LEARNING LOCAL VESSEL APPEARANCE MODELS USING STRUCTURED SPARSITY

Vimal Singh *Ahmed H. Tewfik*

The University of Texas at Austin, Texas, USA

ABSTRACT

Vessel segmentation is a challenging task due to the complexity of vascular networks and limitations of imaging modalities to accurately capture thin structures. Analytical models based on geometric appearance and/or edge-based assumptions have been shown to be sub-optimal in segmenting vessels. In this paper, a novel approach for learning vessel appearance models from localized-vessel image patches is presented. This approach uses subspace clustering methods based on sparse representation of signals to identify local vessel appearance models from a training dataset. This paper also presents a hierarchical subspace clustering framework, which improves clustering speeds in presence of large number of subspaces. The preliminary results obtained for segmenting retinal vessel images using learned appearance models on the publicly available DRIVE database, yields an accuracy of 0.9268 at 0.1344 false detections and reduces the search space up to 80%.

***Index Terms*—** Sparse Representations, Structured Sparsity, Subspace Clustering, Hierarchical Subspace Clustering, Retinal Vessel Segmentation

1. INTRODUCTION

Vascular segmentation is a fundamental step for accurate visualization of vessels from complex datasets and quantification of pathologies [1]. For example, a careful analysis of change in retinal vasculature is required to diagnose diabetic retinopathy, a leading cause for adult blindness [2]. Similarly, an accurate location of portal vein, along with neighboring small vessels is needed to approximate Couinaud segments in the planning stage of a liver tumor resection surgery [3]. Segmentation of vessels is difficult as vascular networks are complex due to the inherent variability in vessel sizes, shapes, and their orientations. In addition, limitations of imaging modalities and presence of pathologies (aneurysms, stenoses) lead to poor contrast, noise, and artifacts at vessel boundaries. These make the segmentation task all the more challenging, especially for thin vessels. This paper, presents a radically novel approach to the vascular segmentation problem through learning of local vessel appearance models based on structured sparsity.

Most of the automatic vessel segmentation techniques in literature, assume a prior model either on the geometric appearance and/or on the intensity distribution of vessels [1]. Often, these models are limited in their ability to account for the complexity of vascular networks, which drastically impacts the performance of underlying segmentation methods. For example, the popular approach of [4], uses second-order derivative features to distinguish vessels as points exhibiting a locally prominent low curvature orientation direction and a plane of high intensity curvature, based on the eigen-analysis of the Hessian matrix at multiple scales. Hessian-based filters have been shown to be sensitive to local deformations and produce false negatives at locations such as: bifurcations, aneurysms, and stenoses [1]. Similarly, analytical model based techniques such as matched-filters, fitting of super-ellipsoids, and parallel edge-extraction schemes do not generalize well for real vessel appearance models [1]. As a consequence of the failure of aforementioned analytical models, machine learning based methods that select features to accurately distinguish between vessel and non-vessel pixels have become increasingly attractive [2]. However, as these methods rely on training classifiers on features generated from a vessel-approximation model(s), they can be sub-optimal in accounting for complex vessel appearances.

In this paper, a novel approach for learning local vessel appearance models from a training dataset based on structured sparse representations is presented. The proposed approach relies on the following assumption: image patches local to vessels for a specific application (imaging modality + anatomical region) would 'live' in multiple low-dimensional subspaces. Identification of these subspaces would allow representing vessel-containing patches using sparse coefficient vectors, with low approximation errors. Hereafter, the space spanned by all identified subspaces is referred to as, the structured dictionary. Note, the dictionary is non-parametric in form and implicitly models all local vessel appearances in the training data. In this work, a dictionary learning method based on subspace clustering is used [5]. A second major contribution of this paper is a novel hierarchical subspace clustering framework. This framework is aimed at reducing the computational costs of sparse coding (clustering) in large scale problems such as segmenting vessels using learned appearance models. An intrinsic advantage of the presented approach is reduction of the search space where vessel detections are evaluated. Each learned appearance model is associated with a segmentation mask, which is used along with a tracing algorithm to evaluate detections only in the neighborhood of the vascular tree instead of the full image, thus reducing the search space. The preliminary results, obtained for segmenting retinal vessel images using learned local-vessel appearance models are better than the state-of-the-art method, yielding an accuracy of 0.9268 at 0.1344 of false detections, while reducing the search space up to 80%.

