Journal of Computers Vol. 28, No. 5, 2017, pp. 228-245
doi:10.3966/199115992017102805021

Method for Target Recognition and Mobile Manipulation Control of Humanoid Robot

Chen-Xi Liu[1], Lei Zhang[1*] and Jing-Qiang Li[1]

[1] School of Information and Electrical Engineering, Beijing University of Civil Engineering and Architecture
Beijing, China
zhanglei@bucea.edu.cn

Received 25 October 2016; Revised 13 June 2017; Accepted 26 June 2017

Abstract. Humanoid robots have become the best form of human services for human beings by the widespread concern. The mobile manipulation of the humanoid robot means that the humanoid robot moves by the footsteps. At the same time, it also uses the arm and hand to complete the crawling and other operations. In order to complete the mobile manipulation, the robot is required to have target recognition, online path planning and motion control. This paper describes the hardware, software design and implementation of the minitype humanoid robot mobile operating system, and has carried on the experimental verification. Firstly, it has carried out kinematics modeling and inverse kinematics for the NAO experimental platform of minitype humanoid robot. Secondly, the research on path planning and motion control of humanoid robot is studied. Thirdly, the target recognition and positioning based on monocular vision are studied. Finally, it carries out experiments on NAO robots. In the experiment, the robot completes the mobile manipulation task accurately and efficiently. Experiments show that the target recognition and positioning method are efficient and reliable, and can realize the mobile manipulation of the robot completely. In addition, it completes the research topic of minitype humanoid robot target recognition and mobile manipulation.

Keywords: humanoid robot, mobile manipulation, monocular vision, target positioning, target recognition,

1 Introduction

The humanoid robot is a kind of intelligent robot with a similar shape with human. It has the ability to sense, self-improvement, self-control, reverie, image performance and emotional exchange with people. Its walking structure can make the robot go to the place where it is needed at any moment, which including those where is not easy for people to reach, to complete the task. The larger activity and workspace is the main advantage of humanoid robots relative to industrial robots. The biped walking structure of a humanoid robot can make it walk on a relatively large obstacle, climb stairs; walk on poor conditions of road; and can reach many places. Humanoid robots are not only similar to people, but also into our living space. At present, scientists from many developed countries, such as Japan and the United States, have done a lot of work in research and development of humanoid robots, and have made breakthrough progress. Humanoid robots have had a great impact on human society, and a great impact on human society. Further, public hope that the humanoid robot will be able to move as a normal person, that is, the humanoid robot exerts a force on the movable object through the arm during the stable biped walking, such as taking the old man to push the cart, holding heavy objects and so on. Stable, reliable and efficient completion of mobile manipulations without changing the manipulating environment, making the humanoid robot has a potentially strong application prospects, social and economic benefits. So humanoid robot has become the core of the field of robot research and development, occupied the first place of robot research and development.

[*] Corresponding Author

Monocular vision measurement technology in the industrial, agricultural and other fields have important and extensive application, where the determination of the target position information is the key technology [1]. Wang and Jia proposed a method only by using pinhole imaging principle and geometric coordinate transformation to calculate the position of obstacles. This method improved the corresponding point calibration and proposed a solution for the camera when it is changed. Xie and Yang proposed a monocular vision self- position method for autonomous soccer robot [2].This method was to realize the self- position geometric by monocular vision ranging under the premise of the color field information identified. At the same time, the lens of the nonlinear distortion and other issues were analyzed. The parameters applicable to the camera are dynamic, but the method is not efficient. This paper studies target recognition and position based on monocular vision. Using the different color characteristics of the object, the color characteristics of the target are identified, and then the target is positioned by the camera model and parameter calibration. Finally, the experiment of minitype humanoid robot mobile manipulation is done. In the experiment, minitype humanoid robot accurate and efficient to complete the mobile manipulation. Experimental results show that the method of target recognition in this paper is efficient and reliable, and it can completely realize the robot's mobile manipulation.

2 Kinematic Analysis of NAO Robot

At present, most robotic control codes are basically programmed by the coordinates of the positions of the joints. We regard the iron hand as a mechanism connected by some connecting rods in sequence, and set up separate coordinate systems for each connecting rod of the iron hand, and then apply the homogeneous transformation matrix to reflect the position and posture of the iron hand [3].

Kinematics is a subject of research in the state of motion and a part of classical mechanics, which is used to reflect the movement of the target, but the reasons for the formation of the movement are not counted. The control of the robot is based on the premise of kinematics. Many different ways have been put forward to deal with the difficulties of the robot in terms of kinematics [4]. This chapter uses the D-H (Denavit-Hartenberg) method to explore the inverse kinematics of minitype humanoid robotic arms.

2.1 Arm Kinematics Modeling

Robots usually contain continuous joints and linkers, and moving one joint will affect the connection of the joints. The connected joint can have any length (including 0) of the linker and rotate around any axis. Each joint has a standard coordinate system to interpret the link, position and determine the basic procedures for joint conversion. Starting from the reference point, a joint is followed by a joint in the order. Once all conversions have been completed, the transformation matrix is obtained and the position of the joint can be obtained from the reference point [5].

In this paper, we use the Denavit-Hartenberg (D-H) notation in kinematics calculations. The D-H method is used to describe the movement of the robot and to express movements. At the same time, it can be applied to any shape of the robot. This D-H method demonstrates the relationship between the connected joints with four different variables, as shown in Fig. 1.

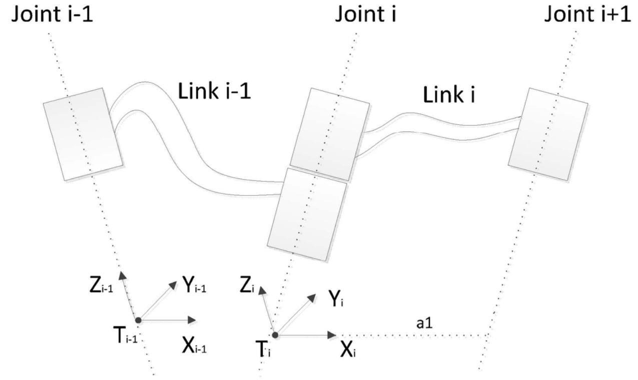


Fig. 1. Denavit-Hartenberg (D-H) method was used to represent the relationship between joints

In order to use the D-H method to represent the relationship of the robot joint, a standard coordinate system is set for each joint. The X axis and the Z axis are assigned to each joint. The Y axis is vertical to the X axis and Z axis, and can be calculated at any time. In order to calculate the coordinate system of the next joint, four standard calculations are required [6].

(1) Rotate until i on the Z_{i-1} axis. This process makes the X_{i-1} axis parallel to the X_i axis.

(2) Move the X_{i-1} axis until d_i follows the Z_{i-1} axis. This process will make X_{i-1} and X_i in the same position.

(3) Move until A_i follow the X_{i-1} axis. This process makes the original coordinates of the two coordinate systems in the same position.

(4) Rotation until i is on the X_i axis. After the process is complete, the two coordinate systems will be exactly same.

The matrix A can be obtained by multiplying the four matrices (the matrix produced by each of the steps above). Formula 1 is their relationship. Each n represents the current joint, and $n+1$ represents the next joint.

$$\begin{aligned} {}^{n+1}_{n}T = {}^{n+1}_{n}A &= R(z,\theta_{n+1}) \times M(0,0,d_{n+1}) \times M(a_{n+1},0,0) \times R(x,\alpha_{n+1}) \\ &= \begin{bmatrix} C\theta_{n+1} & -S\theta_{n+1} & 0 & 0 \\ S\theta_{n+1} & C\theta_{n+1} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d_{n+1} \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & a_{n+1} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & C\alpha_{n+1} & -S\alpha_{n+1} & 0 \\ 0 & S\alpha_{n+1} & C\alpha_{n+1} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad . \tag{1}$$

M : position change (movement), R : coordinate rotation (rotation).

