Research on Static Vision-Based Target Localization for Astronaut Assistant Robots

FENGWEI GU1,2,3, HONGWEI GAO1,2, JINGUO LIU2,3 (Senior Member, IEEE), QING GAO2,3,4, AND ZHAOJIE JU2,3,5 (Senior Member, IEEE)

1School of Automation and Electrical Engineering, Shenyang Ligong University, Shenyang 110159, China
2State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China
3Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110016, China
4Shenyang Institute of Automation, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
5School of Computing, University of Portsmouth, Portsmouth PO1 3HE, U.K.

Corresponding authors: Hongwei Gao (ghw1978@sohu.com) and Jinguo Liu (liujuinguo@sia.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1304600, in part by the Natural Science Foundation of China under Grant 51775541, in part by the Chinese Academy of Sciences (CAS) Interdisciplinary Innovation Team under Grant JCTD-2018-11, and in part by the College Innovative Talents Program Support Project in Liaoning Province.

ABSTRACT Aiming at the problem of localizing and grasping targets by an astronaut assistant robot, developing assisting staff to complete specific tasks, and using a self-built visual system as an experimental platform, this paper proposes two improved static target localization methods based on binocular vision and the traditional localization method: (a) a localization method based on region of interest matching and (b) a localization method based on a pixel coordinate offset. Under the same conditions, the traditional binocular vision localization method and the two improved localization methods are used to locate a spherical target. The experimental results are compared and analyzed, and the effectiveness of the improved methods is verified by two performance indicators: the localization accuracy and the localization time efficiency. The experimental results show that to a certain extent, the two improved localization methods can improve the localization efficiency of the vision system and positively affect the performance and flexibility of the astronaut assistant robots.

INDEX TERMS Astronaut assistant robot, binocular vision, region of interest, localization method.

I. INTRODUCTION

A. BACKGROUND AND MEANING

In space exploration, the astronaut assistant robot are playing an increasingly important role in assisting astronauts to complete the designated work and can also replace astronauts to complete some dangerous tasks [1]. To complete these tasks, the robot should be able to effectively perceive the environment [2] and determine the location of a target. At this time, visual technology has a significant advantage in that it can be used as the “eye” of the astronaut assistant robot to realize the localization function and provide location information [3].

In recent years, with the gradual expansion of the intelligence requirements of all walks of life, machine vision has also been rapidly developed and gradually applied to various fields [4], [5]. For example, the application of machine vision to vehicle detection in the transportation industry [6] can improve the reliability of advanced driver assistance systems (ADASs). In the aerospace field, machine vision and neural network are combined for apply robot gesture recognition and body pose estimation [7] so that the robot can perform corresponding actions and realize human-computer interaction. In industry, machine vision is applied to pointer-type instrument detection [8] to help workers perform calibration tasks for certain instruments, which can solve the problem of the low efficiency and subjective error caused by manual calibration and can improve the measurement accuracy. Machine vision plays a particularly important role in many of the above applications.

B. RELATED WORK

At present, to solve the defects of traditional monocular vision [9], researchers are studying the multi-eye vision system, carrying out many studies of the corresponding measurement methods, and are obtaining many meaningful research
results. In the literature [10], a measurement method for non-cooperative spacecraft based on binocular vision has been proposed. Although this method has high accuracy, the time consumption of the measurement process is high. The reason is that the stereo matching in this method is very time-consuming. In another study [11], a binocular vision target localization method based on a point cloud was presented. This method generates a point cloud by combining the calibration result and the disparity map obtained by stereo matching, and performs noise reduction processing on the point cloud. A three-dimensional reconstruction of the target is achieved within an appropriate distance [12], but the reconstruction takes a long time and generates a large amount of randomly distributed noise during the generation of the point cloud, even if it is processed by de-drying, which is likely to produce a larger impact on the localization accuracy [13]. Another work [14] introduced a method for extracting three-dimensional navigation lines of field roads by binocular vision. In this method, first, the roads are distinguished from the surrounding environment by enhancing the recognition of image shadows [15] and the method of information fusion [16]. Second, the stereo matching image primitives are extracted [17], then a three-dimensional measurement is realized after stereo matching, and finally the autonomous navigation line of the road is obtained [18]. The matching primitive extraction part in this method is very complicated, and the stereo matching requires a long time. The experimental results of the method satisfy the accuracy requirements, but the error is still relatively larger. In [19] and [20], three-eye and four-eye visual techniques were employed to achieve the target measurement. Among them, [19] designed a three-mesh camera model with a variable baseline and three binocular subsystems composed of three-eye camera. The depth information could be obtained from different angles, and the appropriate baseline distance could be adjusted according to different measurement targets [21]. Finally, the size of the measurement target was obtained by comprehensively processing the three binocular subsystems, and the method had good precision and stability. In [20], a four-dimensional reconstruction of a target was realized by the four-eye visual measurement system (FMS). Compared with the traditional binocular vision system, the FMS can form six binocular vision systems to increase the matching constraint and reduce the mismatch probability, which improves the accuracy of the stereo matching while also improving the measurement accuracy. However, the hardware and software of the three-eye and four-eye vision systems are complex, time-consuming, costly, and difficult to implement.

To solve the above problems and meet the localization requirements of astronaut assistant robots, two improved binocular vision localization methods are proposed in this paper: (a) a localization method based on region of interest matching, which uses the ROI to remove the image background and determine the image analysis range and (b) a localization method based on a pixel coordinate offset, where the pixel coordinates of the target are obtained through the ROI region, and a coordinate transformation is realized by the pixel coordinate offset. Both of these methods utilize the ROI region, which solves the problem of the traditional monocular vision method not effectively obtaining the depth information, can effectively obtain three-dimensional information of the measurement target, and can improve the measurement accuracy while avoiding the high complexity and high time consumption of the multi-vision system in the hardware and software. This method improves the localization efficiency of the astronaut assistant robot vision system.

The structure of the rest of the paper is as follows: Chapter 2 introduces the structure of the visual localization system. Chapter 3 introduces two improved localization methods based on the traditional localization method. Chapter 4 uses the above methods for localization experiments, and the two results of the localization accuracy and time efficiency are compared and analysed. Finally, Chapter 5 presents the conclusions and future work directions.

