JARQ 31, 257-264 (1997)

3-D Vision Sensor for Cherry Tomato Harvesting Robot

I Dewa Made SUBRATA*[1], Tateshi FUJIURA*[1], Seiji NAKAO*[1], Hisaya YAMADA*[2], Masaru HIDA*[1] and Takuji YUKAWA*[1]

*[1] Faculty of Life and Environmental Science, Shimane University (Matsue, Shimane, 690 Japan)
*[2] Yanmar Agricultural Equipment Co., Ltd. (Oyamazaki, Kyoto, 618 Japan)

Abstract

A cherry tomato harvesting robot equipped with a-D vision sensor was constructed for experimental purposes. The 3-D sensor emitted 3 laser beams with 2 components of wavelengths: a red (wavelength 685 nm) and an infrared (830 nm). Three Position-sensitive devices (PSDs) were used to detect the light beams which were reflected from the crop. The shape of the crop was determined by scanning the laser beams. Red ripe tomatoes were detected based on the ratio of the red to the infrared reflected signals. To avoid a collision, the location of the stems and leaves was also recognized using the image processing software.

Discipline: Agricultural machinery
Additional key words: agricultural robot, fruit harvesting, recognition of obstacle

Introduction

Cherry tomato is a cluster-shaped fruit in which every cluster usually contains one or more branches of the cluster stem. The fruit form is globular with an average diameter of about 30 mm and is attached to the cluster's stem through a peduncle and a calyx (Fig. 1). Contact point between a peduncle and a calyx is called joint. Since every cluster usually contains more than 10 individual fruits, there is a high probability that the image of the fruit may overlap with the image of another fruit. The degree of maturity of the fruit is expressed by the change of color from green to red. The fruits in the upper part of the cluster become mature before the fruits in the lower part of the cluster.

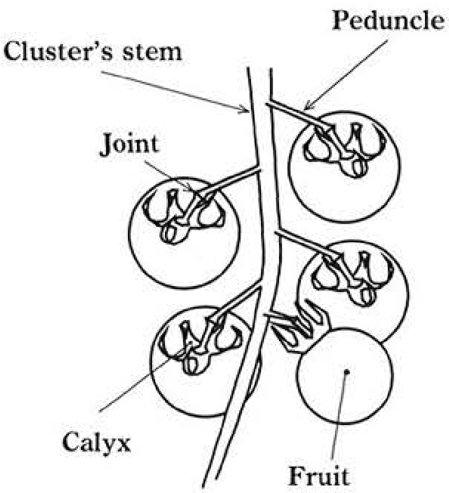

Fig. 1. Cluster of cherry tomato

For mechanical harvesting, the recognition device needs to be built so that it can recognize the position of the mature fruit as precisely as possible. Recognition device was usually constructed by using a video camera with various arrangements[2]. Such a vision system, however, did not enable to detect the leaves and the stems of the fruit plant. Therefore, if the leaves, the stems or the other fruits were located near the aimed fruit, they obstructed the harvesting motion[1]. To obtain a successful harvesting motion, a 3-D vision sensor was considered to be necessary to detect the obstacles.

The main purpose of this study was to develop a 3-D vision sensor to detect the 3-D location of the mature fruits, to separate the overlapping fruits, and to detect the location of obstacles.

Construction of 3-D vision sensor

Based on the reflectance characteristics of the cherry tomato plant (Fig. 2), a 3-D vision sensor (Fig. 3) was constructed by using 2 laser beams i.e. a red laser beam (685 nm) and an infrared laser beam (830 nm). Both beams were installed perpendicularly to each other, and by using a cold filter they were

