

Received 18 November 2022, accepted 24 November 2022, date of publication 28 November 2022, date of current version 6 January 2023.

*Digital Object Identifier 10.1109/ACCESS.2022.3225421*

# **RESEARCH ARTICLE**

# Robot Target Location Based on the Difference in Monocular Vision Projection

ZHAOHUI ZHENG<sup>®[1](https://orcid.org/0000-0003-2981-8976)</sup>, YUANYUAN LI<sup>®1</sup>, DEN[G](https://orcid.org/0000-0002-7383-4167) CHEN<sup>2</sup>, LEI WANG<sup>®3</sup>, JIANPING JU $^{\rm 4}$ , AND QIANG ZHANG $^{\rm 5}$ 

<sup>1</sup> School of Mathematics and Physics, Wuhan Institute of Technology, Wuhan 430079, China <sup>2</sup>Hubei Province Key Laboratory of Intelligent Robot, Wuhan Institute of Technology, Wuhan 430079, China

<sup>3</sup>School of Electrical and Information Engineering, Wuhan Institute of Technology, Wuhan 430079, China

<sup>4</sup>School of Artificial Intelligence, Hubei Business College, Wuhan 430070, China

<sup>5</sup>Department of Public Courses, Wuhan Railway Vocational College of Technology, Wuhan 430079, China

Corresponding author: Yuanyuan Li (yuanyuanli\_wit@hotmail.com)

This work was supported in part by the Science and Technology Research Project of Hubei Education Department under Grant B2022056, and in part by the Science Research Project of Wuhan Institute of Technology under Grant K2021083.

**ABSTRACT** Visual guidance is widely used in industrial robots. At present, the traditional method of template matching and positioning is often used in industrial robots under monocular vision guidance. However, for complex workpieces with height differences, if the angle and position of the workpiece are not consistent with the template, the projection will be different, and the positioning accuracy will be obviously reduced. A new method is proposed to solve this problem by dividing the whole contour of the workpiece and using the weighted method of the contour module to correct and match the workpiece image. First, the contour region of the template image is divided according to the position of the grasping points. Then the weight is assigned according to the distance from each contour region to the grasping points. Then, a fast feature point matching method is used to match the initial image of the workpiece to be measured. Finally, accurate contour matching is carried out for each weighted contour region so that the robot can grasp the object accurately. A large number of experiments show that the method has the design requirements of high stability and high accuracy.

**INDEX TERMS** Robot grapping, projection difference, contour matching, weighted module decision.

#### **I. INTRODUCTION**

With the rapid development of robot technology, robots have been widely used in industrial production. At present, there are robots in the fields of automobiles, electronics, metals and machinery; robots will eventually replace artificial production and become an important development trend in the manufacturing industry and the foundation of industrial intelligence. The traditional industrial robot realizes repetitive action through teaching and solving the problem of efficiency in mass production, but emerging manufacturing industries are moving toward more varieties and smaller batches. Therefore, traditional manual teaching

The associate editor coordinating th[e re](https://orcid.org/0000-0001-9909-0664)view of this manuscript and approving it for publication was Shuai Liu<sup>D</sup>.

and fixed-action robots cannot meet the needs of the emerging manufacturing industry [1], [2]. Machine vision can not only replace artificial vision, but also analyze and explain through visual processing [3], [11]. As a result, the cooperation between machine vision and robots is developed under a background, such as automatic grasping of mechanical parts and automatic sorting of logistics.

Bin-picking systems generally use 2D or 3D sensors, each of which has advantages and disadvantages: 1) 2D sensors are much cheaper than 3D sensors. 2) 2D sensors are more robust than 3D sensors in material suitability. 3) 3D sensors can obtain information on target height, but 2D sensors cannot. Because 2D sensors cannot obtain the target height, a projection difference will occur in the imaging process of 2D sensors for complex three-dimensional workpieces in



**FIGURE 1.** The robot grabbing system and imaging effect. The right red box area obviously displays more image information than the left red box area in the figure b.

industry; the difference leads to a decrease in the grasping precision of the robot, as shown in Fig. 1. Especially when the grab point is the inclined plane of the workpiece, the influence is more obvious.