II. THE COMPOSITION OF THE VISUAL LOCALIZATION SYSTEM

The astronaut assistant robot adopts AAR-2 [22] developed by the Shenyang Institute of Automation, Chinese Academy of Sciences. The structure of the system is shown in Figure 1. Part a is the visual system experimental platform, and part b is the astronaut assistant robot system. The experimental part of this paper is carried out by means of platform a, and the visual platform is transplanted to system b to provide a visual localization function for the robot. The designed vision system provides location information for AAR-2, enabling the crawling system to complete the crawling task. The method uses two industrial cameras to form a binocular vision system for the localization measurement. Compared with a monocular vision system, a binocular vision system can better obtain the depth information of a measurement target in space [23] and obtain the three-dimensional coordinate value of the target in the world coordinate system. The main components of the vision system are as follows: two Basler gigabit network industrial cameras, two kowa 12mm fixed-focus lens, two C-mount interfaces between cameras and lens, and a computer with an Intel Core i5-4460 CPU, 4GB of RAM, and a 64-bit Win7 operating system.
The operation flow of the binocular vision system is shown in Figure 2. First, the camera driver and software development kit (SDK) are installed in the computer to complete the configuration of the camera IP address, control the binocular camera for image acquisition, and then perform image processing, such as filter denoising, image enhancement and feature extraction. Finally, the image processing analysis result is input into the binocular localization scheme, and the localization result of the target in the world coordinate system can be obtained.

III. INTRODUCTION TO THE PRICIPLE OF TWO IMPROVED BINOCULAR LOCALIZATION METHODS

In this section, we will introduce the relevant content of the binocular localization methods in the visual system, i.e., the traditional localization method and the two improved localization methods, including the principle and related derivation process, which lay the foundation for subsequent experimental verification and result analysis; for the purpose of this paper, the binocular localization method mainly has the following three requirements:

1) Ensure the real-time requirements of the binocular localization method;
2) Satisfy the accuracy of the binocular localization method and ensure the measurement accuracy;
3) Ensure the ability to locate measurement targets in a more complex context.

A. IMPLEMENTATION OF THE TRADITIONAL LOCALIZATION METHOD

To localize a measurement target in space and obtain its three-dimensional coordinates, the general flow of the binocular localization method [24] is shown in Figure 3. The methods used to implement the various parts of this paper are described in 1) to 4).

1) CAMERA OFFLINE CALIBRATION

First, this paper performs a camera calibration of the binocular vision system and establishes a correspondence between the image coordinate system and the world coordinate system, as shown in equation (1). The calibration of the binocular cameras is based on a single camera calibration of the left and right cameras after the calibration images are acquired, and the internal parameter matrix $M_1$ and $M_2$ and the outer parameter matrix $R_1$, $T_1$, $R_2$, and $T_2$ of each camera are obtained; then, the calibration result is used. The binocular camera is calibrated to obtain the structural parameters of the binocular cameras, namely, the rotation matrix $R$ and the translation matrix $T$. The process is an offline calibration and is completed by the Zhang Zhengyou calibration method [25].

$$sp = M[R_R; T_T]P$$

where $s$ is a scale factor, $p$ is the image coordinate vector, $P$ is the world coordinate vector, $R_R$ and $T_T$ are the rotation matrix and translation matrix of the world coordinate system to the image coordinate system, respectively, and $M$ is the inner parameter matrix. The inner parameter matrix $M$ can be represented by equation (2):

$$M = \begin{bmatrix} f_x & \mu & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

$f_x$ and $f_y$ represents the focal length normalized by the x-axis and the y-axis, respectively, $\mu$ is the tilt factor, and $u_0$ and $v_0$ are the coordinates of the image principal point.

The structural parameters (the rotation matrix $R$ and translation matrix $T$) of the binocular cameras can be obtained by equation (3):

$$\begin{cases} R = R_1R_2^{-1} \\ T = T_1 - RT_2 \end{cases}$$

$R_1$ and $T_1$ are the rotation and translation matrix of the left camera, respectively, and $R_2$ and $T_2$ are the rotation and translation matrix of the right camera, respectively.

Stereo calibration images of the binocular vision system are shown in Figure 4. A 7 * 10 checkerboard calibration plate is used with a side length of 22.5mm. Twenty sets of calibration images of different poses are collected. The stereo calibration results are shown in Table 1 and Table 2, wherein Table 1 shows the internal parameters of the left and right cameras in the binocular vision system. The main parameters involved include the focal lengths $f_x$ and $f_y$. 

![FIGURE 2. The configuration and processing flow of the binocular vision system.](image-url)

![FIGURE 3. The general flow of the binocular localization method.](image-url)
TABLE 1. Parameters of the binocular vision system.

<table>
<thead>
<tr>
<th>Internal parameters</th>
<th>$f_x$</th>
<th>$f_y$</th>
<th>$u_0$</th>
<th>$v_0$</th>
<th>$k_1$</th>
<th>$k_2$</th>
<th>$k_3$</th>
<th>$p_1$</th>
<th>$p_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left camera</td>
<td>3743.6</td>
<td>3778.1</td>
<td>1315.3</td>
<td>1026.9</td>
<td>-0.03</td>
<td>-0.37</td>
<td>8.81</td>
<td>-0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Right camera</td>
<td>1532</td>
<td>1805</td>
<td>9716</td>
<td>0406</td>
<td>665</td>
<td>991</td>
<td>158</td>
<td>109</td>
<td>561</td>
</tr>
<tr>
<td></td>
<td>3755.3</td>
<td>3783.9</td>
<td>1279.5</td>
<td>967.8</td>
<td>-0.00</td>
<td>-2.17</td>
<td>3.37</td>
<td>-0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>2497</td>
<td>5799</td>
<td>9656</td>
<td>2109</td>
<td>364</td>
<td>539</td>
<td>889</td>
<td>372</td>
<td>542</td>
</tr>
</tbody>
</table>

FIGURE 4. Calibration images of the binocular vision system.

