International Journal of Control and Automation
Vol. 9, No. 4 (2016), pp. 367-380
http://dx.doi.org/10.14257/ijca.2016.9.4.36

The Master Manipulator Control System in a Master-Slave Robot Based on Trinocular Stereo Vision

Qiong Liu*, YaGuang Zhu, Xu Zhang, YanPu Yang and Ke Li

Chang'An University, Key Laboratory of the Ministry of Education On Road Works Technology & Equipment, Xi'an, ShaanXi
48062164@qq.com

Abstract

In this paper, a master manipulator control system in a master-slave robot which uses trinocular stereo vision is presented. The system utilizes three real-time cameras to identify and track the movements of the master handle, while also measuring the master handle position. The system follows a basic four step process. First, the master handle in the image is identified using the contrast between the handle and the background. Second, the speed and direction of the next position for the master handle is predicted in real-time by using the Kalman filter. Third, the position of the master handle is measured in real-time using stereo matching, which utilizes the method of combining feature matching and region gray matching to consistently identify the markers on the master handle and eliminate error matching. The experimental results in this study conclude that the master-handle manipulator control system is fast and exact.

Keywords*: Trinocular vision, Master-slave robot, Kalman filter, Pose measurement*

1. Introduction

The master-slave robot is an important research direction in the field of robotics, particularly because of its potential use and advantages in hazardous or difficult to access environments. In general, a master-slave robot consists of the master handle and the slave robot. The master handle is an interactive tool through which the operator exchanges information to the slave robot. Domestic and foreign scholars have proposed a variety of master handle configurations. Farrington and Gerling [1] proposed a Phantom-type master handle with a multi-dimensional force feedback which can simulate the forces in a virtual environment. Hung [2, 3] proposed the 6-DOF master handle based on the parallel connection mechanism which is known for its fast motion in a larger workspace. Baser [4] proposed a high dynamics master handle based on the dual parallel connection mechanism, which utilizes the method of combining the rigid drive and flexible drive to improve flexibility. Berkelman [5] proposed a non-contact magnetic levitation master handle that can achieve 6-DOF motion and force feedback. All these master handle configurations have a complex mechanical structure which limits their functionality in a limited working space.

In recent years, there have been significant developments in visual detection technology. More specifically, vision detection has become faster and more accurate in its measurements, and become increasingly vital in tracking moving targets. Gao [6] studied human fingertip motion, which can be used for real-time tracking and recognition by the dexterous master-slave operation. Nash [7] studied the tracking of moving objects through 3D space, the tracking system can adjust automatically the position and focus of camera, and can find the target that moves beyond the field of view. Visual detection have some advantages over other tracking methods because it is non-contact, high speed, and highly flexible, which make it easier to measure a moving object through a large space.

ISSN: 2005-4297 IJCA
Copyright⓪ 2016 SERSC

In this study, the movement of the master handle is identified and tracked by three cameras. The position of the master handle is measured by trinocular stereo vision, and then the position and orientation of the master handle is sent to the slave robot. Therefore, the master handle can control the movement of slave robot. In this way, the complex mechanical structure of the master-slave robot can be simplified, and the workspace can be broadened.

2. The Structure of the Master-Slave Robot System

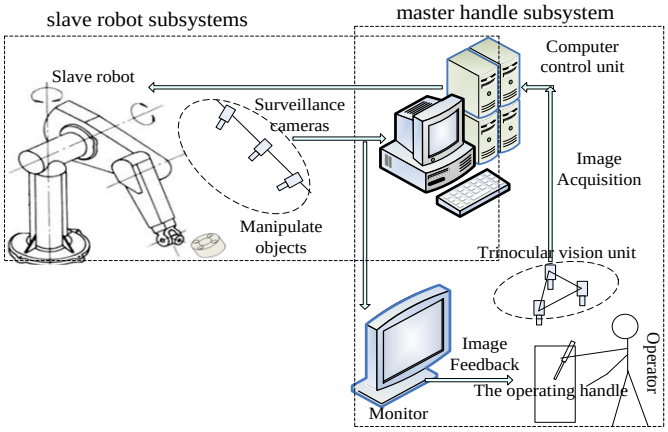


Figure 1. The Structure of Master-Slave Robot System

The structure of the master-slave robot system is shown in Figure 1. The operator controls the movement of the operating handle. The position and posture of the operating handle must be tracked and measured in real-time by the visual detection unit. Through the move mapping function of master-slave system, the position and posture of the operating handle is converted into joint movement, and then the joint movement are used for driving the robot as a corresponding movement of the operating handle. Meanwhile, the operator can make a decision on the operating handle's next movement after reviewing the real-time video. Thus, this operational step forms a work cycle.

