Toward Automated Model Building from Video in Computer-Assisted Diagnoses in Colonoscopy

Dan Koppel[1], Chao-I Chen[1], Yuan-Fang Wang[1], Hua Lee[2]
[1]Department of Computer Science
[2]Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA

Jia Gu, Allen Poirson, Rolf Wolters
STI Medical Systems
733 Bishop Street, Suite 3100
Honolulu, HI

Abstract

A 3D colon model is an essential component of a computer-aided diagnosis (CAD) system in colonoscopy to assist surgeons in visualization, and surgical planning and training. The ability to construct a 3D colon model from endoscopic videos (or images) is thus critical in a CAD system for colonoscopy. This paper summarizes our on-going research in automated model building in colonoscopy. We have developed the mathematical formulations and algorithms for modeling static, localized 3D anatomic structures within a colon that can be rendered from multiple novel view points for close scrutiny and precise dimensioning. This ability is useful for the scenario when a surgeon notices some abnormal tissue growth and wants a close inspection and precise dimensioning. Our modeling system uses only video images and follows a well-established computer-vision paradigm for image-based modeling. We extract prominent features from images and establish their correspondences across multiple images by continuous tracking and discrete matching. We then use these feature correspondences to infer the camera's movement. The camera motion parameters allow us to rectify images into a standard stereo configuration and calculate pixel movements (disparity) in these images. The inferred disparity is then used to recover 3D surface depth. The inferred 3D depth, together with texture information recorded in images, allow us to construct a 3D model with both structure and appearance information that can be rendered from multiple novel view points.

Keyword

Calibration, enhanced reality, image-guided therapy, modeling, visualization

Description of purpose

The problem of modeling the structure and behavior of a colon for computer-assisted colonoscopy is quite challenging. The complexity depends on a variety of factors; the most important of which are the scale and sophistication of the description, the characteristics of the structure, and the amount of available data. For example, a computer description might be sought only for localized anatomical features, such as polyps, tumors, or abnormal cancerous tissue growth. When a small anatomical feature is analyzed, it may be possible to navigate the scope around the feature with negligible disturbance to its structure. Hence, the structure may be considered static, which allows a 3D computer model to be constructed from video images alone. On the other hand, if the description of a whole colon is sought, one has to address significant tissue deformation resulted from navigating the scope through multiple folds and turns of a colon. A 3D model is much more difficult to obtain given that the scope movement likely will disturb the soft colon structure and results in non-rigid deformation. In this case, an image mosaic approach to construct only a 2D appearance model from video data might be more attainable – with 3D structure data supplied from other sensing modalities such as MRI and CT. In this paper, we report our preliminary results on modeling static, localized anatomic structures in a colon (e.g., a polyp) using video images alone.

Methods

Our modeling system uses only video images and follows a well-established computer-vision paradigm for image-based modeling. We extract prominent features from images and establish their correspondences across multiple images by continuous tracking and discrete matching. We then use these feature correspondences to infer the camera's movement. The camera motion parameters allow us to rectify images into a standard stereo configuration and infer pixel movements (disparity) in these images. The inferred disparity is then used to recover 3D surface depth. The inferred 3D depth, together with texture information recorded in images, allow us to construct a 3D model with both structure and appearance information that can be rendered from multiple novel view points.
More precisely, our modeling system comprises the following modalities:

1. *Calibration*: This is an offline process to estimate intrinsic camera parameters and to correct image distortion (e.g., lens distortion), if any,
2. *Feature selection*: This step identifies unique and invariant colon features for ensuing video analysis,
3. *Feature matching*: The step matches image features across multiple images or video frames to establish correspondences of these features,
4. *Camera motion Inference*: The step uses matched image features to infer camera movement between adjacent images or video frames,
5. *Image rectification*: The step is to rearrange image pixels in such a way that corresponding epipolar lines from two images are aligned and stacked as image scan lines. This step allows standard stereo matching techniques to be applicable regardless of camera movement,
6. *Stereo matching*: The step is to compute disparity (image movement) between two rectified images to allow 3D depth inference,
7. *Model construction*: The step is to infer 3D depth from pixel disparity and construct a 3D model that captures both structure and appearance information, and
8. *Model rendering*: The final step is for rendering the computer model from any novel view points.

Results

Figure 1 shows stereo rectification and matching, and model construction results. The first row in Figure 1 shows the stereo pair before rectification while the second row shows the stereo pair after rectification. The third row of Figure 1 shows the stereo matching results. Stereo disparity is displayed as gray scales on the left while the depth profile reconstructed from the stereo images is shown on the right. Figure 2 shows views of the 3D model constructed with both structure and appearance information based on the stereo matching results from Figure 1.

