

Combination Technology of SLAM And VR For Minimally Invasive Surgery

Nuannuan Zhang *

Shenzhen College of International Education, Shenzhen, China

* Corresponding Author Email: s23705.Zhang@stu.scie.com.cn

Abstract. With the development of the internet, the combination technology of SLAM and AR is becoming increasingly important in various fields. Simultaneous localization and mapping (SLAM) and augmented reality (AR) have been used separately in surgeries. AR techniques can help surgeons see the patient's internal anatomy in conjunction with laparoscopic video images and it's convenient for hand-eye coordination. However, only use AR technology in minimally invasive surgery still has some disadvantages, like the view is limited by the camera and geometric deformation. Using SLAM for surgery, on the other hand, offers advantages such as potential improvements in accuracy and the estimation of dynamic information, along with the ability to infer deformation and stiffness. In this paper, the principles of SLAM and AR technology will be introduced first, then the combined technology will be talked about how to use in the minimally invasive surgery and finally about how to improve the recent method in minimally invasive surgery. This essay will have important value about SLAM and AR's research and applying.

Keywords: Augmented reality; visual SLAM; minimally invasive surgery.

1. Introduction

The minimally invasive procedure, also known as "keyhole surgery," involves cutting a small opening in the abdomen or through a natural cavity passage. Small surgical instruments and a tiny camera are inserted into the abdominal cavity. The camera captures an image of the interior, which is then displayed as an enlarged image on a screen. With a minimally invasive approach, large incisions are no longer necessary for abdominal surgery in many cases. Patients have less pain, less scarring, less blood loss, and less trauma to the body than traditional "open" surgery. X-ray fluoroscopic scanning assisted positioning equipment was used to determine the relative position of surgical instruments and improve accuracy. However, its inherent shortcomings cannot be ignored, such as the side effects of radiation. Additionally, the X-ray fluoroscopy system only provides a single plane view during positioning. When the position of the instrument needs to be observed in a multi-plane view, the C-arm position needs to be repeatedly adjusted for scanning and positioning, resulting in the interruption of the X-ray fluoroscopy.

The computational task of tracking an agent's location in an unknown environment while creating or updating a map of it is known as simultaneous localization and mapping (SLAM). When used in minimally invasive surgery, SLAM can provide the real-time model to assist doctors in deciding how to operate.

An interactive experience that blends computer-generated material and the actual world is known as augmented reality (AR). Multiple sensory forms, such as visual, aural, tactile, and olfactory, can be bridged by the content. Augmented Reality (AR) provides experiences with enhanced perception and can be utilized to improve natural environments or situations. Advanced augmented reality technologies (such as object identification, computer vision, and AR camera integration with smartphone apps) enable users to interact with and manipulate their immediate environment digitally.

2. The principle of the combined technology of SLAM and AR

This part will introduce the fundamentals of SLAM and AR technology and explain how they work together. The idea behind SLAM technology is to gather environmental data from sensors like

lidar, cameras, and others. Then, using algorithms combine this data to determine the device's location within the environment and create an environment map. The technology known as augmented reality (AR) blends virtual and physical world information. Multimedia, 3D modelling, real-time tracking and registration, sensing, and other technologies are used in this technology. Its basic idea is to simulate the real world and then apply computer-generated virtual data-text, pictures, music, videos, 3D models, and so on—to it. The comprehension of the floor, walls, and barriers is the most fundamental kind of SLAM technology. These days, the majority of AR SLAM solutions, such as ARKit, locate AR-friendly objects around us using position tracking and floor recognition. They ultimately fail to respond appropriately because they are unable to perceive what is going on around them. On the other hand, sophisticated SLAM technologies, such as Google Tango, build an interactive 3D world and alert us to objects, walls, and floors, making everything in our immediate vicinity interactive.

There are different examples of SLAM, LSD-SLAM, for instance, uses draw graphs to depict the reconstructed 3D world and shows data correlation with probabilistic semi-dense maps. DTAM, on the other hand, can reconstruct a dense map in real time with just one RGB camera [1].

