Robotica (2026), **44**, pp. 684–697
doi:10.1017/S0263574726103270

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

An S-SOM method based on binocular vision for configuration detection of concentric cable-driven manipulators

Zhonghui Wei[1], Weitao Li[1], Jie Li[2], Naijun Zhang[3], Yuxia Li[2] and Zonggao Mu[1,4,5]

[1]School of Mechanical Engineering, Shandong University of Technology, Zibo, China
[2]School of Artificial Intelligence, Shenzhen Polytechnic University, Shenzhen, China
[3]Shandong Jite Industrial Technology Company Limited, China
[4]Institute of Modern Agricultural Equipment, Shandong University of Technology, China
[5]Zibo Key Laboratory of Collaborative Robots, AUBO(Shandong) Intelligent Robot Co, China
Corresponding author: Zonggao Mu; Email: muzonggaook@163.com

Received: 22 May 2025; **Revised:** 21 October 2025; **Accepted:** 11 December 2025; **First published online:** 27 February 2026

Keywords: concentric cable-driven manipulators; the binocular vision; the S-SOM method; PCA

Abstract

The concentric cable-driven manipulator (CCDM) is slender enough to work in narrow oral environments. However, the configuration detection of the CCDM with two segments is the key issue to realize the servo control and force sensing. In this paper, a Segmented Self-Organizing Map (S-SOM) method is presented for configuration detection of the CCDM without preset markers. Firstly, two-dimensional point cloud of the CCDM is acquired by the binocular vision. Secondly, the dimension reduction of point cloud is realized by principal component analysis to avoid noises and eliminate redundant information. Then, the threshold is set to distinguish the segment with different diameters that need to be trained independently. Thirdly, quantization errors and topological errors are designed to select different training parameters of the CCDM with two segments. Finally, the experimental results show that the average point-to-point distance error is less than 2.57 mm. The effectiveness of the proposed S-SOM method for configuration detection of the CCDM is verified.

1. Introduction

With the rapid progress of science and technology, robot technology has gradually penetrated all aspects of life. The rigid robot [1] is often difficult to complete the expected task in the complex working space with the continuous deepening and changes of application scenarios [2, 3]. So flexible robots play an important role in the medical field. Due to its flexibility, flexible robots can minimally enter the confined space of the human body that traditional surgical instruments cannot reach [4, 5]. This makes the CCDM show great potential in oral diagnosis and treatment. The CCDM adopts a concentric tube structure design, which can still maintain flexible work in a small space. Compared with traditional industrial robots, the CCDM can be continuously bent along the central skeleton line. It also has the ability to carry out flexible deformation in space. The multiple degrees of freedom give it a significant advantage in terms of operational flexibility. However, because the CCDM is a strong-coupling, multiple-input-multiple-output, nonlinear system, its servo control and force sensing is difficult to be achieved without suitable configuration detection methods [6, 7].

Many researchers have carried out various researches on manipulators in medical treatment fields. A novel flexible shape memory alloy wire-based gripper is developed by Zhou [8]. This design improves the stability and motion accuracy of the target clamp. Li [9] designs a method to control the length and curvature of the bending section of the flexible robot for better application in minimally invasive surgery. However, the operation of some robots is overly dependent on manual control, which may affect

© The Author(s) 2026. Published by Cambridge University Press.

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

the execution of tasks. Some researchers have found that the installation of a vision system on the CCDM can better assist operations. With the continuous progress of vision technology, the binocular vision has received great attention in the field of medical robot surgery [10, 11]. Rivoir [12] improves a method that uses the binocular vision to predict the completion time of surgery and provide necessary assistance to doctors. Lei [13] establishes a dual-arm robotic needle insertion system with the guidance of intraoperative two-dimensional ultrasound and preoperative three-dimensional computed tomography fusion. This enabled surgeons to achieve precise spatial positioning and direction adjustment for bone grafting. The experimental results show that the average puncture error is 2.48 mm, which meets the basic working requirements. Although the aforementioned methods apply the binocular vision to flexible manipulators, they cannot obtain configurations of flexible manipulators and have limitations in surgical operations. Wang [14] combined computer image navigation with robot technology and proposed an improved iterative closest point algorithm and a new hand–eye calibration method for multi-arm robots and free bone block spatial registration to assist in mandible reconstruction surgery. Zhao [15] proposed a rigid-flexible composite puncture (RFCP) strategy and developed a compact 7-DOF venipuncture robot to assist with venous blood collection. The system employs a differential evolution algorithm for its inverse kinematics solution and quintic polynomial interpolation for trajectory planning. Experiments on the prototype verified the correctness of the robot's kinematics and trajectory planning, as well as the feasibility of the RFCP strategy in preventing double-vessel-wall penetration during blood collection.

In practical applications of servo control and force sensing, it is necessary to determine the configuration of the CCDM. In order to obtain an effective three-dimensional configuration, William [16] develops a dynamic model for continuum robots based on the principle of virtual work. The configuration of manipulator obtained in this way requires the establishment of complex mechanical models. The model takes into account the effects of inertia, driving forces, friction, elasticity, and gravity, allowing for more accurate acquisition of the manipulator's three-dimensional information. This method is to measure the configuration of the CCDM by establishing a mechanical model of manipulators [17], while the model will increase the complexity of the operation and the amount of calculation.

Many researchers have adopted different strategies to realize the configuration detection of manipulators. Among them, the shape-based detection method is an effective detection scheme [18]. Hannan uses a camera to detect the shape of the robotic arm [19]. This method is to obtain the center point coordinates of the cross-section of the robot arm through the camera and then get the complete configuration of the robot arm. However, the disadvantage of this method is that when the manipulator moves to the place where the camera cannot obtain the center point of the section, the camera cannot obtain the corresponding coordinate point. Gao [20] proposes a dynamic convergence adjustment algorithm for improving visual effects. This method provides depth perception by rectifying the imaging plane. In order to avoid the collision between the manipulator and the abdominal cavity, Li [21] proposes a method to solve the remote perception of the manipulator shape. In this method, the error compensation of the manipulator is realized by using markers. This strategy solves the occlusion problem caused by the movement of the manipulator.

