Proceedings of 2014 IEEE
International Conference on Mechatronics and Automation
August 3 - 6, Tianjin, China

# The Identifier-based Relative Position Estimation for Leader-follower Robotic System

Lingyi Xu[1, 2], Zhiqiang Cao[1], Peng Zhao[1]

[1]*State Key Laboratory of Management and Control for Complex Systems*, *Institute of Automation, Chinese Academy of Sciences*, *Beijing* 100190, *China*

lingyixu@yahoo.com, {zhiqiang.cao, peng.zhao}@ia.ac.cn

Yixin Yin[2]

[2]*School of Automation and Electrical Engineering University of Science and Technology Beijing* *Beijing* 100083, *China*

yyx@ies.ustb.edu.cn

***Abstract* - A monocular vision measurement approach is presented for leader-follower robotic system to estimate the position of the leader robot relative to the follower robot. The color identifier with green and red blocks is attached on the back of the leader robot, and it may be captured by the camera mounted on the follower robot. A coarse-to-fine searching method is given to extract the center point of the identifier. The parameters of the camera with the combination of the intrinsic and extrinsic ones are calibrated via least square method with given relative positions and the corresponding image coordinates of the identifier's center. The parameters are further optimized with Levenberg-Marquardt method. Then the position of the leader robot relative to the follower robot can be estimated only with single feature point. The experimental results verify the validity of the given approach.**

***Index Terms –Visual measurement, feature extraction, camera calibration, monocular vision, leader-follower robotic system.***

## I. Introduction

Multi-robot system has many merits such as redundancy, flexibility, parallelism as well as swarm intelligence, which attracts more and more attentions [1]-[8]. For some tasks, especially complex ones, similar to human society, it is advantageous to task execution if the robots are assigned different roles, e.g. leader, follower. For such leader-follower robotic system, the position estimation of the leader robot relative to the follower robot is a basis to realize the given task corporately.

The visual measurement approach is a good selection for the leader-follower robotic system working in indoor environments. The existing visual localization methods for indoor mobile robots including global camera based method [10] and on-board method [9] as well as the combination of them [11] according to the positions of the cameras mounted. It is common for a global camera fixed on the ceiling monitoring downward to the floor. Nascimento *et al*. use a ceiling camera to locate the mobile robot on the floor [10]. The mapping from the floor to the image plane is a linear function. The calibration of the camera is to determine the proportional factor from the image coordinates to the floor coordinates. The positions of the leader and follower robots can be measured simultaneously. The visual model is simple and the measurement has satisfied accuracy. However, the moving area of the robots is limited by the effective view field of the fixed camera. The robots with on-board camera may extend its working area effectively. In this case, the position estimation of the leader robot relative to follower robot may be achieved by certain identifiers. [12] adopts a two-color cylindrical color coded identifier for robot identity recognition. For the purpose of the richness and recognition adaptability, a general symbol identifier is designed [13]. It is composed of a central area and a peripheral area, and there exists radial spokes in the central area.

In this work, a leader-follower robotic system with two robots is considered. One robot is taken as the leader, which is attached a color identifier. The other one is considered to the follower with a camera fixed on the front of its body. The leader and follower robots are individually controlled by their own controllers. The monocular vision measurement is considered, and the relative position of the leader robot is estimated with single feature point of the identifier center, which is extracted by a coarse-to-fine searching.

The rest of the paper is organized as follows. In section II, the feature extraction with coarse and fine searching stages is given. The calibration method for the monocular vision system is given in section III, and the visual measurement is also presented. Experiments and results are provided in section IV. Finally, the paper is concluded in section V.

## II. Feature Extraction

The color identifier is designed as shown in Fig. 1. It consists of green and red color rectangles. The center point of the identifier is used as the feature point. The follower robot with a CMOS camera calculates the coordinates of the leader robot relative to it using the feature point perceived by its vision, which provides the basis for following coordination.