after the x-axis and y-axis normalization, the image main point coordinates $u_0$ and $v_0$, the radial distortion parameters $k_1$, $k_2$ and $k_3$, and the tangential distortion parameters $p_1$ and $p_2$. Table 2 shows the structural parameters of the binocular vision system and the relative positional relationship between the binocular camera coordinate systems, namely, the rotation matrix R and the translation matrix T. After the stereo calibration is completed, the positional relationship diagram between the binocular camera and the calibration plate as shown in Figure 5 can be obtained, and the calibration error map of the binocular vision system shown in Figure 6 can also be obtained. The change in the position of the calibration plate and the relative position of the binocular cameras can be seen within a certain range of the cameras. Figure 5 shows that the relative distance between the two cameras is approximately 100 mm, and the absolute value of the first element of the translation matrix T in the calibration result is approximately 97.89 mm, which are only 2.11 mm apart. From Figure 6, it can be seen that the minimum and maximum values of the calibration error are 0.35 pixels and 0.73 pixels, respectively, and the average error is 0.50 pixels, which satisfies the calibration requirements.

<table>
<thead>
<tr>
<th>Structural parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R$</td>
</tr>
<tr>
<td>0.99931</td>
</tr>
<tr>
<td>-0.00114</td>
</tr>
<tr>
<td>0.03715</td>
</tr>
<tr>
<td>$T$</td>
</tr>
<tr>
<td>-0.00090</td>
</tr>
<tr>
<td>1.00000</td>
</tr>
<tr>
<td>0.00637</td>
</tr>
<tr>
<td>-0.03716</td>
</tr>
<tr>
<td>-0.00633</td>
</tr>
<tr>
<td>0.99929</td>
</tr>
<tr>
<td>[-97.88860, -0.12019, 1.11478]</td>
</tr>
</tbody>
</table>

FIGURE 5. Positional relationship between the binocular vision system and calibration plate.

2) ORIGINAL IMAGE ACQUISITION AND STEREO CORRECTION

In general, circular targets are useful for detection and identification. Therefore, it is usually necessary to design a suitable circular marker point according to the need to attach a marker to the appropriate position at the measurement target to facilitate the target localization. At this point, the measurement target can be any shape, which is more universal. Therefore, it is of great significance to use a spherical target as the measurement object to measure and locate the target.

After the camera calibration is completed, spherical target images are acquired by the binocular vision system (a ping-pong ball is used as the spherical target, wherein the function of the water cup is to increase the image background complexity), and one set of images is shown in Figure 7. Image processing is performed, which involves denoising and an image enhancement. Denoising can remove various forms of noise contained in the image acquisition process. Image enhancement can increase the contrast of the image to make the measurement target clearer. After the image
In practical applications, for the stereo correction problem [26], the optical axes of the binocular cameras are required to be strictly parallel, the acquired images are completely co-planar, and the pixels are line aligned [27]. Therefore distortion correction is also performed in the image stereo correction process [28], which reduces the distortion of the image and makes the image as close as possible to the original undistorted image; this is beneficial for the line alignment of the pixels, but it is almost impossible to achieve a complete coplanar and precise alignment of the imaging plane due to the placement of the binocular camera. To achieve this goal, in this paper, the Bouget polar line correction algorithm is used to correct the binocular images while correcting the distortion so that the binocular images are mathematically aligned to the same observation plane, instead of using a physical alignment. The mathematical model of the Bouget polar correction algorithm [29] is as follows:

(a) Decompose the rotation matrix R of the right camera relative to the left camera into two composite matrices $r_l$ and $r_r$, so that each camera is rotated in part, the two cameras are imaged in parallel, but the pixel rows corresponding to each image are not aligned;

(b) Construct the transformation matrix $R_{rect}$ by the translation matrix $T$ of the right camera relative to the left camera, so that the baseline and the imaging plane are parallel, and the transformation matrix $R_{rect}$ is constructed as follows:

First, construct $e_1$. The transformation matrix transforms the pole of the left image to infinity, causing the polar line to reach a level. The final image plane should be parallel to the line connecting the origins of the left and right camera coordinate systems. In this case, the translation vector $T$ between the cameras can be taken as the direction of the left pole. The vector $T$ is normalized to the unit vector $e_1$:

$$e_1 = \frac{T}{\|T\|} = \left[ \frac{T_x}{\sqrt{T_x^2 + T_y^2 + T_z^2}}, \frac{T_y}{\sqrt{T_x^2 + T_y^2 + T_z^2}}, \frac{T_z}{\sqrt{T_x^2 + T_y^2 + T_z^2}} \right]$$

(4)

Second, construct $e_2$. $e_2$ is required to be orthogonal to $e_1$, the direction vector $e_0 = (0, 0, 1)$ of the optical axis is selected as the medium, and the $e_1$ outer product is normalized to obtain the unit vector $e_2$:

$$e_2 = \frac{e_1 \times e_0}{\|e_1 \times e_0\|} = \left[ -T_y, T_x, 0 \right]^T$$

(5)

Then construct $e_3$. $e_3$ is required to be orthogonal to $e_1$ and $e_2$. After constructing $e_1$ and $e_2$, $e_3$ can be obtained by the
outer product of $e_1$ and $e_2$:

$$e_3 = e_1 \times e_2$$  (6)

Finally, a transformation matrix $R_{\text{rect}}$ is obtained, which converts the left camera image pole to infinity, causing the image to rotate and the polar line to level. $R_{\text{rect}}$ consists of $e_1$, $e_2$ and $e_3$:

$$R_{\text{rect}} = \begin{bmatrix} e_1^T & e_2^T & e_3^T \end{bmatrix}^T$$  (7)

(c) The final correction matrix is obtained by transforming the matrix $R_{\text{rect}}$ by multiplying by $R$’s composite matrices $r_l$ and $r_r$ to achieve an image pixel row alignment. The correction matrices are as follows:

$$\begin{cases} R_l = R_{\text{rect}}r_l \\ R_r = R_{\text{rect}}r_r \end{cases}$$  (8)

A schematic diagram of the stereo correction process is shown in Figure 8. After the stereo correction is completed, the final result is obtained by image cropping, as shown in Figure 9. At this time, the left and right images of the binocular vision system are completely coplanar, and the rows of the pixels are precisely aligned, thus reducing the search range of subsequent stereo matching from a two-dimensional search to a one-dimensional search [27] and filtering out the no-matching points. In this way, a certain line of pixels in one image can be directly searched for a matching point corresponding to another image, which improves the stereo matching efficiency.

