

# Visual Positioning and Grasping Application of Industrial Robot for Casting Parts

Guoyang Wan<sup>1,\*</sup>, Fudong Li<sup>2</sup>, Guofeng Wang<sup>1</sup>

<sup>1</sup>Department of Electronic Engineering School of Automation and Information Engineering, Dalian Maritime University, Dalian, China

<sup>2</sup>Department of Information Engineering, Yangzhou University, Yangzhou, China

## Email address:

wanguoyang@evoluc.com.cn (Guoyang Wan)

\*Corresponding author

## To cite this article:

Guoyang Wan, Fudong Li, Guofeng Wang. Visual Positioning and Grasping Application of Industrial Robot for Casting Parts. *Automation, Control and Intelligent Systems*. Vol. 7, No. 1, 2019, pp. 18-24. doi: 10.11648/j.acis.20190701.13

**Received:** March 13, 2019; **Accepted:** April 29, 2019; **Published:** May 23, 2019

---

**Abstract:** In order to guide an industrial robot to pick a rough casting object, a visual servoing system have been designed to perform the task. This solution present an efficient method to perform robot and vision system's hand eye calibration. Shape matching method is used to find a rough casting object, by combining Line2d algorithm and ROI selection. This vision system can find rough objects with high computational speed and it is robust to environmental noise. The experiment shows that the system repeatability is within 2mm and verify the feasibility of the method and robustness of the algorithm.

**Keywords:** Industrial Robot, Hand Eye Calibration, Template Matching, Machine Vision

---

## 1. Introduction

Manual grinding, in which safety and health of human are all at risk, has always been regarded as a difficult and dangerous job in the industry. Grinding is a process of abrasive machining that produces small particle and leads to environment pollution. In this paper, the necessity to enclose the process and replace human with machine is further discussed. As industrial robots are more flexible than machines, an anthropomorphic robot is employed in this application. Due to the property of flexibility reflected by grinding cells, robots are guided by a vision system instead of establishing a fixed position of the object. Thus, the application of industrial robots offers a new way to deal with the problem that has existed for a long period of time. The method of traditional robotic grabbing refers to manually transporting the workpiece to a special mechanical fixture, and the industrial robot is responsible for performing the work of workpiece grabbing every time it reaches the specified position [2]. In addition to being time-consuming, this method generates a large amount of work, and a new mechanical fixture must be designed when the work line needs to change the type of the workpiece.

If add vision to the industrial robot, great changes will

occur accordingly. The technology of visual grasping employed by industrial robot is regarded as one of the key research points in the industry [3]. In the present market, many visual positioning solutions are available on the basis of industrial robots [2, 3, 4]. Nevertheless, the employment of vision system in the foundry processing industry, castings, especially unprocessed parts are subject to problems such as excessive volume, poor consistency of shape, and harsh application environment, which as a consequence leads to the difficulty in the visual positioning of workpiece [4]. A grinding project for the embedded parts of high-speed rail is jointly developed by this paper, and it helps solve the problem of visual grabbing employed for industrial robots in the process of manufacturing embedded parts by using a high-efficiency method, which undoubtedly provides a new solution for the industry.

At present, the technology of visual positioning and grabbing employed for industrial robots have three kind of main techniques, which are namely 2D based on monocular vision [5], Binocular stereo vision system [6] and 3D vision based on profile or structured light 3D vision [7]. Presently, 2D vision is known as the mainstream. The advantage of image positioning based on 2D vision is low in cost, simple in structure and high in the efficiency of system operation.