The rest of this paper is organized as follows. The proposed approach for learning local vessel appearance models is presented in section 2. The hierarchical subspace clustering framework is presented in section 3. Section 4, summarizes the results of segmenting retinal vessels on the publicly available DRIVE [6] database using learned models. Finally, section 5 concludes the paper citing future directions.

2. LEARNING LOCAL VESSEL APPEARANCE MODELS

In this section, the proposed approach to learn local vessel appearance models based on sparse representation of image patches is presented. Let, $\{x_i\}_{i=1}^{K}$ be N-dimensional K vectorized training image patches, which are fed into the Iterative Subspace Identification (ISI) algorithm to identify the low-dimensional subspaces where the

training set resides [5]. The ISI algorithm recursively extracts the dominant subspaces, through sparse approximation of a randomly selected vector in terms of the training data. For more details on the ISI algorithm, the reader is referred to [5]. Let, $\{\mathbf{A}_i\}_{i=1}^{J}$ denote the J subspaces identified by the ISI algorithm. Each subspace $\mathbf{A}_i$, represents a local vessel appearance model conforming to a unique vessel shape and an intensity distribution. Note, various subspaces can represent similar geometric shapes, however, they will differ in terms of their intensity distributions. Let, $\mathbf{D}$ denote the structured dictionary obtained by concatenating all subspaces $\{\mathbf{A}_i\}_{i=1}^{J}$. Then, a vectorized patch $\mathbf{Y}$, residing in the i^{th} appearance model $\mathbf{A}_i$, can be represented using a unit-block sparse vector p as :

$$\begin{aligned} \mathbf{Y} \cong \quad \mathbf{A}_i * c_i \quad &= \quad \mathbf{D} * p \\ p = \quad [\ \underline{0}^T \quad \ldots \quad {c_i}^T \quad \ldots \quad \underline{0}^T\ ]^T & \end{aligned} \tag{1}$$

where, c_i are its non-zero description coefficients in $\mathbf{A}_i$. During the vessel detection stage, for a candidate patch $\mathbf{Y}$, its structured sparse representation p in (1), is evaluated by solving the following over-determined system :

$$\begin{aligned} &\min_{p} \quad \| \mathbf{Y} - \mathbf{D} * p \|_2 + \xi \\ &s.t. \| p \|_{0,B} \leq 1 \end{aligned} \tag{2}$$

where, $\| p \|_{0,B} \leq 1$ represents the unit-block sparsity constraint, and ξ represents the euclidean approximation error. The process of extracting p, is more commonly known as the sparse coding problem. Problem (2), can be reformulated as solving J least square (LS) problems and picking the solution yielding the minimum error ξ. For more details on the sparse coding algorithm, the reader is referred to [7]. To summarize, given a vessel appearance dictionary D, a test image patch $\mathbf{Y}$, is classified as a vessel-containing patch, only if its sparse approximation error (ξ) is less than a pre-determined threshold.

3. HIERARCHICAL SUBSPACE CLUSTERING

In the previous section, a mathematical formulation to learn vessel appearance models in terms of a structured dictionary D is presented. Learned models simultaneously account for various vessel shapes and intensity distributions. In addition, due to the large scale nature of the patch-based problem the dictionary D is an over-complete basis, i.e., $J \gg N$. During the vessel detection (sparse coding) stage, solving J LS problems for each candidate patch $\mathbf{Y}$ is computationally expensive. In this section, a hierarchical subspace clustering framework focused at improving sparse coding speeds in presence of large number of subspaces (appearance models) is presented. This framework is inspired from divide-and-conquer algorithms such as the bisection method for root finding.

