A Mixed-line Inspection Robot System for Electrical Home Appliances Based on Multi-vision Cooperation

Shenghong Liu

College of Information Science and Technology, Jinan University, Guangzhou 510632, China.

isshenghongliu@163.com

Abstract

To solve the problem of vision-based automated hybrid detection of incorrect assembly, missing assembly, and mislabeling of household appliances, a robot detection method based on multi-vision was proposed. One industrial camera was fixed in the robot workspace to identify the position and pose of the overall product. With the cooperation of the fixed camera, another camera fixed in the gripper of the robot collects images of the target areas on the product to be detected for inspection. A neural network was used to identify and locate the corners of the top plate of the product, which is in the visual field of camera fixed in the working environment. With the location information of corners and the physical dimension information of the product, the position and pose of the home appliance is identified by PnP (Perspective-n-Point) method. Product pose recognition and robot position computing tasks are processed synchronously in the master control unit, meanwhile, non-real-time region recognition calculation and real-time on-line control are processed and realized through the shared memory communication mechanism. In order to reduce the complexity and assembly interference of multiple hand-eye calibrations, after completing the Eye-in-Hand robot calibration of the gripper camera, the Eye-to-Hand robot calibration of global camera is performed immediately based on its results. Eventually, taking the label detection of the front plate of the air conditioner as an example to conduct a test experiment, the experimental results show the effectiveness and accuracy of the system.

Keywords

Robotics Vision; Electrical Home Appliances Detection and Positioning; Production and Detection Synchronization; Mixed-line Detection.

1. Introduction

With the development of machine vision technology and automation technology, intelligent detection methods using machine vision technology are widely used in different manufacturing fields, such as fabric defect detection [1], tablet inspection [2], paint defects detection on cars bodies [3], defect detection on aluminum surfaces [4], etc. Inspection system usually uses a camera to obtain images of target object, and obtains information by performing various processing and analysis on these images. Collecting reasonable high-quality pictures is not only the first step but also the most critical step of automatic visual inspection, which determines the quality of product inspection.

In the most automated vision inspection systems, for the product to be detected is single-type and simple-structure, images to be analyzed are usually acquired by only one industrial camera placed at the inspection workspace [5]. However, on the household appliance products, there are multiple target areas which need to be detected separately, such as wrong or missing assembly of parts, trademark labeling, and surface defects. In addition, there is not only one type of product in the inspection line. One common solution is to detect each area of the product successively by different cameras fixed in the continuous working station on the assembly line, and every camera is matched for each target area. Another common way is to fix multiple cameras at different position of the same station to collect images simultaneously. Due to the limitation of the complex industrial environment structure and object consistency, those two methods above are difficult to meet the requirements of mixed-line production and rapid line-changing detection.

Benefiting from the robot flexibility, many home appliance enterprises use robot integrated vision module to realize the quality inspection of microwave ovens, refrigerators, air conditioners and other products nowadays. The position and pose of the camera relative to the workspace when acquiring image of the target is preset and fixed, and the movement of the robot is planned by offline programing or manual teaching [6]. S. Akhtar et al. [7,8] planned the working path of the robot by offline programing to realize preset trajectory movements, so that the follow-up camera at the gripper of the robot could obtain static images of the workpiece surface defect. In such method there are limitations: a) When a product arrives at the testing station, there are strict requirements for its position and pose. Once the position of the product is offset, the detection target area may be out of the camera's visual field, leading to detection failure. b) It interferes with the normal logistics rhythm, resulting in poor flexibility and low efficiency.

Therefore, this paper proposes a robot detection system based on multi-vision working mode for household appliances mixed-line inspection. Different from the integration of general robot control system and vision, the self-developed robot controller and two types of industrial camera are used to form the Eye-in-Hand (EIH) configuration and the Eye-to-Hand (ETH) configuration. The ETH vision module is used to realize environmental logistics perception and product pose recognition under time series. The EIH vision module moves along the gripper of the robot and obtains images in the spatial sequence for corresponding detection. In the actual production-detection synchronization environment, the system could collect high-quality images of different target areas of different types of products according to the logistics rhythm, and thereby those images are suit for the corresponding detection steps.

2. Principle of Detection System

In this paper, the detection robot system adopts both ETH and EIH robot-camera configuration. The global ETH camera is used for visual guidance. The EIH camera fixed on the gripper of the robot is used as the detection tool, and it moves with the gripper as the robot moves. For purpose of obtain camera model parameters, conventional camera calibration [9] was performed for these two cameras.

In order to coordinate the operation of various parts of the system, it is necessary to clarify the relationship between multi-vision modules and various coordinate frames of the system. The coordinate frames of the key components of the system and their mutual relationships are shown in Fig. 1.

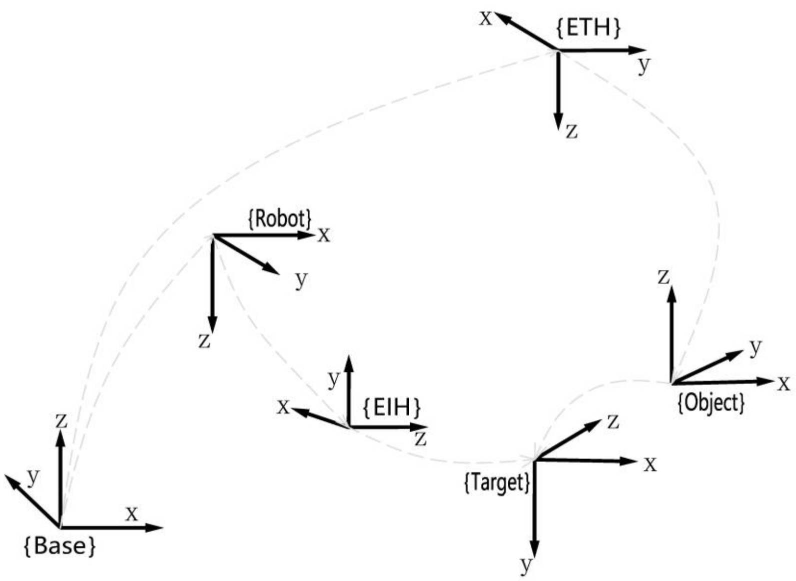

Fig. 1 Definition and interrelation of coordinate frames

We use the following nomenclatures:

• {Base}: The robot base coordinate frame;
• {Robot}: The robot gripper coordinate frame;
• {EIH}: The coordinate frame of camera fixed on the gripper of robot;
• {ETH}: The coordinate frame of global camera;
• {Object}: The coordinate frame of product to be inspected;
• {Target}: ;The coordinate frame of target area on product.

