Positioning Methods and Development of Different Methods of Biomimetic Robot Based on Endoscope

Xinrui Li [1], Wenzhe Pu [2,*] and Yifei Wang [3]

[1] College of Computer and Information Science College of Software, Southwest University, chongqing, 400715, China

[2] School of Physics and Technology, Wuhan University, Wuhan, 430027, China

[3]School of Information Science and Engineering, Wuhan University of Science and Technology, wuhan, 430065, China

lncs@springer.com

These authors contributed equally

Abstract. Endoscopic technology continues to evolve. With the advent of digital imaging technology and miniaturization, the application of endoscopy in diagnosis and treatment has been further expanded. Up to now, it has developed capsule endoscopy, insertion of single fiber endoscopy and other more mature technical achievements. Modern intelligent transformation forms endoscopic robotics. This process has an irreplaceable important role in the field of clinical minimally invasive surgery and great significance for the field of medical science. This paper mainly analyzes and compares the information acquisition and processing technological achievements of endoscopic robots such as vision tracking, SLAM (Simultaneous localization and mapping), monocular ranging, camera calibration, etc. according to the degree of efficiency and ubiquity by data analysis and lateral comparison in recent years. Discussion of more research space and research value, and expectation to the future of endoscopic robots are also included. With the application of artificial intelligence and machine learning, endoscopic robots are also expected to be automated and intelligent. Robots can learn and optimize operational skills, improve surgical efficiency, and can make more accurate diagnosis and treatment recommendations based on a patient's specific situation.

Keywords: Surgical Robot; Endoscopy; Robotic System.

1 Introduction

Compared with traditional surgery, microsurgery is less invasive, has a faster recovery time, and can also greatly reduces the pain of patients. Endoscopic bionic robot is one of the typical medical devices used in microsurgery, and after nearly two hundred years of development history, it is now used in a variety of clinical medical fields. At present, endoscopic bionic robots can realize the basic examination, diagnosis, and treatment functions of endoscopes in medicine, as well as the stability of insertion and the adaptability to complex environments when moving.

Y. Wang (ed.), *Proceedings of the 2024 International Conference on Artificial Intelligence and Communication (ICAIC 2024)*, Advances in Intelligent Systems Research 185,
https://doi.org/10.2991/978-94-6463-512-6_17

At present, several papers have been published to continuously improve the endoscopic bionic robot in different fields with different methods. Among them, localization is one of the basic functions of endoscopic bionic robots. Since endoscopic bionic robots are often used to perform miniature movements and any error in medical treatment may lead to tragedy, it is crucial to ensure the accuracy and precision of localization.

Currently, in terms of endoscopic bionic robot localization, Wang et al. proposed a fast position measurement algorithm and designed a surgical instrument position measurement scheme for laparoscopic surgical instrument position measurement based on existing surgical instrument position measurement methods [1][2]; Hu et al. based on the endoscopic monocular vision to establish a pinhole model for monocular machine vision measurements, which is used to derive the distance between the endoscope and the target as well as the target size for localization [3]; Guo et al. proposed a method for modeling the internal structure of the intestine based on the triangular dissection method and the SLAM algorithm improved by the OctoMap method [4].

Due to the large number of papers related to endoscopic bionic robot positioning and the relatively sparse number of papers used to organize the content.

This paper will focus on organizing and summarizing the three methods mentioned above. It aims to analyze the advantages and disadvantages of these three methods and the localization requirements of the application areas, to propose a general direction for improving the accuracy of the procedure and to predict the development prospects of endoscopic bionic robots. Expect that this study could provide a certain direction for the development of endoscopic bionic robot localization.

2 Main Body

2.1 SLAM Synchronized Positioning and Map Creation

Introduction to SLAM. SLAM, also known as CML (Concurrent Mapping and Localization), is the Synchronized Localization and Map Building, or concurrent mapping and localization. It describes the procedure by which a moving item uses data from sensors to determine its position and create a map of its surroundings.

It is the process of constructing a map of the environment while calculating its own position based on the information from sensors. If a robot is put into an unknown location in an unknown environment, SLAM can enable the robot to map out the environment step by step while it is being placed in the unknown location.

The development of SLAM technology has made it a new solution for the reconstruction of interior scopes called Visual Sensor-based SLAM.

Visual SLAM methods based on visual sensors are mainly categorized into monocular SAM, binocular SLAM, and RGB-D (color and depth) SLAM depending on the sensors.

SLAM systems can collect data from the environment using various sensors. These sensors are laser-based, acoustic and vision-based. There are several types of vision-based sensors, including monocular, stereo, event-based, omnidirectional, and (RGB-D) cameras.

Advantages and Disadvantages of Slam Technology. Robots with visual sensors use visual data from the camera to estimate the position and orientation of the robot relative to its surroundings.

The process of SLAM using visual sensors is known as visual SLAM (VSLAM). The visual data hardware used in SLAM is cheaper, more intuitive for target detection and tracking, and provides rich visual and semantic information.

