

A Method of Acquiring Surgical Instrument Trajectories in Surgery

Based on Micro Checkerboard Markers



Ruofan Bian^{1,a}, Bin Chen^{2,b}, Fa Wu^{3,c}, Ningzi Zhang^{3,d}, Dexing Kong^{1,*}

¹School of Mathematical Sciences, Zhejiang University, Hangzhou, China

²The First Affiliated Hospital, Zhejiang University School Of Medicine, Hangzhou, China

³Demetics Medical Technology, Zhejiang University, Hangzhou, China

^abradbrf@zju.edu.cn

^bchenbin72@163.com

^cmanwufa@zju.edu.cn

^dzhangningzi@demetics-medical.com

*Correspondence: dxkong@zju.edu.cn

Abstract

This paper proposes a binocular vision positioning method using micro checkerboard markers for identifying the posture and position of surgical instruments during surgery. Compared to existing methods that use optical spheres for instrument localization, the approach based on micro checkerboard markers requires minimal modification to conventional surgical instruments. It is applicable to instruments such as scalpels, needle holders, and hemostatic forceps, which are small in size, have high similarity, and are easily obscured during surgery. Even when the surface of target surgical instrument is mostly obscured, this method can precisely identify and locate the instrument based on only 4 marker points.

Keywords: micro checkerboard markers, conventional surgical instruments, trajectory acquisition, occlusion issues.

1. Introduction

Usually, hospitals use text to record patient information, surgical procedures, surgical results and other information as surgical records to preserve operative data. Such records lack a description of the surgical techniques employed by the doctors, making it challenging for computers to quantitatively analyze the surgical process. Addressing the limitation of existing surgical navigation systems in capturing all surgical motion trajectories in routine surgeries, this paper introduces a binocular vision positioning

method based on miniature checkerboard markers. This method digitizes the surgical process, providing data support for establishing surgical databases, computer-assisted surgeries, and robotic surgeries. Traditional surgical navigation systems are designed for specialized procedures like tumor ablation surgeries and spinal pedicle screw fixation surgeries, tracking and positioning large specialized instruments to guide doctors or robotic surgeries. However, positioning studies for commonly used small, conventional surgical instruments such as scalpels, hemostatic forceps, and needle holders are yet to yield mature products.

There are three major challenges in conventional surgical instrument positioning:

1. High difficulty in modification: Altering the appearance of surgical instruments affects the doctor's regular surgical techniques, resulting in the inability to obtain normal surgical trajectories.
2. High similarity among instruments: Conventional cutting instruments have nearly identical shapes, except for their tips. During use, when the tips are obscured as they enter the lesion area, distinguishing between different surgical instruments becomes challenging.
3. Multiple instruments' trajectories: Inevitable intersections and obstructions occur among instruments during surgeries, posing challenges for target identification and tracking.

To address these issues, this paper proposes a high-precision binocular vision positioning method for surgical instruments using densely packed micro checkerboard markers. The markers used in this scheme require minimal modification to surgical instruments. Without compromising the doctor's regular surgical procedures, this approach effectively addresses challenges arising from instrument similarity and tracking issues caused by intersections and obstructions.

2. Related Works

The positioning systems used for trajectory tracking can be divided into two types: electromagnetic navigation systems and visual positioning systems.

The electromagnetic navigation system utilizes the principle of electromagnetic

induction, achieving the positioning of the target by having the detector receive signals from the magnetic field generator. For instance, the team led by Kroger, BJ ^[1], conducted research on speech generation using NDI's Aurora system, while the team led by Wigderowitz, C^[2], utilized Polhemus' FASTRAK system to record wrist movements in various cases, assessing the outcomes of wrist surgeries. Electromagnetic methods offer high positioning accuracy but are susceptible to interference from electromagnetic devices, making it challenging to achieve robust positioning in the complex electromagnetic environment of the operating room.

Visual positioning systems use three-dimensional imaging devices such as infrared, structured light, and binocular cameras. Cameras serve as sensors, analyzing optical signals to track target objects. Optical positioning methods are currently the mainstream in surgical navigation. For example, Stein, KM^[3], applied Medtronic surgical navigation system in dental surgery. Jackson, P^[4], and others compared the accuracy of NDI Polaris Spectra optical tracking system with NDI Aurora electromagnetic tracking system in surgical positioning, revealing that optical positioning significantly outperforms electromagnetic positioning accuracy.

The advantage of visual positioning systems lies in their high accuracy and immunity to environmental electromagnetic interference. However, they rely on positioning markers and cannot handle positioning failures caused by line-of-sight obstructions. Taking NDI's Polaris system as an example, the commonly used markers in optical positioning systems are standard optical spheres. These spheres come in two sizes, with diameters of 11.5mm and 13mm. A positioning component is formed by using four spheres to achieve three-dimensional spatial positioning. This component is securely attached to the distal end of the target surgical instrument, calculating the actual spatial position and orientation of the surgical instrument. While such positioning markers ensure accuracy, they are impractical for instruments like surgical knives and scissors due to significant modifications required for the surgical instruments.