Using the homogeneous transformation (rotation and translation) matrix H to describe the relationship between the coordinate frames, the closed chain formed between coordinate relations can be obtained as the following formula:

$$ {}^{Base}H_{Robot}\,{}^{Robot}H_{EIH}\,{}^{EIH}H_{Object} = {}^{Base}H_{ETH}\,{}^{ETH}H_{Object} \tag{1} $$

After simple matrix operation on formula (1), the position and pose of the robot gripper relative to the robot base coordinate frame can be obtained as formula (2).

$$ {}^{Base}H_{Robot} = {}^{Basse}H_{ETH}\,{}^{ETH}H_{Object}\left({}^{EIH}H_{Object}\right)^{-1}({}^{Robot}H_{EIH})^{-1} \tag{2} $$

In formula (2), ${}^{Basse}H_{ETH}$ is the transformation from frame {Base} to frame {ETH}, and ${}^{Robot}H_{EIH}$ is the transformation from frame {Robot} to frame {EIH}. These two matrices can be obtained by robot hand-eye calibration method. ${}^{ETH}H_{Object}$ is the transformation from frame {ETH} to frame {Object}, and it also represents the position and pose of the product relative to the fixed global camera. ${}^{EIH}H_{Object}$ is the transformation from frame {EIH} to frame {Object}, and it can be obtained by formula (3).

$$ {}^{EIH}H_{Object} = {}^{EIH}H_{Target}\left({}^{Object}H_{Target}\right)^{-1} \tag{3} $$

${}^{Object}H_{Target}$ is the position and pose of the target area (such as labeling position, nut position of fan blade, power line position, etc.) on the detected product, which is set according to the parameters of the detection operation instruction. ${}^{EIH}H_{Target}$ represents the position and pose of the target area in the visual field of the EIH camera when the camera is collecting the image. Because of the different sizes and shooting angles of the target areas, the system needs to set the relative position information of the camera and the target area in advance to ensure that it can acquire high-quality images for subsequent corresponding image processing.

3. Mixed-line Inspection of Household Applications

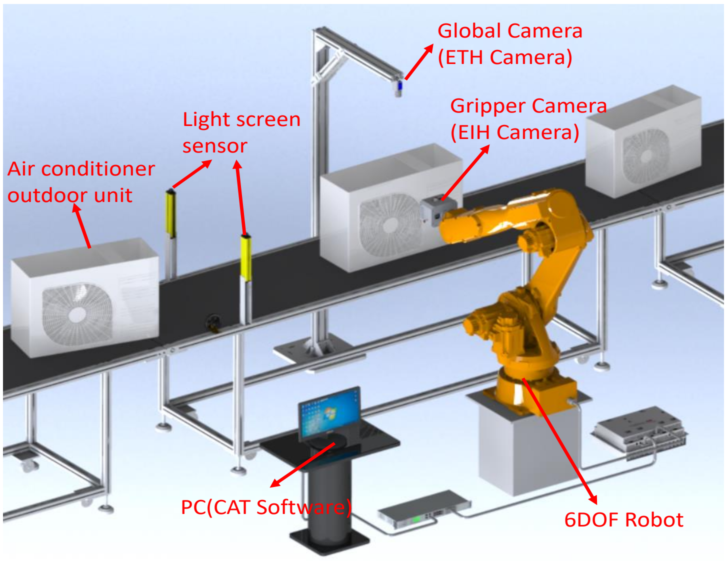

Fig. 2 Air conditioner outdoor unit mixed-line detection scene

In this paper, the detection robot system based on multi-vision cooperation is implemented by taking the detection of mixed-line of the air conditioner outdoor unit as an example, mainly to solve the requirements of appearance assembly and labeling inspection, as shown in Fig. 2.

3.1 Overview of Detection System

The detection robot system is mainly composed of five parts, as shown in Fig. 3:

1) Light screen sensor: used to measure the height and width of the air conditioner outdoor unit and send the measurement results to the computer software, so as to determine the type of the outdoor unit.

2) Robot control system: according to the given position and pose information of the robot terminal, a reasonable trajectory is planned, and the 6 DOF (Degree of Freedom) joint robot executes the trajectory movement to make the EIH camera acquire images of target areas with preset relative posture.

3) Image acquisition system: including two industrial cameras. One is fixed above the conveyor belt and forms the ETH configuration with the robot, used to obtain the overall vision field of the workspace to identify the position and pose of the product. Another camera is fixed on the gripper of the robot and forms the EIH configuration with the robot, which is used to obtain the images of the target areas of the products.

4) Material transfer system: to send different types of air conditioner outdoor units to the inspection workspace, and to send them out after all inspection operation is completed.

5) Computer Aided Test (CAT) software: to control other components of the system and to coordinate the whole system to make the system run stably and effectively.

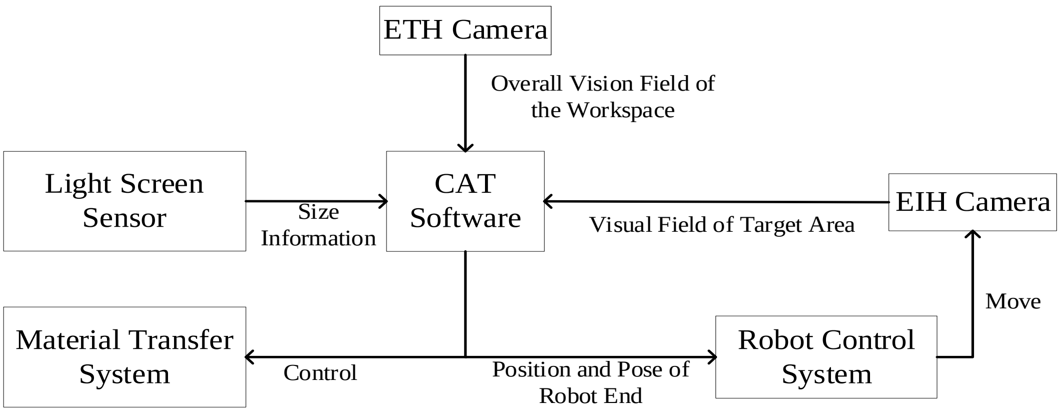

Fig. 3 Composition of detection system

There are various types of products on the mix-line. For each type, the target area to be inspected and the posture of tool camera relative to each target area are usually not the same. So it is necessary to store the size information, the number of target areas to be detected and relative posture while shooting for each type of outdoor unit according to operation instruction in advance. The inspection process is shown in Fig. 4.

Step1 The material transfer system conveys an air conditioner outdoor unit into the operating space. During this process, the width and height of the outdoor unit are measured by the light screen sensor, to provide information for the type judgment.