This transformation matrix is calculated and sorted as shown in formula 2.

$$A_{n+1} = \begin{bmatrix} C\theta_{n+1} & -S\theta_{n+1}C\alpha_{n+1} & S\theta_{n+1}S\alpha_{n+1} & a_{n+1}C\theta_{n+1} \\ S\theta_{n+1} & C\theta_{n+1}C\alpha_{n+1} & -C\theta_{n+1}S\alpha_{n+1} & a_{n+1}S\theta_{n+1} \\ 0 & S\alpha_{n+1} & C\alpha_{n+1} & d_{n+1} \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2}$$

Table 1 shows the information for all joint lengths. These values can be used in the motion calculation of each joint.

Table 1. The length of the link for minitype humanoid robot

Name	Length (mm)
NeckOffsetZ	126.50
ShoulderOffsetY	98.00
UpperArmLength	90.00
LowArmLength	50.55
ShoulderOffsetZ	100.00
HandOffsetZ	58.00
HipOffsetZ	85.00
HipOffsetY	50.00
ThighLength	100.00
TibiaLength	102.74
FootHeight	45.11
HandOffsetZ	15.90

The minitype humanoid robot has a total of 25 joints: the head joint has 2, each arm has five (a total of 10), each leg has five (a total of 10), one in the pelvis, and two used in the implementation of the hand opening and closing movement [7-8]. Each joint can be controlled independently, but both joints in the pelvis need to be controlled at the same time. Each joint has an angle limit, and the limits are shown in the range of the table above. As for the legs, the robot’s falling touch is also included in the angle limit.

2.2 Arm’s Kinematic

In Fig. 2, the position of the hand is calculated using the arm joint value and linker data of the minitype humanoid robot. The center point of the minitype humanoid robot is the base point of the whole system (0,0,0). Use the following equation to indicate the position change.

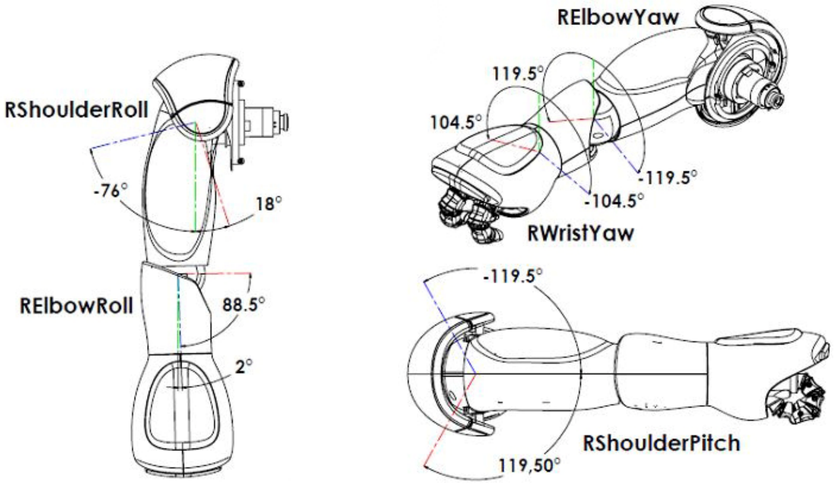


Fig. 2. The name and limit of the joint of the right hand of the minitype humanoid robot

$$ {}^{h}Pos = {}^{h}_{0}T * {}^{0}Pos \,. \tag{3}$$

Pos represents the position of the Cartesian coordinate system. T represents the transformation matrix. The symbol in the upper left corner represents the target to be converted, and the lower left corner represents where it is now. h is the hand, and o is the center point. The formula 4 (position change) can be represented by coordinate action and angle.

$$T = M * R \,. \tag{4}$$

Right hand contains five joints, from the top to the bottom, are RshoulderPitch, RShoulderRoll, RElbowRoll, RElbowYaw and RWristYaw. If the position of the hand changes from the center point in the order, the position information of the hand can be got. The following is the order of the hand to change the position from the center point to the hand: center point -> shoulder -> elbow -> wrist -> hand.

$$ {}^{h}_{0}T = {}^{s}_{0}T * {}^{e}_{s}T * {}^{w}_{e}T * {}^{h}_{w}T \,. \tag{5}$$

s : shoulder, e : elbow, and w : wrist. RShoulderPitch, RShoulderRoll, RElbowRoll, RElbowYaw, RWristYaw, are five joints of minitype humanoid robot used to calculate the D-H method.

The distance from the wrist to the hand and the movement from the center point to the shoulder are not included in the above formula. The transformation matrix T is the transformation matrix of the above five joints, and can be represented by A as follows. For the D-H method, when analyzing the arm joints of a minitype humanoid robot, it can be simplified to the formula 6.

$$A_{n+1} = \begin{bmatrix} C\theta_{n+1} & 0 & S\theta_{n+1}S\alpha_{n+1} & 0 \\ S\theta_{n+1} & 0 & -C\theta_{n+1}S\alpha_{n+1} & 0 \\ 0 & S\alpha_{n+1} & 0 & d_{n+1} \\ 0 & 0 & 0 & 1 \end{bmatrix} . \tag{6}$$

Using the D-H method to using A replace the T transformation matrix to get the position change, you can use the following formula.

$$ {}^{h}_{0}T = {}^{s}_{0}T * A_1 * A_2 * A_3 * A_4 * A_5 * {}^{h}_{w}T \, . \quad \textbf{(7)} $$

The change in the distance between the fulcrum and the shoulders without rotation is indicated by only the displacement operation, and the change of position can be shown in the following formula.

$$ {}^{s}_{0}T = {}^{s}_{0}M = {}^{s_y}_{0}M + {}^{s_z}_{0}M \, . \quad \textbf{(8)} $$

S_y represents the movement to the Y axis, S_z represents the movement to the Z axis and the sum matrix of the two movements. In the Cartesian coordinate system, move to the Y axis by -98 mm and move to the Z axis by100 mm from the center point to the right shoulder. Formula 9 shows the transformation matrix described above.

$$ {}^{s}_{0}T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -98 \\ 0 & 0 & 1 & 100 \\ 0 & 0 & 0 & 1 \end{bmatrix} . \quad \textbf{(9)} $$

The rotation transformation has been included in the shoulder-to-wrist position change, and the transformation matrix is determined according to the D-H method's d, a, θ, α. The change of the shoulder joint is represented by ${}^{e}_{s}T$, which contains two fulcrum joints, as shown in the equation. ${}^{e}_{s}R_{yp}$ means rotates around the Y axis. ${}^{e}_{s}R_{xr}$ represents a roll that rotates around the X axis.

$$ {}^{e}_{s}T = {}^{e}_{s}R = {}^{e}_{s}R_{yp} * {}^{e}_{s}R_{xr} \, . \quad \textbf{(10)} $$

${}^{e}_{s}R_{yp}$ obtained from the right shoulder anterior joint (the first fulcrum joint) is A_1 in the D-H method. A_1 is $\alpha_1 = -90°$, $d_1 = 0$. The transformation matrix is shown in formula 11.

$$ A_1 = \begin{bmatrix} C\theta_1 & 0 & -S\theta_1 & 0 \\ S\theta_1 & 0 & C\theta_1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} . \quad \textbf{(11)} $$

${}^{e}_{s}R_{xy}$ obtained from the right shoulder rolling joint (the second key joint) is A_2 in the D-H method. A_2 is $\alpha_2 = 90°$, $d_2 = 0$. The transformation matrix is shown in formula 12.

$$ A_2 = \begin{bmatrix} C\theta_2 & 0 & S\theta_2 & 0 \\ S\theta_2 & 0 & -C\theta_2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} . \quad \textbf{(12)} $$

