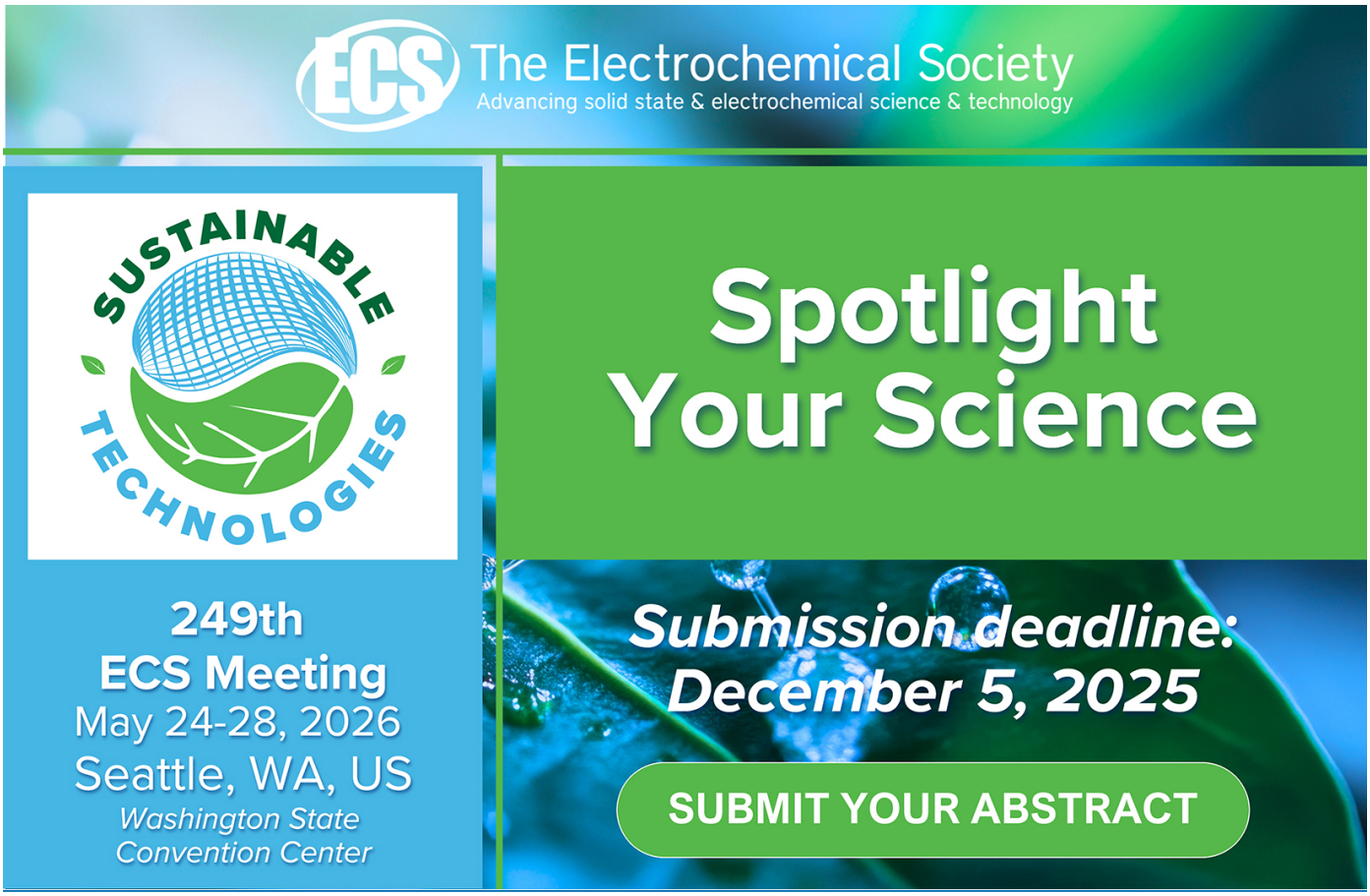
Smart Materials and Structures

PURPOSE-LED PUBLISHING™

PAPER

Teleoperation control of soft pneumatic fingers based on visual target recognition of hand bending features

To cite this article: Jiwen Fang *et al* 2025 *Smart Mater. Struct.* **34** 105009



IOP Publishing Smart Materials and Structures
Smart Mater. Struct. **34** (2025) 105009 (11pp) https://doi.org/10.1088/1361-665X/ae0d19

Teleoperation control of soft pneumatic fingers based on visual target recognition of hand bending features

Jiwen Fang*, Shuangfa Qin, Jinlei Song, Chong Li and Mingming Lv

School of Mechanical Engineering, Jiangsu University of Science and Technology, Zhenjiang 212100, People's Republic of China

E-mail: fjw617@just.edu.cn

Received 4 March 2025, revised 25 August 2025
Accepted for publication 28 September 2025
Published 9 October 2025

Abstract

Soft robots possess unique operational advantages in unstructured environments. The teleoperation control of soft pneumatic fingers, incorporating visual target recognition based on hand bending characteristics, provides a technological foundation for scenarios such as medical assistance and hazardous environment operations. It further represents a new paradigm of contact-free teleoperation control. A multi-cavity soft finger was designed and fabricated using the silicone rubber material Smooth-Sil 950. A Yeoh model was utilized to establish the relationship between the input pneumatic pressure and the bending angle, and the bending characteristics of the soft finger were simulated by the finite element method. A spatial gesture feature recognition system based on OpenCV is designed by analyzing the HSV (H-Hue, *S*-Saturation, V–Value) color space. Analyze the image features of the human hand, identify 21 key points of the human hand based on the MediaPipe framework, and perform theoretical modeling to design a finger bending angle recognition algorithm. The fuzzy matching mapping mechanism between the human hand and the soft finger is realized by the different identification values of the matrix. The activation matrix enables independent control of individual fingers while maintaining compatibility for multi-finger coordinated control. Integrate the feedback information from the pressure sensor and the flexible bending sensor to construct the closed-loop control of the soft finger. Experimental results demonstrate that the multi-chamber type of soft pneumatic finger achieves precise controllability within 110° bending range. High-precision remote control of the soft robot is realized using only a monocular camera, while visual recognition of gesture features maintains good robustness even in complex backgrounds. It validates that closed-loop control of soft robots based on vision is an effective approach for achieving contactless teleoperation.

Keywords: teleoperation control, soft fingers, visual target recognition, fuzzy matching mapping

1. Introduction

Due to the complexity and unstructured nature of the work scenarios, higher demands are placed on the dexterity of the fingers. Rigid fingers require more complex structures and more precise control to meet such requirements. The degrees of freedom of rigid fingers cannot be increased all the time, and the complexity of their hardware will be difficult to realize. Compared to rigid fingers, soft fingers have nearly infinite degrees of freedom and excellent flexibility and safety [1], which can provide excellent human-computer interaction as well as damage-free grasping, better environmental adaptability, and can be applied in complex working environments such as medical aid, disaster rescue, and smart agriculture. Commonly used materials for soft fingers include

* Author to whom any correspondence should be addressed.

silicone rubber [2, 3], electroactive polymers [4, 5], and shape memory alloys [6]. With the rapid development of 3D printing technology and material science, pneumatic soft actuators made of silicone rubber materials have been widely used. Currently, commonly used pneumatic soft actuator structures can be divided into multi-chamber type actuators [7] and fiber-reinforced actuators [8].

