A Point-to-Point Registration Method Using Deep Learning

Jin Kim[1], Da-El Kim[1], MinHyuk Choi[4], Su Yang[3], Se-Ryong Kang[4], So-Young Chun[1], Ji-Yong Yoo[4],

Siyeon Noh[5], Sehun Kim[5], and Won-Jin Yi[2]

1. Interdisciplinary Program of Bioengineering, Seoul National University, Seoul, South Korea

2. Department of Oral and Maxillofacial Radiology, School of Dentistry, Seoul National University, Seoul, South Korea

3. Department of Applied Bioengineering, Graduate School of Convergence Science and Technology, Seoul National University, Seoul, South Korea

4. Department of Biomedical Radiation Sciences, Graduate School of Convergence Science and Technology, Seoul National University, Seoul, South Korea

5. Medial Image Innovation Laboratory, Seoul National University, Seoul, South Korea

kimjin116@snu.ac.kr

Abstract

In Augmented reality-assisted surgery (ARAS), achieving high registration accuracy is very important to avoid damaging high-risk tissues. However, traditional registration techniques for ARAS are based on external fiducial markers with a certain pattern as a reference, which sometimes lowers the accuracy and complicates the process. In this study, we introduce a method for an automatic registration between an optical tracking system (OTS) and a stereo camera using deep learning. In our investigation, the suggested registration approach achieved the successful Fiducial Registration Error FRE of 0.53±0.08 mm. Although the suggested registration technique's FRE satisfied clinical standards, its accuracy was worse(0.21±0.09) than that of the marker-based registration approach.

Background

Surgical navigation systems, augmented reality, virtual reality, and other applications that need exact location values make use of position tracking systems. These technologies track an object in real time and display its location on a screen. Various tracking approaches have been presented, including mechanical, inertial, and radio frequency-based tracking devices.To visualize augmented object, registration between the OTS and the camera is necessary. However, it is tough to get the identical registration marker fixture. Furthermore, it is projected that more complex registration process would make the accuracy assessment more problematic. As a result, the current work recommends utilizing deep learning model to achieve the fast and efficient registration.

Related Works

Detection tasks are treated as a regression issue in YOLOv5 as depicted in Fig. 1. Through a single inference, it predicts bounding box coordinates and class confidence. In order to extract features, YOLOv5 uses Cross Stage Partial (CSP) Darknet as well as multi-scale pooling to have a wider receptive field without adding computing complexity. YOLOv5 also separates the key context characteristics by cascading backbone and Spatial Pyramid Pooling (SPP) blocks. Path Aggregation Network was the first portion that was modified for YOLOv5. In comparison to Feature Pyramid Network (FPN), YOLOv5 includes a bottom-up route augmentation to improve the utilization of low-level features.

Methods

Deep Learning Model Setup. with a probability of 0.9 and a standard deviation of 0.2, YOLO-v5 could execute hyperparameter optimization using a genetic algorithm, which mostly employed mutation to develop offspring based on the ideal combination of all predecessors. A mutation interval was set to avoid hyperparameters from mutating to inappropriate ranges. The F1 and mAP of the model were utilized to assess and establish the optimal hyperparameters in this study.

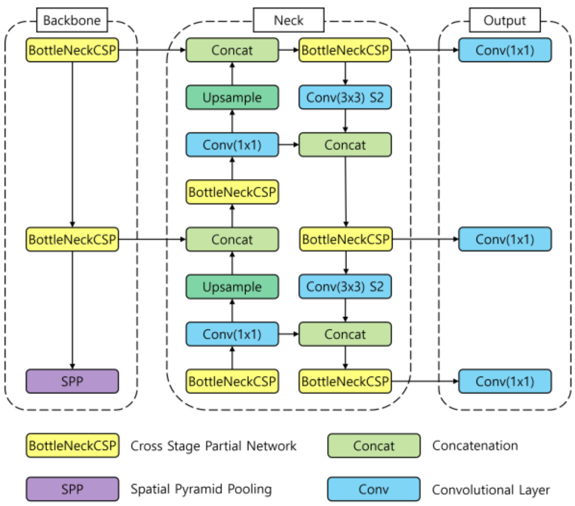

Figure 1) YOLOv5 network architecture

Model Training. In this study, we used a total of 500 RGB images of the optical markers for network training and testing. The ground truths were annotated using LabelBox. The number of training, validation, and test sets was randomly split into 300, 100, and 100 images, respectively. We used mean Average Precision (mAP) for evaluating the model's performance and FRE for the registration performance.

Registration. The transform matrix between the OTS and the stereo camera can be approximated after acquiring the image frame's 3D coordinates from both modality. Fig. 2(a) shows the overall schematics for the suggested method. After training the deep learning model, each optical marker's 3D position is obtained using OTS, and 2D position applying the deep learning model on the stereo camera. Afterwards, 2D pixel position can be translated to 3D position using the characteristics of the stereo camera, as described in Fig. 2(b). The standard procedure is to match the feature points first, then construct the transformation matrix using an iterative closest point (ICP) scheme. The entire registration process is automated and has to be completed 24 times to find the smallest FRE among the every possible permutations of four markers.

TABLE I. FIDUCIAL REGISTRATION ERROR (FRE) AND THE REGISTRATION TIME

	Proposed registration	***Traditional registration***
FRE	0.53±0.08	0.21±0.09
Registration time(s)	0.04±0.01	82.3±17.3

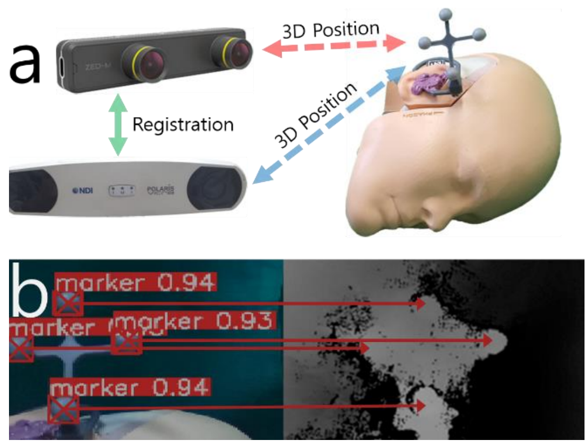

Figure 2) (a) Schematic representation of the suggested registration (b) Depth information acquisition from the depth view

Results and Conclusion

We compared the performance of novel registration method with traditional registration method. The experimental results show that our method yields slightly larger FRE than the traditional registration method, but it is much faster than the traditional method (Table 1). Experimental results demonstrate that our method can obtain efficient performance for ARAS.

Acknowledgements

This work was supported by the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (1711137883, KMDF_PR_20200901_0011, 1711138289, RS-2020-KD00014).

References

1. Lai, Marco, Caifeng Shan, and Peter HN de With. "Hand-eye camera calibration with an optical tracking system." Proceedings of the 12th International Conference on Distributed Smart Cameras. 2018.

2. Jocher, Glenn, et al. "yolov5." Code repository https://github.com/ultralytics/yolov5 (2020).