Lupine Online Journal of Sciences

Review Article

Design and application of touchless interaction system in medical imaging for thoracoscopic surgery

Yuzhuo Zhang[1], **Zhiyuan Xi**[2], **Jun Liu**[3]* **and Fei Cui**[3]*

[1] *Nanshan School, Guangzhou Medical University, China*

[2] *China-Japan Union Hospital of Jilin University, China*

[3] *Department of Thoracic Surgery, the First Affiliated Hospital of Guangzhou Medical University, China*

***Corresponding author:** Fei Cui, Department of Thoracic Surgery, the First Affiliated Hospital of Guangzhou Medical University, China

Jun Liu, Department of Thoracic Surgery, the First Affiliated Hospital of Guangzhou Medical University, China

Received: 📅 June 12, 2023 **Published:** 📅 **June 23, 2023**

Abstract

The primary difficulty in performing precision thoracoscopic surgery lies in the low resolution of fine anatomical structures in a local and deformed setting, lacking a global contextual mapping to pre-operative medical imaging-based models. A gesture-recognition based touchless interaction system is designed to allow surgeons to manipulate the 3D models of patients in an intuitive and timely manner. It greatly enhances the surgeons' cognitive efficiency in understanding anatomy under operations. This system is implemented with the Microsoft Kinect sensor and tested in a real video-assisted thoracoscopic surgery environment.

Keywords: Gesture-recognition; thoracoscopic surgery; cognitive efficiency; kinect sensor

Introduction

Rapid technological progress has made thoracic surgery more minimally invasive and postoperative rehabilitation faster. With the wide spread of early screening by using low-dose spiral CT and automatic detection means, such as deep learning, an increasing number of small lung nodules are able to be detected. Meanwhile, the surgical treatment of lung tumor has also developed towards less aggressive procedures, evolving from pneumonectomy, lobectomy, and segmentectomy to wedge resection. These considerably increase the challenges of a thoracic surgeon in the recognition, tracing, and localization of the lesion buried in the labyrinth of the lung arteriovenous and bronchial tree structures. The deformation of the lung parenchyma due to deflation and the loss of tangible force feedback further add to the difficulties of thoracoscopic surgery. The pre-operative plan, using 3D reconstruction technology based on CT images, is shown to be useful for robot-assisted seg mentectomy, and helps reduce complications and improve surgical efficiency [1].

By using a deep neural network, the precise 3D reconstruction of lungs based on automatic segmentation can achieve intuitive and accurate visualization and identify and analyze the structure and spatial distribution of lung tissues, lesions, pulmonary blood vessels, and bronchial tubes among others [2]. 3D reconstruction based on pre-operative lung surgery planning has been reported to be used on a regular basis, aiding as many as 20~30 operations per month within one department [3]. Despite the value of 3D reconstruction based pre-operative planning, cognition gaps caused by the dissociation of local anatomical structures and global 3D reconstruction model is the main challenge solved to achieve full practicality in a realistic surgery setting. During typical thoracoscopic surgery, operators rely on pre-operatively planned anatomical markers as spatial references to gradually incise the lung parenchyma until removal of lesions. Those anatomical markers commonly consist of pulmonary vascular, bronchial structures and lung fissures, which are recognized with the help of medical imaging-based models.

However, as the operation progresses, the surgeon easily loses the global contextual information of local anatomy under operation in the middle of complex anatomical structures and is obliged to refer to the 3D model of patient regularly. 3D models should be oriented, zoomed and positioned according to the will of the surgeon in a precise and timely manner to enhance the cognition efficiency of surgeon. Classical pointing devices, such as mouse or keyboard, are beyond consideration as they are unsterile and pose a risk of infection. Surgeon can also give instructions to an assistant, who then can operate a mouse or keyboard, but surgeon's concentration is interrupted each time saying each instruction, which furthermore can hardly represent his/her precise intention and increase misconception risks. A few innovative alternatives are being explored.