Silicone rubber is usually chosen as the material for soft fingers because of its simple manufacture, low toxicity, excellent elasticity, and small mechanical damping coefficient value [9]. Pneumatic mode is widely used because of its green, lightweight, and easy accessibility. Soft robots have excellent adaptability and flexibility and show great potential for application in the fields of agriculture, medical treatment, rescue, and home service [10].

At present, soft robots have received extensive attention and in-depth research from scholars at home and abroad. Generally, the deformation of a soft finger is mainly in the form of in-plane bending, focusing on the change of angular displacement at the tip point of the finger [11, 12]. Feifei Chen employs b-splines to generate free boundary surfaces for pneumatic soft robots and uses nonlinear mechanics modeling and shape derivative-based optimization to navigate the high-dimensional design space [13]. The method can be used to solve the design problem of multiple cavities by co-optimization of morphological shapes and pneumatic pressures to achieve multiple target deformation behaviors. Marc Peral introduces a new deep learning framework that utilizes visual information to recognize commands from humans predefined in a gesture database, allowing people to use gesture recognition to communicate with robots for teleoperation purposes [14]. A data-driven visual servo control strategy is developed based on the integrated model to transfer the feedback error from the rigid soft robot to the robot terminal coordinate system, eliminating systematic errors and attenuating the effects of model inaccuracies [15]. Liu proposed a theoretical model to evaluate the bending deformation, gripping force, and loading capacity of hybrid pneumatic soft fingers, and investigated the effects of various design parameters on their performance to achieve a balance between the required flexibility and essential stiffness [16].

Compared with rigid robots, the difficulties in controlling soft robots are mainly due to the use of soft materials, different drive technologies [17], and diverse structural designs, which make soft robots have the characteristics of large deformation. Therefore, it needs to involve multidisciplinary cross-analysis research, which leads to many difficulties in modeling, feedback, and control algorithms for soft robots. In addition to the challenges of precision feedback control of soft robots that need to be overcome, the pressing need for applications extends the research on soft robots in the direction of teleoperation.

The teleoperation control technology is currently a hot spot in robot control research. Combining teleoperation with soft robot control not only improves the intuition and naturalness of robot operation but also reduces the user learning cost and expands the application scenarios of soft robots. Depending on visual or haptic (or force) feedback is a common approach to robot teleoperation control [18–20]. Among them, remote operation based on gesture recognition is gradually becoming a program that has attracted much attention.

Gesture recognition technology, as a natural and intuitive human-computer interaction method, has been widely used in the fields of smartphones and game control. Gesture recognition requires the use of cameras, sensors, and other devices to capture user gestures and identify and classify gestures through image processing, machine learning, and other technologies. Mapping the recognized gestures to the corresponding robot control commands requires the design of suitable mapping algorithms and control strategies to achieve accurate and flexible remote control operations. In addition, the requirements of real-time, stability, and user experience of the system need to be taken into account to ensure the effectiveness and performance of remote control.

Soft fingers are developing rapidly in the direction of tactile feedback, force perception, and high-precision control. However, in the face of special environmental operations (such as deep-sea exploration and space operations) and other inappropriate work or inaccessible working conditions cannot meet the needs of its use. Therefore, it is an effective measure to accomplish the related work by using teleoperated robots.

Currently, the majority of remote control regarding soft robots is based on interaction devices such as data gloves to achieve human-robot collaboration and enable them to complete complex work tasks. Ma proposes a hand movement system that simplifies the definition of human hand movements to facilitate detection and uses inertial sensors mounted on the hand to obtain additional information about hand movements. Sign language movements can be enabled to control soft hands to interact with people [21]. Chua employed a single RGB camera for gesture recognition and localization, utilized the target detection algorithm YOLOv3 to achieve recognition from static gestures to dynamic gesture control, and designed a dataset of all gestures and their corresponding commands for training. The gestures can be accurately recognized and show excellent human-computer interaction response during the execution of gesture commands. However, the method has the limitation of not being able to accurately output the position of the hand pose [22].

Research on the remote control of soft fingers is still in its infancy. The current research is characterized by the difficulty of control, serious time delay, and the need for a large number of datasets to be collected. In this paper, a simple and fast remote control method is proposed, which utilizes vision to detect the bending angle of the finger, and completes the control matching with the soft finger through the fuzzy mapping mechanism to realize remote control.

2. Visual object recognition of gestures

The OpenCV visual library uses Python to realize real-time gesture recognition functions through the camera and relies on the Mediapipe framework to carry out human hand feature extraction.

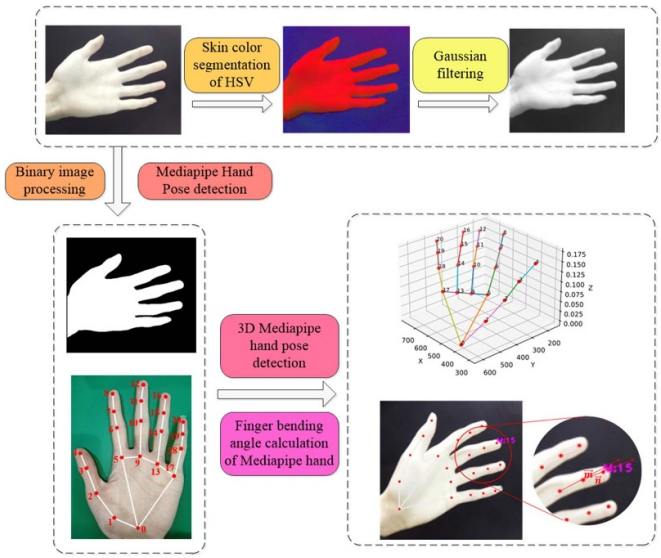


Figure 1. Gesture image processing and recognition algorithms.

Video streaming is used to achieve the acquisition of hand images and the real-time image to perform color space conversion and other image pre-processing. The interference of color information is removed to improve image clarity and recognition. According to the details of the noise, the Gaussian filtering algorithm is used to reduce the impact of noise on the image information, so that the image is smooth to facilitate the subsequent analysis of the image processing. The next step is to binarize the image in real-time to retain important image features. The gesture image processing and recognition are shown in figure 1.

Although human hand skin color is different, skin color is not affected by light and ethnicity in color space, and skin color based on gesture segmentation is widely used in the computer vision field. The HSV (H-Hue, *S*-Saturation, V–Value) color space model can deal with luminance and color information separately, which can effectively reduce the complexity of computation, and the calculation is simple. In HSV color space, the separation of luminance information and color information occurs, and human perception of color, lightness, and darkness in the picture information is also independent of each other. The color space changes with black and white, so the HSV color space model is more intuitive and conforms to the subjective feeling of the user. Therefore, the skin color segmentation method of the HSV color space is used to extract gesture-related information.