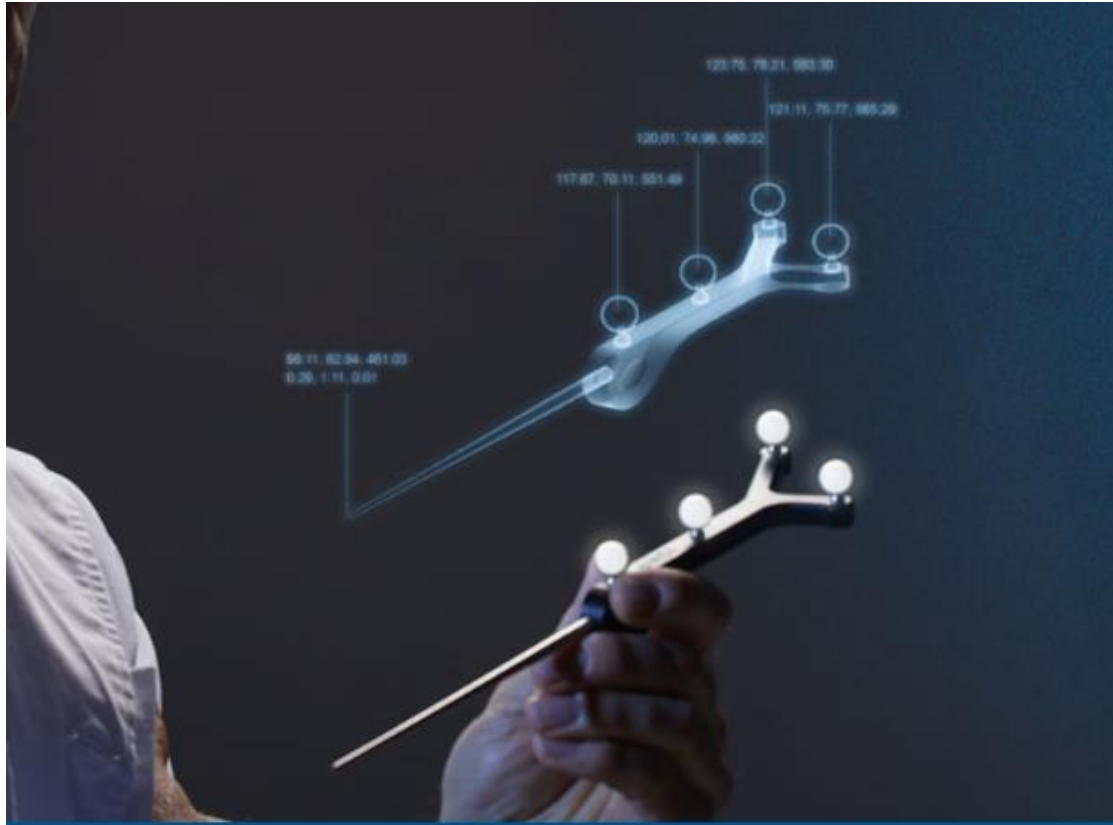


Figure1. Illustration of NDI optical positioning markers. A complete marker module consists of four standard optical spheres, fixed to the tip of an ablation needle using a structural component. Such markers have a voluminous size and are not suitable for marking conventional surgical instruments (such as surgical knives, scissors, forceps, etc.).

3. Methods

This article indirectly obtains the pose and position information of the target surgical instrument by locating a miniature checkerboard. During the trajectory tracking phase, it is necessary to know the intrinsic and extrinsic parameters of the binocular camera, the model of the surgical instrument, and the distribution of miniature checkerboard markers on it. The intrinsic matrix for the left camera is denoted as Int_L , for the right camera as Int_R , the rotation matrix from the coordinate system of the left camera to the coordinate system of the right camera as R , and the translation matrix as T . The set of markers on the target surgical instrument is denoted as M . These parameters can be obtained through a standardized process and can be extended to any binocular

camera system and target object.

The calculation of the posture and position of the target surgical instrument takes images captured by the binocular camera as input, yielding the rotation and translation matrices $Tr(t) = (R(t) \ T(t))$ that vary with time. Based on the set M of markers, the specific coordinates of each marker point and the tip on the target surgical instrument can be calculated.

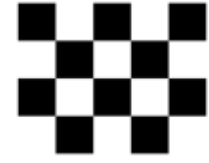
3.1 Micro checkerboard markers

This paper proposes the use of miniature checkerboard markers as a replacement for optical ball combinations with external structural components. These markers are employed for the positioning and tracking of conventional small surgical instruments (such as scalpels, scissors, forceps), achieving intraoperative multi-target surgical instrument tracking and positioning. The micro checkerboard, as a marker, offers the following advantages:

- 1.Small Occupied Area: The marker occupies a small area, allowing for multiple markers to be placed on the target surgical instrument, as illustrated in Figure 2. This addresses the issue of existing markers being unsuitable for small surgical instruments.
- 2.Simple Structure and Easy Recognition: The structure of the micro checkerboard is simple, making it easy to identify. It contains sufficient information for feature point recognition, matching, and posture and position calculations.
- 3.Dense Point Placement: The micro checkerboard allows for a dense point placement, resolving issues of marker obstruction due to gripping, tip entry into the lesion, and intersections during surgical instrument manipulation, which can lead to positioning failures.



(a)



3 × 4 checkerboard

(b)

Figure2 (a) Surgical instruments marked with micro checkerboard markers. (b) Checkerboard pattern.

