Semiautomated Motion Tracking for Objective Skills Assessment in Otologic Surgery: A Pilot Study

Vivek V. Kanumuri, MD1,2, Bishoy Ameen1, Osama Tarabichi, MD1, Elliott D. Kozin, MD1,2, and Daniel J. Lee, MD1,2

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Abstract
Perioperative teaching and feedback of technical performance are essential during surgical training but are limited by competing demands on faculty time, resident work-hour restrictions, and desire for efficient operating room utilization. The increasing use of high-definition video microscopy and endoscopy in otolaryngology offers opportunities for trainees and faculty to evaluate performance outside the operating room but still requires faculty time. Our hypothesis is that automated motion tracking via video analysis offers a way forward to provide more consistent and objective feedback for surgical trainees. In this study, otolaryngology trainees at various levels were recorded performing a cortical mastoidectomy on cadaveric temporal bones using standard surgical instrumentation and high-definition video cameras coupled to an operating microscope. Videos were postprocessed to automatically track the tip of otologic dissection instruments. Data were analyzed for key metrics potentially applicable to the global rating scale used in the Accreditation Council for Graduate Medical Education’s Objective Structured Assessments of Technical Skills.

Keywords
motion tracking, objective surgical skills tracking

Received January 29, 2018; revised November 9, 2018; accepted January 23, 2019.

Evaluation of surgical resident performance has traditionally relied on informal and subjective perioperative feedback from attending surgeons.1 Unfortunately, this method of assessment often occurs sporadically, especially in the context of increasing demands on faculty time and work-hour restrictions.2-5 There has been a greater focus on surgical simulation, retrospective video analyses, and other methods to provide feedback in a less resource-intensive setting.4,6,7 Motion tracking technology has shown promise and has been successfully used to follow surgical stroke motion and predict technical expertise.8-10 However, studies to date have required tagged instruments, custom simulators, or high-speed cameras.

Modern otolaryngology, with its widespread use of high-definition (HD) cameras integrated with endoscopes and microscopes, is well positioned to take advantage of video-based motion tracking. Otologic surgery, especially cortical mastoidectomy, represents an ideal use case, given a largely stable field of view and extensive use of a single instrument (otologic drill). The aim of this pilot study was to assess the utility of commonly used motion tracking software to assess surgical skills with standard dissection instruments and HD cameras coupled to a binocular microscope during cortical mastoidectomy.

Materials and Methods
Subjects were recorded performing cortical mastoidectomies on cadaveric fixed temporal bone specimens with standard HD cameras (Karl Storz, Tuttlingen, Germany) coupled to binocular microscopes (Olympus Corp, Tokyo, Japan) in the otolaryngology surgical training laboratory. Subjects included those at the novice level (medical students), middle-level (postgraduate year 2/3 residents), and advanced level (postgraduate year 5, neurotology fellow). In 3 cases, trainees were recorded performing a simple cortical mastoidectomy intraoperatively under the direct supervision of a senior attending neurotologist with a comparable HD camera integrated to the operative microscope (OH4; Leica, Wetzlar, Germany). All videos were recorded under 30 Hz and 720

1Department of Otolaryngology–Head and Neck Surgery, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts, USA
2Department of Otology and Laryngology, Harvard Medical School, Boston, Massachusetts, USA

Corresponding Author:
Vivek V. Kanumuri, MD, Harvard Medical School, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, MA 02114, USA.
Email: vivek_kanumuri@meei.harvard.edu

This Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (http://www.creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Representative segments 20 to 30 seconds long were then selected where the otologic drill was active and remained within the field of view of the recording camera. Videos were postprocessed with ProAnalyst Motion Tracking Software (Xcitex, Woburn, Massachusetts) to automatically track the tip of the drill across the surgical field. User intervention was provided when it was clear that the tracking deviated from the tip of the drill for $3$ frames. Tracking data were then exported into Matlab (2016a; MathWorks, Natick, Massachusetts) and analyzed for key metrics. Linear regression and analysis of variance with these statistics (cadaveric only) were performed with Microsoft Excel. This study was deemed exempt by our Massachusetts Eye and Ear Institutional Review Board (1040003-1, 1045566-1).

