Context-Dependent Segmentation of Retinal Blood Vessels Using Hidden Markov Models

[1]A. Pourmorteza, [1]S. H. R. Tofighi, [1]A. Roodaki, [1]A. Yazdani, [1-3]H. Soltanian-Zadeh

[1] Control and Intelligent Processing Center of Excellence, Department of Electrical and Computer Engineering, University of Tehran, Tehran, Iran
a.pourmorteza@ece.ut.ac.ir, hszadeh@ut.ac.ir

[2] School of Cognitive Sciences, Institute for Studies in Theoretical Physics and Mathematics (IPM), Tehran, Iran

[3] Image Analysis Lab., Department of Radiology, Henry Ford Hospital, Detroit, Michigan, USA
hamids@rad.hfh.edu

Abstract

Recently, Hidden Markov Models (HMMs) have proven valuable in segmentation of brain MR images. In this paper, a combination of HMMs-based segmentation and morphological and spatial image processing techniques is proposed for the segmentation of retinal blood vessels in optic fundus images. First, a smoothing kernel is convolved with the image and the result is subtracted from the green channel image to reduce the background variations. After a simple gray-level stretching, aimed to enhance the contrast of the image, the feature vectors are extracted. The feature vector of a pixel is formed from the gray-level intensity of that pixel and those of its neighbors in a predefined neighborhood. The ability of the HHMs to build knowledge about the transitions of the elements of the feature vectors is exploited here for the classification of the vectors. Erosion and dilation are then used to delete the erroneously classified pixels.

The performance of the algorithm is tested on the DRIVE database using the accuracy, true positive ratio and false positive ratio as test criteria. The performance of our method is comparable with those of the previous works. It is, however, computationally more efficient and is approximately 9 times faster than previous classifier-based methods.

Keywords

Hidden Markov models (HMM), context-dependent image segmentation, retinal images, morphological processing.

1. Introduction

Early symptoms of systematic diseases can be detected by the assessment of retinal images. Eye is considered as a window to the retinal vasculature, where the influence of the factors that affect the human body vasculature can be observed *in vivo* non-invasively. Furthermore, inspection of optic fundus photographs [1,4], and flourocein images [2] may help to diagnose and monitor the progress of general diseases such as diabetes, hypertension, arteriosclerosis, cardiovascular diseases, stroke and eye diseases like retinopathy of prematurity [1,2,3]. Thus, the measurement and analysis of retinal blood vessels is of diagnostic value for a wide range of pathological states.

However, manual analysis of the complex retinal blood vessel trees in fundus images is a tiresome and laborious task, and as the number of images increases, it may even be impossible. Segmentation of the vessels from the background is an initial requirement for the automatic assessment of retinal images. Many automatic algorithms are proposed for the segmentation of retinal blood vessels.

According to [9], methods for detecting blood vessels generally fall into three categories: kernel-based, tracking based, and classifier-based. Kernel-based methods convolve a kernel with the image based on a predefined model. For instance, in [10] Gaussian shaped curves are used to model the cross section of a vessel, and a matched filter is used for detection. These techniques are fast when used with small kernels. However, as the kernels become larger and since the rotational invariance requires several rotated kernels, the performance of the algorithms become slower. Tracking methods use a model to track the vessels. These methods start from certain seed points which may be set manually, by simple thresholding or by morphological operations [1].

In classifier-based methods, various types of features are extracted from the image and the feature vectors are then introduced to a classifier to assign the

corresponding pixel to a certain class or segment of the image. For example, in [4], wavelet transform features of an image are extracted and a Gaussian mixture model classifier is used for classification.

A combination of aforementioned categories is also possible. In [1], the centerlines of the vessels are detected by a kernel-based method. Then these centerlines are used as seeds or starting points, in a multi-resolution region growing algorithm. In [2] a method based on multi-scale feature extraction is introduced to automatically segment the retinal blood vessels from the background. The method uses first and second derivatives of the intensity image, in a multiple pass region growing algorithm. A model-based method for segmentation of retinal blood vessels is proposed in [3], which is based on Laplace and thresholding steps followed by a classification step to improve the results.