The micro checkerboard markers used in this paper consist of 3 rows and 4 columns, forming a grid of 12 points. Each single black and white square has a side length ranging from 0.45mm to 0.75mm. The size of each marker, including the white border around the checkerboard, ranges from 2.25mm × 2.7mm to 3.75mm × 4.5mm. These markers can adhere to the surfaces of small surgical instruments without altering their overall shape. The micro checkerboard markers do not require additional structural components, making it possible to mark any rigid surgical instrument. Furthermore, they can mark the curved surfaces of surgical instruments, making them suitable for marking conventional instruments such as scalpels, scissors, and forceps.

Meanwhile, the structure of the micro checkerboard is simple, and it has a significant contrast with both surgical instruments and the surgical background, making it easily detectable by recognition algorithms. In specific marker recognition tasks, detection algorithms, based on calculating the position of the checkerboard, can further identify the coordinates of each grid point on the miniature checkerboard. Using the average values of multiple grid points to represent the center position of the checkerboard allows for sub-pixel-level coordinate calculations. Additionally, by considering the order and vertical relationships between grid points, a three-dimensional pose representation of the two-dimensional checkerboard can be calculated, further enhancing the accuracy of

feature point matching in binocular camera images.

Unlike the marking method using optical spheres, the proposed micro checkerboard markers in this paper do not require complete observation of the marker information to locate the posture and position of the target surgical instrument. In extreme cases, the algorithm can recognize and position the target surgical instrument using only 4 feature points. This method of positioning the target instrument addresses the recognition and positioning challenges arising from obscured markers due to gripping, tip insertion into the lesion area, and mutual instrument intersections during surgery.

3.2 Calculation of feature point coordinates

The calculation of feature point coordinates based on the binocular camera involves the following steps:

Step One: Camera Calibration

Camera calibration is the foundation of the positioning system's operation. It involves capturing multiple images of a calibration checkerboard in the target area to obtain the spatial relationship between the left and right cameras in the binocular system, known as intrinsic and extrinsic parameters. These parameters represent the relationships between the world coordinate system, left and right camera coordinate systems, and left and right camera pixel coordinate systems. In this paper, a standard checkerboard with a size of 12×9 and a grid point length of 20mm is used. Images of the checkerboard are taken within a target area of 300mm×400mm×400mm at distances ranging from 900mm to 1200mm from the camera. Approximately 50-70 pairs of checkerboard images are captured. The `OpenCV::StereoCalibration` function is then employed to calculate the intrinsic and extrinsic parameters of the binocular camera system.

The intrinsic parameters pertain to the concept of a single camera and represent the relationship between an object's coordinates in the camera coordinate system and its pixel coordinates in the camera. The intrinsic matrix is given by:

$$Intrinsic = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}.$$

Where f_x, f_y represents how many pixels are contained within each focal length on the focal plane, and c_x, c_y represents the pixel position of the camera's center point. The expression for projecting a point in space onto the camera plane is:

$$sP_c = Intrinsic \times (R \times P_w + T).$$

Where $P_c = (u, v, 1)^T$ represents the generalized pixel coordinates, $P_w = (x, y, z)^T$ represents the world point coordinates, R is the rotation matrix from the world coordinate system to the pixel coordinate system, T is the translation vector from the world coordinate system to the pixel coordinate system. s denotes the scale transformation, and represents the actual directional distance z . Using generalized coordinates $P_w = (x, y, z, 1)^T$ to represent world points, the expression can be written as:

$$sP_c = Intrinsic \times (R \ T) \times P_w.$$

The extrinsic parameters represent the spatial relationship between the two cameras, yielding the rotation matrix and translation vector between the left and right cameras. This can also be viewed as the coordinate transformation between the coordinate systems of the left and right cameras.

Step Two: Feature Point Recognition and Matching

In this paper, feature point recognition refers to the identification of micro checkerboard markers. Firstly, a small segmentation network is employed to identify points in the overall image that may be markers. EfficientNet-B0 [5] is used as the backbone for training on over 2000 images with a total of 27,000 marker point masks, resulting in stable segmentation results. Using the segmentation network's output, individual images of single checkerboards are cropped from the original image. The `cv::findchessboardcorners` function is then applied to each of these images to obtain grid point coordinates and the center point coordinates of the micro checkerboard, as illustrated in Figure 3.

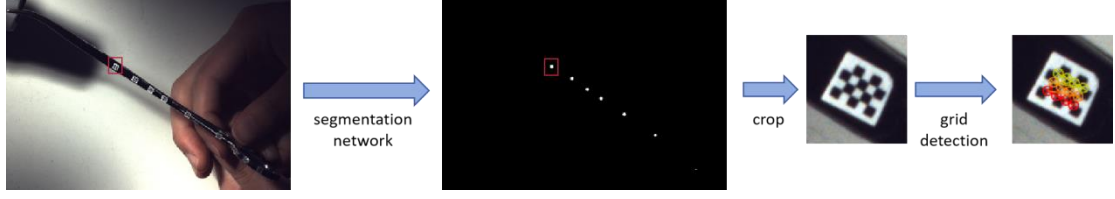


Figure 3. The process of feature point recognition. Firstly, individual markers are segmented using a deep learning network, and then the positions of grid points and the center points on each checkerboard marker are calculated.

