Image Processing of Video Data on Robot-Assisted Intervention: Towards Data Driven-Surgery

L. Yang, *Member, IEEE*, S.K. Singh, Y. Lei, T. Kesavadas, *Member, IEEE*

***Abstract*—There has been enormous interest in Data-driven Surgery, especially in the context of robot-assisted interventional procedures, fueling a need for acquiring and analyzing surgical tasks to improve the efficacy of modern surgical procedures. This work proposed a camera-based framework using image processing technique to acquire trajectory data of surgical instruments. This image processing approach enables marker-less acquisition of trajectory information without joint sensor information of the robot joints. Leveraging the wealth of surgical video data using data analytics, it is hoped that the information extracted lay a foundation for task analysis and eventually contributes towards the advancement of Surgical Data Science by via a video data-driven approach.**

I. Introduction

The need for instrument trajectory data analysis during interventional procedures is motivated by a strong interest in Surgical Data Science [1, 2] fueled by the wealth of data in modern interventional procedures. Despite modern data-driven interventional strategies, the increasingly complex surgical techniques continue to rely highly on manual competency. To bridge the gap between contemporary digital technology and the surgical practice, there is a need to tap on the wealth of digital data available with a ubiquitous camera-based approach for trajectory tracking and analysis.

II. Method

We present a camera-based approach using image processing to acquire surgical instrument trajectories from video data. The method consists of a planar visual tracking technique using passive features to extract features for the tool trajectory. The focus is on the extraction of the planar and point features for recovering positional information of the rigid distal end of the instrument without any external sensor.

The development methodology of the vision-based trajectory analysis approach includes design conceptualization, technical implementation, and evaluation study with experiments and validations. This is graphically summarized in Figure 1. The design concept of the proposed development includes a pipeline of image processing methods and photogrammetry techniques to extract information of the tool trajectories from the surgical video data. This development is proposed based on a problem statement that identifies the needs in existing practices, the defined gaps in contemporary state-of-the-art technology for camera-based tracking and the specified requirements. This formalized design concept ensures systematic development of a solution that is relevant, technically feasible, and contributes towards existing level of technology.

Based on the conceptualized solution, functional capabilities of the tracking solutions will be implemented for testing with existing video data of robot-assisted minimally invasive surgeries. A homography-based pose estimation approach coupled with a model-based 2D-3D registration technique will be proposed. This approach will further combine an optimization step on the sequential estimates based on the kinematics constraint of the pivoted motion in minimally invasive surgery. The proposed workflow is digitizing the surgical video data to Tool Trajectory Data for information extraction as illustrated in Figure. 1 middle cluster.

Apart from the development aspect, experiment and validation of the developed system will be carried out to analytically to study the proposed method. The analytical study will include both qualitative and quantitative evaluation of the developed system in terms of consistency, accuracy and robotustness. Apart from inspecting the feature extraction of actual surgical video, motion estimation accuracy and robustness will be quantitative evaluated through user subject test on a robot-assisted surgical simulation platform as illustrated in the bottom cluster of Figure. 1. Benchmarking against gold standard measurement can be performed in the test platform environment.

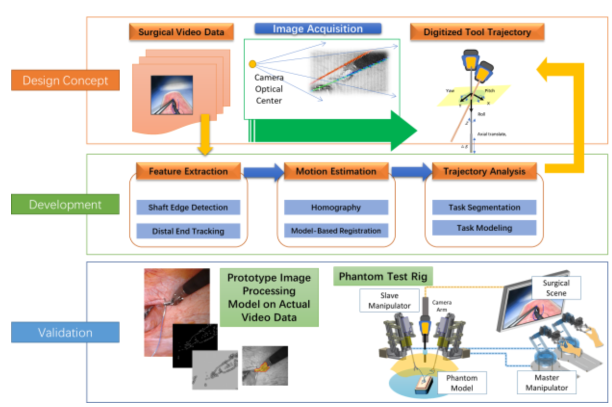

Figure 1. Overview of the proposed development of the framework towards data drive surgery

* This work was supported by the Zhejiang University/University of Illinois at Urbana-Champaign Institute, and was led by Principal Supervisor Liangjing Yang.

L Yang and S.K. Singh with the Zhejiang University/University of Illinois at Urbana-Champaign Institute, Zhejiang 314400 PRC (phone: +86 571-87572560; fax: 571-87572500; e-mail: liangjingyang@intl.zju.edu.cn).

Y. Lei is with School of Mechanical Engineering, Zhejiang University, Zhejiang Province 310058,

T Kesavadas is with the University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA

As illustrated in Fig. 2, the 2D path of the tool shaft can be recovered from the image frames of the endoscopic video data. This work focuses on feature extraction to be extended to the pose estimation. The distal end of the surgical instruments in tracked by matching a template over the image sequence [3] and the shaft edge localized using Canny Edge Detection [4] the Hough transformation [5].

