Fingertip blood sampling robot navigation system based on binocular vision

Xi Li¹,²*, Jiajia Ji², Qin Zhou²

¹ Foundation Department, Chongqing Medical and Pharmaceutical College, Chongqing, 401331, China
² School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China

*Corresponding author’s e-mail: d190101005@stu.cqupt.edu.cn

Abstract. In the fingertip intelligent blood collection robot, we need to move the blood collection needle to the designated position quickly and accurately, but it is difficult to meet the two indicators of speed and accuracy at the same time. This paper proposes a fast navigation method based on binocular vision. This method obtains the coordinates of the fingertip blood sampling point through three-dimensional reconstruction and contour extraction, and guides the robotic arm to quickly move the blood sampling needle to a position near the blood sampling point according to the coordinate position. The method in this paper is used for the rough navigation of blood collection robots, and it is verified by experiments that this method can meet the needs of blood collection robots.

1. Introduction

Fingertip blood sampling has the advantages of light pain and convenient and quick blood sampling, which makes it highly acceptable to people. In addition, under the current COVID-19 epidemic, medical staff can easily lead to infectious diseases when manually collecting blood. Compared with traditional manual blood collection, the automated fingertip blood collection robot has outstanding advantages in blood collection efficiency, quality and prevention of infection of medical staff.

Binocular vision is a very important part of the intelligent blood sampling robot. Only under the guidance of vision, the robot can obtain the correct information of the target, so as to control the robotic arm to complete the specified task. The vision system must recognize and track the position and posture of the target object on the basis of calibration[1-3]. In order to obtain the spatial position of the target object through binocular vision and provide the position information of the target for the subsequent manipulator operation, we need to study the correspondence between the pixel coordinates of the plane image and the spatial coordinates of the points in the scene. Through the optical model of the camera, we can obtain the relationship between these two coordinates. Intelligent blood collection robots need to complete the blood collection process quickly and accurately[4-7], and the two indicators of fast and accurate are in conflict with each other. Therefore, we use binocular vision navigation to quickly guide the blood collection needle to a position near the target point, and then use laser precise guidance to achieve The blood collection process.
2. Navigation system

2.1. Three-dimensional reconstruction

The main content of three-dimensional reconstruction based on binocular vision includes five parts: camera calibration, stereo correction, stereo matching, parallax calculation and three-dimensional reconstruction. The calibration of a single camera is mainly to obtain the internal parameters and distortion parameters of the camera. According to the internal parameters, the imaging plane coordinates can be converted to the camera coordinates, so that the space object point corresponds to the image point[8-9]. Bi-target positioning is mainly based on the internal parameters of the two cameras to obtain the external parameters of the camera. According to the external parameters, we can know the positional relationship between the two cameras. The obtained calibration data establishes the connection between the space object and the image plane, and then calculates the three-dimensional reconstruction information. After the camera calibration is completed, the parallax calculation must be performed. Stereo matching is to find the matching feature points from the left and right images collected by the binocular camera and establish the corresponding relationship to calculate the parallax of the matching points, therefore, the stereo matching is the most critical step[10-12] in the entire three-dimensional reconstruction work.

This paper proposes an improved Scale Invariant Feature Transform operator to describe feature points, which has higher operating efficiency and stronger stability. The Hessian matrix is the core of the SURF feature extraction algorithm and its determinant can be used as the basis for judging extreme points. The Hessian matrix expression of a certain point is:

\[ H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{yx}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \]  

(1)

Where \( L_{xx}(x, \sigma) \), \( L_{xy}(x, \sigma) \), \( L_{yx}(x, \sigma) \), \( L_{yy}(x, \sigma) \) is the convolution value of the second-order partial derivative, and the determinant of the Hessian matrix is:

\[ \text{det}H = L_{xx}L_{yy} - L_{xy}^2 \]  

(2)

We compare the pixels processed by the Hessian matrix with the other 26 pixels in the surrounding neighborhood. If the pixel is the extreme value of these 26 pixels, it will be retained as the preliminary feature point, then, the three-dimensional linear interpolation method is used to obtain the feature points at the sub-pixel level, and the feature points less than a certain threshold are removed, so that the point with the strongest feature can be selected.

