Proceedings of 2020 IEEE
International Conference on Mechatronics and Automation
October 13 - 16, Beijing, China

An Improved Visual Auxiliary Algorithm for the Vascular Interventional Surgical Robot based on Neural Network

Jian Guo[1], Suxiang Feng[1]
[1]Tianjin Key Laboratory for Control Theory & Applications in Complicated Systems and Intelligent Robot Laboratory
Tianjin University of Technology
Binshui Xidao 391, Tianjin, China
jianguo@tjut.edu.cn；1549158571@qq.com;

Shuxiang Guo[1,2*]
[2] Department of Intelligent Mechanical Systems Engineering, Faculty of Engineering,
Kagawa University
2217-20, Hayashi-cho, Takamatsu 761-0396, Japan
*Corresponding author: guo@eng.kagawa-u.ac.jp；

Abstract **- Given the various shortcomings of traditional interventional surgery, it has become a trend to develop vascular interventional robots to assist doctors in completing the surgery. However, the existing visual auxiliary for vascular interventional robots rely on doctors' experience to make judgments based on images. Therefore, our team proposed a method of transforming image coordinates and spatial coordinates based on the visual auxiliary of existing platforms. This paper shows the platform briefly at first. Then we introduce the traditional image coordinate transformation and use the neural network to simplify calculation and parameter preparation process based on the traditional way. Finally, the accuracy and practicability of the transformation between the vision-assisted image coordinates and the space coordinates of the vascular interventional surgical robot based on the neural network are proved through experiments. We can apply this algorithm to the visual assistant and give detailed numerical feedback to the doctor intuitively, greatly improving the accuracy of the operation.**

Index Terms -Vascular interventional surgical robot system, Visual -auxiliary, Coordinate transformation, neural network.

I. INTRODUCTION

With the acceleration of social aging and urbanization, and the background of unhealthy lifestyles of residents, the individual exposure to national cardiovascular disease risk factors has increased significantly, leading to a continuous increase in the number of cardiovascular diseases in China. The number of cardiovascular disease patients in China continues to increase, and it is estimated that the current number of cardiovascular disease patients in China is 290 million[1]. The most direct and effective method for vascular diseases is minimally invasive vascular interventional surgery. Minimally invasive vascular interventional surgery refers to the doctor's guidance of medical images to push the catheter/guidewire to the lesion to complete the operation for related diseases. Compared with traditional surgery, minimally invasive vascular interventional surgery has been widely used because of its advantages such as less injury, lighter pain[2], and faster postoperative recovery.

However, there are several disadvantages of traditional vascular interventional surgery. For example, doctors need to use the X-ray imaging system to complete the intubation operation in the operating room. Long-term accumulated radiation will cause harm to the doctor's body. And highly concentrated operating doctors are prone to fatigue and affect the quality of the operation. Due to the low accuracy of manual intubation, when the force is too great or the operation is improper, dangerous situations such as puncture of blood vessels may occur, causing great harm to patients[3]. Because of these problems, the combination of robot technology and minimally invasive surgery to develop surgical robots to assist cardiovascular disease has huge clinical needs.

Since 2012, Kagawa University in Japan has developed a robotic system for vascular interventional surgery based on force-visual feedback[4]. This system combines to force feedback and visual feedback to better guarantee the safety of surgery[5]. In 2016, the United States-based Corindus Vascular Robotics developed a new vascular interventional robot system called CorPath GRX. The system has an extended arm, a touch screen display and is equipped with an automatic catheter guidance function. Work flow greatly improves the clinical capabilities of the system[6]. In 2018, Professor Guo Shuxiang's team at the Beijing Institute of Technology developed a robotic system for vascular interventional surgery that can be used in the clinic. The master of the system uses two Touch X haptic devices from Geomagic Company to control the catheter and guidewire and provide force in real-time The feedback function, and real-time monitoring of the surgical situation through visual auxiliary has good safety and stability[7].

However, these systems also have some disadvantages. In terms of the visual auxiliary, doctors monitor the progress of the operation with the naked eye and use DSA radiographic images to determine the position of the catheter and guidewire in the blood vessel based on experience. This requires the doctor to accumulate a lot of experience to judge[8]. Therefore, it is very important to help the doctor to accurately locate the target position based on the visual auxiliary system.

