Automated Image Rectification in Video-Endoscopy

Dan Koppel[1], Yuan-Fang Wang[1], and Hua Lee[2]

[1] Department of Computer Science,
{dkoppel,yfwang}@cs.ucsb.edu
[2] Department of Electrical and Computer Engineering,
University of California, Santa Barbara, CA, 93106, USA
hualee@ece.ucsb.edu

Abstract. Video-Endoscopy has proven to be significantly less invasive to the patient. However, it also creates a more complex and difficult operating environment that requires the surgeon to operate through a video interface. Visual feedback control and image interpretation in this operating environment can be troublesome. **Automated image analysis has tremendous potential in improving the surgeon's visual feedback, resulting in better patient safety, reduced operation time, and savings in health care.** In this paper, we present our design of an image rectification algorithm for maintaining the head-up display in video-endoscopy and report some preliminary results.

Video-endoscopy [1, 2] are minimally invasive surgical procedures where small incisions are made on the patient to accommodate surgical instruments such as scalpels, scissors, staple guns, and a video endoscope. The scope acquires video images of the bodily cavity that are displayed on a monitor to provide visual feedback to the surgeon.

Though video-endoscopy has proven to be tremendously beneficial in shortening the recuperation time and lowering the treatment cost [1, 2], this patient-oriented technology has increased the difficulty of performing the procedures for the surgeon. E.g. in open surgery, though the surgeon may move about the operating table and view the body anatomy from different viewpoints, the surgeon's sense of up and down is maintained. Or in the computer graphics terminology, the "head-up" vector, which indicates one's general perception of the "up" direction in the environment, is unchanged.

But in video-endoscopy, the bodily cavity is not exposed and the surgeon's perception of the "up" direction is established through the understanding of the anatomy. A large, about-the-axis rotation of the scope and the camera will change the orientation of the body anatomy, which often times results in the loss of one's orientation and bearing. Hence, how to compensate for the about-the-axis rotation of the video scope to maintain the right "head-up" display (i.e., a display where the image is rectified and re-oriented in such a way that the view corresponds to that taken with an upright camera) is important.

We propose a new formulation for rotation sensing and compensation based on novel image analysis techniques. The image analysis tasks proposed here can be stated succinctly as follows: Given a video sequence taken during an endoscopy surgery, with large panning and rotation of the scope, deduce the amount of the about-the-axis rotation relative to a specified reference frame. The images are then rectified and displayed in such a way as if the camera is held in an upright orientation.

The essence of our rotation estimation and compensation scheme is to track anatomical features in an endoscopic video sequence and use the tracked features over time to deduce the camera motion parameters. The recovered motion parameters then allow rotation of the camera's head-up vector relative to that of a reference frame computed. Images can then be re-rendered by compensating for the rotation of the head-up vector.

Fig. 1 shows some image segmentation and tracking results. This preliminary result shows our ability to distinguish instrument regions from organ regions (see caption for explanation). Fig. 2 shows 2D and 3D tracking results over a sequence of about 200 frames. Fig. 2 (a) shows typical 2D feature location errors as a function of time, while Fig. 2 (b) shows the absolute error in estimated angle of rotation in 3D camera motion. Both results are quite good.
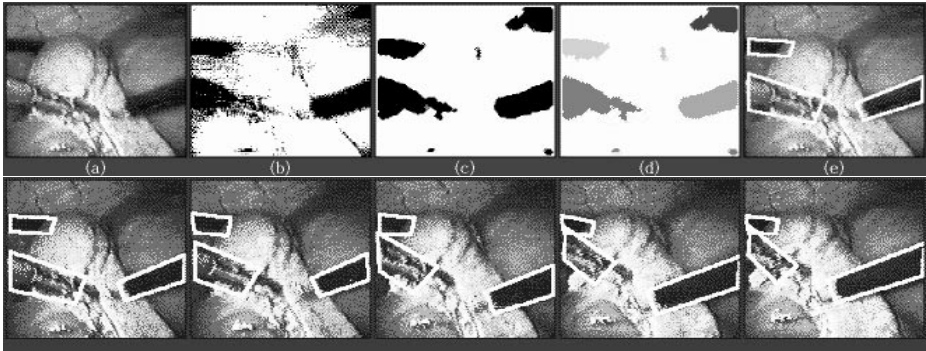
Fig. 1. Upper row: **(a)** Original image. **(b)** Binary image from color classification. Black and white pixels represent surgical instruments and organs, respectively. **(c)** Directional median filters were used to suppress noise. **(d)** Labeled image, different colors represent isolated instrument regions. **(e)** Computed bounding boxes that delineate instrument regions in images. Lower row: Tracking of instrument region.
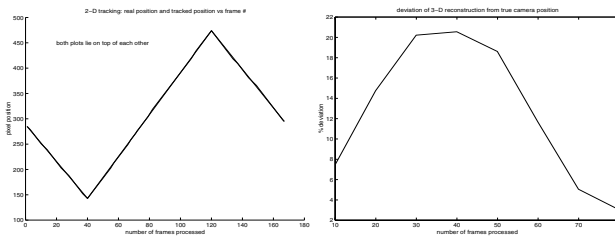
Fig. 2. **(a)** shows the 2D tracking error positions as function of time over about 200 frames. **(b)** shows the 3D tracking error (deviation in estimated angle of rotation from ground truth).

References

[1] J. F. Hulka and H. Reich. *Textbook of Laparoscopy, 2nd Ed.* W. B. Saunders, Philadelphia, PA, 1994.

[2] J. G. Hunter and J. M. Sackier (eds.). *Minimally Invasive Surgery*. McGraw-Hill, New York, 1993.