STUDY ON MACHINE VISION SYSTEM FOR TOMATO PICKING ROBOT

Jou-Ching Li¹, Suming Chen¹*, Yu-Liang Chen¹, Yi-Chich Chiu²,
Wei-Ting Tu¹, Pei-Ju Pan¹

¹Dept. of Bio-Industrial Mechatronics Engineering, National Taiwan University, Taiwan
²Dept. of Biomechatronics Engineering, National Ilan University, Taiwan
*Corresponding Author, Tel +886-2-33665350, Email: schen@ntu.edu.tw

Abstract: In recent years, the agricultural technology development in Taiwan was gradually in the process of reforming. The use of robots in agriculture industry has been regarded as an important trend. The research aims to construct a machine vision system for tomato picking using a robot in greenhouses. The system was composed of two parts, which were the image processing program and the binocular stereo system. In the first part, the image processing program graded the tomatoes, and searched their positions individually on two-dimensional coordinates. In addition, the program also worked when tomatoes were overlapping in an image. The second part is the binocular stereo system. The system obtained the three-dimensional coordinates of each tomato feature for the claw path planning. Within the operation area of 600~800 mm, the results showed that the errors were less than 1.50 mm in the X and Y axes, and the errors in the Z axis were under 4.4 mm. Furthermore, the integration tests of the claw and the machine vision system were conducted under a distance of 700 mm. The accuracy rate of picking was as high as 93.3%.


INTRODUCTION

Automation of agricultural production is a significant trend in the present and future. With the advancement of technology and the impact of our accedence to the World Trade Organization (WTO), Taiwan's agricultural pattern is transforming gradually. In order to compete in the
world market, the mechatronic technology in agriculture has been applied more and more for raising the crop yield and promoting its quality. By the efforts in the past decades, Taiwan has laid a good foundation for the production automation in agriculture. The application of robots for harvesting is an important trend in the world, and the technology is maturing.

To detect the crops, the harvesting robots often have two main strategies, which are color and shape (Hannan and Burks, 2004). The first part is the color, which is the most commonly used strategy to detect the characteristics of the crop; for example, setting a number of colors as standard colors, which represent the subject (Jimenez et al., 2000). A different method is using the RGB color layers of the color image to enhance and recognize the subject, such as subtracting the green layer from the red layer or the red divides by the green (Ling et al., 2004; Arima et al., 2004). However, there remain some problems unsolved in using color to detect subjects. One of the problems is how to identify subjects in different colors as crops in different levels of maturity, and a more difficult one is the change of light source. The second part is the shape. A commonly used method is to set a standard shape from a template at first, and then searched the image for blocks similar with the template (Jimenez et al., 2000). However, the most serious problem of this method is the shadowing. If the crops grow into clusters, the fruit occlusion can cause multiple fruits to appear as a single one in an image. Besides, leaf occlusion decreases the fruit area visibility and disrupts the shape of the fruit.

The main purpose of this study is to develop a machine vision system that is suitable for tomato harvesting robot; thus, an image processing program has been developed to determine the external characteristics of tomatoes and a binocular stereo system has been built to provide the actual three-dimensional coordinates of the tomatoes for the claw path planning.

**MATERIALS AND METHODS**

The main components of the machine vision system were a computer (Pentium IV 2.66 GHz, 512 MB RAM), an IEEE-1394 video capture card, and a CCD color video camera (ALLIED, F145C) with a mobile platform (IAI, M100) for the establishment of the binocular stereo vision. For repeating and facilitating the experiment, FRP models of actual tomatoes (c.v. Momotaro) were used as the experimental objects.

The flowchart of the machine vision system is shown in Fig. 1. In the beginning, the coordinates of the reference points and the image calibration were applied to capture an image of the left eye in a good quality. The acquired image was processed with white balance adjustment and spatial calibration, and then the tomatoes in the image were separated from their background. If there was no tomato existed, the carrier would move to the next stop and capture a left image again; otherwise, the program would determine whether the tomatoes can
be harvested or not. If the tomatoes were not mature yet, the carrier would move to the next stop as well; or the camera would capture an image of the right eye. Then the image processing program determined the tomatoes’ external characteristics of the left image, and searched for the relative points of them in the right image. Finally, the system calculated the three-dimensional coordinates of the reference points for the control system of claw.

![Figure 1 The flowchart of the machine vision system.](image)

1. **Image processing**

The image of tomatoes was transformed into the HSI color coordinate system, which is easy for decreasing the influence of a changing light source. After the transformation, the system would be able to separate different tomatoes by color information. Then the system utilized Boundary Extraction Algorithm (BEA) to acquire edge information and the relationship between a pixel and its adjacent pixels (Liu, 2003). The external characteristics of a tomato included the pedicle, the center, and the bottom. To get the center position, the program randomly selected three points to calculate the cores, and repeated according to the size of the tomato’s area. After dilating the image for connecting the near pixels, there were some blocks formed. By picking the biggest block, the centroid of it was regarded as the center position of the tomato. Because the tomato’s shape was like a heart, the circle gradually enlarged from the center until the radius contacted the circumference, and the point was regarded as the pedicle position. For the bottom point, the more distinct the heart shape was, the larger the change of its slope was. Therefore, the slope search method was applied, which took advantage of the edge information from BEA.