To reduce the effect of noise and improve the detection, a Gaussian filter is used to further process the image. The standard deviation affects the smoothness of the image after processing by the Gaussian filter algorithm, and the closer to the pixel point, the higher the corresponding pixel weight, and the greater the impact on the pixel point. In conjunction with the two-dimensional image information to be processed, remember that the size of the two-dimensional image is m^*n, then the formula corresponding to a certain element (x, y) on this image is:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e \frac{(x-m/2)^2 + (y-n/2)^2}{2\sigma^2}. \tag{1}$$

To get the obvious effect of the gesture or palm contour, the image information is binarized. Because of the complexity of the image information captured by the camera, after completing the preliminary image processing, it is necessary to perform the final pixel processing on the image. That is black-and-white processing of the image information, eliminating pixels above or below a certain threshold. Firstly, a threshold value T is fixed, so that the pixels are compared with the set threshold value. When it is less than the threshold value, it is the background, and vice versa. Remember the original image as f(x,y), then the binarised image is written as w(x,y), which leads to the following equation:

$$w(x,y) = \begin{cases} 1, \mathrm{f}(x,y) \geqslant T \\ 0, \mathrm{f}(x,y) \leqslant T \end{cases}. \tag{2}$$

The hand detection model is created by the MediaPipe framework on the real-time video captured by the camera, which enables real-time hand tracking and recognizes gestures. The average accuracy of its recognition is 95.7% [23]. At the same time, 21 key points within the hand are accurately located, and 3D coordinates are output in real-time. By mathematically calculating the coordinate information of the key points, the relevant information of the gesture is obtained. For example, by calculating the angle between the vectors connecting some key points, the bending angle of the finger can be derived, as shown in figure 2.

To facilitate the derivation of finger bending angles, the skeletal model of the hand is simplified to the extent that each phalanx of the four fingers except the thumb has a degree of freedom to bend toward the wrist point, and the four joints from the palm to the fingertips are defined as the metacarpophalangeal joints (MCP), the proximal interphalangeal (PIP), the distal interphalangeal joints (DIP), and the tip points of the fingers (TIP), respectively. Among these joints, the PIP and DIP have 1 degree of freedom respectively, the MCP has 2 degrees of freedom, and the degree of freedom of the TIP is 0. At this point, to conveniently express the properties of the local coordinate system, the wrist point position is denoted as *W*, the middle finger metacarpophalangeal joint position is denoted as *M*, and the ring finger metacarpophalangeal joint position is denoted as *R*. The unit direction vectors of the axes of the wrist coordinate system are denoted as x, y, and z. The conversion can be derived as

$$\begin{aligned} x &= y \times z \\ y &= \frac{\overrightarrow{WR} \times \overrightarrow{WM}}{\left|\overrightarrow{WR} \times \overrightarrow{WM}\right|} \\ z &= \frac{\overrightarrow{WR} + \overrightarrow{WM}}{\left|\overrightarrow{WR} + \overrightarrow{WM}\right|}. \end{aligned} \tag{3}$$

Since each joint on the finger is in the plane where its finger is located, for the convenience of calculating the angular data of the MCP, PIP, and DIP joints, the coordinates of the MCP,

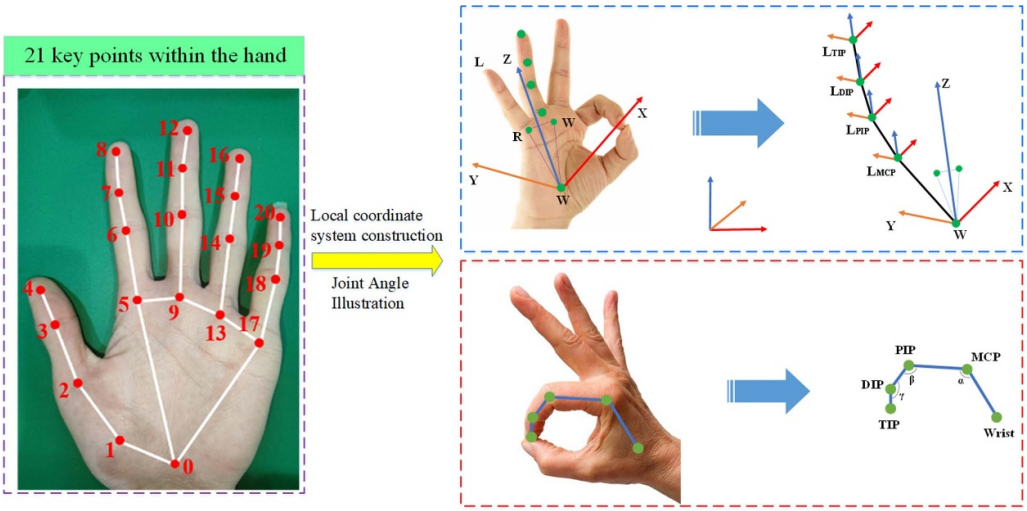


Figure 2. Establishment of local coordinate system and analysis of joint angles.

PIP, DIP, and TIP joints are denoted as M, P, D, and T, respectively. The angle between the adjacent phalanges of each finger is denoted as α, β, and γ, respectively. Then the pinch angle equation can be written as

$$\alpha = \arccos \frac{\overrightarrow{MW} \cdot \overrightarrow{MP}}{\left|\overrightarrow{MW}\right| \left|\overrightarrow{MP}\right|}$$
$$\beta = \arccos \frac{\overrightarrow{PM} \cdot \overrightarrow{PD}}{\left|\overrightarrow{PM}\right| \left|\overrightarrow{PD}\right|}$$
$$\gamma = \arccos \frac{\overrightarrow{DP} \cdot \overrightarrow{DT}}{\left|\overrightarrow{DP}\right| \left|\overrightarrow{DT}\right|}. \tag{4}$$

The average of the bending angle of each joint is taken to replace the actual bending angle. To obtain the actual bending angle, further calculations are required to obtain the actual bending angle of each joint. The bending angles of each joint of the finger are denoted as α', β', and γ', which are calculated as follows,

$$\alpha' = \pi - \alpha$$
$$\beta' = \pi - \beta$$
$$\gamma' = \pi - \gamma$$
$$\theta = \frac{\alpha' + \beta' + \gamma'}{3}. \tag{5}$$

After completing the hand detection throughout the image, MediaPipe utilizes a hand keypoint model for fine-grained processing. The model uses a regression algorithm to accurately locate the position of the 21 3D hand keypoint coordinates of the hand within the detected region [24]. To ensure that the model accurately obtains the locations of the 21 keypoints of the hand, about 30 000 real images have been manually labeled, each of which contains 21 sets of 3D coordinates. The MediaPipe hands module not only labels the locations of these keypoints but also assigns corresponding labels to them, as shown in figure 2.