New or breakthrough work to be presented

The novelty lies in proposing a comprehensive paradigm for 3D model construction, developing novel, efficient, and robust algorithms for 3D motion inference, and validating the modeling system using real-world colonoscopic images.

Publication Status

This work is not being, and has not been, submitted for publication or presentation elsewhere

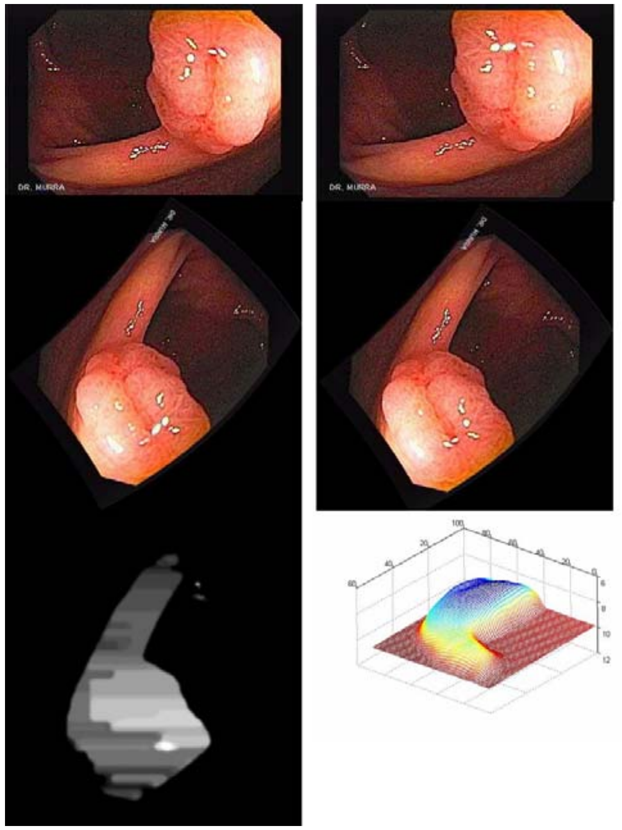

Figure 1 Sample results on stereo rectification, matching, and depth inference.

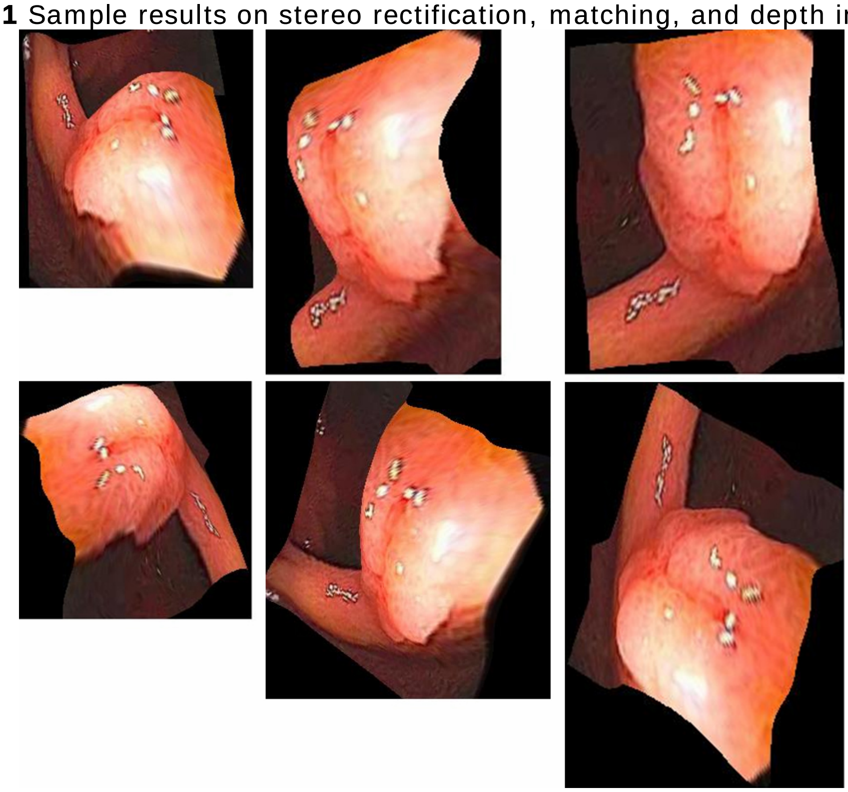

Figure 2 Novel view rendering using models constructed based on dense stereo matching (the input images are those in Figure 1)