In this paper, hierarchical clustering corresponds to the identification of the hierarchy in the structural information of subspaces at any given level. Given, a set of subspaces $\{\mathbf{S}_{L,i}\}_{i=1}^{I}$ at level L, a subset of these, which are at least a measure M_{thres} away from each other, are selected as hierarchical subspaces at level $L+1$. In a subsequent step, for each subspace $\{\mathbf{S}_{L+1,q}\}_{q=1}^{Q}$ at level $L+1$, a link is formed with all subspaces at level L, which are closer than an euclidean distance E_{thres} to $\mathbf{S}_{L+1,q}$. M_{thres} is an application dependent measure. For extracting vessel models, structural similarity index measure (SSIM) is used. SSIM provides a measure to assess the closeness of structural information, it compares the local pixel-intensity patterns that are normalized for luminance and contrast [8]. Using SSIM guarantees that various structurally different subspaces are retained even at the highest hierarchical level. The algorithmic flow of the hierarchical clustering framework is as follows:

Initialization: $Ind = \{1, 2, \ldots, I\}, \quad p = 1$

- Compute the pairwise similarity measure matrix $\mathbf{M}_L = [m_{i,j}]_{I \times I}$

Step 1: $\mathbf{S}_{L+1,p} = \mathbf{S}_{L,Ind(1)}$

- $rInd = \{Ind \mid \mathbf{M}_L\,(1, Ind) \geq M_{thres}\}$
- $Ind = Ind \setminus rInd$
- $\mathbf{M}_L = \mathbf{M}_L\,(Ind : Ind)$

Step 2: $p = p + 1$, go to Step 1 until $Ind = \emptyset$

Step 3: Let, $|\mathbf{S}_{L+1}| = Q$, $Ind = \{1, 2, \ldots, I\}$

- for each $q = 1, 2, \ldots, Q$
  $\mathbf{C}_{L+1,q} = \{Ind \mid \|\mathbf{S}_{L+1,q} - \mathbf{S}_{L,Ind}\|_2 \leq E_{thres}\}$

where, $|\ .\ |$ represents the cardinality of a set, and $\mathbf{C}_{L,q}$ contains indexes of subspaces at level L which are close to the subspace $\mathbf{S}_{L+1,q}$ in the euclidean sense. Detection using trained hierarchical subspaces, requires solving LS problems for all subspaces, only at the highest level. At lower levels, LS problems are solved, only for subspaces located closely to the detected subspace at the next higher level, which are identified by $\mathbf{C}_{L,q}$. In effect, the maximum number of LS problems to be solved for a trained hierarchical subspace cluster is

$$|\mathbf{S}_{L_{high}}| + \sum \max_{p} |\mathbf{C}_{l,p}| \ \forall\ l = 2, \ldots, L_{high} \tag{3}$$

where, L_{high} is the highest hierarchical level.

4. RETINAL VESSEL SEGMENTATION

The proposed approach is used to generate local vessel appearance models for the purpose of segmenting retinal images. It is tested on the publicly available DRIVE database of retinal images [6]. The database consists of 40 24-bit RGB-images captured by a Canon CR5 camera at 45° FOV, with a spatial resolution of 565 x 584 pixels. For our experiments, only the green channel information is used as it provides the best contrast between vessel and background pixels [2]. The manual segmentations available in the database are used as reference segmentations. The 40 images were divided by the authors of the database, into a training set and a test set of 20 images each [6]. For learning local vessel appearance models, $5\text{x}10^4$ image patches of size 15 x 15 pixels centered on the vessel centerlines were extracted randomly from the training images. The ISI method identified 1447 low-dimensional subspaces. Using 947 of the largest subspaces (cluster size $>$ 4), a 3-level hierarchical subspace framework is trained. Table 1 summarizes the details of the hierarchical subspace clusters obtained along with the parameters used. Fig. 1 shows the first vector from each subspace at the hierarchical level $L = 3$ as block images, along with the corresponding segmentation masks. Intensity of each block-image has been shifted linearly for visual enhancement.