The change in the elbow position is done by moving the shoulder to the elbow. The rotation is also done by the movement of *1* and the rotation of *2*. ${}^{w}_{e}R_{xr}$ represents the rolling rotation of the X axis, while ${}^{w}_{e}R_{zy}$ represents the rotation of the forward and backward rotation of the Y axis.

$$ {}^{w}_{e}T = {}^{e}_{e}M * {}^{w}_{e}R_{xr} * {}^{w}_{e}R_{zy} \, . \quad \textbf{(13)} $$

The movement from the shoulder to the elbow in the right elbow rolling joint (the first elbow) is 90mm on the X axis and expressed as "d". Because $\theta_3 = 90°$, $\alpha_3 = 90°$, $d_3 = 90°$, the transformation matrix is shown in formula 14.

$$A_3 = \begin{bmatrix} C_{\theta_3+\theta_3} & 0 & S_{\theta_3+\theta_3} & 0 \\ S_{\theta_3+\theta_3} & 0 & -C_{\theta_3+\theta_3} & 0 \\ 0 & 1 & 0 & 90 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{14}$$

${}^{w}_{e}R_{zy}$ indicates the position of the right elbow anterior joint (the elbow's fulcrum joint), which represents the rotation of the forward and backward rotation of the Y axis and is expressed as A_4. This transformation matrix is shown in formula 15.

$$A_4 = \begin{bmatrix} C_{\theta_4} & 0 & -S_{\theta_4} & 0 \\ S_{\theta_4} & 0 & C_{\theta_4} & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{15}$$

The position change of the right wrist anterior joint (wrist joint) occurs at 1 movement and rotation. The movement of the wrist is 50.55mm on the X axis. ${}^{h}_{w}R$ represents the rotation of the forward and backward rotation of the Y axis.

The transformation matrix of A_5 is shown in formula 16.

$$A_5 = \begin{bmatrix} C_{\theta_5} & 0 & -S_{\theta_5} & 0 \\ S_{\theta_5} & 0 & C_{\theta_5} & 0 \\ 0 & -1 & 0 & 50.55 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{16}$$

The change from the shoulder to the wrist position (including the rotation matrix) is expressed as ${}^{w}_{s}T = A_1 \times A_2 \times A_3 \times A_4 \times A_5$.

Movement from the wrist to hand is 58mm in the X axis, 15.9mm in the Y axis. The transformation matrix of ${}^{h}_{o}T$ is shown in formula 17.

$${}^{h}_{0}T = {}^{s}_{0}T * A * {}^{h}_{w}T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -98 \\ 0 & 0 & 1 & 100 \\ 0 & 0 & 0 & 1 \end{bmatrix} * A * \begin{bmatrix} 1 & 0 & 0 & 58 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -15.90 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{17}$$

The hand based on the center point in the Cartesian coordinate system can be calculated by the following matrix.

$${}^{h}Pos(x,y,z) = {}^{h}_{0}T * {}^{0}Pos(0,0,0). \tag{18}$$

2.3 Arm Inverse Kinematics and Control the Action

Inverse kinematics is the study of how to find the Angle and the moving distance of each joint, when the position of the hand has been determined. It also controls the action by calculating the continuous change in the angle value of the joint [9].

It is necessary to use a plurality of combinations to move the robot's hand to a specific position. It is impossible to find a matrix that can calculate the angle of all joint angles to get the trigonometric function equation required for the conversion. Therefore, in order to separate the equation to obtain the value of each joint, inverse matrix of the transformation matrix A_n must be multiplied by the left side of the relational equation to obtain the element that calculates the angle value.

The transformation matrix A of the data obtained from the arm of the minitype humanoid robot in the previous section. Although the wrist's front yaw joint can be used to determine the direction of the hand, the position is meaningless, so the calculation of the joint is not required in the calculation of the position of the hand. Therefore, for A, only need to get θ_1, θ_2, θ_3, θ_4 (the joint value of A_1, A_2, A_3, A_4), you can calculate the position of the hand. The connection value is removed from the front joint.

Using the algebraic operation, the angular value(θ_1, θ_2, θ_3, θ_4) of each joint can be obtained, and the position of the robot's hand in the Cartesian coordinate system (P_x, P_y, P_z) is pointed out.

The Cartesian coordinate system determines the hand position and sequentially enters the desired joint angle value to the joint action plane, and then passes the commands to move the minitype humanoid robot.

Fig. 3 presents simulation of the movement of minitype humanoid robots in the Choregraphe window. Received nine target points and the joint angles have been calculated, then the robot will receive the angle value to move.

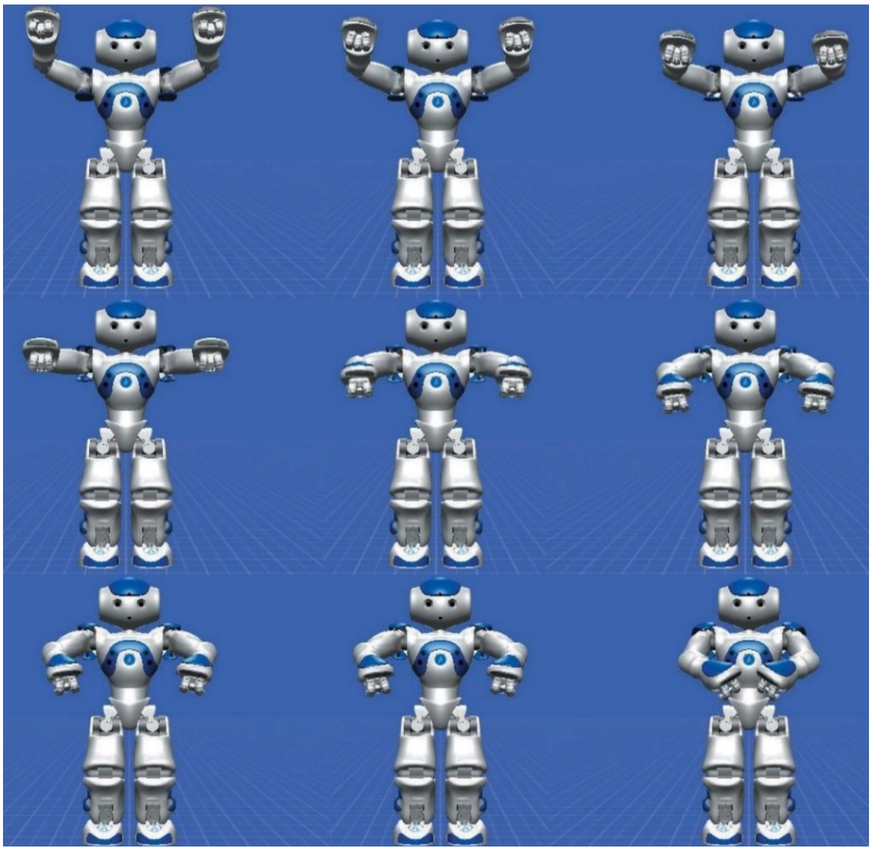


Fig. 3. Using inverse kinematics to control the continuous motion of minitype humanoid robots

3 Target Recognition Method Based on Monocular Vision

The image feature refers to the significant difference between the identified target and other objects, and the need to quantify the difference between them [10]. Usually a type of target can extract a number of features, they formed a number of quantitative values can be composed of a vector, that is the target eigenvector. So identify the target by determining whether the detected feature matches the target model. The color, corner, edge and so on of the monocular vision robot are concerned. The algorithm is optimized for the visual path tracking of the indoor robot.

3.1 Characteristics of Image Color

Color features can be described from subjective sensations and physical quantities. The color feature is used to express the color model. The commonly used color models include RGB model, CMY model and HSV model. Feature extraction methods based on different models have their own advantages and disadvantages.

