Machine Visual Location Method Research in Robot Unstacking

Jin-Quan LI, Fang CUI*, Qi GUAN and Rui ZHANG

Automation School, Beijing University of Post and Telecommunication, 100876, China

Keywords: Machine visual; Image segmentation; Camera calibration; Work pieces' location

Abstract. Work pieces’ location based on machine visual in robot unstacking job has been researched. Adaptive threshold segmentation and modified Hough Transform are used to extract the target region needs further processing. A new method is proposed to choose the appropriate threshold to binary different images with different brightness. Watershed algorithm and Morphology algorithm are used to detect and connect the edges of work pieces. With the distance sensor and camera calibration to complete the location of pallet and work pieces, guiding robot to finish unstacking job successfully.

Introduction

Nowadays, methods of robotic technology with machine visual technology in detecting work pieces' position and shape is rarely used in the industrial robot unstacking operations. Zheng Kui-Jing [1] and others used region growing method to extract the spline area and analyze geometry to achieve precise position based on feature points matching spline, guiding robotic Assembly. Wang Huan-Huan [2] and others adopted MORLC multi directional length encoding method to achieve rapid detection and location of object characteristics, which is suitable for moving objects with changing perspective. Above methods are not suitable for positioning workpieces and pallets under light with changing illumination. A new method is proposed to search the appropriate threshold for images with different illumination to improve the detection accuracy in some extent.

Vision System Design

A machine visual experiment platform have built based on Nachi industrial robots, including work pieces placed module, vision module, robot control module and unstacking module. Work pieces placed module is consistent of conveyor belt and pallet. Vision module is consistent of camera mount, Basler Gigabit Ethernet industrial cameras acA3800-10gm, laser distance sensors DMG-30 and general lighting lamps in factories. Machine control module receives data from computer to control unstacking module which is consistent of robotic arm and hand to finish unstacking operation.

Image Pre-processing

Image segmentation

As is shown in Fig. 1 (a), the original image of packaging - bag placed pallet, including much noise such as conveyor belt, robot pedestal and wooden boxes which have much bad impact on detecting and positioning workpieces. Considering the illumination of images in different time, adaptive threshold method is used to finish initial segmentation to weaken influence caused by different illumination, the segmentation result is shown in Fig. 1 (b). Then modified Hough transform is used...
to detect straight lines in Fig. 1 (b). The difference between modified and traditional Hough transform is that modified Hough transform detect all straight lines in image first and then calculate slope $k$ of each straight line, if slope $k$ satisfy the inequality,

$$0.176 < |k| < 5.67,$$

which means the angle between the straight line and the horizontal direction or the vertical direction is more than 10°, in other words, the straight line deviates much from the horizontal direction or the vertical direction, which means the straight is unqualified and should be deleted. Table 1 shows the quantity of straight line and the time spent in represent these straight lines in two different methods.

<table>
<thead>
<tr>
<th></th>
<th>traditional</th>
<th>modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>quantity</td>
<td>1457</td>
<td>308</td>
</tr>
<tr>
<td>Time[s]</td>
<td>0.507</td>
<td>0.357</td>
</tr>
</tbody>
</table>

To compare the traditional and modified Hough transform, detect straight lines in traditional and modified Hough transform method respectively, then draw the straight lines in a new image which has same size as original image, the result is shown in Fig. 2. Obviously, modified Hough transform is more suitable for detecting the boundary of target connected region.

**Extract the Target Region**

After using modified Hough transform, all connected regions in Fig.1 (b) are detected. Taking area as the criterion to search the target region $A$ which has the biggest area, meaning work pieces placed area, as shown in Fig 3 (a). The result image $B$ of AND operation between Fig. 3 (a) and Fig. 1 (b) is

![Figure 1](image1.png)

**Figure 1.** Result of adaptive threshold segmentation. (a) Before segmentation. (b) After segmentation.

![Table 1](image2.png)

**Table 1.** Comparison of Hough transform modified before and after.

![Figure 2](image3.png)

**Figure 2.** Result of Hough transform. (a) Traditional method; (b) Modified method.

![Figure 3](image4.png)

**Figure 3.** Target region. (a) Outline image of $A$; (b) Target region $B$. 
shown in Fig. 3 (b), apparently through above image processing steps, the amount of data and influence of noise greatly reduced, which makes much easier for post image processing.

**Dynamic Global Threshold Segmentation**

Considering the illumination variation of light is relatively large, the effect of using fixed threshold to binary target region B will be poor, so a new method is proposed to search the appropriate threshold for images with different illumination to improve the detection accuracy in some extent. The specific steps are shown in Fig. 4. Where mean is the average gray value of the target area B, Table 2 lists the value of changing a and b in the dynamic threshold iterations process of region B in Fig. 3 (b).

<table>
<thead>
<tr>
<th>Serial number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>78</td>
<td>58</td>
<td>38</td>
<td>41</td>
<td>44</td>
<td>47</td>
</tr>
<tr>
<td>b</td>
<td>0.726</td>
<td>0.770</td>
<td>0.884</td>
<td>0.858</td>
<td>0.838</td>
<td>0.819</td>
</tr>
</tbody>
</table>

**Table 2. Change of a and b in dynamic threshold iterative process**

![Diagram](image)

**Figure 4. Dynamic threshold iterative process.**

Fig. 5 shows results of global threshold segmentation of images with different illumination, where m is the average gray level of image and t is the global threshold obtained by iterative method in Fig. 4.