Nevertheless, only the angle of rotation of the object perpendicular to the camera can be obtained through the process. Binocular stereo vision system, as two cameras are used by the stereoscopic image coordinates of the images acquired by the left and right cameras are extracted respectively and the feature points are also extracted to acquire the spatial coordinates of the points on the measured object, which thereby helps establish the spatial coordinate system of the measured object. The advantage of this method is that the three axial rotation angles of the product  $x$ ,  $y$ ,  $z$  can be obtained in comparison with the ordinary 2D vision. but the cost of hardware spent on this method is higher than that of monocular vision system, and the accuracy of positioning is greatly affected by the extraction accuracy of the surface feature points of the product in binocular camera images. That is to say, if one of them performs not well, it may fail to give the right position of object. Some 3D sensors can obtain the 3D model through profile and structured light. It means that it can get the information of object's surface in complete depth, and the method is more resistant to interference of external light. However, 3D sensor brands are not large in amount in the market, and the price of sensors is high. In addition, it takes

several seconds to get some objects' 3D model. Thus, not only the difficulty but also the cost are increased.

The grabbing object in this paper refers to the embedded parts of a high-speed rail guide. Before robot grabbing, all the workpieces are placed on the pallet in the same orientation and order. With a fixed number of layers, the pallets are transported to the upper parts of robot by forklift. The 6-axis industrial robot employs a vision system to position and grab the workpieces. Considering the economics of the project and the efficiency of the running cycle, 2D vision is chosen by us to match with the positioning technology employed by the robot.

## 2. System Composition

The vision system adopted for industrial robot is mainly composed of the following parts

- (1). The system of visual acquisition (including industrial camera, lens and light source);
- (2). The system of image processing;
- (3). The system of PLC logic control;
- (4). The system of visual grasping for industrial robot.

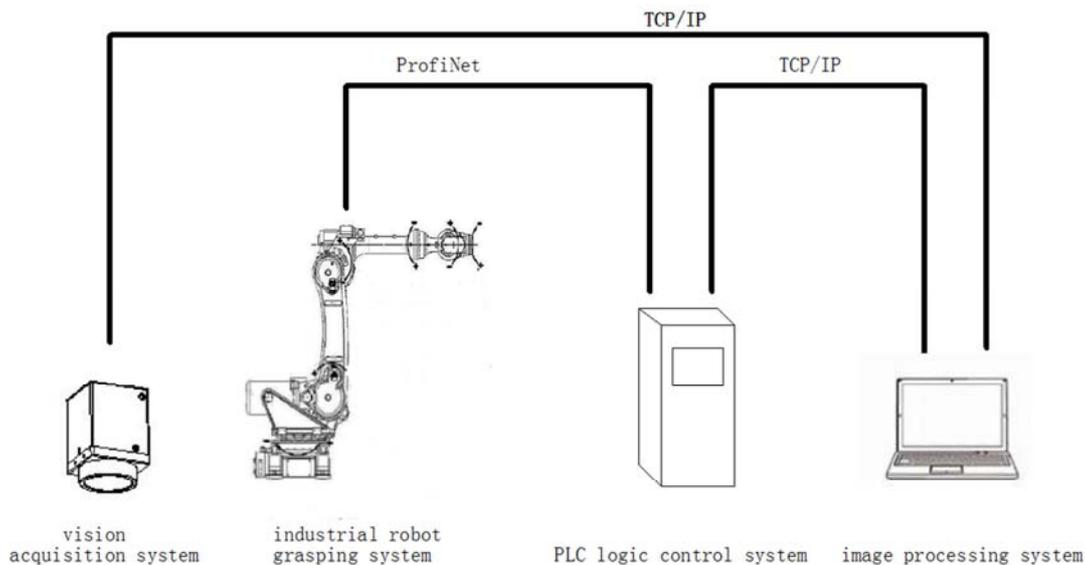


Figure 1. Vision grasping system for industrial robots.

