Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 9279345, 12 pages
https://doi.org/10.1155/2020/9279345

Research Article

6DOF Object Positioning and Grasping Approach for Industrial Robots Based on Boundary Point Cloud Features

Guoyang Wan,[1] **Guofeng Wang**, **Kaisheng Xing**,[2] **Tinghao Yi**,[3] and **Yunsheng Fan**[1]

[1]*Dalian Maritime University, Department of Marine Electrical Engineering, Linhai Road No. 1, Dalian 11600, China*
[2]*Anhui Institute of Information Technology, Yonghe Road No. 1, Xinwu Economic Development Zone, Wuhu, Anhui 241000, China*
[3]*University of Science and Technology of China, JinZhai Road Baohe District No. 96, Hefei, Anhui 230026, China*

Correspondence should be addressed to Guoyang Wan; 704610266@qq.com

Received 12 August 2020; Revised 3 November 2020; Accepted 27 November 2020; Published 14 December 2020

Academic Editor: Alessandro Gasparetto

For the three-dimensional (3D) pose estimation of metal blank casts estimate in industrial production process, a novel 6-degree-of-freedom (6DOF) positioning and grasping approach for industrial robots based on boundary features and combining structured light 3D vision and the point cloud matching method is proposed. The proposed approach first uses the Gray code plus grating phase-shift algorithm to reconstruct the 3D surface information of the object. Subsequently, an improved point pair feature (PPF) matching location method based on point cloud boundary extraction is proposed to realize the accurate 6DOF pose estimation of the object. In this method, the point cloud boundary feature is used to optimize the PPF algorithm. Finally, by combining with the point cloud preprocessing process and the improved PPF matching location method, the 6DOF pose positioning of metal blank casting by industrial robots is realized. The proposed approach can accurately complete pose measurement and positioning of objects placed randomly in an open environment. The experimental results demonstrate that the proposed PPF matching location method significantly improves both matching speed and accuracy compared to the traditional PPF algorithm.

1. Introduction

In recent years, industrial robot vision positioning in industrial production processes has been extensively researched. The industrial robot positioning and grasping workstation combined with machine vision technology not only saves labor cost but also ensures better safety and reliability compared with traditional manual operation. In the production and processing of metal castings, many problems, such as high work intensity and poor production environment, exist. The advantage of applying industrial robot vision positioning technology to the casting industry is obvious.

The 6-degree-of-freedom (6DOF) pose measurement algorithm is the basis of a robotic bin picking task and has been studied by many researchers. Hinterstoisser et al. [1] proposed a line-MOD template matching algorithm that can be used for the 6DOF pose estimation of nontextured objects. On the basis of this work, Ye et al. incorporated the multilevel matching positioning strategy of an image pyramid and realized the real-time detection of 6DOF pose of objects [2]. Le and Len [3] proposed a deep learning-based method to realize a robot's disordered grasping. By establishing the target object database, the model was trained by using a Mask-RCNN detector; thus, the function of robotic bin picking was realized. A special three-finger gripper was used in [4], and its implementation in disordered grasping was studied.

Compared with the common problems in the 6DOF object pose estimation, such as occlusion, cluttered background, and object texture interference [5], in industrial production, the pose estimation of rough metal castings faces more complex problems. First, unprocessed rough metal castings have burrs on the surface, which can

introduce errors in the pose estimation. Compared with processed products, the surface profile of unprocessed castings is smoother, which is not conducive to visual recognition. Second, the production environment of metal castings is harsh, which is not conducive to the maintenance of vision system hardware equipment. Finally, metal castings with a large number of repetitive or similar features (large planes) present more complicated problems in the pose estimation [6]. Such features easily conceal the local features of the object surface used for direction recognition, causing the failure of the direction recognition of the positioned object.

A three-dimensional (3D) matching method based on the point pair feature (PPF) is a widely concerned 3D pose measurement method [7]. This method shows successful object recognition performance on different 3D datasets [8]. In terms of recognition performance and recognition accuracy, PPF-based algorithms are comparable to learning class 3D object detection algorithms that gained popularity recently [9–12]. However, when the object itself has many repetitive features (such as large planes), the matching performance of the traditional PPF method will be reduced [6]. The disordered sorting of industrial applications also faces this problem. To solve this problem, some scholars have improved the PPF algorithm. In [6, 13], the edge points of the point cloud are calculated using the depth map of the measured object, and matching calculation is performed to improve the 3D positioning accuracy. However, this method of calculation of the edge points of the point cloud is not accurate in complex environments. In [14], the edge points of the point cloud are calculated from multiple perspectives using CAD technology, and the matching calculation method of the point pairs is improved to realize high-precision 3D point cloud matching; however, this algorithm is complex and not applicable to industrial objects.

In view of the abovementioned problems, a combination of the binocular structured light vision technology and the point cloud registration algorithm with a 6DOF pose estimation approach for robots based on 3D matching is proposed. The proposed approach mainly includes three phases: 3D point cloud reconstruction based on binocular structured light, an improved PPF matching method based on boundary features of point cloud, and an industrial robot positioning strategy based on the PPF algorithm. This approach realizes the positioning and grasping of 6DOF pose of industrial robots for rough casting objects, and it achieves favorable processing speed and positioning accuracy. The main focus of the study is as follows: first, the existing binocular structured light technology is optimized, and fast 3D reconstruction of the measured object is realized. Second, an improved PPF registration algorithm based on principal component analysis (PCA) boundary extraction is proposed. Then, through the improved PPF method, point cloud preprocessing, and robotic coordinate translation, the 6DOF pose positioning of metal blank casting by industrial robot is realized. Finally, a comparative experiment is conducted to prove the robustness and accuracy of the proposed method.

The rest of the paper is organized as follows: Section 2 introduces the architecture and hardware composition of the proposed 6DOF measurement approach. Section 3 analyzes the key technologies of the proposed strategy. First, the point cloud generation technology based on binocular structured light is introduced; second, an improved 3D reconstruction method based on binocular structured light for metal blank castings is used. Then, the traditional PPF 3D matching algorithm is optimized by combining point cloud boundary extraction algorithm with PCA. Finally, the object positioning online measurement process and robot coordinate transformation are introduced. Section 4 introduces the relevant experimental equipment and presents the experimental analysis. Section 5, finally, summarizes the study.

2. Approach Architecture

The proposed grasping system mainly consists of three parts: the binocular structured light vision system, industrial robot grasping system, and control system. The binocular structured light vision system collects image information of the object and realizes 3D reconstruction. The robot can cooperate with the vision system to achieve an execution mechanism for image acquisition and positioning. Based on the 3D information reconstructed from the vision system, the central control system calculates the grasping position and pose of the robot and plans the grasping path.

2.1. Operation Process. The operation process of our approach is divided into three stages: 3D reconstruction, offline calibration, and online measurement. In the 3D reconstruction process, the 3D information of the object is generated. The offline calibration process completes the calibration of the binocular structured light vision system and creates the template required in the 3D matching process. In the online measurement process, the data of the offline calibration process are used to directly calculate the pose of the object and realize grasping. The operation logic of the proposed grasp system is shown in Figure 1.