However, the marker is not suitable for mounting on the manipulator in some confined or multi-occlusion applications. Moreover, the easy drop of the markers will also affect the practical application of the existing detection method. The traditional SOM method can sense the shape of a soft manipulator without the aid of markers [22]. However, the traditional SOM method mainly applies to the configuration detection of manipulators with a constant diameter. The diameter of the manipulator with single segment is relatively uniform, and its two-dimensional point cloud distribution in the image is relatively uniform. In this case, the central skeleton line of the manipulator can be obtained by using the traditional SOM method to cluster two-dimensional point cloud. As shown in Figure 1, the object of configuration detection in this paper is the CCDM with two segments. The CCDM consists of an inner cable-driven mechanism and an outer cable-driven mechanism.

Because diameters of the inner cable-driven mechanism and the outer cable-driven mechanism are different, the traditional SOM method will cause the clustering results to move to the part of the manipulator with the larger diameter, which will affect the accuracy of the configuration detection. Therefore,

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

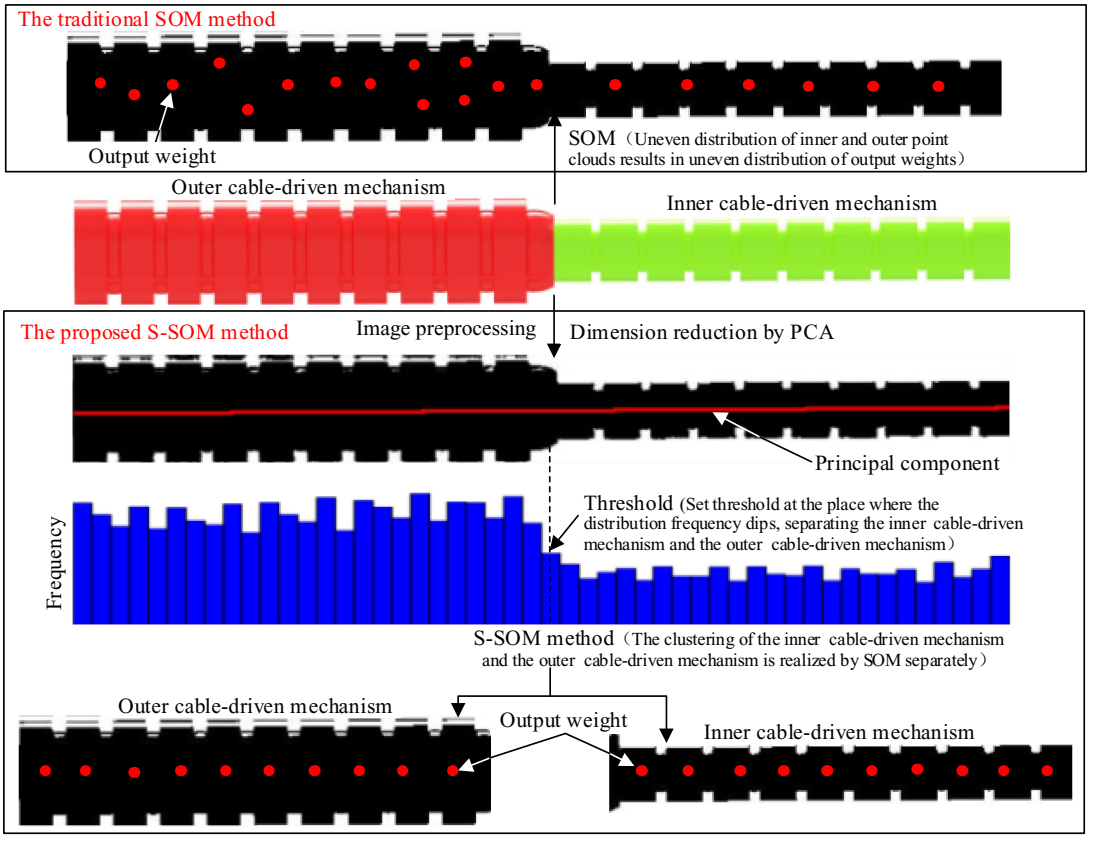


Figure 1. *Modifications of the S-SOM method over the SOM method for the CCDM with two segments.*

a Segmented Self-Organizing Map (S-SOM) method is proposed to detect the configuration of the CCDM with two segments with different diameters.

The "segmented" means that the "Self-Organizing Map" is trained separately for the inner cable-driven mechanism and the outer cable-driven mechanism according to the result of PCA. In order to represent more indicators with the least number of features, the dimension of the point cloud needs to be reduced. The dimension reduction allows the input point cloud to be compressed while retaining most of the original input point cloud. The PCA ensures that the original input point cloud can be recovered using the proposed method. Quantization errors and topological errors are used to determine the selection of various training parameters. Based on the structure of the CCDM with two segments, the S-SOM method is proposed and verified.

2. The S-SOM method for configuration detection of concentric cable-driven manipulators with two segments

The CCDM comprises an inner cable-driven mechanism, an outer cable-driven mechanism, and a restrained tube. During the movement, the outer cable-driven mechanism constrains the overlapped segment of the inner cable-driven mechanism to match its bending shape. This constraint enables the concentric manipulator to attain different configurations, such as "C-shaped" or "S-shaped." As shown in Figure 2, the joints between the inner and outer cable-driven mechanisms are connected by concentric ball pairs. The CCDM is composed of independent joints. The joint of the CCDM is an axisymmetric structure, with the axis located at the center of the central cylindrical hole. The center point of the joint is located on the axis.

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

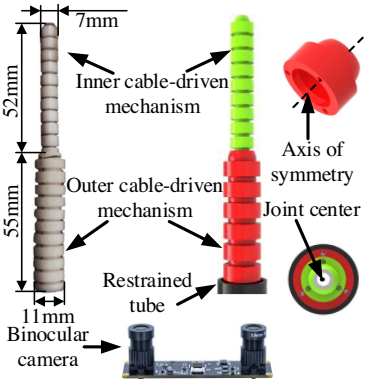


Figure 2. *Structure of the concentric cable-driven mechanism.*

The inner cable-driven mechanism can move forward and backward freely through the outer cable-driven mechanism during operations. The length of the inner cable-driven mechanism is 52 mm with a maximum diameter of 7 mm. The length of the outer cable-driven mechanism is 55 mm with a maximum diameter of 11 mm. A central elastic skeleton runs through the inner cable-driven mechanism. The total length of the CCDM is 107 mm.

