Navigation system for oral surgery based on slam technology

Chengen Cen [1, 4, †], Hao Wang [2, †] and Shiqi Wu [3, †]

1School of International Education, Wuhan University of Technology, Wuhan, 430070, China
2School of Mechanical Design Manufacture and Automation Xi'an Technological University, Xi'an, China
3School of Optoelectronic Engineering, Taiyuan University of Technology, Taiyuan, China

4324148@whut.edu.cn
†These authors contributed equally.

Abstract. As a combination of modern imaging technology, stereotactic technology, computer technology, and artificial intelligence technology, the surgical navigation system has been rapidly developed and applied in recent years. In recent decades, navigation systems have been used to aid in maxillofacial surgery, but rarely in the oral area. By comparing the applications of navigation systems in oral and maxillofacial surgery, we summarize their advantages and problems and list the advantages and disadvantages of the development process and methods of Simultaneous Localization and Mapping(SLAM). We propose to use tooth contours as natural markers to construct a model of AR oral structure, subtracting the operational time problems caused by making and using artificial markers. The AR structure model of the oral cavity was constructed, the method of applying SLAM technology to the oral surgery was proposed reasonably, and the method optimization scheme was proposed for the more complicated oral root canal surgery.

Keywords: oral cavity, navigation, Simultaneous Localization and Mapping.

1. Introduction

In recent years, the exploration of technical improvement methods for oral surgery has been increasing. Oral surgery, due to its small operating space and more concentrated neurovascular, makes its surgical operation more difficult and the patient's postoperative hidden danger higher, causing greater distress to the patient's postoperative life, which makes it urgently need relevant technical support assistance to improve the success rate of surgery and postoperative healing rate.

Among many auxiliary technologies, the surgical navigation system as a combination of modern imaging technology, stereo positioning technology, computer technology, and artificial intelligence technology has been rapidly developed and applied in recent years. Navigation system is being used by various surgical specialists, such as otolaryngologists, craniofacial surgeons, and plastic surgeons. As navigation technology has matured, oral and maxillofacial surgeons have become increasingly interested in this minimally invasive technological tool that can provide reduced trauma. In the current technological context, some surgical navigation systems are already available, but some of their specific

features were developed for other surgical specialties or are mainly aimed at maxillofacial surgery, and few are aimed only at the dental part of the oral cavity. And visual SLAM, as a hot development in recent years, provides a more advanced exploration direction for improved methods of oral surgery.

Visual SLAM has the function of localization and map building, and it provides a reference for the robot to localize itself and build a map by carrying the corresponding sensors and collecting the feature points of the relevant environment, which is like the navigation system, and we think it can play a bigger role in the field of oral modelling. To make the corresponding technology more widely used in the dental part of the oral cavity, in this paper, by analysing the related literature, we summarize and analyse the technical defects existing in the current oral surgery and the feasibility of the SLAM technology in solving them; secondly, we introduce different visual SLAM methods for different feature detection methods of SLAM, back-end optimization and whether they have closed-loop detection, and The advantages and disadvantages of the relevant methods are compared, and it is reasonable to propose the application of ORBSLAM2 to the dental part of the oral cavity; thirdly, the specific operation methods of SLAM application in the oral cavity are introduced, and its development prospects in the direction of dental root canal treatment are proposed.