4.1. Segmenting The Retinal Images

Using learned local vessel appearance models obtained in the previous step in the form of hierarchical subspaces, each image in the test set is segmented. Segmentation is initialized by randomly selecting

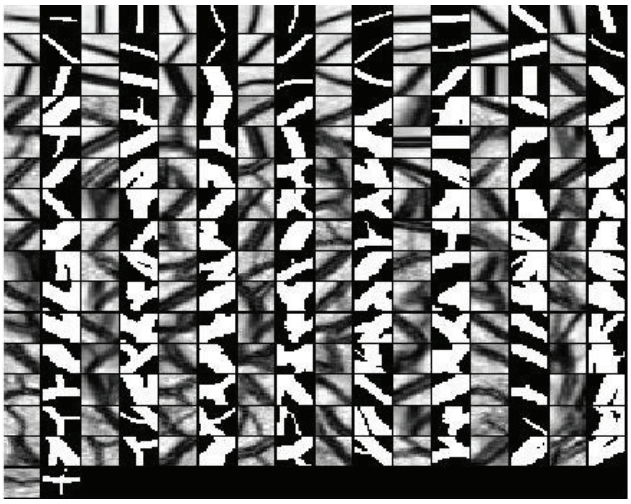


Fig. 1: Local vessel appearance models at the hierarchical level $L = 3$ as block images for the DRIVE database, along with the corresponding segmentation masks.

Level (L)	# Subspaces ($\lvert\mathbf{S}_L\rvert$)	Cluster Size (max $\lvert\mathbf{C}_{L,q}\rvert$)	SSIM Thres. (M_{thres})
1	947	-	0.80
2	428	51	0.65
3	121	62	-

Table 1: Details of the hierarchical subspace clusters derived for retinal images of the DRIVE database.

$2\text{x}10^3$ pixel locations as centers of 15 x 15 image patches for detection. A pixel location is detected as a vessel center based on the criterion described in section 2. Beyond a positive detection at a pixel, a basic tracing algorithm uses the associated segmentation mask (with the appearance model) to predict the next set of pixels, where vessel detections will be evaluated. This procedure is iterated until no further segmentation is achieved or a set number of maximum iterations is reached. Theoretically, to detect the complete vascular tree, positive detections are needed only at the vessel centerlines. For the DRIVE database, $\sim 3\%$ of pixels correspond to centerline pixels for retinal vessels. Using the aforementioned segmentation procedure, detections are performed at $\sim$16.75% of pixel locations. This detection rate is low, however, as detections are localized in the neighborhood of the vascular tree, the segmentation procedure exhibits high positive detection rates. For evaluation and comparison of results, an image with confidence measure of detections is generated as follows: for each detected vessel-centerline pixel, all pixels corresponding to the detected segmentation mask are updated with a confidence value defined as $(1 - \xi)$.

4.2. Results

The performance of the proposed approach is compared with those of: 1) Soares *et. al.* [2], and 2) Frangi *et. al.* [4]. To the best of our knowledge, the method of Soares *et. al.* is the current state-of-the-art method. It trains a Bayesian classifier with class-conditional probability density functions (likelihoods) described as Gaussian mixtures, on feature vectors extracted using 2D-gabor wavelets at multiple scales. For implementation details the reader is referred to [2]. The approach of Frangi *et. al.*, based on multi-scale eigen-analysis of the Hessian-matrix is often used as the minimum benchmark for evaluating vessel segmentation algorithms. Hessian analysis is evaluated at single-step scales between 1 to 10 voxels.