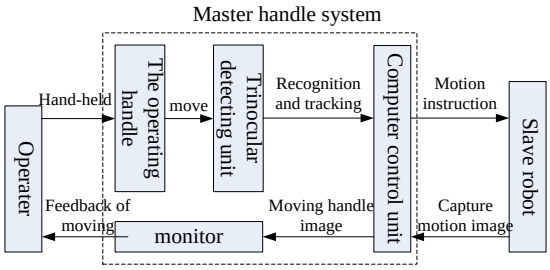


Figure 2. The Block Diagram of the Master Handle Subsystem

3. The Design of the Master Handle Subsystem

The main function of the master handle subsystem is to facilitate a human-computer interaction. The visual detection unit collects information from the motion of operating handle, which is then used to instruct the motion of the robot through the computer control unit. The operator can move the operating handle into the next position using the real-time feedback of the moving robot. The working block diagram of the master handle subsystem is shown in Figure 2.

 Copyright⓪ 2016 SERSC

3.1. The Workspace of the Master Handle

The workspace of the master handle is formed by the field of view of multi-cameras. In order to avoid concealing the operating handle because of the operator's arm, it's better for us to adopt three cameras from three different views to track and locate the motion of the operating handle, as shown in Figure 3.

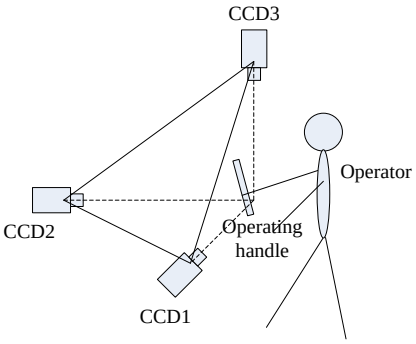


Figure 3. The Multi-Camera Layout

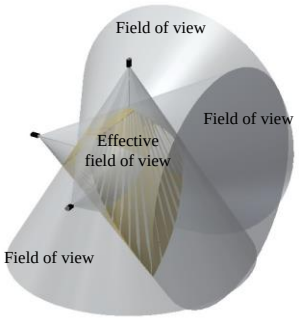


Figure 4. The Overlap Field of Three Camera Views

Each camera exists one field of view, three fields of view overlapped each other at different views, as shown in Figure 4. Thus, the workspace of the operating handle is formed by the overlap, and that is a 3D measure space with a special shape and an unknown position and size [8, 9]. Therefore, it is difficult to calculate the location and size of the workspace. In order to simplify this process, the field of view of camera is expressed in conical space, as shown in Figure 5.

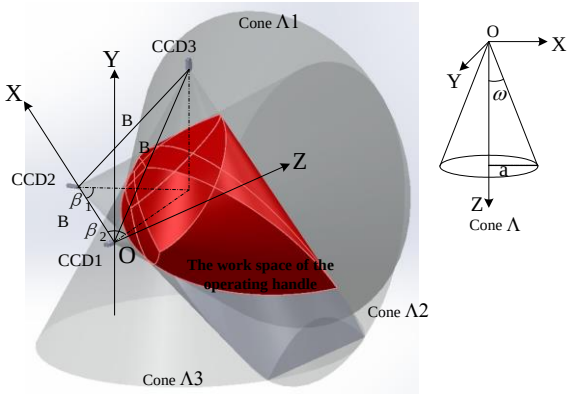


Figure 5. The Shape and Position of the Workspace

As shown in the right of Figure 5, the conical space is the field of view of one camera. The conical surface parametric equation is $\Lambda:\begin{cases} x = a\cdot\cos t \\ y = a\cdot\sin t \\ z = a\cdot\cot\omega \end{cases}$ with the directrix $x^2 + y^2 = a^2$, $z = a\cdot\cot\omega$, including $t \in R$, $\omega = \arctan(\sqrt{T_x^2 + T_y^2}/f$.

Then, the cone rotates around Y-axis through a clockwise (90°-β_1) degrees. So, the rotated factor matrix of CCD1conical surface in the coordinate system OXYZ is:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} \cos(90° - \beta_1) & 0 & -\sin(90° - \beta_1) \\ 0 & 1 & 0 \\ \sin(90° - \beta_1) & 0 & \cos(90° - \beta_1) \end{bmatrix} \cdot \begin{pmatrix} a\cos t \\ a\sin t \\ a\cot\omega \end{pmatrix}$$

Therefore, the CCD1 conical surface parametric equation is:

$$\Lambda 1: \begin{cases} x = a\sin\beta_1\cos t - a\cos\beta_1\cot\omega \\ y = a\sin t \\ z = a\cos\beta_1\cos t + a\sin\beta_1\cot\omega \end{cases} \tag{1}$$

Similarly, the cone rotates around the Y-axis through a counter-clockwise (90°-β2) degrees, and shifts B along the X-axis. Thus, it is understood that in the coordinate system OXYZ the rotated factor matrix of the CCD2 conical surface is:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} \cos(-(90° - \beta_2)) & 0 & -\sin(-(90° - \beta_2)) \\ 0 & 1 & 0 \\ \sin(-(90° - \beta_2)) & 0 & \cos(-(90° - \beta_2)) \end{bmatrix} \cdot \begin{pmatrix} a\cos t \\ a\sin t \\ a\cot\omega \end{pmatrix} + \begin{pmatrix} B \\ 0 \\ 0 \end{pmatrix}$$