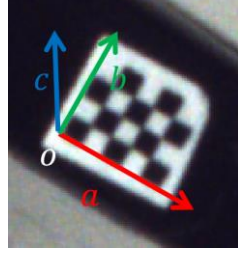


Figure 4. Representation of the three-dimensional pose of a checkerboard in pixel coordinate system. Denoting o as the origin, abc as three-axis directions, and the single checkerboard grid length as the unit length which structures a 3×3 pose matrix to represent the posture information of the checkerboard in the pixel coordinate system.

Based on the coordinates of the 12 grid points, the three-dimensional orientation information of the checkerboard in the pixel coordinate system can be calculated. As shown in Figure 4, we denote the coordinates of the long axis a as (x_a, y_a) , the coordinates of the short axis b as (x_b, y_b) , and the three-dimensional coordinates of the long axis a and short axis b as (x_a, y_a, z_a) and (x_b, y_b, z_b) , respectively. Additionally, the direction perpendicular to the checkerboard plane (axis c) is represented as (x_c, y_c, z_c) . Based on the orthogonality of the three axes and their equal lengths (lengths taken as the spacing between adjacent grid points), we can derive the following equations:

$$\begin{cases} x_a^2 + y_a^2 + z_a^2 = x_b^2 + y_b^2 + z_b^2 \\ x_a x_b + y_a y_b + z_a z_b = 0 \end{cases}.$$

Denote $u = x_a^2 + y_a^2 - x_b^2 + y_b^2$, $v = -x_a x_b - y_a y_b$, then :

$$z_a = \pm \sqrt{\frac{-u + \sqrt{u^2 + 4v^2}}{2}}.$$

$$z_b = \begin{cases} \frac{v}{z_a} & , \quad \text{if } z_a \neq 0 \\ 0 & , \quad \text{if } z_a = 0 \end{cases}.$$

The represent of axis c can be derived by $a \otimes b$.

Thus, we obtain the pose matrices $\begin{pmatrix} x_a & x_b & x_c \\ y_a & y_b & y_c \\ z_a & z_b & z_c \end{pmatrix}$ for each checkerboard marker point

in the left and right cameras. Combining this with the rotation matrix obtained in Section 3.1 for the extrinsic parameters, we can map the posture information of feature points in the left camera coordinate system to the right camera coordinate system.

Assuming the pose matrix P_L for the marker point in the left image is $(L_a \ L_b \ L_c)$, where $L_a = (x_{la} \ y_{la} \ z_{la})'$, $L_b = (x_{lb} \ y_{lb} \ z_{lb})'$, $L_c = (x_{lc} \ y_{lc} \ z_{lc})'$,. Similarly, for the marker point in the right image, the pose matrix P_R is $(R_a \ R_b \ R_c)$, where $R_a = (x_{ra} \ y_{ra} \ z_{ra})'$, $R_b = (x_{rb} \ y_{rb} \ z_{rb})'$, $R_c = (x_{rc} \ y_{rc} \ z_{rc})'$. The coordinates of P_L in the right image can be represented as $R \times P_L = (R \times L_a \ R \times L_b \ R \times L_c) \triangleq (L_a' \ L_b' \ L_c')$.

Define the cosine similarity between two pose matrices as below:

$$\cos - \text{similarity}(P_L, P_R) = \frac{L_a' \cdot R_a}{||L_a'|| \times ||R_a||} + \frac{L_b' \cdot R_b}{||L_b'|| \times ||R_b||} + \frac{L_c' \cdot R_c}{||L_c'|| \times ||R_c||}$$

By considering the cosine similarity between the pose matrices of feature points and whether their center positions are at the same y-coordinate, it is possible to accurately determine whether the two feature points in the left and right images match. Unlike the typical matching approach in 3D reconstruction, the method used in this paper significantly reduces the number of feature points, leading to improved matching accuracy and feature point localization precision.

Step Three: Calculation of Feature Point Coordinates

In stereoscopic vision 3D reconstruction, the depth information of the point cloud is

typically calculated based on the disparity obtained from feature point matching. The advantage of this disparity-based approach is its ability to quickly acquire overall depth information for rapid reconstruction. However, this method assumes that the binocular cameras are on the same plane with parallel optical axes, which may not hold true in real-world three-dimensional space. In surgical navigation applications where high positioning accuracy is required, the errors introduced by this method cannot be ignored. Therefore, in this paper, we improve the formula for calculating feature point coordinates based on the fundamental principles of camera perspective imaging.

Based on the camera calibration results from Step One, we obtain the following relationship between the pixel coordinates of the left and right cameras:

$$\begin{cases} R_{LR} = R_R \times R_L^{-1} = R \\ T_{LR} = T_R - R_{LR}T_L = T \end{cases}.$$

Where R_{LR}, T_{LR} represents the rotation matrix and the translation vector, respectively, indicating the transformation from the left camera coordinate system to the right camera coordinate system. R, T are the results of the extrinsic parameter calibration.

From Section 3.1, we can express the coordinates of the same world point captured by the left and right cameras as:

$$\begin{cases} s_L P_L = Int_L \times (R_L \times P_w + T_L) \\ s_R P_R = Int_R \times (R_R \times P_w + T_R) \end{cases}.$$

Thus,

$$s_L \times R \times Int_L^{-1} P_L - s_R Int_R^{-1} P_R = RT_L - T_R = -T.$$

Where P_L, P_R represent pixel coordinates, Int_L, Int_R represent the intrinsic matrix, R, T represent the extrinsic matrix, and the equation forms an over determined system of linear equations with respect to s_L, s_R . The value of s_L, s_R can be derived the least squares method:

$$\begin{pmatrix} s_L \\ -s_R \end{pmatrix} = (A^T A)^{-1} (-A^T T).$$

Where $A = (R Int_L^{-1} P_L | Int_R^{-1} P_R)$.

