments are needed. To achieve these goals vessel enhancement and segmentation are required. While there are already many vasculature specific vessel segmentation algorithms, we describe in this contribution a more general supervised segmentation method which includes a feature extraction step followed by feature selection and segmentation based on the hysteresis classification paradigm. The method was tested on retina photographs. The rates of false positives and correct classifications were comparable with dedicated methods on similar data sets while it needed less time for both training and providing a segmentation result.

1 Introduction

Vessel imaging and analysis is required in medical applications related to both clinical research and clinical routine. In, e.g., blood vessel and retinal pathology, intervention planning, blood-flow analysis or drug evaluation, 2D projection imaging such as angiography or retinal photography is often the method of choice as it returns results of sufficient quality without the overhead required by 3D techniques. There are many machine vision methods addressing the problem of vessel segmentation in 2D-projection images [11]. They typically begin by a vessel enhancement step which uses e.g. matched filters [14, 9], the eigenvalues of the Hessian matrix at different scales [8], the wavelet transform [3] or other methods [5, 6]. In [17, 6, 4], for an optimal description of vessels multidimensional pixel-feature vectors are used. Vessels can be then segmented in a supervised manner based on a pre-labeled set of examples e.g. using k-nearest neighbor classification [17], or in an unsu-

pervised manner directly on the analyzed image, without any training set by e.g. tracking [3], clustering [6, 4], as well as by multi-thresholding methods [10, 5, 13]. In particular such multi-threshold methods can achieve good results being especially designed to deal with the large overlap of the vessel and background pixel classes. Semi-automatic vessel classification algorithms rely on user intervention [16, 6]. Such a strategy is appropriate only when the number of segmented images is low, otherwise an automatic algorithm is needed. If large data sets need to be analyzed then a fast method is required in which case tracking and clustering are less suited. There are automatic techniques specifically designed to permit the embedding of prior knowledge [5] which are fast enough, but less adaptable as their scope is limited by the prior information. Supervised segmentation is then the method of choice as it can adapt by using a properly chosen pre-labeled training set.

We propose in this contribution an automatic vessel segmentation algorithm based on the hysteresis classification paradigm which extends the method described in [5] to multidimensional pixel-feature vectors in a supervised manner. Our final goal is the computation of vessel tree characteristics like vessel area, number of vessel segments, etc.

2 Vessel enhancement and feature extraction

In 2D-projection imaging, the blood either by itself (e.g. in retinal photography) or by the contrast agent it carries (e.g. in angiography) is responsible for the vessel contrast to the background. Consequently, the more blood present, the better the vessel contrast. Thus there is a significant variance of the vessel pixel class between large well contrasted vessels and small weakly contrasted ones. Such images typically show also vessel-like artefacts which cause a significant overlap between vessel and background classes. Thus vessel enhancement is needed before segmentation to improve the separability of vessels and background.

We describe vessels as tubular, contrasted structures of

a certain size. During enhancement we increase the vessel contrast and improve the homogeneity of their gray-level representation. The result of such vessel enhancement is a vessel map. The design of a vessel map is typically based on a single particular vessel property. To increase separability, such vessel maps can be combined so that each pixel is described by a vector containing all the gray-levels that pixel has in each of the considered vessel maps [6]. We have first computed several vessel maps and then selected those which together yield the best feature space (Section 3.2.1).

We have used the five vessel maps proposed in [6]: 1) The result of a Tophat transform used to select contrasted structures of a certain size. 2) The result of the analysis of the eigenvalues of the Hessian matrix extended in a multi-scale approach to also reach the small vessels and improve the homogeneity of the vessel pixel class. 3) The result of a homomorphic filter to analyze two signals combined in a multiplicative way and then select only the high frequency vessels. 4) The result of a multi-resolution analysis using the Laplace pyramid, to select only high frequency vessel structures and improve homogeneity of the vessel pixel class. 5) The results of a nonlinear high-pass filter to improve the homogeneity of the vessel pixel class.