This paper mainly uses different methods to achieve the calibration between image-assisted image coordinates and spatial coordinates. Finally, it judges the feasibility through experiments. This paper is organized as follows: the first part introduces the research situation of vascular interventional surgical robot; the second part shows our team's vascular interventional robot platform; the third part describes the method to realize the transformation of image coordinates and spatial coordinates; the fourth part verifies the coordinate transformation through experiments. The feasibility and pros and cons of the method of judging the accuracy of the method; the fifth part summarizes the full text.

II. THE OVERVIEW OF PLATFORM

The vascular interventional operation platform of our team is divided into a master side and a slave side, where the master includes a master manipulator, a master controller, and a display screen; the slave terminal includes a slave manipulator, a slave controller, and a camera. The entire system operation process: the manipulator controls the master manipulator, and the master manipulator collects the manipulator's axial displacement and radial rotation information. The collected information is processed by the master controller and the information is input to the slave controller. The slave controller controls the slave manipulator to control the catheter and the guidewire according to the received information[9]. During the operation, the slave manipulator can collect the resistance information about the catheter and the guidewire and send it to the slave controller of processing, and then feed it back to the master. The master controller controls the force feedback damper on the mastrmanipulator to provide tactile feedback on the doctor based on the received feedback information[10]. During the operation, the slave side is also equipped with a camera to monitor the scene in real-time. The manipulator can observe the actual operation of the catheter and guidewire through the display screen on the master side, providing visual feedback to the doctor[11].

Vascular interventional surgery is surgery with very high accuracy requirements[12]. However, in terms of visual auxiliary on existing platforms, doctors only rely the rich experience to grasp the progress of surgery through images. Therefore, based on the team's visual auxiliary platform, I have established the correspondence between image coordinates and spatial coordinates, which can help doctors accurately grasp the target position through data and achieve precise surgical operations[13].

Besides, the establishment of precise correspondence can also make medical images more accurate and assist control. Therefore, it is very necessary to establish the corresponding relationship between image coordinates and spatial coordinates on the existing equipment to enhance

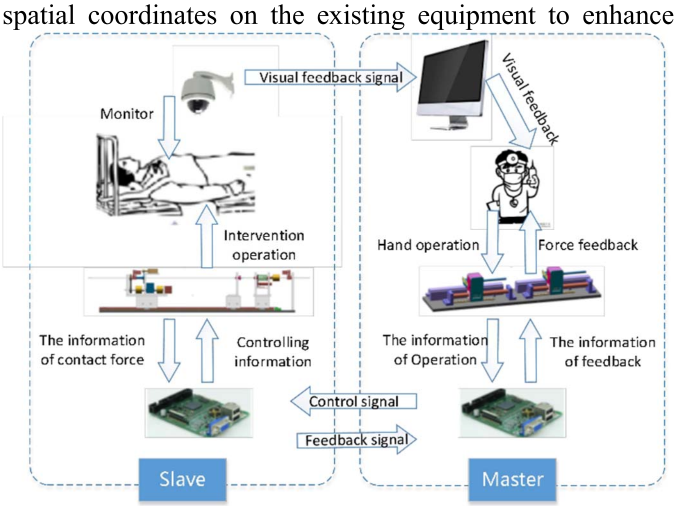

Fig 1 Concept map of vascular interventional surgical robot

the role of visual auxiliary. In the following paper, firstly we use the traditional coordinate calibration method for derivation, and then use the neural network method to optimize based on the traditional method, and finally verify it through experiments.

III. THE THEORY OF VISUAL AUXILIARY SYSTEM

The transformation of two-dimensional image coordinates and three-dimensional space coordinates generally involves four coordinate systems[14], as shown in Fig 2:

