A Computer-Vision Based Solution for Capturing Kinematic Surgeon Skill Metrics

Mona Fathollahi Ghezelghieh, PhD, Hasan Sarhan, PhD, Ramon Pena, PhD, Lela DiMonte, MS, Anshu Gupta, PhD, Aishani Ataliwala, MS, Jocelyn Barker, PhD

Johnson & Johnson

ETHICON

PART OF THE Johnson & Johnson FAMILY OF COMPANIES

INTRODUCTION

The way a surgeon moves their instruments during a case can uncover critical insights such as their level of experience [1] and their progression along the learning curve [2]. In some cases, motion-based metrics have even been shown to correlate with patient outcome measures [3].

While shown to be valuable, these metrics calculated from tool trajectories have often been limited to robot-assisted MIS cases where kinematic data streams are accessed from the robotic system. Here, a solution is proposed for calculating motion metrics from laparoscopic video feed so that all MIS surgeons can access these metrics.

We have developed this broadly-applicable solution using computer vision techniques to generate tool positional data directly from laparoscopic video recordings, so that surgeons can be evaluated directly based on their performance whether using a laparoscopic or robotic-assisted approach.

MATERIALS AND METHODS

The proposed computer vision framework, Multiple Surgical Tool Tracking (MSTT), aims to estimate the trajectories of several tools present within a video segment.

The MSTT framework consists of two models- one for tool detection and another for reidentification of tools (in instances when they reappear after having left the screen). An overview of MSTT can be seen in Figure 1. MSTT can estimate the trajectory (x & y image locations) of both the right and left tools that the surgeon uses throughout each surgical step or task.

After trajectories are estimated by the framework, the surgeon performance metrics such as path length, derivatives of displacement, and curvature are calculated.

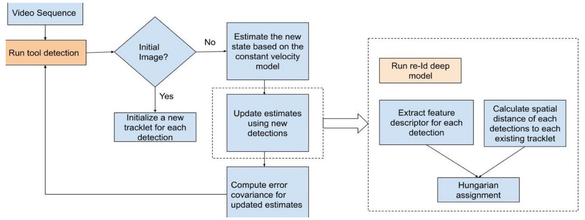


Figure 1. Overview of the MSTT method for generating tool tracks from endoscopic video data. Key components include detection of tools in the individual frames, characterization spatial and visual features of the detected tools, and track assignment using a Hungarian algorithm.

Here, MSTT is run on the ROSMA dataset, which contains video and kinematic data collected for twelve participants completing dry lab training tasks using the da Vinci Research Kit (dVRK) robot. Figure 2 shows results of tracking two tools in an example video and comparing the model results to those generated from kinematic data.

Surgeon performance metrics are then calculated for the ROSMA dataset in multiple ways: (1) in three-dimensional (3D) space using kinematic data at three sampling rates, (2) in two-dimensional (2D) space using projected kinematic data at three sampling rates, and (3) using output of MSTT at 15 fps. These experiments allow us to examine potential differences in computed metrics between kinematic and computer vision analysis.

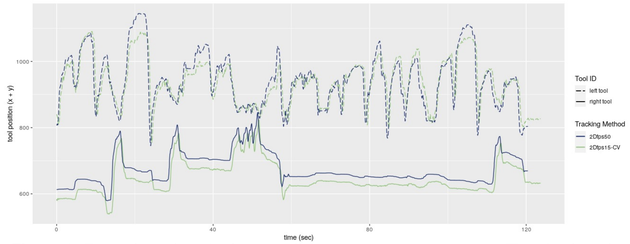


Figure 2. Example tool tracks for experiment X01_Pea_on_a_Peg_01. Time is plotted on the x-axis with the video x+y position on the y-axis. The image shows both the CV tracking data as well as the kinematic data which has been transformed to the camera view. Offset between the two tracks may be due to the location on the tool being tracked (yaw for kinematic, center of end effector for CV)

