Team 9
Tommy Liang
Jintan Zhang
Mike Fan
CIS II

Critical Review: Virtual Reality for Synergistic Surgical Training and Data Generation

Our Project Summary

The project of group 9 “VR-Guided Skull Base Surgery” is focused on Mastoidectomy which is a surgical procedure that requires the surgeon to drill out a portion of the patient’s skull to reach the inner ear. This surgery is often assisted by “steady-hand” surgical robots. These robot systems, however, have little to no awareness of the context of the surgical state, hence, can only provide limited feedback to the surgeon. An ideal accommodation for this challenging procedure is to provide the robot system with the ability of situation context-awareness so that the robot system can provide higher-level constraints to protect the patient.

To achieve “situation context-awareness”, we proposed a combination of the actual robot state and the robot state in a simulator (AMBF, Asynchronous Multi-Body Framework). The real-time surgical situation, including the patient’s pose and the robot’s status will be reconstructed in the simulator. The simulator, on the other hand, will analyze the current context, and send critical information back to the robot controller to enforce constraints.

Paper Background

In medical related studies, especially those topics that involve surgical skill analysis, two kinds of important tools -- training platforms and data generation engines have not been offered together if at all relevant. This paper proposed Asynchronous Multibody Framework Plus (AMBF+) as a solution to the problem mentioned above. AMBF+ is a very capable framework for robotics kinematics/dynamics simulations. It supports stereoscopic display on a virtual reality (VR) headset and haptic feedback for immersive surgical simulation. In this paper, specifically, a VR environment is developed with volumetric drilling features to simulate mastoidectomy. The framework can generate various kinds of data such as object poses, depth, stereo RGB images, and camera parameters, during the procedure of immersive surgical training.

Goal

The goal of this paper is to demonstrate the overall structure, methods, and workflow of the software, and evaluate its outcome as a surgical training tool and synergistic data generator.

Relevance

Our project is relevant to this paper in three main areas:

The framework. Same as the software described in the paper, our project will be using AMBF to construct a virtual surgical environment. Our project will use similar setups and workflows for VR support, haptics support, and data acquisition.

The surgical procedure. Our project also focuses on skull-based surgery, or more specifically, Mastoidectomy. One of our project goals is to reconstruct real-time Mastoidectomy surgical states in the simulator.

The technical methods. The volumetric drilling algorithm in the paper will be slightly modified and integrated into our simulator to reconstruct the surgical state of Mastoidectomy.

Hence this paper provides both technical and logistical foresights to the development of our project.

Technical Overview and Methods

Key Components:

AMBF Framework: A simulation environment that used CHAI3D and Bullet Physics for convenient surgical robot simulation. Objects are loaded according to definition inside the AMBF description file. The advantage of using ADF files is that the objects are modularized, which makes it easier to simulate surgical robots that have multiple interchangeable tools.

AMBF Plugin Design: afSimulator, afWorld, afModel, and afObject are the four main components of the simulation. In AMBF+, each component has a plugin interface that supports custom methods like graphics and physics update, and other component-dependent features.

Rendering Pipeline: OpenGL libraries are used for rendering and segmentation mask generation. Custom shader files are supported to define depth value linearization and normalization logic. Outputs are superimposed on the point cloud volume.

Data Streaming and Recording: Simulation data like stereo images, depth, segmentation mask, object poses, and camera parameters are published in real-time through Robot Operating System (ROS).

Stereo Display and VR support: Stereoscopic images (depth perception info) are displayed on a Virtual-Reality head-mounted display using Occulus Rift VR device. Virtual camera control is also supported by monitoring head-mounted display angular movement in real-time.

Haptic Feedback: Finger proxy collision algorithm is used to provide haptic feedback. A proxy sphere collides with volumes according to physics law while a goal sphere is allowed to penetrate volumes to reach commanded position. The distance between two spheres is used to compute haptic feedback.

Drilling Simulation: User's CT scan are processed by a segmentation pipeline to obtain segmented volume, which is later on rendered with 2D images to build realistic volume. Drilling tool interaction logic between the volume is defined in CHAI3D haptic feedback library. Voxel intensity is set to zero is the tip of the tool cursor collide with any non-zero-intensity voxel.

Evaluations:

Two experiments are conducted to evaluate the effectiveness of the data quality generated inside the simulator for algorithm development.

1. Anatomy Tracking:

   A state-of-the-art SLAM V3 algorithm is used to track patient's anatomy based on geometric consistency. ORB features are extracted from the stereoscopic images to perform tracking. Two sets of scenarios are considered here: static tool with camera movement, and static camera with tool movement and anatomy modification. Ground truth poses are directly available inside AMBF.

2. Depth Estimation using Transformer Networks:

   Stereoscopic images are recorded are fed into a transformer-based depth estimation network (STTR) developed by (Li et al. (2021)). Ground truth depth images are directly available inside AMBF.

Results

Two modes of evaluation were discussed as use cases of the AMBF+ simulator in the skull-base surgery, both using data generated by the simulator. In the case of the vision-based pose tracking task, the authors used simulator data to critique the tracking accuracy of ORB SLAM V3. As seen in Table 1 [1] below, the results were that the algorithm performs significantly better with a stationary camera, and a moving drill, rather than with a stationary drill and a moving camera.

