3D Gaze Tracking based on Eye and Head Pose Tracking

D. García-Mato[1,3], A. Lasso[1], A. Szulewski[2], J. Pascau[3], G. Fichtinger[1]

[1] Laboratory for Percutaneous Surgery, School of Computing, Queen's University, Canada

[2] Department of Emergency Medicine, Kingston General Hospital, Canada

[3] Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Instituto de Investigación Sanitaria Gregorio Marañón, Spain

dgmato@hggm.es

INTRODUCTION

Psychological studies on eye movements have documented a temporal and spatial relation of gaze direction with the requirements of performed motor tasks. The oculomotor system directs the gaze towards the spot of greatest interest in the scene, the point providing the most information for the task at hand [1].

Eye tracking has been used in radiology for the evaluation of the visual search process to determine the effectiveness of displays in radiology workstations [2]. More recently eye tracking technology has started being used in surgeries, mainly for the quantitative assessment of surgical skills during minimally invasive surgeries. Significant differences have been found in the eye movements of novices and experts during the performance of surgical tasks [3] and it has been demonstrated that skill assessment is improved when eye-gaze data is added to surgical tool motion data [4].

The eye tracking data used in the previously mentioned studies is two-dimensional, since most applications are based on 2D images or videos. However, the increased importance of 3D imaging and image-guided surgical procedures in the operating room demands the analysis of gaze data in 3D space. The purpose of this work is to assess the feasibility of combining both eye tracking and head positioning to estimate gaze in the 3D space.

MATERIALS AND METHODS

The proposed framework (Fig.1) combines the use of a wearable eye tracker device (Tobii Pro Glasses 2, Tobii Technology, Danderyd, Sweden), an optical tracking system (Polaris®, NDI, Waterloo, Canada) for the real-time positioning of the user's head and a 3D scanner (Artec Eva™, Artec 3D, Luxembourg) for the tridimensional modeling of the objects in the workspace.

In addition, an application was developed in 3D Slicer [5], a free open-source platform for the analysis and visualization of medical images, which receives the gaze data from the eye tracking glasses through a wireless Ethernet connection and the head positioning data from the optical tracking system through OpenIGTLink communication protocol [6] using PLUS open-source software [7].

The proposed methodology consists of four steps:

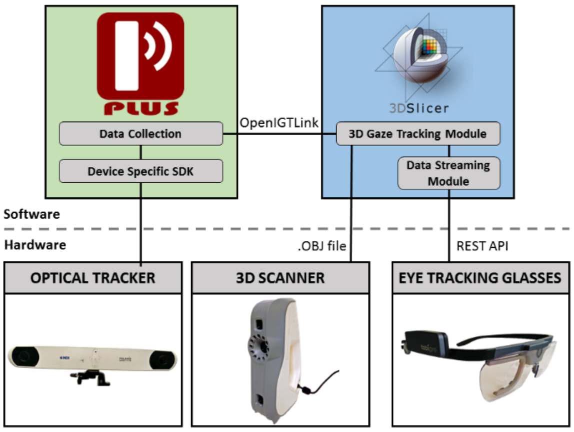

Fig. 1 System overview

1) **3D modeling:** A model of the workspace is generated using the 3D scanner. Several visual markers must be attached to the scene before scanning for calibration and registration purposes.
2) **User preparation:** The wearable eye tracker and a set of three reflective spheres, visible by the optical tracking system, are attached to the user's head using elastic and adjustable straps. Gaze tracking is also possible for those users with eye vision problems, either using glasses or contact lenses.
3) **Calibration:** A calibration procedure is required to compute the relationship between the eye tracker and the optical tracker coordinate systems. During this process the user is asked to focus on 6 visual markers attached to the workspace while gaze data and head position are recorded. A minimization of the shortest distance between each gaze line and corresponding marker 3D position is performed using L- BFGS -B algorithm [8].
4) **Navigation:** Once calibrated, the system is able to display the 3D gaze line in real-time and it is possible to visualize which region of the 3D workspace (generated model) the user is looking at.

The experimental setup for the performance evaluation of the proposed system consists of a room simulating a simple surgical scenario (Fig. 2). A set of 16 visual markers were attached to the scene: 6 of these markers are used for the calibration of the system (calibration markers) and the remaining 10 for accuracy evaluation purposes (evaluation markers).

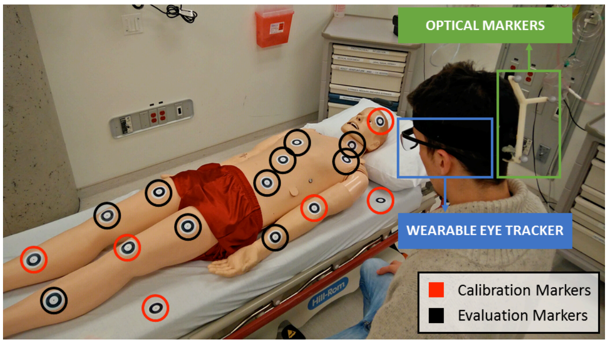

Fig. 2 Distribution of visual markers in workspace

First, the system was calibrated asking users to focus on each of the calibration markers during 3 seconds. Then, users were asked to look at each of the 10 evaluation markers in the scene for 3 seconds in order to assess the accuracy of the 3D gaze tracking system. During this time, samples of the gaze direction and the head position were recorded and 3D gaze lines were estimated. A total of 20 repetitions of this experiment were performed. Gaze tracking error was measured as the shortest distance between gaze lines and marker positions, and as the angular deviation between real and estimated gaze lines.

