SUBJECT: MI102 - I. Išgum and B. Landman; Image Processing

1. TITLE: Blood Vessel Segmentation in Narrow Band Imaging Bronchoscopic Video

2. AUTHORS:

William E. Higgins[1]	Saptarashmi Bandyopadhyay[1]	Vahid Daneshpajooh[1]
weh2@psu.edu	sapta.band59@gmail.com	vvd5140@psu.edu
Danish Ahmad[2]	Jennifer Toth[2]	Rebecca Bascom[2]
dxa557@psu.edu	jtoth@hmc.psu.edu	rbascom@hmc.psu.edu

Penn State University, School of Electrical Engineering and Computer Science,[1] College of Medicine[2]
University Park and Hershey, PA; Phone: 814-865-0186, Fax: 814-863-5341

3. PRESENTATION PREFERENCE: poster presentation

4. PRINCIPAL AUTHOR'S BIOGRAPHY:

William E. Higgins received the B.S. degree in electrical engineering from the Massachusetts Institute of Technology and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois, Urbana-Champaign. He held positions previously at the Honeywell Systems and Research Center, Minneapolis, MN, and the Mayo Clinic, Rochester, MN. He is currently a distinguished professor in the School of Electrical Engineering and Computer Science and Biomedical Engineering at the Pennsylvania State University. His research interests are in multidimensional medical image processing and image-guided intervention (IGI) systems.

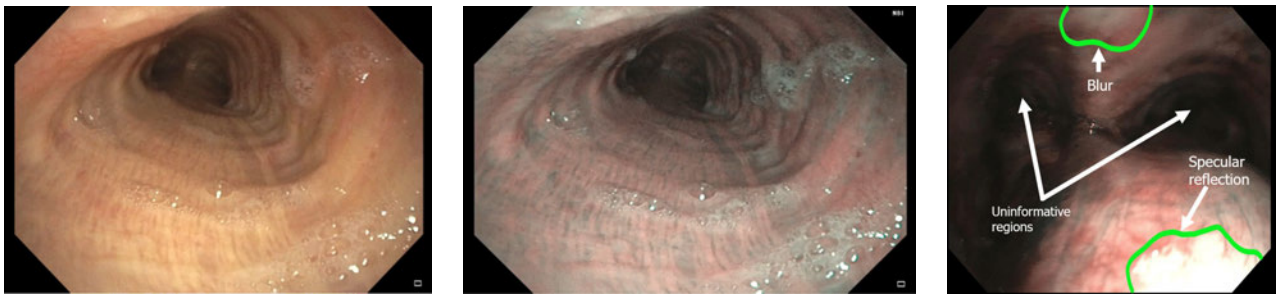
5. ABSTRACT: Lung cancer is the leading cause of cancer fatalities in the world. A recent trend has begun to focus on the idea of using bronchoscopy for early detection of suspect cancerous lesions developing along the airway walls. Because standard white-light bronchoscopy has insufficient sensitivity in locating suspect lesions, researchers are turning to the promising modality referred to as narrow band imaging (NBI). NBI bronchoscopy has the advantage of highlighting the blood vessels contained in the lung mucosa, with cancer lesions exhibiting excessive vessel structure. Unfortunately, the task of locating lesions, and their vessel patterns, in an NBI bronchoscopy video stream proves to be exceedingly tedious for the physician. We present automatic methods for enhancing and segmenting the major blood vessels depicted in NBI bronchoscopic video. Results with ground-truth data indicate that our methods can achieve superior results to an existing vessel segmentation method. We also consider a preliminary application of deep learning to this task. While the approach gives low sensitivity compared to the other approaches, it achieves high specificity and accuracy.

6. KEYWORDS: bronchoscopy, lung cancer, narrow band imaging, vessel segmentation, video processing

Topic areas for MI102 review: computer vision, machine learning and pattern recognition, segmentation methodologies

7. SUMMARY:

Description Of Purpose: Lung cancer is the leading cause of cancer fatalities in the world. To effectively manage lung cancer, it is important to detect the disease early. A recent trend has begun to focus on the early detection of suspect cancerous lesions that develop along the airway walls (lung mucosa), with bronchoscopy being the tool used to locate these lesions [1]. Standard white light bronchoscopy has been considered for this purpose, but researchers have found the modality to be insufficiently sensitive for lesion detection. A state-of-the-art modality found to be far more promising for lesion detection is narrow band imaging (NBI) bronchoscopy (Fig. 1b) [1, 2, 3]. NBI bronchoscopy uses filtered narrow band light to illuminate the airway walls. The illumination wavelengths passed by NBI match the absorption spectrum peaks of hemoglobin, thereby making the vascular structure of the airway mucosa much more visible than in white light bronchoscopy. Using NBI bronchoscopy, the physician interactively searches for lesions by carefully scanning the airways and noting abnormal vessel patterns appearing in the video. Unfortunately, given the enormity of the bronchoscopic video stream and the subtleties entailed in deciding what constitutes an airway lesion, the task of locating vessel patterns is challenging, time consuming, and error prone [2]. The task is also complicated by corruptions degrading the images, such as saliva, specular reflections, blurred regions, blood, and extraneous uninformative regions (Fig 1c).