In this work, the Windows Kinect sensor's capability is exploited in people tracking and gesture recognition and interfaced it to our own medical imaging system to allow touchless control of patient 3D models with gestures. Our contribution is twofold: First of all, in order to design a domain adapted touchless control system, the complete life cycle of a typical thoracoscopic operation is examined, including operating rooms arrangement, the natural body posture of a surgeon in regard of gesture-based model control, interferences that crew members might cause, as well as assistances they might properly bring about, etc. Secondly, we concretized the more general idea of cognitive operating rooms in well-defined use cases. Future operating rooms will be characterized by their cognitive and situation aware abilities acting as human assistants rather than a tool set. The short-term impacts of such conceptual operations are expected from improvement of man-machine-interfaces based on speech or gesture recognition-based systems [4]. This work provides a concrete testing fields and innovation base of the concept.

Design of gesture-based touchless interaction system

Gesture tracking sensor

Microsoft Kinect v2 sensors are used for motion capture. With TOF (Time of flight) and machine learning techniques, the device constructed the depth map of scene and inferred different body parts of an individual fully contained in frame. It was able to track up to six peoples' skeletons simultaneously and each skeleton consists of 25 joints: shoulder, elbow, wrist, hand, hand tip, and thumb for both arms specifically. Detected joints were represented by their (x, y, z) coordinates in the 3D space, which were used to define different gestures.

Gesture design principles

A successful touchless interaction system relies on several key points, including natural gesture definition, robust gesture recognition and intuitive function mapping. We followed the general principles and common practices in the design of gesture recognition system as outlined in Human Interface Guidelines v2.0, KinectHIG.2.0 and enhanced it with the specific requirements of operating theaters. According to operating room sterilization principles, the sterile area of a properly gowned and gloved healthcare worker extends from chest to the level of patient's surgical site; sleeves are sterile from 5cm above elbows to cuffs [5]. This translates roughly to a rectangular surface in front of body, spanning from upper chest to waist from shoulder to shoulder, constrained in which operator forms gestural command. As a result, naturally lifted forearms accompanied with different hand gestures/movements forms candidate gesture pool for clinical use cases.

The gesture design principles are listed as follows:

a) Simple, unambiguous, free of false alarm way for people to engage and disengage.

b) Reliable and stable gesture recognition: robust to interference, robust to jitter, stabilization of controlled model during hand position reset, unambiguous gesture boundary, and easy recover from error.

c) Effortless gesture realization and fluent gesture transition: gestures should be naturally achievable, and the adjacent gestures should be fluently transmittable especially for highly repeating situations.

According to the above items, a gesture table is defined as well as corresponding function mapping in Table 1. Figure 1 illustrates machine state design for the gesture recognition system. The circles colored brown are two major control states which indicate the deactivation or activation of gesture commands respectively. The triggering gestures are designed according to Principle1, such that similar gestures are rarely met in a casual or unintentional manner. Moreover, to lift one's hand over head reaches across the unsterile field, which never happens in a procedure compliant with the sterilization protocol, thus avoiding any interference with model manipulation gesture. The command is to turn on/off the recognition is clear and can easily be done by any appropriate crew member under oral instructions. Continuous gestures for model zoom, pan and rotation send corresponding transformation commands in the same rhythm as the refresh rate of the Kinect event monitor thread. No state transitions are required for these gestures.

Table 1: Gesture definition and function mapping (The "Description" column lists the major design principles which the corresponding gesture adheres to).

SI No	Left hand	Right hand	Function	Description
1	N/A	Palm over head	Gesture engagement	P1
2	N/A	Fist over head	Gesture disengagement	P1
3	N/A	Below elbow	Freeze	P2: robust to interference
4	Below elbow	N/A	Freeze	P2: robust to interference
5	Fist/above elbow	N/A	Freeze	P2: stabilization during hand position reset
6	Palm/above elbow	Fist/above elbow	Rotation	P2: robust to interference
7	Palm/above elbow	Palm/above elbow	Zoom	P2: robust to interference
8	Palm/above elbow	Finger/above elbow	Pan	P2: robust to interference
9	Finger/above elbow	Forearm swing (vertical to horizontal)	Switch active organ	P2, P3
10	Finger/above elbow	Close followed by Open	Switch active organ on/off	P2, P3