Fig. 1 The designed identifier

Feature extraction means to find the center point of the identifier in the image given in Fig. 1. First of all, the dynamic threshold value $G_T(i, j)$ is calculated from the RGB value of a pixel as given, which is used to judge whether the pixel is red, green or not.

$$G_T(i,j) = [612R(i,j) + 1202G(i,j) + 233B(i,j)]/2^{11}$$

where $R(i, j)$, $G(i, j)$, and $B(i, j)$ are the red, green, and blue value of a pixel $(i, j)$.

In the following, a coarse-to-fine searching method is designed for feature extraction, which includes two stages: coarse and fine searching stages. The former searches the image in large interval to find all ROIs (Region of Interest), whereas the latter searches the ROIs in small interval to obtain the result.

The image is searched with the interval of $k_s$ (these points constitute a set $\boldsymbol{\Phi}_s$) to find the red pixel point which is labelled as (i, j). On this basis, the pixel $(i+k_i, j)$ is judged whether it is green. If it is satisfied, the pixels $(i+k_i, j-k_i)$ and $(i, j-k_i)$ are checked successively as red and green, respectively, as shown in Fig. 2. If all fundamental aspects are satisfied, the pixel $(i+k_i/2, j-k_i/2)$ is recorded as mark point (see Function $P_{mark}$) that is utilized to obtain a ROI whose maximal size is 47×47 with the constraint from the borderline condition.

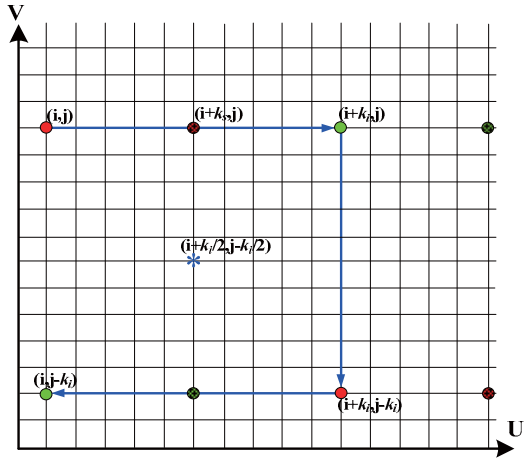



Fig. 2 The searching of mark point

**Function** $C=P_{mark}((i, j), \boldsymbol{\Phi}, k_i)$

```
C←(0, 0);
for each pixel point (i, j) in Φ do
  if (match_r((i, j), red)) then
    if (match_g((i+k_i, j), green)) then
      if (match_r((i+k_i, j-k_i), red)) then
        if (match_g((i, j-k_i), green)) then
          C←(i+k_i/2, j-k_i/2);
        end
      end
    end
  end
end
```

After all ROIs corresponding to the mark points are acquired, coarse searching stage finishes. In fine searching stage, only the points in ROIs will be considered. A new-round searching with minor $k_s$ and $k_i$ similar to that in coarse stage is conducted to obtain a series of new mark points. The average value of the mark points corresponding to the ROI that contains maximum number of mark points is regarded as the coordinate of the feature point.

The detailed algorithm is given in Algorithm 1, where the pixel point is red when $R(i,j) > 1.15G_T(i,j) \cap G(i,j) \le 1.15G_T(i,j) \cap B(i,j) < 1.15G_T(i,j)$, and it is green if $G(i,j) > 1.15G_T(i,j) \cap R(i,j) < 1.15G_T(i,j) \cap B(i,j) < 1.15G_T(i,j)$. It is noted that $G_T(i, j)$ is enlarged to 1.15 times to reduce the influence of large weighted green component caused by the CMOS camera itself.