3) STEREO MATCHING

In binocular vision measurement, stereo matching is a difficult and key problem. Stereo matching technologies can be divided into three categories [30]: feature-based stereo matching technology, region-based stereo matching technology and phase-based stereo matching technology. The stereo matching technology of binocular vision acquires two different two-dimensional images through the two cameras and then obtains the parallax information by finding the correspondence between the pixels of the two images. Furthermore, a disparity map of the two images can be obtained. Finally, the coordinate values in the world coordinate system in the three-dimensional space corresponding to the pixel points in the two-dimensional image can be obtained by using the triangulation principle and the disparity information [31].

On the basis of the stereo correction, the two images have already achieved a line alignment of the pixels; then, when stereo matching is performed, the corresponding two pixel points can be searched in the corresponding pixel row for stereo matching, and the corresponding disparity value is obtained. A schematic diagram of the matching process is shown in Figure 10, wherein 1 to 5 of the left image correspond to 1′ to 5′ of the right image, respectively. The box position of 1′ to 5′ of the same pixel row in the right image is searched based on frame position 5 of the left image until the corresponding box position is found, so it is not necessary to find matching points for the entire image, and it is only necessary to search one by one in the corresponding pixel row. To a certain extent, this approach not only reduces the possibility of mismatching in the stereo matching process but also improves the accuracy and speed of stereo matching at a certain level.

It can be assumed that the pixel coordinates of the left and right images in Figure 10 are $(u_1, v_1)$ and $(u_2, v_2)$, respectively; after the stereo correction is completed, the imaging point of the same target point in the world coordinate system in the left camera is at the right position of the image, whereas the imaging point in the right camera is at the left position of the image; the pixel row numbers of $v_1$ and $v_2$ are the same.
The disparity map of the binocular images. and the following relationship is obtained:

\[
\begin{align*}
&u_1 > u_2 \\
v_1 = v_2
\end{align*}
\]  

At this time, stereo matching is performed, \((u_1, v_1)\) and \((u_2, v_2)\) are obtained as a pair of matching points, and the disparity value can be expressed by the following formula:

\[
d = (u_1 - u_2)\Delta u
\]

where \(d\) is the disparity value between the matching point pairs, and \(\Delta u\) is the physical size of the pixel.

At this point, the disparity values of a pair of matching points are obtained, the disparity values of other matching point pairs can be obtained by the same approach, and the disparity map can be obtained from this disparity information, thereby using the disparity map as a basis for the measurement.

For the stereo matching problem, this paper uses a semi-global block matching (SGBM) algorithm to achieve the stereo matching of binocular images [32], thus obtaining the disparity map of the two images, as shown in Figure 11. As a result, we can obtain the depth information of the target.

4) BINOCULAR LOCALIZATION METHOD

In this paper, three-dimensional reconstruction of the target point is first carried out by using the principle of triangulation. The positional coordinates in the world coordinate system are obtained by combining the disparity, and binocular localization is achieved [33]. The imaging principle of the binocular vision system is shown in Figure 12. The imaging planes of the left and right cameras are stereoscopically corrected and are already in the same plane. One point \(P(X_w, Y_w, Z_w)\) in the world coordinate system obtains a pair of corresponding image points in the left and right images. The binocular cameras by a projection transformation, and the coordinates of the left and right image points in the imaging plane are \(P_1(x'_1, y'_1)\) and \(P_2(x'_2, y'_2)\), respectively; \(O_wX_wY_wZ_w\) is the world coordinate system, \(O_c1X_c1Y_c1Z_c1\) and \(O_c2X_c2Y_c2Z_c2\) are the left and right camera coordinate systems, respectively, and \(O_1X_1Y_1\) and \(O_2X_2Y_2\) are the image physical coordinate systems of the left and right cameras, respectively. \(f_1\) and \(f_2\) are the focal lengths of the left and right cameras, respectively. \(f_1 = f_2 = f\), the baseline length is \(T_s\), the segments \(O_1P\) and \(O_2P\) are projected onto the horizontal plane of the optical axis, and the corresponding points for points \(P, P_1, \text{ and } P_2\) are \(P', P'_1, \text{ and } P'_2\), respectively. After the stereo correction, the \(x\)-axis of the coordinate system \(O_c1X_c1Y_c1Z_c1\) and the \(x\)-axis of the coordinate system \(O_c2X_c2Y_c2Z_c2\) are on the same straight line \(O_1O_2\). The crossing points \(P'_1, P' \text{ and } P'_2\) are perpendicular to the straight line \(O_1O_2\) and intersect at points \(a, b \text{ and } c\), respectively. Thus, a series of related triangles such as \(\Delta PO_1O_2, \Delta P'O_1O_2a \text{ and } \Delta P'O_1Ob\) are obtained, and the principle of triangulation [34] can be derived:

\[
\begin{align*}
\frac{x'_1}{X_w} &= \frac{f}{Z_w} = \frac{-x'_2}{T_s - X_w} \\
\frac{y'_1}{Y_w} &= \frac{f}{Z_w} = \frac{y'_2}{f} \\
\frac{T_s}{Z_w} &= \frac{T - x'_1 + x'_2}{Z_w - f}
\end{align*}
\]

The three-dimensional coordinates of the target point can be obtained from the above equations (11), (12) and (13), however, a term \(x'_1 - x'_2\) is generated during the calculation, indicating the magnitude of the disparity value. Let \(x'_1 - x'_2 = d\), the following results are obtained:

\[
\begin{align*}
X_w &= T_s \frac{x'_1}{d} \\
Y_w &= T_s \frac{y'_1}{d} \\
Z_w &= T_s \frac{f}{d}
\end{align*}
\]

In this way, the three-dimensional coordinate values of the measurement target in the world coordinate system are obtained; the specific measurement results will be given in the experimental results and analysis section.