RESULTS

Our results show high correlation across many metrics useful for surgeon skill assessment [4, 5] between 3D kinematic and computer vision data (Figure 4). In many cases, computer vision correlation approaches theoretical upper limit of 2D kinematic data at 15 fps (Figure 3). In some cases (e.g., time in motion, average velocity) when correlation decreased, the primary limiter was data dimensionality whereas for others (max velocity, average tortuosity) the primary limiter was temporal resolution. It should be noted that the data collection rate for the ROSMA video was abnormally low at 15fps, so those metrics limited by temporal resolution should be improved in a more typical 30fps sample. One other potential source of differences was the lack of exact temporal synchronization between the kinematic and video data which may account for some of the slightly reduced correlation compared to the 2D, 15fps kinematic data.

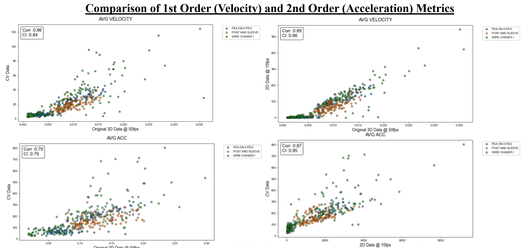


Figure 3. 3D Kinematic vs CV (left), 3D Kinematic vs Reduced Framerate 2D Kinematic vs CV (right)

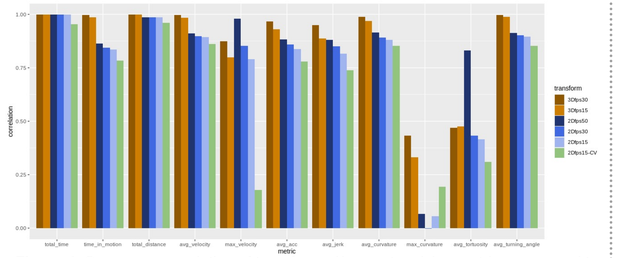


Figure 4. Spearman correlation of transformed/computer vision metrics compared to the full resolution 3D kinematic metrics across ROSMA experiments.

CONCLUSION

This work shows the ability to use computer vision methods to replicate prior findings in kinematic data. These findings greatly expand the potential scope for tool motion metrics by making them a potential target for video-based assessment. Future work would include applying these computer vision methods to both laparoscopic and robotic patients video data to demonstrate the ability to distinguish surgeons of various skill levels.

REFERENCES

1. Fard, M.J., Ameri, S., Chinnam, R.B., Ellis, R.D.: Soft boundary approach for unsupervised gesture segmentation in robotic-assisted surgery. IEEE Robotics and Automation Letters 2 , 171178 (1 2017). https://doi.org/10.1109/LRA.2016.2585303
2. Shaei, S.B., Guru, K.A., Esfahani, E.T.: Using two-third power law for segmentation of hand movement in robotic assisted surgery. vol. 5C-2015. American Society of Mechanical Engineers (ASME) (2015). https://doi.org/10.1115/DETC2015-47813
3. Hung, A.J., Chen, J., Gill, I.S.: Automated performance metrics and machine learning algorithms tomeasure surgeon performance and anticipate clinical outcomes in robotic surgery (8 2018). https://doi.org/10.1001/jamasurg.2018.1512, https://jamanetwork.com/journals/jamasurgery/fullarticle/2685266
4. Fard, M.J., Ameri, S., Ellis, R.D., Chinnam, R.B., Pandya, A.K., Klein, M.D.: Automated robot-assisted surgical skill evaluation: Predictive analytics approach. International Journal of Medical Robotics and Computer Assisted Surgery 14 (2018). https://doi.org/10.1002/rcs.1850
5. Azari, D.P., Frasier, L.L., Quamme, S.R., Greenberg, C.C., Pugh, C.M., Greenberg, J.A., Radwin, R.G.: Modeling surgical technical skill using expert assessment for automated computer rating. Annals of surgery 269 , 574581 (3 2019). https://doi.org/10.1097/SLA.0000000000002478

ACKNOWLEDGEMENTS

We would like to thank Derek Peyton (Johnson & Johnson) for orchestrating our video annotation making this project possible.

www.PosterPresentations.com