2.2. Test Object. The object used in this study for testing the grasping system is cast aluminum blank with a size of 100 mm × 100 mm × 5 mm. The surface of the object consists of many burrs and stains as well as many planar features. In addition, there are many similar features on the object surface, which hinders orientation. Figure 2 shows the status of a single object and the actual multiple objects to be grasped randomly.

In the actual production process, the upper part of the object randomly placed must be detected and grasped by the robot accurately.

3. System Description

To realized high-precision 6DOF positioning of industrial objects, the main technique of the novel approach consists of binocular structured light 3D reconstruction, an improved

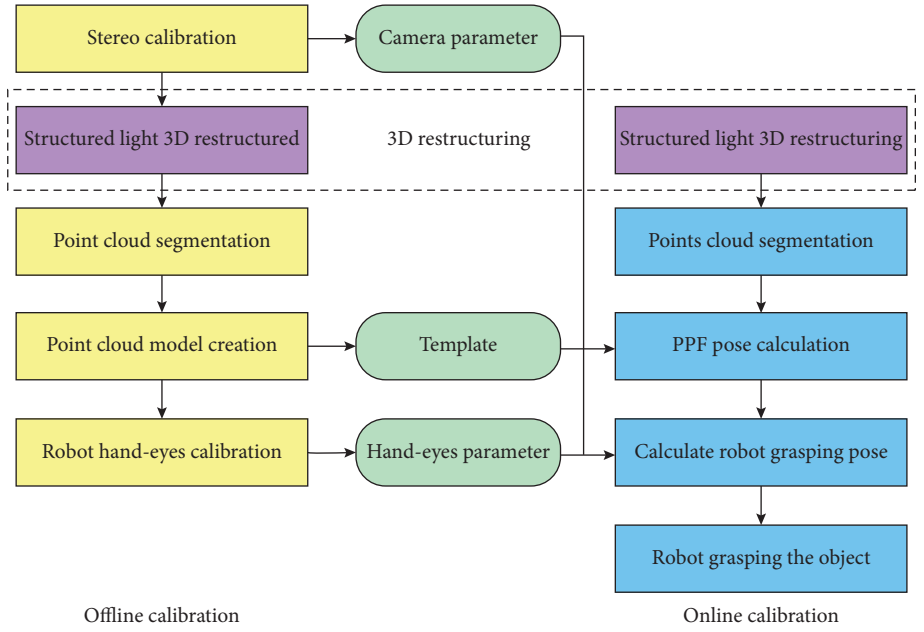
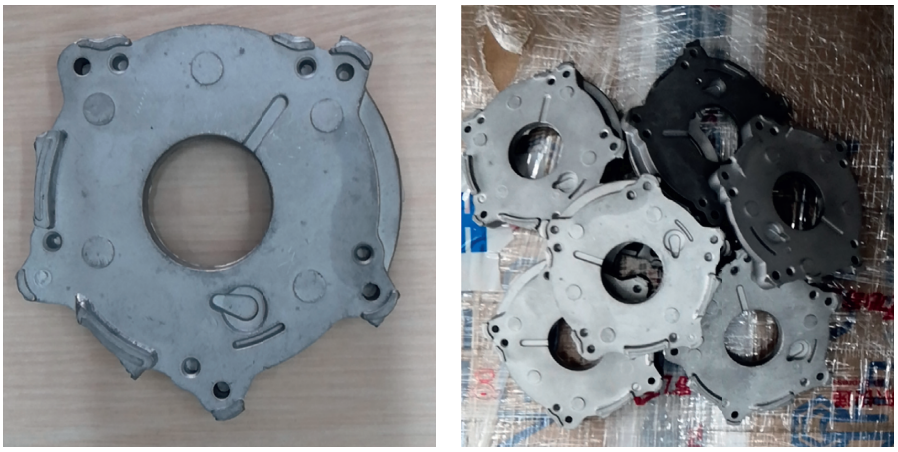


Figure 1: Robot grasping system operation logic diagram.



(a) (b)

Figure 2: (a) Single object; (b) several objects randomly placed on a pallet.

PPF point cloud registration method based on boundary features, and an industrial robot positioning strategy.

3.1. 3D Reconstruction Based on Binocular Vision and Structured Light

3.1.1. Generation of Structured Light by Gray Code and Phase-Shift Method. The 3D reconstruction based on binocular structured light avoids the problem that a binocular stereo vision method cannot obtain the detailed 3D information of an object and solves the problem of complex calibration of the monocular structured light vision technology. Compared with the linear structured light technology, this method has the advantages of fast imaging speed and low cost.

The projection method of the Gray code and phase-shift method is used for 3D reconstruction. With this method, Gray code coding does not require multiple image reconstruction, the high-frequency code image is not easily affected by noise, and the phase error of the phase-shift method is suppressed.

The color projection image is used to improve the efficiency of the system. According to the test, the reflection of blue-green light by the object is good, so the projection image is composed of blue-green light (see Figure 3). By introducing the color projection image, the projection effect of 10 monochromatic light images can be achieved with 5 images. In the 3D reconstruction process, RGB images must be decomposed into single channel images. To simultaneously improve the calculation efficiency, shadow mask calculation is introduced into the method [15].