The captured image information can be processed by MediaPipe to output real-time information on 21 key points, which all have 3D (X, Y, Z) coordinates. Because the distance of key points between fingers is determined by the human hand joints, the distance of key points can be derived from the following formula,

$$l_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}. \tag{6}$$

Based on the direction vectors of the lines between the keypoints on each finger, the bending angle of the finger can be calculated from the angle between the two space vectors. For example, the direction vector of the line connecting the 10th and 11th key points is $\boldsymbol{m} = (m_1, m_2, m_3)$, and the direction vector of the line connecting the 11th and 12th key points is $\boldsymbol{n} = (n_1, n_2, n_3)$, then the angle of the space vectors of the two line segments can be calculated from the following formula,

$$\cos\varphi_{ij} = \frac{m_1 n_1 + m_2 n_2 + m_3 n_3}{\sqrt{m_1^2 + m_2^2 + m_3^2}\sqrt{n_1^2 + n_2^2 + n_3^2}}. \tag{7}$$

Gesture recognition starts with the acquisition of data from the human hand, and the transfer of information is completed by recognizing the image information of the human hand through mathematical algorithms, as shown in figure 3. Keypoints are identified by combining the skeletal structure of the human hand. The meaning expressed by a particular gesture is then recognized by a particular combination of these key points, for example, the bending of a particular finger.

In the process of gesture recognition, to reduce the adverse effects of hand movement, the sampling data of a fixed interval is averaged. Therefore, the bending angle of the finger obtained can be written as

$$\varphi_i = \frac{\sum_{j=1}^{n} \varphi_{ij}}{n}. \tag{8}$$

Figure 3. Hand keypoints detected using MediaPipe.

Table 1. The input and output based on the fuzzy matching mechanism.

Input	p	1	2	3	4	5	6
Output	k	1	3	5	7	9	11

The fuzzy matching mechanism is used to realize the mapping between the hand and the soft finger. The fuzzy input p_i was constructed from the averaged bending angle φ_i of the finger as a variable.

$$p_i = \left[\left|\frac{\varphi_i + a}{a}\right|\right] \tag{9}$$

where a is the value of the mapping interval. For the convenience of subsequent experiments, the value of a is 20.

The output of fuzzy matching is a scale factor k_i. The relationship between the input and output of the fuzzy mechanism is shown in table 1. The reference angle value θ_i of the corresponding soft finger can be obtained by multiplying the scale factor k_i with the set value b,

$$\theta_i = k_i{}^* b \tag{10}$$

where i is the finger number ($i = 1, 2, 3, 4, 5$). For experimental convenience, the value of b is set to 10.

The five-finger reference input of the soft finger can be obtained by multiplying the angle matrix $\boldsymbol{A}$ with the activation matrix $\boldsymbol{S}$. The main diagonal value of the angle matrix $\boldsymbol{A}$ is the input angle value corresponding to the soft finger. All other values are 0,

$$A_{ij} = \begin{cases} \theta_i & i = j \\ 0 & i \neq j \end{cases}. \tag{11}$$

The activation matrix $\boldsymbol{S}$ can be regarded as a switch for subsequent bending control. When the variable s_i in the activation matrix is 1, it means that the corresponding soft finger will perform bending control. When the value of the variable s_i is 0, it means that the corresponding finger is not deformed,

$$S = \begin{bmatrix} s_1 & s_2 & s_3 & s_4 & s_5 \end{bmatrix}^T$$
$$s_i = \begin{cases} 0 & \text{off} \\ 1 & \text{active} \end{cases}. \tag{12}$$

3. Fabrication and simulation analysis of soft finger

The designed five-finger soft robot is mainly divided into soft actuator and hand support parts. The soft actuator adopts a multi-cavity type drive structure, which is made through 3D mold design, silicone casting, and actuator curing and molding processes, as shown in figure 4. The soft actuator is modeled after the shape of a finger, which mainly includes a pneumatic inner cavity, a restriction layer, a finger connector, and an air channel. The outer side of the drive layer is semi-cylindrical and increases the equal spacing of the raised structure. The inner side is a semicircular thin-walled cavity structure with a semicircular groove ventilation channel. The air chambers inside the soft finger are independent of each other, and the air channels in the restriction layer connect the air chambers in series into a closed whole. The hand support is completed by direct 3D printing using additive manufacturing. The detailed structural parameters of the soft robotic finger are described in table 2.

The support member is provided with a boss at each finger joint to act as a fixation to prevent axial rotation or sliding of the soft finger. The interior is provided with an air channel, and a groove structure is designed on its front side to enhance friction. The back side is provided with grooves of a certain depth and a cover plate to facilitate the rational arrangement and fixation of the air tubes.

The soft finger is made of Smooth-Sil 950 silicone rubber in a 10:1 ratio of uniformly mixed Silicone A and Silicone B. It is subjected to vacuum defoaming and thermostatic oven curing. Silicone adhesive is used to bond the driver layer, limiting layer, and air tubing, which are also fastened with tie wraps. The bending sensor is fixed in the inner groove on the outer

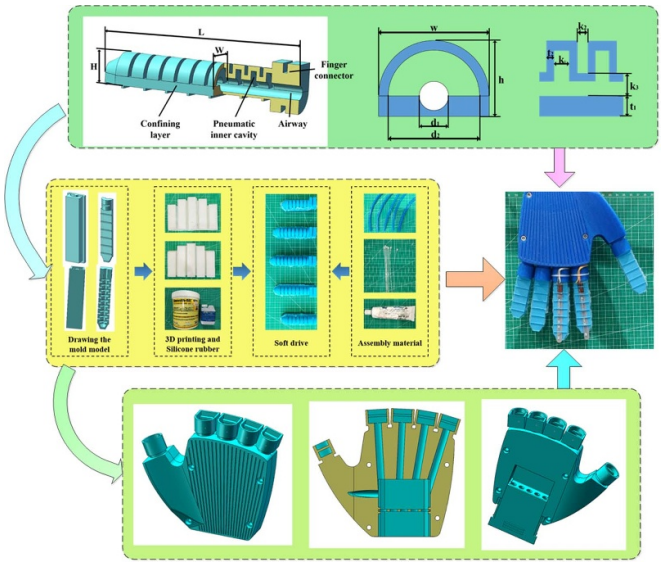


Figure 4. Design and fabrication of soft finger.

Table 2. Soft finger structure parameters.

Parameters	Value (mm)	Parameters	Value (mm)
Total length (l)	87.1	Thickness of bottom (t_1)	4
Total width (w)	18	Thickness of top (t_2)	1.5
Total height (h)	13	Inner width (k_1)	2.5
Inner diameter (d_1)	6	Groove width (k_2)	1.5
Outer diameter (d_2)	14	Groove radius (k_3)	5

side of the limiting layer of the soft finger, and the end of the bending sensor is fixed with a tie-wrap so that it can be disassembled repeatedly without affecting usability. Each of the five soft fingers is assembled to the support member according to the manual structure, and the support plate secures each of the five air tubes separately and independently of each other.

The bending deformation of the soft finger is simulated based on the Yeoh model with the help of ABAQUS finite element analysis software. The influence of the bending sensor on the bending deformation of the soft finger is ignored. The fixed end of the soft finger is clamped to the support, and air pressure of different size values is applied to observe the bending deformation of the soft finger. The relationship between the simulation and experimental results about the bending angle of the finger is shown in figure 5.

The experimental variation trend is consistent with the results obtained by the simulation. Because of measurement errors and the presence of bending sensors installed in the limiting layer, the experimental data curve does not coincide with the simulation result curve completely.