Semantic SLAM can be utilized to improve the geometric map of SLAM by offering a comprehensive understanding of the surrounding environment. This approach has significant potential for usage in robotic applications, including AR, task planning, and object interaction [1]. The latest advances in deep learning have made it possible to use a neural network for Semantic SLAM, which has resulted in a model with better accuracy, faster performance, and lower power consumption. Some apply semantic SLAM to identify and eliminate dynamic items (e.g., DetectSLAM) to achieve a more exact rebuilt map and a more accurate camera pose. Visual SLAM, or VSLAM for short, is SLAM using a camera as a sensor. As seen in Figure 1, the camera operates as a sensor that gathers a great deal of data, recognizes objects well, and senses at a considerably better resolution than radar [2]. It is also less expensive, easier to take around, and easier to use.

In 2009, Grasa et al. calculated the three-dimensional information of the scene from successive photographs after using monocular laparoscopy to create a sparse map of the environment from SLAM. The Mountney P team proposed the use of SLAM systems for endoscope positioning and 3D modeling in minimally invasive surgery in 2009 and 2010, respectively. The monocular system was based on the SLAM framework to establish a motion model to accurately estimate the motion of the camera and soft tissue [3].

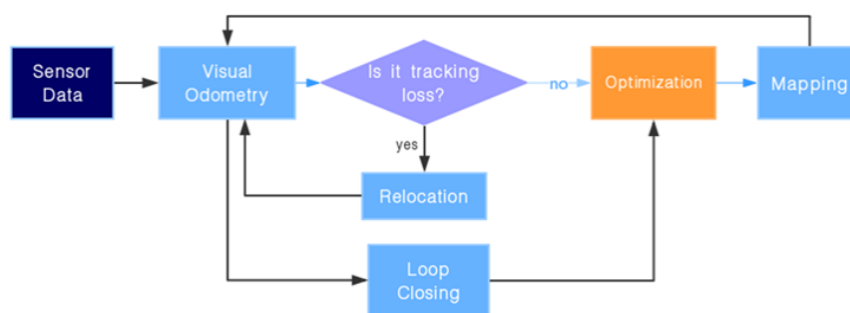


Fig. 1 Technical Framework of VSLAM, A Review of VSLAM Technology Applied in Augmented Reality

Because augmented reality (AR) can provide patients vivid visualizations of medical data at the site of their procedure, it has already become more and more common in computer-integrated surgeries. For three-dimensional reconstruction, it identifies and marks the boundaries and layers of tissues and organs. Then, using SLAM technology, the colored objects may be located and recreated as models. Early surgical augmented reality research mostly concentrated on multi-modal information combining and offered overlays based on displays. Consequently, AR made it possible to use 3D data before and after interventions [4]. A robotic operating room (OR) system using RGB-D semantic SLAM is evaluated. In a dynamic OR environment, the system tracks robot trajectories and produces

a dense semantic reconstruction of the operating table. Thanks to its precise trajectory estimate and semantic scene interpretation, the SLAM system significantly enhances robotic perception of its surroundings. The technology may help with surgical navigation, robot setup, and human-robot interaction [5]. Leveton et al. employed augmented reality pictures in 1997 to help surgeons do surgery. Fuchs et al. successfully visualized laparoscopic surgery in 1998 using augmented reality technology.

The usage of AR in operating rooms has the drawback that magnetic or metallic objects can interfere with electromagnetic signal transmission. The placement and quantity of markers or sensors, the size of the navigation area, and soft tissue deformation can all have a substantial impact on the accuracy of OTSs and EMTSs in complex surgical navigation scenarios. When using AR alone for surgery is insufficient, wireless imagery such as SLAM can be employed. Additionally, with those two technologies, surgeons can respond more quickly in the event of an emergency during surgery.

3. Combination of AR and SLAM used in minimally invasive surgery

Hospitals are gradually reconstructing the surgical region and localizing the endoscope in a minimally invasive surgical setting by employing the SLAM technique. The feature information of the image in the lumen environment is found using the K-Means algorithm in conjunction with the Super Point technique. The proportion of effective points grew by 25.28%, the mistake matching rate dropped by 0.64%, the extraction time dropped by 1.98%, and the logarithm of successful matching points increased by 32.69% with Super points [6].