Results

Seven subjects performed mastoidectomies in cadaveric specimens with the motion tracking software, and $18$ discrete video segments were analyzed (Figure 1, Video 1). Three additional trainees were recorded performing segments of mastoidectomies intraoperatively. The motion tracking adequately followed the tip of the otologic drill as it moved across the surgical field without any user intervention in the majority of frames (and in the case of the higher-quality intraoperative videos in $>99.7\%$ of frames).

The tracking data were then analyzed for key metrics potentially applicable to the global rating scale used in the Accreditation Council for Graduate Medical Education’s Objective Structured Assessments of Technical Skills (OSATS). Specifically, we aimed to use metrics that could help define the “time and motion” or “instrument handling” criteria of the OSATS (Table 1). The metrics chosen included speed (mean ± SD: $12.1 ± 9.1$ pixels/s), acceleration ($11.1 ± 10.3$ pixels/s$^2$), total distance ($7436.2 ± 4863.5$ pixels), and total stoppage time ($7.0 ± 4.9$ seconds; Table 2). Total distance traveled, speed, and acceleration trended with level of training, although this did not achieve significance.

Discussion

This pilot study explores the utility of using video-based motion tracking software to track instruments in otologic surgery post hoc and quantify operative skills of surgical trainees. We found that in the specific case of a cortical mastoidectomy, using standard dissection instruments and HD cameras coupled to an operative microscope was sufficient to enable automated tracking of an otologic drill as it

### Table 1. Comparison of Relevant Measures Obtained from Motion Tracking and the Global Rating Scale Used in the OSATS.

| OSATS Measures and Scale$^a$ | Time and motion | Instrument handling |
|-----------------------------|-----------------|---------------------|
| 1 | Many unnecessary moves | Repeatedly makes tentative or awkward movements with instruments |
| 2 | Efficient time/motion but some unnecessary moves | Competent use of instruments although occasionally appeared stiff or awkward |
| 3 | Economy of movement and maximum efficiency | Fluid moves with instruments and no awkwardness |

| Potential Metrics Extracted from Motion Tracking Software |
|----------------------------------|
| Overall speed | Mean velocity over duration of video in pixels/s |
| Acceleration | Mean acceleration over duration of video in pixels/s |
| Total distance traveled | Total distance traveled in x and y |
| Stoppage time | Amount of time drill was not in motion |

Abbreviation: OSATS, Objective Structured Assessments of Technical Skills.

$^a$Adapted from original description of OSATS global rating scale from Martin et al. $^{12}$
moved across the surgical field. This is a significant finding, as using recording equipment that is analogous to what is already commonly found in otolaryngology operating theaters removes a crucial barrier to wider adoption of motion tracking technology in surgical training.6,8-10

It is important to note that this motion tracking was not fully automated, requiring users to select their instrument of interest and presegment videos into 20- to 30-second clips, in which the drill burr remained within the field of view of the microscope. Moreover, especially fast movements of the instrument occasionally required the user to reorient the tracking cursor. These limitations may be overcome in the near future by incorporating artificial intelligence models, which have already shown promise in surgical instrument detection in simulated procedures.11

We also attempted to extract key metrics from the tracking data potentially applicable to the “time and motion” or “instrument handling” sections of the OSATS scale for surgical trainees.12 We based these in part on simplified versions of the metrics used by McGoldrick et al in their study of motion tracking in microvascular surgery. Total distance traveled, speed, and acceleration trended with level of training, but this should be interpreted with caution given the relatively small size of our initial pilot cohort (eg, a single outlier resulted in midlevel trainees having higher speed/acceleration than advanced-level participants). Future work will focus on expanding our study to a larger well-controlled cohort and careful validation and contextualization of these measures by expert surgeons.