The color model is to establish a mathematical model to represent the color. The basic idea is to establish a suitable three-dimensional coordinate system with different color components. For example, the RGB model can be used to create spatial coordinate systems by R, G, and B components. The type of color model comes from different needs. The demand for color models varies from human, electronic, and printing devices. People's perception of color is more suitable for using HSV color models. The display is more suitable for RGB mode. Therefore, by the RGB color model through a series of matrix conversion can get more color models to meet the needs of different algorithms. The appropriate color space affects the performance of the edge detection algorithm to a large extent.

In the cube of the RGB model, the color corresponding to the origin is black, and the color corresponding to the furthest from the origin is white. In the machine vision, the collected image pixels are compared with the color of the target in the RGB model, If the distance is within a sphere radius, the color of the pixel is considered to match the target color. Let the color of the pixel in the image be $\left(R_p, G_p, B_p\right)$, the color of the target is $\left(R_C, G_C, B_C\right)$. The radius from the center of the sphere is r. As shown in Fig. 4, the model is represented by a Cartesian coordinate system.

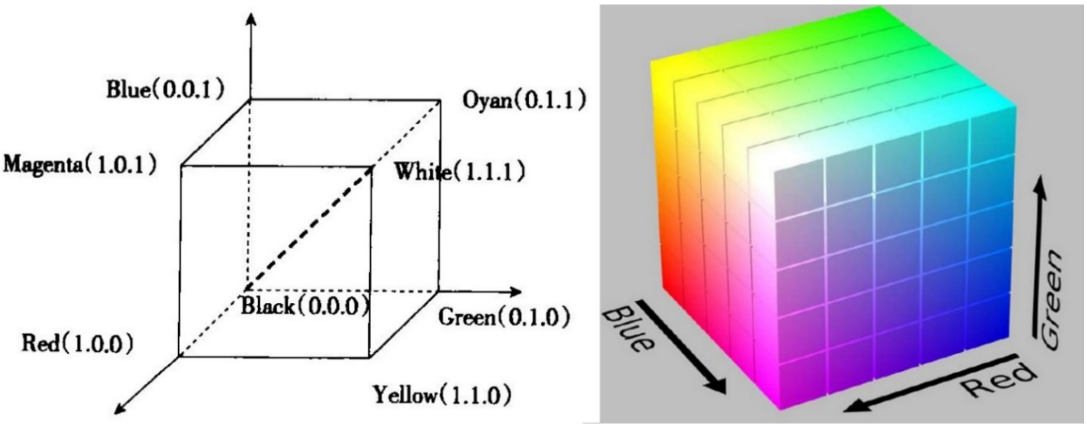


Fig. 4. RGB color model

$$y = \begin{cases} 1(\sqrt{(R_p - R_c)^2 + (G_p - G_c)^2 + (B_p - B_c)^2} \le r) \\ 0(\sqrt{(R_p - R_c)^2 + (G_p - G_c)^2 + (B_p - B_c)^2} \succ r) \end{cases}. \quad \textbf{(19)}$$

$y = 1$ indicates that the pixel color is the same as the target point color, and vice versa.

Although the RGB model is simple, most of the image device supports this mode. But the model is greatly affected by the light, any changes in light intensity will lead to complete failure of identification [11].

HSV (Hue, Saturation, Value) is a color space created by A. R. Smith in 1978 based on the intuitive nature of color, also known as Hexcone Model. The parameters of the color in this model are hue (H), saturation (S) and value (V).

The RGB and CMY color models are hardware-oriented, while the HSV (Hue Saturation Value) color model is user-oriented. The three-dimensional representation of the HSV model evolved from the RGB cube. Imagine that you can see the hexagonal shape of the cube from the white vertex of the cube to the black vertex. The hexagonal boundary represents the color, the horizontal axis represents the purity, and the brightness is measured along the vertical axis [12].

HSV models are commonly used in computer graphics applications [13]. The user must choose a color applied to a particular graphic element in various application environments often using HSV color wheel, as shown in Fig. 5. In which the hue is expressed as a ring. A separate triangle can be used to represent the saturation and brightness. Typically, the vertical axis of the triangle indicates saturation, while the horizontal axis represents the brightness. As shown in Fig. 6, the choice of color can first select the hue in the ring, then select the desired saturation and brightness from the triangle. Another visual method of the HSV model is the cone. In this representation, the hue is expressed as the angle about the central axis of the cone. The saturation is expressed as the distance from the center of the cross section of the cone to this point [14]. The brightness is expressed as the distance from the center of the cross section of the cone to the vertex. This method is more suitable for displaying this HSV color space in a single object. Because of its three-dimensional nature, it is not suitable for selecting colors in a two-dimensional computer interface. The HSV color space may also be represented as a cylinder similar to the above-mentioned cones. The hue varies along the outer circumference of the cylinder. The saturation varies along the distance from the center to the outside of the cross section. The brightness varies along the distance from the cross-section to the bottom and top. This representation may be considered a more accurate mathematical model of the HSV color space [15]. However, the number of levels of saturation and hue that can be distinguished in practice decreases as the brightness approaches black. In addition, the computer typically stores the RGB value with a limited precision range, which is limited the accuracy. Considering human color-aware constraints, making the cone representation more practical in most cases [16].

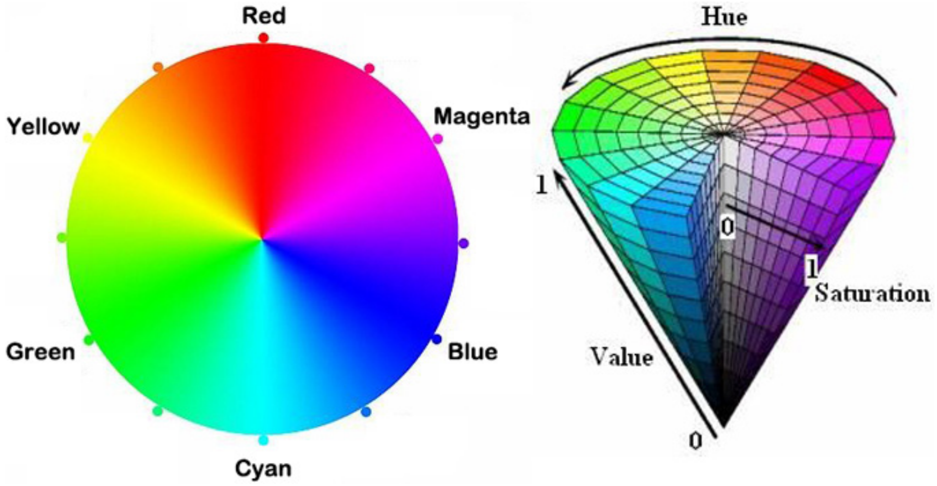


Fig. 5. HSV model

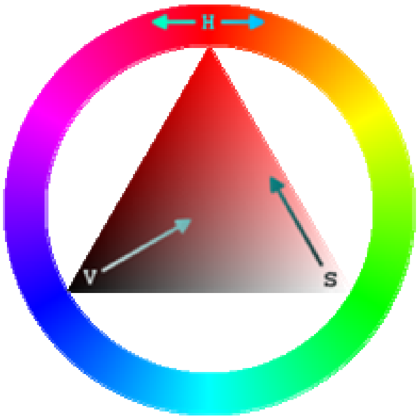


Fig. 6. HSV color wheel allows users to quickly select many colors

3.2 RGB to HSV

Let (R,G,B) are the red, green and blue coordinates of a color, and their values are real numbers between 0 and 1. Let max be the largest of r, g, and b. Let min be equal to the smallest of these values. To find the (h,s,v) value in the HSV space, where $h\in[0,360]$ is the hue angle of the angle, and $s,v\in[0,1]$ is the saturation and the brightness.