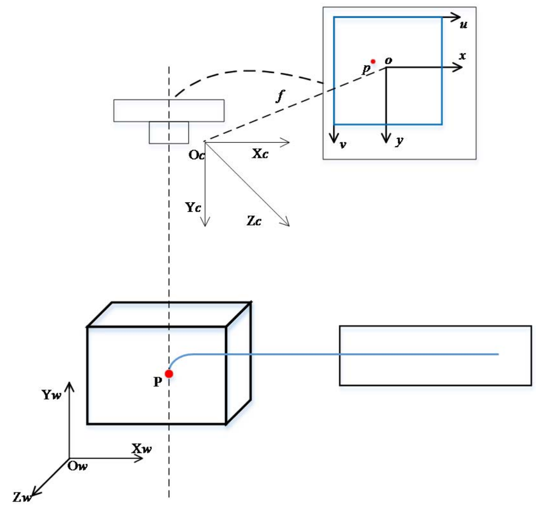

Fig. 2 Four coordinate systems of vascular interventional robot assist system

TABLE I
Explanation Table of Each Coordinate System

Coordinate System	Explanation
O_w-$X_wY_wZ_w$	World coordinate system, describing the target position
O_c-$X_cY_cZ_c$	Camera coordinate system with the light center as origin
o-xy	Image coordinate system, with the midpoint of the imaging plane as the origin
uv	Pixel coordinate system, the origin is the upper left corner of the image
P	A point in the world coordinate system is the position of the target in real coordinates
p	Point P imaging point in the image
f	Camera focal length, equal to the distance between o and O_c, $f=\|o-O_c\|$

A. Traditional coordinate transformation

In order to determine the correlation between the three-dimensional geometric position of a point in space and its corresponding point in the image, a mathematical model of the camera imaging must be established.

The first thing is to complete the conversion between the world coordinate system and the camera coordinate system. From the world coordinate system to the camera coordinate system is a rigid body transformation, the object will not be deformed[15], only rotation and translation (*RT*) are needed, as shown in formula (1)：

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} r11 & r12 & r13 & t1 \\ r21 & r22 & r23 & t2 \\ r31 & r32 & r33 & t3 \end{bmatrix} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (1)$$

$[X_c, Y_c, Z_c]$ represents the coordinates of the object in the camera coordinate system, $[X_w, Y_w, Z_w]$ represents the world coordinates where the object is located, R is the rotation matrix (3 degrees of freedom), T is the translation matrix, and the two form a 3 × 4 matrix.

Considering the ideal case, the correspondence between the camera coordinate system and the image coordinate system points can be seen as two sets of similar $\Delta ABO_C \sim \Delta oCO_C$ and $\Delta PBO_C \sim \Delta pCO_C$, as shown in Fig. 3.

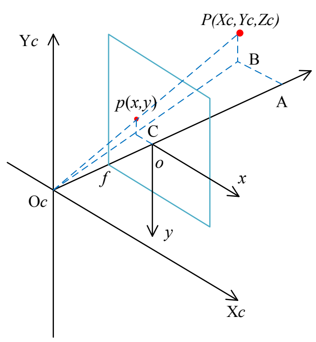

Fig. 3 Correspondence between camera coordinate system and image coordinate system

So the correspondence between (x,y) and $[X_c, Y_c, Z_c]$ is formula (2):

$$\begin{cases} x = f\dfrac{X_c}{Z_c} \\ y = f\dfrac{Y_c}{Z_c} \end{cases} \Rightarrow Z_C \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (2)$$

However, there are distortions in the camera imaging process, so the distortion parameters need to be taken during the calibration process. Distortion is generally divided into radial distortion and tangential distortion. Distortion is the change in distance, and the change in distance is reflected in the coordinate value is the relationship of addition and subtraction[16], so to get the specific distortion parameters to need to rely on equation (3):

$$\begin{cases} x' = X_c / Z_c \\ y' = Y_c / Z_c \\ r = x'^2 + y'^2 \\ x'' = x'(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1(x'y') + p_2(r + 2x'^2) \\ y'' = y'(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_2(x'y') + p_1(r + 2y'^2) \end{cases} \quad (3)$$

In the equation(3), $k_1 k_2 k_2$ are radial distortion coefficients, $p_1 p_2$ are tangential distortion coefficients. The corrected camera coordinate system coordinates at this time are:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} x''Z_c \\ y''Z_c \\ Z_c \end{bmatrix} \quad (4)$$

Finally, the image coordinate system coordinates are transformed into the pixel coordinate system. These two coordinate systems are in the same plane, but the representation unit and the coordinate origin are changed. The conversion relationship is shown in equation (5):

$$\begin{cases} u = \dfrac{x}{dx} + u0 \\ v = \dfrac{y}{dy} + v0 \end{cases} \Rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{dx} & \gamma & u0 \\ 0 & \dfrac{1}{dy} & v0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (5)$$

γ is the distortion factor, generally is 0[17].