In this paper, we will introduce a novel method for the automated segmentation of blood vessels. We start with simple kernel-based preprocessing operations followed by a contrast enhancement step. Then a Hidden Markov Model-based classification is used to assign the pixels to vessel or background classes.

Recently, Hidden Markov Models (HMMs) are used for MRI brain segmentation in [5]. Markov Random Fields (MRF), Hidden Markov Random Fields (HMRF), and Hidden Markov Models, take advantage of the relative information of the vectors and fall under the category of context-dependent classifiers. Kernel-based and thresholding methods, simply assume no relation between different classes, i.e. once a feature vector (corresponding to a pixel) is assigned to a class, the next vector may be assigned to any other class. Whereas in context-dependent classification, the class to which a vector is assigned, depends (a) on its own value, (b) on the value of other feature vectors and (c) on the existing relation among various classes.

MRFs and HMRFs were introduced in several segmentation frameworks [11-14]. While these frameworks encode the dependency between the pixel to be segmented and its first-degree neighbors, they are computationally intensive and therefore, they are not welcome in medical environments [5]. In contrast, HMMs have proven valuable in Automatic Speech Recognition (ASR) and MRI brain segmentation tasks [5,6].

The primary goal of this paper is to report a method which combines kernel-based techniques with state-of-the-art HMM-based segmentation. The algorithm is evaluated using the images obtained from the publicly available DRIVE database [16]. This combined method allows for results comparable with manually segmented images as well as with those reported by other authors.

A brief introduction to HMMs and its training algorithms is provided in section II. Section III details the preprocessing and feature extraction algorithm. Section IV includes the segmentation algorithm and necessary post-processing step to improve the results. Finally, experimental results of the proposed method are presented in section V.

2. Hidden Markov Models

Similar to the work done in [5], the fundamental idea of this paper relies on the ability of the underlying HMM to encode the knowledge about the input data vectors or sequences that reflect the characteristics of the image to be segmented e.g. intensity information about the pixel and its neighborhood.

HMMs are made of different states statically bound by transition probabilities. An HMM is characterized by a set of internal states, the transition probabilities among the states in response to an input symbol from the sequence, and the emission probabilities of symbols from the different states[5]. The knowledge is built in the form of the transition and the emission probabilities of the states that are trained during the learning stage.

A Hidden Markov model is specified by a set of *states* $S = \{s_1, s_2, \dots, s_{Ns}\}$, and a set of parameters $\Theta = \{\boldsymbol{\pi}, \mathbf{A}, \mathbf{B}\}$ [7]. The *prior probability* $\pi_i = P(q_1 = s_i)$ is the probability of s_i being the *first state* of a state sequence. Element of the transition matrix **A** are the transition probabilities, that is, the probabilities of going from state i to state j.

$$a_{i,j} = P(q_{n+1} = s_j | q_n = s_i) \qquad (1)$$

Where $a_{i,j}$ are the elements of **A**. The *emission probabilities* characterize the likelihood of a certain observation x, given that the model is in state s_i. For discrete observations,

$$x_n \in \{v_1, \dots, v_K\}$$

$$b_{i,k} = P(x_n = v_k | q_n = s_i) \qquad (2)$$

$b_{i,k}$ is the probability to observe v_k if the current state is s_i. Hence, matrix **B** is consisted of numbers $b_{i,k}$.

On the other hand, for continuous observations, the emission probabilities are defined by a set of functions describing the probability densities over the observation space given that the system is in state s_i,

$$b_i(x) = p(x_n | q_n = s_i) \qquad (3)$$

and are collected in the vector $\mathbf{B}(x)$ of functions. These probability density functions (pdf) are often parameterized, e.g. by Gaussians [6] or mixtures of Gaussians [5].

