

IMPERIAL

SURG70006 Group Project
2024/10

Project 19 Surgical Robot Instrument Pose Estimation

Group Number: Group 8

Group Members: Jie Li, Jinling Qiu, Leen AIShekh, Yanrui Liu, Yulin Huang

Supervisor Name: Dr Stamatia (Matina) Giannarou

DEPARTMENT OF
Surgery and Cancer

Imperial College London

Contents

1	Introduction	2
1.1	Minimally invasive surgery and Robotic surgery	2
1.2	Engineering Background	2
1.2.1	DoF of da Vinci Surgical Robot	2
1.2.2	Difficulties of Surgical Tools Tracking	2
1.2.3	Limitations of OTS and EMTS	2
1.2.4	Vision-based Methods	2
1.3	Clinical aspects(TBD)	3
1.4	Problem Statement	3
2	Related Work	4
3	Proposed Methodlogy	4
4	Goals and Objectives	4
4.1	Objectives of the project	4
4.2	Hardware and software requirements(TBD)	5
5	Risk Assessment	5
6	Project Timeline	6
7	Project Management	6
7.1	Progress Monitoring	6
	Glossary	7

1 Introduction

1.1 Minimally invasive surgery and Robotic surgery

1.2 Engineering Background

1.2.1 DoF of da Vinci Surgical Robot

The da Vinci surgical robot operates with [Six Degrees of Freedom \(6DoF\)](#). These [6DoF](#) refer to the robot's ability to move and rotate in three-dimensional space, including three translational movements (up/down, left/right, forward/backward) and three rotational movements (pitch, yaw, and roll)[1]. This range of motion allows the da Vinci system to replicate the complex dexterity of a surgeon's hand for precise control over surgical instruments in confined spaces.

1.2.2 Difficulties of Surgical Tools Tracking

Pose estimation refers to finding the transformation (translation and rotation) that relates the object (or camera) coordinates in 3D space to its projection on a 2D image. The pose estimation of surgical tools has emerged as a critical job in [Robot-assisted minimally invasive surgery \(RMIS\)](#). The majority of robots in [RMIS](#) are driven by cables, resulting in kinematic input that is not always precise, since the kinematic data describes the positions of the motor rather than the real position of the joints connected to the motor via a cable[2].

1.2.3 Limitations of OTS and EMTS

[Optical Tracking System \(OTS\)](#) and [Electromagnetic Tracking System \(EMTS\)](#) are well-established methods for tracking in medical applications. [OTS](#) offers high accuracy but requires a clear line-of-sight, making it prone to errors when obstructed. [EMTS](#), while effective without line-of-sight, suffers from interference caused by metal objects and electronic devices in the operating room, leading to reduced accuracy[3].

1.2.4 Vision-based Methods

1. PnP Problem and Solvers

[Perspective-n-Point \(PnP\)](#) problem was proposed by Fischler in 1980s, which aims at estimating the position and orientation of a calibrated camera based on known 3D-to-2D point correspondences between a 3D model and their image projections[4]. The [PnP](#) is a fundamental problem of many computer vision applications, among which self-motion estimation for robots is a problem of interest. [Perspective-3-Point \(P3P\)](#) and [Efficient Perspective-n-Point](#)

(EPnP) are two common solutions to the PnP problem in camera pose estimation[5]. P3P calculates the camera pose using a 3D-2D correspondences, which provides up to four possible solutions and can be disambiguated by adding a fourth point. It is suitable for minimal data scenarios. In contrast, EPnP handles large datasets more efficiently by representing n 3D points as a weighted sum of four virtual control points, reducing computational complexity while maintaining accuracy[6].