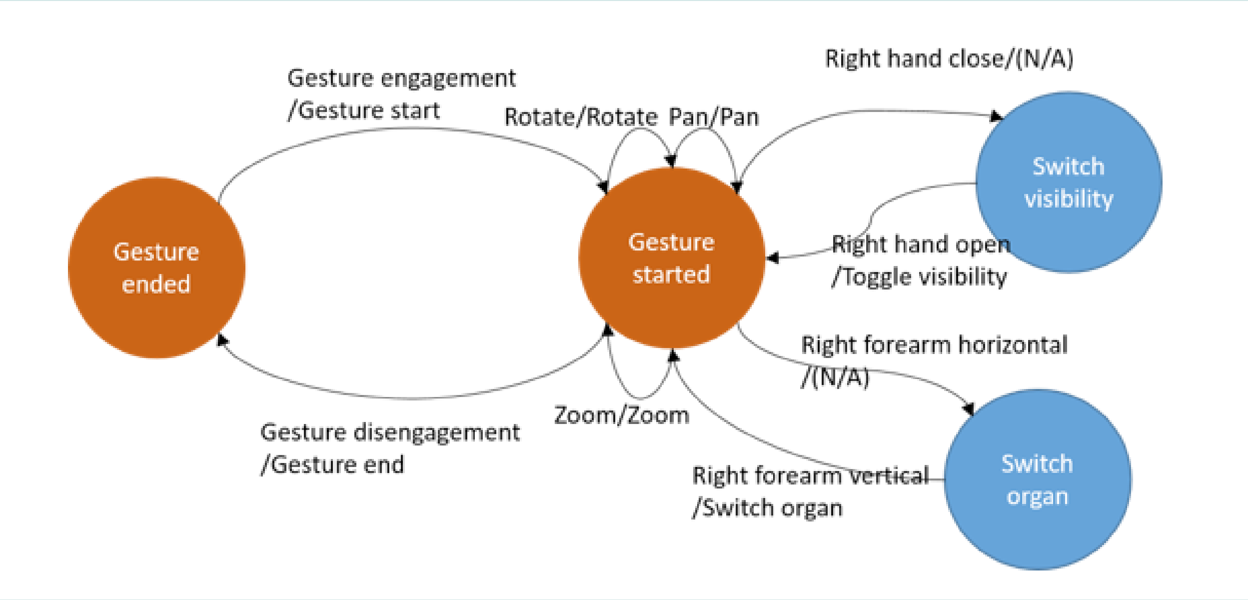

Figure 1: Mealy state machine for gesture design. Circle: current state; annotation: gesture/command pair.

On the other hand, for discrete functions, such as model switching and visibility toggling, we need to introduce an assistant state to avoid repeated reaction to a single trailing gesture. In order to minimize interference from other crew members, only natural but intentional gestures are recognized as valid. Casual swings of hands where both hands are not above the elbow levels are considered invalid. Accumulating gestures like rotation (as defined in Line 6), require continuous movement of the hand along one direction until arriving at the desired model orientation during which one might need to temporarily reset the hand to a comfortable starting point before resuming. To avoid negating the already accumulated movements, a freezing mechanism should be provided that stabilizes the model during hand reset. In our design, whenever a closed left hand is above elbow level, the model stops performing any transformation.

A precise one to one map of the model movement to the operator's motion is crucial for pleasing user experience. Filtering techniques were employed to cancel joint jitters which causes jumping or wandering effect on the model. The shoulder joint is used as reference point for model pan which moves around slightly with the movement of the forearm. Accumulated average filter is used for anchor position filtering which locks the position of the anchor point near the shoulder. It solves the "joint jumping" effect as well if somebody walks in front and causes an occlusion. The detected hand joint moves slightly around even if the hand is holding still which causes model jitter effect. A double exponential filter is used to suppress the motion within a radius. The filter dampens as well the movements outside of the primary motion direction [6].

Citation: Yuzhuo Zhang, Zhiyuan Xi, Jun Liu* and Fei Cui*. Design and application of touchless interaction system in medical imaging for thoracoscopic surgery. LOJ Sci 1(2)- 2023. LOJS.MS.ID.000109.