3 The hysteresis classification paradigm

In a binary classification problem, if the two classes are not linearly separable in the feature space, and under the assumption that no new features can be added, nonlinear techniques need to be used to obtain satisfactory results. Such nonlinear techniques are typically expensive in training time and can easily overfit. However if the components of one of the classes (object class) exhibit certain connectivity in some additional space different from the feature space, while at the same time vectors of the two classes which are close in this space are different enough in the feature space, then two linear classifiers working in the original feature space suffice to achieve a good classification. One of these classifiers, the “pessimist” should select only object points, having thus practically a zero false positives rate and the other one, the “optimist” should select practically all object points irrespective of the number of false positives. We make here the realistic assumption that neither class has an infinite support. Then based on the connectivity property from the additional space, the points which with a high confidence belong to the object class and which were selected by the “pessimist” can be used to choose true object points from among those selected by the “optimist” as long as they are linked with one another over a vicinity [2], [12]. The pair of linear classifiers builds then a linear hysteresis classifier. Here we concentrate on linear classifiers but other classifiers can be used also as long as they are trained to fit the role of a “pessimist” and an “optimist”, for example

by training on specially chosen sets. Of course if the connectivity can be described by a feature or set of features, these can then be added to the original feature space and a standard classifier can be used.

3.1 Hysteresis threshold

The simplest form of a linear hysteresis classifier working in a 1D feature space is a bi-threshold procedure called hysteresis thresholding [2, 13, 5, 1]. Assuming that object points are described by high values of the feature and non-object points by low values, a high threshold has the role of the “pessimist” and selects only object-points and a low threshold is the “optimist” and selects all object-points. Points whose feature value is below the low threshold are most probably not object-points and points with a feature value above the high threshold are most probably object-points. Those in between are object-points only if they are connected to a certain object-point.

The two thresholds are established rather heuristically. Canny [2] recommends that their ratio should be between two and three. In [13, 1] they are chosen in an application-dependent way. In [5] they are set by hypothesis testing based on specific error probabilities (significance levels) for both high and low threshold. These methods are designed to achieve good results in the absence of a training set. This observation justifies the heuristic manner in which the thresholds are set, as the user needs a way to adapt the hysteresis threshold to his or her needs.

3.2 Supervised multidimensional hysteresis classifier

We propose to use as “pessimist” and “optimist” two Fisher linear classifiers [7], respectively. The parameters of such a classifier are a set of weights \vec{w} defining a transformation from the original input feature space to a 1D space where the two classes are optimally separated and a threshold T to discriminate between the two classes. The separability criterion is: $F = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}$ where $\mu_{1,2}$ and $\sigma_{1,2}$ are the means and variances in the transformed space respectively. Then the weights vector is: $\vec{w} = (\vec{m}_1 - \vec{m}_2)^T \left[\frac{1}{2}(\mathbf{K}_1 + \mathbf{K}_2) \right]^{-1}$ where $\vec{m}_{1,2}$ and $\mathbf{K}_{1,2}$ are the means and covariance matrices in the input feature space respectively. Thus the classifier has the form: $\vec{w}^T \vec{y} \stackrel{\omega_1}{>} \stackrel{\omega_2}{<} T$ with \vec{y} the input feature vector. If the left side is larger than the threshold we decide for the class ω_1 , conversely for ω_2 .

For supervised classification with hysteresis, the weights \vec{w} are the same for both “pessimist” and “optimist” but the thresholds are different. To compute the two thresholds we

use the receiver operating characteristic (ROC) of the decision rule computed based on the training set. The ROC is determined by varying the threshold between the minimum and the maximum values in the transformed 1D space and computing each time the percentage of correct classifications (cc) and that of false positives (fp) on the pre-labeled training set. The thresholds are:

$$T_p = \arg \max(cc|fp = \alpha_p) \quad (1)$$

for the “pessimistic” classifier, and:

$$T_o = \arg \min(fp|cc = \alpha_o) \quad (2)$$

for the “optimistic” classifier, where α_p and α_o are as close as reasonably possible to 0% and 100% respectively.

α_p and α_o can be imposed by the user based on some prior-knowledge or they can be trained: starting from a standard Fisher classifier – with T corresponding to the ROC-point most distant to the base-line linking the ROC-points with $fp = 0\%$ and $fp = 100\%$ – used as “optimist”, try all “pessimists” – building thus a hysteresis-ROC – and choose again the one corresponding to the point most distant to the base-line. Repeat, this time keeping the “pessimist” constant and continue then for a predetermined number of steps or until the thresholds do not change anymore. It remains however to be investigated if this is a convex training procedure. Sometimes, in practice (e.g. for a sub-optimal feature space) the selection rule (2) leads to over-segmentation due to the large fp rate associated with 100% cc . In such cases it is better to define the “optimist” over a maximum fp rate: $T_o = \arg \max(cc|fp < \alpha_M)$.