$$h=\begin{cases} 0^{\circ}, \mathit{if}\ \max=\min \\ 60^{\circ}\times\frac{g-b}{\max-\min}+0^{\circ}, \mathit{if}\ \max=randg\geq b \\ 60^{\circ}\times\frac{g-b}{\max-\min}+360^{\circ}, \mathit{if}\ \max=randg\prec b \\ 60^{\circ}\times\frac{g-b}{\max-\min}+120^{\circ}, \mathit{if}\ \max=g \\ 60^{\circ}\times\frac{g-b}{\max-\min}+240^{\circ}, \mathit{if}\ \max=b \end{cases}. \quad \textbf{(20)}$$

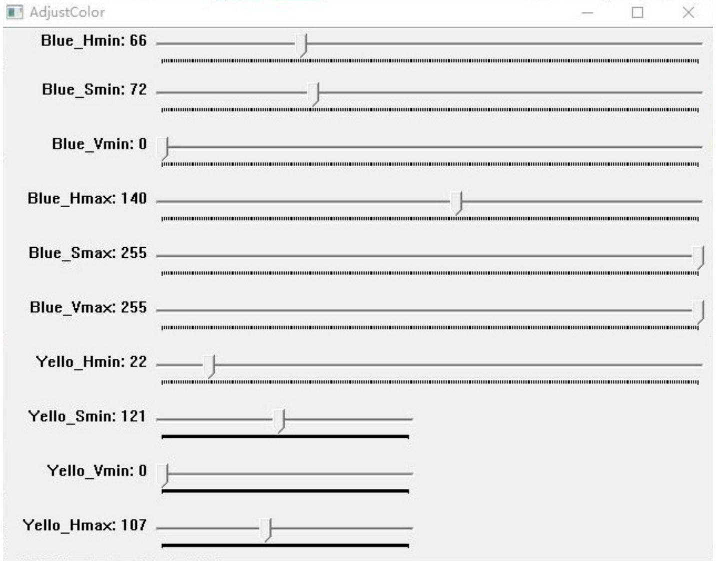
$$s=\begin{cases} 0, \mathit{if}\ \max=0 \\ \frac{\max-\min}{\max}=1-\frac{\min}{\max}, otherwise \end{cases}. \quad \textbf{(21)}$$

OpenCV has a function can be directly converted RGB model to HSV model, but in OpenCV $h\in[0,180]$, $s\in[0,255]$, $v\in[0,255]$. We know that the h component basically represents the color of an object, but the value of s and v must be within a certain range. s represents the degree of mixing of white and the color represented by h. The smaller the s, the whiter the color. v represents the degree of mixing of black and the color represented by h. The smaller the v, the darker the color. Experiments show that the value of blue is h between 100 and 140, and s and v are between 90 and 255.

4 Target Positioning Method Based on Monocular Vision

The optical imaging process of the camera is to process by simplifying it into a mathematical model. Therefore, the mathematical model (geometric imaging model) are set up to ensure a higher degree of accuracy [17]. At the same time, the establishment of the model will be affected by the parameters of the camera itself and the camera relative to the world coordinate system of geographical impact. The model is usually built by the principle of pinhole imaging. Moreover, the model of pinhole imaging is the linear imaging model, which does not meet the lens distortion caused by linear and non-linear distortion. The position of the camera in the world coordinate system also needs to be clear [17]. The internal parameters in the calibration are the parameters of the camera itself, such as the scale factor, the focal length and the lens distortion coefficient, which directly affect the effect of the camera to obtain the image. The external parameters can be understood as the process of determining the orientation and angle of the camera, where the spin matrix and the translation matrix are needed to be precisely known for the establishment of the model. The aim of the camera calibration is to establish the imaging mathematical model for the optical imaging process of the camera. Based on the model, the three-dimensional information in the world coordinate system can be deduced by the two-dimensional information in the image coordinates; can also be reversed.

Camera calibration refers to the given camera model for the process of obtaining the camera’s internal and external parameters. Camera imaging model is a mathematical model. It is an abstract vision-ranging algorithm for optical imaging process, and is implemented based on imaging. In the ideal case, the camera only need to consider the linear model. However, in fact, due to the camera production process or other reasons, the camera will cause linear or nonlinear distortion. In the daily modeling, we must consider the impact of these factors. As shown in Fig. 7.



Blue_Hmin: 66

Blue_Smin: 72

Blue_Vmin: 0

Blue_Hmax: 140

Blue_Smax: 255

Blue_Vmax: 255

Yello_Hmin: 22

Yello_Smin: 121

Yello_Vmin: 0

Yello_Hmax: 107

Fig. 7. RGB to HSV experiment screenshots

4.1 Pinhole Camera Model

When the object passes through the center point of the camera axis, the model projected onto the imaging surface is called the pinhole-imaging model. The center point of the camera's optical axis is the heart of the camera lens. As shown in Fig. 8, it is the center of the camera's optical axis, the plane of the object, the imaging plane of the camera. Since the plane looks like an inverted image, it converts its direction in the signal processing of the camera so that it can reverse again, the same as the objects up and down, left and right. Therefore, it can be considered that the imaging plane is equivalent to the image plane.

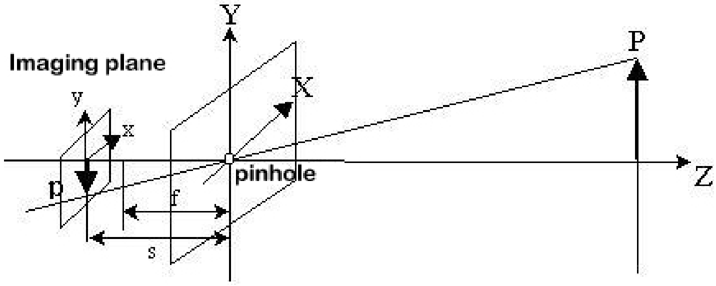


Fig. 8. Principle of pinhole imaging

The camera coordinate system is set up and the center point of the camera's optical axis is taken as the origin; the Z axis is parallel to the camera's optical axis, from the camera to the direction of the object is positive; the X axis is parallel to the vertical plane, with the upward direction positive. In the coordinates of the camera, set the object point P coordinates to (x, y, z), P in the plane C_2 of the imaging point P_2 coordinates to (x_2, y_2, z_2).

$$\begin{cases} \dfrac{x_1}{z_1} = \dfrac{x_2}{z_2} = \dfrac{x_2}{f} \\ \dfrac{y_1}{z_1} = \dfrac{y_2}{z_2} = \dfrac{y_2}{f} \end{cases} . \tag{22}$$

In the formula: f is the focal length of the camera, $f = P_2$.

4.2 The Internal Parameter Model of Camera

The camera's internal parameter model describes the relationship between the object points and the image pixels, while the equation 22 is the relationship between the object points and the imaging points in the camera coordinate system. The image on the imaging plane is enlarged and processed to obtain a digital image where the imaging point (x_2, y_2) is mapped to the pixel in the digital image (u, v). Let the intersection of the center of the optical axis and the imaging plane correspond to the coordinate in the digital image is (u_0, v_0).

$$\begin{cases} u - u_0 = a_x x_2 \\ v - v_0 = a_y y_2 \end{cases} . \tag{23}$$

a_x and a_y are the magnification coefficients of the plane-to-digital image plane in the X and Y directions, respectively.

Substituting formula 22 into formula 23.

$$\begin{cases} u - u_0 = a_x f \dfrac{x_1}{z_1} \\ v - v_0 = a_y f \dfrac{y_1}{z_1} \end{cases} . \tag{24}$$

Rewriting formula 24 into a matrix form 25.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1/z_1 \\ y_1/z_1 \\ 1 \end{bmatrix} = M_{in} \begin{bmatrix} x_c/z_c \\ y_c/z_c \\ 1 \end{bmatrix} . \tag{25}$$

In the formula 25, $k_x = a_x f$ and $k_y = a_y f$ are the magnification coefficients of the X and Y axis direction. M_m is called the internal matrix. (x_1, y_1, z_1) is the object point in the camera coordinate system coordinates, where the subscript is changed to (x_c, y_c, z_c).