Clinical applications of the gesture-based touchless interaction system

The gesture-based touchless interaction system has been deployed in our center for 6 months and has demonstrated good efficiency in assisting complicated thoracoscopic surgeries. As we all know, further progresses in thoracoscopic surgery are linked with improvements in medical instruments and accurate preoperative diagnostics [7]. Besides, intraoperative image interaction systems are critical. Lately, we retrospectively compared the roles of gesture-based touchless and traditional thoracoscopic surgeries. We find that with all the advantages of our system, it avoids side effects of traditional ones. Before the deployment of gesture control system, pulmonary 3D reconstruction models have already been made use of in operating rooms for real time surgery assistance. 3D models were rendered on a 2D medical class screen, or a stereoscopic screen placed 1.8 to 2.5 meters away from the surgeon across the surgical table. Different tissues were by default selected and rendered, namely the lung parenchyma, pulmonary vessel, pulmonary artery, bronchus, and tumors. Surgeons referred to 3D models regularly for locating anatomical landscapes.

Dedicated technical assistants were needed for manipulation of 3D models under surgeons' instructions. According to our experience, assistants received around 10 instructions during surgery. Some of instructions were definite and unmistakable, such as hide the vessel and display only the bronchus and the nodule, or any other combinations of tissues. While others were less clear and often required multiple intermediate steps, for example, rotating the model towards the nodule. For unclear instructions, an average of 3 instructions/tries were needed for the model to arrive at the desired orientation, which was energy consuming and cognitively inefficient. Besides, some specific tips and tricks, which are time-consuming, need to be taken into consideration to ensure surgeon's comfort and patient's safety. Otherwise, during these procedures, there may be a crash between surgeon and assistant, and assistant's shoulder might get injured. In addition, some born aspects of VATS surgery could limit surgeons' ability to deal with local factors safely which lead to intraoperative accidents. Losing data is an obvious disadvantage of 3D modeling, because in most cases these applications focus on the visualization of specifically selected and requested anatomic structures.

Consequently, information on specific structure (for example, lymph nodes, thoracic cage, the heart, diaphragm, and other intrathoracic vessels) is lost. Likewise, as 3D segmentation is manually, not only does it increase the risk for error, but it is also time-consuming. Therefore, 3D modeling increases health risks for patients. Furthermore, the intra-operative identification and surgical treatment of small pulmonary lesions are still challenging. Due to the need to repeatedly check the anatomical structure, the operation process is often interrupted. Longer operative time was the most reported disadvantage for video-assisted thoracoscopic surgery [8], which increasing the risk of likelihood of inattention. It was irreversible and unacceptable for patients, once surgeons got wrong information served by assistants and cut off blood vessels or bronchus mistakenly. The gesture-based touchless interaction system can ease these troubles mentioned above as it has the following advantages: The gesture control system shows the 3D structures of the chest cavity, which makes thoracoscopic surgery more accurate, efficient and timesaving.

Traditional thoracoscopic approaches rely on CT images to locate the lesion. With the advent of our gesture control system, the situation is largely improved, as it can assist localizing lesion before surgery. And surgeons can use this tool to view and interact with anatomical locations of interest. Moreover, this is very efficient and timesaving, because the operation will not be interrupted by frequently checking the anatomical structure and communicating with assistants. Touchless interface with gesture recognition is developed to enhance operating room control system [9]. A touchless interaction system is an ideal solution in the operating room, as it requires nonphysical contact, decreases the risk of contamination, and can still provide a cleansed and sterilized environment. Since our system can provide a low-cost and accurate control by just using hand gestures, the surgeon is able to manipulate the model in his will and event achieve fuzzy controls which are impossible in the past, for example to slightly oscillate the model around a fixed pivot. 3D is correlated with a much stronger sense of depth perception, spatial location of the target and adjacent anatomic structures, leading to higher precision of surgical performance.