Step2 The ETH camera obtains overhead visual field of the operation space, from which corners of top plane of the outdoor unit is located. Then CAT software recognizes the position and pose of the outdoor unit.

Step3 For each target area of the outdoor unit to be measured, the terminal pose of the robot to be reached is computed according to the relative relation between the coordinate systems. The robot movement instruction sequence is sent to the robot control system by refreshing the exchange data area within a fixed task cycle.

Step4 The robot controller plans the robot trajectory according to the CAT calculation command list, and the EIH camera follows the robot gripper to collect the images of the area. The detection is completed by the corresponding analysis of the image and the detection results are labeled.

Step5 After completing inspection of all areas of the outdoor unit, the material transmission system drives next outdoor unit into the detection working space for inspection.

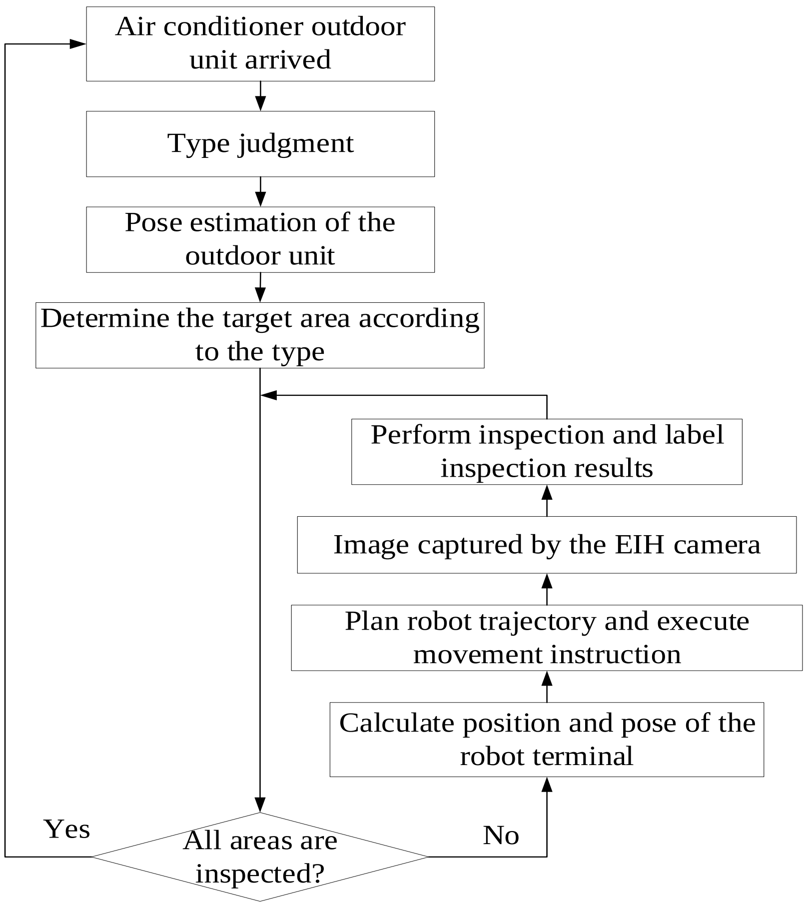

Fig. 4 Flow chart of the system inspection

3.2 Corner location algorithm of air conditioner top plane based on neural network

Neural network have been widely applied in object detection, image classification, image segmentation, human pose estimation and other fields. Restricted to the unstructured industrial environment, changeable illumination and complex background information, it is difficult to identify corners of outdoor unit top plane accurately by using the traditional method of artificially designing image features and graphic models. Inspired by A. Newell [10], this paper utilizes a convolutional neural network to infer the collected digital images from bottom to top (high resolution to low resolution) and from top to bottom (low resolution to high resolution), using multi-scale features to learn the spatial relationship between the corners of outdoor unit top plane, so as to improve the recognition accuracy of corners.

In the process of bottom-up and top-down inference using neural network, feature modules as shown in Fig. 5 are used for feature extraction. The number of input channels is reduced through the 1x1 filters. Then 1x1 filters and 3x3 filters are respectively taken to extract the higher-level semantic features of the output. The two output are concatenated and added with the original input to get the output.

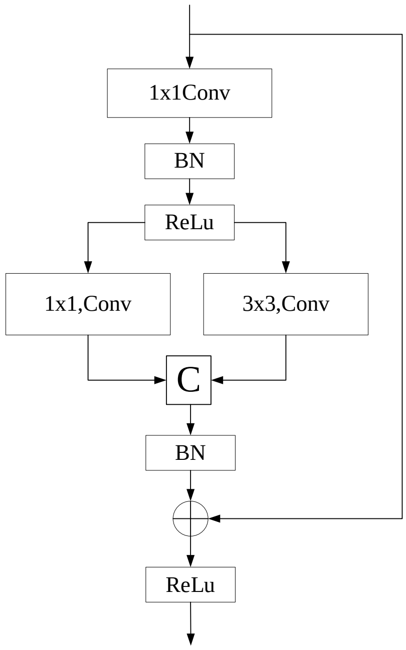

Fig. 5 Feature Module

The network structure for positioning corners is shown in Fig. 6, and the heatmap is predicted by the variable scale feature maps. Firstly, the network takes 3x3 filters to extract feature maps, which have high resolution but low semantic information from grayscale images. Then, the feature extraction and down-sampling are performed sequentially through Feature Module and pooling layer to obtain feature maps with low resolution and high semantic information. Before the pooling layer for down-sampling, the original scale information is retained by branching out, which is added and fused with the same-scale feature maps obtained by up-sampling on the right side of the network. In order to match the channel number of the same scale feature maps, Feature Module also avails extracting features and adjusting the channel number before up-sampling. Afterwards, the final fusion feature map was successively passed through the 3x3 Conv-ReLu module and the 1x1 Conv-Sigmoid module to get the final predicted heatmap. At last, the network outputs four heatmaps, which are respectively used to predict the top-left, top-right, bottom-right, and bottom-left corners of the air conditioner outdoor unit top plane.