2. Marker-based and Marker-less Methods

1.3 Clinical aspects(TBD)

1.4 Problem Statement

RMIS has come significantly in the last decade due to advances in surgical robotics such as artificial intelligence and the Da Vinci Platform. Pose estimation of surgical instruments has become an important task in RMIS. Nowadays there are many external devices like depth camera, electromagnetic trackers etc. available for space estimation in surgical instruments but they are not practical in in vivo surgeries because of space and hardware constraints[7]. There are some vision-based methods that use external markers to track the instruments. However, these methods have major limitations; the markers must always be visible in the camera's field of view and are sensitive to background changes and occlusions[2]. In this case, a vision-based markerless instrument tracking method that does not require any modifications to the hardware setup or external markers is necessary. The main aim of this project is to develop a deep learning based markerless 6DoF surgical instrument pose estimation system. The system will be designed to provide highly accurate surgical instrument 6DoF estimation without relying on external markers or complex hardware.

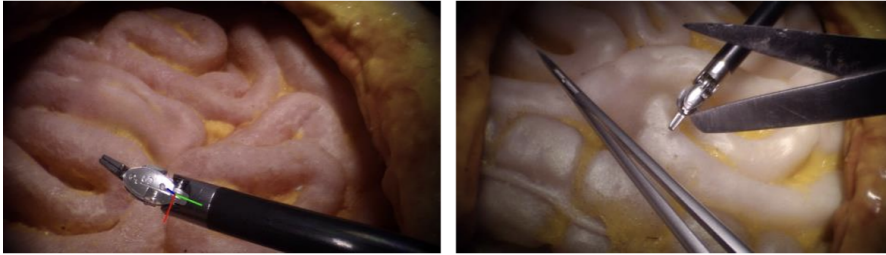


Figure 1: 6DoF surgical instrument pose estimation with (left) and without occlusion (right). [8]

2 Related Work

3 Proposed Methodology

4 Goals and Objectives

4.1 Objectives of the project

This project aims to leverage state-of-the-art deep learning models for pose estimation to accurately determine the rotation and position of surgical tools present in the surgical environment during [RMIS](#). The detailed objectives of the project are as follows:

1. **Dataset Analysis**

The first objective is to analyze the datasets which include high-quality images captured by the Da Vinci Si endoscopic stereo camera and accurate and consistent ground truth data obtained from the Hamlyn Centre.

2. **Model Development**

Based on the available datasets, we will develop deep learning-based models to detect and estimate the pose of surgical instruments in [RMIS](#).

3. **Robust Pose Estimation**

We also need to devise novel approaches to ensure accurate and robust pose estimation in the presence of challenges such as partial tool visibility, occlusions, and other variations encountered during surgery.

4. **Performance Evaluation**

Finally, the project will evaluate and validate the performance of the applied models under various degrees of occlusion, ensuring their reliability in practical surgical scenarios.

4.2 Hardware and software requirements(TBD)

5 Risk Assessment

Risk	Contingencies	Likelihood	Impact
Pose estimation unstable in complex scenes (e.g., occlusion, dynamic background)	Optimizing the dataset with more occlusion scene data	High	Very High
Low Frame Rates	Optimizing model efficiency, invest in high-quality hardware	High	Medium
Model computation time is too long for real-time simulation	Optimizing model efficiency and using hardware acceleration	High	Medium
Reliance on specific deep learning frameworks, leading to migration difficulties	Reduced binding to specific frameworks	Medium	High
Data damage or lost	Using GitHub or external storage devices to make backup	Medium	Very High
Data Privacy	Implementing robust data encryption and comply with data protection laws such as GDPR	Medium	Very High
Ethical and Regulatory	Consulting with regulatory experts and following related regulations	Very Low	High
Timeline delay	Making a detailed timeline and a thorough monitoring plan	High	High

Table 1: Risk Assessment Table

6 Project Timeline

7 Project Management

7.1 Progress Monitoring

For code part, we use GitHub for monitoring and progress management. We use GitHub's version control to branch the code (each team member manages a branch independently) to ensure smooth team collaboration. We also manually record project logs (such as meeting minutes and group activities) and combine them with timelines to ensure the project run smoothly.