**Algorithm 1**. **The Center Point of the Identifier**

**Input**: $R$, $G$ and $B$ component of all pixels.
**Output**: the coordinate $(u_f, v_f)$ of the feature point.

```
1  Φ_{c,m}←0;
2  C_m←(0, 0);
3  N_temp=0;
4  Φ_temp←0;
5  for each pixel point (i, j) in Φ_s do
6    if (!match(C_m←P_mark((i, j), Φ_s, 10), (0, 0))) then
7    Φ_{c,m}←C_m;
8    end
9  end
10 for each pixel point P_{c,m} in Φ_{c,m} do
11   N←0;
12   Φ_{f,m}←0;
13   for each pixel point (p, q) in Φ_ls|R_OI(P_{c,m})
14     if (!match(C_m←P_mark((p, q), Φ_ls, 5), (0, 0))) then
15       Φ_{f,m}←C_m;
16       N←N+1;
17     end
18   end
19   if (N_temp<N) then
20       N_temp=N; Φ_temp=Φ_{f,m};
21   end
22 end
23 (u_f, v_f)=average(Φ_temp).
```

## III. Vision System Calibration and Measurement

### A. Model and Calibration

As is known to all, the intrinsic parameters of a camera can be expressed as

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \quad (1)$$

where $k_x$, $k_y$ are the magnification factors along X- and Y-axis, $(u, v)$ is the image coordinates of the feature point, $(u_0, v_0)$ is the image coordinates of primary point, $(x_c, y_c, z_c)$ is the Cartesian coordinates of the feature point in the camera frame.

The extrinsic parameters of the camera describe the relationship of the camera frame and the world frame in Cartesian space. It can be expressed as

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} n_x & o_x & a_x & p_x \\ n_y & o_y & a_y & p_y \\ n_z & o_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{2}$$

where $[n_x\ n_y\ n_z]^T$, $[o_x\ o_y\ o_z]^T$, $[a_x\ a_y\ a_z]^T$ are the unit vectors of the $X_w$-, $Y_w$- and $Z_w$-axis of the world frame expressed in the camera frame, $[p_x\ p_y\ p_z]^T$ is the position vector of the origin of the world frame described in the camera frame, $(x_w, y_w, z_w)$ is the coordinates of the feature point in the world frame.

With the consideration of the leader and follower robot moving at the same plane, combining (1) and (2), we have [14]:

$$\begin{bmatrix} x_w & y_w & 1 & 0 & 0 & 0 & -x_w u & -y_w u \\ 0 & 0 & 0 & x_w & y_w & 1 & -x_w v & -y_w v \end{bmatrix} m = \begin{bmatrix} u \\ v \end{bmatrix} \tag{3}$$

where $m = \begin{bmatrix} m_1 & m_2 & m_3 & m_4 & m_5 & m_6 & m_7 & m_8 \end{bmatrix}^T$

***m*** in (3) is the vision system parameters vector, including the intrinsic and extrinsic parameters. It is conspicuous that each position of a feature point can offer two equations with ***m***. Thus ***m*** can be resolved with at least four different positions of the feature point. Of course, more positions are helpful to improve the estimation accuracy of ***m***.

Let

$$A = \begin{bmatrix} x_w & y_w & 1 & 0 & 0 & 0 & -x_w u & -y_w u \\ 0 & 0 & 0 & x_w & y_w & 1 & -x_w v & -y_w v \end{bmatrix}, B = \begin{bmatrix} u \\ v \end{bmatrix}$$

then (3) can be expressed as (4):

$$Am = B \tag{4}$$

Least square method is employed to solve $m$, as given in (5) [14].

$$m = (A^T A)^{-1} A^T B \tag{5}$$

*B. Measurement*

Once the vector ***m*** is known, the relative position can be calculated with the image coordinates of the feature point. The measurement formula is from (3).

$$\begin{bmatrix} m_1 - u_i m_7 & m_2 - u_i m_8 \\ m_4 - v_i m_7 & m_5 - v_i m_8 \end{bmatrix} \begin{bmatrix} X_{mi} \\ Y_{mi} \end{bmatrix} = \begin{bmatrix} u_i - m_3 \\ v_i - m_6 \end{bmatrix} \tag{6}$$

where $(X_{mi}, Y_{mi})$ is the measured coordinates of the leader robot relative to the follower robot, $(u_i, v_i)$ is the feature point's image coordinates at the i-th measurement.