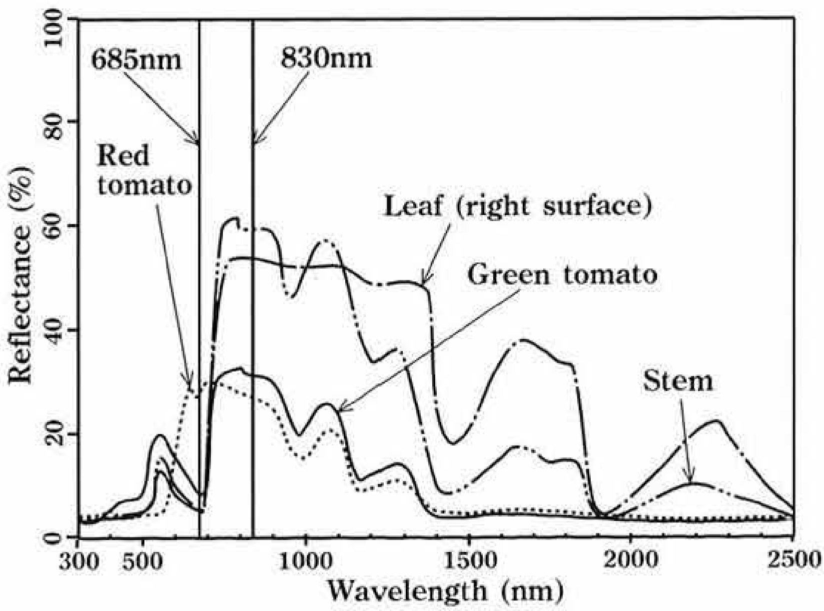

Fig. 2. Reflectance characteristics of tomato plant

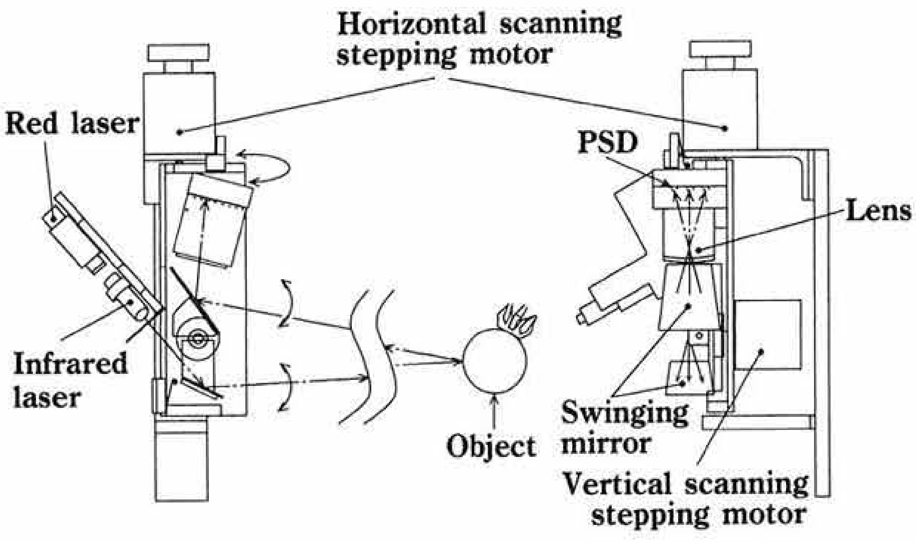

Fig. 3. 3-D vision sensor

united into the same optical axis. The united beam was then split into 3 beams by using half mirrors and a full reflecting mirror in such a way that they formed an angle of 14.4° with each other (Fig. 4). The split beams were transmitted into the crop through a swinging mirror, and the reflected beams from the crop were focused onto the sensitive surfaces of 3 PSDs (position-sensitive device) through another swinging mirror and a focusing lens. A PSD has 2 anodes and the ratio of the 2 anode currents changes with the position of the light spot. The distance can be measured by the ratio of 2 anode currents (the distances to 3 points of the crop were measured simultaneously). The ambient light with a wavelength of less than 580 nm was cut off by using a red glass filter (R-63) that was installed in front of the focusing lens. Two stepping motors were used to scan with a resolution of 120 pixels horizontally and 120 pixels vertically. The vision angle was 43.2° in both the horizontal and vertical directions.

The red light beam was modulated by using a 6.5 kHz square wavelength and the infrared light beam was modulated by using a 13 kHz square wavelength (Fig. 5). The reflected beams from the crop were converted by each PSD into 2 output currents through anode A and anode B. These output currents were pre-amplified by using operational amplifiers and the output signals were demodulated into red ($A_R + B_R$), infrared A_{IR}, and infrared B_{IR} signals by using lock-in amplifiers. These signals were converted into digital signals by using an A/D converter and were then transferred into a note-computer.