2.1. The dimension reduction by PCA

The dimension reduction by PCA is to map the input data of point cloud with n-dimensional features to the output data with m-dimensional features. The m-dimensional features which are regarded as principal components are completely new orthogonal features. The principal components are reconstructed based on the original n-dimensional features of the input data through a series of transformations to achieve the purpose of dimension reduction.

The specific dimension reduction process by PCA is as follows. First, the input data of point cloud are processed centrally. Second, the covariance matrix $\boldsymbol{XX}^{\mathrm{T}}$ of the input data is calculated, and the covariance matrix is decomposed by eigenvalue. Taking the left image of the CCDM captured by the binocular camera as an example, the coordinate value of the black pixel is reduced. Since pixels are two-dimensional information in the image plane, for a pixel point (u_i, v_i), the covariance calculation of the input data U and input data V is as follows.

$$\begin{aligned} Cov\,(U,V) &= E\left[(U-E(U))(V-E(V))\right] \\ &= \frac{1}{n-1}\sum_{i=1}^{n}(u_i-\overline{u})(v_i-\overline{v}) \end{aligned} \tag{1}$$

where $E(U)$ and $E(V)$ represent mathematical expectations. $\overline{u}$ and $\overline{v}$ represent average values.

For the two-dimensional data of point cloud, the calculation process of the covariance matrix $\boldsymbol{XX}^{\mathrm{T}}$ can be specified as:

$$\boldsymbol{XX}^{\mathrm{T}} = \begin{pmatrix} Cov(x_1,x_1) & Cov(x_1,x_2) \\ Cov(x_2,x_1) & Cov(x_2,x_2) \end{pmatrix} \tag{2}$$

Through the above steps, the covariance matrix $\boldsymbol{XX}^{\mathrm{T}}$ of the input data of point cloud is shown in Figure 3. The pixel coordinate information of the CCDM is reduced from two dimensions to one dimension. Its retained principal component is [0.0197, −0.9999].

The explained variance ratio EVR_i can be used to measure how much information a principal component retains in the input data of point cloud. It is used to quantify the degree to which a principal

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

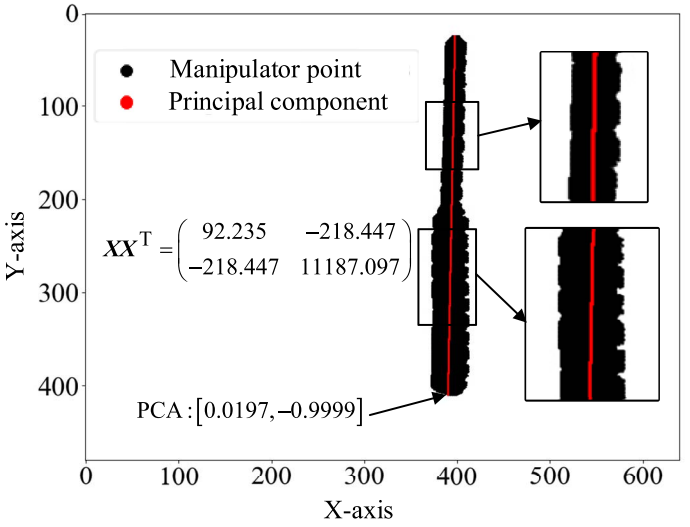


Figure 3. *Results of the dimension reduction by PCA.*

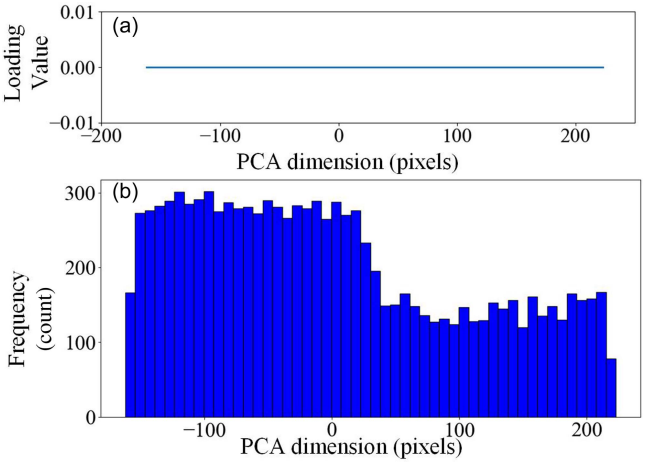


Figure 4. *The dimension reduction by PCA. (a) The point cloud reduced to one-dimensional data by PCA (b) Frequency histogram.*

component explains the variance of the input data of point cloud. The calculation method is:

$$EVR_i = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j} \tag{3}$$

where λ_i is the ith eigenvalue. n is the amount of the input data.

The variance interpretation ratio of the principal component is 0.992. The greater the proportion of variance explained by the principal component, the more variation that the principal component can explain in the input data of point cloud.

Results of the point cloud reduced to one-dimensional data by PCA are shown in Figure 4(a). PCA converts each black pixel point in the original image from a two-dimensional pixel plane coordinate to a one-dimensional coordinate. Because the point cloud is reduced to one dimension, the ordinate in Figure 4(a) no longer has practical physical significance. So, the value of the ordinate is set to 0. The frequency histogram result is shown in Figure 4(b). The horizontal coordinate represents the projected value of the point cloud on the principal component by PCA. The ordinate represents the frequency of

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

each projected value on the principal component. The frequency histogram can more intuitively show the distribution characteristics of point cloud in the direction of principal components.

According to the analysis of histogram results, it can be intuitively seen that the frequency of the CCDM changes obviously at the connection between the inner cable-driven mechanism and the outer cable-driven mechanism. The outer cable-driven mechanism should have more input data than the inner cable-driven mechanism. However, the inhomogeneity of the point cloud will result in the formation of superior neurons more easily in the data of the outer cable-driven mechanism.