Since there is little tissue movement and deformation during minimally invasive surgery on the abdomen, SLAM is widely used in these procedures. SLAM was initially incorporated into the algorithm as a positioning auxiliary at the outset of the application. In 2009, Grasa evaluated the three-dimensional information of the scene from photos obtained from SLAM and created a map of the environment using monocular laparoscopy. The Mountney P team proposed the use of SLAM systems based on monocular and binocular endoscopes in minimally invasive surgery for 3D modelling in 2009 and 2010, respectively [6]. As a result, the monocular system, which was based on the SLAM framework, established a motion model to accurately estimate the motion of the camera and soft tissue.

Forming a map of the endoscope's position information is challenging due to its irregular moving procedure. Nevertheless, the camera's SLAM system can display the estimate algorithm utilised in endoscopic minimally invasive surgery by employing the lumen mapping results of image feature points as illustrated in Figure 2 and image data from the actual minimally invasive surgery environment for the endoscopic frame.

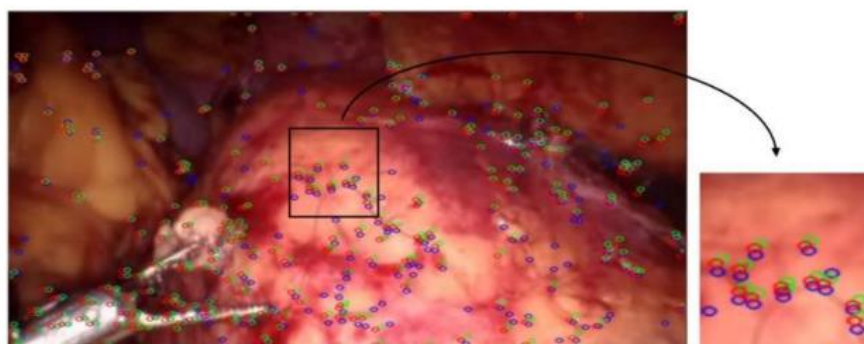


Fig. 2 Feature point mapping of cavity image, study of visual SLAM methods in minimally invasive surgery

The feature point of the current frame is shown as a blue point in Figure 2, while the feature point of the following frame is shown as a green point. The red point represents the location of the green point following the rotation matrix and leftward motion vector, which explains why the camera moves

when shooting in different frames. With an average inaccuracy of 5.43 pixels, the gap between blue and red dots is less than that between green dots [6]. The lumen model is built more efficiently.

Providing navigational information during intervention can be solved by using AR technology. When it comes to minimally invasive surgery, doctors constantly have difficulties in checking the entire view of the organs or the wounds. With the help of AR system, doctors can easily solve the problem by looking at the 3D-model built in the computer. AR systems often rely on tracking of external markers, which can make them less useful as they require line-of-sight and invasive tissue implantation [4]. In the context of hybrid operating rooms that allow real-time input of all instruments in space and in connection to anatomical structures, augmented reality (AR) seeks to overlay simulated bony structures and preplanned screw courses on patients [7]. For instance, the colour of synovial cysts makes it easy to distinguish them from the surrounding bone and neural tissue. It also enables the best possible trajectory and laminotomy site to be reached by the pathologist (Figure 3) [8].