4. Experiments and results

The experimental platform is mainly composed of a drive module, control module, and sensing module, as shown in figure 6. Among them, the drive system is mainly composed

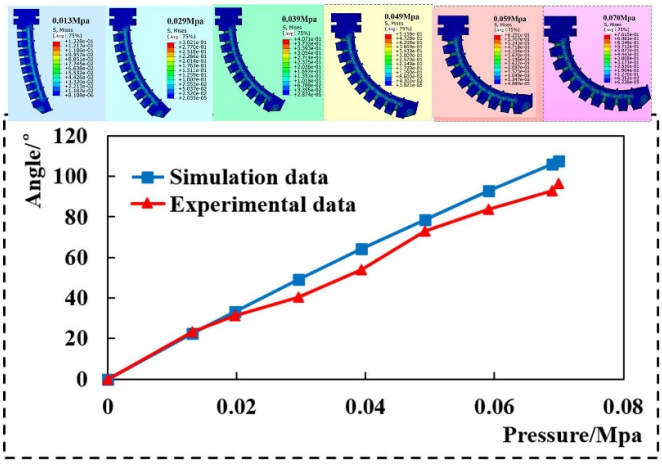


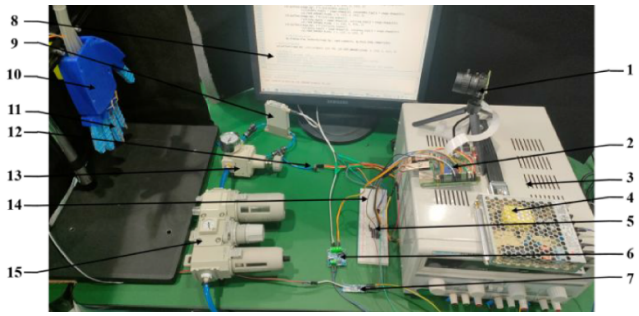
Figure 5. Simulation data and experimental data.

of an air pump, an oil mist separator, an electromagnetic proportional valve, a manual regulator, a 24 V power supply, and a soft finger. The control system mainly includes a Raspberry Pi 4B, an ADC module, an air pressure sensor, and a bending sensor. The experimental platform can complete the bending experiments of the soft finger and the remote operation control experiments.

The remote operation system mainly includes a gesture recognition processing unit and a soft finger control processing unit. The teleoperation system operates the operator independently from the closed-loop control of the finger. The first-level processing unit only needs to process the human hand image, read the human finger bending angle information, and transfer the data to the control system processing unit. The second-level processing unit needs to be based on the hand bending angle information sent by the previous level processing unit as the reference input of the soft finger, and through the return of the bending sensor and air pressure sensor to form a closed-loop control. Thus, the remote operation control of the soft finger based on visual gesture recognition is realized.

Based on the computer vision technology, remote operation control experiment is mainly through the visual input device, which will send gesturing image input to the upper computer, after a series of image processing, and then extract based on the video stream of the key information of the hand and calculate the bending angle of the finger. Finally, the recognized information will be sent to the control unit. The final realization of human gestures and the computer interaction control effect. The processing unit of the gesture recognition system uses Pycharm programming software as the upper computer and Raspberry Pi4B as the lower computer to communicate with the upper computer, using a camera to obtain the video data stream and receive human gesture information through OpenCV.

In the experiment, the human hand is fixed in position, facing the camera and adjusting the finger bending state, and the camera transmits the captured image information to the Raspberry Pi shown in figure 7. The bending sensor on the limiting layer of the soft body driver transmits the angle information to the Raspberry Pi through the ADC module, and the



1-Camera, 2-Raspberry Pi 4B, 3-.24V power supply, 4-12V power supply, 5-ADC conversion module, 6-PWM Signal Converter, 7-Sensor Module, 8-Display, 9-Solenoid Proportional Valve, 10-Soft Finger, 11-Bending Sensor, 12-Air Pressure Sensor, 13-Manual Regulator, 14-Breadboard, 15-Oil Mist Separator

Figure 6. Experimental platform.

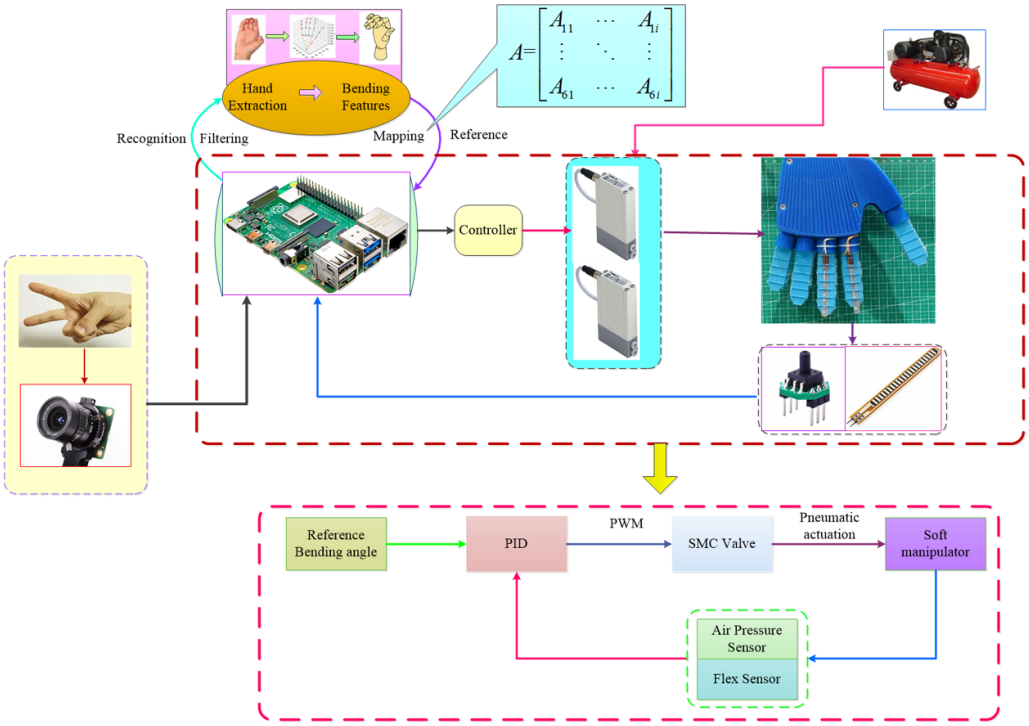


Figure 7. Control flow of the remote operation system.

Raspberry Pi compares the received gesture recognition angle signal with the bending sensor angle information. When there is a difference between the two data, the Raspberry Pi adjusts the PWM signal to control the opening of the electromagnetic proportional valve, and at this time, the barometric pressure sensor transmits the detected barometric pressure data back to the Raspberry Pi through the ADC module, precise control of the electromagnetic proportional valve is achieved and closed-loop control is formed. The control system processing unit uses a Raspberry Pi 4B to control the whole system. Raspberry Pi adjusts the gas input of the real electromagnetic proportional valve to the soft finger by outputting PWM signals through the PID control algorithm. Through the bending sensor and air pressure sensor, the soft finger can realize closed-loop control.