(a) White light bronchoscopy (b) Narrow band imaging (c) Image corruptions

Figure 1: Bronchoscopic imaging modalities and possible corruptions. (a-b) depict the same location for different modalities.

Researchers have evaluated the use of NBI bronchoscopy for mucosal lesion detection [3]. But this effort requires painstaking manual assessment of video frames, with no means for delineating lesion regions and their associated vascular structures [2]. On another front, much research has applied automated image analysis to the extraction of blood vessels in medical images. A classic work on blood vessel enhancement by Frangi *et al.* considers multi-scale second-order structure based on the local Hessian across an image [4]. Subsequent work by Jerman *et al.* extended this approach with a blob enhancement filter employing a modified volume ratio derived from the eigenvalues of a 3×3 Hessian matrix [5]. For endoscopy, Moccia *et al.* considered the problem of tissue classification as depicting in narrow band laryngeal video [6]. No research, however, has focused on enhancing and segmenting the blood vessels as seen in NBI bronchoscopic video of the airways. *We present automatic methods for enhancing and segmenting the major blood vessels depicted in NBI bronchoscopic video.*

Methods: Using a 3-channel RGB color NBI video frame as input, our method consists of three steps:

1. Image preprocessing converts the RGB input into the HSV form (hue, saturation, value), with anisotropic diffusion filtering the value channel to reduce noise.
2. Vessel enhancement applies a Hessian-based operation to the preprocessed image to produce a filtered image highlighting significant vascular structures.
3. Segmentation applies an image mask, derived through a color clustering operation, to isolate the desired vessel pixels while eliminating non-vascular regions.

We elaborate more on these steps below.

The vessel enhancement filter is the key operation in our method. We draw upon two distinct filters for our work, motivated in part by the work of Jerman *et al.* [5]. To begin, we note that the vessel structure appearing in 2D NBI video frames are generally elongated, while rounded structures and bifurcations also appear. An analysis of Jerman's enhancement filter, which uses a volume ratio for 3D images, led us to conclude that their filter is not directly suitable for our 2D imaging scenario [7]. Given 2D images, the structures of interest are better described by a 2×2 Hessian matrix having 2 eigenvalues, instead of the 3×3 Hessian matrix and 3 associated eigenvalues required to properly model vascular structures as appearing in 3D images.

We propose two modified vessel enhancement filters. A filter begin by applying a multi-scale Gaussian filter on the preprocessed value channel image to give a series of images $\{I_\sigma\}$ filtered at different scales σ as specified by the Gaussian kernel. Next, for each filtered image I_σ, we compute the 2D Hessian matrix

$$H_\sigma = \begin{pmatrix} \frac{\partial I_\sigma^2}{\partial x^2} & \frac{\partial I_\sigma^2}{\partial x \partial y} \\ \frac{\partial I_\sigma^2}{\partial y \partial x} & \frac{\partial I_\sigma^2}{\partial y^2} \end{pmatrix}$$

at each image pixel (x, y), as described by Frangi *et al.* [4]. The two eigenvalues, λ_1 and λ_2, for each Hessian H_σ are then found, with $\lambda_2 > \lambda_1$. This leads to two distinct vessel enhancement filters: proposal 1, given by

$$\nu_P^1 = \begin{cases} 0, & \text{if } \lambda_2 \leq 0 \text{ or } \lambda_\rho \geq 0 \\ 1, & \text{if } \lambda_2 \geq \lambda_\rho > 0 \\ \lambda_2(\lambda_\rho - \lambda_2)\left[\frac{2}{\lambda_\rho}\right]^2, & \text{otherwise} \end{cases} ; \tag{1}$$

and proposal 2, given by

$$\nu_P^2 = \begin{cases} 0, & \text{if } \lambda_2 \leq 0 \text{ or } \lambda_\rho \geq 0 \\ 1, & \text{if } \lambda_2 \geq \lambda_\rho > 0 \\ e^{\frac{-R_B^2}{2\beta^2}} \lambda_2(\lambda_\rho - \lambda_2)\left[\frac{2}{\lambda_\rho}\right]^2, & \text{otherwise} \end{cases} . \tag{2}$$