Therefore, the CCD2 conical surface parametric equation is:

$$\Lambda 2: \begin{cases} x = a\sin\beta_2\cos t + a\cos\beta_2\cot\omega + B \\ y = a\sin t \\ z = -a\cos\beta_2\cos t + a\sin\beta_2\cot\omega \end{cases} \tag{2}$$

Then, the cone rotates around X-axis through a clockwise 90°degrees, shifts B/2 along the X-axis, shifts B/2 along the Z-axis and shifts B/$\sqrt{2}$ along the Y-axis. It is understood, then, that in the coordinate system OXYZ the rotated factor matrix of the CCD3 conical surface is:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} \cos 90° & 0 & -\sin 90° \\ 0 & 1 & 0 \\ \sin 90° & 0 & \cos 90° \end{bmatrix} \cdot \begin{pmatrix} a\cos t \\ a\sin t \\ a\cot\omega \end{pmatrix} + \begin{pmatrix} B/2 \\ B/\sqrt{2} \\ B/2 \end{pmatrix}$$

Therefore, the CCD3 conical surface parametric equation is:

$$\Lambda 3: \begin{cases} x = -a\cot\omega + B/2 \\ y = a\sin t + B/\sqrt{2} \\ z = a\cos t + B/2 \end{cases} \tag{3}$$

Assuming that the depth of field is in perfect condition for the three cameras CCD1, CCD2 CCD3, the effective field of view for the trinocular vision is consists of a cluster of points Ω which lie in the overlapping field of the three camera views. The cluster of points Ω is expressed by the formula (1) to (3):

$$\Omega = \{(x, y, z) / (x, y, z) \subseteq (\Lambda 1 \cap \Lambda 2 \cap \Lambda 3)\} \tag{4}$$

To review, the workspace of the operating handle is determined by the set of points in the overlap zone.

4. Recognition and Tracking Algorithm of the Master Handle

The background difference method is an efficient recognition algorithm used to track a moving object. In general this process begins by selecting one or several images as a background template and then subtracting in the chosen image from the background

 Copyright⓪ 2016 SERSC

template by using this formula (5). If the number of nonzero pixels in the differential image is greater than the preset threshold value, it shows a moving object in the image. Let $f_k(x,y,t)$ and $f_b(x,y,t)$ denote the current frame image at time t and the background template separately, so that the differential image D_k is:

$$D_k(x,y,t) = \begin{cases} f_k(x,y,t) & \left|f_k(x,y,t) - f_b(x,y,t)\right| > T \\ 0 & others \end{cases} \quad (5)$$

In each of these formulas, T is the set threshold value; the nonzero region in D_k is the region of the moving target. Because the background difference method can be influenced by environmental change, light intensity, and other factors, it is necessary to promptly update the background templates in an effort to nullify these influential effects. The background-updating algorithm proposed by Lamard[10] is:

$$B_k(x,y) = \alpha f_k(x,y) + (1-\alpha)B_{k-1}(x,y) \quad (6)$$

Wherein α is the updating weights of the background templates that is used for adjusting the updating rate of the background templates.

4.1. Motion Tracking Based on Kalman Filter

After using background difference method to determining the moving target, a maximum bounding rectangle W_{ROI} of the target contour is used as the window for the motion tracking. Then, in order to narrow the moving target search area, the next location of the moving target must be predicted with the Kalman filter according to the current status (position, velocity, acceleration) of the moving target.

Suppose the status vector of the Kalman filter is $X=(x,y,v_x,v_y,\alpha_x,\alpha_y)^T$, which separately denotes position, velocity and acceleration for the x, y-direction in the image; and the observation vector is $Z=(x,y)^T$. The prediction algorithm is shown, then, as the following steps:

(1) Initialization: the model of the Kalman filter is:

State equation:

$$X_{k+1} = \Phi_k X_k + \omega_{k+1}, \quad \omega_{k+1} \in N(0,Q) \quad (7)$$

where $ø_k$ is a state transition matrix; w_k is a random interference vector.
Observation equation:

$$Z_{k+1} = H_k X_k + \upsilon_{k+1}, \quad \upsilon_{k+1} \in N(0,R) \quad (8)$$

where H_k is an observation matrix; v_{k+1} is a observation noise vector.

The Kalman filter should be initialized when it is used for the first time. The detected initial state X_0 of the moving target is assigned to the Kalman filter as the initial value. If the velocity and acceleration of the moving target is unknown, the assignment is zero. Meanwhile, the time of the current image is recorded and the covariance of the initial error is set to zero.