In conjunction with light sources, industrial cameras and lenses work together to offer images with a higher resolution to the vision system. The system of image processing not only provides real-time man-machine interaction and data traceability, but also performs image processing and template matching. In addition, it also does the job related to coordinate translation. The value of the measured object acquired by the camera in the camera coordinate system is converted into the reachable grab position of the robot, and then transmitted to the PLC logic control system of the entire system via the communication bus. The system of PLC logic control is known as the center for logical control of the entire workstation. It is responsible for completing the logic control of all peripherals while completing the data transfer between the vision system and the industrial robot. The industrial

system of robot grabbing is mainly composed of industrial robots and robotic grippers. The industrial robots complete the grasping and placing of the workpieces by combining the gripping work.

Based on the requirements of practical application, the system of robot vision can be generally divided into two types: eye to hand and eye in hand. Eye to hand refers to the situation that the camera is installed outside the industrial robot. In such case, both the position of camera and attitude calibration is relatively simple, but the field of vision possessed by camera is fixed and is difficult to be changed. The eye in hand refers to the situation that the camera is mounted on the body of robot. The camera is usually fixed on the end flange of the robot's 6 axis, and then the view of camera varies with the different postures of the robot. This

method allows the field of vision possessed by the camera to be free from being limited to a fixed range, but the accuracy reflected by positioning of relative camera is affected by the relative posture of robot body.

The role played by the robot vision system in this paper goes like this: the parts placed on the pallet based on certain rules will be taken by industrial robots and then sent to the post process. In accordance with the requirements of the working environment, the way of eye in hand is employed in this paper. Besides, both the camera and the light source are fixed at the end of robot. Cooperating with the vision system through the demolition function of robot, the robot can grab workpiece automatically and send it to the next process.

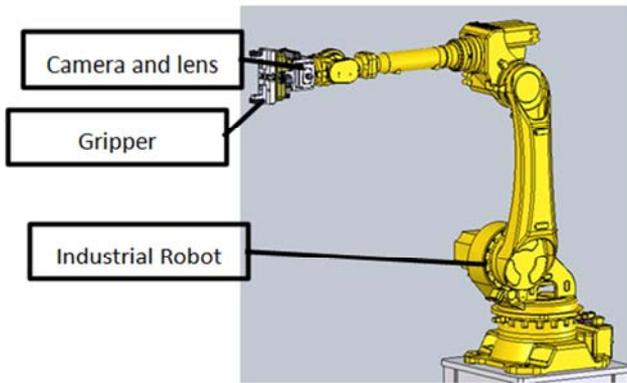


Figure 2. Contour matching process.

### 3. Work flow

The main work flow of the visual system shown by the grinding workstation is displayed as follows:

#### 3.1. Camera Calibration and Hand Eye Calibration of Robot

##### 3.1.1. Camera Calibration

Camera calibration is composed of internal parameter calibration and external parameter calibration. In this paper, the [8, 9] calibration method proposed by Zhang is employed to complete the internal parameter of camera, and then the fractional model with single distortion coefficient is adopted as the distortion model of the camera. With the calibration plate fixed on a plane, the robot is moved to take pictures, and a total of 24 images showing the calibration plates of different attitudes are collected. Table 1 shows the result obtained from calibration. The average error obtained from calibration is 0.45 pixels.

Table 1. Camera Calibration Result.

f/mm	k/mm	Sx/um	Sy/um	Cx	Cy
12.2594	-510.88	1.66967	1.67	1919.07	1329.35

##### 3.1.2. Calibration of Robot Hand Eye and Coordinate Conversion

The hand-eye calibration of the robot is aimed at establishing the relationship between the systems of camera image coordinate and robot coordinate. The traditional

method of hand-eye calibration is employed mainly to control the robot to perform the transformations of pose. In the process, the method of calibration proposed by Zhang is adopted to obtain the calibration board position to camera. Taking into account the motion parameters of the robot controller, the homogeneous equation  $CX = XD$  is established. After multiple translations, the rotation is then conducted to obtain the transformation matrix  $X$  of the camera to the end effector, that is, the matrix of hand-eye transformation.

