Projector-based Surgeon-computer Interaction on Deformable Surfaces

B. Kocev[1,2], F. Ritter[2], L. Linsen[1]
[1] Jacobs University, Bremen, Germany
[2] Fraunhofer MEVIS, Bremen, Germany

Keywords: Surgeon-computer Interaction, Multi-touch Gestures, Projector-based Medical Data Visualization, Image Processing.

Purpose

Providing intuitive and easy to operate interaction mechanisms for medical augmented reality applications is one of the most crucial aspects for accepting these applications for use in the operating room. Commonly, intra-operative navigation information is displayed on an installed monitor, requiring the operating surgeon to change focus from the monitor to the surgical site and vice versa during navigation. Projector-based augmented reality has the potential to alleviate this problem. The aim of our work is to use a projector for visualization, and provide intuitive means for direct interaction with the projected information.

Methods

A consumer-grade projector is used to visualize preoperatively defined surgical planning data. The projection of the virtual information is possible on any deformable surface and a surgeon can interact with the presented virtual information. A Microsoft Kinect camera is used to capture both the surface interactions and the deformations of the surface over time. The retrieved 3D point cloud is processed using the Point Cloud Library to create a 3D virtual model of the interaction zone.

Both the depth and the RGB image from the Kinect camera are used to recognize gesture- and touch-based interactions. The detection of the surgeon's interaction is based on the positions of his/her fingertips. The image processing pipeline that needs to be applied for a robust fingertip detection is presented in Figure 1. We measure the contour's curvature at the convex hull points of the detected hand mask (in form of angles) and put a threshold on the minimum curvature expected for a point to be accepted as a fingertip [1].

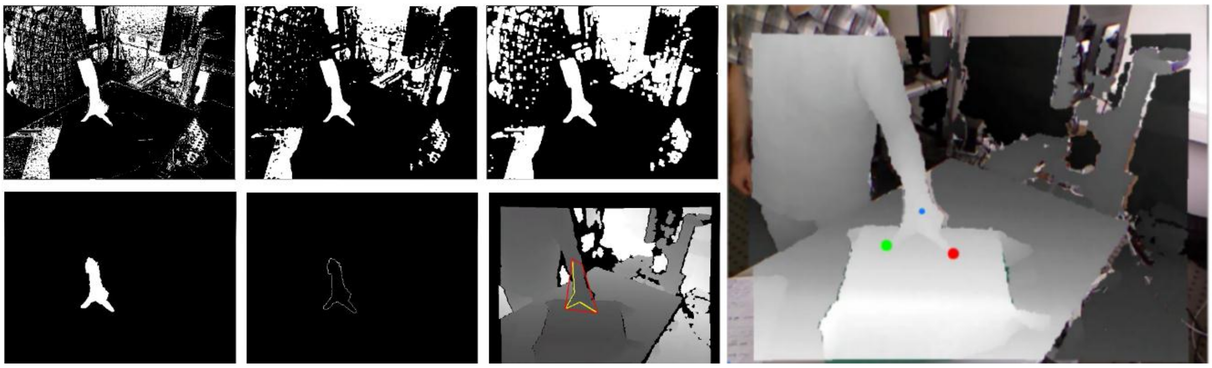

Fig. 1. Two fingertips in contact with the interaction zone: (top-down, left to right) HSI-profile-based hand segmentation mask; median-filtered hand segmentation mask; dilated median filtered hand segmentation mask; final hand segmentation mask obtained by filtering out pixels based on distance to hand palm point and depth differences; extracted hand contour [4]; coarsened hand contour [2] and its convex hull [3]; aligned Kinect RGB/depth image together with two detected fingertips (red and green) and hand palm point (blue).