The PID control algorithm is utilized to realize the precise control of the soft finger. By setting different PID control parameters, the bending measurement results are shown in figure 8. The control performance is evaluated based on finger response (rise time) and stability (steady-state error), with greater emphasis on angular error during steady-state control. Experimental results indicate that the rise time is minimized when $K\mathrm{p} = 0.16$ or 0.17, while the steady-state error is minimized when $K\mathrm{i} = 0.06$. Although $K\mathrm{d}$ has a relatively minor influence on the system, it suppresses overshoot during large-angle bending. Therefore, comprehensive analysis suggests that the configuration with coefficients $K\mathrm{p} = 0.16$, $K\mathrm{i} = 0.06$, and $K\mathrm{d} = 0.01$ yields optimal performance comparatively.

The tip displacement of a soft finger is most pronounced under pneumatic actuation. Take the end of the soft finger as

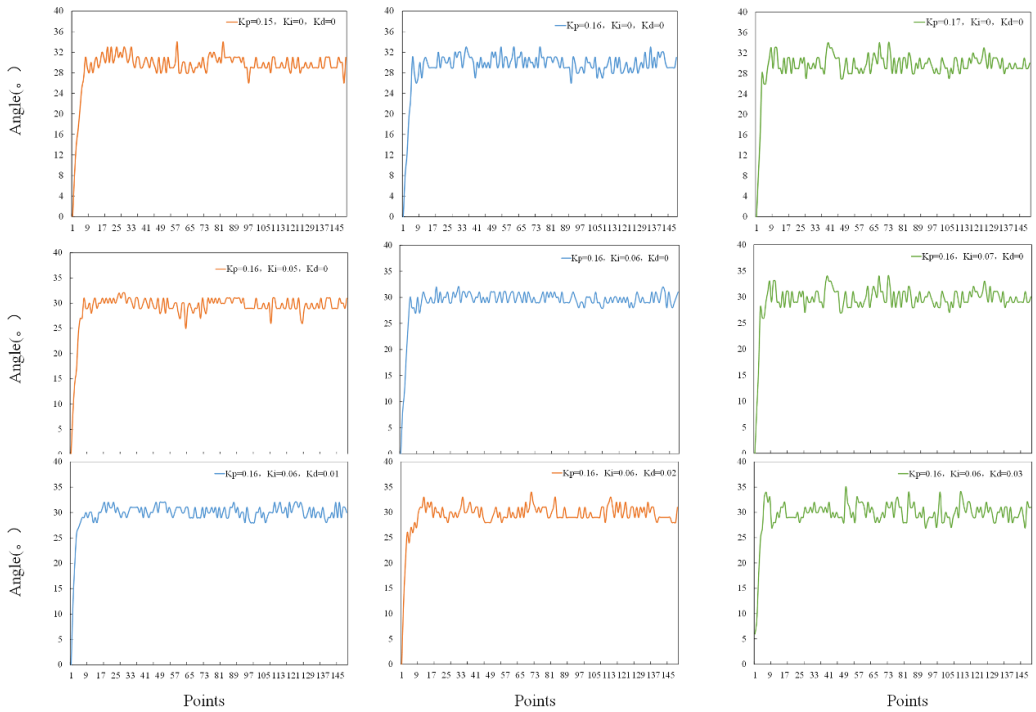


Figure 8. PID control experiment based on different parameters.

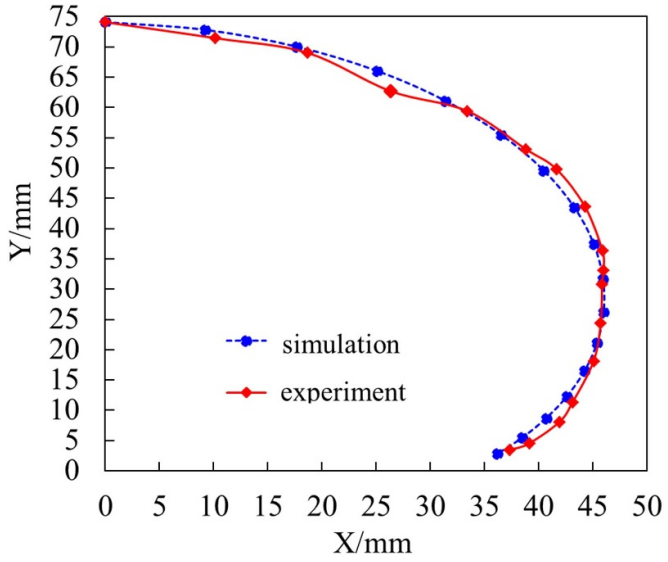


Figure 9. Simulation and experimental comparison of soft fingertip motion trajectories.

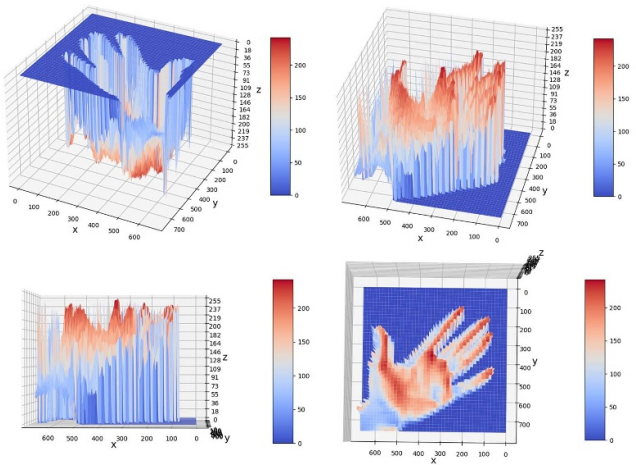
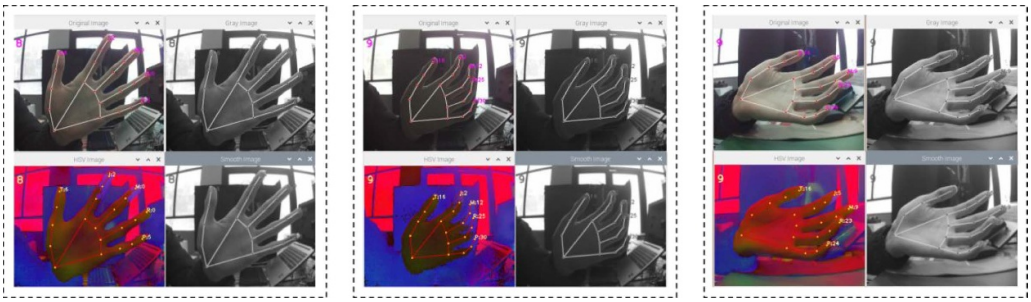


Figure 10. 3D grayscale image of the human hand.

the origin (0,0), and its side as the plane to establish a 2D plane right-angle coordinate system, measure the coordinates of the end of the soft finger, to get the specific situation of its bending, and the simulation and experimental comparisons of the coordinates of the end trajectory are shown in figure 9. From the figure, it can be observed that the moving distance of the soft finger in the *y*-axis direction is more than that in the *x*-axis direction, which indicates that it has a superior upward moving ability. It also confirms the validity of the bending deformation theory from the side. This deformation trend implies that the soft finger has a strong load-bearing capacity when grasping objects.