In terms of the eye in hand mode adopted by this project, a more effective hand-eye calibration method is designed in this paper. The calibration process is shown as follows:

At first, a mirror is placed in the center of the measured plane, and the camera is employed as well to obtain the real-time image. After that, the cross line ( $x=image.width/2$ ,  $y=image.height/2$ ) is added to the center in the vision field of the camera. Then, the mirror is used to check the position of camera lens, and move the robot to the place where the camera is located in the center of the cross line. At this point, we admit that the  $Z$  axis of the camera coordinate system is perpendicular to the measured plane.

Secondly, put a calibration board which can be adopted to distinguish and identify the  $X$  and  $Y$  directions to the center of cross line, and then adjust the position of the calibration board, so that the  $X$  and  $Y$  directions of the calibration board coincide with the cross line in the camera field. In such situation, the coordinate system of camera is calibrated.

After completing the camera calibration, do not move the calibration board. Create a coordinate system for robot workpiece in the calibration board, and put it in the same place as the coordinate system of camera.

Finally, remove the demarcation board, and then place the workpiece at the center of the field of vision. Without moving the workpiece, the robot takes the position  $PI$  of the work piece in the coordinate system of the workpiece, and then the position  $VI$  of the current workpiece is obtained by employing the camera. At this point,  $PI$  refers to the offset value obtained from the tool coordinate relative to the coordinate system of robot workpiece when the robot grabs the workpiece, and  $VI$  denotes the offset value of the workpiece relative to the coordinate system of camera.

As the coordinate system of the camera coordinates with the coordinate system of the robot workpiece, the offset value obtained from the workpiece relative to the camera coordinate system  ${}^C T_p^1$  is equal to the offset value obtained from the workpiece relative to the coordinate system of the robot  ${}^U T_p^1$ . Therefore, the equation is obtained as follows:

$${}^C T_p^1 = {}^U T_p^1 \quad (1)$$

In terms of  ${}^U T_p^1$ , the equation is obtained as follows:

$${}^U T_p^1 = {}^U T_T^1 * {}^T T_p \quad (2)$$

In terms of the robot, the offset of the workpiece is fixed, which is in line with the fixed posture  ${}^T T_p$ . When the robot grasps the posture of the workpiece which is invariable, it is

fixed value, and therefore it can be obtained from the upper form:

$${}^rT_p = [{}^uT_T^1]^{-1} * {}^cT_p^1 \quad (3)$$

When the position of workpiece moves, the offset value obtained from the workpiece in the camera coordinate system is  ${}^cT_p^2$ ,

Therefore, the equation can be obtained as follows:

$${}^cT_p^2 = {}^uT_p^2 = {}^uT_T^2 * {}^rT_p \quad (4)$$

From this, the coordinates of the robot can be obtained after moving the workpiece:

$${}^uT_T^2 = {}^cT_p^2 * [{}^uT_T^1]^{-1} * {}^cT_p^1 \quad (5)$$

### 3.2. Template Matching

Template matching refers to a common method employed for machine vision for the purpose of solving the problem related to positioning in the industrial field. The process of establishing the appropriate ROI (region of interest) through the algorithm of image processing, creating the template, finding the ROI and placing it in the image coordinate system is regarded as the process of template matching.

The method of template matching can be divided into grayscale-based and feature-based matching algorithms, respectively. According to the grayscale statistics, the image is viewed as a digital signal, and the matching relation is found based on the gray value of the pixel. This method is so mature that it is widely applied in a variety of fields, but it has a large amount of calculation in actual use and poor real-time response to speed. Moreover, this method can be easily influenced by the outside light.

On the basis of the algorithm employed for feature matching, the feature of the image is extracted and the feature description factor is generated accordingly. In the end, the features of the two images are matched based on the similarity existing in the description factor. In the process of match, commonly used features can be divided into points, lines (edges), area (surface) and other characteristics. As this method is less influenced by light, it has good performance in robustness and fast speed in response.