In the early stage, we proposed an ADK algorithm based on feature matching for robot grasping [5], which uses the distance difference and angle difference of feature points for clustering to obtain accurate registration, and proposed a non-linear multi-parallel surface calibration method based on adaptive k-segment master curve algorithm, but did not consider the difference in monocular vision projection. In order to solve the problem, we propose a new method based on Reference [5], which combines feature matching and contour matching and uses the weighted method of the contour module to correct the image matching. Our contribution in this paper includes the following aspects:

(i) The combination of feature point matching and contour matching is used for fast location.

(ii) The weighted decision-making module is constructed to correct the matching results of the grab point region.

# **II. RELATED WORK**

Machine vision is widely used in various fields [6], [7], [8], [9]. Vision recognition and positioning technology is a typical application of machine vision in industrial automation, which is widely used in automatic assembly, moving objects on production lines and defect detection in products [10], [11], [12]. The use of visual sensors for object detection and attitude estimation is still an active research area, especially in the application of industrial robots.

Dekel et al. [13] proposed a template matching algorithm for the best similarity points, which can effectively solve the problem of image matching and positioning in the case of similar area interference and partial occlusion. However, the matching algorithm is complex, slow and obtains poor results in real time. Xu et al. [14] proposed an intelligent watermelon recognition and positioning method in a natural scene and designed a squint binocular positioning algorithm by taking the left camera as the coordinate origin to obtain the actual three-dimensional spatial coordinates of the watermelon. However, the recognition error was approximately 15%. In [15], Yang et al. proposed a new learning framework that segmented the learned skills into a sequence of subskills automatically; then, each individual subskill was encoded and regulated accordingly. However, skill representation, trajectory alignment, and skill segmentation could not be addressed satisfactorily using such skill transfer modeling. Gu et al. [16] developed a portable assembly demonstration (PAD) system by using a RGB-D camera, which automatically identified the involved objects (parts/tools), the actions performed and the assembly state representing the spatial relationship between parts. Fazeli et al. [17] proposed a method of simulating hierarchical reasoning and multisensory fusion in robots. This method used a camera and wrist strap to receive visual and tactile feedback and compared these measured values with the previous actions of the robot. Meanwhile, it learned in real time whether to keep pushing the wood block or steer it to a new target in case the tower collapsed. Kalashnikov et al. [18] addressed the problem of learning vision-based dynamic manipulation skills using a scalable reinforcement learning approach. This study introduced QT-Opten and realized closed-loop vision-based control, whereby the robot continuously updates its grasp strategy based on the most recent observations to optimize long-horizon grasp success. Chen et al. [19] proposed a simple model-free framework that can learn to redirect objects when the robot hand is upward and downward and has strong zero sample migration performance. Pertsch et al. [20] proposed a deep latent variable model, which combines the embedding space of learning skills with the a priori skill from the experience of off-line agents. This method can efficiently migrate skills from rich datasets. Zeng et al. [21] proposed a transporter network, which can represent a complex multimodal strategy distribution and be extended to multistep sequential tasks and 6DOF picking and placement. Li et al. [22] proposed a skill learning method based on multimodal information description. Using these multimodal information parameters, the robot can acquire the skills of searching, position determination and attitude adjustment. Furthermore, the robot can achieve the assembly goal by analyzing the two-dimensional image without position constraints.

Although current research uses deep learning to enable the robot to grasp the target object accurately, most algorithms have difficulty completing the accurate recognition and positioning in various cases for the high-precision grasping of the robot, especially the application of robot grasping and welding. Therefore, for the problem of the difference in projections produced by vision, we focus on the recognition and localization of multitype workpieces by a monocular vision system and industrial robot to realize the recognition and localization of multipose workpieces and guide the robot to complete the grasping of the workpiece. The rest of the paper is structured as follows. The third section elaborates on our proposed method. The fourth section presents experimental results and comparative evaluation to verify the proposed method. Finally, the fifth section summarizes the results of the current work and proposes future research directions.