B. IMPLEMENTATION OF THE IMPROVED LOCALIZATION METHOD

In the visual system, under certain conditions, the traditional binocular localization method often takes more time to complete the localization task. The main reason for this is that the stereo matching process of the binocular images in the traditional binocular localization method is very time-consuming. The number of matching points is large, which results in a slower matching speed. Shortening the time of stereo matching and speeding up the stereo matching is key to solving the above problems. The improved binocular localization method described below aims to solve the problem with the two ideas of shortening the time of stereo matching and increasing the speed of stereo matching.

1) LOCALIZATION METHOD BASED ON REGION OF INTEREST MATCHING

In the image processing analysis we usually focus on the range of the target in the image, that is, the region of
interest (ROI), rather than the whole image. If the region of interest is introduced in the binocular vision localization method and the ROI is used to specify the range of the region to be analysed in the image, the image processing accuracy is increased to a certain extent, the processing time is reduced, and the efficiency of the image analysis is improved.

Inspired by the above points, a method of merging the stereo matching part in the traditional binocular localization method with the ROI is proposed, and a schematic diagram of the matching process is shown in Figure 13. First, the target range is determined in the pixel coordinate system of the left and right original images, leaving a certain margin, the ROI is obtained by the rectangular extraction method, and the extraction regions of the two images are defined as ROI$_1$ and ROI$_2$. Second, two new blank images with the same size as the original image are created and stored in a matrix. Again, the ROI is copied to the corresponding ROI$_1$ and ROI$_2$ in the new image by loading the image mask, and the ROI image with the background removed is obtained. Finally, the new ROI image is stereo-matched to obtain the disparity information, and the three-dimensional information of the measurement target can be obtained by using the triangulation principle.

At this point, the ROI matching process can be achieved by the following function relationship:

The ROI extraction equation is: $\mu : S \rightarrow R$.

\[
\begin{align*}
\text{ROI}_1 &= \mu(S_1) \\
\text{ROI}_2 &= \mu(S_2)
\end{align*}
\]  

(15)

The ROI image mask equation is: $\lambda : R \rightarrow F$.

\[
\begin{align*}
F_1 &= \lambda(\text{ROI}_1) \\
F_2 &= \lambda(\text{ROI}_2)
\end{align*}
\]  

(16)

The ROI image matching equation: $\delta : F \rightarrow D$.

\[
D = \delta(F_1, F_2)
\]  

(17)

where $S_1$ and $S_2$ are extraction parameters, $S_1, S_2 \in \mathbb{R}^{2048 \times 2048}$, ROI$_1$ and ROI$_2$ are the extraction results of the
region of interest. \( ROI_1, ROI_2 \in R^{h \times w}, w \in [0, 2448], h \in [0, 2048] \), \( F_1 \) and \( F_2 \) are the mask results, which uses \( w = 600, h = 400 \) and \( F_1, F_2 \in R^{2448 \times 2048} \). D is a disparity map, and \( D \in R^{2448 \times 2048} \). The ROI extraction equation achieves the extraction of the binocular image ROIs. The extraction result is used as the input of the ROI image masking equation, and a mask image with the same resolution as the original image is obtained. Then, the mask image is input into the ROI image matching equation to obtain the disparity information of the binocular images.

The disparity map of the binocular images obtained with the localization method using region of interest matching is shown in Figure 14.

2) LOCALIZATION METHOD BASED ON A PIXEL COORDINATE OFFSET

It can be seen from the analysis that the abovementioned localization method based on region of interest matching accelerates the speed and improves the efficiency of stereo matching to some extent, but the impact on the accuracy does not seem to be appreciable. Therefore, based on the ROI employed in the above method, a localization method based on a pixel coordinate offset is proposed, the principle of which is shown in Figure 10. First, we determine the approximate range of the measurement target in the pixel coordinate systems of the left and right images, determine the rectangular area where the target is located by using the ROI, and record the coordinates of the upper left corner vertex of the rectangular area in the left image in the pixel coordinate system \( o_1u_1v_1 \) as \( A_1(\Delta u_1, \Delta v_1) \); the coordinates of the upper left corner vertex of the rectangular area in the right image in the pixel coordinate system \( o_2u_2v_2 \) are \( A_2(\Delta u_2, \Delta v_2) \), where \( \Delta u_1, \Delta v_1, \Delta u_2, \) and \( \Delta v_2 \) are called offsets of pixels. Second, the ROIs in the left and right images are extracted as \( ROI_{11} \) and \( ROI_{22} \), respectively, to reduce the search domain. The target detected by the Hough transform determines that the spherical target has a pixel coordinate of \( x_1(u'_1, v'_1) \) in the \( ROI_{11} \) pixel coordinate system \( o_{11}u_{11}v_{11} \); and the pixel coordinate in the \( ROI_{22} \) pixel coordinate system \( o_{22}u_{22}v_{22} \) is \( x_2(u'_2, v'_2) \).

Then, the coordinate system is transformed: the coordinate \((u'_1, v'_1)\) in the \( o_{11}u_{11}v_{11} \) coordinate system is converted to \((u'_1 + \Delta u_1, v'_1 + \Delta v_1)\) in the \( o_1u_1v_1 \) coordinate system, and the coordinate \((u'_2, v'_2)\) in the \( o_{22}u_{22}v_{22} \) coordinate system is converted to \((u'_2 + \Delta u_2, v'_2 + \Delta v_2)\) in the \( o_2u_2v_2 \) coordinate system. Finally, combined with the imaging model, the three-dimensional information of the spherical target in the world coordinate system is obtained by the least square method. The reconstruction principle used to achieve a three-dimensional reconstruction of the target point is shown in Figure 15.