Given an observation sequence $X = \{x_1, x_2, \dots, x_N\}$ and a state sequence $Q = \{q_1, q_2, \dots, q_N\}$ the likelihood of the X along the states Q is given by the product of the emission probabilities computed along the path Q.

$$P(X|Q,\Theta) = \sum_{n=1}^{N} P(x_n|q_n,\Theta) = b_{q_1 x_1} \cdot b_{q_2 x_2} \cdot \dots \cdot b_{q_N x_N} \qquad (4)$$

According to Bayes, the probability that X and Q occur simultaneously is given by (5).

$$P(X,Q|\Theta) = P(X|Q,\Theta) \cdot P(Q|\Theta) \qquad (5)$$

Finally, the likelihood of an observation sequence X with respect to a HMM with parameters Θ expands as follows:

$$P(X|\Theta) = \sum_{\text{all } Q} P(X,Q|\Theta) \qquad (6)$$

HMMs assume that the states are hidden and cannot be observed at the output stage. Instead, only the output

emitted from the states are observable, without knowing which states emitted those outputs. From this point of view, the HMM is regarded as a symbol generating process in which the observations are viewed from the outside without knowing which state emitted them.
Here, similar to what has been previously done in [5] HMMs are viewed from a different perspective. In using HMMs for blood vessel segmentation, the goal is to find the best state sequence that might have produced a specific output. Thus the output probability in (6) is calculated only using the most probable path, that is, without the summation over all paths.

Hence, during the training, the goal of the training algorithm is to increase the output probability of the input sequences representing a class of tissue [5].
The transition and emission probabilities are updated in a manner that maximizes the output probability of a given class. As a result, the relationship between the elements of a sequence (voxels of a neighborhood) is encoded in the transition and emission matrices; in some cases, the vector of prior probabilities π is also updated to facilitate the learning process.

The concept of Minimum Classification Error is used to update the parameters of both discrete and continuous mixture of Gaussian distributions in [5]. In [6], an Expectation Maximization (EM) algorithm is utilized to optimize the parameters of a HMM which has Gaussian pdfs for emission probabilities.

Here, the algorithm used in [6] is selected. In our model, the sequence of observations is nothing more than a sequence of pixel intensities. There are many ways to generate this observation sequence. One reasonable way is to consider a 3×3 neighborhood of a certain pixel and put the intensity of the pixel along with the intensities of its neighbors in a 9×1 feature vector. Other choices of neighborhoods are also possible. The increase in size of the neighborhood may adversely affect the accuracy due to smoothing effect of large neighborhoods [5]. Thus, a suggestion is to weight the contribution of pixels inversely proportional to their distance from the center pixel.

Our proposed method involves representing each pixel by a vector. To classify a vector into one of two classes, two HMMs are trained, one per class, and then the likelihood that each model gives to the test sequence is computed; if the first model is the most likely, then the class of the vector is declared to be class 1, otherwise it is classified as class 2.

3. Preprocessing

The goal of preprocessing phase is to reduce the unwanted effects of noise and background intensity variations, and enhance the contrast between the vessel and non-vessel pixels prior to segmentation. Since the HMMs will have to learn respond to feature vectors from different images, the images must share nearly similar brightness and contrast characteristics. By compensating for these through and across data set variations, one can improve the performance of the segmentation process.

3.1. Selection of Monochromatic Image Representation

As stated earlier, this phase aims to enhance the contrast between vessel and non-vessel pixels. The first step is to choose a monochromatic image that shows high contrast between vessel and non-vessel pixels. A Monte Carlo simulation of retinal vessel profiles [15] predicts that light that is observed from a retinal vessel at green band on a RGB color image is predominantly backscattered from the vessel or transmitted once through the vessel [2]. This is confirmed by experimental results [1,4,8] where the green channel is selected among other monochromatic representations. However, it is stated that the a^* component of $L^*a^*b^*$ representation is a slightly better choice whenever the original images were not in JPEG format [1]. The a^* component is derived from green and blue channels of RGB representation. In this work, we have selected the green channel representation because of its previous good performance.