In (1-2), $R_B = \lambda_1/\lambda_2$ as defined in Frangi *et al.* [4] and

$$\lambda_\rho(\sigma) = \begin{cases} \lambda_2, & \text{if } \lambda_2 > \tau \cdot \max_x \lambda_2(x,\sigma) \\ \tau \cdot \max_x \lambda_2(x,\sigma), & \text{if } 0 < \lambda_2 \leq \tau \cdot \max_x \lambda_2(x,\sigma) \\ 0, & \text{otherwise} \end{cases}, \tag{3}$$

where $\tau = 0.35$ [7]. Quantity ν_P^2 adds an exponential term, as used by Frangi *et al.*, over ν_P^1. This helps suppress rounded vascular structures. The final vessel enhanced image, as given by ν_P^1 or ν_P^2, is derived by applying λ_ρ over all scales σ and retaining the maximum over all σ.

Color clustering next masks out non-vascular regions by using the original RGB video frame's color values. Color clustering applies k-means clustering in 3D RGB color space to produce k=8 distinct pixel clusters. This leads to clusters that contain the bulk of the pixels constituting unwanted specular reflections, saliva, and other non-vessel regions. Thus, we use the union of preselected clusters to create two binary masks: a conservative mask, made up of two clusters; and a liberal mask, made up of two additional clusters. This then leads to two distinct vessel extraction methods, rejecting varying amounts of unwanted regions: a liberal method, drawing upon ν_P^1 and the liberal mask, and a conservative method, drawing upon ν_P^2 and the more restrictive conservative mask. The final segmented image is derived by applying the designated mask.

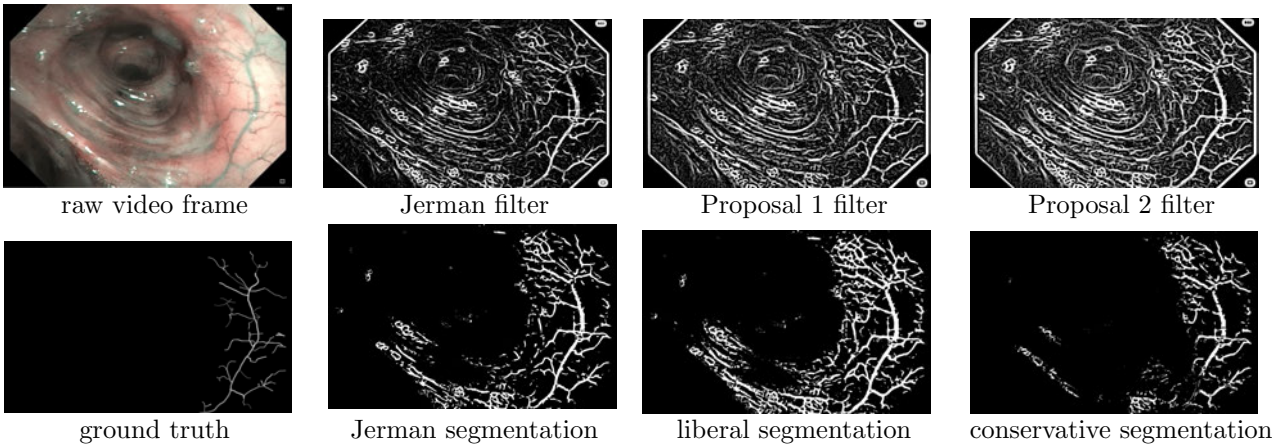


Figure 2: Filtering and segmentation results for frame 1367 in case 21405.168

Results: We used 4 NBI video frames from 3 different patient studies to pretrain our methods. For each method, we first labeled the ground truth vessels of interest. We then ran parameter sensitivity tests to preset τ and the clusters constituting the masks derived by color clustering [7]. Over the four training images, we achieved the following aggregate sensitivity, specificity, and accuracy metrics, respectively [5]: Jerman — 0.636, 0.961, 0.954; liberal method — 0.805, 0.923, 0.919; and conservative, 0.530, 0.943, 0.939. For the training images, the liberal method has the highest sensitivity, while the conservative method has specificity and accuracy comparable to the Jerman method. Fig. 2 illustrates example filtered and segmented results for all methods.

We next tested the tuned methods on 15 other test NBI video frames (7 patients), with ground truth vessels of interest labeled. Table 1 gives aggregate segmentation results. The liberal method has the highest sensitivity, while the conservative method has the highest specificity and accuracy. Notably, our methods show higher sensitivity than the Jerman method, while the Jerman method exceeds the liberal method's specificity and accuracy.