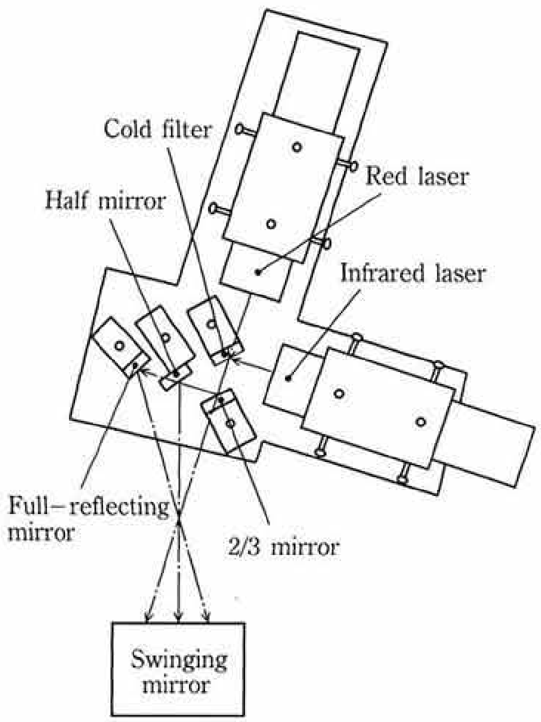

Fig. 4. Optical arrangement

Experimental procedure

Fundamental experiment was carried out at Shimane University to determine the reflectance characteristics of the cherry tomato crop, the accuracy of the distance from the vision sensor, and the threshold value that was used to extract the red fruit from other objects. The accuracy of the