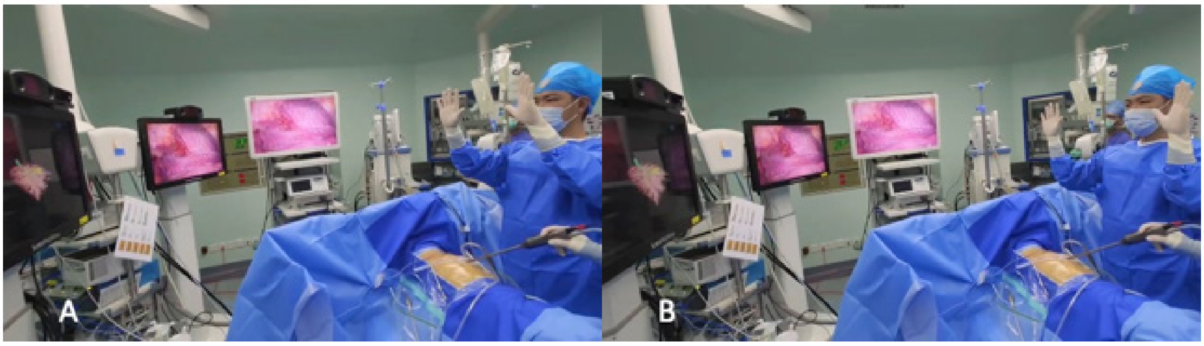

Figure 2: The surgeon scales the 3D model during surgery. Before scaling (A), after scaling (B).

Citation: Yuzhuo Zhang, Zhiyuan Xi, Jun Liu* and Fei Cui*. Design and application of touchless interaction system in medical imaging for thoracoscopic surgery. LOJ Sci 1(2)- 2023. LOJS.MS.ID.000109.

Thus, with our system, accuracy has been improved. It can facilitate the surgeons performing their tasks quicker and more efficiently, compared to the use of an assistant. The accuracy of hand gesture recognition with Microsoft Kinect v2 sensor is better, which could be used to navigate and manipulate various types of computer-aided devices and applications through touchless gesture interaction. The emergence of our touchless interaction system contributes to improvements of precise lesion resection, and it will become important adjuncts to thoracoscopic surgery. With this touchless interaction system, the surgeon can remove the lesions more willingly. And patients are more satisfied, as smaller incisions and less pain lead to shorter hospital stays, earlier functional recovery. On June 5th, for the first time, we successfully performed thoracoscopic surgery with gesture control system (Figure 2). Since then, we've completed more than 6 cases. The surgeries were smooth in all these cases, without death or severe complications. Consequently, in those cases, the outcomes are better than the ones of traditional thoracoscopic surgeries, and the quality of life will be improved.

Additionally, our system is easy and ready to use. Before using, what needs to be finished is just a single demonstration and a tutorial. The system can be installed on each computer with sufficient graphics and processing capacity. It is foreseeable that, with the availability of gesture control system, the increase in the anesthetic and surgical techniques, and the acceptance of the enhanced recovery after surgery, the thoracoscopic surgery will be much more accurate and efficient. The gesture control system has a variety of gesture control modes. Model zooming and model rotation controls are both among the most frequently used gestures combinations. Thus, a more friendly operating system is developed, which is suitable to use. The ability to switch between them influences user experience smoothly and unambiguously to a great extent. In our design, the two gestures switch by simply closing or opening the right hand, which proves to be very intuitive. Firstly, our system provides a 3D construction of the nodule location relative to the chest anatomy, which could be zoomed, rotated etc. Secondly, it gives the surgeon access to reset and maneuver the model. Thirdly, this dynamic interaction can help surgeons plan the surgical approach.

Fourth, to eliminate false model manipulation caused by interference from the crew members or by casual or unintentional gestures of the surgeon recognized as a valid command by the sensor, both hands are obliged to be lifted above the elbow level for a valid gesture control. This design proves to be effective in noise reduction while keeping the gestures as effortless as possible, which helps the surgeon to perform much more efficiently. As surgeons are operating with both hands, hand gestures are used to control the zoom and rotation of the lens, so controls of lens are on surgeons' own will without worrying about contamination. The surgeons can orient, zoom, and position the 3D model whenever they want with low latency. Besides, they can view the chest from a certain angle which gives an exceptional vision to the structure of thoracic cavity and improves the understanding of patient anatomy. Thus, surgeons gain better insight into patients' anatomy. And active organ can be switched by simply swinging their forearm, which allows surgeons to focus more on the parts that need surgery. By zooming in on the lesion, surgeons could resect small lung nodules that used to be impalpable.