The captured images (or video frames) can also be used for vision-based applications, including semantic segmentation and target detection. The above features have made VSLAM a popular direction in machine-anthropology and have led to the emergence of a new field of machine-anthropology and computer vision (CV). It has led to a lot of research and studies by experts in robotics and computer vision over the past decades. As a result, VSLAM has been used in various applications that require reconstruction of 3D (three-dimensional) models of environments, such as autonomous driving, augmented reality (AR), and services robots [5]. Laser SLAM based on the particle filtering framework, the RBpf (rao-blackwellised particle filtering) algorithm, separates the localization and mapping processes. Each particle carries a map along with the odometry and laser information, and the RBpf particle filtering method separates the localization and mapping processes, first localization and then mapping. Each particle carries a map, which requires less computation and higher accuracy to construct a map of a small scene.

However, in order to simplify the position solving problem for mobile robots, most SLAM algorithms usually treat the scene as static, which results in the robot treating dynamic objects as static in a dynamic scene. This causes the robot in a dynamic scene to treat the dynamic objects as static as the background environment, so the features extracted from the dynamic objects are treated as static feature points. As a result, the dynamic object's features that were extracted will be considered static feature points, which introduces a large amount of dynamic error and leads to inaccurate localization results and the inaccurate positioning result of the robot [6]. Meanwhile, in the high-definition mapping, the positional updating is not good in the stationary state. The problem is that there is too much noise and vibration. Currently, slam technology is widely used in all aspects of life, such as AR, robotics, robotics, and unmanned driving. It is just like cell phone positioning in the era of mobile Internet.

Endoscopic Slam Technology Development Prospects. With the continuous improvement of SLAM algorithms and hardware technology, endoscopic SLAM technology will be able to achieve more accurate navigation and localization to help doctors find and treat lesions. Meanwhile, endoscopic SLAM technology is expected to realize real-time 3D reconstruction of internal organs. This will provide doctors with a more intuitive surgical field of view and a better sense of the environment, which will help to improve the efficiency and success of surgery. Combined with robotic technology, endoscopic SLAM can realize self-guided navigation and operation, and provide a better solution for robotic surgery.

With robotic technology, endoscopic SLAM can realize self-navigation and operation, providing more reliable localization and environmental awareness for robotic-assisted surgery.

This further enhances the accuracy and safety of surgery. In addition, endoscopic SLAM technology can be combined with augmented reality technology to provide real-time virtual information for doctors.

In addition, endoscopic SLAM technology can be combined with augmented reality to provide real-time virtual information deficit, such as lesion location, blood vessel distribution, etc. This will help doctors to better navigate and make decisions during surgery.

In summary, the development of SLAM technology for endoscopy has a very promising future. It is expected to bring more accurate, intelligent and efficient diagnostic and treatment tools to the medical field. It can also provide better medical services to patients [7].

Enhancement of Slam Technology. In view of the challenges faced by deep learning-based visual SLAM systems when applied in real life, there are several key development directions.

First, real-time performance improvement is crucial. Although deep learning methods can improve the robustness and accuracy of the system in dynamic environments, their incorporation will lead to a reduction in the real-time performance of the system. Therefore, it is important to reduce the time consumption of the system to improve the real-time performance.

Secondly, more precise object segmentation is needed, especially for small objects. Methods based on deep learning often lose details when dealing with small objects in complex dynamic environments.

This affects the accuracy of posture estimation, so it is necessary to improve the accuracy of network segmentation. In addition, the retention of more static features is also an important consideration. In deep learning networks, static feature points in dynamic areas are often mistakenly eliminated when dynamic feature points are eliminated. The information contained in these static points is beneficial to the follow-up camera tracking and map construction.

Therefore, reducing the waste of static feature points in the dynamic area by using techniques such as background restoration can help improve the accuracy of the system. The system will be more accurate. Finally, hardware upgrades are essential. The deep learning visual SLAM system relies heavily on the hardware performance of computers, and hardware upgrade can greatly improve the speed of image processing in the network, thus effectively improving the real-time performance of the system [7].

2.2 Visual Analysis and Image Processing

Visual analysis and image processing is an important link in the clinical application of endoscopy, which is different from the map creation function of slam. Unlike the map creation function of slam, it is realized by extracting and calculating the feature points from the camera-acquired images to analyze the three-dimensional model of the intestine based on the two-dimensional planar samples visualized by the camera.

The three-dimensional model of the intestine is analyzed based on the two-dimensional planar samples visualized by the camera, so as to complete the accurate measurement of the intestine.

Common visual analysis algorithms of endoscopy, such as the PNP (Perspective-n-Point) equation, can accomplish the measurement of the specific position of the feature points in three-dimensional space by using the feature points in the image. There is no obvious advantage or disadvantage in measuring the location of feature points in three-dimensional space by means of image feature points.

Triangulation is the most widely used endoscopy visual analysis algorithm. Its advantage is that it can efficiently and concisely determine the location of feature points in three-dimensional space from the linked map feature points [8]. Nevertheless, in theory, the inaccurate triangulation of monocular SLAM algorithm will lead to errors in the map construction process; In practice, due to the fog and darkness in the actual intestinal environment, there are fewer feature points [1], which may require manual interpolation to make up for the errors.