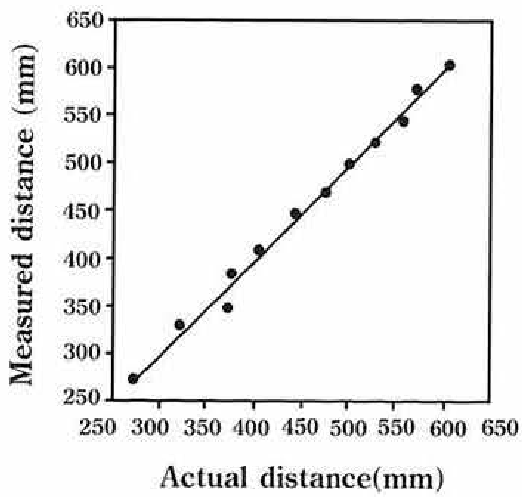

Fig. 6. Measuring accuracy of distance

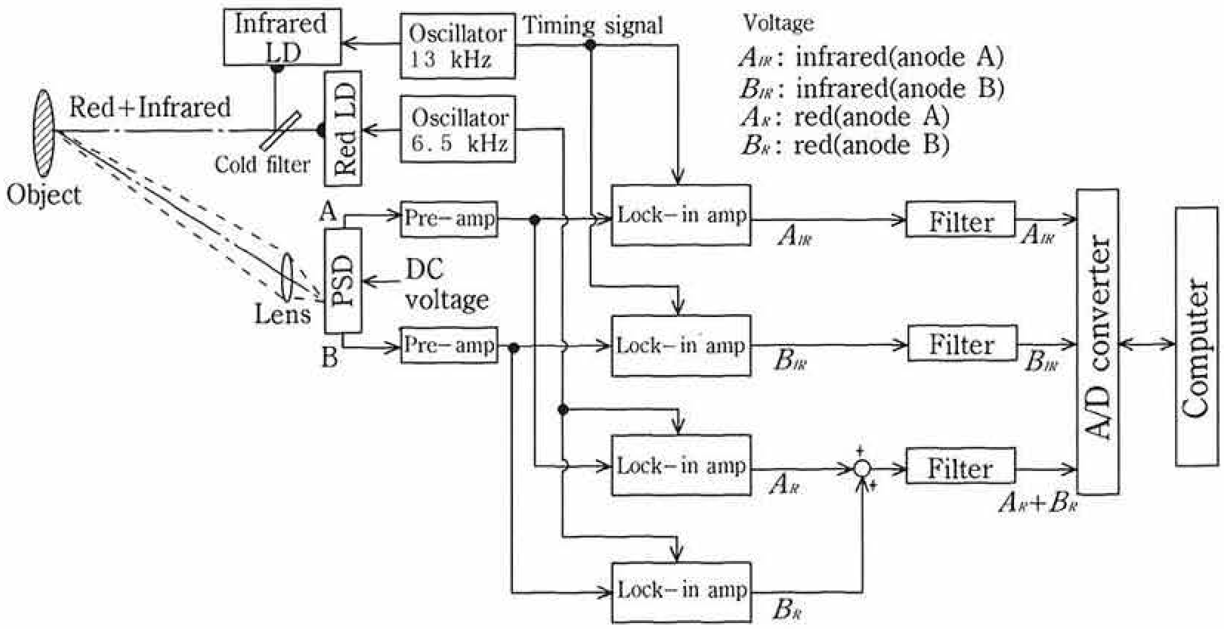

Fig. 5. Block diagram

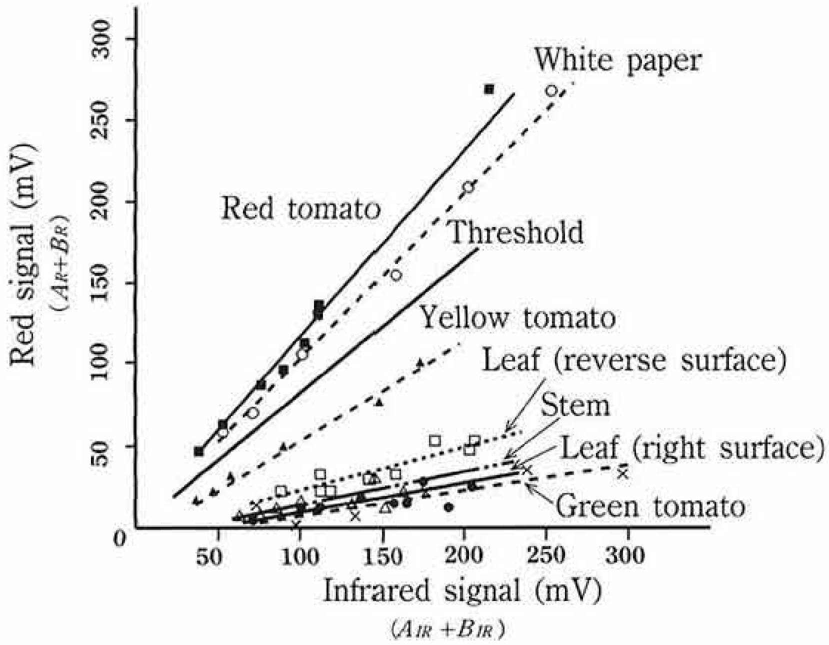

Fig. 7. Relationship between infrared and red signals

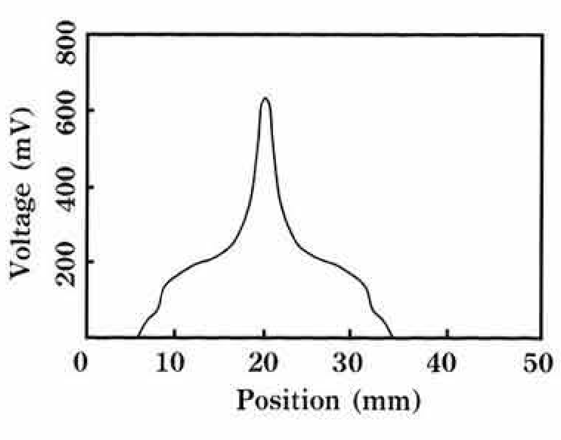

Fig. 8. Red signal

distance (Fig. 6) was examined by scanning a white paper 30 mm in diameter that was pasted in front of the cherry tomato fruit located within a certain distance in front of the 3-D vision sensor.

For predicting the threshold value of the mature fruit, the united beam was transmitted into approximately the center point of the objects that were located within distances of 300, 350, 450, 500, and 550 mm from the center point of the swinging mirrors. Fig. 7 shows the relation between infrared (A_{IR} + B_{IR}) and red (A_R + B_R) signals.

For extracting the individual fruits from their cluster, the mirror reflectance characteristic (MRC) of the fruit surface was measured. The red tomato fruit was fixed at the x-axis of the x-y pen recorder that was located 350 mm in front of the 3-D vision sensor. The laser beam was transmitted through the horizontal plane of the center of the fruit. The red output signal of the lock-in amplifier was connected to the y-axis of the x-y pen recorder. When the pen recorder was turned on, the red fruit moved horizontally with the movement of the x-axis, and the laser beam followed the fruit surface horizontally. The red output signal was recorded by the movement of the y-axis (Fig. 8).