Therefore, the weight vector of the outer cable-driven mechanism will be updated frequently, while the update frequency of the weight vector of the inner cable-driven mechanism is lower than that of the outer cable-driven mechanism. This will lead to the final clustering results more inclined to the outer cable-driven mechanism and then affect the detection of the configuration of the inner cable-driven mechanism. Therefore, in this paper, the point cloud of the CCDM is trained by the S-SOM method by PCA. At the place where the intermediate frequency value drops sharply, the threshold is set to distinguish the segment with different diameters that need to be trained independently. At the end of the training by the S-SOM method, the manipulator with two segments with different diameters trained independently is combined into a complete the CCDM.

2.2. Training parameters of the S-SOM method

The different selection of training parameters directly affects the training results of the S-SOM method. Training parameters include number of weight vectors, initial learning rate, learning rate function, winner neighborhoods, and number of training epochs.

The weight vector of the S-SOM method is initialized. The weight vector initialization mode is random initialization. The random initialization is to assign the random number to the initial weight vector so that the weight vector can be distributed in the whole point cloud. However, setting the position of the initial weight vector in this way makes it difficult for weight vectors far away from the input data of point cloud to be adjusted by the S-SOM method, while the weight vector in the center of point cloud is adjusted several times. This further affects the effect of clustering.

Another way to initialize weight vectors is linear initialization. The linear initialization is achieved by initializing the weight vector to a straight line. Compared with the random initialization, the linear initialization can avoid the problem that the weight vector is difficult to adjust. If there is no information about the point cloud, the random initialization is usually chosen. According to the prior knowledge, the configuration information of the CCDM is closer to the linear structure. Therefore, this paper chooses linear initialization.

For the S-SOM method, the selection of each training parameter will greatly affect the effect of training. Faced with many training parameters of the S-SOM method, the training parameters are usually adjusted blindly according to the clustering results. This will also affect the performance of the S-SOM method and the effect of training. Therefore, this paper uses topological errors and quantization errors to measure the performance of different training parameters.

The proportion of the input data of point cloud containing the wrong topology to the total input data of point cloud is defined as the topological error ε_{te}, that is:

$$\varepsilon_{te} = \frac{1}{N}\sum_{i=1}^{N}\mu(\hat{P}_i) \tag{4}$$

where $\hat{P}_i$ is the ith input vector. N indicates the total number of input data of point cloud. $\sum_{i=1}^{N}\mu(\hat{P}_i)$ indicates the input data of point cloud containing the incorrect topology.

For input data of point cloud $\hat{P}_i$, find the winning neurons and the second winning neurons. They are judged to be adjacent in the linear array structure of the S-SOM method. If they are adjacent, there is no wrong topology, and $\mu(\hat{P}_i)$ of the input data of point cloud is equal to 0. On the contrary, if they are not

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

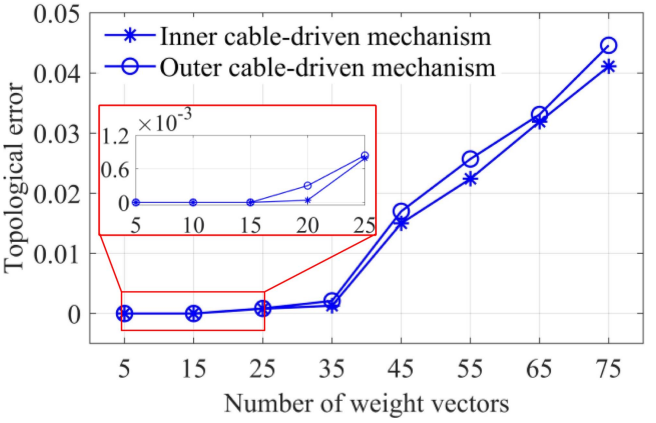


Figure 5. *Topological errors corresponding to different numbers of weight vectors.*

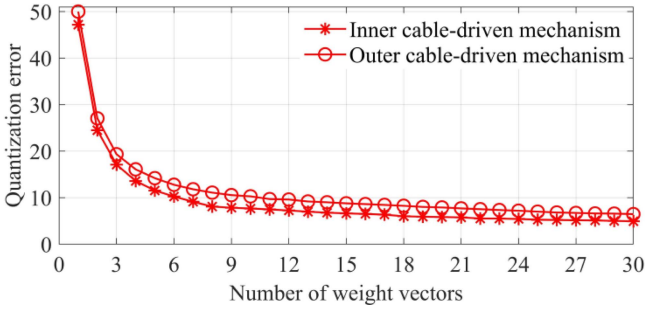


Figure 6. *Quantization errors corresponding to different numbers of weight vectors.*

adjacent in the linear array structure, the topological relationship of the winning neurons representing the output data is not related. In this case, $\mu(\hat{P}_i)$ of the input data of point cloud is equal to 1.

The average Euclidean distance of each input mode vector $\hat{P}_{ij}$ and its corresponding weight vector $\hat{W}_j$ is defined as the quantization error ε_{qe} of the S-SOM method.

$$\varepsilon_{qe} = \frac{1}{N}\sum_{i=1}^{n_i}\sum_{j=1}^{m}\left\|\hat{P}_{ij} - \hat{W}_j\right\| \tag{5}$$

where N is the total number of input data of point cloud. n_i is the number of vectors.

With the continuous learning and iteration of the S-SOM method, the variation trend of quantization errors must be smaller and smaller. It can be concluded that the significance of quantization errors is to reflect the clustering effect of the CCDM. When the quantization error is smaller, the distance between the input vector and its corresponding weight vector is closer. The result of the S-SOM method can better represent the configuration information of the CCDM.

The S-SOM method is trained with different weights. To find the appropriate number of weights, a large interval is selected to calculate the corresponding topological errors. Then, according to the results, the interval is narrowed and the experiment is carried out again. The inner and outer cable-driven mechanism of the CCDM is trained independently. To ensure the consistency of the results as much as possible, the same training parameters are used for simultaneous training, as shown in Figure 5.