Since there are three times of down-sampling in the network, the input image needs to be adjusted into an integer multiple of 16. Consequently, the periphery of each input training image is expanded by zero elements to adapt to the size of the network input. The ground-truth heatmaps of the training process is generated by a two-dimensional Gaussian function. As the proportion of positive and negative samples in the heatmap is seriously unbalanced, the loss is calculated by the improved focal loss proposed by H. Law and D. Jia [11].

$$L_{fl} = -\sum_1^N \sum_1^H \sum_1^W \begin{cases} (1-p_{cij})^{\alpha} log(1-p_{cij}), y_{cij}=1 \\ (1-y_{cij})^{\beta} (p_{cij})^{\alpha} log(1-p_{cij}), y_{cij} \neq 1 \end{cases} \quad (4)$$

Where, p_{cij} is the predicted value at the coordinate (x, y) of the predicted heatmap and y_{cij} is the probability value at the coordinate (x, y) of the ground-truth heatmap. N is the number of heatmaps. H and W are the height and width of the heat map respectively. The factor α is introduced to reduce the loss of easily classified samples, and the factor β is introduced to reduce the penalty of the region near the Gaussian peak.

In the prediction process, non-maximum suppression is performed on each heatmap. Then maximum value of the output is obtained to determine the existence of corner by compared with the confidence threshold. If the result tells that the corner exists, the coordinate of the maximum value on the heatmap means the corner coordinate.

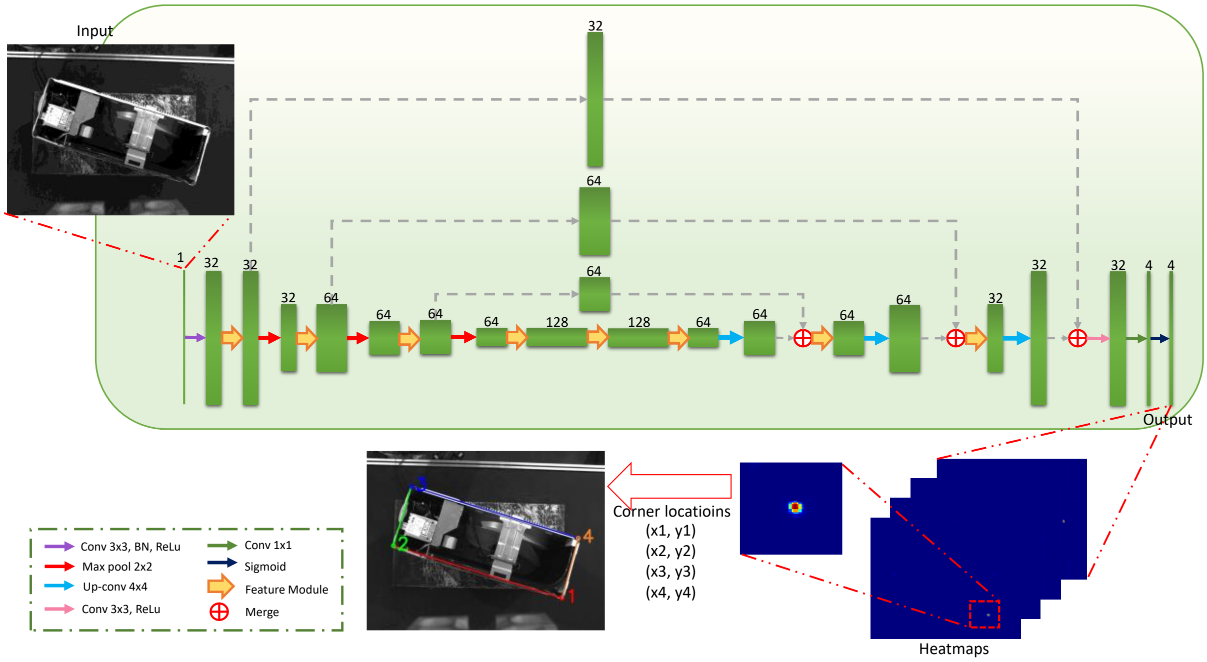

Fig. 6 Neural network structure for positioning the corners of top plane of the air conditioner outdoor unit

In this paper, the position and pose of outdoor unit are identified based on the geometric correspondence relation of projection. The product coordinate system that is fixed on the outdoor unit. The original point coincides with the bottom-left corner; the X axis is directed from the bottom-left corner to the bottom-right corner; the Y axis is directed perpendicularly to the X axis on the top plane of the outdoor unit; the Z axis is directed perpendicularly to the top plane of the outdoor unit.With 2D pixel coordinates of corners and the corresponding 3D coordinates relative to product coordinate frame, the PnP method [12] was utilized to compute ${}^{ETH}H_{Object}$, which describes the position and pose of the outdoor unit relative to the ETH camera coordinate frame.

3.3 Robot hand-eye calibration

The Eye-in-Hand robot calibration refers to solving the relationship between the gripper camera coordinate system and the base coordinate system of robot. And the Eye-to-Hand robot calibration refers to figuring out the relationship between the global camera coordinate system and the robot base coordinate system. Thus the calibration of these two configurations can be abstractly described as solving the homogeneous matrix equation $AX = XB$, of which the solution is the result of hand-eye robot calibration.

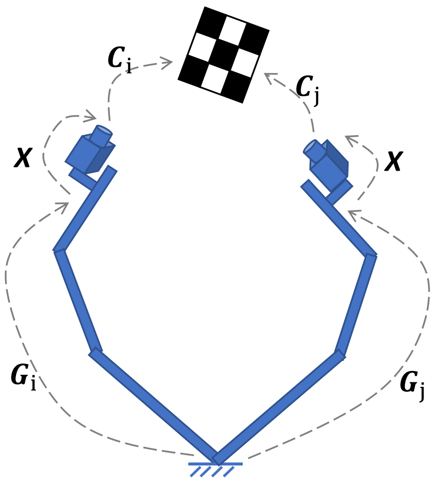

Fig. 7 Eye-in-Hand calibration

The Eye-in-Hand calibration is shown in Fig. 7. Placing a calibration plate in the workspace, the robot is manipulated to move from position i to position j. The moment the robot get to the position i or position j, the calibration plate should not be out of the camera's view of field. Because of the unchanged relationship between calibration plate and robot base, matrix equation $G_i X G_i = G_j X G_j$ can be acquired. By matrix transformation, the equation can be transformed into $AX = XB$, where $A = G_j^{-1} G_i$ and $B = C_j C_i^{-1}$.

Similarly, to the Eye-to-Hand configuration, H. Pan et al. [13] placed the calibration plate on the gripper of the robot and made the robot move from position i to position j, as shown in Fig. 8. Based on the invariability of the relationship between camera and gripper, the homogeneous equation $G_i^{-1} X C_i = G_j^{-1} X C_j$ was derived. By matrix transformation, the equation transformed into $AX = XB$, where $A = G_j G_i^{-1}$ and $B = C_j C_i^{-1}$.