![Images](image)

**Figure 5. Global threshold segmentation result of different illumination image.**
Workpiece Positioning

Workpiece Separation

In order to separate all workpieces, first erode target area \( B \) in Fig.5, making gaps between workpieces more obvious to improve the separation accuracy, while reducing the impact of noise. Then extract the all contours in \( B \) after erosion. Contours inversion has been dilated, as is shown in Fig. 6 (a).

![Figure 6. Workpiece separation results. (a) First step; (b) Second step; (c) Final separation result.](image)

In fact, after above steps all workpieces may not completely separated, as is shown in Fig. 6 (a). Using modified Hough transform mentioned above can connect the break of white boundary and the separation result of the workpieces is shown in Fig. 6 (b). Extracting the contours of each separated workpiece to complete the separation of all workpieces, calculate the centroid of each separated workpiece. As shown in Fig. 6 (c), the centroid’s location is marked with white spots.

After the experiment, it is found that the workpiece surface pattern and incomplete pile type has a great influence on the accuracy of the positioning of the workpiece. Therefore, the accurate rate of tray positioning is proposed.

Pallet Positioning

Pallet Contour Extraction

After global threshold segmentation of target area \( B \), there is still noise in \( B \), as is shown in Fig. 5. In order to eliminate noise, using morphological open operation for target area \( B \) in Fig. 5 and then searching the contour which has the biggest area in Fig. 5. Take a rectangular window which has certain width, slide from the four directions to center point of the image with the appropriate step, as shown in Fig. 7 (a). Where the width of window is 20 pixels and the step is 5 pixels.

![Figure 7. Pallet positioning. (a) Searching points; (b) Fitting points’ extraction result; (c) Pallet positioning result.](image)

Computing the number of white pixels in the window and donated the number, denoted as \( N \), searching the position of window where has the biggest \( N \) and collect all white pixels in the window of the specific position in point set \( U1,U2,U3 \) and \( U4 \). As is shown in Fig. 7 (b), the gray thickened part is the points to be fitted in the four directions of the detected contour.

Linear Fitting Locate the Pallet

Considering that the tray is rectangle, fitting points of the four directions respectively is not desirable.
First fit an edge as a reference, then fit the adjacent edges in the vertical direction and fit the other edges in turn. Specific implementation process is as following:
1) Compare N of four point sets and sort them in descending order assumed as U1, U2, U3 and U4;
2) Fitting U1 gives the line L1 as the reference edge;
3) Fitting the points set vertical to U1 gives line L2 and $L_2 \perp L_1$;
4) Fit others point sets in same way.
Fig. 7 (c) shows the result of fitting points sets based on above steps.

**Camera Calibration**

The module in Zhang’s camera calibration [3] is

$$s \begin{bmatrix} u_c \\ v_c \\ 1 \end{bmatrix} = k_c \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + t_c,$$

where $s$ is scale factor, $k_c = \begin{bmatrix} f_x & 0 & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$ is the intrinsic matrix of the camera, $R_c$, $t_c$ are the extrinsic parameters with $R_c$ the rotation matrix and $t_c$ the translation vector. Considering influence of the camera lens, there is a derivation between the actual and the ideal image pixel coordinates, which is called lens distortion. A lens distortion mathematic model is

$$u' = (1 + a_0 r^2 + a_1 r^4 + a_2 r^6)u + s_0 r^2 + (p_0 + p_2 r^2)(r^2 + 2u^2),$$
$$v' = (1 + a_0 r^2 + a_1 r^4 + a_2 r^6)v + s_1 r^2 + (p_1 + p_3 r^2)(r^2 + 2v^2),$$
$$r^2 = u^2 + v^2,$$

where $(u, v)$ is the ideal image pixel coordinate without distortion, and $(u', v')$ is the distorted actual image pixel coordinate, and $(a_0, a_1, a_2), (s_0, s_1), (p_0, p_1, p_2, p_3)$ are coefficients describing radial, prism, and tangential lens distortions, respectively. Use the calibration plate to calibrate the camera's intrinsic and extrinsic parameters and the distortion factor, joined with the distance measured by the distance sensor and formula (1) and (2), the image pixels coordinates of centroid of work pieces and the four corner points of detected pallet can be transformed into world coordinate.

**Conclusion**

The experiment was carried out for four periods of time and 800 pictures was taken in different weathers. As it can be seen from Table 3, woven bag pattern and stack-type have more influence in positioning the workpiece. Comparing with positioning the workpiece, positioning the pallet is more stable and accurate. The whole program for the application of robot vison using in Robot unstacking is an innovation and needs further study.

<table>
<thead>
<tr>
<th>success rate</th>
<th>complete stack</th>
<th>uncomplete stack</th>
</tr>
</thead>
<tbody>
<tr>
<td>workpiece</td>
<td>95%</td>
<td>80%</td>
</tr>
<tr>
<td>Pallet</td>
<td>100%</td>
<td>99%</td>
</tr>
</tbody>
</table>

Table 3. Success rate of different situation.
References