Translational and Rotational Arrow Cues (TRAC) Outperforms Triplanar Display for use in 6-DOF IGS Navigation Tasks

David E. Usevitch and Jake J. Abbott

Department of Mechanical Engineering and the Robotics Center, University of Utah

{*david.usevitch, jake.abbott*}@*utah.edu*

INTRODUCTION

Many image-guided surgery (IGS) tasks require a human to manually position an object in space with 6 degrees-of-freedom (6-DOF). We are particularly interested in positioning robotic [1] and magnetic [2] devices during robot-assisted cochlear-implant surgery.[1] Displaying the current and desired pose of the object on a 2D display is straightforward, however, providing guidance to accurately and rapidly navigate the object in 6-DOF is challenging [3]. Visual guidance is typically accomplished using a triplanar display (Fig. 1) that shows three orthogonal views of the workspace [1,3–5]. This method is quite unintuitive, particularly before training.

We propose an intuitive method of visually communicating navigation instructions using a single principal view that approximates the human's egocentric view of the actual object, using translational and rotational arrow cues (TRAC) defined in an object-centric frame (Fig. 2). We show the arrow corresponding to the single DOF with the largest error until convergence in that DOF, at which point we switch to the new arrow corresponding to the largest error, alternating between rotation and translation. Copies of the moving virtual object and the respective arrow cues, which have small rotational offsets from the central virtual object but fixed positions on the screen, ensure that the arrow cues do not become occluded. With the TRAC method, the target pose is not actually required, and is useful to quickly perform the initial gross alignment. We show that the TRAC method outperforms the state-of-the-art triplanar method in terms of time to complete 6-DOF navigation tasks. Preliminary results also suggest this method enables improved accuracy.

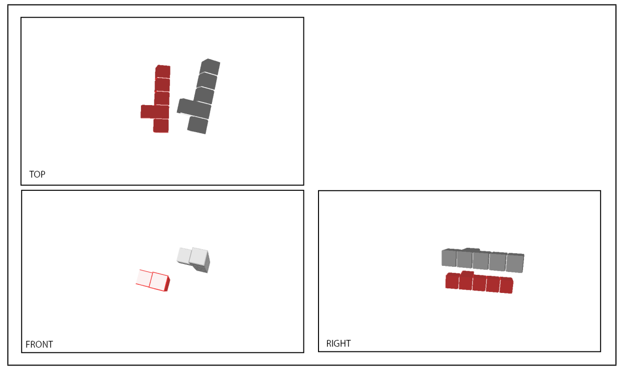

Fig. 1: Triplanar display showing three orthogonal views. The red stationary object is the target pose, and the gray object moves with the hand-held object.

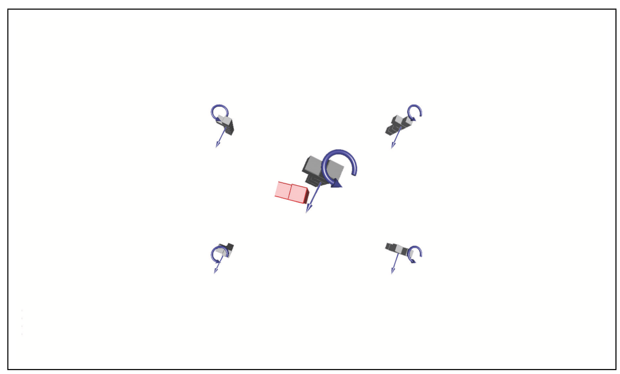

Fig. 2: Proposed TRAC method. The red stationary object is the target pose, and the gray objects move with the hand-held object. We show both translational and rotational arrows here for display purposes, but in practice only one arrow cue (either translational or rotational) is displayed at a time during navigation. Note that we maintained object sizes consistent with the triplanar display to avoid potential biases due to the increased object sizes enabled by the TRAC method (i.e., a single view filling the entire screen).

MATERIALS AND METHODS

To quantify the benefits of TRAC relative to a triplanar display, we conducted a human-subject study with IRB approval, comprising 4 males and 4 females, aged 21–31 years, from the engineering-student population at the University of Utah. The completion time (T) of 6-DOF pose matching was measured in two sessions, separated by 2 days to mitigate effects of learning and fatigue. Session A measured T using the triplanar display. There is no standard arrangement used in prior work, so we chose an arrangement that correlates closely with [1], in which we position a front view at the lower left corner, a side view at the lower right, and a top view at the upper left (Fig.1). This arrangement is also consistent with common mechanical-drawing techniques. Session B measured T using the TRAC method. In this method a single translational (straight) or rotational

[1]Research reported in this publication was supported in part by the National Institute of Deafness and Other Communication Disorders of the National Institutes of Health under Award Number R01DC013168.