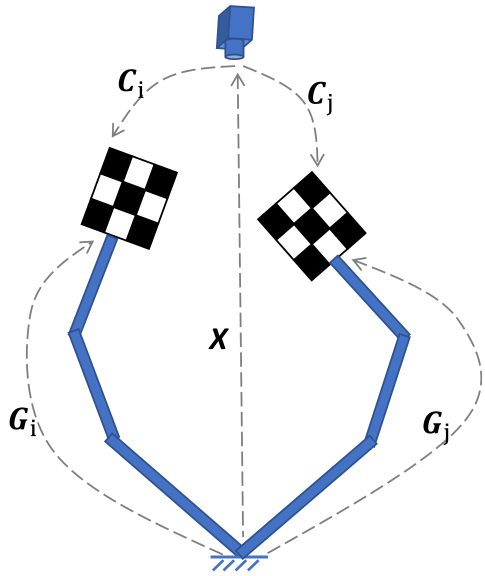

Fig. 8 Eye-to-Hand calibration

The detection system designed in this paper needs to know the hand-eye matrixs of both EIH and ETH configuration. If the above methods are used for hand-eye calibration in turn, there are the following disadvantages:

a) Increase the complexity of mechanical structure configuration: Since fixing the calibration plate on the gripper of the robot will affect EIH calibration, the ETH hand-eye calibration should be conducted in priority. After the ETH calibration is completed, the calibration plate fixed on the gripper of the robot should be removed to continue EIH calibration. The whole calibration process requires an additional mechanical structure to hold the calibration plate in place.

b) Increase the burden of calibration procedure: HTH calibration requires no less than two robot motions (generally 9-20 motions are needed in actual production to ensure calibration accuracy). Moreover, after each robot movement, the calibration plate needs to be completely present in the visual field of camera, which also increases the operational burden of the calibration procedure.

Therefore, this paper adopts the above method for EIH calibration and uses the analytical method proposed by R.Y. Tsai and R. Lenz [14] to solve $AX = XB$ to obtain the EIH matrix ${}^{Robot}H_{EIH}$. However, for ETH calibration, instead of adopting the above method, EIH calibration result is used to solve the problem, making the calibration process more simple. That is, after EIH calibration is completed, the calibration plate and the robot are adjusted so that the calibration plate is in the visual field of both ETH camera and EIH camera, as shown in Fig. 9. According to the homogeneous matrix transfer closed loop, the ETH hand-eye matrix can be obtained:

$$ {}^{Base}H_{ETH} = G({}^{Robot}H_{EIH}) C_{EIH} C_{ETH}^{-1} \tag{5} $$

G is the position and pose of the robot gripper, which can be obtained from the robot controller; C_{EIH} is the calibration plate coordinate frame relative to the EIH camera coordinate frame; C_{ETH} is the calibration plate coordinate frame relative to the ETH camera coordinate frame.

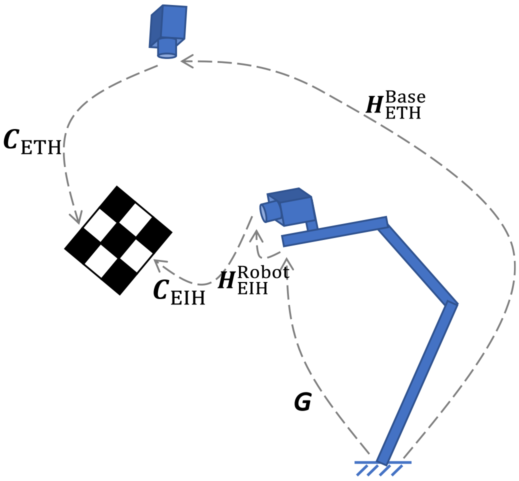

Fig. 9 Modified Eye-to-Hand calibration

4. Implementation and Performance Aanlysis

4.1 Experimental system configuration

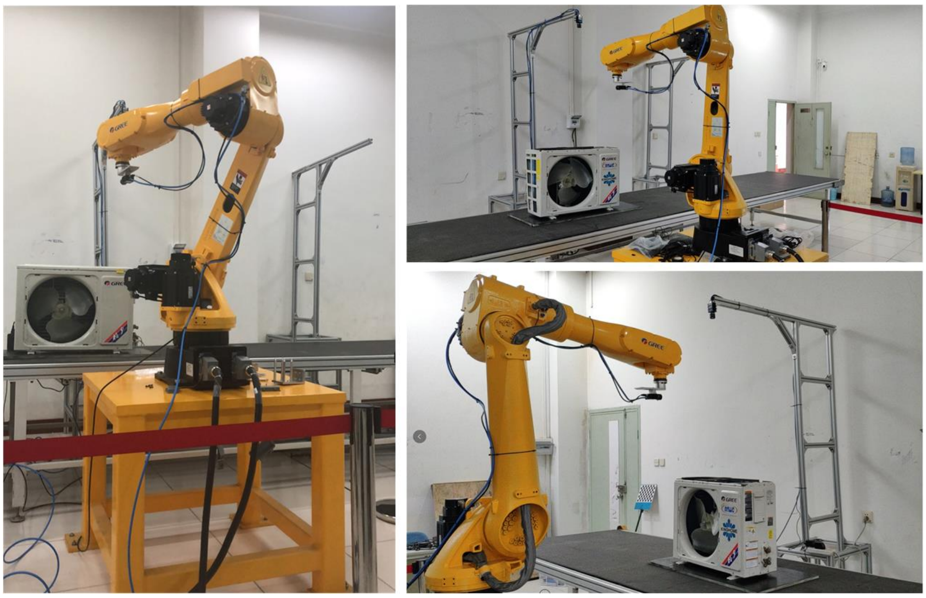

Fig. 10 The robot-based detection setup of air conditioner outdoor unit

In order to verify the feasibility and accuracy of the detection system, an experimental platform for the detection and positioning of air conditioner outdoor unit labels was built, as shown in Fig. 10. System configuration is shown in Table 1.

Table 1. Detection system configuration

Device	Type specification
Robot	Gree
Servo amplifier	TSINO DYNATRON CoolDrive R6 CDR6-A0502-T0-v2
Control system master station	Intel® Core™ i5-3610ME CPU
ETH Camera	IMAGINGSOURCE DMK33GP031
EIH Camera	IMAGINGSOURCE DFK33GP006

The experimental platform applies 6 DOF robot as the detection actuator with arm span of 1813 mm and repeated positioning accuracy of ±0.04 mm. Its D-H parameters are shown in Table 2.