**FIGURE 2.** The workflow diagram of target location.



**FIGURE 3.** Binary preprocessing of the image to be tested.



**FIGURE 4.** Matching results of feature points.

**III. PROPOSED METHOD**

In this section, we describe the proposed target location algorithm in detail, as shown in Fig. 2. First, we segment the image of the workpiece to be measured by using binary processing. Then, we adopt the feature matching algorithm to quickly obtain the initial rotation angle and displacement. Finally, we use weighted modular contour matching for accurate positioning by using the initial rotation angle and displacement to avoid the influence of the projection error of the workpiece.

#### A. INITIAL TARGET MATCHING

Since multiple artifacts may appear in the image to be tested, binary preprocessing is used to segment the artifacts, as shown in Fig. 3. Then a feature point matching algorithm is used to match each region. First, the OBR feature points [23] of the sample image and the image to be registered are extracted; then the feature point matching algorithm [24], [25] is used to quickly match and screen the area to be measured. Finally, the rotation angle and displacement are calculated.

ORB features combine FAST feature point detection with BRIEF feature descriptors in an improved and optimized manner. ORB uses the FAST feature point detection method to detect the feature points and then selects N feature points with the largest Harris corner response value from the FAST feature points. The Harris corner response function is defined as:

$$
R = detM - a (traceM)^2 \tag{1}
$$

where *M* is the covariance matrix of the gradient,  $det = \lambda_1 *$  $\lambda_2$  is the determinant of the matrix, *trace* =  $\lambda_1 + \lambda_2$  is the trace of the matrix,  $\lambda_1$  and  $\lambda_2$  is the eigenvalue of matrix M, and a is an empirical constant in the range (0.04, 0.06).

At the same time, the ORB takes the connection between the feature point and the particle as the feature point direction and improves the BRIEF feature descriptors by using the connection between the feature point and the

particle obtained previously to rotate the image and obtain the descriptor.

After extracting orb feature points, we use the FUGC algorithm [25] for feature point matching. First, we use brute force to construct a set of hypothetical matches for orb feature points. Considering the local feature consistency, we use a unilateral grid to segment image features, introduce local clustering constraints to remove the mismatches contained in the set, and use local linear transformation to rescreen the feature point pairs. The screening condition is matching distance error and is defined as follows:

$$
||q_i - H(p_i)|| < \tau \tag{2}
$$

where  $q_i$  and  $p_i$  are the coordinates of the matching feature points corresponding to the two images, and  $Hp_i$  are the corresponding coordinates after the transformation of  $p_i$ ,  $\tau$ is the distance error threshold. This strategy can extract high-accuracy matches from coarse baseline matches of low accuracy, screen out outliers, and quickly obtain accurate matching point pairs, as shown in Fig. 4.

Finally, the rotation angle θ*init* and displacement *dinit* between the sample image and the image to be measured are calculated by using these high-accuracy matching points.

# B. ACCURATE MATCHING OF CONTOUR MODULAR WEIGHTED DECISIONS

Due to the projection difference of the monocular camera, the rotation angle and displacement of grasping points may not be equal to the global rotation angle and displacement of the feature points mentioned above. It can also be simply considered that the projection difference leads to local deformation of the image, which interferes with the accuracy of the overall matching. Therefore, the influence of projection difference can be reduced by capturing the local contour matching around the points.

First, the gradient value and gradient direction of the sample workpiece image are calculated, and the region with a significant gradient value is extracted as the target contour. Projection differences are easy to produce as the grasping points are not at the same height. As shown in Fig. 1, the



**FIGURE 5.** The contour division of the sample image. Green and dark blue are contours of the same height, and light blue is the contour of another height.

contours of different heights need to be segmented according to the changes in the contour and grasping height of the sample image. Manual participation in the division can be required. Additionally, contours of the same height can be divided according to different shapes or positions, as shown in Fig. 5.

