ACS 2022 Surgeons and Engineers: A Dialogue on Surgical Simulation Meeting

Research Abstracts

Automating Context Dependent Gaze Metrics for Evaluation of Laparoscopic Surgery Manual Skills

Shiyu Deng, MS; Chaitanya Kulkarni, MS; Shawn David Safford, MD, FACS, MBA; Sarah Parker, PhD; Laura E. Barnes, PhD; Tianzi Wang, MS; Jacob Hartman-Kenzler, MS; and Nathan Lau

Virginia Tech, Blacksburg, VA, Penn State Health, Hershey, PA, Virginia Tech, Roanoke, VA, University of Virginia, Charlottesville, VA

Introduction: Research has found that feedback gaze behaviors can differentiate performance between experts and novices in the technical surgical skills. However, gaze behaviors revealing feedforward control are absent in research. This study compared feedforward and feedback gaze metrics in assessing skill levels in the peg-transfer task.

Methods: Medical students practiced the peg-transfer task, which was video-recorded with eye-tracking data. The final dataset consisted of 499 practice trials from 9 participants. The feedback metrics were fixation rates on (1) objects, and (2) tool-holding-an-object. The feedforward metrics were fixation rates on areas (1) of objects traversed 0.12-0.6s later, and (2) outside tool-holding-an-object. These two are feedforward metrics because the former suggests looking ahead and the latter suggests looking for information beyond the immediate moving items. Cluster analysis with the 4 metrics was used to group the trials into different skill levels. A random forest model was trained to predict the skill level using the metrics as predictors of the clusters. About two and one third of the data were used for model training and testing, respectively.

Results: Within-cluster-sum of squared errors suggested 3 clusters, identifying 3 skill levels amongst the practice trials. Between the 3 skill levels, ANOVAs indicated significant differences in fixation rates on areas of objects ($F_{2,496}$=383.1,p<.001), tool-holding-an-object ($F_{2.496}$=86.1,p<.001), objects traversed 0.12-0.6s later ($F_{2,496}$=445.1,p<.001), and outside tool-holding-an-object ($F_{2,496}$=32.4,p<.001). The correspondence between the clusters and skill levels was confirmed by differences between cluster completion times ($F_{2,496}$=6.9,p=.001). For the skill level with shortest completion time, trainees exhibited highest scores for all gaze metrics. The random forest model predicted skill level with 4.8% out-of-bag error. The Mean Decrease Accuracy plot indicated that fixation rates on objects, and objects traversed 0.12-0.6s later were the most important for predicting skill levels.

Conclusions: Feature importance in random forest model suggested that fixation rates on objects and objects traversed 0.12-0.6s later were almost equally important. This study provided the first evidence that feedforward can be as sensitive and capable as feedback gaze metrics at assessing surgical skills.