In the formula 26, there are four parameters, which is called four-parameter model. In the formula 27, if the difference is not considered between k_x and k_y, there are only three parameters, which is called three-parameter model. In the formula 28, if the difference between k_x and k_y is considered, the internal model of the camera has five parameters called the five-parameter model.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c/z_c \\ y_c/z_c \\ 1 \end{bmatrix} . \tag{26}$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k & 0 & u_0 \\ 0 & k & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c/z_c \\ y_c/z_c \\ 1 \end{bmatrix} . \tag{27}$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & k_y & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c/z_c \\ y_c/z_c \\ 1 \end{bmatrix}. \tag{28}$$

4.3 The Internal Parameter Model of Camera

The description of the object coordinate system in the camera coordinate system is called the external parameter model of the camera. The object coordinate system $O X_w Y_w Z_w$ is represented in the camera coordinate system $O_c XYZ$, which constitutes the external matrix of the camera.

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} r_x & s_x & t_x & p_x \\ r_y & s_y & t_y & p_y \\ r_z & s_z & t_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \begin{bmatrix} R & P \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = M_w \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}. \tag{29}$$

(x_c, y_c, z_c) is the coordinates of the object in the camera coordinate system $O_c XYZ$. (x_w, y_w, z_w) is the coordinates of the object in the object coordinate system $O X_w Y_w Z_w$. M_w is the external parameter matrix.

5 Conclusions

In this chapter, the experiment of mobile manipulation of a minitype humanoid robot is done based on the study of the various parts of the mobile manipulation. It describes the process of the mobile manipulation of the minitype humanoid robot in detail, and validates the validity and feasibility of the previous chapters.

5.1 Mobile Manipulations Based on Monocular Vision

Target search. Only when minitype humanoid robots can see the target, we can track the target. We put the minitype humanoid robot in an unknown job space and plan to let it search for the target. There are two stages of the search, the first stage is to stand still in the search target, just swing the head, when this stage does not search the target, the minitype humanoid robot begins to circle search target. If the robot still cannot find the target, it will search the way as described above until the target is searched. Search and approach the process is shown in Fig. 9.

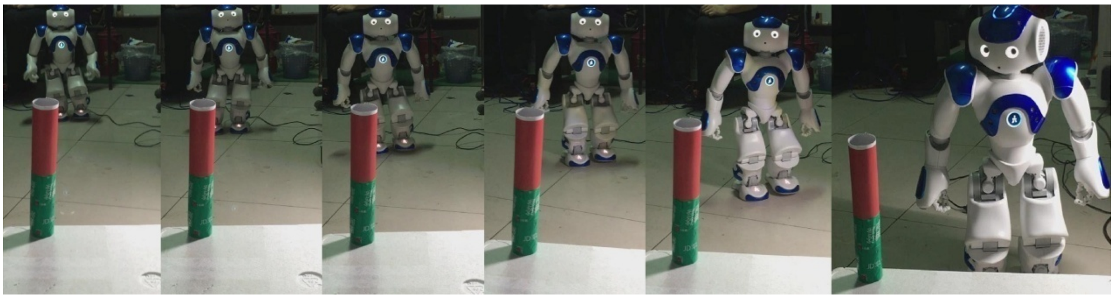


Fig. 9. Process of search and approach

Target positioning. The camera of minitype humanoid robot can measure a wide range. In the vertical direction, it can measure the upper edge to the middle line of 23.82°. In the horizontal direction, it can measure the upper edge to the middle line of 30.45°. It first identifies the target, at this stage, obtains the target spatial coordinates (x, y, z). The coordinates of the center of the camera lens are (320,240), the coordinates of the coordinates come from the resolution of the camera when lens collect images. The solution algorithm is as follows.

The center of the target and the camera's main axis has an angle of α, which is horizontal.

$$\alpha = X\ axis\ direction\ ratio \times Horizontal\ visual\ range \times \frac{\pi}{180} \ . \tag{30}$$

$$X\ axis\ \ ratio = \frac{x-320}{640} \ . \tag{31}$$

$$Horizontal\ \ field\ view = 60.9° \ . \tag{32}$$

$$\alpha = \frac{x-320}{640} \times 60.9° \times \frac{\pi}{180} \ . \tag{33}$$

After the target is identified, the coordinates (x, y, z) of the target are calculated by a series of algorithms, substituting into formula 33, and finally get the camera lens center in the horizontal direction of the angle α. The α is passed to the minitype humanoid robot to control the moving correlation function, and control the minitype humanoid robot towards the target.

$$\beta = \frac{y-240}{480} \times 47.64° \times \frac{\pi}{180} . \tag{34}$$

Y is the y in the center coordinates (x, y, z) of the target recognition through a series of algorithms. It is possible to obtain the angle β of the target at the center of the camera lens by substituting it into the formula 34. The β is in the vertical direction, and specifies the value of the head joint by the β from the head, ensuring that the target is always at the midpoint of the camera lens.

Target Tracking. In the experiment, we used a very simple but very efficient algorithm to track moving objects. This algorithm enhances the autonomy of mobile manipulations.

We use the midpoint of the bottom of the target minimum rectangle to represent the target. As shown in Fig. 10, when the abscissa of the target after moving is greater than the abscissa before the movement, it means moving to the right. When the abscissa of the target after moving is less than the abscissa before the movement, it means moving to the left. When the coordinates of the target after moving is greater than the ordinate before the movement, it represents the upward movement. When the moving target is less than the ordinate before the movement, it represents the downward movement.

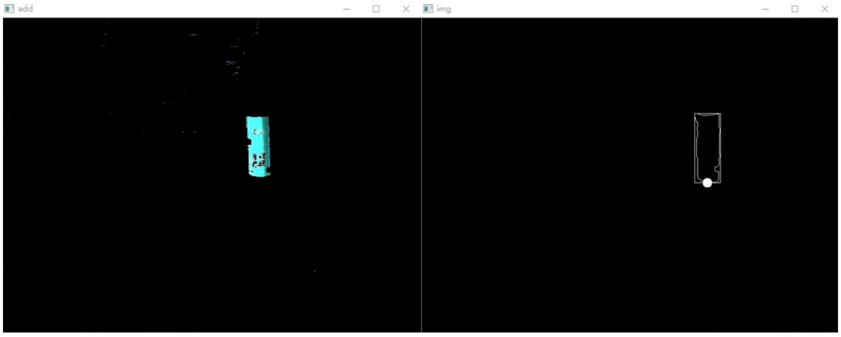


Fig. 10. Identify the midpoint of the smallest rectangular bottom edge

Monocular vision ranging. Other visual measurement methods are based on monocular vision measurements, for example, single cameras take multiple targets from different angles, and the effect is similar to using multidimensional vision. Monocular vision can use simpler devices than multi eye vision, thus it reduces the cost, runs fast and has great flexibility. The most basic method avoids the difficulty of stereo matching, but the multi eye vision cannot avoid this problem. Monocular vision has a better real-time than the multi-vision, so in this context, it is widely used.

With the progress of science and technology, monocular vision is also rapid development. There are many different ways, for example, optical structure method is widely used in the measurement of high precision requirements of the parts, geometric constraint method can measure the posture state of the movement, and even intelligent aircraft, geometric similarity method is used to measure the height of the target.

Referring to the principle of monocular vision ranging, a distance-measuring algorithm is designed based on the hardware structure and parameters of minitype humanoid robot. Distance-measuring algorithm is shown in Fig. 11.