It is necessary to determine the matching weight  $\epsilon$  of each region after dividing the region of the workpiece contour. The closer the region is to the grasping point, the smaller the influence of projection error on the matching angle and displacement is. The weight  $\epsilon_j$  of the *j*<sup>th</sup> region contour is calculated as follows:

$$
w_j = \frac{\left(\sum_{i=1}^m d_i - d_j\right) \times \epsilon_j}{\sum_{j=1}^m \left(\left(\sum_{i=1}^m d_i - d_j\right) \times \epsilon_j\right)}\tag{3}
$$

$$
d_i = \frac{1}{n_i} \sum_{k=1}^{n_i} \sqrt{(x_k - x_g)^2 + (y_k - y_g)^2}
$$
 (4)

where  $(x_g, y_g)$  is the grasping point coordinate,  $(x_k, y_k)$  is the contour point coordinate,  $d_i$  is the average distance from the grasping point to the contour of the  $i^{th}$  region,  $n_i$  is the number of contour points in the  $i<sup>th</sup>$  region,  $m$  is the number of the divided regions and  $\epsilon$  is the height parameter. When the contour and grasping point are at the same height,  $\epsilon = 1$ . Otherwise, the value is far less than 1.

Before contour matching between the workpiece sample image and the image to be tested, the previously calculated rotation angle and displacement are used to reduce the search range of the sample image, which can improve the contour matching speed and reduce mismatching. Then, the similarity measurement method of the contour gradient is used to calculate the matching similarity *S* between the sample contour and the contour to be measured after translation and rotation.

$$
S = \sum_{i=1}^{m} \left( w_i \left( T_{ix} \times B_{ix} + T_{iy} \times B_{iy} \right) \times \sqrt{(T_{ix})^2 + (T_{iy})^2} \sqrt{(B_{ix})^2 + (B_{iy})^2} \right)
$$
(5)

where  $T_{ix}$  and  $T_{iy}$  are the X- and Y-direction gradients, respectively, of the  $i^{th}$  contour of the sample, and  $B_{ix}$  and  $B_{iy}$ 



**FIGURE 6.** Final matching result. The blue line is the contour after matching, and the red cross is the grasping point.

are the X- and Y-direction gradients, respectively, of the *i th* contour of the contours to be tested.

According to the rotation angle θ*init* and displacement  $d_{init}$ , the sample contour will move in the rotation angle range ( $\theta_{init} - r$ ,  $\theta_{init} + r$ ) and the translation range ( $d_{init}$  –  $l$ ,  $d_{init} + l$ ). The similarity of the contour to be measured  $S_{\theta,d}$  is calculated. By comparing all the  $S_{\theta,d}$ , the maximum similarity *Smax* is obtained; *Smax* must be larger than the given threshold  $\sigma$ .  $\sigma$  set according to the number and weight of contour points. Only when the similarity meets the conditions can it be considered the final matching result.

$$
S_{max} = \max \{ S_{\theta, d} \}
$$
  
\n
$$
(\theta_{init} - r) < \theta < (\theta_{init} + r), (d_{init} - l)
$$
  
\n
$$
< d < (d_{init} + l) \tag{6}
$$

where *r* and *l* are the rotation angle and translation range radius, respectively.

$$
\begin{cases}\nS_{max} > \sigma \quad true \\
S_{max} \leq \sigma \quad false\n\end{cases} \tag{7}
$$

When *Smax* meets the threshold, we obtain the rotation angle  $\theta_s$  and the displacement  $d_s$  as the rotationd angle and displacement of the final grasping point. The result is shown in Fig. 6.

#### **IV. EXPERIMENTAL RESULTS AND ANALYSIS**

To verify the proposed robot target positioning method, we evaluate the performance of the algorithm from two aspects: accuracy and time consumption. We compare our algorithm to contour matching based on the Hu moment and RANSAC feature points matching. In the whole experimental process, the parameters of the algorithm are consistent. These experiments were carried out on a notebook with a 3.5-GHz Intel i7 CPU and  $16$ -gb memory using  $c++$  and the open source toolkit OpenCV4.