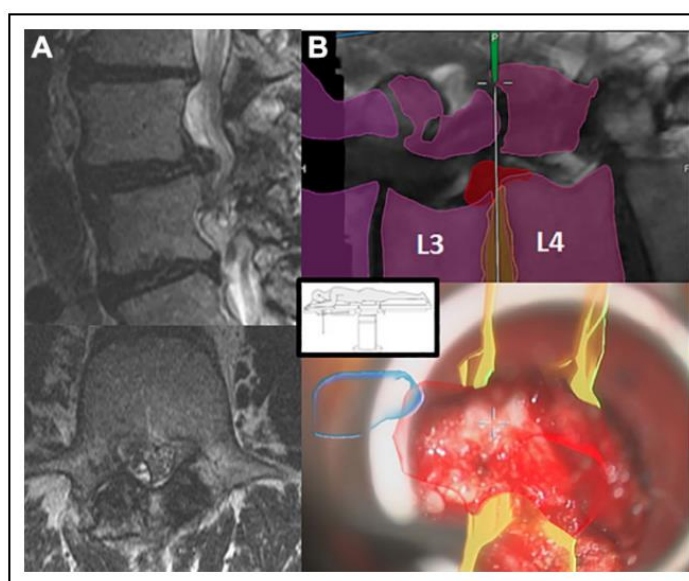


Fig. 3 The organ model built by AR, with shown optimal trajectory and laminotomy site, Evolving Navigation, Robotics, and Augmented Reality in Minimally Invasive Spine Surgery

SLAM algorithm comes from robotics research and can provide geometric positioning for AR systems. The SLAM algorithm can build a 3D map of the environment while tracking the position and orientation of the camera in that environment. The algorithm estimates the location of sensors (built into cameras, phones, or goggles) while modeling the environment to create a map.

By using tubular retractors intraoperatively, AR makes it possible to identify the pars, pedicle, and disc even when they are not in the exact field of vision. This enhances safety and confidence for the surgeon while performing manoeuvres for tissue or bone excision. With the aid of these technologies, the surgeon can concentrate on the current operation rather than constantly switching between the patient and navigation screens. Research on using augmented reality (AR) in real patients is still in progress. It involves evaluating factors including radiation exposure, software planning, operating room layout, image capture, and navigation instruments.

Purely AR system used in tele-medicine still cannot provide or model the location information for operating tools. In this circumstance, SLAM, with the ability of track the position of devices and mapping the 3D model, is necessary for helping doctors to know and operate those machines [9]. In Figure 4, the process of organ tracking of AR applied and based on vision is demonstrated. After having trachea and common carotid artery exposed in lens, SLAM can be applied to build a 3D map of the field, while has the ability to keep AR images. By using a monocular surgical camera, a method called 3D-2D registration method can mix 3D AR image with 2D image. The AR images of those organs were semi-automatically overlaid. Color and axis of those organs used for first overlay. The

location of the AR image can be changed manually by translation, rotation and zoom with a mouse [10].