Let

$$\begin{cases} a_1 = m_1 - u_i m_7, \quad b_1 = m_2 - u_i m_8, \quad c_1 = u_i - m_3 \\ a_2 = m_4 - v_i m_7, \quad b_2 = m_5 - v_i m_8, \quad c_2 = v_i - m_6 \end{cases} \tag{7}$$

then we have

$$\begin{cases} X_{mi} = \dfrac{b_2 c_1 - b_1 c_2}{a_1 b_2 - a_2 b_1} \\ Y_{mi} = \dfrac{a_1 c_2 - a_2 c_1}{a_1 b_2 - a_2 b_1} \end{cases} \tag{8}$$

*C. Parameters Optimization*

A performance function is established with actual and measured positions in order to optimize the parameters.

$$F = \frac{1}{2}\sum_{i=1}^{n}(X_{mi} - X_{wi})^2 + \frac{1}{2}\sum_{i=1}^{n}(Y_{mi} - Y_{wi})^2 \tag{9}$$

where $(X_{wi}, Y_{wi})$ is the i-th actual coordinates of the leader relative to the follower, $n$ is the number of optimization point.

The purpose of optimization is to find the minimal value of $F$ by adjusting parameters in ***m***. The derivations of formula (8) relative to $a_1$, $b_1$, $c_1$, $a_2$, $b_2$, and $c_2$ are given in (10), with $w=a_1b_2 - a_2b_1$.

$$\begin{cases} \dfrac{\partial X_{mi}}{\partial a_1} = -\dfrac{b_2 X_{mi}}{w}, \quad \dfrac{\partial X_{mi}}{\partial a_2} = \dfrac{b_1 X_{mi}}{w}, \quad \dfrac{\partial X_{mi}}{\partial b_1} = \dfrac{a_2 X_{mi}}{w} - \dfrac{c_2}{w} \\ \dfrac{\partial X_{mi}}{\partial b_2} = \dfrac{c_1}{w} - \dfrac{a_1 X_{mi}}{w}, \quad \dfrac{\partial X_{mi}}{\partial c_1} = \dfrac{b_2}{w}, \quad \dfrac{\partial X_{mi}}{\partial c_2} = -\dfrac{b_1}{w} \\ \dfrac{\partial Y_{mi}}{\partial a_1} = \dfrac{c_2}{w} - \dfrac{b_2 Y_{mi}}{w}, \quad \dfrac{\partial Y_{mi}}{\partial a_2} = -\dfrac{c_1}{w} + \dfrac{b_1 Y_{mi}}{w}, \quad \dfrac{\partial Y_{mi}}{\partial b_1} = \dfrac{a_2 Y_{mi}}{w} \\ \dfrac{\partial Y_{mi}}{\partial b_2} = -\dfrac{a_1 Y_{mi}}{w}, \quad \dfrac{\partial Y_{mi}}{\partial c_1} = -\dfrac{a_2}{w}, \quad \dfrac{\partial Y_{mi}}{\partial c_2} = \dfrac{a_1}{w} \end{cases} \tag{10}$$

The derivations of ***m*** are concluded by taking formula (7) and (10) into (9).