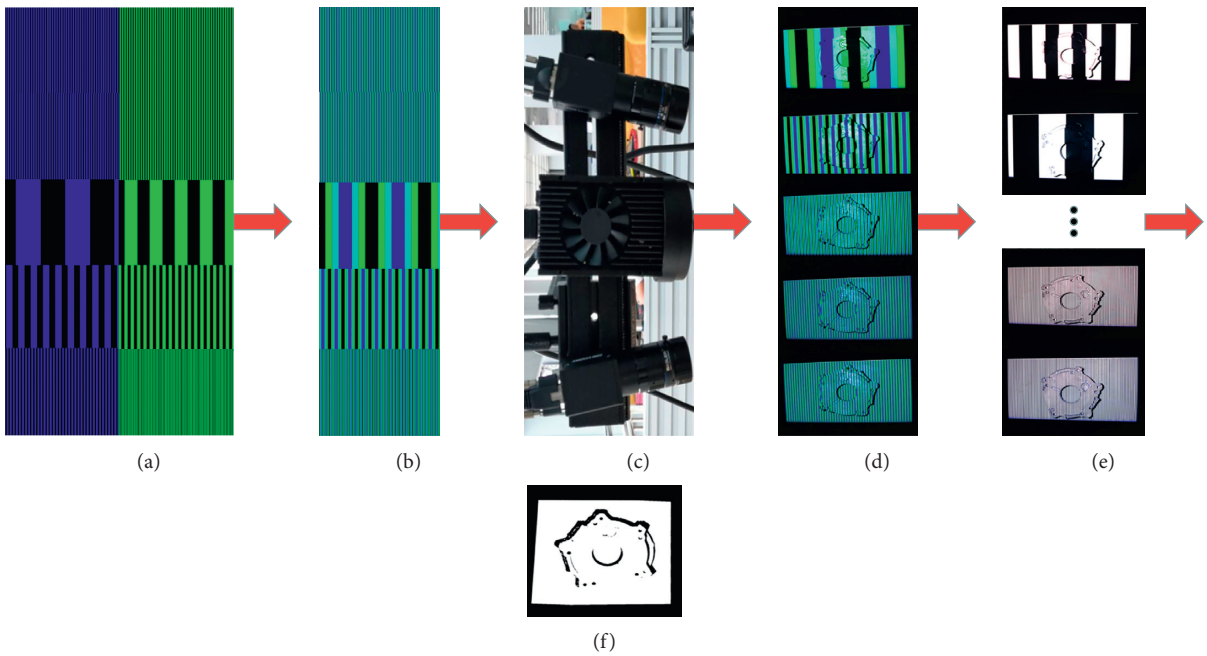


Figure 3: (a) Gray code and phase-shift coded image; (b) projected image after color fusion; (c) binocular camera and the projector; (d) projection image obtained by a binocular camera; (e) decomposed into picture; (f) reconstructed point cloud image after decoding.

The absolute phase diagram is calculated by the decoded code diagram and the dephased phase diagram as follows:

$$\psi(x, y) = 2 \times \pi \times M(x, y) + \Phi(x, y), \tag{1}$$

where $\psi(x, y)$ and $M(x, y)$ are the absolute phase value and Gray code decoding value of pixel position (x, y), respectively.

3.1.2. 3D Point Cloud Generation. Zhang's calibration method [16] is selected to complete the double target determination and obtains the intrinsic and extrinsic parameters of the left and right cameras and their relative positions. The Bouguer stereo correction method [17] is used to eliminate distortion and aligns the left and right views; therefore, the imaging origin coordinates of the left and right views are consistent, the optical axes of the two cameras are parallel, the left and right imaging planes are coplanar, and the epipolar lines are aligned.

Stereo matching aims to find matching points in the left and right absolute phase maps after stereo correction. The effective difference between the left and right phases is $(h - n, h + n)$:

$$T - \text{abs}[LG(x, y) - RG(x, y + i)] > 0, \quad i \in (h - n, h + n), \tag{2}$$

$$T - \text{abs}[LG(x, y + i) - RG(x, y)] > 0, \quad i \in (h - n, h + n), \tag{3}$$

where T is the contrast threshold and LG and RG are the phase values of (x, y) points in the left and right absolute phases. The corresponding point RP1 of LP1 in the left image can be obtained in the right image by satisfying Equation (2). From equation (3), the corresponding point $LP2$ of point $RP1$ in the left image can be obtained. If the distance between $LP1$ and $LP2$ is less than $D1$, then $(LP1 + LP2/2)$ in the left image and $RP1$ in the right image are a pair of matching points.

After obtaining the matching points in the left and right images, they are combined with the camera calibration results; then, using the triangulation method, the 3D reconstruction of the measured object surface can be completed. The reconstruction process is shown in Figure 4.

3.2. 3D Point Cloud Matching and Positioning. Point cloud matching is one of the most popular methods for 6DOF pose estimation of objects. Commonly used methods include the viewpoint feature histogram (VFH) and the PPF-based 3D surface matching method. However, if these algorithms are directly applied to the object in this study, a small number of hollowed out area features in the plane point cloud will be covered by the dense area in the point cloud during the matching calculation. This leads to the failure of the plane point cloud on the Z-direction rotation axis of the object to be matched (Figure 5). To solve this problem, this paper proposes a PPF matching location method based on edge extraction and realizes the accurate 6DOF positioning of the object.

3.2.1. Point Cloud Matching Algorithm Based on PPF. For points $m1$ and $m2$ in the point cloud, the corresponding normal vectors are $n1$ and $n2$ and $d = m2 - m1$, and then, a PPF exists as follows:

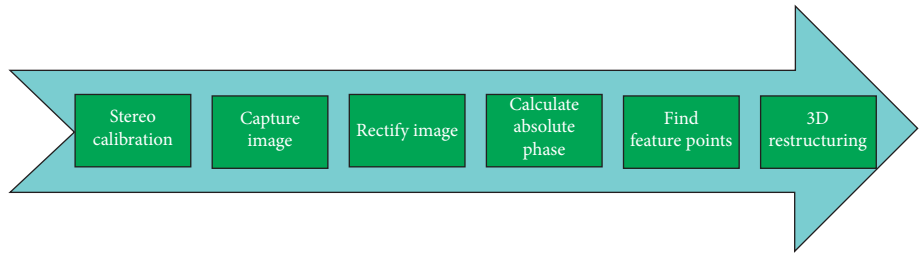
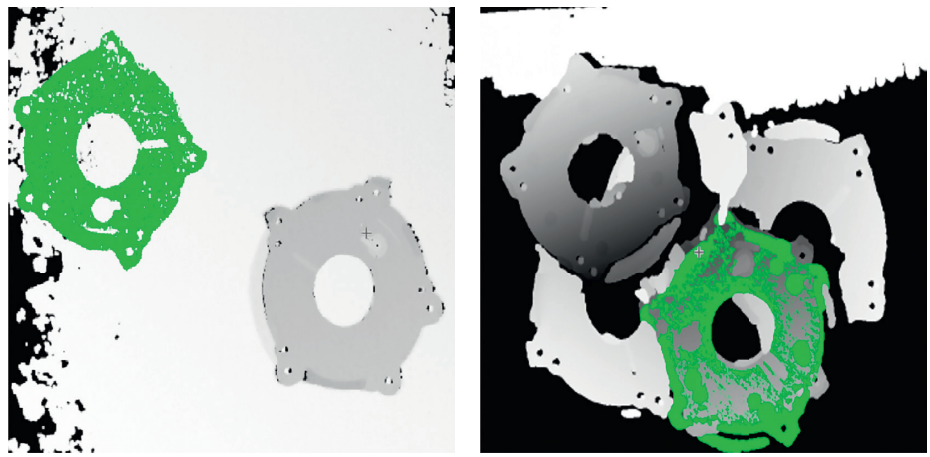


Figure 4: 3D reconstruction flow chart of structured light.



(a) (b)

Figure 5: Failure case of PPF matching algorithm: (a) similar features in the background lead to the failure of object pose estimation; (b) Z-direction rotation angle matching fails due to the plane PPF of the object itself.

$$F(m1, m2) = \left(\|d\|_2, \angle(n1, d), \angle(n2, d), \angle(n1, n2)\right), \quad (4)$$

where $\angle(a, b) \in [0, \pi]$ is the angle between two vectors and F is asymmetric.

The matching process is divided into offline and online processes. In the offline process, the point cloud of the object to be matched is set as a model, and the feature vector F is calculated for all the point pairs on the template surface. Then, a hash table is constructed to record the points with the same PPFs. In the online process, the same PPFs are extracted from the scene point cloud to be matched, and the similar point pairs are found in the template for pose registration. Finally, a voting mechanism is used to calculate the position and pose of the object that falls on most point pairs on the surface of the template.