References

- [1] L. Bai, J. Yang, X. Chen, *et al.*, "Solving the time-varying inverse kinematics problem for the da vinci surgical robot," *Applied Sciences*, vol. 9, no. 3, 2019, ISSN: 2076-3417. DOI: [10.3390/app9030546](https://doi.org/10.3390/app9030546). [Online]. Available: <https://www.mdpi.com/2076-3417/9/3/546>.
- [2] H. Xu, M. Runciman, J. Cartucho, C. Xu, and S. Giannarou, "Graph-based pose estimation of texture-less surgical tools for autonomous robot control," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 2731–2737. DOI: [10.1109/ICRA48891.2023.10160287](https://doi.org/10.1109/ICRA48891.2023.10160287).
- [3] A. Sorriento, M. B. Porfido, S. Mazzoleni, *et al.*, "Optical and electromagnetic tracking systems for biomedical applications: A critical review on potentialities and limitations," *IEEE Reviews in Biomedical Engineering*, vol. 13, pp. 212–232, 2020. DOI: [10.1109/RBME.2019.2939091](https://doi.org/10.1109/RBME.2019.2939091).
- [4] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, pp. 381–395, 1981. [Online]. Available: <https://api.semanticscholar.org/CorpusID:972888>.
- [5] X. X. Lu, "A review of solutions for perspective-n-point problem in camera pose estimation," *Journal of Physics: Conference Series*, vol. 1087, no. 5, p. 052009, Sep. 2018. DOI: [10.1088/1742-6596/1087/5/052009](https://doi.org/10.1088/1742-6596/1087/5/052009). [Online]. Available: <https://dx.doi.org/10.1088/1742-6596/1087/5/052009>.
- [6] V. Lepetit, F. Moreno-Noguer, and P. Fua, "Epnnp: An accurate o(n) solution to the pnp problem," *Int. J. Comput. Vision*, vol. 81, no. 2, pp. 155–166, Feb. 2009, ISSN: 0920-5691. DOI: [10.1007/s11263-008-0152-6](https://doi.org/10.1007/s11263-008-0152-6). [Online]. Available: <https://doi.org/10.1007/s11263-008-0152-6>.

- [7] J. Cartucho, C. Wang, B. Huang, D. S. Elson, A. Darzi, and S. Giannarou, “An enhanced marker pattern that achieves improved accuracy in surgical tool tracking,” *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 10, no. 4, pp. 400–408, 2022. DOI: [10.1080/21681163.2021.1997647](https://doi.org/10.1080/21681163.2021.1997647). eprint: <https://doi.org/10.1080/21681163.2021.1997647>. [Online]. Available: <https://doi.org/10.1080/21681163.2021.1997647>.
- [8] S. R. I. P. E. (SurgRIPE), *Surgical robot instrument pose estimation (surgripe)*, Synapse Project, SynID: syn51471789. Available: <https://www.synapse.org/Synapse:syn51471789/wiki/622255>, Accessed: 2024-10-22, 2024.

Glossary

6DoF Six Degrees of Freedom [2](#), [3](#)

Da Vinci Platform The da Vinci system is composed of three primary components: the patient-side cart, the surgeon console, and the vision cart. Notably, the da Vinci SP and da Vinci 5 systems stand out for offering seven degree-of-freedom (DOF) through their wristed instruments, whereas the da Vinci Xi system utilizes five DOF. The surgeon console, positioned a short distance away from the operating table, enables the surgeon to manipulate the surgical instruments and camera. The purpose of the vision cart is to provide reliable and intuitive control over the instruments, offer six DOF in terms of dexterity, and deliver immersive three-dimensional (3D) visualization. [3](#)

EMTS Electromagnetic Tracking System [2](#)

EPnP Efficient Perspective-n-Point [2](#), [3](#)

OTS Optical Tracking System [2](#)

P3P Perspective-3-Point [2](#), [3](#)

PnP Perspective-n-Point [2](#), [3](#)

RMIS Robot-assisted minimally invasive surgery [2–4](#)