Limitations of gesture-based touchless interaction system

At present, the surgeon-assistant combination is still the mainstream in thoracoscopic surgery. Thus, our system is urgently needed. Despite its promising prospects, there are some potential disadvantages to using this system. It is believed that, with the progress of engineering technology, the gesture control system will become more stable. We acknowledge that the current study has some limitations. First, the main observations in designing the gesture recognition system are the limited gesture choices. Put aside the constraints imposed by the specialties of operating room environment, one reason for the limitation comes from the fact that calculations are based on the joint coordinates calculated by the sensor instead of using the original depth image. One of the future works will be training neural networks basing directly on the combination of the depth image and the RGB image, to recognize richer and finer gestures. Second, surgeons must overcome learning difficulties before they can use hand gestures efficiently. Third, only a small number of patients were included, and only several thoracoscopic surgeons have used and evaluated the system. In order to draw a clear conclusion from the application of this system, more prospective, extensive, and comparative analyses are required.

Conclusion

Though gesture choices are limited due to the limitations of the operating room, this system benefits both patients and surgeons significantly in reducing costs, operating time, complications, and the need for additional procedures. And it allows other clinicians to view the exact anatomical part simultaneously in the operating room, which could benefit greatly in training residents and students. Importantly, it will enhance the surgeons' ability to locate the lesion and finalize the approach to allow surgeons to concentrate during the operation without worrying about getting interrupted. In conclusion, it has the potential to be widely applied across thoracoscopic surgery, especially minimally invasive surgery. Our goal is to contribute to the development of thoracoscopic surgeries for educational and training purposes. With the rapid development of computer processing technology, the thoracoscopic surgeries are making great progress, which promotes the advances in minimally invasive surgery. The gesture control system has been proven to be a promising new technology for thoracoscopic surgery. Although there're some shortages, its greatest advantage is to facilitate surgeons complete the operation without being disturbed. It will dramatically promote the development and application of minimally invasive endoscopic surgery, with great economic and social benefits.

Funding

This study was supported by the Guangzhou Science and Technology Bureau 2022 Basic and Applied Basic Research Project "Visualized Surgical Planning and Navigation System of Glasses-free

Three-Dimensional (3D) for Thoracic Surgery" (grant number: 20220102465).

Conflicts of Interest

The authors have no conflicts of interest to declare.

References

1. Le Moal J, Peillon c, Dacher JN, Baste JM (2018) Three-dimensional computed tomography reconstruction for operative planning in robotic segmentectomy: a pilot study. J Thorac Dis 10(1): 196-201.
2. Zhi Xiuyi (2019) Expert consensus on three-dimensional visualization positioning and surgical planning of pulmonary nodules under the artificial intelligence platform. Chinese Journal of Thoracic and Cardiovascular Surgery 26(12): 1161-1166.
3. Tian G, Zhao Z, Kurtaev D, Liu Y, Zhu L, et al. (2020) Hisense Medical Computer-assisted Surgery (CAS) System for Pulmonary Surgery based on Intel R Architecture. Hisense Medical pp. 1-8.
4. Maier-Hein L, Eisenmann M, Feldmann C, Feussner H, Forestier G, et al. (2018) Surgical Data Science: A Consensus Perspective 1: 1-29.
5. Aseptic Technique & The Sterile Filed.
6. https://channel9.msdn.com/series/Programming-Kinect-for-Windows-v2/07
7. Mikheev AV, SN Trushin, Ogoreltsev AY (2021) The evolution of operative access in lung surgery. Kardiochir Torakochirurgia Pol 18(4): 247-251.
8. Kiblawi R, Zoeller C, Zanini A, Ure BM, Dingemann J, et al. (2021) Video-Assisted Thoracoscopic or Conventional Thoracic Surgery in Infants and Children: Current Evidence. Eur J Pediatr Surg 31(1): 54-64.
9. Cho Y, Lee A, Park J, Ko B, Kim N (2018) Enhancement of gesture recognition for contactless interface using a personalized classifier in the operating room. Comput Methods Programs Biomed 161: 39-44.

Citation: Yuzhuo Zhang, Zhiyuan Xi, Jun Liu* and Fei Cui*. Design and application of touchless interaction system in medical imaging for thoracoscopic surgery. LOJ Sci 1(2)- 2023. LOJS.MS.ID.000109.