3.2.2. PPF Matching Location Method Based on Edge Extraction. To solve the problem that the original PPF matching algorithm cannot estimate the present object pose accurately, a PPF matching location algorithm based on edge extraction is proposed. This algorithm is divided into offline modeling and online matching stages.

In offline modeling, the PPF model file required in the online matching stage is created. The template can be obtained from a CAD file or by directly sampling the object with the 3D vision system.

3.2.3. Point Cloud Edge Extraction. PCA is a common reduction algorithm. Bazazian et al. [18] proposed a non-clustering point cloud edge extraction method based on PCA, which extracts the boundary points of the point cloud according to the changes in the characteristic value of each point in the point cloud. This statistical method eliminates the sensitivity of normal estimation to sharpening edges, eliminates the clustering steps, and simplifies the process of edge estimation.

PCA-based edge detection method must calculate the covariance of each point in the point cloud. Covariance is a measure of how much each dimension changes relative to each other's average. For 3D datasets (x, y, z), the 3×3 covariance matrix C of sample point $\boldsymbol{P}\ (x, y, z)$ is given by the following formula:

$$C = \begin{bmatrix} Cov(x,x) & Cov(x,y) & Cov(x,z) \\ Cov(y,x) & Cov(y,y) & Cov(y,z) \\ Cov(z,x) & Cov(z,y) & Cov(z,z) \end{bmatrix}, \quad (5)$$

where cov (x, y) is the covariance of points x and y and is calculated as follows:

$$Cov(x, y) = \frac{\sum_{i=1}^{k} (x_i - \overline{x})(y_i - \overline{y})}{n - 1}. \quad (6)$$

The characteristic value of C is $\lambda_0 \le \lambda_1 \le \lambda_2$, with

$$\sigma_n(p) = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2}. \tag{7}$$

Here, $\sigma_n(p)$ quantitatively describes the change along the surface normal, that is, the degree of deviation of the estimated point from the tangent plane [19]. λ_0, λ_1, and λ_2 reflect the position relationship between a point in the point cloud and the surrounding points:

$$\sigma' = \frac{\lambda_1}{\lambda_1 + \lambda_2}. \tag{8}$$

When λ_1 is $\sigma' \geq k \times \lambda_{1MAX}$, the point is considered the edge point of the point cloud. k is determined by the shape of the point cloud; for the object shown in Figure 2, let $k = 0.8$.

3.2.4. Point Cloud Model Generation. Through the above calculation, the point cloud boundary of the object can be obtained. To improve the computational efficiency, the Euclidean clustering segmentation method is used to segment the point cloud boundary point set. Suppose the query point is $p_i \in P$ and P is the point cloud set. In addition, suppose there are n points in the neighborhood whose radius is r, that is, $\{p_{ik}, \quad k = 1, 2, 3, \ldots, n\}$. In this study, we set $r = 0.5$. Then, the Euclidean distance between the adjacent point and the query point is calculated as follows:

$$\begin{aligned} \mathrm{d}(p_i, p_{ik}) &= \sqrt{\sum_{k=1}^{n} (p_i - p_{ik})^2}, \\ Q &= \{p_{ik} | \mathrm{d}(p_i, p_{ik}) < r, \quad p_{ik} \in P\}, \end{aligned} \tag{9}$$

where Q is the processed point cloud and $d(.)$ is the distance between two points.

Through the Euclidean segmentation, the contour point cloud Ps of the object can be divided into several subpoint clouds Pn; that is, $Ps \in \{P1, P2, \ldots, Pn\}$. The point cloud is filtered according to the number of points in the subpoint cloud:

$$Pu = Pe1 + \ldots + Pen, \{\mathrm{size}(Pe1), \ldots, \mathrm{size}(Pen) > Kn\}. \tag{10}$$

Pu is the newly generated point cloud after filtering, $Kn = 300$ is the threshold of the number of point clouds, and size $(Pe1)$ represents the number of point clouds in $Pe1$. The model point cloud is created by Pu (see Figure 6).

3.2.5. Online Matching. In the online matching, the PPFs of the scene point cloud are calculated, and the matching object pose is calculated by a voting method (see Section 3.3).

3.3. Robotic Positioning and Grasping Strategy

3.3.1. Point Cloud Preprocessing. To reduce external interference and improve the computational speed and robustness of the PPF matching algorithm, it is necessary to preprocess the point cloud acquired by the vision system (see Figure 7).

Due to the existence of sensor error and measurement error, there is inevitably some noise in the point cloud. The statistical filtering method is used to filter outliers in the target point cloud as follows:

$$(x_n, y_n, z_n) = \begin{cases} \text{savepoint1} & |p(x, y, z) - \mu| < \sigma T \\ \text{outliner} & \text{else} \end{cases}, \tag{11}$$

$$\begin{aligned} \mu &= \frac{1}{k} \sum_{p(x_i, y_i, z_i) \in S}^{i=k} p(x_i, y_i, z_i), \\ \sigma^3 &= \frac{1}{k} \sum_{p(x_i, y_i, z_i) \in S}^{i=k} \left| p(x_i, y_i, z_i) - \mu \right|^3, \end{aligned} \tag{12}$$

where S is the neighborhood point set of point p in the point cloud, $\mu$and $\sigma$are the mean and variance of the distance between the point and other points in S, and $k = 30$ is the number of midpoints of S. From equation (11), the filtered target point cloud can be obtained [20].

Then, the random sample consensus (RANSAC) algorithm is used to segment the target point cloud to remove the plane background as follows:

$$\begin{aligned} d_n &= \frac{\left| ax_n + by_n + cz_n + d \right|}{\sqrt{a^2 + b^2 + c^2}}, \\ (x_n, y_n, z_n) &= \begin{cases} \text{plant point}, & d_n < \tau, \\ \text{outliner}, & d_n < \tau, \end{cases} \end{aligned} \tag{13}$$

where (a, b, c, d) are the plane parameters to be fitted, dn is the distance from any point cloud to the fitting plane, and $\tau = 6$ is the threshold value of outliers in the control point cloud.