2. Literature comparison

As shown in Table 1., Nathalie Pham Dang and others reported the 3D-2D shape matching method, which is mainly: the process of finding a spatial transformation to superimpose the same spatial position in a 2D image and 3D image exactly [1-5]. It is usually necessary to align a pre-know 3D image with a 2D image acquired in real time to achieve precise navigation in real time for surgery. However, it does not consider that if the vision is obscured by soft tissue or bleeding causes contamination of the field of view, the 2D image does not clearly show the advantages of the augmented reality part. And although Qingchuan Ma mentioned the autonomous surgical system with the assistance and monitoring of surgeons can enhance the visual ability during surgical procedures [6]. The clinical cost and the care of the robot will be an expensive expense if the robot is also responsible for the manual part, and the focus of this paper is on the dental part of the mouth, so it is difficult to determine the clinical feasibility of the robot performing the surgery; and the paper showed that the surgical results still. The paper focused on the dental portion of the mouth, so it is difficult to determine the clinical feasibility of robotic execution of the procedure, and the paper shows that surgical outcomes are still dependent on physician experience. Nagi Demian and others used markers to reduce the matching error of the navigation system [5][7][8]. But the process required to produce the markers would be another issue, and the markers would also be time-consuming to produce, increasing the total time spent on the procedure. Qingchuan Ma and others did not use markers but theirs result did not show the soft tissue structures of the face [6][9].A.D. Nijmeh and others illustrated the feasibility of augmented reality to improve the quality and speed of the procedure [10-12]. Endoscope Navigation and 3D Reconstruction of Oral Cavity by Visual SLAM with Mitigated Data Scarcity and 3D Reconstruction of Oral Cavity proposes that laser patterns can be used to help generate more identification points to reconstruct maps by ORBSLAM, which provides the basis for augmented reality (AR) modelling.

Table 1. Comparison of navigation systems in oral and maxillofacial surgery.

organs	Marker	Article work	Method	Reference
Oral and maxillofacial	No manual marking	AR has feasibility and clinical value in the oral cavity.	3D-2D (3Dimensional-2Dimensional)	[1]
Oral and maxillofacial	No manual marking	The feasibility and accuracy of the marker-free registration method for the AR system are proved.	3D-2D	[2]

Table 1. (continued).

Maxillofacial	No mention	CBCT allowed real-time imaging and surgical navigation	3D-2D	[3]
Maxillofacial, teeth	No manual marking	Used to guide the procedure; registration time is drastically reduced	3D-2D	[4]
Maxillofacial	Non-invasive fiducial markers for occlusal substrates	AR-guided surgery can increase the accuracy of surgery	3D-2D	[5]
Mandible	No manual marking	Artificial intelligence will be introduced to replace surgeons' operations and relieve doctors' pressure.	Use a robot instead	[6]
Maxillofacial	No mention	Real-time surgical navigation can make the surgical wound small, and improve the accuracy of surgery.	map, the tracker, and the transmitter	[7]
Oral and maxillofacial	Titanium screw marker	The BeiDou-SNS navigation system was used to assist the operation.	The probe for determining the visual positional relationship	[8]
Maxillofacial	No manual marking	Develop a simple navigation system	Put in an electromagnetic sensor to track the pencil	[9]
Oral and maxillofacial	Summary class	It has demonstrated the feasibility of AR on oral implants.	Summary class	[10]
Oral implant	No mention	Two registration methods are provided to verify the feasibility of the navigation system.	Polaris Vicra optical tracking device	[11]

The authors learned from their study that ORBSLAM2 can synchronize localization and augmented reality in real-time and can change the field of view as needed, solving the above soft tissue occlusion, marker, and detection problems while retaining the advantages of augmented reality and reduced radiation exposure in the above navigation system. Therefore, the authors reasonably believe that ORBSLAM2 has great potential for application in dentistry.