Table 2. D-H parameters of Gree 6DOF robot

#Joint	θ(rad)	d(mm)	α(rad)	a (mm)
1	θ_1	450.1373	π/2	149.7985
2	$\theta_2+\pi/2$	0	0	860.4977
3	θ_3	0	π/2	149.5159
4	θ_4	790.0969	-π/2	0
5	θ_5	0	π/2	0
6	θ_6	140.7968	0	0

4.2 The training and test of the corner location network

The server used for neural network model training in this paper is configured as Ubuntu18.04.4 operating system, Intel(R) Xeon(R) CPU E5-2620 V4 CPU, NVIDIA GeForce RTX 2080TI and 12G running memory.

The dataset is composed of 800 images of air conditioner ' s overlooking angle collected in the workshop. The corner markers are marked in the order of top-left, top-right, bottom-right and bottom-left (looking down at the outdoor unit). The whole dataset is randomly divided into training set, verification set and test set in a ratio of 7:1:2. In order to enhance the training set, every training image is processed randomly during the training procedure, including rotation, horizontal and vertical flipping and variations in brightness and contrast.

The framework of neural network is based on the PyTorch, and the initial learning rate of network training is 0.001. When the training loss does not decrease for 15 consecutive times, the learning rate is reduced by half. Batch size is set to 8, and the epoch is set to 300. Adam algorithm is used to optimize the model. The confidence threshold for inferring the existence of corner points based on the maximum probability value on the heat map is set to 0.8.

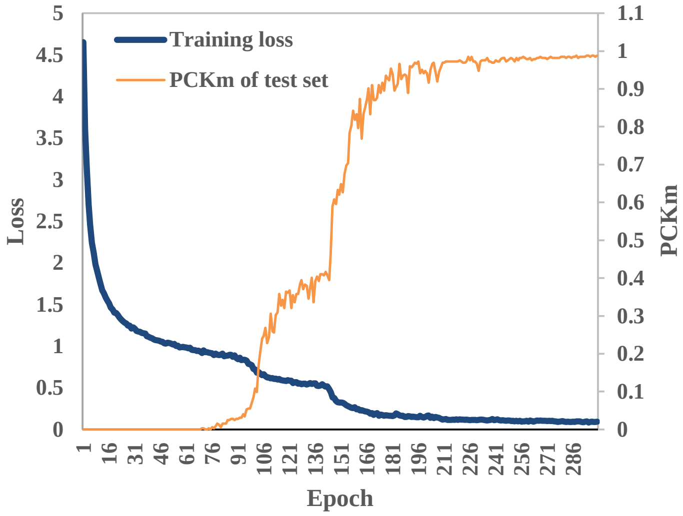

Fig. 11 Training loss and PCKm of test set during training

In the experiment, PCKm (mean Percentage of Correct Keypoints) is used to calculate the proportion of correct corners predicted by the network, which is used as an indicator to evaluate the ability of the network to predict the corners of air conditioner outdoor unit.

$$PCK_m = \frac{\sum_k \sum_i \delta\left(\frac{d_{ki}}{ref_i} < T\right)}{\sum_k \sum_i 1} \tag{6}$$

In the formula, k represents the index of the corners; i represents the image index in the test set; d_{ki} represents the Euclidean distance between the k-th predicted corner point and the artificially labeled corner point in the i-th test picture. Considering the scale difference of the different type of outdoor unit, the Euclidean distance between the bottom-left corner and the top-right corner marked manually is introduced as the scale factor ref to ensure the relative scale invariance of the threshold T.

In the process of network training, the average loss of training set and the change of PCKm indicator of test set corresponding to each epoch are shown in Fig. 11. With the optimization of the network model, the loss of the training set converges and finally converges to the range of 0.085 to 0.1. The PCKm indicator is also significantly improved and finally tends to 0.988, indicating that the network can effectively identify and accurately locate corners of the air conditioner outdoor unit.

When the system performs the inspection, the outdoor unit is completely in the visual field of the ETH camera, and the network should accurately identify all corners of current outer units. The subsequent accuracy of the air conditioner position and pose estimation depends on the corners location. Therefore, formula (7) is used to calculate the average precision (AP) of corner positioning algorithm. If and only if the four corners of the top plane of the outdoor unit are identified and accurately located, the location procedure is determined to be successful. In the network model with 300-th epoch, AP was 96.5%.

$$AP = \frac{\sum_i \delta\left(\sum_k \delta\left(\frac{d_{ki}}{ref_i} < T\right) = 4\right)}{\sum_i 1} \times 100\% \tag{7}$$

Under the framework of Visual Studio 2017 development, the average inference time of each image is 254 ms measured on the control system master station described in Chapter 4.1. Although it is longer than the general computer vision algorithm, it meet the requirements of system dynamic detection along the line.