Finally, the pass-through filter is used to further remove the invalid point cloud as follows:

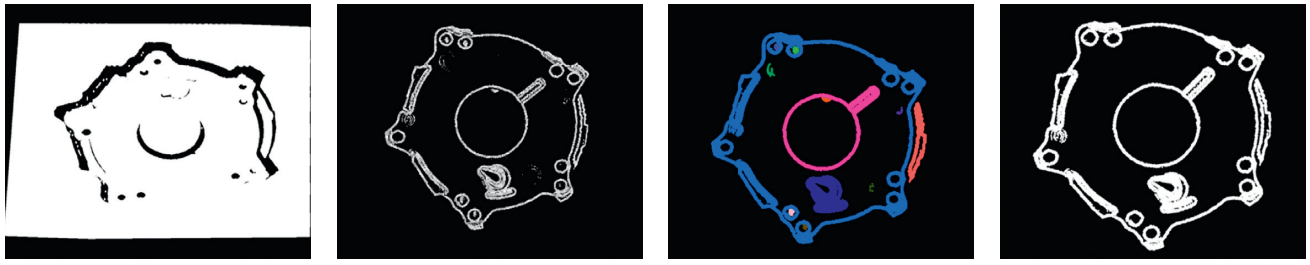
$$(x, y, z) = \begin{cases} \text{savepoint2}, & x_1 < x < x_2, y_1 < y < y_2, z_1 < z < z_2, \\ \text{outliner}, & \text{else}. \end{cases} \tag{14}$$

3.3.2. Point Cloud Matching. If a point pair (sr, sr) in the target point cloud and a point pair (mr, mi) in the model point cloud have the same pair feature F, then the reference point sr in the scene is considered to match the point mr in the model. The transformation from the model point to the scene point can be described by rotation angle α.

The transformation from model point mr to scene point sr is defined as follows:

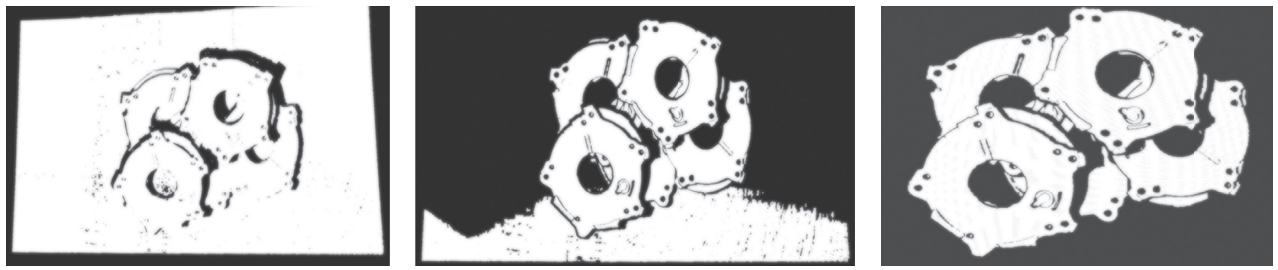
$$Sr = T_{s \longrightarrow g}^{-1} R_x(\alpha) T_{m \longrightarrow g} m_r, \tag{15}$$

where $T_{s \longrightarrow g}$ represents the translation and rotation transformation from the reference point to the coordinate



(a) (b) (c) (d)

Figure 6: Creation process of the point cloud model based on the boundary: (a) object point cloud obtained by a 3D camera; (b) the boundary of the object point cloud obtained by the PCA-based algorithm; (c) multiple subedges obtained by Euclidean clustering; (d) after removing the small edges, the object boundary model is also used for matching.



(a) (b) (c)

Figure 7: Preprocessing and matching process of the target point cloud: (a) target point cloud acquired by the 3D camera; (b) outliers and RANSAC plane segmentation results are removed by statistical filtering; (c) pass-through filtering results.

system in the scene, $T_{m\longrightarrow g}$ represents the translation and rotation transformation from the reference point to the coordinate system in the model, and $R_x(\alpha)$ represents the rotation angle α around the positive half axis of the X axis.

To speed up the calculation, $\alpha$is decomposed into $\alpha = \alpha_m - \alpha_s$. Considering $R_x(\alpha) = R_x(-\alpha_s)R_x(\alpha_m)$ and $R_x^{-1}(-\alpha_s) = R_x(\alpha_s)$ and combining with equation (15), the following is obtained:

$$t = R_x(\alpha_s)T_{s\longrightarrow g}s_i = R_x(\alpha_m)T_{m\longrightarrow g}s_i. \tag{16}$$

In [14], it is pointed out that, in the process of PPF matching, the point pairs whose distance is greater than the maximum size of the model cannot belong to the same target. Therefore, the weight screening criteria are often added to the voting mechanism:

$$\text{weight}(d) = \begin{cases} 1, & 0 \le d \le d(N),\ \mathrm{d}z < \omega, \\ 0, & \text{others}, \end{cases} \tag{17}$$

where $\mathrm{d}(N) = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2}$, where d$x$, d$y$, and d$z$ are the maximum length of the object in the X, Y, and Z directions, respectively, and ω is the threshold value to ensure that the object does not exceed the fixed tilt angle. Voting weight selection criteria can improve the matching speed and accuracy of the PPF matching algorithm.

To improve the accuracy of the Z-direction angle recognition of the target object in the matching process, weight calculation is introduced in the calculation of the matching similarity score of the contour point cloud. For the proposed object, the contour point cloud is divided into boundary point cloud Pb and nonboundary point cloud Pnb:

$$S_{\text{Main}} = w_1 \times S_{Pb} + w_2 \times S_{Pnb}, \tag{18}$$

where S_{Main} is the similarity score of single object matching, S_{Pb} and S_{Pnb} are the matching scores of the boundary point cloud and the nonboundary point cloud, and w_1 and w_2 are their corresponding weights.

3.3.3. Robot Grasping. The position of the object relative to the end of the robot flange can be obtained by hand-eye calibration. The point cloud generated by the 3D camera is relative to the left camera coordinate system.

According to equation (16), the grasping mode of the robot for the object with different positions and postures can be calculated. The specific derivation process is shown as follows [21]:

$${}^{U}T_{T}^{1} \times {}^{T}T_{P} = {}^{U}T_{CT}^{2} \times {}^{CT}T_{C} \times {}^{C}T_{P}, \tag{19}$$

where ${}^{U}T_{T}^{1}$ is the posture of the six-axis flange in the robot coordinate system when the robot grasps the object, that is, the robot grasp posture; ${}^{T}T_{P}$ is the position of the object in the robot grasp coordinate system; ${}^{U}T_{CT}^{2}$ is the posture of the robot when the vision system takes an image; ${}^{CT}T_{C}$ is the posture of the vision system relative to the end of the robot flange, which can be obtained by hand-eye calibration; and ${}^{C}T_{P}$ is the posture of the object relative to the vision system, which is directly acquired by the vision system.

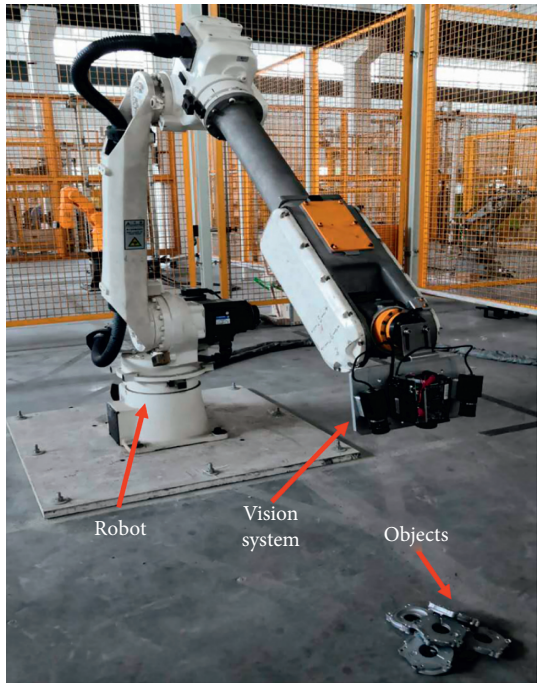


Figure 8: 6DOF pose estimation system.