3.2.1 Feature selection

The ROC can be used to characterize the feature space. Clearly, the larger the area under the ROC (AROC), the more separable the feature space. To the limit the two classes are linearly separable when AROC is one, i.e. 100% cc for 0% fp . If several features have been computed, then those which build the best feature space will also yield the largest AROC. In the case of hysteresis classification we are interested also in the derivative of the ROC curve, especially in the region where the fp rate is small. A hysteresis classifier trained on a ROC curve with a large integral but a comparatively mild increase over the region with small fp rate, will yield rather poor results because the “pessimist” classifier will select too few true vessels. Thus it is better to consider only a partial area under the ROC (pAROC). In our experiments we have computed only the AROC bounded by a 30% fp rate. Two strategies can then be followed: a full search strategy when all possible combinations of features are investigated and a sequential search strategy [15] when first the single best feature is selected then the best combination between that feature and another one and so on until the optimal feature set is found.

Feat.	T	nlf	mLp	TH	4D
pAROC	0.8201	0.6007	0.5904	0.7165	0.8227

Table 1. Partial area under ROC.

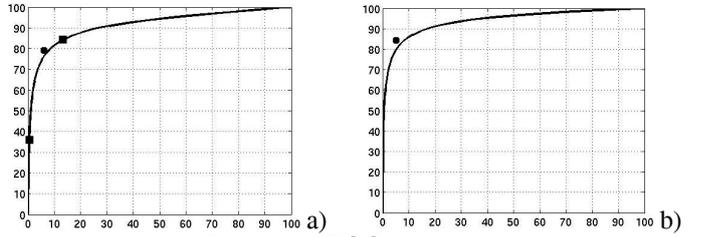


Figure 1. ROC curves.

4 Results

We have tested the linear hysteresis classifier on a publicly available data set containing retina images: the Utrecht database [17]. This data set is divided into a training set and a testing set, each containing 20 images. For each image a manual ground truth is available. For five features there are $N = 31$ possible combinations and a full search feature selection strategy was used. After feature selection we obtained a final four dimensional pixel-feature vector (4D) computed using the Tophat (T), the result of nonlinear filtration (nlf), the Multiscale-Laplacian (mLp) and the Tophat-Hessian (TH) vessel maps. The performance of this feature space is shown in Table 1. The training-ROC curve used to find the thresholds of the “pessimistic” and “optimistic” classifiers is shown in Figure 1a. The two threshold were obtained as: $T_p = \arg \max(cc|fp < 0.5\%)$ and $T_o = \arg \max(cc|fp < 15\%)$. The corresponding points are marked by squares. Using these thresholds, the static functioning point of the hysteresis classifier is 79.18% cc and 6.16% fp . This point is marked by a bullet. The results obtained by the hysteresis classifier on the test set were 83.99% cc and 5.79% fp . Using the weights vector \vec{w} , a ROC curve for the test set can be computed, describing the performance of a linear classifier. For the same false-positives rate, such a procedure will yield only 81.11% cc . This suggests an improvement of 2.88 percentage points over a linear classification. This ROC curve is shown in Figure 1b. The performance-point of the hysteresis classifier is marked by a bullet. Under similar conditions a standard Fisher classifier obtained 88.33% cc and 10.21% fp . Some segmentation results are shown in Figure 2a (from the training set) and c (from the test set). The corresponding ground truth is shown in Figure 2b and d respectively.

5 Discussion

At the heart of the hysteresis classifier is the connectivity which object-points exhibit in some additional space. This information is used to achieve a correct segmentation

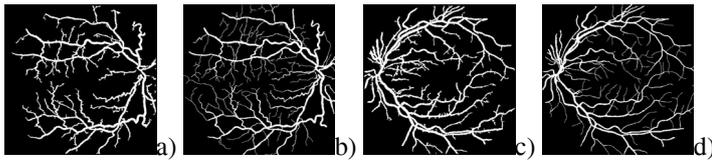


Figure 2. Segmentation results.

in the feature space despite the fact that the data is not linearly separable. At the same time it is important that in the feature space, object-points situated on the border and non-object-points are well separable, otherwise the connectivity constraint will bring false object-points in the segmentation. This should be enforced particularly during feature extraction but also during selection.