(2) The prediction of moving target:

Status prediction equation:

$$X_{k+1} = \Phi_k X_k + \omega_{k+1} \quad (9)$$

Error covariance prediction equation:

$$P_{k+1}^{-} = \Phi_k P_k \Phi_k + Q \quad (10)$$

Copyright ⓒ2016 SERSC

$$\Phi_k = \begin{bmatrix} I & \Delta t \cdot I & \frac{1}{2}\Delta t^2 \cdot I \\ 0 & I & \Delta t \cdot I \\ 0 & 0 & I \end{bmatrix} \tag{11}$$

Where, Δt is the sampling time, ϕ_k is a 6 × 6 state transition matrix, I is a 2 × 2 unit matrix.

(3) Update process:

Observation matrix:

$$H_{k+1} = [1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0] \tag{12}$$

Kalman gain coefficient:

$$K_{k+1} = P_{k+1}^{-} H_{k+1}^{T} (H_{k+1} P_{k+1}^{-} H_{k+1}^{T} + R)^{-1} \tag{13}$$

Covariance update equation:

$$P_{k+1} = (1 - K_{k+1} H_{k+1}) P_k^{-} \tag{14}$$

Status update equation:

$$X_{k+1} = X_{k+1} + K_{k+1} (Z_{k+1} - H_{k+1} X_{k+1}) \tag{15}$$

5. Gesture Measurement Algorithm

The gesture of the operating handle can be determined by measuring the coordinates of certain marks on the handle. However, markers on the handle will be concealed when the operator move the handle. One or two cameras cannot acquire markers image. To accommodate visual measurement, the trinocular detecting unit is adopted to calculate the handle gesture as long as two of cameras are able to detect four or more markers. Here's an example in which any two cameras can be used for the algorithm.

5.1. Recognition and Location of Markers

The Image binarizing process by iterative method: the number of the markers N_0 (Known) is viewed as a condition for the ending iteration. The choice of the threshold T_0 is:

$$T_0 = (T_{\max} + T_{\min}) / 2 \tag{16}$$

where $T_{\max}$, $T_{\min}$ are separately the initial maximum, minimum threshold value. Assume that the max number of markers detected is N. In order to reduce marker loss, it should be that $N=\lambda N_0$, λ is greater than 1. First, the image is binarized by the threshold T_0; and then the contour number n is detected, if $n > N$, it takes $T_{\min}=T_0$; if $n < N$, it takes $T_{\max}=T_0$; if $n = N$, it ends the iteration and the final T_0 is the selected threshold.

After the binarization processing, the contour extraction will begin. Because the markers are round, and the marker's size and the relative position are known quantities, the pseudo markers in the image can be removed according to the shape of the marker's contour, the size, and the position. With this information the real markers can be determined. The real markers are extracted at the sub-pixel level by the centroid method [11]; the image coordinates of the centroid for each marker are denoted by $\{P_{Li} \mid i = 1, 2, \cdots N_0\}$, $\{P_{Rj} \mid i = 1, 2, \cdots N_0\}$.

 Copyright⓪ 2016 SERSC

5.2. The Matching of the Markers

During the marker matching process, the marker centroids are taken as the matching primitive of the feature, and the relevant neighborhood windows of the centroids are taken as the matching primitive of the gray area.

Suppose the relevant neighborhood windows of the centroids are denoted by $\boldsymbol{W}$; the search area is $\boldsymbol{W_{ROI}}$ (which was determined in the previous section); and the matching algorithm is:

(1) The initial matching of the feature points is taken by the epipolar constraint and then the candidate points are obtained. Suppose $\mathbf{P_i}$, $\mathbf{P_j}$ represent separately the corresponding point of the feature in the source image and the target image. $\mathbf{L_i}$ represents the corresponding epipolar of $\mathbf{P_i}$ in the target image. $\mathbf{D_{ij}}$ represents the distance between $\mathbf{P_j}$ and $\mathbf{L_i}$. $\mathbf{A_i}$ represents a cluster of the candidate points for $\mathbf{P_i}$.

$$D_{ij} = f(P_i, P_j, L_i), \tag{17}$$

Suppose the distance threshold value is $\mathbf{D_0}$, if$\mathbf{D_{ij}} < \mathbf{D_0}$, then $\mathbf{P_j}$ will be added to the cluster $\mathbf{A_i}$.