So far the transformation from the world coordinate system to the pixel coordinate system is shown in equation (6):

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{dx} & \gamma & u0 \\ 0 & \dfrac{1}{dy} & v0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (6)$$

B. *Neural Network method*

Firstly, according to formula (6), we can simplify it to formula (7):

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = H^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (7)$$

According to the formula (7), the H represents the complex nonlinear mapping relationship between space coordinate system and the pixel coordinate system, and the neural network is suitable to reaction this nonlinear mapping relationship

In 1986, Rumelhart and McCelland and his team proposed a backpropagation learning algorithm, which realized the concept of Minsky's multi-layer perceptron[18]. Because multi-layer perceptrons are often trained with error back-propagation algorithms, they are often called BP neural networks.

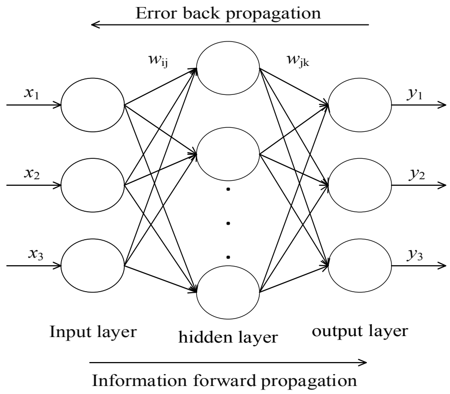

Fig. 4 BP neural network schematic

BP neural network has a multilayer measurement network structure, which is generally divided into layers: input layer, hidden layer, output layer. The neurons in each layer are fully connected, and there is no connection between the layers.

The training process of the BP neural network is divided into two stages: forward signal propagation and error back propagation. In the process of forward propagation, the input samples are propagated through all levels. If the actual output of the output layer does not match the expected output, the error is transferred into the backpropagation of the error. During the back-propagation process, the errors are distributed to each neuron in each layer, and the weights and thresholds are modified according to the errors. These two processes are repeated alternately until convergence.

Assume that the input layer has I neurons and the input vector $X=[x_1,x_2...,x_i]^T$; the hidden layer has J neurons; the output layer has K neurons ,output vector $Y=[y_1,y_2...,y_k]^T$; w_{ij} is the weight between the input layer and the hidden layer, w_{jk} is the weight between hidden layer and output layer; the threshold of neurons in hidden layer is $a_j(j=1,2,...,J)$; the threshold of neurons in output layer is $b_k(k=1,2...,K)$;the mathematical expression of the process is formula (8):

$$y_k=\sum_{j=1}^{J} w_{jk} f\left(\sum_{i=1}^{I} w_{ij} x_i - a_j\right) - b_k \tag{8}$$

Due to the construction and running algorithm of the Bp neural network, BP neural network can theoretically realize the fitting of the connection between any undetermined output vector and input vector. Therefore, the complex non-linear correspondence between the pixel coordinate system and the world coordinate system can be converted to each other and this relationship can be fitted through the BP neural network[19].

After building the neural network through MATLAB, we will bring a large amount of coordinate corresponding data that has been determined into the neural network, continuously iterate and optimize the parameter training model, and until the error reaches the set value, we fix the training model. Its processes as shown in Fig5. We bring the three-dimensional coordinates into the fixed training model to obtain the two-dimensional coordinates of the image and verify the accuracy of the method through experiments.

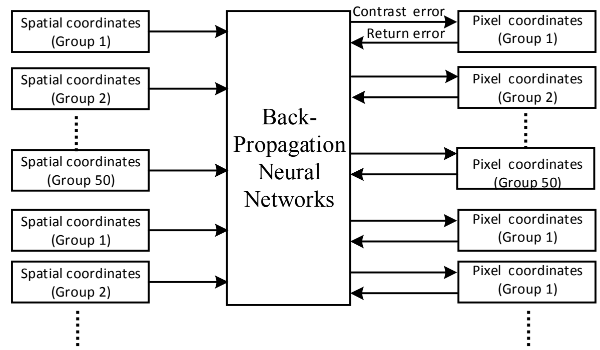

Fig. 5 Training process schematic diagram

IV. Experiments And Results

A. The platform of experiment

First of all, we must first take three-dimensional coordinates and their corresponding image coordinates to obtain their accurate positions, as the source of training data for the neural network and the basis for judging the accuracy of the method.