The feature extraction step allows our algorithm to be used for different types of vasculature as it separates the hysteresis classifier from the raw input data. To segment a new type of vasculature, one or several vessel maps need to be computed. The vessel maps can enhance vessel-like structures in general or use some specific information related to the investigated vasculature. The results on the test set show that in comparison with a linear classifier, the hysteresis classifier can significantly reduce the fp rate while keeping a similar cc rate. The linear hysteresis classifier can achieve results of a quality comparable with that of non-linear techniques almost as fast in training and computation time as a simple linear classifier. To achieve a correct supervised segmentation, a properly labeled training set is needed. For any binary classification there are two types of possible errors: omissions and false inclusions. For vessel segmentation, with respect to the vessel class, omissions are typically encountered for some small vessels and false inclusions for background points at the border of vessels. The training set was affected by such errors, mainly due to the tedious pre-labeling process as manually segmenting an angiogram can take up to two hours. Therefore the accuracy of any vessel segmentation algorithm using such a “bronze” ground truth is negatively affected.

6 Conclusions

We have described new methods for supervised segmentation of vessel-pixel in 2D-projection images. For training, a pre-labeled set of examples is needed. In a first step for all images of the training set a pixel-feature space is computed by gathering the results of several enhancement methods to increase the separability between vessels and background. Then feature selection follows when the best pixel-feature set is computed starting from the initial set of features. On the multidimensional feature space thus obtained two linear classifiers are trained: a “pessimist” which is supposed to select (almost) only vessel points and an “op-

timist” which selects (almost) all vessel points. Together these yield a multidimensional linear supervised hysteresis classifier. When a new image has to be segmented, first a pixel-feature space is computed by applying the same vessel enhancement methods which were used to compute the training space. Then the hysteresis classifier decides for each pixel if it is a vessel or a background pixel. The performance of the hysteresis classifier is linked to the quality of the feature set through the ROC curve. For a successful classification a feature set should be determined to maximize AROC and ensure a good separation between vessels and background at the vessel-borders.

References

- [1] T. Aach and A. P. Condurache. Transformation of adaptive thresholds by significance invariance for change detection. In *Proc. SSP*:500–506, 2005.
- [2] J. Canny. A computational approach to edge detection. *IEEE TPAMI*, 8(6):679–698, 1986.
- [3] Z. Chen and S. Molloi. Multiresolution vessel tracking in angiographic images using valley courses. *OE*, 42:1673–1682, 2003.
- [4] A. P. Condurache, et al. Imaging and analysis of angiogenesis for skin transplantation by microangiography. In *Proc. ICIP*:1250–1253, 2005.
- [5] A. P. Condurache and T. Aach. Vessel segmentation in angiograms using hysteresis thresholding. In *Proc. MVA*:269–272, 2005.
- [6] A. P. Condurache, et al. Vessel segmentation and analysis in laboratory skin transplant micro-angiogram. In *Proc. CBMS*:21–26, 2005.
- [7] R. A. Fisher. *Contributions to Mathematical Statistics*. J. Willey & Sons Inc., 1950.
- [8] A. F. Frangi et al. Multiscale vessel enhancement filtering. *LNCS*, 1496:130–137, 1998.
- [9] M. Franz and R. Schüffny. Segmentation of blood vessels in subtraction angiographic images. In *Proc. DIC*:215–224, 2003.
- [10] X. Jiang and D. Mojon. Adaptive local thresholding by verification based multithreshold probing with application to vessel detection in retinal images. *IEEE TPAMI*, 25(1):131–137, 2003.
- [11] C. Kirbas and F. K. H. Quek. A review of vessel extraction techniques and algorithms. *ACM*, 36(2):81–121, 2004.
- [12] G. Liu and R. Haralick. Two practical issues in Canny’s edge detector implementation. In *Proc. ICPR*:680–682, 2000.
- [13] A. Niemistö, et al. Robust quantification of in vitro angiogenesis through image analysis. *IEEE TMI*, 24(4):549–553, 2005.
- [14] R. Poli and G. Valli. An algorithm for real time vessel enhancement and detection. *CMPB*, 52(1):1–22, 1997.
- [15] P. Pudil et al. Floating search methods in feature selection. *PR*, 15(11):1119–1125, 1994.
- [16] D. Selle, et al. Analysis of vasculature for liver surgical planning. *IEEE TMI*, 21(11):1344–1357, 2002.
- [17] J. Staal, et al. Ridge-based vessel segmentation in color images of the retina. *IEEE TMI*, 23(4):501–509, 2004.