The method of LineMod [10, 11, 12] put forward by Hinterstoisser is regarded as one of the most advanced method employed for template matching. Combining the depth information with the gradient information of the color image, this method is usually employed to generate a template by quantizing the gradient direction to a fixed direction, and then several schemes of optimization are adopted for quick measures of windowing similarity between input images [13]. In this paper, the Line2d module in the LineMod method is slightly modified, and thus it is made suitable for grayscale images and improving the recognition rate of the experimental product. The specific process is shown as follows:

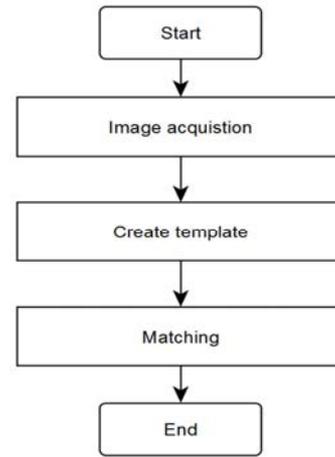


Figure 3. Contour matching process.

#### 3.2.1. Obtain Image

After selecting a standard workpiece, place it in the center of the field of view possessed by the camera according to the imaging posture, and then add a high-contrast background material under the workpiece based on the color and contour of the workpiece to highlight the features of the workpiece. Then, take picture by employing the camera.

#### 3.2.2. Create a Contour-Based Shape Template

On the basis of the characteristics obtained from the workpiece and the actual situation on the spot, the template area is required to have the characters as follows:

- 1) Minimize burrs and impurities;
- 2) Clear outlines and corner features;
- 3) With the directional feature that can be adopted to distinguish the detected workpiece, the template can always be found in the image no matter how the image rotates.

The steps for creating a template are listed as follows:

- 1) After conducting Gaussian blur on the image, use the Sobel Operator to carry out calculation on the gradient, and then solve the gradient direction and amplitude at each pixel to obtain the contour information of the image and retain the contour information which is able to satisfy the threshold.
- 2) After dividing the gradient direction space into  $N$  equal intervals, encode it through the employment of a binary string.
- 3) The direction of the gradient is calculated in the neighborhood of  $M * M$ , and the gradient direction exceeding  $L$  is adopted as well.
- 4) Based on the image, the gradient information independent of the contour is removed. Save the template after reducing the interference between the workpiece and the background.
- 5) After conducting rotation of  $360^\circ$  for the template, save a new template every  $n1^\circ$ .

Figure 4 is a diagram showing the workpiece which are being inspected, and all the workpieces are layered on the pallet in a certain order. Based on the characteristics of the workpiece, the system finally selects a part of the two holes in the center of product as the template by taking into

account the factors such as the mutual dependence of the workpiece and the conditions of on-site illumination. The template is shown in Figure 5 as follows.

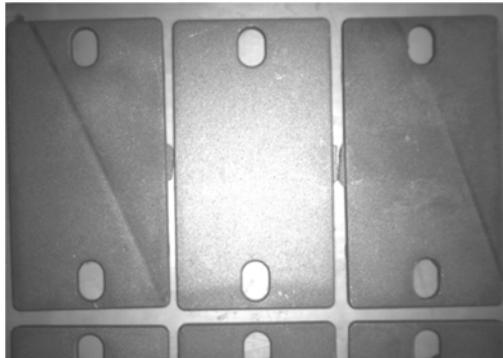


Figure 4. Workpiece put in pallet.

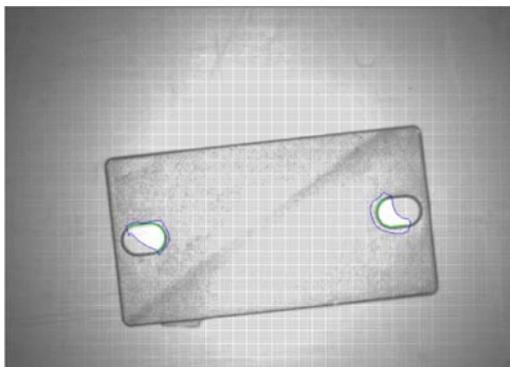


Figure 5. Template.