We use the Polaris Spectra position tracking system from NDI Canada to collect spatial coordinates. The device can accurately record the position and orientation information of a specific marking tool within a specific range. For image coordinate acquisition, we use MATLAB to calculate the pixel coordinates of the marked points by using its "impixelinfo" function. We will operate the catheter to a suitable position, and while collecting the coordinates of the catheter end on the image, use NDI to acquire its corresponding coordinates in

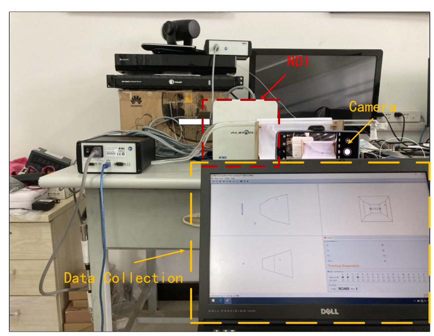

Fig. 6 Spatial coordinate collection


Fig. 7 Pixel coordinate collection

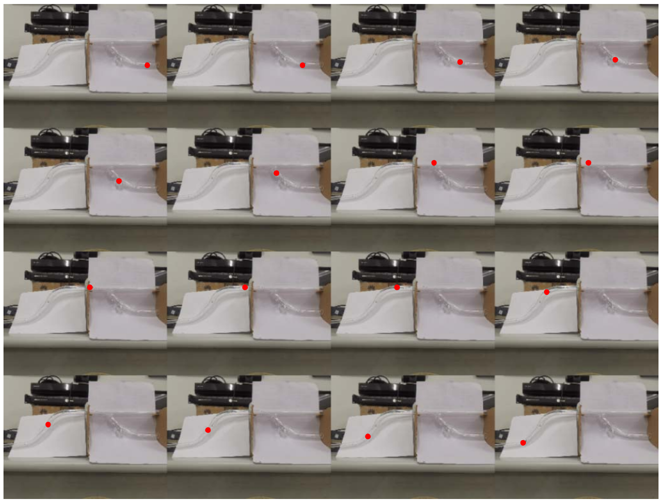

Fig. 8 Image collection of catheters at different positions

space. This process is repeated many times to obtain enough sets of data., and we took 59 sets of data. 50 of them as training sets, 9 of them as test sets, and We use them for cross-validation. The process is shown in Fig.6, Fig.7 and Fig.8.

B. The process of experiment

Firstly, we set the input of the neural network to spatial coordinates and the output to pixel coordinates. After setting the number of hidden layer nodes according to the empirical formula, after 667 rounds of iterations as shown in Fig9, the mean square deviation of the test sample can reach 10^{-5} levels of accuracy. The additionally acquired test space coordinates are brought into the obtained training model, and compared with the original trajectory, as shown in Fig11. The red line represents the two-dimensional coordinate path extracted from the image coordinates, and the blue dots represent the positions where the spatial coordinates of the test set are converted into image coordinates by the neural network.

Then do the same operation, we set the input to pixel coordinates, and the output to space coordinates, as shown in Fig10, and train the model after 1210 rounds. Since the three-dimensional space is not intuitive, we separate each axis to compare the fitted coordinates with the original coordinates, as shown in Fig11. In the figure, the red line represents the path coordinates of the end of the catheter collected through NDI, and the blue dots represent the positions where the image coordinates of the test set are converted into spatial coordinates through the neural network.