The imaging model of the camera [35] can obtain equations (18) and (19), which describe the imaging relationship of a certain point in the world coordinate system in the left and right camera coordinate systems, where \((u_1, v_1)\) and \((u_2, v_2)\) are the pixel coordinates of the left and right camera images, \((X_w, Y_w, Z_w)\) is the world coordinate, \(M_i(i = 1, 2)\) is the projection matrix, \(m_{jk}^n(n = 1, 2; j = 1, 2, 3; k = 1, 2, 3, 4)\) is the element of the \(j\)th row and the \(k\)th column of the projection matrix \(M_i(i = 1, 2)\). \(i\) is used to distinguish the projection matrices of the left and right cameras, and \(n\) is used to distinguish the elements in the left and right camera projection matrices. When \(n = i = 1, M_1\) and \(m_{jk}^n(j = 1, 2, 3; k = 1, 2, 3, 4)\) represent the left camera parameters. When \(n = i = 2, M_2\) and \(m_{jk}^n(j = 1, 2, 3; k = 1, 2, 3, 4)\) represent the right camera parameters. \(Z_{c1}\) and \(Z_{c2}\) are the parameter values of the \(Z\) direction in the left and right camera coordinate systems.

\[
\begin{align*}
Z_{c1} & \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = M_1 \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \\
& = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{18}
\end{align*}
\]

\[
\begin{align*}
Z_{c2} & \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = M_2 \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \\
& = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{19}
\end{align*}
\]

Now the above two equations can be sorted to eliminate \(Z_{c1}\) and \(Z_{c2}\) from the equations; then, introduce \(x'_1(u'_1 + \Delta u_1, v'_1 + \Delta v_1)\) and \(x'_2(u'_2 + \Delta u_2, v'_2 + \Delta v_2)\) into the above two equations:

Equation (20) written in matrix form is (21) and (22), as shown at the bottom of the next page.

The world coordinates of the measurement target obtained by the least square method [36] are:

\[
X = (M^T M)^{-1} M^T N \tag{23}
\]

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The main problem addressed in this paper is the accuracy of the measurement results and large time consumption of binocular visual localization for a large field of view and complex image background. The traditional binocular localization
method can localize a spherical target, but the method can be further improved in terms of localization accuracy and time efficiency. This section will use the binocular vision system described in Chapter 2 to verify the traditional binocular vision localization method and the two improved localization methods introduced in this paper by experiments with the same conditions. Then, we present the experimental results and a comparative analysis. The localization methods are evaluated based on two performance indexes: the localization accuracy and the localization time efficiency.

A. LOCALIZATION ACCURACY
For the localization accuracy, this paper considers a ping-pong ball as the spherical measurement target for an experimental verification by placing the ball, in the measurement field of view. The ball is positioned horizontally in 10 different positions, and the distance between each two adjacent positions is set to 22.5 mm. The world coordinate system can be defined at the position of the camera coordinate system. First, the distance between the front of the camera and the origin of the left camera coordinate system is measured with a ruler, so the coordinate of the Z direction is 700.0 mm; then, the length value of 5.0 mm is measured as the coordinates of the X direction; finally, the vertical height of the origin of the left camera coordinate system and the center of the ball is measured, so the coordinate of the Y direction is 42.0 mm. Thus, the actual coordinates of the first point are (−5.0, 42.0, 700.0), and the measurement error is required to be within 1 mm. The actual coordinates of all points are obtained according to the following three steps: (a) Move

\[
\begin{align*}
&\left[ (u_1' + \Delta u_1)m_{11} - m_{11} \right] X_w + \left[ (u_1' + \Delta u_1)m_{32} - m_{12} \right] Y_w + \left[ (u_1' + \Delta u_1)m_{33} - m_{13} \right] Z_w = m_{14} - m_{34} (u_1' + \Delta u_1) \\
&\left[ (v_1' + \Delta v_1)m_{31} - m_{31} \right] X_w + \left[ (v_1' + \Delta v_1)m_{32} - m_{32} \right] Y_w + \left[ (v_1' + \Delta v_1)m_{33} - m_{33} \right] Z_w = m_{34} (v_1' + \Delta v_1) \\
&\left[ (u_2' + \Delta u_2)m_{31} - m_{31} \right] X_w + \left[ (u_2' + \Delta u_2)m_{32} - m_{32} \right] Y_w + \left[ (u_2' + \Delta u_2)m_{33} - m_{33} \right] Z_w = m_{34} (u_2' + \Delta u_2) \\
&\left[ (v_2' + \Delta v_2)m_{31} - m_{31} \right] X_w + \left[ (v_2' + \Delta v_2)m_{32} - m_{32} \right] Y_w + \left[ (v_2' + \Delta v_2)m_{33} - m_{33} \right] Z_w = m_{34} (v_2' + \Delta v_2)
\end{align*}
\]

(20)

Define: 

\[
M = \begin{bmatrix}
(u_1' + \Delta u_1)m_{31} - m_{31} & (u_1' + \Delta u_1)m_{32} - m_{32} & (u_1' + \Delta u_1)m_{33} - m_{33} & (u_1' + \Delta u_1)m_{34} - m_{34} \\
(v_1' + \Delta v_1)m_{31} - m_{31} & (v_1' + \Delta v_1)m_{32} - m_{32} & (v_1' + \Delta v_1)m_{33} - m_{33} & (v_1' + \Delta v_1)m_{34} - m_{34} \\
(u_2' + \Delta u_2)m_{31} - m_{31} & (u_2' + \Delta u_2)m_{32} - m_{32} & (u_2' + \Delta u_2)m_{33} - m_{33} & (u_2' + \Delta u_2)m_{34} - m_{34} \\
(v_2' + \Delta v_2)m_{31} - m_{31} & (v_2' + \Delta v_2)m_{32} - m_{32} & (v_2' + \Delta v_2)m_{33} - m_{33} & (v_2' + \Delta v_2)m_{34} - m_{34}
\end{bmatrix}
\]

(21)

\[
X = \begin{bmatrix}
X_w \\
Y_w \\
Z_w
\end{bmatrix}, \quad N = \begin{bmatrix}
m_{14} - m_{34} (u_1' + \Delta u_1) \\
m_{14} - m_{34} (v_1' + \Delta v_1) \\
m_{14} - m_{34} (u_2' + \Delta u_2) \\
m_{14} - m_{34} (v_2' + \Delta v_2)
\end{bmatrix}
\]

Then,

\[
MX = N
\]

(22)
the ball on the plane, and (b) fix the three coordinate values in the XYZ direction by two. Then (c) the third coordinate value is incremented by 22.5mm steps. Therefore, the actual coordinate values of all the points can be obtained. Then, a set of binocular images is collected for the spherical targets in the above 10 positions. The resolution of each group of images is 2448*2048 pixels, and the collected partial images for the localization experiments are shown in Figure 16. First a calibration of the binocular cameras is required, the relationship for the conversion of the measurement target with respect to the world coordinates to the image coordinates is obtained, and the measurement model is established. The traditional localization method (T-LM), the localization method based on region of interest matching (ROI-LM) and the localization method based on the pixel coordinate offset (PCO-LM) are applied to perform the localization calculation. As shown in Table 3, the final measurement result of the localization experiment is obtained, the three-dimensional coordinates of the target in the world coordinate system with the three localization methods are obtained, and the error is calculated.