In an attempt to compensate for the contrast and brightness variations among the images of the data set, the mean and standard deviation of the histograms of the images are changed so that they nearly match predefined values of mean = 0.5 and STD = 0.075. Prior to this operation, a simple thresholding is done to remove the pixels with intensities less than 0.2, which correspond to the dark area around the image.

3.2. Background Normalization

To compensate for the background intensity variations, a smoothing kernel is convolved with the green channel after overlapping it with the mask provided in the data set. Consequently, the low intensity pixels of the image, which are the constituents of the black mask around the image, are forced to zero. The kernel convolution gives an approximation of the background variations. Next, this approximation is subtracted from the green image. The result is a background normalized image as can be seen in fig.1(c). Size of the smoothing kernel is empirically selected to be 11×11.

Since pixels in the optic disk normally have very high intensities, the smoothing and subtraction steps cause artifacts that may be selected as vessels in segmentation steps. To overcome this problem, prior to the background normalization step, the optic disk is segmented by a simple thresholding and subtracted from the image. The empty area is then filled with pixels with intensities equal to the mean value of the intensities of other pixels.

3.3. Contrast Enhancement

Numerous methods to enhance the contrast of the retinal images are reported in the literature. A kernel-based local contrast enhancement technique is introduced in [3]. In [8] a number of contrast enhancement techniques, from the classic Histogram Equalization to newly developed Contourlet transform are discussed. Since the proposed algorithms are computationally intensive, a relatively simple method is chosen here. The nature of our segmentation algorithm allows for moderate

contrast enhancement. The significance of superior contrast enhancement, however, is not denied.

Histogram of the background normalized image is similar to a Gaussian distribution with a small variance where most of the pixels have midrange gray-levels. In an attempt to enhance the global contrast, the histogram is stretched by a sigmoid function that is given by

$$sigmoid(x) = \frac{1}{1 + e^{-\alpha(x-\gamma)}} \quad (7)$$

where γ is the inflection point of the sigmoid function and is adaptively chosen to coincide with peak of the histogram, and α is the expansion factor, which is empirically selected to be 40. Result of this step is illustrated in figure 1.

4. Segmentation and Post-processing

The notion behind HMMs-based segmentation [5] is that, by training a HMM with feature vectors representing a special class of data, the model learns to produce higher outputs when a vector of that class is presented to the model as an input sequence.

Hence, the segmentation algorithm is straightforward. A HMM is trained for each class of data, the output probability of every HMM for the feature vector representing the pixel to be classified is then computed using the trained transition matrices and emission vectors that encode the relative dependency of the elements of the feature vector. The segmentation is then done, by simply assigning the pixel to the class associated with the HMM showing the highest output probability.

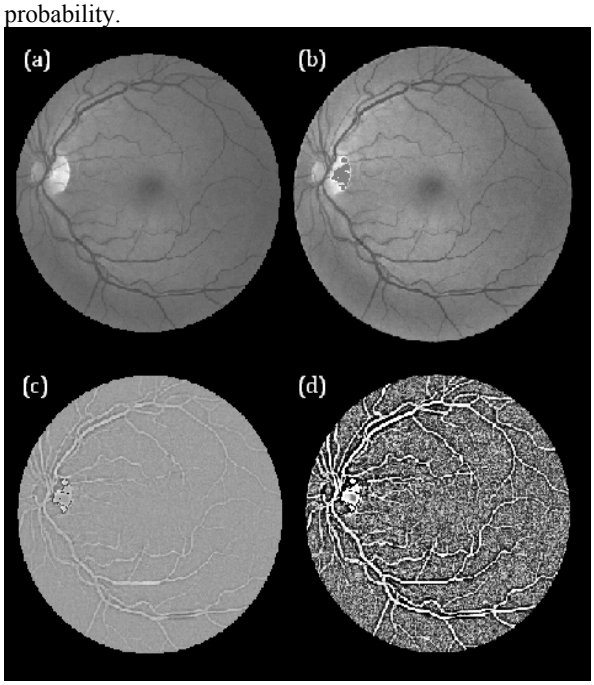

Fig. 1. (a) green channel of original image, (b) image with corrected mean and variance values, (c) background normalized image, and (d) contrast enhanced image.