3. The development of SLAM

The visual SLAM workflow can be divided into five parts: information collection, front-end visual odometer, back-end optimization, closed-loop detection, and mapping. The front-end vision odometer extracts the feature points in the image data to obtain the marker points and calculates the current pose of the camera to achieve real-time positioning. The back-end optimization mainly eliminates and optimizes the error of the output results of the previous section, to improve the accuracy of mapping. In pose estimation, when the similarity of the two frames of images reaches a certain standard, then it is considered that a closed loop is formed, and all the marker points of the two frames of the image are optimized once, and the mark points after optimization are used for mapping, as shown in Figure 1 [13].
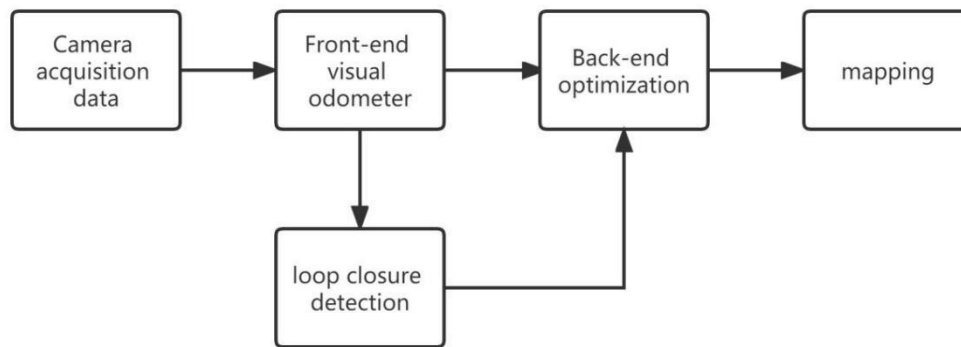

Figure 1. Flow chart of Visual SLAM.

The sensors commonly used in classical vision SLAM are mainly monocular, binocular, and RGB-D cameras [13]. Monocular vision SLAM is based on a monocular camera, and can complete the calculation and mapping of camera pose between continuous images under the condition of scale initialization, which has the advantages of low cost and simple equipment. The disadvantage is that there is a scale drift phenomenon, and the accuracy of mapping is not high. Therefore, binocular vision SLAM and RGB-D SLAM were proposed to solve the problem of not being able to obtain image depth. Binocular vision SLAM uses a binocular camera to collect information, and the binocular camera can calculate the pixel depth through the principle of stereo vision. The advantage is high adaptability to the environment; The disadvantage is that the pixel depth calculation is large (Table 2).

3.1. Mono SLAM—based on extended kalman filtering

In 2007, Andrew J. Davison et al. proposed Mono SLAM (Monocular SLAM), which uses an active-guided method to map and measure features and then extracts sparse feature points for pose estimation, using a general-purpose 3D camera to capture the inherent dynamic information in the continuous video [14]. This method is the first real-time vision SLAM system to be completed using a monocular camera, which uses the "standard" single, fully covariance EKF (Extended Kalman Filter) method for SLAM, rather than using variants with different probability representations.

Mono SLAM uses only a single thread, feature point extraction, and matching, camera poses estimation, and mapping work is all frame-by-frame, and the updated computational complexity is very high, to improve SLAM speed, only about 10 feature points can be processed in each frame of the image, and feature points are easily lost [13].

3.2. PTAM—based on keyframes

PTAM (Parallel Tracking and Mapping) is a camera tracking system for augmented reality. It requires no markers, pre-made maps, known templates, or inertial sensors. Georg Klein and David Murray proposed PTAM in 2007 to solve the problem of the complexity of Mono SLAM calculations [15]. PATM optimizes a part of the visual odometer to select the most representative frame as the keyframe

in a part of the image, reducing the complexity of the calculation.

3.3. ORB-SLAM-based keyframe

ORB-SLAM (Oriented FAST and Rotated BRIEF SLAM) is a three-threaded joint solution by improving the PTAM method and adding a new part of closed-loop detection, in which the three threads are: Tracking, Local Mapping, and Loop Closing. Loop Closing is responsible for the loopback detection and optimization of the global Pose Graph. ORB-SLAM, optimized for the previous monocular vision SLAM, has a fast initialization in binocular and RGB-D conditions and does not require movement to complete its initialization. ORB-SLAM improves the accuracy of localization and map building due to the closed-loop detection part. The selection of keyframes is first chosen to be more lenient to improve the robustness of the tracking and positioning process, and then the keyframes that do not meet the expectations are removed to improve the efficiency of BA optimization and the accuracy of closed-loop detection [13]. In the related experimental report literature, it is shown that the root mean square error of the key frame trajectory of ORB-SLAM2 is only 1.57 cm, while moving the camera rapidly does not affect it to maintain a better tracking state, PTAM, and LSD-SLAM are much lower than ORB-SLAM in terms of method robustness, but ORBSLAM is easy to lose the tracking point in dynamic environment extraction, leading to tracking failure, and because the feature point extraction of ORB-SLAM is relatively sparse, its causes degradation of map accuracy [16].