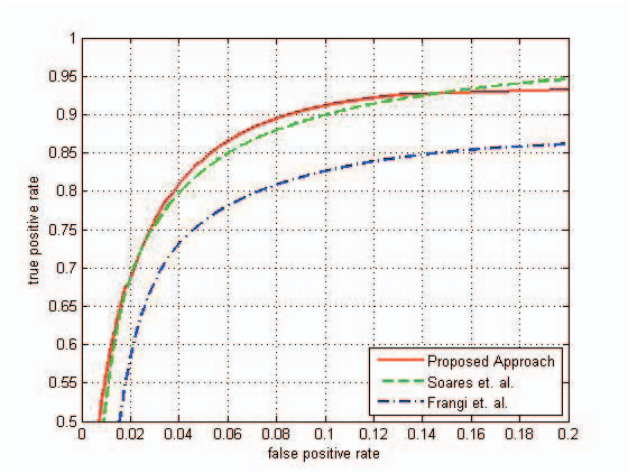


Fig. 2: ROC analysis on the output of vessel segmentation algorithms for the test set of the DRIVE database.

To evaluate the performance of segmentation algorithms, receiver operating characteristic (ROC) analysis is carried out on the probability maps (confidence measures) generated for all images in the test set. Fig. 2 compares the ROC curves for all the three approaches. In addition, to quantitatively compare the performances, following measures from the ROC analysis are used : 1) True Positive Rate (TPR) at a specified False Positive Rate (FPR), and 2) Area under the ROC curve (AUC). Table 2 compares the quantitative measures for all three methods, at the accuracy of the proposed approach. Accuracy is defined as the fraction of correctly detected pixels. The proposed approach performs better than the state-of-the-art method in terms of marginally higher true detections at the same false detection rate, however, the AUC measure is marginally lower. This is acceptable, as detections are performed only at $\sim$16.75% of the total image and a basic vessel tracing algorithm is employed. Note that the reference segmentations in the database are not ground truth but manual segmentations [6]. Thus, the improvement in segmentation performance suggests that learned appearance models are more tuned to the training data than the analytical features used in Soares *et. al.* [2]. Fig. 3 shows a pair of segmentation results obtained for the DRIVE database. Confidence measure images and manual segmentations are also placed alongside for comparison.

	TPR(%)	FPR(%)	AUC
Proposed Approach	**92.68**	**13.44**	**0.9481**
Soares *et. al.*	92.27	13.44	0.9614
Frangi *et. al.*	84.55	13.45	0.8961

Table 2: Quantitative performance comparison of vessel segmentation algorithms on the DRIVE database. TPR - True Positive Rate, FPR - False Positive Rate, AUC - Area under the ROC curve.