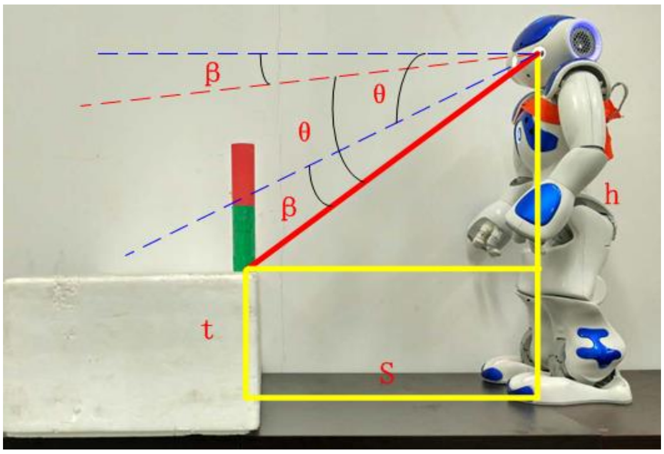


Fig. 11. Distance-measuring algorithm

$$\tan(\beta+\theta)=\frac{h-t}{s}. \tag{35}$$

The transformation formula is available.

$$S=\frac{h-t}{\tan(\beta+\theta)}. \tag{36}$$

h is the height of the robot neck to its sole. β is the current offset angle of the Head Pitch joint. θ is the natural angle of the NAO robot mouth camera and the horizontal line. t is the height of the table. S is the distance between the target center and the robot. $t=459.59mm$, $\theta=39.7°$.

$$S=\frac{0.45959-t}{\tan(\beta+39.7°)}. \tag{37}$$

We can calculate the target center to the minitype humanoid robot near the S. This calculation is based on the robot head angle in the vertical direction. S is obtained only based on the algorithm. Only need to determine the angle can get an order of forward, backward or stop, so that can make the target keep S meters away from the robot. The result of this is to reduce the cost of human time.

5.2 Experimental Results and Analysis

First, the minitype humanoid robot searches and identifies the target. The visual sensor identifies the target by the features of the image color. Secondly, the target is positioned through the camera modeling and parameter calibration. The target position coordinates is got. Then, model the mobile manipulation of the whole body kinematics is established. The robot is moved to the target position by kinematics algorithm. Finally, the kinematics model and control algorithm of arm are used to grab the objects, and the mobile manipulation experiment of minitype humanoid robot is completed.

The robot identifies the target and moves towards the target. The experiment is shown in Fig. 12.

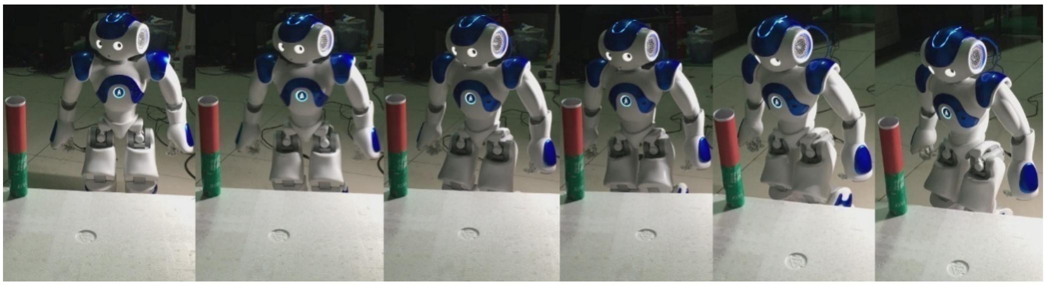
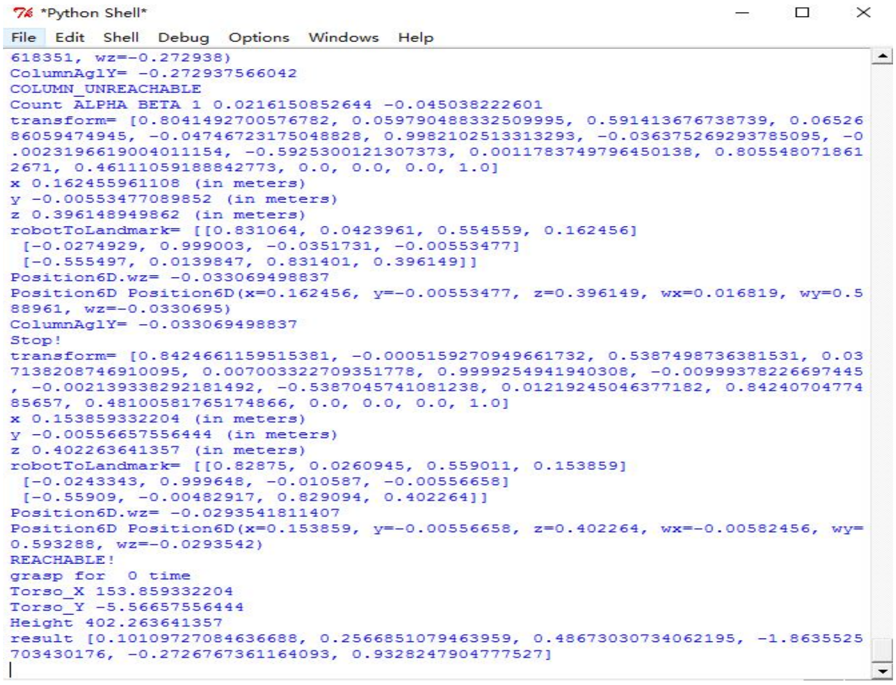


Fig. 12. Identify the goal and move towards the goal

In the research experiment of mobile manipulation control method of the minitype humanoid robot based on monocular vision, the minitype humanoid robot has been adjusted more than sixty times, so that the distance of the minitype humanoid robot to the target is more suitable for the crawling operation. The best distance from the target is S. Fig. 13 shows the real-time data capture of the optimal distance from the target for the minitype humanoid robot in the mobile manipulation. Fig. 14 is the line map of real-time data. It is possible to conclude that the minitype humanoid robot is closer to the target than the distance of 15cm, which is most suitable for crawl distance S of the minitype humanoid robot.



```
618351, wz=-0.272938)
ColumnAglY= -0.272937566042
COLUMN_UNREACHABLE
Count ALPHA BETA 1 0.0216150852644 -0.045038222601
transform= [0.8041492700576782, 0.05979048833250995, 0.591413676738739, 0.06526
86059474945, -0.04746723175048828, 0.9982102513313293, -0.03637526929378509 5, -0
.0023196619004011154, -0.5925300121307373, 0.0011783749796450138, 0.805548071861
2671, 0.4611105918884277 3, 0.0, 0.0, 0.0, 1.0]
x 0.162455961108 (in meters)
y -0.00553477089852 (in meters)
z 0.396148949862 (in meters)
robotToLandmark= [[0.831064, 0.0423961, 0.554559, 0.162456]
 [-0.0274929, 0.999003, -0.0351731, -0.00553477]
 [-0.555497, 0.0139847, 0.831401, 0.396149]]
Position6D.wz= -0.033069498837
Position6D Position6D(x=0.162456, y=-0.00553477, z=0.396149, wx=0.016819, wy=0.5
88961, wz=-0.0330695)
ColumnAglY= -0.033069498837
Stop!
transform= [0.8424661159515381, -0.0005159270949661732, 0.5387498736381531, 0.03
713820874691009 5, 0.0070033227093517 78, 0.9999254941940308, -0.009993782269744 5
, -0.002139338292181492, -0.5387045741081238, 0.012192450463771 82, 0.84240704774
85657, 0.4810058176517486 6, 0.0, 0.0, 0.0, 1.0]
x 0.153859332204 (in meters)
y -0.0055665755644 4 (in meters)
z 0.402263641357 (in meters)
robotToLandmark= [[0.82875, 0.0260945, 0.559011, 0.153859]
 [-0.0243343, 0.999648, -0.010587, -0.00556658]
 [-0.55909, -0.00482917, 0.829094, 0.402264]]
Position6D.wz= -0.0293541811407
Position6D Position6D(x=0.153859, y=-0.00556658, z=0.402264, wx=-0.00582456, wy=
0.593288, wz=-0.0293542)
REACHABLE!
grasp for  0 time
Torso_X 153.859332204
Torso_Y -5.56657556444
Height 402.263641357
result [0.10109727084636688, 0.2566851079463959, 0.48673030734062195, -1.8635525
703430176, -0.2726767361164093, 0.9328247904777527]
|
```