3.4. DSO

In 2016 Jakob Engel proposed DSO (Direct Sparse Odometry), DSO is a visual odometry method that is a sparse direct method [17]. Its computational process does not require finding a feature point and matching it with the feature points in other frames as in traditional SLAM algorithms, DSO is to project all points into each frame, calculate the residuals between each point and fit the residuals to each other by computation, and ignore the correspondence between these points, because its computation is extremely fast, so the superposition of computation will generate a dense DSO is superior to the above two methods in terms of robustness, computational speed, and accuracy, but it cannot calibrate the detected image, environment due to the lack of closed-loop detection, and it will be difficult to reposition once the tracking of the environment is lost.

Table 2. Comparison of advantages and disadvantages of visual slam methods

Method	Advantage	Disadvantage	References
Mono SLAM	Good real-time performance	High computational complexity, Feature points lost	[13]
PTAM	Reduce computational complexity, Higher accuracy	The process of closed-loop detection cannot be optimized	[15]
DTAM (dense tracking and mapping)	Reduce calculation time, Avoid missing feature points	Sensitive to light changes	[13]
LSD-SLAM (large-scale direct monocular SLAM)	Faster computation	Easy to miss the target, No closed-loop detection	[13]

Table 2. (continued).

DSO	Calculation speed reduced	Without back-end optimization and closed-loop detection	[17]
ORBSLAM2	High computational efficiency, good relocation, and closed-loop detection, Wide range of capability	Low accuracy of sparse points, Easy to lose the target point in the dynamic environment	[16]

4. The specific operation method of SLAM application in the oral cavity
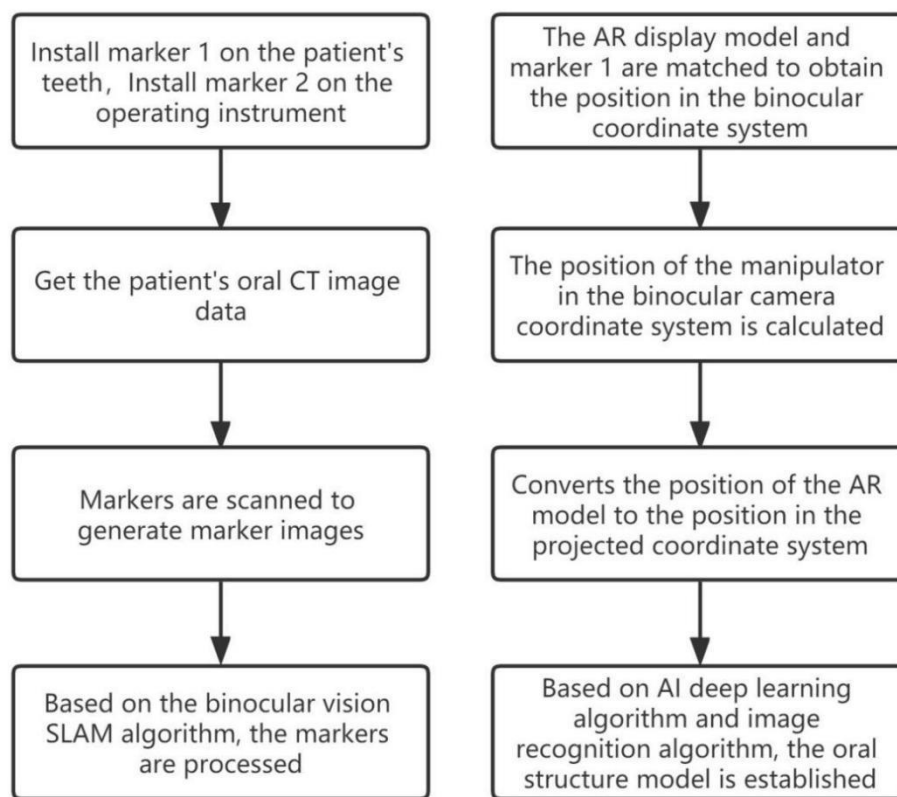

Figure 2. The oral structure model was constructed by the marker method.