In practical application, due to factors such as fog and dimness in the actual intestinal environment, the camera is supposed to capture even fewer characteristic points [3], which may require manual interpolation to make up the points, and the cost and reliability are still open. The cost and reliability are still open to discussion.

The fundamental reason is that most open-source visual SLAM systems are based on scene point features, but when the texture is relatively single or the image is fuzzy, the scene point features can be used as the basis for the SLAM system. However, when the texture is relatively single or the image is blurred, the small number of point features in the above cases will seriously affect the accuracy of the posture estimation. Research has shown that the number of line features in the scene is much larger than the number of point features, and the line features are better able to characterize the image information [9].

The development of triangulation algorithms under dynamic monocular vision is conducive to the precise localization of lesions. If the algorithm that automatically interpolates the complementary points according to the lightness and darkness of the feature points can be developed to reduce the need for manual involvement in endoscope use scenarios, the cost of endoscope use can be further reduced and the efficiency of surgery can be improved. Alternatively, SLAM systems with line feature extraction techniques can be added to the algorithms of endoscopic robots.

2.3 Monocular Distance Measurement and Posture Determination

An important form of binocular stereo vision machine is based on the principle of parallax to obtain 3D geometry information of objects through multi-frame images.
The information is obtained from multiple frames of images. However, under the same conditions, it is obvious that the prospect of smaller and lighter monocular rarefaction mirrors is much broader.

However, due to the lack of multiview information, the problem of positional calibration arises. The technology of monocular endoscopy microscope and its posture calibration problem has made the endoscopy microscope more popular in the field.

The technology of monocular endoscopy and the related technology of its posture calibration problem has become an important issue in the field of endoscopy [10].

3 Conclusion

The three algorithms in this paper have been applied in endoscopy and have the advantages of high efficiency, accuracy and stability for visual inspection. The advantages for visual detection are high efficiency, accuracy, stability, reliability, and visualization. However, the disadvantages are high initial cost, high requirements for environmental conditions, possible adaptation problems, and the possibility of external interference.

The advantages of monocular distance measurement are that it is less costly and less demanding on computational resources. The system is versatile, but it has to constantly update and maintain a large sample database to ensure that the system achieves a high recognition rate.

However, it must constantly update and maintain a large sample database to ensure the system achieves a high recognition rate, be unable to judge non-standard obstacles; the distance in reality is not as measured and the result is of low accuracy. They can be combined when designing for a variety of applications. While today's endoscopic robots are receiving a great deal of attention, it is important to recognize that the current research on endoscopic robots has not yet been completed.

At the same time, it is important to recognize the current state of research and some of the problems that exist. For example, a simple, easy-to-control, and flexible endoscopic system is a good idea. It is easy for the simple, easy-to-control and flexible internal rare microscope system to be influenced by the characteristics of the target surface and has more limitations, while for the relatively complicated and flexible internal rare microscope system, it is more limited.

For the relatively complex and complicated structure, different cooperative targets need to be customized according to different situations, and the requirements for target fabrication and installation are high, and contact or modification with the object to be used is not allowed. The target can be fabricated and mounted with high requirements and does not allow for defects such as contact with or alteration of the object to be used. However, it is believed that endoscopic systems will be able to develop better and have more applications in the medical field in the coming years.

References

1. Wang, X., Chen, B., Ju, F.: Research on visual measurement methods in robot-assisted laparoscopic surgery. Machinery & Electronics **39**(09), 60-65 (2019)

2. Liu, Z, Chen, K, Chen, Z.: Research Status and Development of Endoscopic Bionic Robot . Machinery Design & Manufacture **6**, 161-163 (2009)
3. Hu, T.: Distance measurement based on endoscopic monocular visual surgical navigation . China Tissue Engineering Research and Clinical Rehabilitation **22**, 4241-4245 (2008)
4. Guo, W.: Human body map construction of endoscopic intestinal surgery machine based on triangulation. Robot **43**(04), 395-405 (2019)
5. Ali Tourani, Hriday Bavle, Jose Luis Sanchez-Lopez, Holger Voos.: Visual SLAM: What are the Current Trends and What to Expect. Sensors (2022)
6. Liu, M.: 3D SLAM algorithm based on geometric constraints of feature points in dynamic scenes. Journal of Beihang University (2023)
7. Luo, Y.: Under the dynamic environment based on the depth study of visual SLAM research review . Journal of semiconductor optoelectronic **1**, 1-10 (2024)
8. Miao, D., Chen, N.: A monocular active visual attitude measurement method for small space. Modern Manufacturing Engineering (2024)
9. Jia, S., Ding, M., Zhang, G.: Visual SLAM Algorithm based on point-and-line features in RTM framework. Robot **41**(3), 384-391 (2019)
10. Anita, R.: Bimodal Camera Pose Prediction for Endoscopy. IEEE Transactions on Medical Robotics and Bionics (2023)