In fact, s_L represent the coordinates of the observed point in the left camera coordinate

system. If we choose the left camera coordinate system as the world coordinate system, then the world coordinates of the observed point would be $P_w = s_L \times Int_L^{-1} P_L$.

3.3 Calculation of the Pose and Position of Surgical Instruments

In the localization process, we can transform the problem of locating surgical instruments into a problem of determining the pose and position of the point cloud formed by the markers on the instruments. Then, based on the pose and position information of the point cloud formed by the markers, we can calculate the position information for each point on the surgical instrument, including the tip position.

Firstly, we need to model the feature point cloud of the surgical instrument. We rotate the surgical instrument with markers for several turns in a binocular camera system, capturing a series of images and calculating the three-dimensional coordinates of the feature point cloud in the images. Based on the spatial relationship between partially repeated point clouds in adjacent frames, we calculate the motion transformation of the target surgical instrument between adjacent frames. Then, considering the target surgical instrument as stationary, these observed images can be seen as views of the feature point cloud from different angles. By computing the motion transformations between adjacent frames and their composition, we can calculate the observed angles for each frame. Due to errors in observation and calculation, bundle adjustment ^[6] is employed. With the obtained observed angles as initial conditions, the positions of the observed point clouds are optimized to finally obtain the model of the feature point cloud of the surgical instrument.

During the localization process, the observed local point cloud is matched with the overall point cloud model of the surgical instrument. In cases where the feature marker points are heavily occluded, the number of observed local point clouds is much smaller than the number of model point clouds. In such situations, traditional algorithms using iterative closest point (ICP) algorithms ^[7] by optimizing the energy function

$\operatorname{argmin}_{R,T} \sum (p_i - (R \times q_j + T))^2$ and their variants (where p_i represents the local point cloud, and q_j represents the model point cloud) can lead to multiple solutions and misinterpretations due to the asymmetry in data scale.

To address the above issues, this paper leverages the sparsity of feature points and the high accuracy of the positioning results. For a model point cloud consisting of n points, we calculate the lengths of the sides of any triangle formed by three points (denoted as $m_i = \{a_i, b_i, c_i\}, i = 1, 2, \dots, C_n^3$) among the n points as the feature representation for matching, and this set is denoted as M_{mesh} . Subsequently, based on the three-dimensional coordinates of observed feature points in the images obtained in Section 3.2, we calculate the set D_{mesh} , which consists of the side lengths of triangles formed by any three observed points. Due to the uniqueness of the triangle determined by its three sides, each element in D_{mesh} corresponds to at most one element in M_{mesh} . This establishes a one-to-one correspondence between point pairs. Utilizing the positional variations of the point pairs, we calculate the coordinate transformation representation from the model to the observed point cloud. By incorporating model-based modeling, we obtain the positional information for any point on the surgical instrument, including the tip position.

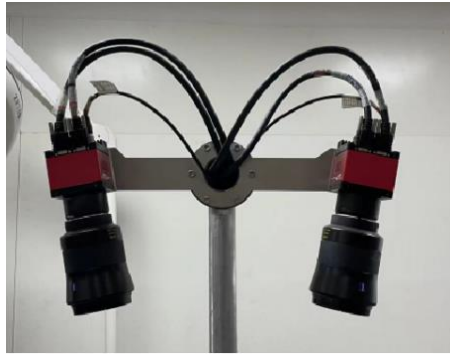
4. Implementation details

This paper uses a high-definition binocular camera system (shown in Fig. 5), with a camera resolution of 9280×6992 , a spacing of 400 mm between the two cameras, an optical axis angle of 22.62° , and a lens focal length of 55 mm. In order to accurately localize the moving surgical instruments, this paper uses a surface array camera and a simultaneous trigger to control the exposure of the two cameras. The localization area is 900mm to 1200mm from the binocular camera baseline, 300mm×400mm×400mm of the effective range of motion of the surgical instruments. The relevant parameter settings in this paper are shown in Table 1.

Table 1 Details of binocular camera parameters

Fig. 5 Binocular camera system

Parameter Name	Parameter Selection	Parameter Name	Parameter Selection
Camera Resolution	9280×6992	Camera Pitch	400mm
Camera Angle	22.62°	Exposure setting	Global exposure
Exposure Trigger	Synchronized trigger	Ambient light intensity	10000lux
Aperture	f8	Exposure time	1.500ms
Exposure Gain	12 times	Lens focal length	55mm fixed focus



5. Experiment & Result

5.1 Single Target Surgical Instrument Positioning

The tweezers are used as an example to show the results of single-target surgical instrument localization experiments. The surgical instrument attitude position localization is a hexadecimal group of (x, y, z, r_x, r_y, r_z) , and after calculating to get this hexadecimal group, the result is reprojected back to the left camera image according to the left camera internal reference in the camera calibration result, and the visualization result as in Fig. 6 is obtained.