4.1. Marker method

The method of realizing oral navigation based on mixed reality technology can well solve the problem that surgery cannot be performed intuitively due to the small oral space [18], and the specific operation flow chart of this method is shown in the Figure 2.

In this method, the oral structure model was constructed by installing markers. But making the markers takes time, increasing the overall surgical time. And the use of artificial markers requires a complex process to make markers, which also adds to the cost of surgery. The oral and dental problems we're talking about are common problems for the general public, which is not appropriate.

4.2. Label-free method

Based on the problems with the 4.1 methods, we propose to use the tooth contour as a natural marker

point to reduce the surgical time caused by the production and use of artificial markers. Moreover, this method is in line with the front-end visual odometer part of S LAM, which meets all the advantages of SLAM application in oral surgery. The no-manual labelling method also means that only preoperative testing is required, which reduces radiation to the patient. According to Junchen Wang and others purposed [4][6][11],we found that these navigation systems without artificial markers are also highly accurate, meeting clinical needs and reducing the cost of reoperation for patients. Therefore, it is reasonable to think that the tooth contour should be used as a natural marker as shown in Figure 3, and then the oral structure model should be constructed by imaging technology. This method reduces the cost of surgery and shortens the operation time, so it can be applied in surgery.
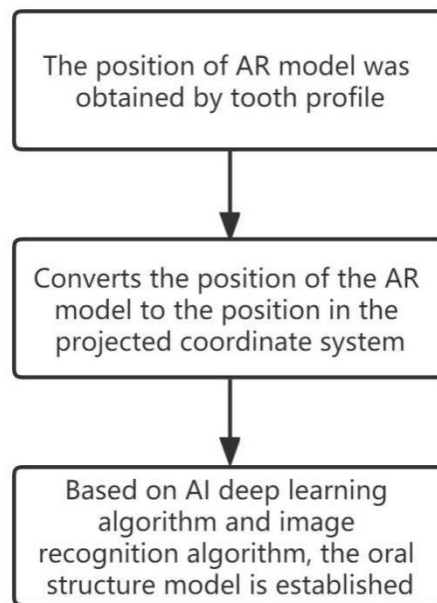

Figure 3. The oral structure model was constructed by a marker-free method.

5. Conclusion

In this article, we summarize and compare the navigation systems used in oral surgery, and by comparing the advantages of the development history of SLAM and its branches, we summarize the advantages of existing navigation systems, and they correspond to the advantages of SLAM application in oral surgery; By comparing and summarizing the advantages of markers and navigation systems in surgery, we reasonably propose the use of non-manual labelling methods to assist doctors in surgery, and the use of technical concepts is carried out around ORBSLAM2. We discovered by learning part of oral surgery: root canal treatment, also known as endodontic treatment, is a type of surgery in dentistry to treat pulp necrosis and root infection. In recent years, root canal treatment has been used by most doctors, but the root canal system is very complex, and no instrument can reach the entire root canal system, so root canal treatment still fails because it cannot see directly into the internal structure of the patient's teeth. Therefore, we propose an idea to use SLAM in oral root canal surgery, use a micro-SLAM robot to synchronously build an oral AR model, and then identify the AR model based on an AI deep learning algorithm. This allows the doctor to see the covering structures in the mouth, such as teeth, superior molar nerves, and inferior molar nerves. The technology proposed is largely superior to the traditional dental operating microscope (DOM), and the miniature SLAM robot can construct the internal structure of the dental root canal system, combining the realism of AR eyepieces with the synchronicity of composition.