**FIGURE 7.** Grasping error sequence of four kinds of workpieces. The first column is the workpiece type, the second column is the grasping distance error, and the third column is the grasping angle error.

#### A. DATASETS AND SETTINGS

To comprehensively evaluate our method, we selected four different workpieces with different heights, each of which has a total of 1000 pieces for a continuous test dataset. The image size is  $2448 \times 2050$ . We use the positioning accuracy as the corresponding test index: the average



**FIGURE 8.** Results of the comparison of the three kinds of algorithms.

distance and angle of the offset error of the target grasping point.

# B. EXPERIMENTAL RESULTS

To compare the performance of each algorithm, we specify that the placement angle of the workpiece is  $-60^{\circ} \sim 60^{\circ}$ , and the grasping point is set on the plane. Some parameters in the paper, such as  $\tau$ ,  $\epsilon$ ,  $r$ ,  $l$  and  $\sigma$ , need to be set according to the image size. According to the experimental test of the data in this paper, we set  $\tau = 3$ ,  $\epsilon = 1$ ,  $r = 5$  and  $l = 0.2$  of the radius of the maximum circumscribed circle of the outer contour.

We quantitatively compare our algorithm to a contour matching algorithm based on the Hu moment algorithm [26] and the feature matching algorithm [25] on the dataset. By calculating the displacement distance and rotation angle of the grasping point of the manipulator, the grasping error of each algorithm is obtained. The grasping error of the same algorithm in the dataset is sorted from small to large and sampled at intervals. The results are shown in Fig. 7. The error of the three algorithms in the first row of Fig. 7 is not large because the workpiece is a plane workpiece and there is no height projection error. However, because our algorithm has a twice-matching effect, it has a smaller error. At the same time, the large error data in the back basically do not exist. However, the second, third and fourth rows of Fig. 7 have different height planes, which will produce a certain height projection error. Therefore, affected by the height projection error, there are many large errors in the contour matching algorithm which uses the Hu moment and the matching algorithm which uses feature points. The matching performance of our algorithm is better and more stable than the first two algorithms.

If we set the threshold of the distance error value to 4 pixels and the threshold of the angle error value to 0.5 degrees, when the distance error is less than or equal to 4 pixels and the angle error is less than or equal to 0.5 degrees, the grasping is successful; otherwise, the grasping is considered to fail. We list the success rate of each algorithm in Fig. 8. Our algorithm has a higher number of success rate than the other two algorithms. Therefore, it is better proposed for our grasp planning approach. The accuracy of both offset distance and angle is relatively high for monocular cameras, which effectively solves the automatic grasping of manipulators. Furthermore, the stability and positioning results of the system can meet the needs of actual production.

# **V. CONCLUSION**

In this paper, we presented a robust, flexible, low-cost robot grasping system that allows the detection and precise location of workpiece targets using only a single camera. In order to solve the problem of positioning error caused by projection difference of monocular camera, We developed a method for workpiece image correction and matching by dividing the full contour of the workpiece and using the contour module weighting method, which improves the robustness, stability and accuracy of robot grasping. The experimental results show that the success rate of the proposed scheme is 92.1%, 10.9% higher than the method based on Hu moment matching, and 12.6% higher than RANSAC feature matching algorithms. The method proposed in this study can be extend to otherfields, such as defect detection, visual navigation and so forth. For future research, we will study how to use the deformation feature to correct when the grasping point is on the inclined plane and apply it to 3D image matching.