### 3.2.3. Template Match

After the template is created successfully, the vision system is able to obtain images of other objects to be tested, and then the saved template will be called to match in the

newly obtained image. The matching process of the image to be matched is divided into the three steps as follows:

1) Image preprocessing: Gaussian blurring is performed on the image by employing the same parameters as the template generated, and then Sobel is employed to calculate the gradient, the spatial division of gradient direction, and the operation of binary coding representation. And the gradient direction is diffused in a certain neighborhood, which means that the gradient direction is diffused to the pixel of the  $t$  neighborhood.

2) Calculate the graph of gradient response. The process of calculating the gradient response refers to calculating the response images of  $n$  gradient directions for the search image after the quantized gradient orientation and the diffusion orientation by using the method of traversing, and calculating the orientation value of each pixel. The similarity of the gradient directions is saved as a list.

3) Taking into account the technology of linearized memory of the computer, the matching of the template and the image to be matched is completed by checking them up in the table.

Through the above-mentioned process of matching, not only the information related to the translation position of the workpiece but also the rotation angle of the workpiece can be determined. After the completion of positioning work, the robot is able to perform the coordinate points for workpiece grabbing by taking into account the above-mentioned method of hand eye calibration and coordinate conversion.

### 3.3. Robotic Grab Control and Trajectory Optimization

In the entire grinding system, the control Commands of robot mainly come from the PLC master system. Figure 6 shows the control logic of the main control system.

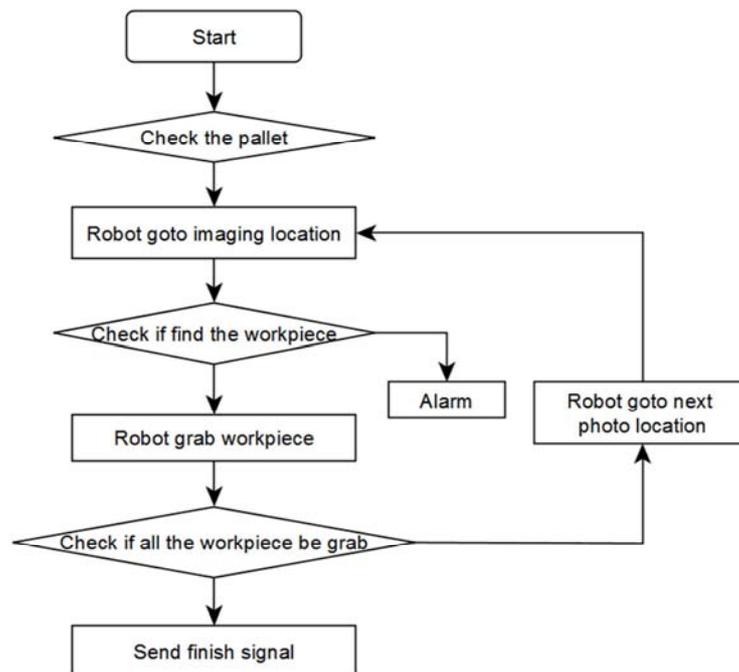


Figure 6. Flow chart of the main control system.

## 4. The Project Is Realized

Figure 7 shows the picture about the actual operation of the workstation. The system focuses mainly on a set of grabbing robots including the vision system, and it employs the electronic control system to complete the loading and unloading of the pallet workpieces.



