Saumya Gurbani

EN.600.446 – CIS II

Paper Critical Review

Paper: Ascari, L.; Bertocchi, U.; Laschi, C.; Stefanini, C.; Starita, A.; Dario, P.; , "A segmentation algorithm for a robotic micro-endoscope for exploration of the spinal cord," *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on* , vol.1, no., pp. 491- 496 Vol.1, 26 April-1 May 2004. doi: 10.1109/ROBOT.2004.1307197.

Background and Relevence to Project:

In this paper, Ascari et al present an algorithm for segmenting images taken from an endoscope inserted into the spinal cord. Using endoscopic images, as opposed to preoperative CT or MRI imaging, has advantages in that it is able to provide more detailed information regarding features of the sub-arachnoid spinal space. The authors mention that such preoperative techniques are "not able to provide sufficient useful information for diagnoses." Visual analysis through an endoscope can help extract location details of components such as the lumen, nerves, and blood vessels; such information is critical for determining safe insertion paths for surgery.

The algorithm uses a combination of image luminance, hue, and saturation in order to separate features from each other. While our project is for the imaging of the cochlear canal, we believe that this algorithm can be adapted for our use. The endoscopic images shown in the paper are similar in nature (in regards to the noise, coloring, and overall quality) to the images we have been able to capture through our borescope exploration of a phantom cochlea. We would use this algorithm to aid in safe-path planning for cochlear implant surgery.

Algorithm Summary

The first task of the algorithm is to find the threshold for the dark regions of the image, a value which will be used later to find lumens. This step begins by applying adaptive histogram pre-processing to the image. This involves going pixel by pixel through the image and sorting pixels by their luminance. A five pixel sliding window is applied to smoothen out the histogram; a histogram function is defined as the number of pixels in each luminance value. Next, the local minima are found in the averaged luminance histogram; a conservative bound is used since it is safer to underestimate the lumen. If overestimated, a surgical tool may make contact with an area thought to be lumen but actually is sensitive tissue. The next step is to reduce the number of luminance minima; this is done by choosing a single minimum if two minima occur close to each other in luminance values. The lower minimum is always kept unless the higher minimum has less than 85% of the pixels as the lowest one (equation (1)); again, this is used for safety to underestimate rather than overestimate lumen size. Finally, if there is a peak between two minima (as defined by equation (2) in the paper), then again only one of minima bounding the peak is kept. The same formula (equation (1)) is used to determine which one.

The next step is to find the lumen in the image. The image is binarized using the highest of the minima who value is less than 65, a value that the paper mentions was optimal for most endoscopic images. Small black and white areas are filtered away so as to remove noise from the image. Darker pixels which form homogenous regions (referred to as “blobs”) represent the lumen of the spinal cord. If no such blobs are found, then a search for a “dirty lumen” must be computed.

For a dirty lumen, the image is binarized again but using a value of 135. However, the lumen in this instance is not labeled. The authors discuss that these dirty lumen represent the likely location of a lumen if the endoscope were inserted a little deeper. Instead, just the contours of the dirty lumen are shown on the image.

Next, the contours of the segments must be found. Traditional methods which use hue and luminance, or edge based segmentation failed to find accurate contours in these spinal images. The authors present a hybrid approach for this purpose. The authors first join all the “lumen” regions together to create a single region. Then, a modification of the Convex Hull algorithm is used to determine the contours. Convex Hull algorithm works by finding the smallest set of points which form a polygon that bounds the region. As the name mentions, normally the region must be convex, which the lumen is often not. Thus, the algorithm is modified to allow for some concavity. Furthermore, down sampling is used in order to make the algorithm run in real time, whereas the original Complex Hull runs in O(n log(n)).

At this point, we have a segmented image with clear definitions of the boundaries between lumen and non-lumen regions. Furthermore, any white areas within the black of the lumen must be from nerves or vessels. The paper explains additional methods for determining whether the objects are vessels by analyzing the HLS color space to check for the red color associated with blood vessels.

Experimental results conducted by the authors *in vivo* on pig spinal cords showed “complete accordance with classification made by medical doctors”. Furthermore, the algorithm runs almost in real time on endoscopic video, requiring only 60ms per frame to yield a rate of approximately 16fps.

Critique

Overall, the algorithm presented in this paper offers a new way to segment and analyze images from an endoscopic video feed. The authors made modifications to traditional, tested techniques (such as the Convex Hull algorithm) but explained the reasoning behind the changes. The algorithm was optimized to run at real time which allows it to be used in live video feeds. Furthermore, the authors always kept safety as the main priority, never willing to compromise it for increased performance or accuracy. Such an approach makes this algorithm viable for our work in cochlear imaging.

There are a few points which should have been explained better in the paper. The first is the thresholding values of 65 (clean lumen) and 135 (dirty lumen) used to binarize the image after pre-processing. The authors mentions that these were selected as the optimal parameters, but never explained what changing these values do. Such information would have been useful, since the environment we plan to use this algorithm in differs from the spinal cavity the authors used. As such, these parameters were optimized for that space and would likely need to be modified for our purpose.

Secondly, there are only a few images shown from experimental trials. While I commend the authors for showing images where they segmentation did not necessarily produce desirable results, I still feel the images shown were too few in number to gain a sense of the robustness of the algorithm.

Finally, the authors did not present a comprehensive analysis of actual computational running time of the algorithm. One of the key motivations in modifying traditional techniques was to bring the cost of computation down so as to be able to run the algorithm in real time. The authors mention that modifications were made but not actually what was changed (except in the case of down-sampling the image). It would have been more beneficial to see what accuracy or precision losses were sacrificed for performance. It has been seven years since this paper was published and the processing power of computer has dramatically increased. Perhaps real time segmentation could still be achieved without compromising on accuracy.

Implications For Our Project

As mentioned, the goal of our project is to develop safe paths for cochlear implant surgery using a micro borescope. The borescope is attached to a camera and produces a similar feed to that of an endoscope. We plan to implement this algorithm into our own program in order to help break down and analyze the stream of images as we navigate the borescope through the cochlea. Doing so would allow us to make more precise predictions for safe paths for the surgeon to follow.

Furthermore, as mentioned in the paper's conclusion, a combination of this visual segmentation and scans from pre-operative imaging could allow for a complete 3D reconstruction of the inner cochlea. The visual segmentation could also aid in registering the probe's current location to this 3D mesh. Doing so would allow for very detailed and precise safe path creation.

Ultimately, it is the safety of the patient that is of utmost priority. With cochlear implant surgery, there are many risks involving damage to peripheral tissue (including the cranial nerve). Creating a precise and accurate safe path would help mitigate such risks.