# **REFERENCES**

- [1] S. Hjorth and D. Chrysostomou, ''Human–robot collaboration in industrial environments: A literature review on non-destructive disassembly,'' *Robot. Comput.-Integr. Manuf.*, vol. 73, Feb. 2022, Art. no. 102208.
- [2] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schäffer, K. Kosuge, and O. Khatib, ''Progress and prospects of the human–robot collaboration,'' *Auto. Robots*, vol. 42, no. 5, pp. 957–975, Jun. 2018.
- [3] S. Liu, S. Wang, X. Liu, C.-T. Lin, and Z. Lv, ''Fuzzy detection aided realtime and robust visual tracking under complex environments,'' *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 1, pp. 90–102, Jan. 2021.
- [4] S. Wang, X. Liu, S. Liu, K. Muhammad, A. A. Heidari, J. D. Ser, and V. H. C. de Albuquerque, ''Human short long-term cognitive memory mechanism for visual monitoring in IoT-assisted smart cities,'' *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7128–7139, May 2022.
- [5] Z. H. Zheng, Y. Ma, H. Zheng, and, ''Industrial part localization and grasping using a robotic arm guided by 2D monocular vision,'' *Industrial Robot*, vol. 45, no. 11, p. 6, 2018.
- [6] S. Liu, S. Wang, X. Liu, J. Dai, K. Muhammad, A. H. Gandomi, W. Ding, M. Hijji, and V. H. C. de Albuquerque, ''Human inertial thinking strategy: A novel fuzzy reasoning mechanism for IoT-assisted visual monitoring," *IEEE Internet Things J.*, early access, Jan. 11, 2022, doi: [10.1109/JIOT.2022.3142115.](http://dx.doi.org/10.1109/JIOT.2022.3142115)
- [7] S. Liu, ''Human memory update strategy: A multi-layer template update mechanism for remote visual monitoring,'' *IEEE Trans. Multimedia*, vol. 23, pp. 2188–2198, 2021.
- [8] Z. Zheng, H. Zheng, J. Ju, D. Chen, X. Li, Z. Guo, C. You, and M. Lin, ''A system for identifying an anti-counterfeiting pattern based on the statistical difference in key image regions,'' *Expert Syst. Appl.*, vol. 183, Nov. 2021, Art. no. 115410.
- [9] Z. Zheng, B. Xu, J. Ju, Z. Guo, C. You, Q. Lei, and Q. Zhang, ''Circumferential local ternary pattern: New and efficient feature descriptors for anticounterfeiting pattern identification,'' *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 970–981, 2022.
- [10] M. Schmitz, *Machine Learning in Industrial Applications: Insights Gained From Selected Studies*. Erlangen, Germany: Friedrich-Alexander-Universit, 2022.
- [11] T. Wang, J. Su, C. Xu, and Y. Zhang, "An intelligent method for detecting surface defects in aluminium profiles based on the improved YOLOv5 algorithm,'' *Electronics*, vol. 11, no. 15, p. 2304, Jul. 2022.
- [12] M. P. Muresan, D. G. Cireap, and I. Giosan, "Automatic vision inspection solution for the manufacturing process of automotive components through plastic injection molding,'' in *Proc. IEEE 16th Int. Conf. Intell. Comput. Commun. Process. (ICCP)*, Sep. 2020, pp. 423–430.
- [13] T. Dekel, S. Qron, and M. Rubinstein, "Best buddies similarity for robust template matching,'' in *IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 2021–2029.
- [14] L. J. Xu, Z. L. Kang, P. Huang, and H. Luo, "Research on intelligent watermelon identification and positioning method in natural scene,'' *J. Intell. Fuzzy Syst.*, vol. 31, no. 4, pp. 2109–2117, Sep. 2016.
- [15] C. Yang, C. Zeng, Y. Cong, N. Wang, and M. Wang, "A learning framework of adaptive manipulative skills from human to robot,'' *IEEE Trans. Ind. Informat.*, vol. 15, no. 2, pp. 1153–1161, Feb. 2019.
- [16] Y. Gu, W. Sheng, C. Crick, and Y. Ou, "Automated assembly skill acquisition and implementation through human demonstration,'' *Robot. Auton. Syst.*, vol. 99, pp. 1–16, Jan. 2018.
- [17] N. Fazeli, M. Oller, J. Wu, Z. Wu, J. B. Tenenbaum, and A. Rodriguez, ''See, feel, act: Hierarchical learning for complex manipulation skills with multisensory fusion,'' *Sci. Robot.*, vol. 4, no. 26, Jan. 2019, Art. no. eaav3123.
- [18] D. Kalashnikov, A. Irpan, P. Pastor, J. Ibarz, A. Herzog, E. Jang, D. Quillen, E. Holly, M. Kalakrishnan, V. Vanhoucke, and S. Levine, ''QT-opt: Scalable deep reinforcement learning for vision-based robotic manipulation,'' 2018, *arXiv:1806.10293*.
- [19] T. Chen, J. Xu, and P. Agrawal, "A system for general in-hand object reorientation,'' 2021, *arXiv:2111.03043*.
- [20] K. Pertsch, Y. Lee, and J. J. Lim, "Accelerating reinforcement learning with learned skill priors,'' 2020, *arXiv:2010.11944*.
- [21] A. Zeng, P. Florence, J. Tompson, S. Welker, J. Chien, M. Attarian, T. Armstrong, I. Krasin, D. Duong, A. Wahid, V. Sindhwani, and J. Lee, ''Transporter networks: Rearranging the visual world for robotic manipulation,'' 2020, *arXiv:2010.14406*.
- [22] F. Li, Q. Jiang, W. Quan, R. Song and Y. Li, ''Manipulation skill acquisition for robotic assembly using deep reinforcement learning,'' in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics (AIM)*, 2019, pp. 13–18, doi: [10.1109/AIM.2019.8868579.](http://dx.doi.org/10.1109/AIM.2019.8868579)
- [23] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF,'' in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 6–13.
- [24] J. Ma, J. Zhao, J. Jiang, H. Zhou, and X. Guo, ''Locality preserving matching,'' *Int. J. Comput. Vis.*, vol. 127, no. 5, pp. 512–531, 2019.
- [25] Z. Zheng, H. Zheng, Y. Ma, F. Fan, J. Ju, B. Xu, M. Lin, and S. Cheng, ''Feedback unilateral grid-based clustering feature matching for remote sensing image registration,'' *Remote Sens.*, vol. 11, no. 12, p. 1418, Jun. 2019.
- [26] W. U. Wenhan, Y. A. N. G. Ming, W. A. N. G. Bing, and W. A. N. G. Chunxiang, ''Pallet detection based on contour matching for warehouse robots,'' *J. Shanghai Jiaotong Univ.*, vol. 53, pp. 197–202, Feb. 2019.