4.3 Hand-eye calibration results

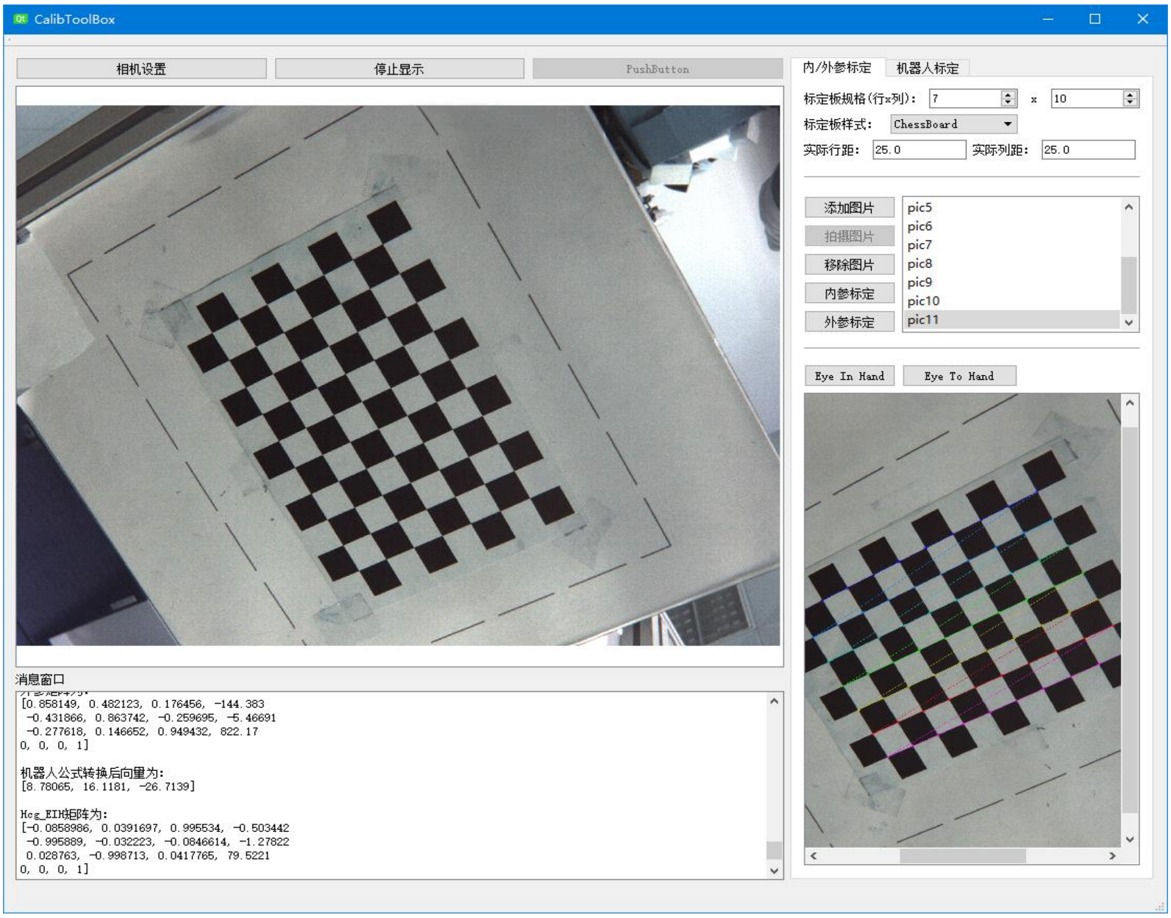

Fig. 12 Robot hand-eye calibration toolbox

Based on the development environment of Visual Studio 2017 and OpenCV 4.1, the hand-eye calibration toolbox for external computer detection system was developed, as shown in Fig. 12. Using the calibration tool , the results of EIH and ETH hand-eye relationship matrix were obtained, as shown below.

$$ {}^{Robot}H_{EIH} = \begin{bmatrix} -0.0859 & -0.0392 & 0.9955 & -0.503 \\ -0.9959 & -0.0322 & -0.0847 & -1.278 \\ 0.0288 & -0.9987 & -0.0418 & 79.522 \\ 0 & 0 & 0 & 1 \end{bmatrix} $$

$$ {}^{Base}H_{ETH} = \begin{bmatrix} -0.0203 & 0.9995 & -0.0203 & 736.631 \\ 0.9996 & -0.0199 & -0.0200 & -87.127 \\ 0.0195 & 0.0207 & -0.9996 & 1637.661 \\ 0 & 0 & 0 & 1 \end{bmatrix} $$

4.4 Performance Aanlysis

In order to verify whether the detection system can arrive at the detection position with the specified shooting pose to acquire clear images of the areas to be inspected, this paper takes the detection of the labeled area of the front panel of outdoor unit as an example to carry out the experiment.

According to the product manual of the external machine, the width and height of the outdoor unit are 700 mm and 530 mm respectively. In addition, relative to the outdoor unit coordinate frame, the central position of the target area that is need to label on the front plate is (560,0, -95) mm, and the pose is to rotate 90 degrees counterclockwise around the x-axis of the outdoor unit coordinate frame. The homogeneous matrix is expressed as:

$$ {}^{Object}H_{Target} = \begin{bmatrix} 1 & 0 & 0 & 560 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & -95 \\ 0 & 0 & 0 & 1 \end{bmatrix} $$

Set the optical axis of the EIH camera to be directly aligned and obtain the image at a distance of 270 mm from the labeling area. The homogeneous matrix between the target area and the camera is:

$$ {}^{EIH}H_{Target} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 270 \\ 0 & 0 & 0 & 1 \end{bmatrix} $$

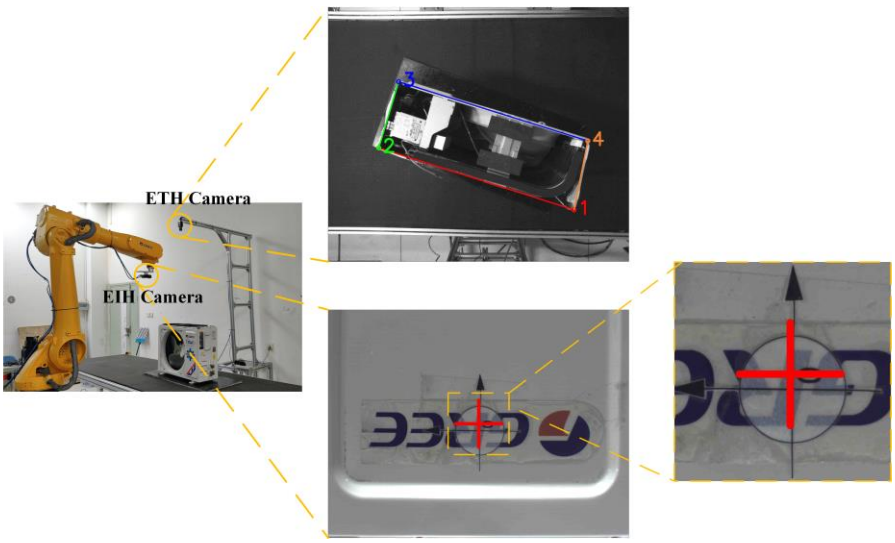

Fig. 13 The processing result during detection

During system operation, the CAT software and the robot controller established a memory sharing data communication mechanism. The air conditioner outdoor unit is placed in a random pose on the conveyor belt and entered the visual field of ETH camera. The pose and position of the outdoor unit is identified by positioning the four corners on the top plane. Then, the position and pose of the robot gripper is calculated by formula (2) and sent to the robot controller to execute the corresponding motion instructions. The EIH camera moves and reaches the suitable position to shoot as the robot moves, see in Fig. 13. Repeating 15 sequential experiments, images of corner positioning result and target area are respectively collected, as shown in Fig. 14. For visual purposes, red marker "+" is applied to mark the center of the EIH camera's visual field.

In order to quantify the relative position and pose of EIH camera and the target area while shooting, an axis paster with circle is pasted in the center of labeling area on the front panel of outdoor unit. The circle on the axis paster is centered at the origin and has a radius of 15 mm. Utilizing the physical coordinates of the point of intersection between the circle and the coordinate axis and its pixel coordinates in visual field of EIH camera, the position and pose of the target area relative to the EIH camera when shooting can be obtained. In this paper, pose was described by RPY angle. To analyze the error, the actual position and pose of the target area are compared with the preset position (0, 0, 270) mm and preset pose (±180, 0, 0)°, and the results are shown in Table 3.