Upon calibration of projector and Kinect camera using the nestk library, having the fingertips extracted, and by using the point cloud representation of the surface, we are able to detect whether the surgeon interacts with the projected virtual information by putting a threshold on the closeness of his/her fingertips to the interaction zone and compute what projected virtual information the surgeon interacts with by using the projector-Kinect calibration information. We then showcase how the system can be deployed for navigation in a surgical context by identifying the required functionality of such an application, defining the available interaction vocabulary in terms of detected gestures and (multi-)touches, and mapping gestures and (multi-)touches to navigation tasks in the application, which modify the displayed virtual data.

Results

Figure 2 shows our interaction and visualization system, targeted for use with any surgical navigation application. The user stands in front of the Microsoft Kinect camera, while relevant medical information is projected on the interaction zone, e.g., a surgical table covered with green sheets. A hand wave gesture initiates the tracking of the hand. The user can, then, interact with the projected virtual information according to the defined multi-touch-based gestures. Thus, all information such as preoperative planning data is provided to the surgeon and his team intra-operatively in an easy and intuitive way in the proximity to the patient.

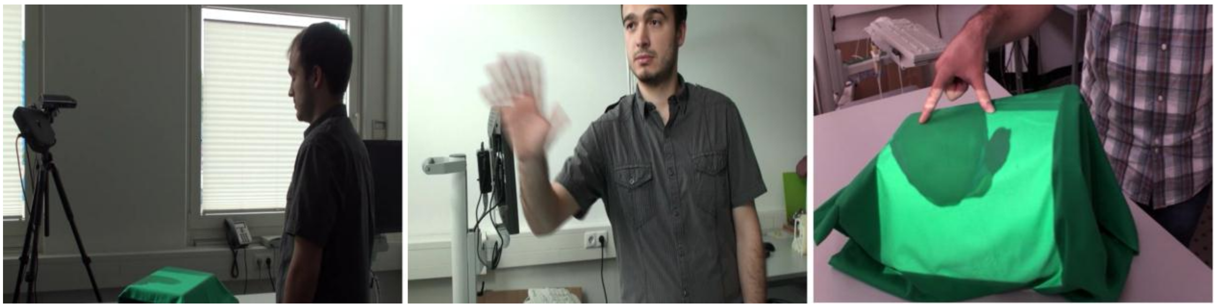

Fig. 2. (left to right) Coupled projector-Kinect device where visual data is projected on interaction zone (green sheet); user initiates tracking with hand wave gesture; pinch-and-pan gesture to interact with the projected virtual information on a curved surface.

In lab studies, we observed that all users got used with the system and the interaction mechanisms very quickly and found it very intuitive. They stated that it seems familiar from what they are used to from standard multi-touch displays and applications. Overall there are no visual mismatches between what is being shown on the interaction zone, and what the user expects to see based on his touch-based gesture input to the system.

Conclusion

We enabled the projection of the virtual information on an arbitrarily shaped surface and used a Microsoft Kinect camera to capture the interaction zone and the surgeon’s actions. Our novel vision-based interaction and projector-based visualization system is targeted for use in medical augmented-reality applications. The system eliminates the burden for the surgeon to have to turn his view from the surgical site to the monitor, eliminates unpleasant distractions and, consequently, may enhance the surgeon’s performance.

References

[1] Bailly, G., Walter, R., Mller, J., Ning, T., and Lecolinet, E. Comparing free hand menu techniques for distant displays using linear, marking and finger-count menus. In Human-Computer Interaction INTERACT 2011, P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, and M. Winckler, Eds., vol. 6947 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2011, 248–262. 10.1007/978 − 3 − 642 − 23771 − 3 19.

[2] Douglas, D., and Peucker, T. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Cartographica: The International Journal for Geographic Information and Geovisualization 10, 2 (1973), 112–122.

[3] Sklansky, J. Finding the convex hull of a simple polygon. Pattern Recognition Letters 1, 2 (1982), 79–83.

[4] Suzuki, S., et al. Topological structural analysis of digitized binary images by border following. Computer Vision, Graphics, and Image Processing 30, 1 (1985), 32–46.