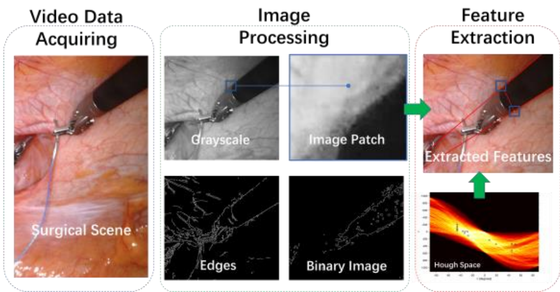

Figure 2. Image Processing from Endoscopic Surgical Scene

The distal end of the surgical tool is tracked by matching a template over the sequence of image frames. The cross-correlation $w_{cc}(u,v)$ at pixel coordinates (u, v) of the template $g(p,q)$ and an image $f(p,q)$ in a particular frame is expressed as

$$w_{cc}(u,v) = \sum_{p=0}^{P}\sum_{q=0}^{Q} g(p,q) f(p+u, q+v) \tag{1}$$

for a P x Q patch and U x V image. To account for intensity variation, we normalize $w_{cc}(u,v)$ to obtain the coefficient

$$w_{ncc}(u,v) = \left(\sum_{p=0}^{P}\sum_{q=0}^{Q}\langle G\rangle\langle F\rangle\right) \Big/ \left[\left(\sum_{p=0}^{P}\sum_{q=0}^{Q}\langle G\rangle^2\right)\left(\sum_{p=0}^{P}\sum_{q=0}^{Q}\langle F\rangle^2\right)\right]^{0.5}, \tag{2}$$

where $\langle G\rangle = \left(g(p,q) - \bar{g}\right)$ and $\langle F\rangle = f(p+u, q+v) - \bar{f}(u,v)$. Notation $\bar{g}$ and $\bar{f}$ represent the mean intensity value in the template and the window overlapping the patch, respectively. The edge of the tool shaft is detected using linear Hough Transformation. Hough space (in polar coordinates). This is a mapping of the parameters (m, b) from the linear equation

$$ax + by + c = 0, \tag{3}$$

to the parameters (r, θ) in the Hough space expressed as

$$r = x\cos\theta + y\sin\theta, \tag{4}$$

A planar feature of the cross-section view of the tool is extracted to recover 3D positional information as illustrated in Figure 3.. Configuration of the articulating end effector could be solved given the geometric model of the specific tool [6,7].

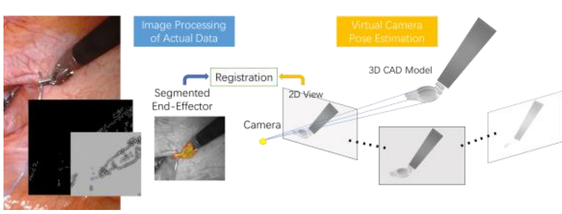

Figure 3. 3d model0-based registration 2d-3d registration

III. Preliminary Results

To evaluate the feature detection capability, quantitative measurement of the localization error is studied. For the detection of the line edges, all involved candidate points (N=4171) in the voted coordinates are used for measurement of geometrical error which is then used for failure rejection. Validation of the rejected lines are further visually inspected. The average error of the line detection is reported in Table I.

Table I. Error of Automatic Line Detection

Detected Line	*Error (Pixels)*	*# of Candidate Points*
Accepted Line	1.31	178.0
Rejected Line[a]	96.2	152.8

a. Validated through visual inspection

IV. Conclusion and Future Work

While this paper features a preliminary conceptual study, it is highly motivated by research interest and practical needs. Future work to incorporate model-based registration will completely solve the recovery of surgical tool configuration for surgical task model and analysis. Further validation of the development framework will be carried out in a controlled simulation environment. This can be done using the Mater-slave system in our laboratory to simulate surgical task procedures on phantom model while collecting video data for analysis.

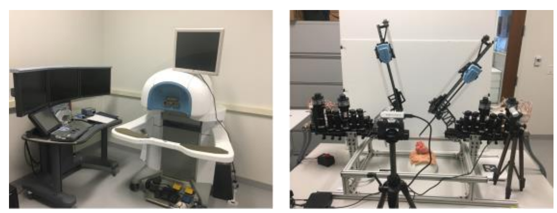

Figure 4. Image Processing from Endoscopic Surgical Scene

Highly motivated by research interests and practical needs, this work proposes an image processing approach to acquire surgical tool motion to establish the provision for data-driven understanding of surgical tasks. The long-term goal is to contribute towards the advancement of Surgical Data Science.

References

[1] L. Maier-Hein et al., "Surgical data science for next-generation interventions," vol. 1, no. 9, p. 691, 2017

[2] S. S. Vedula and G. D. J. I. S. S. Hager, "Surgical data science: the new knowledge domain," vol. 2, no. 3, pp. 109-121, 2017.

[3] L. Matthews, T. Ishikawa, and S. Baker, "The template update problem," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 6, pp. 810-815, 2004.

[4] J. Canny, "A computational approach to edge detection," IEEE Transactions on pattern analysis and machine intelligence, no. 6, pp. 679-698, 1986.

[5] D. H. Ballard, "Generalizing the Hough transform to detect arbitrary shapes," Pattern recognition, vol. 13, no. 2, pp. 111-122, 1981

[6] Sandhu, R., et al. (2009). Non-rigid 2D-3D pose estimation and 2D image segmentation. 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE

[7] J. Wang et al., "Augmented Reality Navigation With Automatic Marker-Free Image Registration Using 3-D Image Overlay for Dental Surgery," Ieee Transactions on Biomedical Engineering, vol. 61, no. 4, pp. 1295-1304, APR 2014 2014.