In this paper, two HMMs are trained. The first model is trained using feature vectors randomly chosen from vessel pixels of manually segmented images. The second HMM is trained using non-vessel feature vectors. Feature vectors corresponding to every non-black pixel of the image is then presented to the HMMs. The output probabilities of the HMMs are represented in two likelihood maps (see figure 2). The maps are then compared pixel by pixel and the pixels whose vessel likelihoods are greater are segmented as vessel. The HMMs do not learn every possible form of blood vessel or background sequences. This is because the training features are chosen randomly.

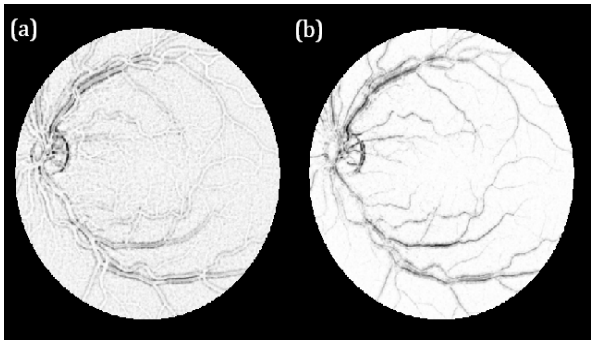

Fig. 2. Likelihood maps of (a) vessel and (b) non-vessel classes

Therefore, some sequences may never show up in the training step. This fact and the presence of noise and contrast variations, result in falsely detected blood vessels. To eliminate these, a post-processing step is necessary.

Morphological operations such as erosion and dilation can be very helpful in this step. First, the image is eroded using a structural element of a certain size. Next, connected components smaller than a specified size are removed. Finally the image is dilated using the same structural element that was used for eroding (see fig. 3.).

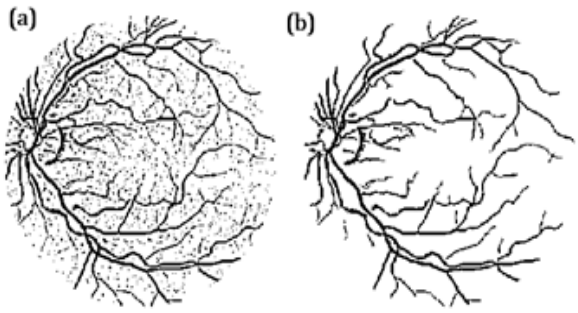

Fig. 3. Result of segmentation step before (a) and after (b) morphological post-processing.

5. Experimental Results

Our proposed method was tested on images of the publicly available DRIVE database [16].

The DRIVE database contains 40 color images of the retina, with 565 × 584 pixels and 8 bits per color channel, in LZW compressed TIFF format. The images were

originally captured from a Canon CR5 nonmydriatic 3 charge-coupled device (CCD) camera and were initially saved in JPEG-format. The database also includes binary images with the results of manual segmentation. These binary images have already been used as ground truth for performance evaluation of several vessel segmentation methods [16]. The 40 images were divided into a training set and a test set by the authors of the database. The results of the manual segmentation are available for all the images of the two sets. For the images of the test set, a second independent manual segmentation also exists.

Segmentation accuracy is selected as performance measure to compare our results with previous retinal vessel segmentation algorithms. The accuracy is estimated by the ratio of the total number of correctly classified points (sum of true positives and true negatives) by the number of non-black points in the image. The ground proof for computing the performance measures was the manual segmentation result. The values for the fraction of pixels erroneously classified as vessel pixels, false positive ratio (FPR), and the percentage of pixels correctly classified as vessel pixels, true positive fraction (TPR), are also reported. Two experiments were carried out to evaluate the performance of the proposed algorithm.