Figure 7. industrial robot and vision system show

Transportation of the tray is carried out by the forklift truck. According to the different products, products of 6-9 layers are stacked on the tray. All the products are separated by wooden board. When the material is finished on the tray, the worker drops corresponding programs to start the command over the robot. After receiving the instructions, the robot runs to the position of the first work piece, obtains the image, and then carries out calculation on the image. The position of the workpiece is transformed to the coordinate where the robot can grab. It is sent to the robot, and then it is grabbed by the robot. After being grabbed by the robot, the workpiece is moved to the material table, the grinding work is completed by the rear robot, and the next workpiece is grabbed by the robot. When one piece of workpiece is completed, the robot then switches the grip and holds the partition board to the designated position. After that, the robot will switch to the next hand and continue the following work of the next layer.

## 5. Conclusion

This paper proposes an effective method employed for the vision of grasping robot in rough casting. This method is characterized by wide adaptability, high efficiency and good embeddedness. By combining the calibration of coordinate system employed in the industrial robot workpiece with the vision camera, the hand eye calibration of the robot is completed efficiently. The method of shape-based template matching provides a way for locating and grasping rough castings.

## References

- [1] Lars Lienenlücke, Lukas Gründel, Simon Storms, Christian Brecher. Model-based Process Planning for Milling Operations Using Industrial Robots: [J], 2018 3rd International Conference on Control and Robotics Engineering (ICCRE).
- [2] Zhengke Qin, Peng Wang, Member, IEEE, Jia Sun, Jinyan Lu, and Hong Qiao, Senior Member, IEEE, "Precise Robotic Assembly for Large-Scale Objects Based on Automatic Guidance and Alignment". IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, VOL. 65, NO. 6, JUNE. 2016.
- [3] P. K. Allen, A. Timcenko, B. Yoshimi, and P. Michelman, "Automated tracking and grasping of a moving object with a robotic hand-eye system", IEEE Trans. Robot. Autom., vol. 9, no. 2, pp. 152-165, Apr. 1993.
- [4] Ke Xia, Zhengxin Weng, "Workpieces sorting system based on industrial robot of machine vision", The 2016 3rd International Conference on Systems and Informatics (ICSAI 2016).
- [5] Frank Cheng. "Robot Manipulation of 3D Cylindrical Objects with a Robot-Mounted 2D Vision Camera", Computing Conference 2017 18-20 July 2017 | London, UK.
- [6] Heiko Koch, Alexander König, Alexandra Weigl-Seitz, Karl Kleinmann, and Jozef Suchý, "Multisensor contour following with vision, force, and acceleration sensors for an industrial robot", IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, VOL. 62, NO. 2, FEBRUARY. 2013.
- [7] Jae Byung Park, Seung Hun Lee and Il Jae Lee, "Precise 3D Lug Pose Detection Sensor for Automatic Robot Welding Using a Structured-Light Vision System". Sensors 2009, 9,7550-7565doi:10.3390/s90907550.

- [8] ZHANG ZY. "Camera calibration with one dimensional objects". Computer Vision-European Conference on Computer Vision 2002. Berlin Heidelberg:Springer, 2002:161-174.
- [9] ZHANG ZY. "A flexible new technique for camera calibration". IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11):1330-1334.
- [10] S. Hinterstoisser, Member, IEEE, Cedric Cagniart, Slobodan Ilic, Member, IEEE, Peter Sturm, Member, IEEE, Nassir Navab, Member, IEEE, Pascal Fua, Fellow, IEEE, and Vincent Lepetit, "Gradient Response Maps for Real-Time Detection of Textureless Objects". IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 34, NO. 5, MAY 2012.
- [11] S. Hinterstoisser, V. Lepetit, S. Ilic, P. Fua, and N. Navab. "Dominant orientation templates for real-time detection of texture-less objects". In CVPR, 2010.
- [12] S. Hinterstoisser, S. Holzer, C. Cagniart, et al. "Multimodal templates for real-time detection of texture-less objects in heavily cluttered scenes". International Conference on Computer Vision. IEEE Computer Society, 2011:858-865.
- [13] Chan, Jacob, J. A. Lee, and Q. Kemao. "BIND: Binary Integrated Net Descriptors for Texture-Less Object Recognition." 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) IEEE Computer Society, 2017.