(2) Because of the similarity of the gray value within the neighborhood windows around the feature points, the matching might not be fully accurate. This can be described the by the normalized cross-correlation coefficient:

$$S(x,y) = \sum_{y'=0}^{n-1}\sum_{x'=0}^{m-1}(T(x',y')-\bar{T})(I(x+x',y+y')-\bar{I}) \Bigg/ \sqrt{\sum_{y'=0}^{n-1}\sum_{x'=0}^{m-1}(T(x',y')-\bar{T})^2 \sum_{y'=0}^{n-1}\sum_{x'=0}^{m-1}(I(x+x',y+y')-\bar{I})^2} \tag{18}$$

Where, S($\boldsymbol{x}$, $\boldsymbol{y}$) is the similarity of two $\boldsymbol{m} \times \boldsymbol{n}$ corresponding neighboring windows, I($\boldsymbol{x}$, $\boldsymbol{y}$), T($\boldsymbol{x}$, $\boldsymbol{y}$) are the gray value in the target image, and the source image at coordinate ($\boldsymbol{x}$, $\boldsymbol{y}$), $\bar{I}$ and $\bar{T}$ are the gray average within the neighborhood window in the target image and the source image. Suppose the similarity threshold is ε_0. If the similarity of the point-to-point (P_{Li}, P_{Rj}) is $S_{ij} > \varepsilon_0$, the candidate points are retained or excluded.

(3) Sequential consistency constraints and unique constraints are introduced to eliminate spurious matches. In the above matching process, a match must be bidirectional, which means that a search of the matching points must be carried out between the left and right image. If the bidirectional match is not a success, the match itself is incorrect and this pair of corresponding points should be excluded.

After the above steps have been completed, the best matching points of each marker can be accurately found.

5.3. The Gesture Calculation of the Operating Handle

The gesture calculation is divided into three steps: 1) 3D coordinates of each marker are found in any two of the trinocular system cameras. The experiment is shown that if the arm conceals a few markers, the marker coordinate can be calculated by one pair of cameras; if the marker isn't concealed, the marker coordinate can be calculated using three groups of the binocular system. 2) The optimal coordinate values of the markers need to be determined, that is the unique gestures of the operating handle. 3) According to the structural parameters of the operating handle and the 3D coordinates of its markers, the gesture of the operating handle can be solved. The algorithm is as follows:

First, suppose the world coordinate system and the CCD1 camera coordinate system coincide with each other in the trinocular system. The internal and external parameters for the three groups of the trinocular system are separately obtained through calibration, including the rotation matrix $\boldsymbol{R_{rli}}$, translation vector $\boldsymbol{T_{rli}}$ and effective focal length $\boldsymbol{f_{li}}$, $\boldsymbol{f_{ri}}$ (i = 1,2,3), when $\boldsymbol{i}$ represents the group number of the binocular system.

Copyright ⓒ2016 SERSC

$$R_{rli} = \begin{pmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{pmatrix}, T_{rli} = \begin{pmatrix} t_x & t_y & t_z \end{pmatrix}^T \tag{19}$$

$\boldsymbol{p_l(X_l,Y_l)}$ and $\boldsymbol{p_r(X_r,Y_r)}$ are the image coordinates of the marker **P** in the left and right image plane. According to the relation of perspective projection for the marker and its corresponding points in the left and right image plane, the 3D coordinates **P**(x,y,z) in the CCD1 camera coordinate system are:

$$x = zX_l / f_{li}, y = zY_l / f_{li}, z = \frac{f_{li}(f_{ri}t_x - X_r t_z)}{X_r(r_7 X_l + r_8 Y_l + r_9 f_{li}) - f_{ri}(r_1 X_l + r_2 Y_l + r_3 f_{li})} \tag{20}$$

Second, as shown in the trinocular system, in Figure 5, the two groups of the binocular system, CCD1-CCD2 and CCD1-CCD3, are used to determine the 3D coordinates of the marker. When using the third group of the binocular system, CCD2-CCD3, the 3D coordinate value $\mathbf{P^3}(x_3,y_3,z_3)$ of the marker needs to coordinate a conversion to the CCD1 camera coordinate system. The rotation matrix $\boldsymbol{R_{rl1}}$, translation vector $\boldsymbol{T_{rl1}}$ can be obtained through calibration; the equation of the coordinate conversion is:

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = R_{rl1} \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix} + T_{rl1} \tag{21}$$

where (x_1,y_1,z_1) is the coordinate value of the point $P^3(x_3,y_3,z_3)$ in the CCD1camera coordinate system.

Third, the marker coordinates are converted through a coordinate conversion, but only after they have been determined by each binocular system. Then, several coordinate values of the same marker are averaged out and the optimal coordinate value is determined. The marker **P** coordinates in the CCD1camera coordinate system are:

$$P^W = P^{CCD1} = (\frac{1}{n}\sum_{i=1}^{n} x_i, \frac{1}{n}\sum_{i=1}^{n} y_i, \frac{1}{n}\sum_{i=1}^{n} z_i), \quad \boldsymbol{n}=1,2,3. \tag{22}$$

Fourth, the coordinate matrix **M** = (P_1^W, P_2^W, ..., P_i^W, ..., P_n^W) is composed of the 3D coordinates of all the markers. Where $P_i^W = (x_i,y_i,z_i)^T$, $\boldsymbol{i}$ = 1 ... $\boldsymbol{n}$, $\boldsymbol{n}$ is the number of markers.