References

[1]   Dang N P, Chandelon K, Barthélémy I, et al. A proof-of-concept augmented reality system in oral

and maxillofacial surgery[J]. Journal of Stomatology, Oral and Maxillofacial Surgery, 2021, 122(4): 338-342.

[2] Shi J, Liu S, Zhu Z, et al. Augmented reality for oral and maxillofacial surgery: The feasibility of a marker-free registration method[J]. The International Journal of Medical Robotics and Computer-Assisted Surgery, 2022, 18(4): e2401.

[3] Goguet Q, Lee S H, Longis J, et al. Intraoperative imaging and navigation with mobile cone beam CT in maxillofacial surgery[J]. Oral and Maxillofacial Surgery, 2019, 23: 487-491.

[4] Wang J, Suenaga H, Yang L, et al. Video see-through augmented reality for oral and maxillofacial surgery[J]. The international journal of medical robotics and computer-assisted surgery, 2017, 13(2): e1754.

[5] Zhu M, Liu F, Chai G, et al. A novel augmented reality system for displaying inferior alveolar nerve bundles in maxillofacial surgery[J]. Scientific Reports, 2017, 7(1): 42365.

[6] Ma Q, Kobayashi E, Wang J, et al. Development and preliminary evaluation of an autonomous surgical system for oral and maxillofacial surgery[J]. The International Journal of Medical Robotics and Computer-Assisted Surgery, 2019, 15(4): e199

[7] Demian N, Pearl C, Woernley T C, et al. Surgical navigation for oral and maxillofacial surgery[J]. Oral and Maxillofacial Surgery Clinics, 2019, 31(4): 531-538.

[8] Qin C, Cao Z, Fan S, et al. An oral and maxillofacial navigation system for implant placement with automatic identification of fiducial points[J]. International Journal of computer-assisted radiology and Surgery, 2019, 14: 281-289.

[9] Bouchard C, Magill J C, Nikonovskiy V, et al. Osteomark: a surgical navigation system for oral and maxillofacial surgery[J]. International Journal of Oral and maxillofacial surgery, 2012, 41(2): 265-270.

[10] Nijmeh A D, Goodger N M, Hawkes D, et al. Image guided navigation in oral and maxillofacial surgery[J]. British Journal of Oral and Maxillofacial Surgery, 2005, 43(4): 294-302.

[11] Chen X, Xu L, Wang H, et al. Development of a surgical navigation system based on 3D Slicer for intraoperative implant placement surgery[J]. Medical Engineering & Physics, 2017, 41: 81-89.

[12] Qiu L, Ren H. Endoscope navigation and 3D reconstruction of oral cavity by visual SLAM with mitigated data scarcity[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2018: 2197-2204.

[13] Wang Peng, Hao Weilong, Ni Cui, Zhang Guangyuan, Gong Hui. Review of visual SLAM methods [J]. Journal of Beijing University of Aeronautics and Astronautics:1-9.

[14] Davison A J, Reid I D, Molton N D, et al. Mono SLAM: Real-time single camera SLAM[J]. IEEE Transactions on pattern analysis and machine intelligence, 2007, 29(6): 1052-1067.

[15] Klein G, Murray D. Parallel tracking and mapping for small AR workspaces[C]//2007 6th IEEE and ACM international symposium on mixed and augmented reality. IEEE, 2007: 225-234.

[16] Şımşek B, Satir S, Bılge H Ş. Performance comparison of direct and feature-based vSLAM algorithms[C]//2021 29th Signal Processing and Communications Applications Conference (SIU). IEEE, 2021: 1-4.

[17] Engel J, Koltun V, Cremers D. Direct sparse odometry[J]. IEEE Transactions on pattern analysis and machine intelligence, 2017, 40(3): 611-625.

[18] Wang Jianhua, Lin Qiong, Zhu Ming. A method and system for oral implant-assisted navigation based on mixed Reality Technology: CN114587657A[P]. 2022.