$$\begin{cases} \dfrac{\partial F}{\partial m_1} = \sum_{i=1}^{n}(X_{mi} - X_{wi})\dfrac{\partial X_{mi}}{\partial a_1} + \sum_{i=1}^{n}(Y_{mi} - Y_{wi})\dfrac{\partial Y_{mi}}{\partial a_1} \\ \dfrac{\partial F}{\partial m_2} = \sum_{i=1}^{n}(X_{mi} - X_{wi})\dfrac{\partial X_{mi}}{\partial b_1} + \sum_{i=1}^{n}(Y_{mi} - Y_{wi})\dfrac{\partial Y_{mi}}{\partial b_1} \\ \dfrac{\partial F}{\partial m_3} = -\sum_{i=1}^{n}(X_{mi} - X_{wi})\dfrac{\partial X_{mi}}{\partial c_1} - \sum_{i=1}^{n}(Y_{mi} - Y_{wi})\dfrac{\partial Y_{mi}}{\partial c_1} \\ \dfrac{\partial F}{\partial m_4} = \sum_{i=1}^{n}(X_{mi} - X_{wi})\dfrac{\partial X_{mi}}{\partial a_2} + \sum_{i=1}^{n}(Y_{mi} - Y_{wi})\dfrac{\partial Y_{mi}}{\partial a_2} \\ \dfrac{\partial F}{\partial m_5} = \sum_{i=1}^{n}(X_{mi} - X_{wi})\dfrac{\partial X_{mi}}{\partial b_2} + \sum_{i=1}^{n}(Y_{mi} - Y_{wi})\dfrac{\partial Y_{mi}}{\partial b_2} \\ \dfrac{\partial F}{\partial m_6} = -\sum_{i=1}^{n}(X_{mi} - X_{wi})\dfrac{\partial X_{mi}}{\partial c_2} - \sum_{i=1}^{n}(Y_{mi} - Y_{wi})\dfrac{\partial Y_{mi}}{\partial c_2} \\ \dfrac{\partial F}{\partial m_7} = -\sum_{i=1}^{n}(X_{mi} - X_{wi})(\dfrac{\partial X_{mi}}{\partial a_1}u_i + \dfrac{\partial X_{mi}}{\partial a_2}v_i) - \sum_{i=1}^{n}(Y_{mi} - Y_{wi})(\dfrac{\partial Y_{mi}}{\partial a_1}u_i + \dfrac{\partial Y_{mi}}{\partial a_2}v_i) \\ \dfrac{\partial F}{\partial m_8} = -\sum_{i=1}^{n}(X_{mi} - X_{wi})(\dfrac{\partial X_{mi}}{\partial b_1}u_i + \dfrac{\partial X_{mi}}{\partial b_2}v_i) - \sum_{i=1}^{n}(Y_{mi} - Y_{wi})(\dfrac{\partial Y_{mi}}{\partial b_1}u_i + \dfrac{\partial Y_{mi}}{\partial b_2}v_i) \end{cases} \tag{11}$$

The adjustment of ***m*** based on Levenberg-Marquardt optimization algorithm [15] is

$$m(j+1) = m(j) + L_s (H + \lambda I)^{-1} \nabla F_m \qquad (12)$$

where $\boldsymbol{m}(j)$ is the $\boldsymbol{m}$ at the $j$-th optimization step, $L_s$ is the step length factor, $\lambda$ is a positive constant, $\boldsymbol{H}$ is Hessian matrix, $\boldsymbol{I}$ is an unit matrix with the same size of $\boldsymbol{H}$, $\nabla \boldsymbol{F}_m$ is the partial derivations of $\boldsymbol{F}$ relative to $\boldsymbol{m}$, as given in (11).

## IV. Experiment and Results

A leader-follower robotic system has been designed, as shown in Fig. 3. A series of experiments are conducted to verify the given approach.

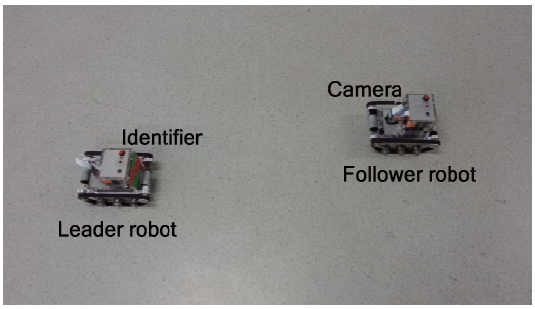



Fig. 3 The leader-follower robotic system

### A. Feature Point Extraction

The designed color identifier is captured by the embedded CMOS camera of follower robot. A series of images containing the identifier are captured in various circumstances where the leader robot is at different positions relative to the follower under different illuminations. The feature extraction results for four frames of images are shown in Fig. 4, where the extracted features are marked with a white square. Fig. 4(a) gives a normal image; Fig. 4(b) is the image with large weighted green component; Fig. 4(c) and Fig. 4(d) give the images with weak illumination and bright illumination, respectively. It can be seen from Fig. 4 that the given approach can extract the feature point accurately and it is robust to color aberration and the illumination changes.