Parallax refers to the difference in the coordinates of the corresponding points of the two images obtained when observing the same point from two different directions. There is also a noun referred to as the parallax angle, which, as the name implies, is the angle between the line between two different positions and the observed point[13-14]. The baseline refers to the distance between two observation points, and it is an indispensable condition for calculating the depth index of the object to be measured, just like the parallax angle. The completion of the previous series of key steps can obtain the camera parameters, the distortion coefficient and the characteristic relationship between the points, and finally the depth information of the object can be obtained based on it. The essence of three-dimensional reconstruction is to calculate the three-dimensional coordinates by obtaining the depth information of the object to be measured.

2.2. Coordinate transformation

We need to convert between various coordinate systems when positioning and navigating. The world coordinate system, camera coordinate system, image physical coordinate system and image pixel coordinate system correspond to the point conversion relationship shown in Figure 1.
The unit of pixel coordinate system is pixel, where \((u, v)\) is a pixel in the image. The unit of the physical coordinate system is millimeter. Assuming that \(dx\) and \(dy\) are the physical size of a single pixel on the x-axis and y-axis, \((u_0, v_0)\) are the coordinates of point \(O_3\) in the pixel coordinate system, and the conversion relationship of the coordinate system in the two-dimensional image plane is as follows:

\[
\begin{align*}
    u &= \frac{x}{dx} + u_0 \\
    v &= \frac{y}{dy} + v_0
\end{align*}
\]

It is expressed in matrix form as Equation (4).

\[
\begin{bmatrix}
    u \\
    v \\
    1
\end{bmatrix} = \begin{bmatrix}
    1/dx & 0 & u_0 \\
    0 & 1/dy & v_0 \\
    0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
    x \\
    y \\
    1
\end{bmatrix}
\]

The point \(P(X_c, Y_c, Z_c)\) is a point in the camera coordinate system, and there is a point \(p(x, y)\) corresponding to it in the image physical coordinate system. Based on the triangular similarity principle, we can obtain the conversion relationship between the camera coordinate system and the image physical coordinate system is:

\[
\begin{align*}
    x &= f \frac{X_c}{Z_c} \\
    y &= f \frac{Y_c}{Z_c}
\end{align*}
\]

Where \(f\) is the focal length of the camera, and the Equation (5) is expressed in homogeneous coordinates as follows:

\[
\begin{bmatrix}
    x \\
    y \\
    1
\end{bmatrix} = \begin{bmatrix}
    f & 0 & 0 & X_c \\
    0 & f & 0 & Y_c \\
    0 & 0 & 1 & Z_c
\end{bmatrix}
\]

There is a transformation relationship between the rotation matrix \(R\) and the translation vector \(t\) between the points in the camera coordinate system and the points in the world coordinate system. The relationship between the camera coordinate system and the world coordinate system is shown in Equation (7):

\[
\begin{bmatrix}
    X_c \\
    Y_c \\
    Z_c
\end{bmatrix} = [R \ t] \begin{bmatrix}
    X_w \\
    Y_w \\
    Z_w
\end{bmatrix}
\]

Where \(R\) represents the rotation matrix of the point in the world coordinate system for coordinate transformation to the camera coordinate system. It is a 3-row and 3-column unit orthogonal matrix, and \(t\) represents the translation vector of coordinate transformation, which is a 3-row 1 column vector.

3. Results and Discussion

The mechanical arm system of the blood collection robot is composed of the mechanical arm body and multi axis motion control card, as shown in Figure 2. The robotic arm is made of aluminum alloy and
has four degrees of freedom, including two rotary joints, one vertical motion joint, and one local degree of freedom to control the end load. The robotic arm control system uses C++ programming to build a software platform for robotic arm motion control, and uses the dynamic library provided by the motion control card to complete the motion control functions of motor position, speed, hardware capture, and comparison output.

Firstly, the finger and its surrounding environment are reconstructed by binocular camera, the fingertip contour is extracted, the contour coordinate points are obtained, and the center point coordinates of fingertip contour are calculated. Then control the mechanical arm to quickly move the blood sampling needle to 2cm above the coordinate of the central point. After 10 experiments, the blood sampling needle moved to the designated position within 1.8s. At the same time, the error range of the three directions of X, Y, and Z is shown in Figure 3. It can be seen that the average error of the three directions does not exceed 2mm, and this indicator can meet the requirements of rough navigation for blood collection robots.