3. **Binocular stereo system**

The binocular camera stereo system used the disparity between two camera images and their similar triangle of imaging to obtain the actual three-dimensional position. In Eq. (1), The $XYZ$ are actual horizontal, vertical, and depth coordinates; $x_l$ and $y_l$ are imaging of the target in left image; $x_r$ is horizontal imaging of the target in the right image; $b$ is the distance of camera’s movement; $f$ is the focal length of the camera.
To improve the accuracy of the binocular stereo vision, this system was established by a single camera with a mobile platform. Its accuracy was tested by calculating the coordinates of 45 grid points at 5 different depths, which were 600 mm, 650 mm, 700 mm, 750 mm, and 800 mm. By discussing the relative positions of grid points in X-Y coordinates and comparing the actual distance measured by a laser distance meter in Z coordinate, the accuracy of the binocular stereo vision was validated.

4. Integration experiment

The robotic fruit harvesting system contained the machine vision system and the claw system. This study used red balls (with diameters of 60 mm, 70 mm, and 80 mm) as the tomatoes in the picking experiments. According to the practical operating area in the green house, the integration experiments were conducted for 45 times in the depth of 700 mm.

RESULTS AND DISCUSSION

1. Image processing

After capturing the image, the system transforms it into an HSI color image to separate the tomatoes, and utilize BEA to get edge information (Fig. 2). This study searched external characteristics which include the pedicle, the center, and the bottom of a tomato, and the results are shown as the white points in Fig. 3. The image processing provided all positions of tomatoes, and it also forsook the immature green tomatoes. To avoid using complicated grouping methods, the image was dilated to get the center position quickly, and this strategy worked even though there was overlapping of tomatoes. The image processing results are shown as the black dots in Fig. 4. In this study, the self-developed program solved the most common problem of crops overlapping.

![Image Processing for Tomatoes Detection](image)

(A) Original image (B) Background-separated (C) Edge-extracted

Figure 2 Image processing for tomatoes detection.
2. **Binocular stereo system**

The mobile platform of the binocular stereo system only consumed 1 second for moving 100 mm, and the error was under 0.02 mm. In the operating area between 600 mm to 800 mm, the system correctly calculated the coordinates. Table 1 gives the result of the mean errors in the X and Y axes, which were all less than 1.50 mm, and comparing to the actual distance, the mean errors in the Z axis were under 4.38 mm. Therefore, the coordinates provided by the binocular stereo system were very close to actual value, and the coordinates can be used for the claw path planning.

<table>
<thead>
<tr>
<th>Distance (mm)</th>
<th>X error (mm)</th>
<th>Y error (mm)</th>
<th>Z error (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>600</td>
<td>1.50</td>
<td>1.25</td>
<td>4.37</td>
</tr>
<tr>
<td>650</td>
<td>0.76</td>
<td>1.44</td>
<td>3.70</td>
</tr>
<tr>
<td>700</td>
<td>1.16</td>
<td>1.10</td>
<td>2.03</td>
</tr>
<tr>
<td>750</td>
<td>1.06</td>
<td>1.10</td>
<td>4.38</td>
</tr>
<tr>
<td>800</td>
<td>1.14</td>
<td>0.90</td>
<td>3.78</td>
</tr>
</tbody>
</table>

3. **Integration experiment**

The accuracy rate of picking balls in the diameter of 60 mm, 70 mm, and 80 mm were 93.3%, 91.1%, and 84.4%. If the location of a ball in an image was too close to the image edge, the right image could not capture the ball to calculate the coordinates. In addition, if the target
was too low for the claw, it would decrease the accuracy rate. These were the main reasons for picking failure. However, the accuracy rate of picking was acceptable and the above problems can be solved by modifying the route of the carrier.

CONCLUSIONS

A machine vision system for the detection of tomatoes has been developed. The image processing was composed of background separation, edge extraction and external characteristics detection. It was designed to obtain the tomato coordinates and solve the problems of fruit occlusion through the edge information and image dilating. The binocular stereo system was established by a single camera with a mobile platform. It avoided the error from the calibration between two cameras, and it consumed less time to complete the work. Evaluation of the system consisted of the tomato models image processing program and picking experiments. The machine vision system was validated in picking experiments, and the results showed that it was effective in detecting tomatoes even in the case of overlapping. Furthermore, more than 84% of the targets were detected and picked.

ACKNOWLEDGEMENTS

We would like to express our thanks to the Council of Agriculture, Taiwan for financial support of research Project 98AS-4.2.3-FD-Z5. We would also like to thank for the assistance from J. F. Lin, who was a graduate student at National Ilan University during this project.

REFERENCES

Liu, Y. H. Feature 2003. Analysis and classifier design and their applications in computer vision and data mining. PhD dissertation. Taipei, Taiwan: National Taiwan University, Department of Mechanical Engineering.