Fig. 14 Corner positioning images of outdoor unit during detecting and the corresponding images of detected target

Table 3. The position and pose of the target to be detected relative to the camera coordinate system

No.	Position (x, y, z)mm	Pose (Rz, Ry, Rx)	Position error (Δx, Δy, Δz) mm	Pose error (ΔRz, ΔRy, ΔRx)
1	(1.71, 1.37, 272.89)	(-179.31, 2.43, -0.06)	(1.71, 1.37, 2.89)	(0.69, 2.43, 0.06)
2	(0.12, -0.83, 266.91)	(-179.41, 1.63, -0.73)	(0.12, 0.83, 3.09)	(0.59, 1.63, 0.73)
3	(-2.61, -2.56, 268.42)	(-178.97, 2.60, 1.22)	(2.61, 2.56, 1.58)	(1.03, 2.60, 1.22)
4	(0.57, -3.39, 268.48)	(-178.87, 1.57, -0.86)	(0.57, 3.39, 1.52)	(1.13, 1.57, 0.86)
5	(7.73, -6.97, 273.18)	(179.47, -0.81, 1.83)	(7.73, 6.97, 3.18)	(0.53, 0.81, 1.83)
6	(5.66, -5.88, 274.84)	(179.64, 1.89, 0.34)	(5.66, 5.88, 4.84)	(0.36, 1.89, 0.34)
7	(-7.37, -1.80, 266.91)	(-179.05, -2.81, -1.71)	(7.37, 1.80, 3.09)	(0.95, 2.81, 1.71)
8	(6.70, -2.74, 269.45)	(-179.33, 2.50, -1.91)	(6.70, 2.74, 0.55)	(0.67, 2.50, 1.91)
9	(0.21, -1.14, 272.62)	(-179.67, -0.89, 2.32)	(0.21, 1.14, 2.62)	(0.33, 0.89, 2.32)
10	(-6.49, 2.97, 267.27)	(-179.61, -0.08, 1.34)	(6.49, 2.97, 2.73)	(0.39, 0.08, 1.34)
11	(0.59, -1.30, 266.93)	(-179.27, 0.48, -1.59)	(0.59, 1.30, 3.07)	(0.73, 0.48, 1.59)
12	(0.85, 4.44, 270.73)	(179.97, 2.83, -2.01)	(0.85, 4.44, 0.73)	(0.03, 2.83, 2.01)
13	(-2.57, -0.81, 263.27)	(-178.93, 2.69, -3.22)	(2.57, 0.81, 6.73)	(1.07, 2.69, 3.22)
14	(-6.55, -3.72, 269.83)	(-178.80, 2.70, -1.07)	(6.55, 3.72, 0.17)	(1.20, 2.70, 1.07)
15	(6.27, -5.10, 274.51)	(-178.69, 3.17, -2.95)	(6.27, 5.10, 4.51)	(1.31, 3.17, 2.95)

According to the statistical analysis on Table 3, the average position errors are (3.73, 3.00, 2.75)mm respectively and the average attitude errors are (0.73, 1.94, 1.54)°. The maximum position deviation does not exceed 7.73 mm, and the maximum angle error does not exceed 2.95°. High-quality images can be collected according to the preset shooting pose, which meets the system detection requirements.

5. Conclusion

In this paper, a vision robot inspection system with multi-vision coordination is proposed to solve the problem that the detection actuator can arrive at different positions flexibly to obtain suitable images in the automatic detection production mixed-line. The system is designed with the example of air conditioner outdoor unit detection. The ETH camera senses the working space environment and guides the robot to drive the camera fixed on the gripper to obtain high quality images. Due to the complex unstructured environment and light transformation, a neural network is designed to identify and locate the top corner information for pose estimation. Because both ETH and EIH configurations are used in the system, this paper adopts a simple method to reduce the complexity of the calibration process by performing the ETH hand-eye calibration based on the results from the EIH hand-eye calibration. The experimental results prove that the neural network in this paper can effectively identify and locate the corners of the air conditioner's top plane, and the system can collect high-quality target pictures with the preset shooting position and pose.

References

[1] D. Schneider, T. Holtermann, D. Merhof: A traverse inspection system for high precision visual on-loom fabric defect detection, Machine Vision and Applications, Vol.25 (2014) No.6, p.1585-1599.

[2] M. Možina, D. Tomaževič, F. Pernuš et al. Automated visual inspection of imprint quality of pharmaceutical tablets, Machine Vision and Applications, Vol.24 (2013) No.1, p.63-73.

[3] L. Armesto, J. Tornero, A. Herraez, et al. Inspection system based on artificial vision for paint defects detection on cars bodies, IEEE International Conference on Robotics and Automation, (2011).

[4] X. Zhang, D.J. Huang: Defect detection on aluminum surfaces based on deep learning, Journal of East China Normal University(Natural Science), (2020) No.6, p.105-114.

[5] E.N. Malamas, E.G.M. Petrakis, M. Zervakis, et al. A survey on industrial vision systems, applications and tools, Image and Vision Computing, Vol.21 (2003) No.2, p.171-188.

[6] T.S. Lembono, Q.L. Yuan, Y.H Zou, et al. Automatic robot taping: system integration, IEEE International Conference on Advanced Intelligent Mechatronics, (2015).

[7] S. Akhtar, A. Tandiya, M. Moussa, et al. A Robotics Inspection System for Detecting Defects on Semi-specular Painted Automotive Surfaces, IEEE International Conference on Robotics and Automation, (2020).

[8] S. Akhtar, A. Tandiya, M. Moussa, et al. An Efficient Automotive Paint Defect Detection System, Advances in Science, Technology and Engineering Systems Journal, Vol.4 (2019) No.3, p.171–182.

[9] Z.Y. Zhang: A Flexible New Technique for Camera Calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22 (2000) No.11, p.1330-1334.

[10] A. Newell, K.Y. Yang, D. Jia: Stacked hourglass networks for human pose estimation, European Conference on Computer Vision (Amsterdam, The Netherlands, October 11-14, 2016). p.483-499.

[11] H. Law, D. Jia: Cornernet: Detecting objects as paired keypoints, European Conference on Computer Vision (Munich, Germany, September 8-14, 2018). p.765-781.

[12] T. Ke, S.I. Roumeliotis: An Efficient Algebraic Solution to the Perspective-Three-Point Problem, IEEE Conference on Computer Vision and Pattern Recognition (Honolulu, HI, USA, July 21-26, 2017). p.4618-4626.

[13] H. Pan, D.L. Wang, Y.S. Qin: A closed-form solution to eye-to-hand calibration towards visual grasping, Industrial Robot, Vol.41 (2014) No.6, p.567-574.

[14] R.Y. Tsai, R. Lenz: A new technique for fully autonomous and efficient 3D robotics hand/eye calibration, IEEE Transactions on Robotics and Automation, Vol.5 (1989) No.3, p.345-358.