![Figure 2 Robotic arm system: (a) Robotic arm; (b) Motion control card.](image)

![Figure 3 Errors in X, Y, and Z directions](image)
4. Conclusions
In the intelligent fingertip blood collection robot, the fast positioning and navigation method is needed
to move the blood collection needle to the fingertip blood collection point quickly. This paper
proposes a three-dimensional navigation method based on binocular vision. Firstly, the finger is three-
dimensionally reconstructed. Then, the coordinate position of the fingertip center is obtained by
contour extraction, and finally, the experiment is verified with the manipulator. The results proves that
the method proposed in this paper can meet the requirements

Acknowledgments
This work was partially supported by the natural science foundation of Chongqing, China (No.
cstc2020jcyj-msxmX0702), the science and technology research program of Chongqing municipal
education commission (Grant No. KJQN20202804).

References
[1] Wen S, Hu X, Ma J, et al. Autonomous robot navigation using Retinex algorithm for multiscale
image adaptability in low-light environment. Intelligent Service Robotics, 2019, 12(4):359-369.
[2] Francisco J, LA Vázquez, Casta?Eda C E, et al. Neural Block Control via Integrator
Backstepping for a Robotic Arm in Real-Time. Neural Processing Letters, 2018:1-17.
[3] Ran, Wang Y, Han M Z, et al. Accuracy study of a binocular-stereo-vision-based navigation
robot for minimally invasive interventional procedures. World Journal of Clinical Cases,
2020, v.8(16):69-78.
[4] Carron A, Arcari E, Wermelinger M, et al. Data-Driven Model Predictive Control for
Trajectory Tracking With a Robotic Arm. IEEE Robotics and Automation Letters, 2019,
4(4):3758-3765.
[5] Dongming G E, Sun G, Zou Y, et al. Impedance control of multi-arm space robot for the
capture of non-cooperative targets. Journal of Systems Engineering and Electronics, 2020,
31(5):1051-1061.
[6] Tseng C S, Chen H H, Wang S S, et al. Image-guided robotic navigation system for
neurosurgery. Journal of Field Robotics, 2015, 17(8):439-447.
[7] Chen T, Chen C Y. A New Viewpoint on Control Algorithms for Anthropomorphic Robotic
Arms. Journal of Intelligent and Robotic Systems, 2020, 99(1).
[8] Kong X Z, Duan X G, Wang Y G. An integrated system for planning, navigation and robotic
assistance for mandible reconstruction surgery. Intelligent Service Robotics, 2016, 9(2):113-121.
[9] Gui H, Zhang S, Luan N, et al. A Novel System for Navigation-and Robot-Assisted
Craniofacial Surgery: Establishment of the Principle Prototype. Journal of Craniofacial
Surgery, 2015, 26.
[10] Yuan, Zhang, Zhicheng, et al. Calibration method for hand-eye system with rotation and
translation couplings.. Applied optics, 2019, 58(20):5375-5387.
[11] Ali I, Suominen O, Gotchev A, et al. Methods for Simultaneous Robot-World-Hand–Eye
Calibration: A Comparative Study. Sensors (Basel, Switzerland), 2019, 19(12).
[12] Sharifzadeh S, Biro I, Kinnell P. Robust hand-eye calibration of 2D laser sensors using a single-
plane calibration artefact. Robotics and Computer-Integrated Manufacturing, 2020,
61(Feb.):101823.1-101823.10.
[13] Lee S, Shim S, H Ha, et al. Simultaneous Optimization of Patient-Image Registration and
Hand-Eye Calibration for Accurate Augmented Reality in Surgery. IEEE Transactions on
Biomedical Engineering, 2020, PP(99):1-1.
[14] Krittin, Pachtrachai, Francisco, et al. Adjoint Transformation Algorithm for Hand-Eye
Calibration with Applications in Robotic Assisted Surgery.. Annals of biomedical
engineering, 2018.