Through the analysis of error results, the topological error is small when the number of weight vectors of the inner and outer cable-driven mechanisms is between 5 and 25. As the number of weight vectors increases, the quantization error becomes smaller and smaller. Therefore, the quantization error corresponding to the quantities of different weights are calculated in the range (1–30) to further determine the appropriate training parameters. The quantization error results are shown in Figure 6.

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

In the selection of the number of weight vectors, it is necessary to synthesize the results of topological errors and quantization errors. When the number of weight vectors is small, the topological error is 0, but the quantization error is large. It is difficult to reflect the information of the CCDM. When the number of weight vectors is larger, the quantization error is smaller, but the topological error is larger. Too many weight vectors will cause the weight vectors distributed along the diameter direction of the manipulator when training the point cloud of the outer cable-driven mechanism. It is not in line with the expected objective of the experiment. To better determine the selection of the number of weight vectors of the inner and outer cable-driven mechanisms, many influencing factors are considered. When it is determined to train the inner cable-driven mechanism, the number of weight vectors is selected to be 8. When training the outer cable-driven mechanism, the number of weight vectors is selected to be 12.

When selecting the learning rate function, it is necessary to meet the change law that the learning rate gradually decreases with the increase of the number of iterations. Therefore, in the selection process, the appropriate monotone decreasing function is selected as the learning rate function. The learning rate function selected in this paper is:

$$\eta(t) = 10 \cdot \eta_0 \times \left(1 + \frac{t}{\tau}\right) \tag{6}$$

where η_0 is the initial learning rate. t is the current number of iterations. τ is a time constant, and $\tau = T/2$. T is the total number of iterations.

The design of the winner neighborhood is the same as the learning rate function. It should also continue to decay with the increase of training epochs, showing a decreasing trend in general. Therefore, the formula (6) is also used for the design of the winner neighborhood, so that the S-SOM method can gradually converge with the continuous training process. This design can effectively ensure that the winner neighborhood is large at the initial stage of training, allowing multiple nodes in the S-SOM method to adapt to the input signal simultaneously. This can speed up the convergence of the network in the initial training. As the training process continues, the winner neighborhood decreases to only a few or a single node, thus fine-tuning the topology of the network.

In addition, the design of the winner neighborhood and the determination of the learning rate also affect the training effect of the S-SOM method. Therefore, the influence of different sizes of winner neighborhoods and learning rates on the S-SOM method is measured by quantization errors. Three sets of simulations are carried out for the inner and outer cable-driven mechanisms, respectively, in which the size of the winner neighborhood is 2, 5, and 7. In the simulation result analysis of the S-SOM method, the solid line is used to represent the simulation result of the inner cable-driven mechanism, and the dashed line is used to represent the simulation result of the outer cable-driven mechanism. The analysis of simulation results is shown in Figure 7.

According to the analysis of simulation results, it can be seen that both a large learning rate and a small learning rate will lead to the accuracy of the clustering results of the CCDM. A large learning rate will increase the difficulty of training and produce a large quantization error. However, a small learning rate will cause the input data of point cloud to fall into the local learning situation, thus increasing the quantization error.

According to the analysis of simulation results, the learning rate in this paper is set to 0.15. By adjusting the size of the winner neighborhood, it can be concluded that the smaller the winner neighborhood, the smaller the quantization error. This is because larger winner neighborhoods cluster more input data of point cloud to the same point during training. This will affect the location distribution of clustering results, which is contrary to the original intention of obtaining the configuration of the CCDM. Therefore, according to the analysis of simulation results, the size of the winner neighborhood is set to 2. By determining the training parameters of the S-SOM method, the number of training epochs is selected by quantization errors and topological errors. The simulation results are shown in Figure 8.

It can be concluded from the simulation results that the topological error will decrease to 0 in fewer training epochs. Although the quantization error will decrease with the increase in the number of training epochs, the changing trend will be gradual with the increase in the number of training epochs. The

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

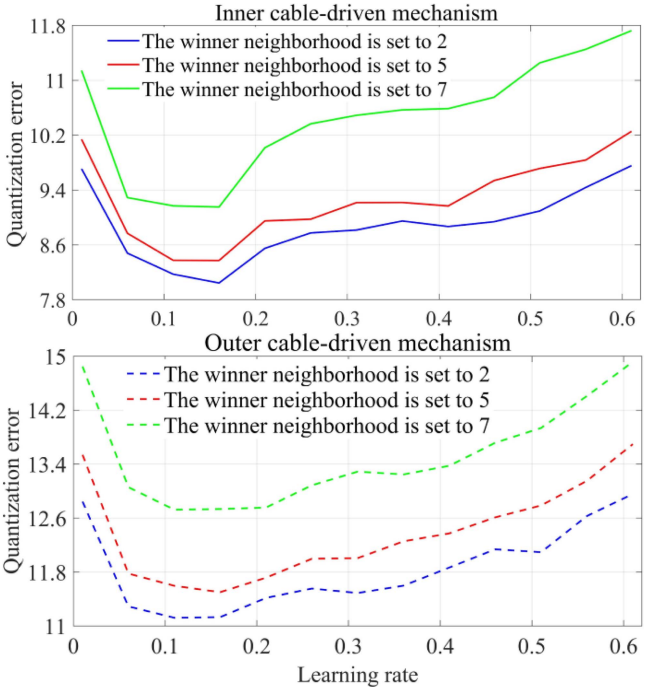


Figure 7. *Quantization errors corresponding to different learning rates and winner neighborhoods.*

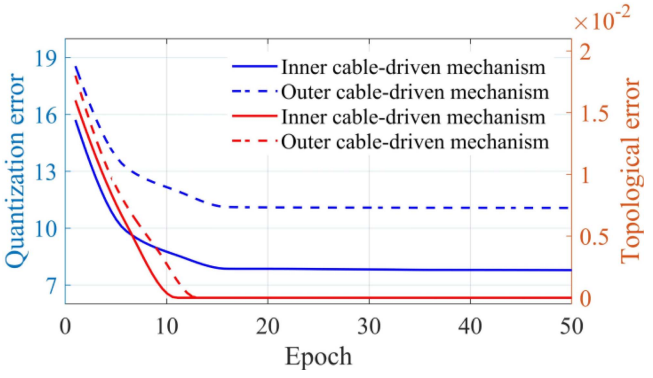


Figure 8. *Quantization errors and topological errors corresponding to different training epochs.*

amount of variation in quantization errors no longer changes significantly. Therefore, the choice of the final number of training epochs is 20.