Table 1. Quantitative result (mean and standard deviation of L1 error) of ORB SLAM V3 applied on the synthetic stereo microscopic data generated by the drilling simulator.

	Translation Error (mm)	**Rotation Error (deg)**
Moving camera	40.97 ± 22.40	8.44 ± 3.07
Moving drill	8.1E-1 $\pm$ 9.1E-1	3.2E-3 $\pm$ 3.6E-3

While the tracking for the moving drill case is acceptable, the magnitude of error for both the translational and the rotational case in the moving camera case is problematic if this algorithm is to be used directly in a clinical setting.

For the second mode of evaluation, the authors used simulator data to analyze a deep learning-based vision algorithm, which uses a stereo depth network to estimate the depth of a scene. After training the network from scratch using the "moving camera" data from the simulator, the "moving drill" data was tested, and the results shown in Figure 8(b) below [1].

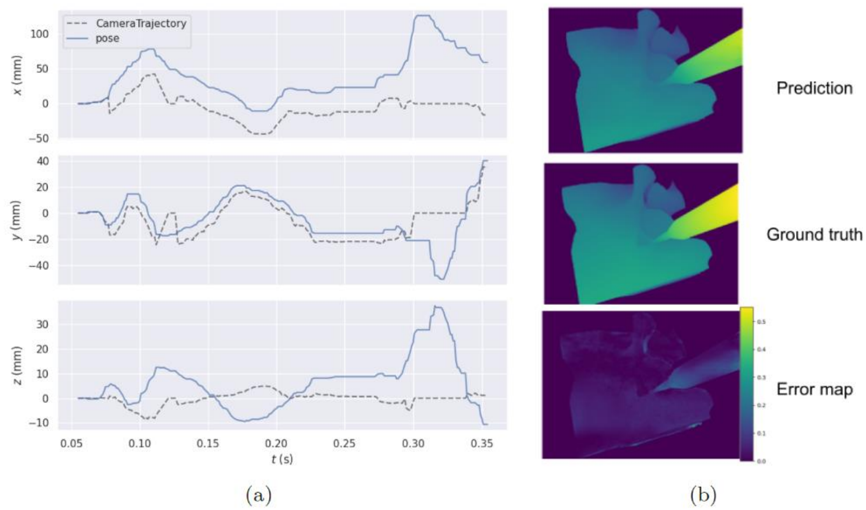

Figure 8. (a) Plot of estimated XYZ trajectory of anatomy pose tracking from ORB SLAM V3. (b) Visualization of depth prediction from STTR.

It is clear to see based on the error map, that the prediction from the deep network made a close estimation to the ground truth. Furthermore, it is important to note that the depth estimation testing data had deformations from the drilling procedure, compared to the training data, in which only the camera was moving. These results show not only the efficacy of the deep neural network, but also the simplicity and cost-effectiveness of the AMBF+ simulator. The simulator provided surgical training for this skull-base drilling procedure, and generated the data required for various modes of analysis.

Conclusion

Strengths

The authors did a good job taking a system developed for general surgical training and data generation and applying it to a specific use case that showcases these functions. In particular, the deep network use case required the drilling and deformation of the bone structure to validate the network, which would alternatively have had to been done with data collected during a surgical procedure. This again highlights the importance of the framework as a cost-effective and simplistic alternative.

Critiques

While it is understood that the scope of the paper is more an overview of a novel simulation tool and a demo of a specific use case, rather than a rigorous analysis of the effectiveness of the framework over the current standards, the paper still falls short of showcasing the key features of the framework. Arguably one of the strongest selling points that the authors mention is the ability of the system to improve surgical outcomes for patients by using the simulation framework in conjunction with a surgical robotics system. The simulation can provide unique information on the structure of the tissue from CT scans that can be used as feedback to the surgeon. While the use-case in the paper does prove the ability to generate this structure from preoperative scans, it does not show that this information can be used to improve the surgeon's capabilities when performing a surgical task. Another key feature of the framework mentioned in the paper is its ability to seamlessly handle kinematically redundant robots in the simulation environment. This is also not addressed in the use-case, as only the surgical drill is modeled in the environment rather than a full parallel-linked robot.

Key Takeaways

This paper does a fantastic job providing an overview of a novel simulation framework demoing some of its use-cases. However, these demonstrations are brief, and do not go into the key use case of integrating the framework into the information-enhanced system that surgical robots can provide to surgeons during a procedure. This ties in nicely with our group project, which seeks to model the Galen Robotics System in the virtual space and facilitate communication between the simulation framework and the robot, allowing a surgeon to use a physical robot in conjunction with the simulator, rather than just a haptics device. This sets up the groundwork necessary to conduct studies to prove the ability of the joint system in improving patient surgical outcomes.

References

1. [1] Adnan Munawar, Zhaoshuo Li, Punit Kunjam, Nimesh Nagururu, Andy S. Ding, Peter Kazanzides, Thomas Looi, Francis X. Creighton, Russell H. Taylor & Mathias Unberath (2021) Virtual reality for synergistic surgical training and data generation, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, DOI: 10.1080/21681163.2021.1999331