ZHAOHUI ZHENG received the Ph.D. degree from Wuhan University. He is currently a Lecturer with the School of Mathematics and Physics, Wuhan Institute of Technology. His research interests include machine vision, image processing, biometrics, intelligent systems, and robotic dynamics.

YUANYUAN LI received the Ph.D. degree from Wuhan University. She is currently an Associate Professor with the School of Mathematics and Physics, Wuhan Institute of Technology. Her research interest includes machine vision.

**DENG CHEN** received the Ph.D. degree from Wuhan University. He is currently an Associate Professor with the Hubei Province Key Laboratory of Intelligent Robot, Wuhan Institute of Technology. His research interests include artificial intelligence, and the Internet of Things.

LEI WANG received the Ph.D. degree from Wuhan University. He is currently a Lecturer with the School of Electrical and Information Engineering, Wuhan Institute of Technology. His research interests include consist of signal processing, and mechanical fault diagnosis.

JIANPING JU received the B.S. degree in optical information science and technology from the Huazhong University of Science and Technology, in 2003, and the M.S. degree in artificial intelligence and pattern recognition from the Wuhan Institute of Technology, Wuhan, China, in 2010. His research interests include machine learning, and pattern recognition.

He is currently an Associate Professor with the School of Artificial Intelligence, Hubei Business College.

QIANG ZHANG received the master's degree in signal and information processing from Nanchang Hangkong University. She is currently a Lecturer with the Department of Public Courses, Wuhan Railway Vocational College of Technology. Her research interests include image processing, and pattern recognition.

 $\sim$   $\sim$   $\sim$