In addition, since the left and right images of the CCDM are trained separately, the final training result may not reflect the position information of the same point in space well. In the Section 2 of this paper, the image has been stereo rectification. This means that the same point on the CCDM can match the projected pixel points in the left and right image, that is, the longitudinal value of the projected pixel points in the left and right image is equal. This will result in additional errors due to the training of the S-SOM method. This error is defined in this paper as a rectification error ε_{re}, which is used to measure the difference in the ordinate values of the corresponding pixel in the left and right images.

$$\varepsilon_{re} = \frac{1}{m}\sum_{i=1}^{m}\left|v_i^l - v_i^r\right| \tag{7}$$

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

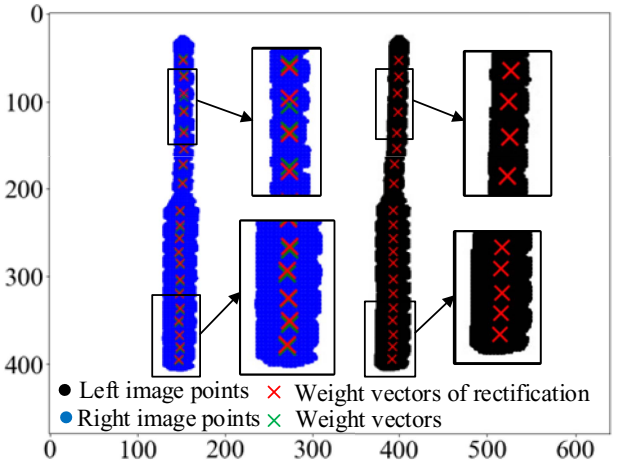


Figure 9. *Weight vectors of left and right images of the CCDM.*

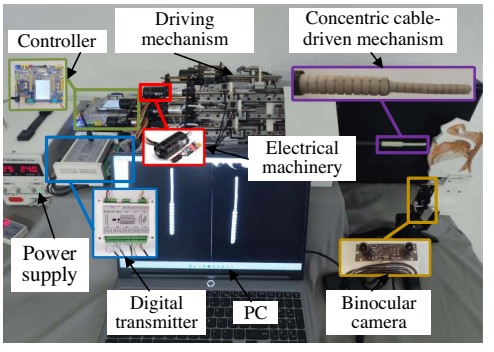


Figure 10. *Experimental system.*

where m is the number of output weight vectors. v_i^l is the ordinate value of the ith projection pixel in the left image. v_i^r is the ordinate value of the ith projected pixel in the right image.

To eliminate the error of matching, the clustering results of the right image are corrected based on the clustering points of the CCDM in the left image. Firstly, a central skeleton line representing the shape of the CCDM in each image is generated by the output weight vector trained by the S-SOM method. Secondly, the right image weight vector corresponding to the left image weight vector is found by way of a polar search. Thirdly, the correction step of the output weight vector of the CCDM is completed. The corrected weight vector is used as the input coordinates for the subsequent calculation of spatial coordinates. The weight vector of the left and right image of the CCDM before and after correction is shown in Figure 9.

3. Configuration detection experiment of the concentric cable-driven manipulator

3.1. Experimental system

To verify the feasibility of the S-SOM method for detecting the configuration of the CCDM, the CCDM prototype system is designed, as shown in Figure 10. The prototype system includes the CCDM, driving mechanism, controller, power supply, digital transmitter, binocular camera, and computer. The prototype system controls the motor to drive the corresponding cable to complete the configuration change of the CCDM. The contraction and bending of the CCDM are also realized.

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

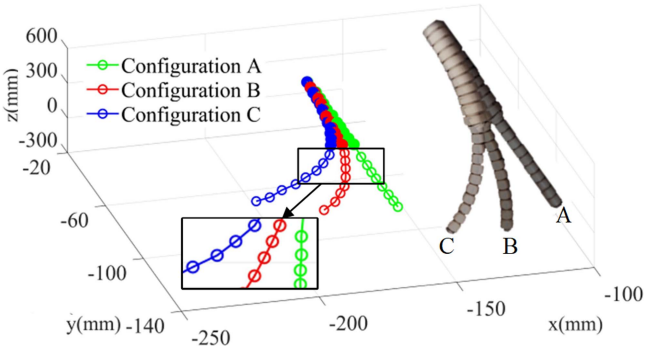


Figure 11. *Detection results of the three configurations.*

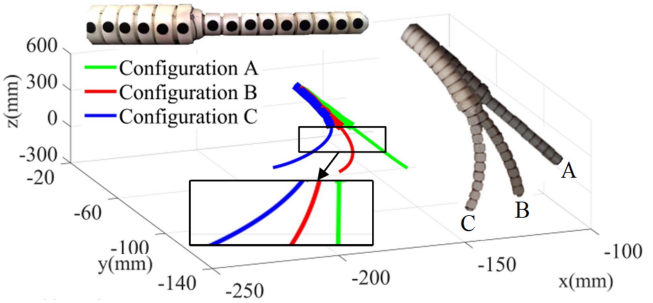


Figure 12. *Polynomial fitting of the curve obtained from the S-SOM method.*

The prototype system detects the configuration information of the CCDM by the binocular camera. The binocular camera is connected to a computer, and the configurations of the CCDM in different motion states are obtained from the computer.

The motor is controlled by the controller and the digital converter, and the cable is pulled to make the CCDM start to move. When the inner cable-driven mechanism and the outer cable-driven mechanism are driven by cables, the CCDM gradually bends. To verify the feasibility of the S-SOM method, a configuration detection experiment of the CCDM is designed. During the gradual bending movement of the CCDM, three different configurations are selected to detect. The experimental results are shown in Figure 11.

3.2. Analysis of experimental results

In order to process the experimental data of the CCDM obtained by the S-SOM method, the scatter plot is simulated into a curve by polynomial. As shown in Figure 12, the three smooth curves in the figure are respectively obtained by fitting the scatter plot drawn by the S-SOM method. The shape information of the curve reflects the configuration information of the CCDM.