Table 1- Performance of Vessel segmentation methods

Method	accuracy	TPR	FPR
2nd human observer	0.9473	0.7761	0.0275
Mendonça et al. [1]	0.9452	0.7344	0.0236
Jiang et al. [4]	0. 8911	not reported	
Staal et al. [1]	0.9442	0.7194	0.0227
Niemeijer [1]	0.9417	0.6898	0.0304
Soares et al. [4]	0.9466	not reported	
First Experiment	0.9388	0.7342	0.0332
Second Experiment	0.9401	0.7492	0.0384

In our first experiment, a simple global thresholding algorithm was applied on the test set of the DRIVE database. The images were initially pre-processed according to the algorithm discussed in section III. Next a global thresholding algorithm was applied on the data set. The threshold level was empirically chosen to be 0.6. Then, we used the first two images of the training set to train two HMMs with 2000 feature vectors randomly selected for each class of pixels using the labels of the results of the manual segmentation. Each HMM had 9 hidden states and the emission probabilities were approximated by mixtures of 15 Gaussian distributions. Next, the images of the test set were segmented followed by erosion, small connected component removal, and dilation to remove the falsely detected vessel pixels. To further improve the results, the union of the globally thresholded image and the result of the HMM-based segmentation were chosen as the final result, whenever the accuracy of the segmentation step was not acceptable.

The second experiment aimed to decrease the computation time of the segmentation step. The whole algorithm was the same as for the first experiment, except that the output likelihood maps of the HMMs were formed in a 4-connected manner, i.e. starting from the first pixel of the image, its 4-connected neighbors had the same likelihood as the pixel. As a result, the computational time was reduced to 20% of the first experiment. Consequently, once a pixel is chosen to be in class 1, its neighbors are also chosen, resulting in a 4 connected segmented image which in turn, leads to better TPR since the number of small gaps between the connected components of the vessel class is reduced.
Table 1 shows the results of our experiments along with results of other authors.

A straightforward MATLAB implementation was used in our test. We used a 1.86 GHz Dell XPS M1210 laptop with 1 GB of RAM. A computer with better configurations along with efficient implementation in C++ will speed up the algorithm.

6. Discussion and Conclusion

Our method proved the ability of Hidden Markov Models in retinal blood vessel segmentation. The results are particularly comparable to previous works in which classifier-based methods are used. In [4], all 20 images of the training set of the DRIVE database were used for training, with about one million feature vectors. Whereas in our work, only 2000 feature vectors, extracted from 2 images, were used. Thus, the training time was reduced to approximately 1 hour, which is 9 times less than the time reported in [4]. However, the accuracy of these two methods are approximately equal. The simplicity of the feature vectors is also remarkable. Unlike [4] in which feature vectors are formed by Gabor transforms, our feature vectors are merely made of intensities of the pixels in a predefined neighborhood.

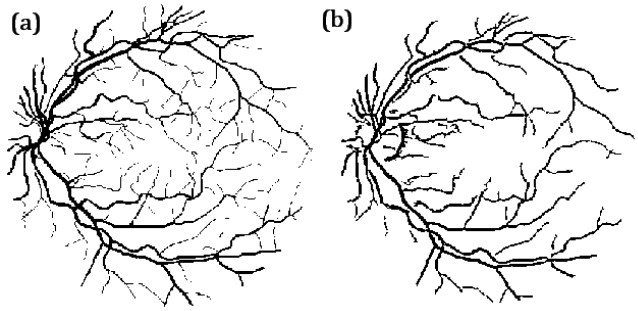

Fig. 4. Results from - (a) human observer, (b) our algorithm

The downside to our algorithm is that the a priori information that is used for training is not accurate. In fact, the TPR and FPR for the human observers are 0.8949 and 0.0610, respectively. Hence, some training

vectors are wrongly presented to the HMMs and the models learn the wrong combination of pixel intensities.

Another drawback of the proposed algorithm is that the EM algorithm used for the training of the models, may converge to local minima, resulting in poorly trained HMMs. In [5] the authors claim that under certain conditions, the MCE algorithm may achieve the global minimum with higher probability. The choice of number of hidden states and Gaussian distributions is also open to further investigation.