RESULTS

Results for the accuracy evaluation of the 3D gaze tracking system indicate an average shortest distance between estimated gaze lines and marker positions of 6.0 ± 3.3 mm and an average angular difference between real and estimated gaze lines of 0.4° ± 0.2°. The mean distance between the user's eyes and the visual markers was 94.1 ± 11.7 cm. The average range of motion of the user's head was 73.7° ± 1.8° rotation, 37.8° ± 4.7° flexion-extension, and 23.5° ± 2.9° lateral flexion. 3D head and gaze tracking information can be visualized in real-time together with the generated models of the objects in the scene (Fig. 3).

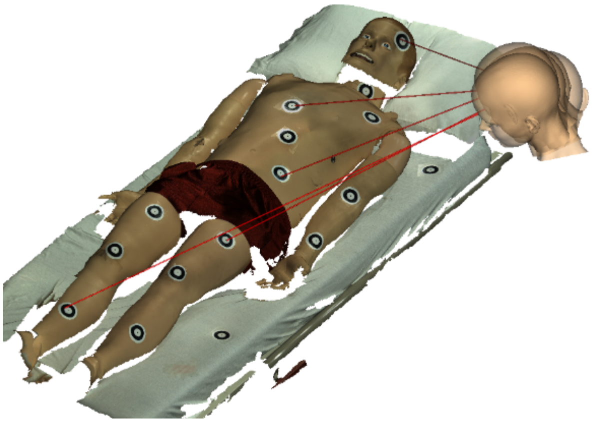

Fig. 3 Visualization of 3D gaze line intersections with model

DISCUSSION

Gaze tracking accuracy is affected by the devices intrinsic errors, the calibration procedure and the attachment of the optical markers to the user's head. The system requires a 6-point calibration procedure which takes less than 30 seconds and assumes the relative position of the eye tracker with respect to the optical markers is maintained. Therefore, during the described experiment the eye tracker was not removed or repositioned after the calibration. However, errors could be reduced by fixing the optical markers directly to the wearable eye tracking glasses.

Although this study was done in a static environment, some applications could involve movable objects in the workspace. For those cases, optical markers could be attached to those objects to know their relative position with respect to the user's head. As a future work, we will study the feasibility of using RGB-D cameras for periodic updates of the 3D scene.

For this work, a 2-camera optical tracking system with limited field-of-view was used, restricting the movement of the user's head. Using multi-camera optical tracking systems would remove this limitation.

The results of this study demonstrate the feasibility of this novel system to provide tridimensional gaze tracking with an average localization error of 6 mm and average angular error in the estimated gaze line of 0.4°. The reported accuracy will enable this system to be used for the analysis and visualization of gaze data in the 3D space for applications requiring the identification of which objects in the scene a person is focusing on.

REFERENCES

[1] M. Land, N. Mennie, and J. Rusted, "The roles of vision and eye movements in the control of activities of daily living," *Perception*, vol. 28, no. 11, pp. 1311–1328, 1999.

[2] M. S. Atkins, A. Moise, and R. Rohling, "An application of eye gaze tracking for designing radiologists' workstations: Insights for comparative visual search tasks," *ACM Trans. Appl. Percept.*, vol. 3, no. 2, pp. 136–151, 2006.

[3] B. Law, M. S. Atkins, A. J. Lomax, and C. L. Mackenzie, "Eye Gaze Patterns Differentiate Novice and Experts in a Virtual Laparoscopic Surgery Training Environment," *ETRA '04 Proc. 2004 Symp. Eye Track. Res. Appl.*, vol. 1, no. 212, pp. 41–48, 2004.

[4] N. Ahmidi, G. D. Hager, L. Ishii, G. Fichtinger, G. L. Gallia, and M. Ishii, "Surgical Task and Skill Classification from Eye Tracking and Tool Motion in Minimally Invasive Surgery," *Lect. notes Comput. Sci.*, vol. 63, no. 63, pp. 295–302, 2010.

[5] S. Pieper, M. Halle, and R. Kikinis, "3D Slicer," in *2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano (IEEE Cat No. 04EX821)*, 2004, vol. 2, pp. 632–635.

[6] J. Tokuda *et al.*, "OpenIGTLink: an open network protocol for image-guided therapy environment," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 5, no. 4, pp. 423–434, Dec. 2009.

[7] A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi, and G. Fichtinger, "PLUS: Open-source toolkit for ultrasound-guided intervention systems," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 10, pp. 2527–2537, 2014.

[8] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A Limited Memory Algorithm For Bound Constrained Optimization," *SIAM J. Sci. Comput.*, vol. 16, no. 5, pp. 1190–1208, 1995.