(a) normal image

(b) Image with large weighted green component

(c) Image with weak illumination

(d) Image with bright illumination

Fig. 4 The feature extraction results

### B. Camera Calibration and Measurement

The world frame is established at the central point in the front of the follower robot, where $X_w$ axis is selected the direction from left to right in bird view; $Y_w$ axis is the forward direction of the follower robot; $Z_w$ axis is upward.

The leader is placed at ten different positions in front of the follower robot. The images are captured and the feature points are extracted. The actual relative positions of the center of the identifier on the leader robot are manually measured. Based on our approach, $\boldsymbol{m}$ is calculated and we have

$$m^T = \begin{bmatrix} -16.4151 & -13.0372 & 363.2342 & 1.0073 \\ -15.2553 & 414.9922 & 0.0037 & -0.0386 \end{bmatrix}.$$

The positions calculated via (8) with the image feature coordinates and the parameters $\boldsymbol{m}$ are taken as the measured positions. The results are shown in Fig. 5. It can be found from Fig. 5 that the measured positions coincide with the actual positions in the central area of $X_w$-axis direction within an acceptable range for images captured by CMOS camera. The deviation tendency is also correct for the positions far away the central area.

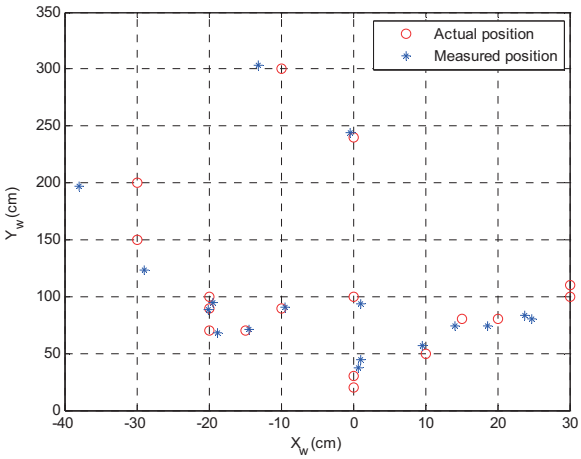
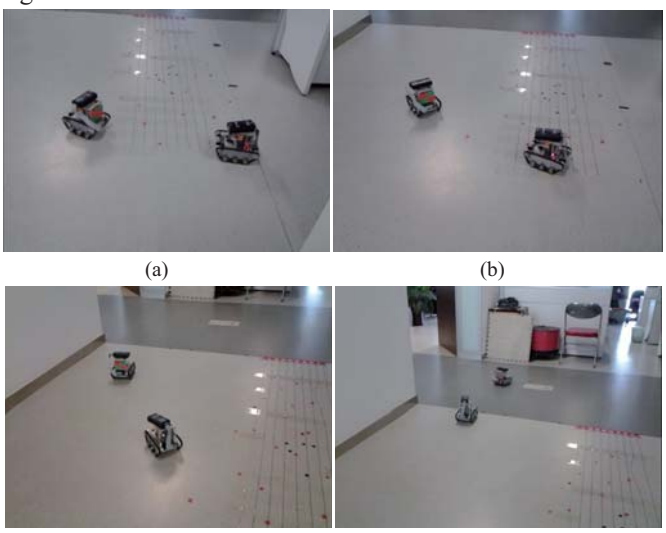



Fig. 5 The actual and measured positions

### C. *The Following Experiment*

To testify the identifier-based relative position estimation, a following experiment is given where the leader is controlled by an operator and the follower is required to follow the leader. The video snapshots of the experiment are shown in Fig. 6.



(a) (b)

(c) (d)

Fig. 6 The video snapshots of the experiment

## V. Conclusion

A visual measurement method based on monocular vision is presented to estimate the position of the leader robot relative to the follower robot with single image feature point of the center of the identifier attached on the leader robot. In order to extract the center of the identifier, a two-stage searching including the coarse and fine searching is designed. It is robust to the disturbances such as the color and illumination variations. The validity of the approach is verified by experiments. In future, the image-based visual servoing combining with the measured relative positions will be investigated.

## Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grants 61273352, 61175111, 60805038.

## References

[1] A. Bicchi, A. Fagiolini, L. Pallottino, "Towards a society of robots," *IEEE Robotics & Automation Magazine*, vol. 17, no. 4, pp. 26-36, 2010.

[2] S. Liu, D. Sun, C. Zhu, "Coordinated motion planning for multiple mobile robots along designed paths with formation requirement," *IEEE/ASME Transactions on Mechatronics*, vol. 16, no. 6, pp. 1021-1031, 2011.

[3] A. Drenner, M. Janssen, A. Kottas, *et al*., "Coordination and longevity in multi-robot teams involving miniature robots," *Journal of Intelligent and Robotic Systems*, vol. 72, no. 2, pp. 263-284, 2013.

[4] L. Parker, "Distributed intelligence: Overview of the field and its application in multi-robot systems," *Journal of Physical Agents*, vol. 2, no. 1, pp. 5-14, 2008.

[5] Z. Cao, M. Tan, L. Li, N. Gu, S. Wang, "Cooperative hunting by distributed mobile robots based on local interaction," *IEEE Transactions on Robotics*, vol. 22, no. 2, 403-407, 2006.

[6] P. Zhao, Z. Cao, X. Chen, D. Xu, "A docking approach for child robot based on vision guiding by mother robot," *Journal of Huazhong University of Science and Technology (Natural Science Edition)*, vol. 41, no. SUPPL.I, pp. 429-431, 435, 2013. (in Chinese)

[7] H. Hourani, P. Wolters, E. Hauck, *et al*., "A marsupial relationship in robotics: a survey," *Lecture Notes in Computer Science*, vol. 7101, pp. 335-345.

[8] J. P. Hsiao, S. S. Yeh, P. L. Hsu, "Target position estimation using multi-vision system implemented on distributed mobile robots," *International Journal of Robotics and Automation*, vol. 28, no. 2, pp. 154-169, 2013.

[9] L. Yang, Z. Cao, M. Tan, G. Liu, "Vision-based fuzzy coordination control for multiple robots," *IEEE International Conference on Robotics, Automation and Mechatronics*, pp. 521-524, 2008.

[10] R. C. A. Nascimento, B. M. F. Silva, L. M. G. Gonçalves, "Real-time localization of mobile robots in indoor environments using a ceiling camera structure," *IEEE Latin American Robotics Symposium*, pp. 61-66, 2013.

[11] Q. X. Yu, W. X. Yan, Z. Fu, Y. Z. Zhao, "Service robot localization based on global vision and stereo vision," *Journal of Donghua University (English Edition)*, vol. 29, no. 3, pp. 197-202, 2012.

[12] W. W. Zhang, J. Wang, Z. Q. Cao, Y. Yuan, C. Zhou, "A Local Interaction Based Multi-robot Hunting Approach with Sensing and Modest Communication," *ICIRA 2009*, LNAI 5928, pp. 90-99, 2009.

[13] H. B. Qian, W. B. Yuan, X. L. Liu, *et al*., "A Symbol Identifier Based Recognition and Relative Positioning Approach Suitable for Multi-robot Systems," C.-Y. Su, S. Rakheja, H. Liu (Eds.): ICIRA 2012, Part III, LNAI 7508, pp. 523-531, 2012.

[14] Y. J. Yin, D. Xu, Z. T. Zhang, X. G. Wang, W. T. Qu, "Plane measurement based on monocular vision," *Journal of Electric Measurement and Instrument*, vol. 27, no. 4, pp. 347-352, 2013. (in Chinese)

[15] D. H. Le, C. K. Pham, T. T. T. Nguyen, T. T. Bui, "Parameter extraction and optimization using Levenberg-Marquardt algorithm," *4th International Conference on Communications and Electronics*, pp. 434-437, 2012.