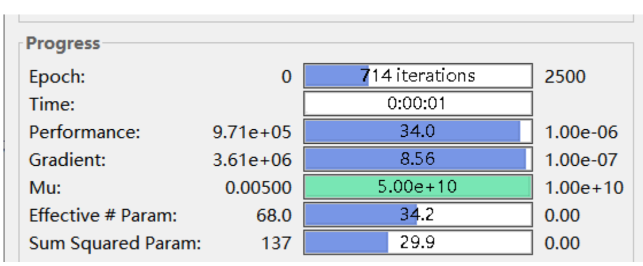

Fig. 9 Spatial to pixel coordinates

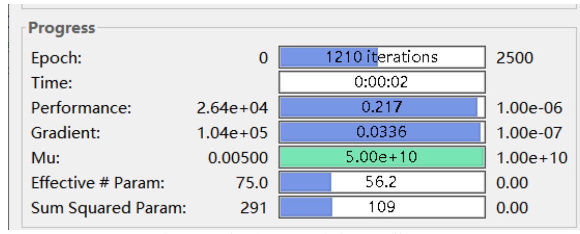

Fig. 10 Pixel to spatial coordinates

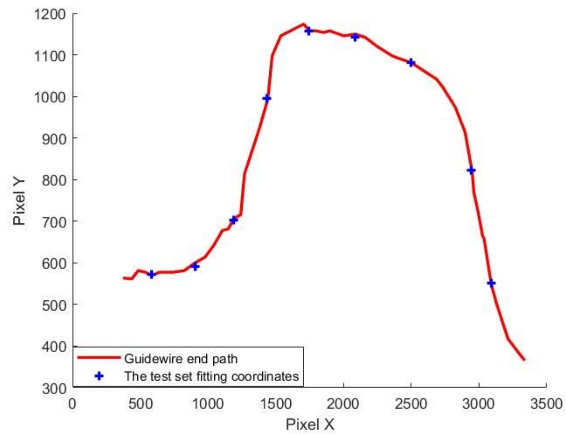

Fig. 11 Comparison of fitted pixel coordinates and the actual path

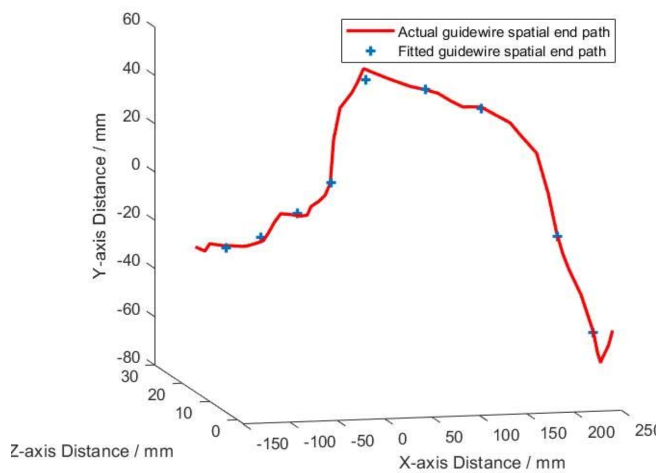

Fig. 12 Comparison of fitted spatial coordinates and actual coordinates

C. The results

After experiments, comparing the pixel coordinates fitted by the BP neural network with the actual coordinates, the average error of the x-axis is 4.81 pixels, and the average error of the y-axis is 4.47 pixels; the comparison of the fitted spatial coordinates and the actual space coordinates, the average error of the x-axis is 0.430mm, the average error of y-axis is 0.528mm, and the average error of z-axis is 0.507mm.

The experimental results prove that the BP neural network is suitable for the mutual conversion between the space coordinates and image coordinates in the visual auxiliary system of the vascular interventional robot, and can realize the accurate grasp of the catheter position.

V. Conclusion And Future Work

The paper first finds the problem of insufficient accuracy of coordinate correspondence in the visual auxiliary system of vascular interventional robot and improves the traditional coordinate calibration method through neural networks to solve this problem. And through the experiment, the fitting coordinates generated by the BP neural network are compared with the actual coordinates, and it is verified that the accuracy of the BP neural network method can complete the coordinate conversion. Experimental results show that the method has high accuracy, which verifies the correctness and accuracy of the BP neural network algorithm calibration, and it is easy and fast to obtain the parameters. This algorithm can be applied to the visual auxiliary system of vascular interventional surgical robot, which can not only improve the image accuracy but also greatly improve the doctor's ability to accurately grasp the surgical process through the image.

In future work, we will apply this algorithm to the operation interface of the vascular robot to help doctors achieve accurate operations.

Acknowledgments

This research is supported by National Natural Science Foundation of China (61703305), Key Research Program of the Natural Science Foundation of Tianjin (18JCZDJC38500) and Innovative Cooperation Project of Tianjin Scientific and Technological (18PTZWHZ00090).

References

[1] L.Ma, Y.Wu , W.Chen, "Chinese Cardiovascular Disease Report 2018", Chinese Journal of Hypertension, vol. 27, no. 8 pp. 712-716, 2019.