The marker detection method is a traditional method used to detect the configuration of the CCDM. The marker detection method has quite high precision and can be used as a comparison of the S-SOM method in this paper. As shown in Figure 12, a black circular marker is pasted on each joint of the CCDM, and the center position of the marker is identified by the binocular vision, so as to detect the configuration of the CCDM.

Because the discrete point information obtained by the marker detection method corresponds to the joint of the CCDM. In other words, the joint points of ten inner and ten outer cable-driven mechanism joints are obtained. However, there are eight weight vectors for training of the inner cable-driven mechanism and twelve weight vectors for training of the outer cable-driven mechanism. Although the number of discrete points obtained by the two methods is the same, the corresponding position of the CCDM

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

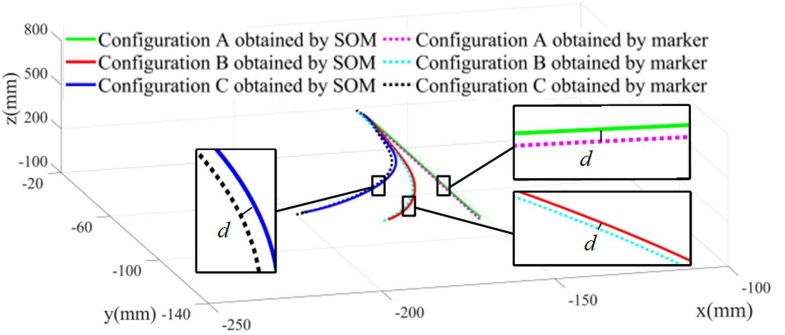


Figure 13. *Configuration curve of the CCDM conclusion.*

is not the same. Therefore, the distance difference of corresponding discrete points cannot be directly compared to measure the feasibility of the S-SOM method. Therefore, in this paper, 20 discrete points obtained by the S-SOM method and the marker detection method are fitted, respectively. Two curves describing the configuration information of the CCDM are obtained. The same number of point pairs are sampled at the upper distance between the two curves. The corresponding three-dimensional coordinate values are obtained. In this paper, the average point-to-point distance error is designed to quantify the accuracy of the CCDM configuration obtained by the S-SOM method. The calculation formula is as follows:

$$\varepsilon_{de} = \frac{1}{n}\sum_{i=1}^{n} d_i \tag{8}$$

where n indicates the number of pairs of sampling points. d_i is the distance between pairs of sampling points in group i.

The starting and ending points of curves obtained by the two methods are different. The curvature of the outer cable-driven mechanism is small, and the variation along the x-axis is small. Therefore, in this paper, the first set of point pairs is set at the position of the configuration curve $y = -30$ of each the CCDM. The remaining pairs are evenly arranged according to the length of the cable, totaling 20 pairs. The difference in the configuration curves of the CCDM obtained by the two methods is quantified by the average point-to-point distance error ε_{de}.

As shown in Figure 13, when the CCDM is at its initial configuration A, the curve obtained by using the S-SOM method is compared with that obtained by using markers. The average point-to-point distance error is $\varepsilon_{de}^{\mathrm{A}} = 1.52\mathrm{mm}$. When the CCDM is at its initial configuration B, the curve obtained by using the S-SOM method is compared with that obtained by using markers, and the average point-to-point distance error is $\varepsilon_{de}^{\mathrm{B}} = 2.34\mathrm{mm}$. When the CCDM is at its initial configuration C, the curve obtained by using the S-SOM method is compared with that obtained by using markers, and the average point-to-point distance error is $\varepsilon_{de}^{\mathrm{C}} = 2.57\mathrm{mm}$.

Conclusion

This paper proposes an S-SOM method for configuration detection of the CCDM without preset markers. Firstly, the images and two-dimensional point cloud of the CCDM are acquired by the binocular vision. Secondly, the dimension reduction of the input point cloud is processed by PCA to avoid noises and eliminate redundant information. Then, the threshold is set to train the CCDM with two segments independently. Thirdly, quantization errors and topological errors are designed to select different training parameters of the CCDM with two segments with different diameters. According to the simulation of the inner and outer cable-driven mechanisms, the appropriate training parameters are obtained, respectively. Finally, experiments are designed to verify the S-SOM method, where results of the marker detection method are used as the reference. Experimental results demonstrating an average point-to-point distance

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

error of less than 2.57 mm verify the effectiveness of the proposed S-SOM method for configuration detection of the CCDM. In the future, the detected configuration can be used to achieve the servo control and force sensing of the CCDM.

Data availability statement. The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Author contributions. Zhonghui Wei: Conceptualization, Writing original draft, and Resources. Weitao Li: Data curation, Writing review and editing. Jie Li: Writing review and Editing. Naijun Zhang: Writing review and Editing. Yuxia Li: Methodology & Supervision. Zonggao Mu: Writing review, Editing, and Funding acquisition. (Zhonghui Wei, Weitao Li, and Jie Li contributed equally to this work.)

Financial support. This work was supported in part by the Key R&D Program of Shandong Province, China under Grant 2024TSGC0751, in part by the Open Scientific Research Project of Institute of Modern Agricultural Equipment under Grant NZY-2025-12, in part by the Taishan Scholar Project of Shandong Province under Grant tsqn202312185, in part by the National Natural Science Foundation of China under Grant 62473238 and 61903100, in part by Guangdong Basic and Applied Basic Research Foundation under Grant 2023A1515110729, in part by Shenzhen Science and Technology Program under Grant 20231128093642002, in part by the Research Foundation of Shenzhen Polytechnic University under Grant 6023312007K.

Competing interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical standard. Not applicable.

References

[1] L. Yan, T. Stouraitis, J. Moura, W. Xu, M. Gienger and S. Vijayakumar, "Impact-aware bimanual catching of large-momentum objects," *IEEE Trans. Robot*. **40**(1), 2543–2563 (2024). doi: 10.1109/TRO.2024.3381551.