Finally, a coordinate system O-xyz is built for the operating handle, in which the position and orientation of the handle is indicated by $\boldsymbol{^wF}$,

$${}^wF = \begin{pmatrix} n_x & o_x & a_x & x_O \\ n_y & o_y & a_y & y_O \\ n_z & o_z & a_z & z_O \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{23}$$

Where $(n_x, n_y, n_z)^T$, $(o_x, o_y, o_z)^T$ and $(a_x, a_y, a_z)^T$ is the unit direction vector, which separately represents the orientation of three axes for the coordinate system O-xyz. $(x_o, y_o, z_o)^T$ is the coordinate value of the origin O in the world coordinate system.

According to the structure parameters of the handle, the coordinate value of each marker in the coordinate system O-xyz can be calculated, which is show as $P_i' = (x_i', y_i', z_i')^T$, $\boldsymbol{i}$ = 1 ... $\boldsymbol{n}$, $\boldsymbol{n}$ is the number of markers. Then, the coordinate matrix M'= (P_1', P_2', ..., P_i', ..., P_n') is composed of 3D coordinates of all the markers in the coordinate system O-xyz, therefore,

 Copyright⓪ 2016 SERSC

$$\mathrm{M} = {}^{w}F \cdot M' \tag{24}$$

Substituting **M** and ***M'***, the gesture of the handle can be expressed by ${}^{w}\boldsymbol{F}$.
The attitude angle of the handle can be expressed by the Euler angles Z-Y-X. Therefore, the gesture of the handle ${}^{w}F$ is the vector $((x, y, z), \alpha, \beta, \gamma)$, where (x, y, z) represents the position of the handle and α, β, γ is the angle which separately represents the rotations for the handles relative to the world coordinate system's Z-axis, Y-axis, X-axis.

6. Experimental Analysis

The hardware of the experimental system consists of the following devices. Three industrial cameras with a maximum frame rate of 30 frames/s. The camera lens is 12.5 mm focal length; the computer configuration is CPU frequency of 2.4GHz and 2GB RAM; a precise 3D motion platform whose positioning accuracy is 0.02mm with a workspace of 200mm×200mm×150mm. Also, a binocular system was used to carry out the experimental verification of the vision measurement for the handle gesture.

6.1. The Camera Calibration of the Binocular Vision System

A familiar calibration method[12] was adopted to calibrate two of the cameras of the binocular system. These calibrations yielded the camera's Internal and external parameter.

Internal parameter matrix of the left, right camera CCD1 ,CCD2:

$$A_l = \begin{pmatrix} 2956.2067 & 0 & 627.9930 \\ 0 & 2940.0975 & 515.2024 \\ 0 & 0 & 1.0000 \end{pmatrix} \quad A_r = \begin{pmatrix} 2939.5353 & 0 & 667.5260 \\ 0 & 2922.2544 & 502.6779 \\ 0 & 0 & 1.0000 \end{pmatrix};$$

Suppose the external parameter $\boldsymbol{R_l, T_l}$ or $\boldsymbol{R_r, T_r}$ separately represent the corresponding position relation between the left or right camera coordinate system and the world coordinate system. Then, the position relation between the left, right camera coordinate system is:

$$R_{rl} = R_l R_r^{-1}, T_{rl} = T_l - R_l R_r^{-1} T_r \tag{25}$$

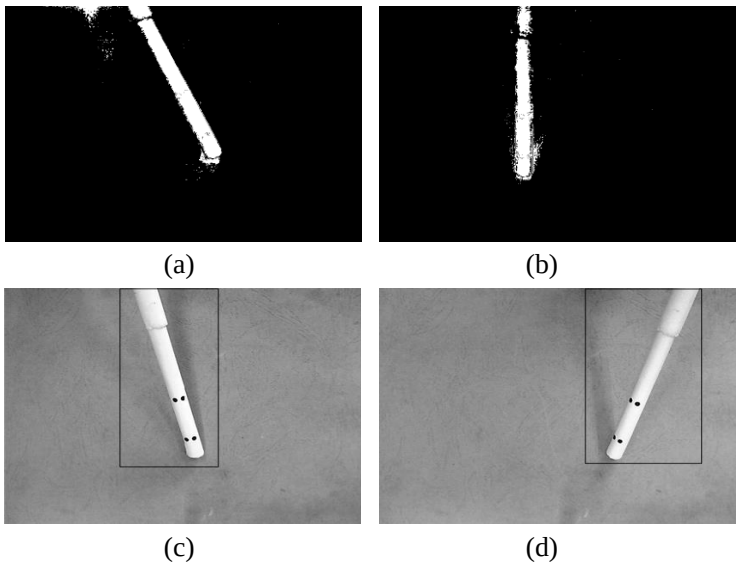
The rotation matrix $\boldsymbol{R_{rl}}$ and translation vector $\boldsymbol{T_{rl}}$ can be obtained by the iterative optimization of 14 positions parameters for calibration target, namely,

$$R_{rl} = \begin{bmatrix} 0.9838 & -0.0045 & -0.1799 \\ 0.0092 & 1.0000 & 0.0224 \\ 0.1801 & -0.0241 & 0.9837 \end{bmatrix}, \quad T_{rl} = \begin{bmatrix} 98.7723 & 0.8410 & 9.1032 \end{bmatrix}$$