[2] T.Hu, T.Tian, K.Wang , L.Li, “Comparative study on the effect of craniotomy and intravascular interventional aneurysm embolization on the rupture of high-risk intracranial and eurysms in young and middle-aged people”, Practical Journal of Cardio-Cerebro-Pulmonary Vascular Disease, vol. 23, no. 7, pp. 81-84, 2015.

[3] X.Qian, “Observation on the effect of endovascular interventional therapy in patients with failed hemostasis under endoscope”, Chinese Health Standard Management, vol. 9, no. 11, pp. 52-54, 2018.

[4] Y.Wang, S.Guo, T.Tamiya, et al. A virtual-reality simulator and force sensation combined catheter operation training system and its preliminary evaluation[J]. International Journal of Medical Robotics and Computer Assisted Surgery, 13(3), 2017.

[5] Y.Song, S.Guo, X.Yin, “Performance evaluation of a robot-assisted catheter operating system with haptic feedback”, Biomedical Microdevices, vol. 20, no. 2, 2018.

[6] Smitson C. C., Ang L., Pourdjabbar A., “Safety and Feasibility of a Novel, Second-Generation Robotic Assisted System for Percutaneous Coronary Intervention: First in Human Report”, Journal of Invasive Cardiology, vol. 30, no.4, pp. 152-156, 2018.

[7] X.Bao, S.Guo, N.Xiao, “ A cooperation of catheters and guidewires-based novel remote-controlled vascular interventional robot”, Biomed Microdevices, vol. 20, no.1, pp.20, 2018.

[8] M.Deng, C.Lin, Z.Qin, C.Huang, “The value of three-dimensional digital subtraction angiography in the diagnosis of cerebrovascular diseases”, China Digital Medicine, vol.12, no. 12, pp. 41-43, 2017.

[9] C.Meng, J.Guo,S. Guo,Q.Zhan, L.Qi, “Design and Evaluation of a Novel Slave Manipulator for the Vascular Interventional Robotic System” , Proceedings of 2019 IEEE International Conference on Mechatronics and Automation, pp.1350-1355, August 4.7,Tianjin, China, 2019.

[10]S.Guo, Q.Zhan, J.Guo, C.Meng, L.Qi, “A Novel Vascular Interventional Surgeon Training System with Cooperation between Catheter and Guidewire” ,Proceedings of 2019 IEEE International Conference on Mechatronics and Automation, pp. 1403-1408, August 4-7, Tianjin, China, 2019.

[11]J.Guo, X.Jin, S. Guo, “Study of the Operational Safety of a Vascular Interventional Surgical Robotic System”, Micromachines, 9(3):119, 2018.

[12]J.Guo, Y.Cheng, S.Guo, W.Du, “A novel path planning algorithm for the vascular interventionalsurgical robotic doctor training system”, Proceedings of 2017 IEEE International Conference on Mechatronics and Automation, pp. 45-50,2017.

[13]K.Chen, “Research on key technologies of X-ray image guidance in vascular interventional surgery”, University of Chinese Academy of Sciences (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), 2019.

[14]L.Chen.“Research on Key Technologies of Image Processing in Assisted Vision ”, Xi'an University of Electronic Technology, 2016.

[15]J.Zhen, X.Cheng, X.Liu Source, “Templateless quasi-rigid shape modeling with implicit loop-closure”, Proceedings of 2013 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 145-152, 2013.

[16]Y.Zhu, S.Li, “Lens distortion and its correction technology ”, Optical Technology, no. 1, pp.136-138, 2015.

[17]X.Tong , “Research on Nonlinear Epipolar Geometry Based on RTcam ”, Xidian University, 2014.

[18]Rumelhart D, Mcclelland J, “Learning and Relearning in Boltzmann Machines”, Parallel distributed processing: explorations in the microstructure of cognition, vol.1, 1986.

[19]Y.Li, “Research Analysis and Application of BP Neural Network”, Anhui University of Science and Technology, 2016.

[20]D.Gao, P.Wang, Z.Cai, “Optimization of hidden layer nodes and training times in artificial neural networks ”, Journal of Harbin Institute of Technology, vol.2, pp.80-82, 2016.

[21]Y.Wu, M.Zhao, “Neuron Model with Adjustable Activation Function and Its Supervised Learning and Application”, SCIENCE CHINA: Technical Sciences, vol. 31, no. 3, pp. 263-272, 2001.