Video Summarization of Laparoscopic Cholecystectomies

[1,2]Michael A. Grasso, MD, PhD, [1]Tim Finin, PhD, [1]Xianshu Zhu, [1]Anupam Joshi, PhD, [1]Yelena Yesha, PhD

[1]Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD
[2]University of Maryland School of Medicine, Baltimore, MD
mikegrasso@umbc.edu

ABSTRACT

We compared image features with a distance metric and support vector machine to identify the critical view of a laparoscopic cholecystectomy. Our accuracy was up to 91%. We are currently experimenting with particle analysis, edge analysis, and feature clustering to create a more robust image classifier.

BACKGROUND

Laparoscopic surgery is a minimally invasive technique that is the method of choice for a number of surgical procedures. Patients who undergo laparoscopic surgery have smaller scars, reduced pain, and a quicker recovery. The laparoscopic approach, however, is more technical challenging and has more demanding training requirements [i].

Our overall goal is to develop a software tool to assist with video-based assessment of surgical trainees. We present an initial feasibility study, where we attempted to identify the critical view during a laparoscopic cholecystectomy (surgical procedure to remove the gall bladder). The critical view is an important validation step in the surgery when the essential anatomy has been identified. Related efforts include the segmentation of hysteroscopy video [ii] and echocardiogram video [iii].

Support vector machines are machine learning methods. They are used for image classification by constructing an *N*-dimensional hyperplane, which optimally separates images into two categories. This mapping is performed by a set of mathematical functions, known as kernels.

METHODS

We randomly selected 378 representative images from five laparoscopic cholecystectomy videos, with 104 of the images collected near the critical view. We analyzed 49 spectral and textural features [iv]. A surgeon reviewed the videos to identify the critical view images. We used FFmpeg (http://ffmpeg.org/) and ImageJ (http://rsbweb.nih.gov/ij) to extract images and features. We used the Jeffrey Divergence with an experimentally derived threshold to identify the best image features. These features were combined to create an image classifier based on a support vector machine. We used LIBSVM (http://www.csie.ntu.-edu.tw/~cjlin/libsvm/) to create the support vector machine.

RESULTS

The best textural image features were energy (pixel uniformity), entropy (pixel complexity), contrast (local variation), and correlation (linear patterns). Accuracy of the image classifier was 91% using these textural features combined with color distribution.

DISCUSSION

Our initial results show promise with accuracy up to 91%. The linear and Gaussian kernels performed better than the polynomial kernel.

When interpreting our results it is important to consider several limitations: the study was small in size, the cases were restricted to a single academic medical center, and comparisons were limited to one feature from one image at a time.

We are currently working on a more robust image classifier using particle analysis, edge analysis, and feature clustering.

DISCUSSION

i. Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. Br J Surg. 2004 Dec;91(12):1549-58.

ii. Scharcanski JN, Neto WG, Cunha-Filho JS. Diagnostic hysteroscopy video summarization and browsing. Proceedings 27th IEEE-EMBS. 2005;:5680-5683.

iii. Roy A, Sural S, Mukherjee J, Majumdar AK. State-based modeling and object extraction from echocardiogram video. IEEE Transaction on Information Technology in Biomedicine. 2008 May;12(3):366-376.

iv. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Transaction on Systems, Man, and Cybernetics. 1973 Nov; 3(6):610-621.