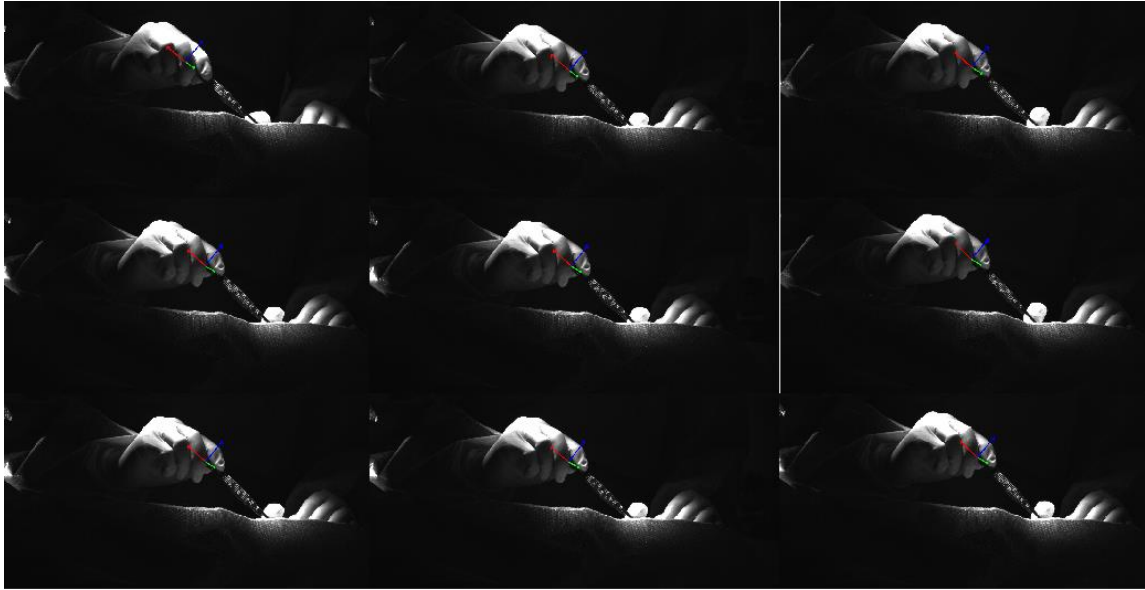


Fig. 6 Schematic diagram of single-target surgical instrument (forceps)

localization. The attitude position localization results obtained in this paper are six-degree-of-freedom localization results in three-dimensional space. In the visualization results, the camera imaging's projective formula is used to map to a feature point on the target surgical instrument, and its attitude and rotation are marked with red, green and blue colors.

5.2 Multi-Target Surgical Instrument Positioning

Conventional surgical instruments such as scalpel, hemostat, scissors, and electro-surgical knife are taken as examples to demonstrate the effect of multi-target surgical instrument localization, and the results are shown in Figures 7 and 8.



Fig. 7 Demonstration of the effect of multi-target localization using an electric knife, screwdriver, and hemostat as the target surgical instruments. A feature point on the target instrument is taken to indicate positional localization, and three axes above it indicate attitude localization.

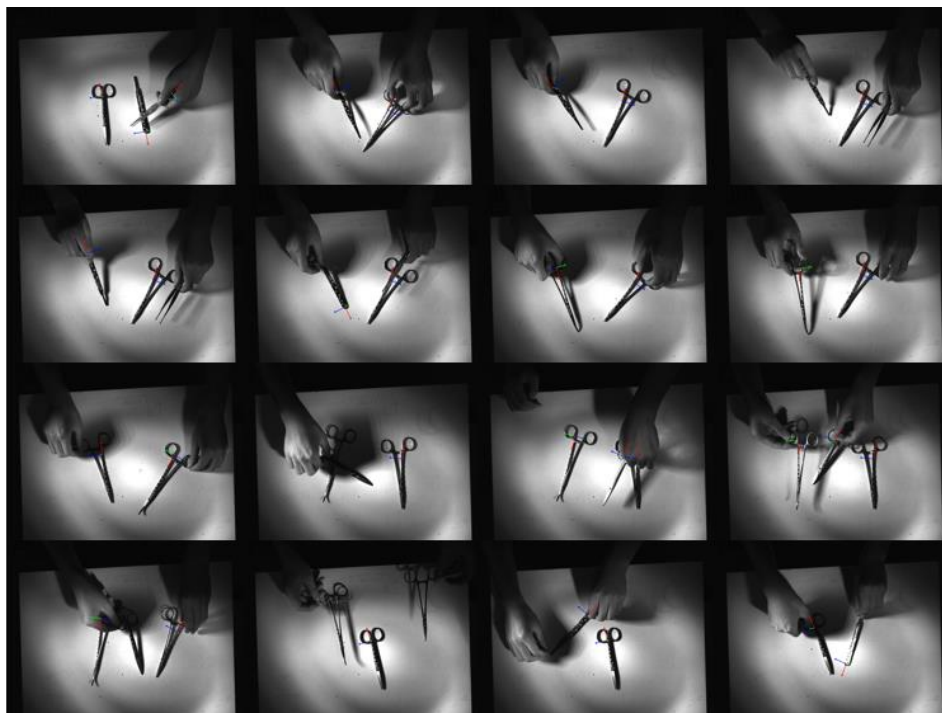


Fig. 8 Demonstration of the effect of multi-target localization with hemostat, forceps, needle holder, scissors and scalpel as target surgical instruments. A feature point on the target instrument is taken to indicate positional localization, and three axes above it indicate attitude localization.