In this part, an industrial robot positioning and grasping strategy which includes the PPF matching algorithm based on point cloud boundary extraction method and robot coordinate translation method is proposed. The proposed method considers the point cloud boundary as the matching object and introduces weight optimization in the similarity calculation process. Compared with the traditional PPF matching method, our method has a higher matching efficiency and accuracy.

4. Experiment

4.1. Experimental Platform. A six-axis industrial robot with 20 kg load was selected for image acquisition and object grasping. A high-speed projector with a resolution of 1028×720 was selected for structured light projection, two cameras with a resolution of 1280×1024 were used to complete image acquisition, and an industrial computer was used for 3D reconstruction and 3D matching. Figure 8 is the photograph of the pose estimation system.

The hand-eyes calibration parameter is shown in Table 1.

Table 1: Hand-eyes calibration parameter.

x (mm)	y (mm)	z (mm)	a (°)	b (°)	c (°)
−178.086	−167.2359	402.44602	359.471	0.543	0.317

4.2. Analysis of 3D Reconstruction Data of Binocular Structured Light. 3D reconstruction is the process of calculating the point cloud image through phase unwrapping and triangulation (see Figure 9).

The point cloud information of the object surface is obtained by 3D reconstruction. The color projection can achieve the effect of 10 projection images of the monochromatic projection model with 5 images; thus, the efficiency of 3D reconstruction is improved.

Due to the hardware limitation of the structure light projector and the binocular color camera, only the blue-green light spectrum can be well separated. If better equipment is used, less 3D image reconstruction can be achieved.

4.3. Comparative Analysis of Point Cloud Matching Based on PPF. The improved PPF matching algorithm realizes the disordered grasping of the object, which is a rough metal casting, by the industrial robot through the measurement of the position and orientation of the target object. A method of target recognition using CAD technology was proposed by Ulrich et al. [22] and optimized by Zhu et al. [23]. In this paper, Ulrich's method, the original PPF matching algorithm, and the proposed point cloud boundary PPF matching method are compared.

4.3.1. Comparison of Recognition Robustness. The three methods were applied to the task of identification of a single object and multiple objects.

The comparison the recognition effects of the three methods, shown in Figure 10, indicates that, with a good light source, Ulrich's method has better recognition effect for a single object; however, it has higher requirements for the light source and the recognition effect of disordered objects is poor. Hence, this method is not suitable for positioning randomly placed objects. The traditional PPF algorithm has a good positioning effect, except for the rotation axis in the Z-direction, but its accuracy is still insufficient. In this study,

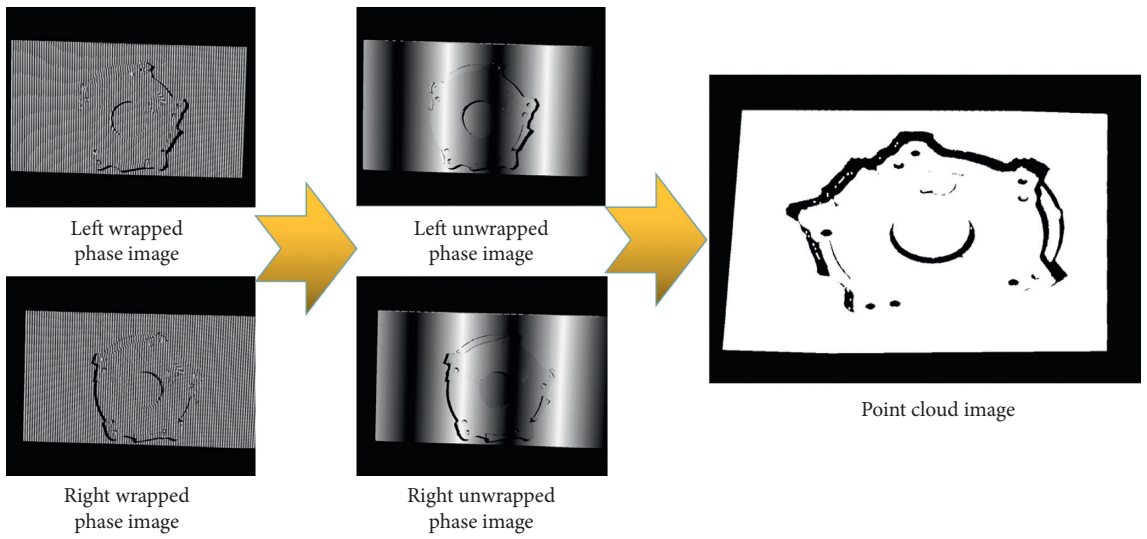


Figure 9: 3D reconstruction of binocular structured light.