To measure the ability of the methods to reject non-vascular regions, we also conducted a test with 13 NBI video frames deemed not to contain any significant vessels; hence, for these images, the ideal segmentation result should contain no output pixels. A comparison of the percentage of non-zero segmented pixels for the 13 non-vascular and 15 vascular images were as follows: non-vascular, 0.39% $\pm$ 0.002%; vascular, 8.71% $\pm$ 4.28%.

We finally considered a preliminary application of deep learning to the problem. We applied the U-net architecture, which has been widely used for semantic segmentation tasks in medical imaging [8]. Combining the 19 ground-truthed images used for training and testing the vessel enhancement filters, we now use 15 images for training the network and the remaining 4 for testing, to give 256×256 network output images. Since the

training set is still small, we utilized data augmentation techniques (e.g., random cropping along with rotation, scale, translation, etc.) to enhance the size of our training set. We used the Adam optimizer at a learning rate of 0.001 and binary cross-entropy as the loss function. We trained the model for 100 epochs with mini-batch size 15. We obtained the following aggregate segmentation sensitivity, specificity, and accuracy results, respectively: training set — 41%, 99%, 98%; test set — 21%, 99%, 97%. While the approach's sensitivity is especially low, it achieved superior specificity and accuracy results to the previous methods. (We strongly caution that the two sets of results are not strictly speaking comparable, but do offer promise to the deep learning approach!)

Segmentation Method	Sensitivity	Specificity	Accuracy
liberal	0.507 ± 0.116	0.883 ± 0.058	0.877 ± 0.057
	[0.352,0.694]	[0.808,0.964]	[0.832,0.955]
conservative	0.418 ± 0.093	0.958 ± 0.03	0.937 ± 0.031
	[0.299,0.625]	[0.889,0.986]	[0.865,0.978]
Jerman	0.357 ± 0.137	0.934 ± 0.034	0.928 ± 0.036
	[0.162,0.654]	[0.88,0.983]	[0.861,0.976]

Table 1: Aggregate segmentation performance for 15 test images.

New or Breakthrough Work: Automated methods have been proposed for the first time for extracting the vasculature depicted in NBI bronchoscopic video frames.

Conclusion: A multi-scale Hessian-based filtering formulation has been devised to enhance and segment the blood vessels in NBI bronchoscopic image frames. Our methods exceed the performance of vessel extraction methods proposed for other domains. In addition, we demonstrated the potential of deep learning for the vessel extraction problem; yet, further effort is needed in producing ground truth data, perhaps by using our new filtering methods as a base for producing refined ground truth data over many frames. Our final paper will detail the derivation of the enhancement filters and elaborate further on the results.

Submission History: This work has not been submitted for publication or presentation elsewhere, nor is it currently in the process of being submitted elsewhere.

References

[1] Inage, T., Nakajima, T., Yoshino, I., and Yasufuku, K., "Early lung cancer detection," *Clin. Chest Med.* **39**, 45–55 (2018).

[2] Dumas, C., Fielding, D., Coles, T., and Good, N., "Development of a novel image-based program to teach narrow-band imaging," *Ther. Adv. Respir. Dis.* **10**, 300–309 (Aug 2016).

[3] Zaric, B., Perin, B., Stojsic, V., et al., "Relation between vascular patterns visualized by narrow band imaging (NBI) videobronchoscopy and histological type of lung cancer," *Medical Oncology* **30**, 1–8 (March 2013).

[4] Frangi, A. F., Niessen, W. J., Vincken, K. L., and Viergever, M. A., "Multiscale vessel enhancement filtering," in [*Int. Conf. Med. Image Comput. Computer-Assisted Interv.*], Wells, W. M., Colchester, A., and Delp, S., eds., 130–137, Springer, Berlin, Heidelberg (1998).

[5] Jerman, T., Pernus, F., Likar, B., and Spiclin, Z., "Enhancement of vascular structures in 3D and 2D angiographic images," *IEEE Trans. Med. Imaging* **35**, 2107–2118 (April 2016).

[6] Moccia, S., Momi, E. D., Guarnaschelli, M., Savazzi, M., Laborai, A., Guastini, L., Peretti, G., and Mattos, L. S., "Confident texture-based laryngeal tissue classification for early stage diagnosis support," *J. Med. Imaging* **4**, Sept.

[7] Bandyopadhyay, S., *Vessel analysis in narrow band imaging bronchoscopic video*, Master's thesis, The Pennsylvania State University, Department of Computer Science and Engineering (2020).

[8] Ronneberger, O., Fischer, P., and Brox, T., "U-net: Convolutional networks for biomedical image segmentation," in [*Int. Conf. Med. Image Comp. Computer-assisted Interv.*], 234–241 (2015).