The error in the measured data in this experiment is measured by the Euclidean distance. The Euclidean distance between each set of measured values and the actual value is calculated separately, and the corresponding measurement error data are obtained. An error comparison of the three localization methods is shown in Figure 17, and the average error of each localization method for the experimental conditions in this paper is statistically analysed. The average error is shown in Figure 18.

Figure 17 shows that from the overall distribution of the error, the traditional localization method and the localization method based on region of interest matching exhibit large error fluctuations ranging from 3.99 mm to 12.48 mm and 3.71 mm to 11.69 mm, respectively. The localization method based on the pixel coordinate offset exhibits a small error fluctuation, fluctuating within the range of 1.50mm to 6.33mm and is relatively stable. The measurement method based on the pixel coordinate offset has the smallest measurement error. At the 4th, 5th and 6th measurement points, the measurement error of the traditional localization method is lower than that of the localization method based on region of interest matching, and at the 6th measurement point, the measurement error of the traditional localization method is almost the same as that of the localization method based on the pixel coordinate offset. The reason for this result may be that the effect of stereo matching is better, and a more accurate disparity value is obtained, thereby improving the accuracy of the measurement. As shown in Figure 18, the average errors of the above three methods are 8.63 mm, 8.36 mm, and 3.62 mm, respectively. The error in the traditional localization method is slightly larger than that of the localization method based on region of interest matching, and the errors of the two methods are relatively similar. The localization method based on the pixel coordinate offset exhibits less error, the accuracy
TABLE 3. Localization experiment results (unit:mm).

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Actual value</th>
<th>T-LM</th>
<th>Measurement results</th>
<th>ROI-LM</th>
<th>PCO-LM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(−5.0,42.0,700.0)</td>
<td>(−6.38,39.90,696.89)</td>
<td>(−6.29,40.50,696.86)</td>
<td>(−5.89,42.10,701.20)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>(17.5,42.0,700.0)</td>
<td>(14.67,36.61,692.56)</td>
<td>(16.88,36.50,693.86)</td>
<td>(16.65,40.90,697.88)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>(40.0,42.0,700.0)</td>
<td>(32.72,37.86,690.89)</td>
<td>(33.32,37.36,691.61)</td>
<td>(41.12,43.30,702.32)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>(62.5,42.0,700.0)</td>
<td>(69.11,47.37,704.89)</td>
<td>(68.41,46.39,706.82)</td>
<td>(63.45,43.52,704.31)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>(85.0,42.0,700.0)</td>
<td>(87.52,45.84,702.11)</td>
<td>(86.91,46.65,703.09)</td>
<td>(83.60,44.15,697.89)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>(85.0,42.0,722.5)</td>
<td>(82.54,40.96,719.21)</td>
<td>(79.91,39.36,713.66)</td>
<td>(83.51,45.06,720.18)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>(62.5,42.0,722.5)</td>
<td>(68.76,46.74,728.29)</td>
<td>(65.56,46.20,729.45)</td>
<td>(61.63,43.56,722.98)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>(62.5,42.0,745.0)</td>
<td>(58.68,42.17,755.92)</td>
<td>(56.16,46.31,753.19)</td>
<td>(64.58,45.30,748.23)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>(40.0,42.0,745.0)</td>
<td>(46.53,42.67,755.61)</td>
<td>(43.62,39.98,739.25)</td>
<td>(43.60,40.76,739.94)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>(17.5,42.0,745.0)</td>
<td>(12.85,37.22,743.13)</td>
<td>(12.42,38.45,746.74)</td>
<td>(15.79,40.34,741.86)</td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 19. Comparison of the time consumption of the three localization methods.

is improved by 56.8% compared with the accuracy of the traditional localization method, and the localization accuracy is significantly better than that of the other two methods. The localization method based on the pixel coordinate offset has obvious advantages in terms of localization accuracy, and its measurement accuracy is significantly improved compared with that of the traditional localization method and the localization method based on region of interest matching. The localization result within the measurement range is relatively stable, which can meet the measurement accuracy requirements of the astronaut assistant robot binocular vision system.

B. LOCALIZATION TIME EFFICIENCY

To verify that the improved binocular localization methods in this paper can effectively improve the localization efficiency of a spherical measurement target and to some extent shorten the localization time and improve the overall performance of the system, we use the three methods to locate the ping-pong ball in the above 10 different positions and also consider the time consumption problem. The time consumption of the three localization methods is shown in Figure 19. In theory, the fastest of the three methods should be the localization method based on the pixel coordinate offset, followed by the localization method based on region of interest matching and finally the traditional localization method. This order results from the fact that compared with the traditional localization method, fundamentally, the localization method based on the pixel coordinate offset saves a lot of time when the pixel coordinate of the spherical target position is determined after extracting the ROI on the image and is converted to the original image coordinates by the coordinate system. The localization method based on region of interest matching reduces the range of the stereo matching regions and background complexity.