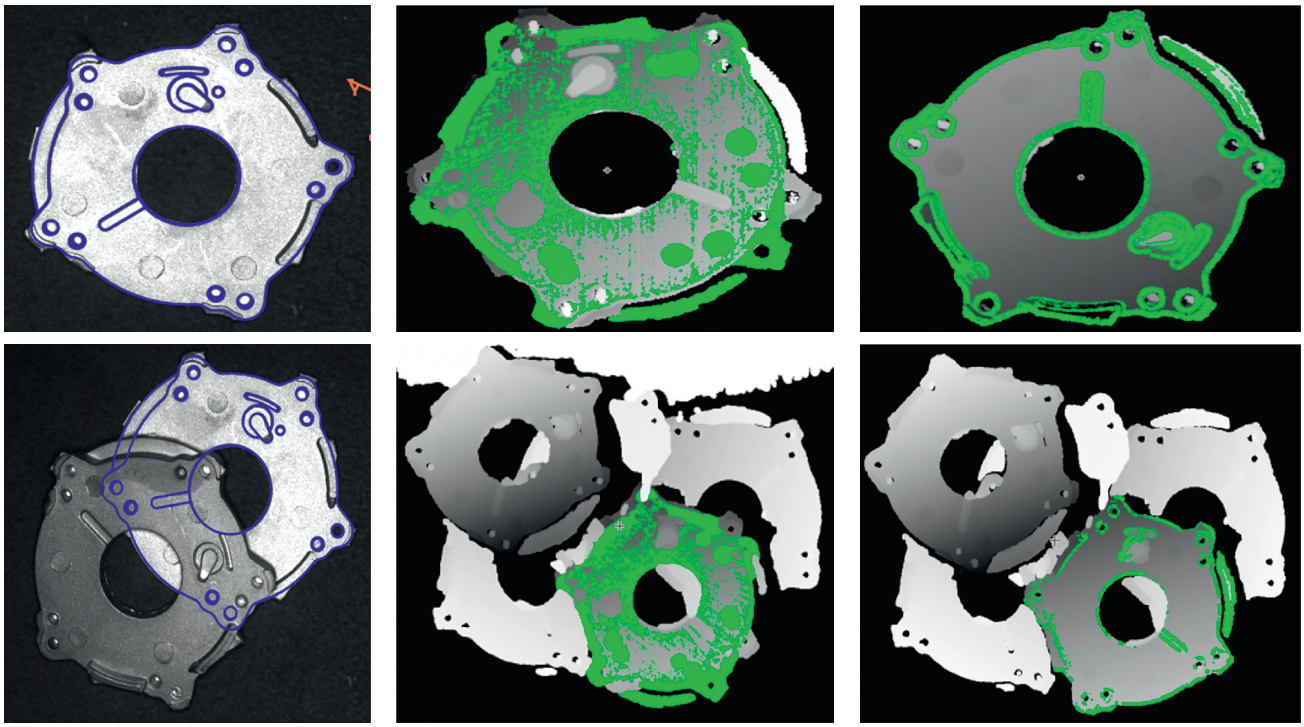


Figure 10: Comparison of measurement results of three methods. Recognition effect of (a) Ulrich's method, (b) the traditional PPF matching algorithm, and (c) the proposed method.

the edge-based object matching method is better than the other two methods in both single and multiple object recognition.

4.3.2. Recognition Speed Comparison. Ten objects were tested with the three methods, and the results are shown in Figure 11.

Figure 11 indicates that the performance of the proposed method is better than that of the first two methods in terms of the matching time and similarity score for the metal castings tested in this study. Unlike the PPF matching algorithm, which only creates one model for the same object, Ulrich's method requires several models, which increase the matching time. The similarity score is used to measure the

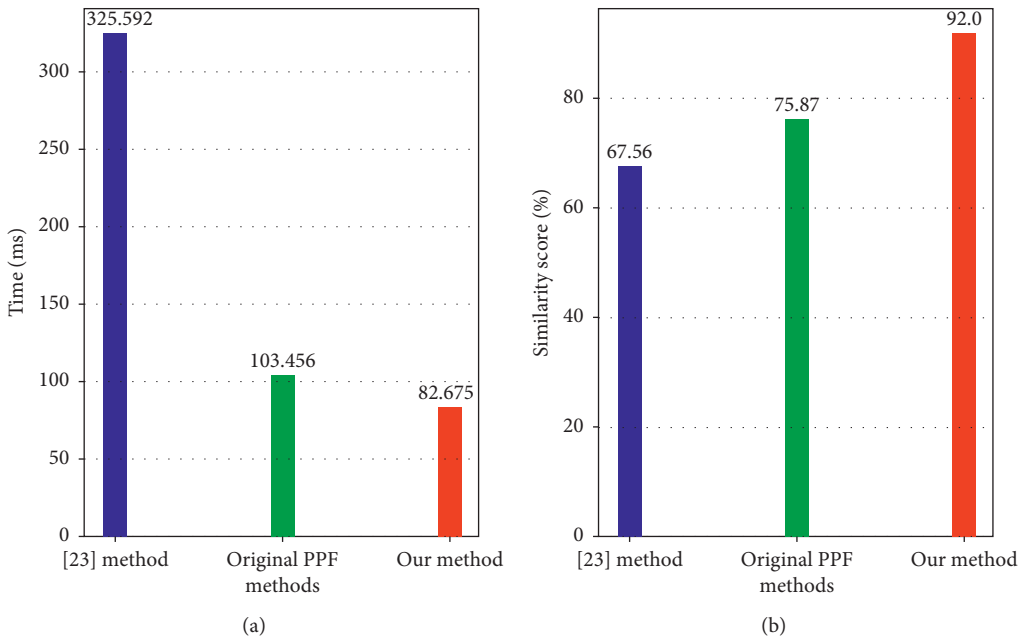


Figure 11: Comparison of three measurement methods.

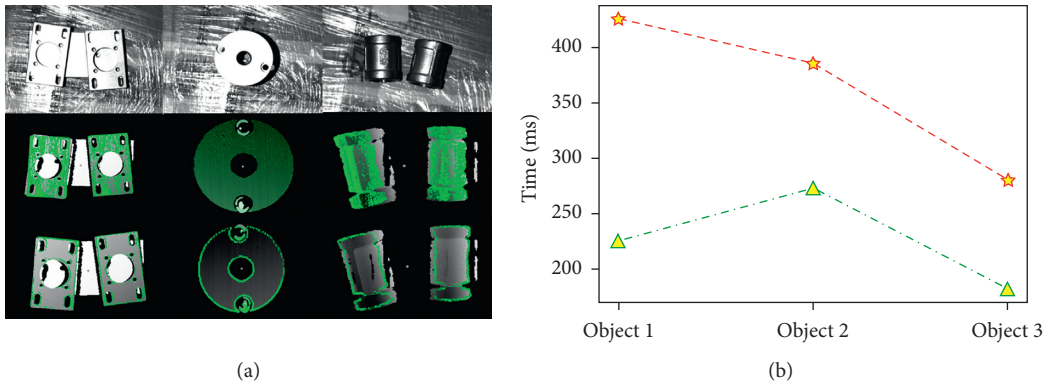


Figure 12: Comparison between the improved PPF and the original PPF matching methods: (a) matching effect; (b) matching time.

similarity between the tested object and the model object. Because the matching feature of the proposed algorithm can better reflect the position relationship between the model object and the object to be tested, the similarity score of our method is also better.

The proposed method was applied to the location test of different shape metal casts with different shapes (see Figure 12).

Figure 12 indicates that the proposed improved PPF matching method is advantageous in both precision and speed for conventional metal castings. The PPF matching algorithm based on the boundary can extract the boundary points that are more representative of the object features in the point cloud and reflect the pose information of the object to be measured with fewer points. Therefore, compared with the original method, the matching accuracy and matching speed are improved.

To verify the stability of the proposed method, a grasping measurement experiment was conducted as follows: the object was placed in the area for robot grasping pose estimation. The object grasping pose of the robot was obtained by hand-eye calibration, which was used as the pose of the robot grasping model. With the movement of the robot, the vision system measured the selected object from different positions many times, and the robot grasping position of the object was calculated after each measurement. The stability of the system was demonstrated by comparing the recorded values with the position and pose of the grasping model. Fifty measurements were performed, and the results are shown in Table 2.

As shown in Table 2, we can find when the number of experiments increases, the positioning accuracy of Ulrich's method is not high. This is because the interference of external light and other factors in the industrial environment

Table 2: Maximum errors of the robot grasping pose for the same object image at different poses.

	x (mm)	y (mm)	z (mm)	a (°)	b (°)	c (°)
Ulrich's method	±3.3	±3.5	±5.5	±2.7	±3.5	NA
PPF matching method	±2.3	±2.5	±2.3	±1.8	±1.8	NA
Edge extraction-based PPF matching method	±2.0	±2.5	±2.5	±1.5	±1.8	±1.6

leads to the 2D vision system's inaccurate extraction of the feature information of the object surface, resulting in the decline of matching accuracy and stability.