Acknowledgment

The authors would like to thank the authors of DRIVE database for making their data publicly available. We would also like to thank Kevin Murphy for making his HMM-toolbox for MATLAB available.

References

[1] M. Mendonça, and A. Campilho, "Segmentation of Retinal Blood Vessels by Combining the Detection of Centerlines and Morphological Reconstruction," *IEEE Trans. Med. Imag.,* vol. 25, no. 9, pp. 1200-1213, 2006.

[2] M. E. Martinez-Prez, A. D. Hughes, S. A. Thom, A. A. Bharath, and K. H. Parker, "Segmentation of Blood Vessels from Red-Free and Fluoroscein Retinal Images," *Medical Image Analysis,* vol. 11, pp. 47-61, 2007.

[3] K. A. Vermeer, F. M. Vos, H. G. Lemij, and A. M. Vossepoel, "A Model Based Method for Retinal Blood Vessel Detection," *Computers in Biology and Medicine,* vol. 34, pp. 209-219, 2004.

[4] J. V. B. Soares, J. J. G. Leandro, R. M. Cesar Jr., H. F. Jelinek, and M. J. Cree, "Retinal Vessel Segmentation Using the 2-D Gabor Wavelet and Supervised Classification," *IEEE Trans. Med. Imag.,* vol. 25, no. 9, pp. 1214-1222, 2006.

[5] M. Ibrahim, N. John, M. Kabuka, and A. Younis, "Hidden Markov Models-Based 3D MRI Brain Segmentation," *Image and Vision Computing*, vol. 24, pp. 1065-1079, 2006.

[6] J. Solomon, J. A. Butamen, and A. Sood, "Segmentation of Brain Tumors in 4D MR Images Using the Hidden Markov Model," *Computer Methods and Programs in Biomedicine,* vol. 84, pp. 76-85, 2006.

[7] B. Resch, "Hidden Markov Models," *A Tutorial for the Course Computational Intelligence*, [Online]. Available: http://www.igi.tugraz.at/lehre/CI

[8] P. Feng, Y. Pan, B. Wei, W. Jin, and D. Mi, "Enhancing Retinal Images by the Contourlet Transform," *Pattern Recognition Letters*, vol. 28, pp. 516 – 522, 2007.

[9] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating Blood Vessels in Retinal Images by Piecewise Threshold Probing of a Matched Filter Response," *IEEE Trans. Medical Imaging,* vol. 19, no. 3, pp. 203-219, 2000.

[10] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of Blood Vessels in Retinal Images Using Two-Dimensional Matched Filters," *IEEE Trans. Med. Imag.*, vol. 8, no. 3, pp. 263-269, 1989.

[11] Y. Zhang, M. Brady, and S. Smith, "Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation–Maximization Algorithm," *IEEE Trans. Med. Imag.,* vol. 20, no. 1, pp. 45–57, 2001.

[12] J. C. Rajapakse, and J. Piyaratna, "Bayesian Approach to Segmentation of Statistical Parametric Maps," *IEEE Trans. Biomed. Eng.,* vol. 48, no. 10, pp. 1186–1194, 2001.

[13] J. C. Rajapakse, and F. Kruggel, "Segmentation of MR Images with Intensity Inhomogeneities," *Image and Vision Computing,* vol. 16, pp. 165–180, 1998.

[14] J. C. Rajapakse, J. N. Giedd, and J. L. Rapoport, "Statistical Approach to Segmentation of Single-Channel Cerebral MR Images," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 176–186, 1997.

[15] M. Hammer, S. Leistritz, L. Leistritz, and D. Schweitzer, "Light Paths in Retinal Vessel Oximetry," *IEEE Trans. Biomed. Eng.,* vol. 48, pp. 592–598, 2001.

[16] M. Niemeijer, and B. van Ginneken, 2002 [Online]. Available:http://www.isi.uu.nl/Reseach/Databases/DRIVE/results.php