Fig. 13. Minitype humanoid robot adjusts the real-time data of the best distance process

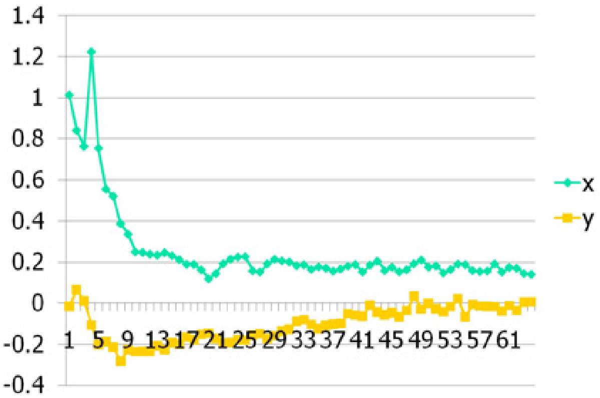


Fig. 14. Minitype humanoid robot adjusts the real-time data of the best distance process

The minitype humanoid robot moves to the target and performs the crawling operation. The robot has successfully captured the target and verified the accuracy of the kinematic model of the minitype humanoid robot we established, as shown in Fig. 15. At the same time, it also verifies the correctness from target recognition, positioning to the whole body kinematics analysis, arm kinematics modeling and inverse kinematics solving of the minitype humanoid robot. It shows that the research of target recognition and mobile manipulation control method of minitype humanoid robot is successful.

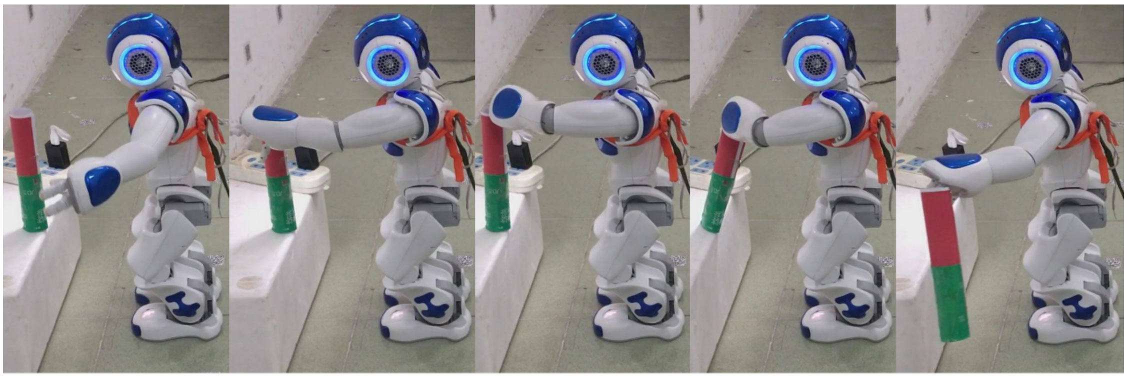


Fig. 15. Minitype humanoid robot completes the mobile manipulation successfully

Acknowledgements

This work is supported by NSFC under Grant NO.61473027, also by Beijing Key Laboratory for Biomimetic and Function of Robot Grant BZ0337.

References

[1] S. Keizer, P. Kastoris, M.E. Foster, A. Deshmukh, O. Lemon, Evaluating a social multi-user interaction model using a Nao robot, in: Proc. IEEE International Symposium on Robot and Human Interactive Communication, 2014.

[2] I. Rodriguez, A. Astigarraga, E. Jauregi, T. Ruiz, E. Lazkano, Humanizing NAO robot teleoperation using ROS, in: Proc. International Conference on Humanoid Robots. IEEE, 2014.

[3] Z. Sun, N. Roos, An energy efficient dynamic gait for a Nao robot, in: Proc. IEEE International Conference on Autonomous Robot Systems and Competitions, 2014.

[4] Nguyen T L, Boukezzoula R, Coquin D, E. Benoit, S. Perrin, Interaction between humans, NAO robot and multiple cameras for colored objects recognition using information fusion, in: Proc. International Conference on Human System Interactions, 2015.

[5] O. Melinte, L. Vladareanu, L. Munteanu, H. Yu, S. Cang, Z.-G. Hou, G.-B. Bian, H. Wang, Haptic intelligent interfaces for NAO robot hand control, in: Proc. International Conference on Advanced Mechatronic Systems, 2015.

[6] S. Wen, Z. Ma, S. Wen, Y. Zhao, J. Yao, The study of NAO robot arm based on direct kinematics by using D-H method, in: Proc. Control (CONTROL), 2014 UKACC International Conference on. IEEE, 2014.

[7] W. Dudek, K. Banachowicz, W. Szynkiewicz, T. Winiarski, Distributed NAO robot navigation system in the hazard detection application, in: Proc. IEEE International Conference on Methods and MODELS in Automation and Robotics, 2016.

[8] C. Li, C. Yang, P. Liang, A. Cangelosi, J. Wan, Development of Kinect based teleoperation of Nao robot, in: Proc. International Conference on Advanced Robotics and Mechatronics, 2016.

[9] O. Tutsoy, D.E. Barkana, S. Colak, Learning to balance an NAO robot using reinforcement learning with symbolic inverse

kinematic, Transactions of the Institute of Measurement & Control, 2016.

[10] E. Dong, D. Wang, C. Chen, J. Tong, Realization of biped robot gait planning based on NAO robot development platform, in: Proc. IEEE International Conference on Mechatronics and Automation, 2016.

[11] Rodríguez I, Aguado A., O. Parra, E. Lazkano, B. Sierra, NAO Robot as Rehabilitation Assistant in a Kinect Controlled System, in: Proc. the 3rd International Conference on NeuroRehabilitation, 2016.

[12] J.A. Rincon, A. Costa, P. Novais, V. Julian, C. Carrascosa, Detecting Social Emotions with a NAO Robot, in: Proc. 14th International Conference, PAAMS 2016, 2016.

[13] J. Avalos, S. Cortez, K. Vasquez, V. Murray, O.E. Ramos, Telepresence using the kinect sensor and the NAO robot, in: Proc. IEEE, Latin American Symposium on Circuits & Systems, 2016.

[14] A.A.B. Ka, A.K. Mullapudi, D. Ebert, N. Phadnis, R. Middha, HTKS Game for executive functions disorder using NAO robot, in: Proc. ACM International Conference on Pervasive Technologies Related to Assistive Environments, 2016.

[15] S.M. Hadfield, C.S. Coulston, M.G. Hadfield, L.B. Warner, Adventures in K-5 STEM outreach using the NAO robot, in: Proc. ACM Technical Symposium on Computing Science Education, 2016.

[16] A. Bolotnikova, K. Tarvas, G. Anbarjafari, Goalpost and Ball Detection and Self-Localisation Systems of NAO Robot, LAMBERT Academic Publishing, Düsseldorf, 2016.

[17] Nguyen T L, Boukezzoula R, Coquin D, S. Perrin, Color recognition for NAO robot using Sugeno fuzzy system and evidence theory, in: Proc. 2015 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology (IFSA-EUSFLAT-15), 2015.