(circular) arrow is displayed along one object-centric axis of the moving virtual object (Fig. 2). The object is first rotated to be within rotation thresholds for all three axes; translation is then performed until the position is within positioning thresholds for all three axes; during rotation and translation, the arrow corresponding with the maximum rotational or translational distance to the goal is displayed; the display alternates between these two cues until the pose is within all specified thresholds. Half of the subjects (balanced by gender) performed Session A first, and the other half performed Session B first.

A set of 30 poses was randomly generated from a uniform distribution with translational offsets along each of the three principal axes within ± 40 mm of the home position, and with each pose having a rotational offset of 15° (using the angle-axis formulation). These offsets were designed to represent the precision-alignment stage of pose matching. The same 30 poses were used in both sessions for all subjects, but the order of the 30 poses was randomized each time.

A 1.02 m LCD monitor was positioned above a Polaris Spectra optical tracker, both located 1.83 m in front of a subject seated at a table. The subject manually manipulated an unconstrained hand-held object with three distinct axes, which was equipped with motion-tracking markers. For each of the 30 poses in each session, the subject homed the object to a specified homing position, and then moved the object to match the target pose as quickly as possible. When the object was within 5° and 3 mm of the target (see Discussion), the object changed colors to signal a match, and the subject was required to maintain the object within this threshold for 1.5 sec, at which point the T was recorded and the 1.5 sec hold time was subtracted. The object was then homed in preparation for the next target pose.

A generalized linear regression analysis is used to determine statistical significance of treatment factor *navigation method* (treated as a fixed-effect variable) on response T, using blocking factors *subject* (treated as random-effect variable) and *session order* (treated as fixed-effect variable). The conventional significance for the entire analysis was determined at $\alpha = 0.05$.

RESULTS

In Fig. 3 we present the results of our study. The T data was skewed, and a gamma distribution was used to represent the data. We find that 6-DOF alignment is significantly faster using the TRAC method than using the triplanar method. With the triplanar method, the median T was 78% longer than the median T of the TRAC method. The worst-case T when using the triplanar method was more than six times longer than the worst-case T using the TRAC method. Additionally, subjects declared a unanimous preference for the TRAC method.

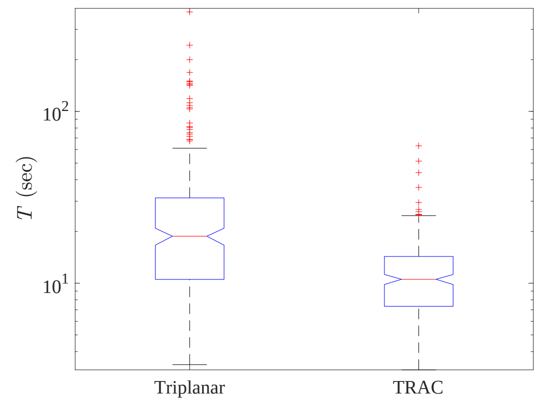

Fig. 3: Notched box-whisker plot comparing the completion time T for both methods, using a logarithmic scale. The red line represents the medium datum, the lower and upper edges of the box represent the first and third quartiles of the data, respectively, the lower/upper whisker represents the smallest/largest datum within 1.5 IQR, the notches represent the confidence interval of the median, and the red crosses represent outliers.

DISCUSSION

We chose a quite large convergence tolerance for our study because in pilot testing we found many subjects struggled with completing pose alignment in a reasonable amount of time when using the triplanar display if we used a tighter tolerance. Consequently, we believe our results under-represent the true relative benefit of TRAC. Anecdotally, we have been able to rapidly achieve 6-DOF alignment with sub-mm and sub-degree accuracy using TRAC. Quantifying the accuracy of TRAC is left as future work.

REFERENCES

[1] T. L. Bruns and R. J. Webster III, An image guidance system for positioning robotic cochlear implant insertion tools, *SPIE Medical Imaging*. 10135:1-6, 2017.

[2] L. Leon, F. M. Warren, and J. J. Abbott. Optimizing the magnetic dipole-field source for magnetically guided cochlear-implant electrode-array insertions, *J. Med. Robot. Res.* 3(1):1-15, 2018.

[3] R. F. Labadie and J. M. Fitzpatrick. What Does the Future Hold? *Image-guided Surgery : Fundamentals and Clinical Applications in Otolaryngology*. San Diego: Plural Publishing, 185–188, 2016.

[4] G. Dagnino, I. Georgilas, S. Morad, P. Gibbons, P. Tarassoli, R. Atkins, and S. Dogramadzi. Intraoperative fiducial-based CT/fluoroscope image registration framework for image-guided robot-assisted joint fracture surgery. *Int. J. Comput. Assist. Radiol. Surg.*, 12:1383-1397, 2017.

[5] L. Adams, W. Krybus, D. Meyer-Ebrecht, R. Rueger, J. M. Gilsbach, R. Moesges, and G. Schloendor. Computer-assisted surgery. *IEEE Comput. Graphics Appl.*, 10:43–51, 1990.