Compared with the pose information of other axes, the original PPF matching method does not extract sufficient feature information for the Z-direction rotation of the object. As a result, the original PPF point cloud matching algorithm has small measurement errors for the A- and B-axis rotation in the X-, Y-, and Z-directions but fails to match the C-axis rotation. By contrast, the C-axis matching error of our proposed algorithm is within ±2°, and the matching accuracy of the other axis is also good, which proves the effectiveness of the proposed method.

5. Summary

In this study, a novel 6DOF positioning and grasping approach for industrial robots based on 3D matching was proposed. This approach can guide the industrial robot in realizing high-precision grasping of blank castings placed randomly and achieve a high recognition rate for repetitive and similar features, which is not possible with the traditional PPF algorithm. The main contributions of this paper are as follows. First, the 3D reconstruction method based on binocular structured light was used to realize the 3D reconstruction of objects with rough surfaces. Then, a PPF matching location method based on edge optimization was proposed. The proposed method not only overcomes the ineffective matching of planar features by the traditional algorithm but also improves the matching accuracy and speed. Furthermore, the proposed method was used in an industrial robot grasp strategy to realize metal object positioning and grasping. Finally, the stability and accuracy of the proposed method were confirmed by the comparative experiments conducted for various objects.

At present, the proposed algorithm uses only the surface profile features of the object and does not sufficiently use the radian features of the noncontour information of the object surface. This aspect will be investigated in our future work to further improve the pose estimation accuracy of the proposed system.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (Grant no. 51609033), Natural Science Foundation of Liaoning Province (Grant no. 20180520005), Key Development Guidance Program of Liaoning Province of China (Grant no. 2019JH8/10100100), Soft Science Research Program of Dalian City of China (Grant no. 2019J11CY014), and Fundamental Research Funds for the Central Universities (Grant nos. 3132019005 and 3132019311).

References

[1] S. Hinterstoisser, C. Cagniart, S. Ilic et al., "Gradient response maps for real-time detection of textureless objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 5, pp. 876–888, 2012.

[2] C. Ye, K. Li, L. Jia, C. Zhuang, and Z. Xiong, "Fast hierarchical template matching strategy for real-time pose estimation of texture-less objects," in *Proceedings of the International Conference on Intelligent Robotics and Applications, Lecture Notes in Computer Science*, pp. 225–236, Tokyo, Japan, August 2016.

[3] T.-T. Le and C.-Y. Lin, "Bin-picking for planar objects based on a deep learning network: a case study of USB packs," *Sensors*, vol. 19, no. 16, p. 3602, 2019.

[4] S. Hasegawa, K. Wada, K. Wada, K. Okada, and M. Inaba, "A three-fingered hand with a suction gripping system for warehouse automation," *Journal of Robotics and Mechatronics*, vol. 31, no. 2, pp. 289–304, 2019.

[5] C. Sahin and T. Kim, "Recovering 6D object pose: a review and multi-modal analysis," in *Proceedings of the European Conference on Computer Vision*, pp. 15–31, Munich, Germany, September 2018.

[6] B. Drost and S. Ilic, "3D object detection and localization using multimodal point pair features," in *Proceedings of the International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*, pp. 9–16, Zurich, Switzerland, October 2012.

[7] O. Tuzel, M. Liu, Y. Taguchi et al., "Learning to rank 3D features," in *Proceedings of the European Conference on Computer Vision*, pp. 520–535, Zurich, Switzerland, September 2014.

[8] M. Li and K. Hashimoto, "Fast and robust pose estimation algorithm for bin picking using point pair feature," in *Proceedings of the International Conference on Pattern Recognition*, pp. 1604–1609, Beijing, China, August 2018.

[9] E. Brachmann, A. Krull, F. Michel et al., "Learning 6D object pose estimation using 3d object coordinates," in *Proceedings of the European Conference on Computer Vision*, pp. 536–551, Zurich, Switzerland, September 2014.

[10] E. Brachmann, F. Michel, A. Krull et al., "Uncertainty-driven 6D pose estimation of objects and scenes from a single RGB image," in *Proceedings of the Computer Vision and Pattern Recognition*, pp. 3364–3372, Las Vegas, NV, USA, June 2016.

[11] W. Kehl, F. Manhardt, F. Tombari et al., "SSD-6D: making RGB-based 3D detection and 6D pose estimation great again," in *Proceedings of the International Conference on Computer Vision*, pp. 1530–1538, Venice, Italy, October 2017.

[12] B. Tekin, S. N. Sinha, P. Fua et al., "Real-time seamless single shot 6D object pose prediction," in *Proceedings of the Computer Vision and Pattern Recognition*, pp. 292–301, Salt Lake City, UT, USA, June 2018.

[13] M. Li and K. Hashimoto, "Accurate object pose estimation using depth only," *Sensors*, vol. 18, no. 4, p. 1045, 2018.

[14] D. Liu, S. Arai, J. Miao et al., "Point pair feature-based pose estimation with multiple edge appearance models (PPF-meam) for robotic bin picking," *Sensors*, vol. 18, no. 8, p. 2719, 2018.

[15] Q. Gu, K. Herakleous, and C. Poullis, "3DUNDERWORLD-SLS: An Open-Source Structured-Light Scanning System for Rapid Geometry Acquisition," 2014, https://arxiv.org/abs/1406.6595.

[16] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[17] J.-Y. Bouguet, "Pyramidal implementation of the lucas kanade feature tracker description of the algorithm," *Intel Corporation Microprocessor Research Labs*, 1999.

[18] D. Bazazian, J. R. Casas, and J. Ruiz-Hidalgo, "Fast and robust edge extraction in unorganized point clouds," in *Proceedings of the International Conference on Digital Image Computing: Techniques & Applications*, Queensland, Australia, December 2016.

[19] P. Mark, K. Richard, and G. Markus, "Multi-scale feature extraction on point-sampled surfaces," *EUROGRAPHICS*, vol. 22, no. 3, 2003.

[20] X.-F. Han, J. S. Jin, M.-J. Wang, W. Jiang, L. Gao, and L. Xiao, "A review of algorithms for filtering the 3D point cloud," *Signal Processing: Image Communication*, vol. 57, pp. 103–112, 2017.

[21] G. Wan, F. Li, W. Zhu, and G. Wang, "High-precision six-degree-of-freedom pose measurement and grasping system for large-size object based on binocular vision," *Sensor Review*, vol. 40, no. 1, pp. 71–80, 2020.

[22] M. Ulrich, C. Wiedemann, and C. Steger, "Combining Scale-Space and Similarity-Based Aspect Graphs for Fast 3D Object Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 10, pp. 1902–1914, 2012.

[23] W. Zhu, P. Wang, R. Li, and X. Nie, "Real-time 3D work-piece tracking with monocular camera based on static and dynamic model libraries," *Assembly Automation*, vol. 37, no. 2, pp. 219–229, 2017.