The field experiment was carried out at Yanmar Green Farm where the robot was located between the crop rows in such a way that the horizontal axis of the 3-D vision sensor was perpendicular to the vertical plane of the plant. The fruit was located within the vision area of the sensor. The scanning data consisting of a red (A_R + B_R) signal, an infrared (A_{IR} + B_{IR}) signal, the distance calculated based on the ratio of A_{IR} and (A_{IR} + B_{IR}), and a binary image were stored in a floppy disk for improving the image processing software. The tomato plant shown in Plate 1 was measured by the sensor and was used to evaluate the accuracy of the image processing algorithm. The stored data were plotted by using an x-y plotter (Fig. 9) and the image processing results were displayed on the screen of the computer (Plate 2).

Results and discussion

Fig. 6 shows the experimental results of the measuring accuracy. The measured value of the distance was nearly the same as that of the actual value and it was possible to use the sensor for a distance from 280 to 600 mm.

Plate 1. Cluster for experiment

The 3-D vision sensor was designed to extract mature fruits from others such as immature fruits, leaves, stems, and objects that were used to support the plant. Extraction of the fruits was performed using the threshold value of the ratio between a red signal and an infrared signal that were reflected by the crop. Fig. 7 shows that the ratio was small for green tomato, leaf, and stem containing a large amount of chlorophyll. For the tomato fruit, since the ratio increased with the increment of the degree of maturity, by selecting the ratio between red fruit and green fruit, the red fruit could be extracted from the other parts of the plant. However, other objects such as white polypropylene rope and vinyl chloride pipe that were used to support the crop also gave a high ratio value that was almost the same as that of the red fruit. Therefore, it was deemed necessary to develop an algorithm for extracting such objects. Extraction could be achieved through the identification of their basic properties i.e. slender for the rope and the pipe and globular for the fruit. However, this method was only applicable when the rope or the pipe was not covered by the leaves. The rope or the pipe that was partially covered by the leaves often produced an image very similar to that of the red fruit (identified as red fruit).

Image processing was carried out as shown in Fig. 10. At first, the plant was scanned by using the 3-D vision sensor. The output of the scanning signals consisting of a red ($A_R + B_R$) signal and an infrared ($A_{IR} + B_{IR}$) signal was measured and analyzed to obtain the binarization of the image. When the values of the red and infrared signals were small, it was estimated that there was no object near the 3-D sensor. The ratio of the red signal (A_R +

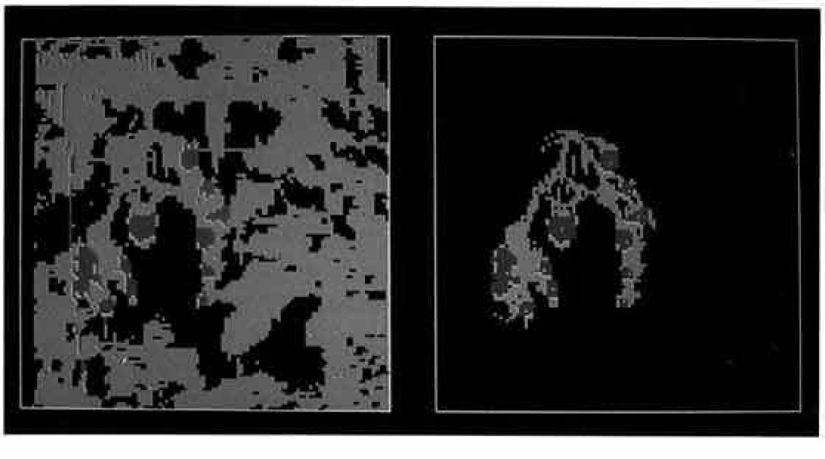

Plate 2. Results of measurement (left), extracted fruit and obstacles (right)