To improve the accuracy of the capture of the human hand, the image processing of the human hand is carried out to improve the efficiency of the Raspberry Pi image processing, and the three-dimensional grayscale image of the human hand after the enhancement of the features is shown in figure 10, which shows that it can be seen that the features of the palm and the five-fingered joints points are prominent. The acquired images contain various noises that interfere with gesture recognition. The Gaussian filtering algorithm is employed to perform smooth processing on the images. In the processed images, hand contours are well-defined, thereby improving the accuracy of gesture recognition.



(a)High light intensity (b) Medium intensity (c)Low light intensity

Figure 11. Hand recognition in different light-intensity conditions.

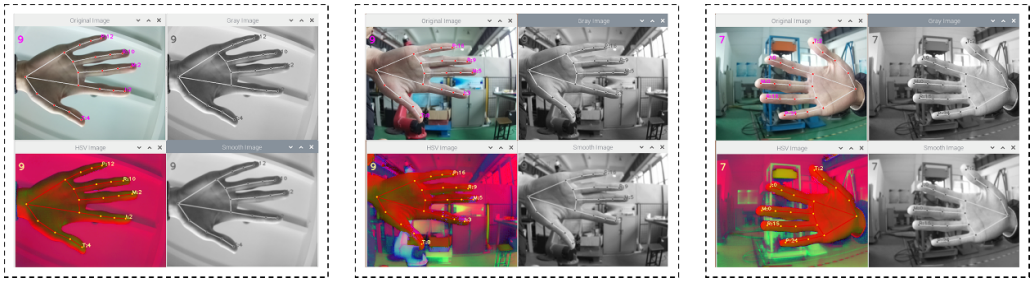


Figure 12. Hand recognition in different complex background conditions.

To detect the accuracy of gesture recognition, the human hand in different lighting conditions for gesture recognition experiments and recognition results are shown in figure 11. It can be seen that the human hand can still be captured and recognized by the camera under different strong and weak lighting conditions. The human hand is in different backgrounds for hand recognition experiments, and the recognition results are shown in figure 12. The recognition system is still able to track hand gestures and perform data computation promptly under complex background conditions.

To test the effectiveness of teleoperation control of soft finger movement and to reduce the difficulty of the experiment, the middle finger was selected as the object of the experiment, and several groups of bending angles (30°, 50°, 70°, 90°, 110°) were set to conduct teleoperation experiments, and the experimental results are shown in figure 13. The system latency in the remote closed-loop control of soft pneumatic fingers using vision-based gesture recognition primarily originates from image processing and pneumatic actuation processes. The significant hysteresis inherent in pneumatic actuation constitutes the main factor contributing to the experimental delay. Experimental results demonstrate that a soft finger, employing HSV-based skin segmentation, MediaPipe keypoint extraction, fuzzy mapping, and PID closed-loop control, accurately tracks the bending angle of a human finger. Based on the closed-loop control, the real-time bending steady-state errors are within 5°. This reveals the precise controllability of the multi-chamber soft pneumatic finger within the 0–110° bending range, validating the effectiveness of the teleoperation strategy integrating vision recognition with pneumatic control.

Furthermore, this vision-based remote control method achieves high-precision operation relying solely on a monocular camera, eliminating the need for marker-free control. It overcomes the limitations of traditional data gloves, avoids dependence on their complex hardware, and reduces experimental costs and control complexity. Experimental results confirm the system's feasibility and robustness within the 110° range, providing a technical foundation for applications such as medical assistance and operations in hazardous environments. Decoupled control via activation matrices enables independent finger actuation (only the middle finger was activated in this experiment), laying the groundwork for multi-finger coordinated manipulation.

5. Conclusion

A finger based on a multi-cavity structure soft actuator is proposed, and its bending deformation is analyzed by finite element simulation. Silicone rubber material 950 and 3D printing technology are utilized to complete the fabrication of the soft robot. Through the bending control experiments on a single soft driver, the experimental data and simulation data results are similar. Through Python and OpenCV to build a gesture recognition system based on visual recognition, to improve the method of image recognition to effectively solve the deficiencies of high latency and large errors in the common gesture recognition, so that the recognition judgment of the gesture and the extraction of data are more accurate and efficient. Combining computer vision technology, sensor technology, and pneumatic soft body drive control, a teleoperation system based on soft finger closed-loop control is established. The soft