6.2. Recognition and Tracking of the Handle

The length of the cylindrical handle used in the experiment is 200mm; the diameter is 10mm. Six columns markers are distributed evenly along the circle of the cylindrical surface atop the handle, and the angle of the two adjoining markers along the circle is 60°. Every column has two markers. Figure 6 shows part of the processing result of the recognition and tracking for the handle. Figure 6(a) and (b) are the identification results of the handle after using the background difference method, while 6(c) and (d) are the tracking process of the handle after using the Kalman filter. According to the figures, the algorithm can track the motion of the handle precisely. The result shows that the algorithm introduced in this paper can adapt well to changing ambient light.

Copyright ⓒ2016 SERSC



(a) (b)

(c) (d)

Figure 6. Results for the Recognition and Tracking of the Handle

6.3. Gesture Measurement of the Handle

The precise 3D motion platform is used for driving the handle move within the workspace of the handle in the experiment. The handle is fixed on the motion platform according to the setting attitude angle, and then the parameters of reality gesture for the handle can be obtained, and the reality parameters of handle gesture are used as theoretical value. As shown in Table 1. Meanwhile, the handle gesture is measured by three group of binocular system. Compare the parameters of visual measurement with the theoretical parameters, visual measure precision can be analyzed. Thus, Table 1 provide part of the experiment result, from the data in the table, the algorithm can measure the handle gesture precisely.

6.4. Error Analysis

The mean absolute error is introduced to represent the measure precision of algorithm, and carry out quantitative analysis of the error. The mean absolute error of position measurement can be defined as:

$$E_t = \sum_{i=1}^{N} \sqrt{(x_i' - x_i)^2 + (y_i' - y_i)^2 + (z_i' - z_i)^2} \Big/ N \tag{26}$$

The mean absolute error of attitude angle can be defined as:

$$E_\theta = \sum_{i=1}^{N} |\theta_i' - \theta_i| \Big/ N \tag{27}$$

where (x, y, z) is the reality 3D coordinate, while (x', y', z') is the measure 3D coordinate. And $\theta' = (\alpha', \beta', \gamma')$ is the measure attitude angle, while $\theta = (\alpha, \beta, \gamma)$ is the reality attitude angle. N is the number of sample, N=20. Through counting the data in the Table1, the absolute error of the position measurement can be obtained: E_t =0.31mm; and the absolute error of the angle measurement: E_α =0.67°, E_β =0.56°, E_γ =0.65°. So the measurement error can be influenced by the factors as lens marginal distortion, the measuring distance, the size of markers, the motion speed of handle, and these problems need to further resolve in the next research.

 Copyrightⓒ 2016 SERSC

Table 1. Result of Handle Gesture Measurement

	Measured position coordinate(x', y', z')/mm	Theoretic position coordinate(x, y, z)/mm	Measured attitude angle(α', β', γ')/°	Theoretic attitude angle(α, β, γ)/°
1	(-114.59,36.91,598.83)	(-114.8,36.81,598.66)	(27.1, 77.3,-88.2)	(26.4,76.7,-89.1)
2	(-90.78,41.93,586.25)	(-90.57,41.90, 586.25)	(-123.0,77.1,-100.1)	(-123.6,76.8,-101.)
3	(-77.54,45.73,587.22)	(-77.61,45.67,587.04)	(-102.9,67.0,-111.9)	(-103.6,66.1,-112.)
....	.....	.....	.....	.....
7	(13.65,47.37,565.65)	(13.68,47.29,565.45)	(-85.4,68.8,-155.3)	(-86.3,68.4,-156.1)
8	(37.06,48.25,561.22)	(37.27,48.4289,561.5)	(-97.7,39.2,-178.5)	(-98.3,38.5,-179.0)
9	(45.70,42.58,559.16)	(46.09,42.41,559.35)	(-103.2,55.3,-167.6)	(-103.7,55.4,-168.)
....	.....	.....	.....	.....
20	(-84.07,-21.32,549.06)	(-84.06,-21.40, 549.0)	(-56.8,72.0,-172.4)	(-57.2,71.4,-172.6)

7. Conclusions

In order to solve the problem of the handle concealed, the moving handle can be recognized and tracked by three groups of binocular system in this paper. By the background difference method, the moving handle can be found out in image. For introducing the renew strategy, the influence of the change of the light to the detecting result is decreased. Then, through Kalman filter, the position, speed and acceleration of the handle can be forecasted, which narrow the hunting zone of target effectively and improve the detecting speed. For the stereo matching of markers, the methods of feature matching and regional gray matching are combined. By introducing the uniqueness and sequential consistency restriction, the pseudo matching can be rejected. At last, the spatial gesture of handle can be obtained by 3D coordinate analyser of marker points.