Multi-target surgical instrument localization can be divided into multiple single-target surgical instrument trajectory tracking tasks using different specifications of checkerboard grid markers to mark different surgical instruments. Multi-target surgical instrument localization can be accomplished by using different channels to output the segmentation results of different checkerboard grids in the segmentation network mentioned in subsection 3.2.

5.3 Precision Analysis

Three-dimensional surgical instrument positioning involves six degrees of freedom, including three degrees of freedom for the rotation matrix and three degrees of freedom for the translation vector. We use a high-precision slide table and a three-axis rotary table to test the positioning accuracy of rotation and translation, respectively. The displacement accuracy of the slide table is 0.001mm, and the rotation angle accuracy of the three-axis rotary table is 0.001° .

In the displacement test, the surgical instruments were fixed to the sliding table platform and the moving position of the sliding table was controlled. The displacement of the sliding table was used as the reference value of the distance of the relative movement of the surgical instrument, and the difference of the relative displacement, compared with the relative movement distance calculated by the method of this study, was used to indicate the positional positioning error. The displacement accuracy test set one test point per 30 mm grid in the x-axis and y-axis in the $200\text{mm} \times 150\text{mm}$ target area, totaling 48 test points, to calculate the displacement deviation. The test result is that the average error of positioning is 0.153mm, the standard deviation is 0.161mm, and the maximum deviation is 0.467mm. The specific displacement accuracy test program and

results are shown in Appendix I .

In the stance test, the surgical instrument is fixed to the rotary table platform, and the rotation angle of the rotary table is controlled. The rotation angle of the rotary table was used as the distance reference value for the relative movement of the surgical instrument, and the difference of the relative rotation angle, compared with the relative rotation angle calculated by the method of this paper, was used to indicate the attitude error. Attitude accuracy test were tested on three mutually perpendicular planes of the rotation angle, with 5° as the interval, each taking 10 test points, to obtain the average error of the attitude of 0.076° , the standard deviation of 0.095° , the maximum deviation of 0.207° . The specific attitude accuracy test program and the results are shown in Appendix II.

Accuracy tests show that the intraoperative surgical instrument positioning accuracy proposed in this paper fully meets the needs of high-precision surgery.

6. Conclusion

The binocular vision intraoperative surgical instrument trajectory acquisition scheme based on miniature checkerboard grid markers proposed in this paper solves the problems of data acquisition in the complex environment of the operating room, high-precision localization computation of surgical instruments, difficult labeling of conventional surgical instruments, and identification and tracking of surgical instruments under large-area occlusion. The localization accuracy test shows that the average localization error of the localization scheme in this paper is sub-millimeter, which meets the demand of high-precision surgical trajectory localization.

In the experimental stage, this paper uses tessellated markers with stickers, which have not yet met the high-temperature sterilization conditions in the operating room. We further tested the micro tessellated lattice markers with zirconia ceramic sheets, which can realize intraoperative localization of surgical instruments under the localization and tracking process of this paper, and in the future, the ceramic sheets can be embedded into the surface of the surgical instruments and be applied to the real surgical scenarios.

Based on the research achievement in this paper, the reconstruction of intraoperative surgical instrument motion trajectories can be realized, which will lay the foundation for collecting surgical trajectory data, establishing a surgical trajectory database, artificial intelligence-assisted surgical navigation, digitalized surgery, and many other fields.

Reference

- [1]Kroger BJ, Pouplier M, Tiede MK.An Evaluation of the Aurora System as a Flesh-Point Tracking Tool for Speech Production Research[J].Journal of Speech Language & Hearing Research, 2008, 51(4):914-921.DOI:10.1044/1092-4388(2008/067).
- [2]Wigderowitz C , Scott I , Jariwala A ,et al.Adapting the Fastrak System for Three-Dimensional Measurement of the Motion of the Wrist[J].Journal of Hand Surgery European Volume, 2007, 32(6):700-704.DOI:10.1016/j.jhse.2007.06.019.
- [3]Stein, Kyle M .Use of Intraoperative Navigation for Minimally Invasive Retrieval of a Broken Dental Needle[J].Journal of Oral & Maxillofacial Surgery, 2015, 73(10):1911-1916.DOI:10.1016/j.joms.2015.04.033.
- [4]Jackson P , Simon R , Linte C .Surgical Tracking, Registration, and Navigation Characterization for Image-guided Renal Interventions[J].IEEE, 2020.DOI:10.1109/EMBC44109.2020.9175270.
- [5]Tan, Mingxing and Le Quoc, V. EfficientNet: rethinking model scaling for convolutional neural networks. ICML 2019.
- [6]Triggs B, McLauchlan P F, Hartley R I, et al. Bundle adjustment-a modern synthesis[C]. International workshop on vision algorithms. Springer Berlin Heidelberg,1999: 298-372.
- [7]Raúl San José Estépar, Brun A , Westin C F . Robust Generalized Total Least Squares Iterative Closest Point Registration[J]. DBLP,2004.DOI:10.1007/978-3-540-30135-6_29.

Appendix I Displacement Accuracy Test

3D surgical instrument positioning involves 6 degrees of freedom, including 3 degrees

of freedom for the rotation matrix and 3 degrees of freedom for the translation vector. We used a high-precision Newport slip table to test the positioning accuracy of the translation. The displacement accuracy of the Newport slip table is 0.001mm.