### Author Contributions

Vivek V. Kanumuri, study design, data acquisition, analysis and interpretation, revising the manuscript; Bishoy Ameen, study design, data acquisition, analysis and interpretation, revising the manuscript; Osama Tarabichi, study design, data acquisition, revising the manuscript; Elliott D. Kozin, study design, revising the manuscript; Daniel J. Lee, study design, analysis and interpretation, revising the manuscript.

### Disclosures

**Competing interests:** None.

**Sponsorships:** None.

**Funding source:** National Institutes of Health (T32 training grant for V.V.K.).

### Table 2. Mean Values from Metrics Extracted from Motion Tracking Software.α

| Metric                      | Overall       | Novice (n = 2) | Midlevel (n = 3) | Advanced (n = 2) |
|-----------------------------|---------------|----------------|------------------|------------------|
| Overall speed, pixels/s     | 8.4 ± 0.7     | 4.9 ± 1.5      | 12.4 ± 10.3      | 7.1 ± 1.6        |
| Acceleration, pixels/s²     | 8.1 ± 8.5     | 4.8 ± 1.8      | 12.5 ± 12.8      | 6.0 ± 1.7        |
| Total distance traveled, pixels | 7192.3 ± 4972.8 | 4794.5 ± 1837.8 | 8194.3 ± 3994.8 | 8666.9 ± 7989.0 |
| Stoppage time, s            | 15.8 ± 14.5   | 20.8 ± 6.5     | 7.2 ± 4.3        | 21.9 ± 24.5      |

αValues are presented as mean ± SD.

### References

1. Mercier E, Chagnon-Monarque S, Lavigne F, Ayad T. Objective assessment of technical skills in otolaryngology–head and neck surgery residents: a systematic review. Otolaryngol Head Neck Surg. 2018;158:54-61.
2. Carlile GS. Teaching within the operating theater. Perspect Biol Med. 2012;55:127-136.
3. Haluck RS, Krummel TM. Computers and virtual reality for surgical education in the 21st century. Arch Surg. 2000;135:786-792.
4. Scott DJ, Cendan JC, Pugh CM, Minter RM, Dunnington GL, Kozar RA. The changing face of surgical education: simulation as the new paradigm. J Surg Res. 2008;147:189-193.
5. Moulton CA, Dubrowski A, Macrae H, Graham B, Grober E, Reznick R. Teaching surgical skills: what kind of practice makes perfect? A randomized, controlled trial. Ann Surg. 2006;244:400-409.
6. Ahmidi N, Poddar P, Jones JD, et al. Automated objective surgical skill assessment in the operating room from unstructured tool motion in septoplasty. Int J Comput Assist Radiol Surg. 2015;10:981-991.
7. Kerwin T, Wiet G, Stredney D, Shen HW. Automatic scoring of virtual mastoidectomies using expert examples. Int J Comput Assist Radiol Surg. 2012;7:1-11.
8. McGoldrick RB, Davis CR, Paro J, Hui K, Nguyen D, Lee GK. Motion analysis for microsurgical training: objective measures of dexterity, economy of movement, and ability. Plast Reconstr Surg. 2015;136:231e-240e.
9. Scholl L, Schmidig G, Faizan A, TenHuisen K, Nevelos J. Evaluation of surgical impaction technique and how it affects locking strength of the head-stem taper junction. Proc Inst Mech Eng H. 2016;230:661-667.
10. Conroy E, Surender K, Geng Z, Chen T, Dailey S, Jiang J. Video-based method of quantifying performance and instrument motion during simulated phonomurgery. Laryngoscope. 2014;124:2332-2337.
11. Zisimopoulos O, Flouty E, Stacey M, et al. Can surgical simulation be used to train detection and classification of neural networks? Healthc Technol Lett. 2017;4:216-222.
12. Martin JA, Regehr G, Reznick R, et al. Objective Structured Assessment of Technical Skill (OSATS) for surgical residents. Br J Surg. 1997;84:273-278.