The experiment prove that the method presented in the paper can identify in real time and track the handle, measure its spatial gesture, and the measurement precision is 0.31mm, while the measurement precision of attitude angle is 0.67°. However, the measurement precision of the system and the robustness of the algorithm require to be improved, besides, how to introduce force feedback into the main operating handle need to do further study.

Acknowledgments

This research is supported by Fundamental Research Funds for the Central Universities (2014G1251029, 2014G1251030, 2014G1251032, 310825151039), Natural Science Foundation of Shaanxi Province (2015JM5216). The authors gratefully acknowledge the partial support from Key Laboratory of the Ministry of Education On Road Works Technology & Equipment, and ShaanXi Engineering & Technology Research Center of Special Digital Manufacturing Equipment.

References

[1] M. W. Farrington, G. J. Gerling, L. R. Kohan and S. L. Johnson, "A multisensory illusion with haptic interaction to treat phantom limb pain", Haptics Symposium, Vancouver, BC, Canada, **(2014)** March 4-7.
[2] V. M. Hung and U. J. Na, "Tele-operation of a 6-DOF serial robot using a new 6-DOF haptic interface", 9th IEEE International Symposium on Haptic Audio-Visual Environments and Games, Phoenix, AZ, United States, **(2013)** October 16 -17.
[3] V. M. Hung and U. J. Na, "A New 6-DOF Haptic Device for Teleoperation of 6-DOF Serial Robots", IEEE Transactions On Instrumentation And Measurement, vol.11, **(2012)**, pp. 60.
[4] O.Baser and E. I. Konukseven, "Kinematic calibration of PHANTOM Premium 1.5/6DOF haptic device", Key Engineering Materials, Advanced Design and Manufacture, vol. 4, **(2013)**, pp. 486.
[5] P. Berkelman and M. Dzadovsky, "Actuation model for control of a long range Lorentz force magnetic levitation device", 23rd IEEE/RSJ International Conference on Intelligent Robots and Systems. Taipei, Taiwan, **(2012)** October 18-22.

Copyright ⓒ2016 SERSC

[6] Y.Q. Gao, J. H. Li, H.Su and J.M. Li, "Development of a teleoperation system based on virtual environment", IEEE International Conference on Robotics and Biomimetics, Phuket, Thailand, **(2015)** December 7-11.
[7] J.Nash, K.Atanassov, S.Goma, V.Ramachandra and H.Siddiqui, "Real-time object tracking and dimension measurement using stereo cameras", Real-Time Image and Video Processing. Burlingame, United States, **(2013)** February 6-7.
[8] Q.Liu, X.S.Qin, S.Yin and F. HE, "Structural Parameters Design and Accuracy Analysis of Binocular Vision Measure System", China Mechanical Engineering, vol. 19, **(2008)**, pp. 22.
[9] L. S. Garcia, D. C. Albarran and B. T. Ferrer, "Visual performance comparison between contact lens-based pinhole and simultaneous vision contact lenses", Clinical And Experimental Optometry. vol. 1, **(2013)**, pp. 96.
[10] L.Lamard, R.Chapuis and J.Boyer, "Dealing with occlusions with multi targets tracking algorithms for the real road context", IEEE Intelligent Vehicles Symposium, Alcal de Henares, Madrid, Spain, **(2012)** June 3-7.
[11] P.Sandoz, I.Elhechmi and T.Gharbi, "Toward stereo visual monitoring of three-dimensional translations with submicrometer resolution", Journal of the Optical Society of America A: Optics and Image Science, and Vision, vol. 29, **(2012)**, pp.11.
[12] M. Xie, Z.Wei, G. Zhang and X. Wei, "A flexible technique for calibrating relative position and orientation of two cameras with no-overlapping Fov", Journal of the International Measurement Confederation, vol. 1, **(2013)**, pp. 46.

Authors

Qiong Liu, (1976-), PhD, she graduated from Northwestern Polytechnical University and major in machine vision and advanced manufacturing technology.

YaGuang Zhu, (1986-), PhD, she graduated from Zhe Jiang University and major in Mechanical and Electronic Engineering, advanced manufacturing technology.

Xu Zhang (1983-), PhD, he graduated from Northwestern Polytechnical University and major in Mechanical automation and Radio Communications.

YanPu Yang (1986-), PhD, he graduated from Northwestern Polytechnical University and major in the finite element analysis, Industrial Designs.

 Copyrightⓒ 2016 SERSC

Ke Li (1983-), PhD, she graduated from Chang'An University and major in Mechanical manufacturing and automation.

Copyright ⓒ2016 SERSC

Copyrightⓒ 2016 SERSC