The surgical instruments were fixed to the sliding table platform, and the moving position of the sliding table was controlled. The displacement of the sliding table was used as a reference value for the distance of the relative movement of the surgical instrument, and the difference of the relative displacement, compared with the distance of the relative movement calculated by the method of this study, was used to indicate the positional positioning error.

The experiment was divided into 8 groups, covering all positions of the pair target area. The experimental procedure is shown in Table 2 and Figure 6.

Table 2 Displacement Test Record Sheet

	Direction	Starting position/mm	Ending position/mm	Interval distance/mm	No. of pickup points
Group 1	x	100	490	30	14
Group 2	x	100	490	30	14
Group 3	x	100	460	30	13
Group 4	y	200	470	30	10
Group 5	y	200	470	30	10
Group 6	y	200	470	30	10
Group 7	xy	100	520	30	15
Group 8	xy	150	510	30	13

Group 1



Group 2



Group 3



Group 4



Group 5



Group 6



Group 7



Group 8



Figure 9 Positional accuracy test chart

The test results are shown in Table 3.

Table 3 Accuracy test results. Shows the reference value, measured value and error (in mm) for each group of 6 test points. Where the error is the error of relative displacement as (absolute value of measured value - absolute value of reference value).

Groups		1	2	3	4	5	6
Group 1	Reference value	-180	-120	-60	60	120	180
	Measured value	-179.655	-119.941	-59.99	59.972	119.921	179.746
	Error	-0.345	-0.059	-0.01	-0.028	-0.079	-0.254
Group 2	Reference value	-180	-120	-60	60	120	180
	Measured value	-179.629	-119.971	-59.998	59.987	119.844	179.662
	Error	-0.371	-0.029	-0.002	-0.013	-0.156	-0.338
Group 3	Reference value	-180	-120	-60	60	120	180
	Measured value	-179.52	-119.783	-59.92	59.913	119.715	179.533
	Error	-0.479	-0.217	-0.08	-0.087	-0.285	-0.467
Group 4	Reference value	-150	-90	-30	30	90	120
	Measured value	-149.658	-89.838	-29.964	29.948	89.784	119.62
	Error	-0.342	-0.162	-0.036	-0.052	-0.216	-0.38
Group 5	Reference value	-150	-90	-30	30	90	120
	Measured value	-149.849	-89.96	-29.996	29.991	89.976	119.923
	Error	-0.151	-0.04	-0.004	-0.009	-0.024	-0.077
Group 6	Reference value	-150	-90	-30	30	90	120
	Measured value	-149.716	-89.922	-29.989	29.982	89.921	119.844
	Error	-0.284	-0.078	-0.011	0.018	0.079	0.156
Group 7	Reference value	-150	-90	-30	30	90	150
	Measured value	-149.611	-89.767	-29.965	29.964	89.905	149.878
	Error	-0.389	-0.233	-0.035	-0.036	-0.095	-0.122
Group 8	Reference value	-150	-90	-30	30	90	150
	Measured value	-149.673	-89.977	-30.023	29.982	89.94	149.872
	Error	-0.327	-0.023	0.023	-0.018	-0.06	-0.128

The test results were a mean error in positioning of 0.153mm, a standard deviation of 0.161mm and a maximum deviation of 0.467mm.

Appendix II Attitude Accuracy Test

Three-dimensional surgical instrument positioning involves six degrees of freedom, of which the rotation matrix has three degrees of freedom. We tested the rotational accuracy using a high-precision three-axis rotary table. The rotational angular accuracy of the three-axis rotary table is 0.001° .

The surgical instruments are fixed to the turntable platform, and the rotation angle of the turntable is controlled. The rotation angle of the rotary table was used as a distance reference value for the relative movement of the surgical instrument, and the difference of the relative rotation angle, compared with the relative rotation angle calculated by the method of this study, was used to represent the attitude error.

The experiments were divided into three groups to test the rotation angles on three mutually perpendicular planes. The test procedure is shown in Table 4 and Figure 7.

Table 4 Attitude Accuracy Test Record Sheet

	Direction	Starting position/ $^{\circ}$	End position/ $^{\circ}$	Interval angle/ $^{\circ}$	Number of points
Group 1	xy	0	360	30	13
Group 2	yz	30	120	10	10
Group 3	xz	50	130	10	9



Fig. 7 Attitude accuracy test plot

The test results are shown in Table 5.

Table 5 Accuracy test results. Shows the reference value, measured value and error (in $^{\circ}$) for each group of 6 test points. Where the error is the error of the relative rotation angle as (absolute value of the measured value - absolute value of the reference value).

Groups		1	2	3	4	5	6
Group 1	Reference value	30	90	150	150	90	30
	Measured value	30.01	90.006	149.978	149.996	89.985	29.963
	Error	0.01	0.006	-0.022	-0.004	-0.015	-0.037
Group 2	Reference value	10	20	30	40	50	60
	Measured value	9.934	20.014	29.953	40.046	50.092	60.036
	Error	-0.064	0.014	-0.047	0.046	0.092	0.036
Group 3	Reference value	10	20	30	40	50	60
	Measured value	9.954	19.939	30.057	40.073	50.11	60.207
	Error	-0.046	-0.061	0.057	0.073	0.11	0.207

The mean error of the attitude was obtained to be 0.076° with a standard deviation of 0.095° and a maximum deviation of 0.207° .