Neural Radiance Fields and Robotics for the 3D Reconstruction of Subglottic Stenosis

Jinjie Sun and Julia Wiercigroch

Motivation:

Subglottic stenosis (SGS) is a rare recurrent disease characterized by the gradual narrowing of the airway between the vocal cords and trachea. The current standard of care is routine laryngeal examinations with a monoscopic flexible endoscope to monitor the length and width of the stenosis and to determine when surgical intervention will be required. However, conventional exams provide clinicians only with a top-down view of the stenosis, making it difficult to visualize and quantify the extent of the disease progression. Computer-assisted diagnosis with continuum robotics and endoscopic 3D reconstruction of the affected area has the potential to provide clinicians with an appropriate measurement tool of SGS.

Related Work:

Most research on surgical scene reconstruction has been focused on improving existing stereo methods such as structure from motion (SfM) or simultaneous localization and mapping (SLAM) but has been limited due to the sparse camera viewpoints, uniform tissue texture and minimal features [1]. Neural radiance fields (NeRF) are an emerging method for continuous scene representations using fully-connected deep networks [2]. This technique uses a set of images along with their spatial location and viewing angles to train the network to output view-dependent emitted radiance and volume density [2]. Researchers have tested NeRFs on robotic surgical scenes (EndoNeRF) and have suggested superior performance to other reconstruction methods when implemented with stereo cameras for added depth estimation [3]. However, these methods have not been tested with monoscopic laryngoscopes with extremely restricted viewpoints in the airway.

Overview of Project

For the course project, we propose to create a method for 3D reconstruction of the subglottic region with sparse viewpoints captured by a continuum robot in a virtual environment. We aim to apply regularized NeRFs (RegNeRF) which have been optimized for sparse inputs [4] and will incorporate methods to deduce depth information from monocular geometric cues (MonoSDF) [5]. By replacing the traditional laryngoscope with a continuum robot, we hope to increase the accessible camera viewpoints for 3D reconstruction and provide the algorithm with the camera location and direction.

The overall goals of our project are summarized as the following:

1. Modify mesh models to include textures from routine endoscopic exams
2. Create a virtual environment with subglottic region and continuum robotic endoscope in Unity
3. Develop 3D reconstruction methods for SGS in simulation based on RegNeRF [4], EndoNeRF [3] and MonoSDF [5]

Milestones and Timeline:

Table 1: Overview of Expected Tasks and their Deadlines

Date	Description	Deadline
Nov 16	Unity scene set up: subglottic region and robotic controlled endoscopic simulation. Mesh model modification. Explore RegNeRF, MonoSDF and EndoNeRF with existing datasets.	Nov 23
Nov 23	Improve simulation result: Combine the modified mesh with Unity models, set texture, light and other parameters to make a more realistic simulation. Collect data used for 3D reconstruction method. Test collected data on modified RegNeRF / EndoNeRF method.	Nov 30
Nov 30	Implement and test the 3D reconstruction methods performance on the dataset.	Dec 7

References:

[1] R. Wei *et al.*, "Stereo Dense Scene Reconstruction and Accurate Laparoscope Localization for Learning-Based Navigation in Robot-Assisted Surgery".

[2] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12346 LNCS, pp. 405–421, Mar. 2020, doi: 10.48550/arxiv.2003.08934.

[3] Y. Wang, Y. Long, S. H. Fan, and Q. Dou, "Neural Rendering for Stereo 3D Reconstruction of Deformable Tissues in Robotic Surgery," Jun. 2022, doi: 10.48550/arxiv.2206.15255.

[4] M. Niemeyer, J. T. Barron, B. Mildenhall, M. S. M. Sajjadi, A. Geiger, and N. Radwan, "RegNeRF: Regularizing Neural Radiance Fields for View Synthesis from Sparse Inputs," pp. 5470–5480, Dec. 2021, doi: 10.48550/arxiv.2112.00724.

[5] Z. Yu, S. Peng, M. Niemeyer, T. Sattler, and A. Geiger, "MonoSDF: Exploring Monocular Geometric Cues for Neural Implicit Surface Reconstruction," Jun. 2022, doi: 10.48550/arxiv.2206.00665.