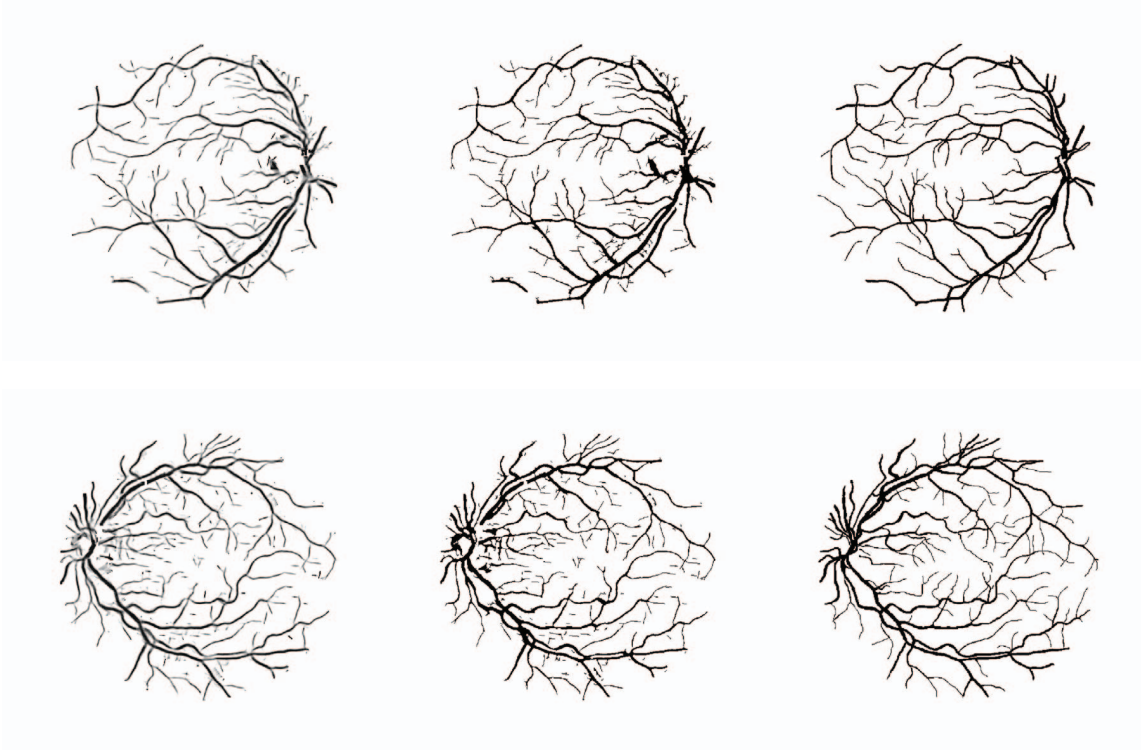


Fig. 3: Retinal vessel segmentations produced by the proposed approach for two test images of the DRIVE database. Left Column: Confidence Measure Image (CMI) representing the approximation error in the learned structured dictionary. Middle Column: Final segmentations obtained by thresholding the CMI. Right Column: Reference segmentations available in the database.

5. CONCLUSION AND FUTURE WORK

In this paper, a local vessel appearance model learning technique based on sparse representations of image patches is proposed. The novelty of the technique lies in its assumption that localized-vessel image patches for an application live in multiple low-dimensional subspaces. Identification of these dominant subspaces using sparse signal representations allows for extracting non-parametric vessel appearance models. These models inherently capture the complexity of the underlying vascular networks and account for deviations associated with the pathologies. A second contribution of this paper is a hierarchical subspace clustering framework, which reduces the computational costs of sparse coding in large scale clustering problems. The preliminary results obtained using learned appearance models for retinal vessel segmentation on the DRIVE database are comparably better than the state-of-the-art method. An extension of the presented approach for vessel segmentation in volumetric Computed Tomography and Magnetic Resonance images forms the future scope of this project.

6. REFERENCES

[1] D. Lesage, E.D. Angelini, I. Bloch, and G. Funka-Lea, "A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes," *Medical Image Analysis*, vol. 13, no. 6, pp. 819 – 845, 2009.

[2] J.V.B. Soares, J.J.G. Leandro, R.M. Cesar, H.F. Jelinek, and M.J. Cree, "Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification," *Medical Imaging, IEEE Transactions on*, vol. 25, no. 9, pp. 1214 –1222, sept. 2006.

[3] B. Reitinger, A. Bornik, R. Beichel, and D. Schmalstieg, "Liver Surgery Planning Using Virtual Reality," *Computer Graphics and Applications, IEEE*, vol. 26, no. 6, pp. 36 –47, nov.-dec. 2006.

[4] A. Frangi, W. Niessen, K. Vincken, and M. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Interventation MICCAI98*, vol. 1496 of *Lecture Notes in Computer Science*, pp. 130–137. 1998.

[5] B.V. Gowreesunker and A.H. Tewfik, "Learning sparse representation using iterative subspace identification," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 3055 –3065, june 2010.

[6] J. Staal, M.D. Abramoff, M. Niemeijer, M.A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *Medical Imaging, IEEE Transactions on*, vol. 23, no. 4, pp. 501 –509, april 2004.

[7] D. Wang and A.H. Tewfik, "In vivo tracking of 3D organs using spherical harmonics and subspace clustering," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, nov. 2009, pp. 817 –820.

[8] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600 –612, april 2004.