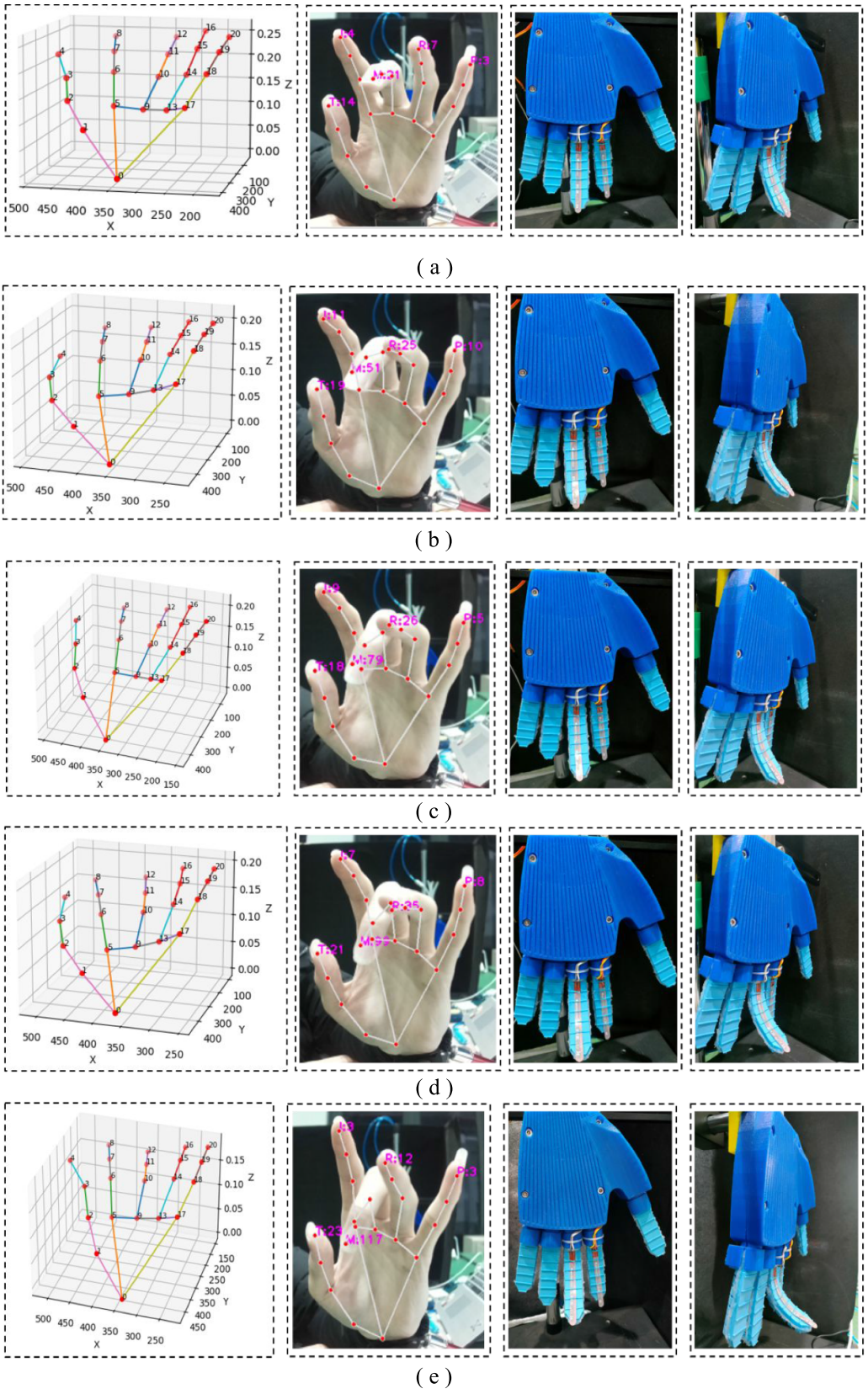


Figure 13. Remote operation control experiment for different angle bending: (a) 30°, (b) 50°, (c) 70°, (d) 90°, (e) 110°.

finger teleoperation control experiment shows that the teleoperation system has good robustness and feasibility.

Soft robot teleoperation based on gesture recognition still faces some challenges. For example, the accuracy and stability of gesture recognition need to be further improved, especially in complex environments where the recognition performance still needs to be improved. The teleoperation system of this robot is user-friendly, but it still exhibits certain delays. The latency primarily originates from two aspects: one is the process of recognizing hand gestures through visual methods, while the other is the process of pneumatic actuation. In addition, the real-time and latency problems of remote control manipulation need to be effectively solved to ensure the smoothness and responsiveness of user operations. This method eliminates the need for traditional array sensors, avoids processing signals from multiple sensors, increases signal processing speed, reduces control complexity, and offers a novel approach to contact-free teleoperation for soft robots. Despite the many challenges, soft robot remote control technology based on gesture recognition is still full of great development potential. In the future, with the continuous progress of artificial intelligence, machine learning, and other technologies, it is believed that this field will make more significant breakthroughs and progress, bringing more application scenarios.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgement

This work was supported by: the National Natural Science Foundation of China (Grant Number 51705217), Key Research and Development Program of Jiangsu Province (Grant Number BE2022062).

ORCID iDs

Jiwen Fang 0000-0002-3497-9398
Chong Li 0000-0002-2380-7777

References

[1] Wang T L, Jiao W, Sun Z and Zhang X 2023 Design and gesture optimization of a soft-rigid robotic hand for adaptive grasping *Soft Robot.* **10** 580–9
[2] Polygerinos P, Wang Z, Overvelde J T B, Galloway K C, Wood R J, Bertoldi K and Walsh C J 2015 Modeling of soft fiber-reinforced bending actuators *IEEE Trans. Robot.* **31** 778–89
[3] Sparrman B, du Pasquier C, Thomsen C, Darbari S, Rustom R, Laucks J, Shea K and Tibbits S 2021 Printed silicone pneumatic actuators for soft robotics *Addit. Manuf.* **40** 101860
[4] O'halloran A, O'Malley F and McHugh P 2008 A review on dielectric elastomer actuators, technology, applications, and challenges *J. Appl. Phys.* **104** 071101
[5] Shao J, Yu L, Skov A L and Daugaard A E 2020 Highly stretchable conductive MWCNT–PDMS composite with self-enhanced conductivity *J. Mater. Chem. C* **8** 13389–95
[6] Liang X R *et al* 2024 Soft self-healing robot driven by new micro two-way shape memory alloy spring *Adv. Sci.* **11** 2305163
[7] Liu X, Zhang J, Gu S, Zhao L and Li Z 2023 Modelling and angle tracking control for multi-chamber soft bending pneumatic muscle *IEEE Robot. Autom. Lett.* **8** 7647–54
[8] Ye Y, Scharff R B N, Long S, Han C and Du D 2024 Modelling of soft fiber-reinforced bending actuators through transfer learning from a machine learning algorithm trained from FEM data *Sens. Actuators A* **368** 115095
[9] Sut D J and Sethuramalingam P 2023 Soft manipulator for soft robotic applications: a review *J. Intell. Robot. Syst.* **108** 10
[10] Cao M, Sun Y, Zhang J and Ying Z 2023 A novel pneumatic gripper driven by combination of soft fingers and bellows actuator for flexible grasping *Sens. Actuators A* **355** 114335
[11] Yan J, Xu Z, Shi P and Zhao J 2022 A human-inspired soft finger with dual-mode morphing enabled by variable stiffness mechanism *Soft Robot.* **9** 399–411
[12] Liu X, Zhao Y, Geng D, Chen S, Tan X and Cao C 2021 Soft humanoid hands with large grasping force enabled by flexible hybrid pneumatic actuators *Soft Robot.* **8** 175–85
[13] Chen F, Song Z, Chen S, Gu G and Zhu X 2023 Morphological design for pneumatic soft actuators and robots with desired deformation behavior *IEEE Trans. Robot.* **39** 4408–28
[14] Peral M, Sanfeliu A and Garrell A 2022 Efficient hand gesture recognition for human-robot interaction *IEEE Robot. Autom. Lett.* **7** 10272–9
[15] He S *et al* 2023 A modeling and data-driven control framework for rigid-soft hybrid robot with visual servoing *IEEE Robot. Autom. Lett.* **8** 7281–8
[16] Liu X, Zhao Y, Geng D, Chen S, Tan X and Cao C 2020 Soft humanoid hands with large grasping force enabled by flexible hybrid pneumatic actuators *Soft Robot.* **8** 175–85
[17] Zhou S, Li Y, Wang Q and Lyu Z 2024 Integrated actuation and sensing: toward intelligent soft robots *Cyborg Bionic Syst.* **5** 0105
[18] Liu L, Zhang Y, Liu G and Xu W 2018 Variable motion mapping to enhance stiffness discrimination and identification in robot hand teleoperation *Robot. Comput.-Integr. Manuf.* **51** 202–8
[19] Chen B *et al* 2024 Teleoperation of an anthropomorphic robot hand with a metamorphic palm and tunable-stiffness soft fingers *Soft Robot.* **11** 508–18
[20] Mizrahi A and Sintov A 2024 TeleFMG: a wearable force-myography device for natural teleoperation of multi-finger robotic hands *IEEE Robot. Autom. Lett.* **9** 2933–40
[21] Ma H *et al* 2020 A sign language interaction system based on pneumatic soft hand *Proc. 2020 15th IEEE Conf. on Industrial Electronics and Applications (ICIEA)*
[22] Chua S N D, Chin K Y R, Lim S F and Jain P 2022 Hand gesture control for human–computer interaction with deep learning *J. Electr. Eng. Technol.* **17** 1961–70
[23] Zhang F *et al* 2020 MediaPipe hands: on-device real-time hand tracking (arXiv:2006.10214)
[24] Abdallah M S, Samaan G H, Wadie A R, Makhmudov F and Cho Y-I 2023 Light-weight deep learning techniques with advanced processing for real-time hand gesture recognition *Sensors* **23** 2