Figure 19 shows that the localization method based on the pixel coordinate offset consumes the least amount of time, and the time consumption curve is located at the bottom of the distribution positions of the three curve. The localization method based on region of interest matching is second in terms of time consumption, and the traditional localization method consumes the most time. The actual time consumption measurement result is the same as the theoretical estimation result. At the 4th measurement point, the time consumption of the traditional localization method is slightly lower than that of the localization method based on region of interest matching. The possible reasons for this phenomenon are (a) interference of random noise and (b) the impact of
the extraction of the ROI area range inappropriate impact. The localization method based on the pixel coordinate offset suddenly increases the time consumption at the sixth measurement point, while the time consumption fluctuation at other measurement points is relatively stable. The possible reason is that the binocular image feature extraction process consumes more time. As shown in Figure 20, the average time consumptions of the three localization methods are 6.91 s, 6.61 s, and 4.66 s, respectively. The localization method based on region of interest matching improves the efficiency by approximately 43% compared with that of the traditional localization method. The localization method based on the pixel coordinate offset improves the efficiency by approximately 32.6% compared with that of the traditional localization method and has been greatly improved. In addition, noise interference, image resolution, and the size of the ROI area all affect the complexity of the image background and the time consumption to some extent. Therefore, our method can effectively improve the localization time efficiency.

It can be seen from the above analysis that the localization method based on the pixel coordinate offset has obvious advantages in terms of time efficiency. In the larger measurement field of view, the size of the ROI region is unchanged, and the noise is random. Compared with the conventional localization method and the localization method with region of interest matching, the measurement efficiency of the above method exhibits a large increase, which can meet the measurement efficiency requirements of the astronaut assistant robot binocular vision system.

V. CONCLUSION

In summary, this paper mainly solves the localization problem of the binocular vision system of an astronaut assistant robot. Based on the traditional localization method, two improved binocular localization methods are proposed: a localization method based on region of interest matching and a localization method based on a pixel coordinate offset. The innovations and contributions of this paper are as follows: (a) we develop a binocular vision experiment platform for an astronaut assistant robot, which is conducive to the verification and implementation of the method. (b) The localization method based on region of interest matching achieves the localization by removing the complex background of the image by the ROI and determining the stereo matching region, which improves the localization accuracy and reduces the time consumption to some extent. (c) On the basis of the ROI, the localization method based on the pixel coordinate offset uses the ROI to determine the search range to extract the target pixel coordinates, and the coordinate system is transformed by extracting the pixel coordinate offset to complete the localization. The method has a high localization accuracy and low time consumption. The experimental results show that our methods can effectively improve the localization accuracy and reduce the time consumption and improve the localization efficiency and meet the localization requirements. In addition, the localization method based on the pixel coordinate offset is especially prominent in terms of localization accuracy and time efficiency. Therefore, the method is applied to an astronaut assistant robot binocular vision system to provide position information.

Future work will mainly includes (a) optimizing the localization methods of this paper, (b) identifying, tracking and locating dynamic targets, focusing on the localization accuracy and efficiency, and performing a dynamic target experiment verification.

REFERENCES


FENGWEI GU was born in Heilongjiang, China, in 1992. He received the B.S. degree from Shenyang Ligong University, China, in 2016, where he is currently pursuing the M.S. degree. His research interests include digital image processing and machine vision.

HONGWEI GAO received the Ph.D. degree in pattern recognition and intelligent system from the Shenyang Institute of Automation (SIA), Chinese Academy of Sciences (CAS), in 2007. Since September 2015, he has been a Professor of the School of Automation and Electrical Engineering, Shenyang Ligong University. He is currently the leader of academic direction for optical and electrical measuring technology and systems. His research interests include digital image processing and analysis, stereo vision, and intelligent computation. He has published more than 60 technical articles in these areas as first author or coauthor.

JINGUO LIU (M’07–SM’18) received the Ph.D. degree in mechatronics from the Shenyang Institute of Automation (SIA), Chinese Academy of Sciences (CAS), in 2007, where he has been a Full Professor, since January 2011, has also been the Assistant Director of the State Key Laboratory of Robotics, since 2008, and has also been the Associate Director of the Center for Space Automation Technologies and Systems, since 2015. His research interests include bio-inspired robotics and space robot. He has authored or coauthored three books, over 100 articles and holds 50 patents in above areas.

He is a Senior Member of the IEEE Technical Committee on Safety, Security, and Rescue Robotics, the IEEE Technical Committee on Marine Robotics, and the Senior Member of the Chinese Mechanical Engineering Society. He was a recipient of the T. J. TARN Best Paper Award in Robotics from the 2005 IEEE International Conference on Robotics and Biomimetics, the Best Paper Award of the Chinese Mechanical Engineering Society, in 2007, the Best Paper Nomination Award from the 2008 International Symposium on Intelligent Unmanned Systems, the Best Paper Award from the 2016 China Manned Space Academic Conference, the Outstanding Paper Award from the 2017 International Conference on Intelligent Robotics and Applications, and the Best Paper Award from the 2018 International Conference on Electrical Machines and Systems. He serves as the Associate Editor for several journals, such as IEEE ACCESS, Mechanical Sciences, Advance in Mechanical Engineering, and the Chinese Journal of Mechanical Engineering.

QING GAO was born in Tangshan, China. He received the B.S. degree in automation from the Electrical Engineering and Automation School, Liaoning Technology University, China, in 2013. He is currently pursuing the Ph.D. degree with the State Key Laboratory of Robotics, Shenyang Institute of Automation (SIA), Chinese Academy of Sciences (CAS), Shenyang, China. His research interests include space robot, artificial intelligence, machine vision, and human–robot interaction. He has authored or coauthored more than ten publications in journals and conference proceedings in above areas and received One Outstanding Paper Award.

ZHAOJIE JU (M’08–SM’16) received the B.S. degree in automatic control and the M.S. degree in intelligent robotics from the Huazhong University of Science and Technology, China, and the Ph.D. degree in intelligent robotics from the University of Portsmouth, U.K. He held research appointments at University College London, London, U.K., before he started his independent academic position at the University of Portsmouth, in 2012. He has authored or coauthored over 170 publications in journals, book chapters, and conference proceedings and received four Best Paper Awards and one Best AE Award in ICRA2018. His research interests include machine intelligence, pattern recognition and their applications on human motion analysis, multi-fingered robotic hand control, human–robot interaction and collaboration, and robot skill learning.


* * *