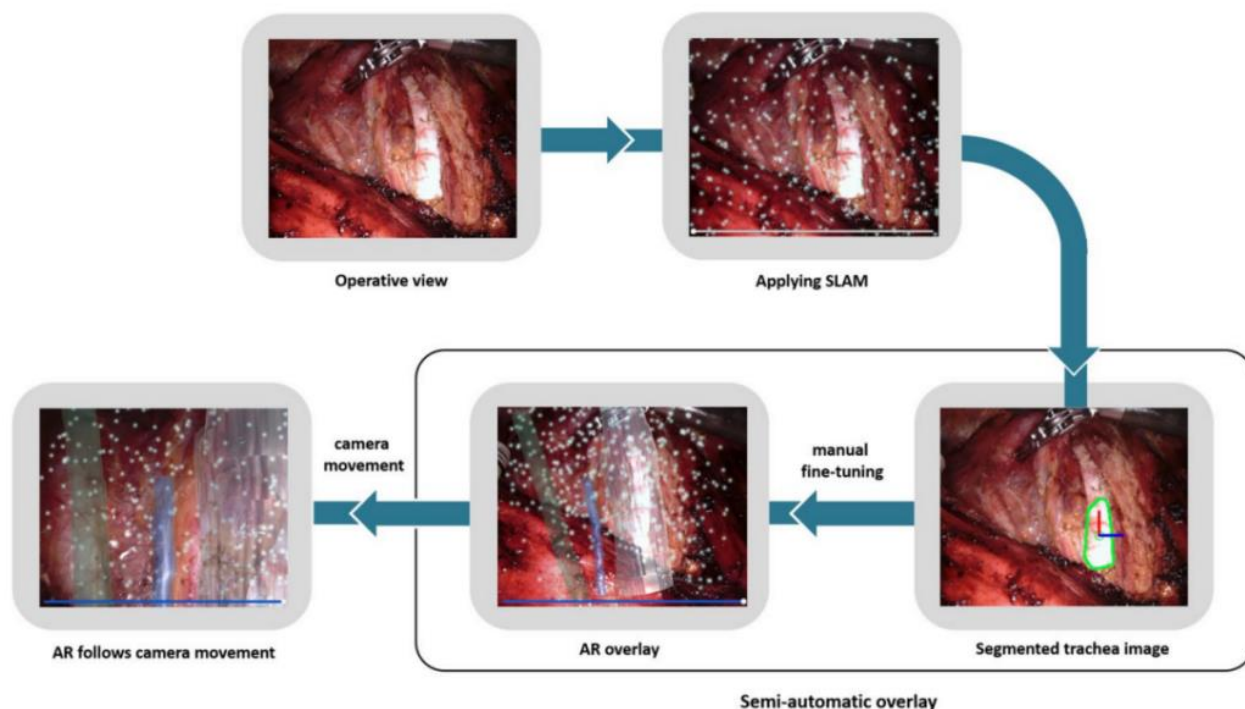


Fig. 4 The process of organ tracking, Vision-based tracking system for augmented reality to localize recurrent laryngeal nerve during robotic thyroid surgery

4. Conclusion

This paper presents how the combined technology works in minimally invasive surgery. It shows how AR and SLAM works separately in surgeries in recent times. By superimposing preoperative models—such as tumours with resection margins and vessels—on laparoscopic pictures, the created AR system can offer intuitive visualisation for navigation, enabling surgeons to see the patient's anatomy under the surface during operation. And SLAM can provide real-time modeling, track the position and orientation of the camera in the environment. After combining those two technologies, doctors can not only see the organs in an AR headset, but also react to the changes that occur in the organs as well, which decrease the difficulties of the operation.

For Future prospection, with the help of AR and SLAM technology, doctors can operate even far-distant tele-machine based on the model built by SLAM, also with a higher accuracy when curing. With improved navigation technologies combined with current technology, these systems have enormous potential to augment the expertise of spine surgeons, improving patient outcomes and safety while lowering expenses and difficulties. With further research in clinic operation and technology innovation, the combined technology of AR and SLAM have the potential to become timesaving and accuracy-improving tool for minimally invasive surgery.

References

- [1] Marchesi G, Eichhorn C, Plecher D A, et al. Envsam: Combining slam systems and neural networks to improve the environment fusion in AR applications[J]. ISPRS International Journal of Geo-Information, 2021, 10(11): 772.
- [2] Tang B, Cao S. A review of VSLAM technology applied in augmented reality[C]//IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2020, 782(4): 042014.

- [3] Lin B. Visual SLAM and Surface Reconstruction for Abdominal Minimally Invasive Surgery[M]. University of South Florida, 2015.
- [4] Fotouhi J, Mehrfard A, Song T, et al. Development and pre-clinical analysis of spatiotemporal-aware augmented reality in orthopedic interventions [J]. IEEE transactions on medical imaging, 2020, 40(2): 765-778.
- [5] Gao C, Rabindran D, Mohareri O. RGB-D Semantic SLAM for Surgical Robot Navigation in the Operating Room [J]. arXiv preprint arXiv:2204.05467, 2022.
- [6] Deng L, Liu Z, Zhang T, et al. Study of visual SLAM methods in minimally invasive surgery [J]. Mathematical Biosciences and Engineering, 2023, 20(3): 4388-4402.
- [7] Burström G, Nachabe R, Persson O, et al. Augmented and virtual reality instrument tracking for minimally invasive spine surgery: a feasibility and accuracy study [J]. Spine, 2019, 44(15): 1097-1104.
- [8] Hussain I, Cosar M, Kirnaz S, et al. Evolving navigation, robotics, and augmented reality in minimally invasive spine surgery[J]. Global Spine Journal, 2020, 10(2): 22S-33S.
- [9] Reis G, Yilmaz M, Rambach J, et al. Mixed reality applications in urology: Requirements and future potential[J]. Annals of Medicine and Surgery, 2021, 66.
- [10] Lee D, Yu H W, Kim S, et al. Vision-based tracking system for augmented reality to localize recurrent laryngeal nerve during robotic thyroid surgery [J]. Scientific reports, 2020, 10(1): 8437.