[2] Y. Hu, J. Li, Y. Chen, Q. Wang, C. Chi, H. Zhang, Q. Gao, Y. Lan, Z. Li, Z. Mu, Z. Sun and A. Knoll, "Design and control of a highly redundant rigid-flexible coupling robot to assist the COVID-19 oropharyngeal-swab sampling," *IEEE Robot. Autom. Lett*. **7**(2), 1856–1863 (2022). doi: 10.1109/LRA.2021.3062336.

[3] H. Wang, X. Wang, W. Yang, Z. Du and Z. Yan, "Construction of controller model of notch continuum manipulator for laryngeal surgery based on hybrid method," *IEEE/ASME Trans. Mechatron*. **26**(2), 1022–1032 (2021). doi: 10.1109/TMECH.2020.3015133.

[4] Q. Jiang, F. Wang and S. Jiang, "Modeling and experimental analysis of wire-driven continuum surgical robot," *J Robot Surg*. **18**(1), 98 (2024). doi: 10.1007/s11701-024-01839-4.

[5] P. Tumlin, B. Dugan, S. Hassid, G. Lawson and M. Turner, "Robotic and endoscopic trans-oral total laryngectomy, a systematic review and meta-analysis," *J Robot Surg*. **18**(1), 214 (2024). doi: 10.1007/s11701-024-01970-2.

[6] J. Burgner-Kahrs, D. C. Rucker and H. Choset, "Continuum robots for medical applications: A survey," *IEEE Trans. Robot*. **31**(6), 1261–1280 (2015). doi: 10.1109/TRO.2015.2489500.

[7] O. M. Omisore, S. Han, J. Xiong, H. Li, Z. Li and L. Wang, "A review on flexible robotic systems for minimally invasive surgery," *IEEE Trans. Syst. Man Cybern. Syst*. **52**(1), 631–644 (2022). doi: 10.1109/TSMC.2020.3026174.

[8] H. Zhou and N. Ma, "Modeling and experimental implementation of a flexible SMA wire-based gripper for confined space operation," *J Intel Mat Syst Str*. **17**(33), 2161–2175 (2022).

[9] Z. Li, H. Ren, P. W. Y. Chiu, R. Du and H. Yu, "A novel constrained wire-driven flexible mechanism and its kinematic analysis," *Mech Mach Theory* **95**(94), 59–75 (2016).

[10] G. Dong and Z. H. Zhu, "Kinematics-based incremental visual servo for robotic capture of non-cooperative target," *Robot Auton Syst*. **112**(1), 221–228 (2019). doi: 10.1016/j.robot.2018.10.011.

[11] Z. Mu, L. Zhang and L. Yan, "Hyper-redundant manipulators for operations in confined space: Typical applications, key technologies, and grand challenges," *IEEE T Aero Elec Sys*. **58**(6), 4928–4937 (2022).

[12] D. Rivoir, S. Bodenstedt, F. von Bechtolsheim and M. Distler, "Unsupervised Temporal Video Segmentation as an Auxiliary Task for Predicting the Remaining Surgery Duration," **In:** *OR. 2.0 Context-Aware Operating Theaters and Machine Learning in Clinical Neuroimaging* (L. Zhou and D. Sarikaya, eds.) (Springer, Shenzhen, 2019) pp. 29–37.

[13] L. Lei, B. Zhao, X. Qi, R. Mi and H. Ye, "Robotic needle insertion with 2D ultrasound-3D CT fusion guidance," *Ieee T Autom Sci Eng*. **21**(4), 6152–6164 (2023). doi: 10.1109/TASE.2023.3322710.

[14] X. Wang, Z. Qiu, X. Liu, C. Guo, X. Liu, J. Wang, X. Duan, C. Li, J. Deng and Y. Xue, "Research on the robot-assisted navigation technology in mandibular reconstruction surgery," *Robotica* **43**(9), 3316–3331 (2025). doi: 10.1017/S0263574725102221.

[15] Z. Yang, S. Wen, Q. Qi, Z. Lv and A. Ji, "A trocar puncture robot for assisting venipuncture blood collection," *Robotica* **42**(5), 1597–1613 (2024). doi: 10.1017/S0263574724000407.

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press

[16] W. S. Rone and P. Ben-Tzvi, "Continuum robot dynamics utilizing the principle of virtual power," *IEEE T Robot*. **30**(1), 275–287 (2014).
[17] Z. Mu, Y. Gao, W. Li, G. Lv, S. Zhao and R. Dong, "A static modeling method considering friction for the shape prediction or force sensing of concentric cable-driven manipulators," *IEEE T Autom Sci Eng*. **42**(1), 1–12 (2024). doi: 10.1109/TASE.2024.3408916.
[18] Y. Wang, H. Wang, Z. Liu and W. Chen, "Visual servo-collision avoidance hybrid task by considering detection and localization of contact for a soft manipulator," *IEEE/ASME Trans. Mechatronics* **25**(3), 1310–1321 (2020). doi: 10.1109/TMECH.2020.2974296.
[19] M. W. Hannan and I. D. Walker, "Real-time shape estimation for continuum robots using vision," *Robotic* **23**(5), 645–651 (2005).
[20] Y. Gao, J. Li, J. Li and S. Wang, "Modeling the convergence accommodation of stereo vision for binocular endoscopy," *Int J Med Robot*. **14**(1), 1–13 (2018). doi: 10.1002/rcs.1866.
[21] J. Li, Y. Sun, H. Su, G. Zhang and C. Shi, "Marker-based shape estimation of a continuum manipulator using binocular vision and its error compensation," *IEEE ICMA*. **20**(1), 1745–1750 (2020). doi: 10.1109/ICMA49215.2020.9233616.
[22] S. Xu, G. Li, D. Song, L. Sun and J. Liu, *Real-time shape recognition of a deformable link by using self-organizing map*, (2018).

Cite this article: Z. Wei, W. Li, J. Li, N. Zhang, Y. Li and Z. Mu (2026). "An S-SOM method based on binocular vision for configuration detection of concentric cable-driven manipulators", Robotica **44**, 684–697. https://doi.org/10.1017/S0263574726103270

https://doi.org/10.1017/S0263574726103270 Published online by Cambridge University Press