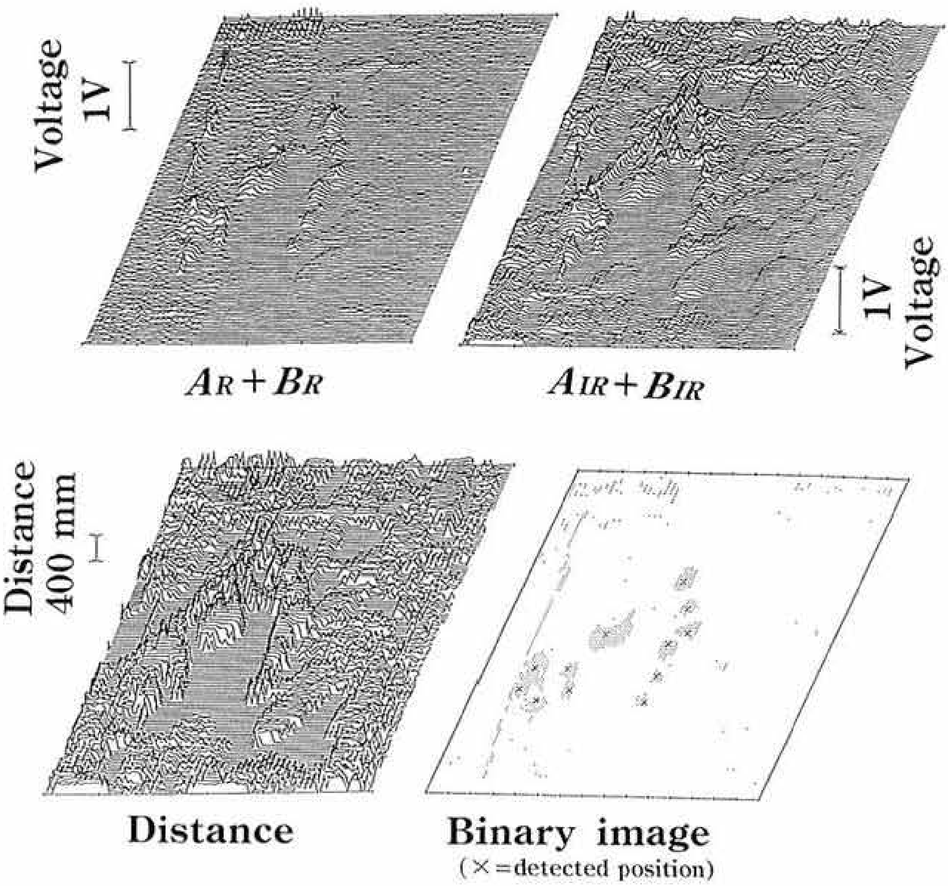

Fig. 9. Results of experiment

B_R) to the infrared signal (A_{IR} + B_{IR}) was calculated and compared to the threshold value of mature fruit to obtain the binarization of the mature fruit pixels. Binary value of the mature fruit pixel was denoted by 1 and that of non-mature fruit pixel by 0. If the sum of the values of the infrared signals (A_{IR} + B_{IR}) was small, the pixel was denoted by 0. Distance of each pixel from the sensor to the crop was calculated by using the ratio between the infrared (A_{IR}) signal to the infrared (A_{IR} + B_{IR}) signal. After binarization of the image and calculation of the distance of each pixel, the red tomato pixels were then grouped to determine the area and size of the fruit image. The red tomato pixels that were detected near another red tomato pixel were incorporated into the same group. The vertical height (h) and the horizontal width (w) of each group were calculated in pixel units. The center point of each group was calculated in pixel units from the coordinate (0, 0) i.e. upper left corner of the image. All these calculations were performed by using an assembler language.

The individual fruit was extracted from each group through the MRC phenomenon. Due to this phenomenon, the light beam was reflected strongly at the center point of the fruit (Fig. 8), indicating that the output signal of the center point of the fruit was higher than that of neighboring pixels. The MRC algorithm was analyzed using the signal of the red signals of (A_R + B_R) as follows: in the group area (the square area with h × w pixels), the pixels that could correspond to a center point of the fruit were screened row by row. If red pixels were present within a square area of 3 × 3 pixels in which the value of the (A_R + B_R) signal of the center pixel was higher by at least 2 (decimal value) than that of other pixels, the area was considered to correspond to the center of the individual fruit (PCIF). After all the PCIFs of one group were detected, the distance between each PCIF was calculated in pixel units. Two PCIFs separated by a distance less than certain values (n was assumed to be constant) were considered to correspond to the same fruit. In such a case, the PCIFs were united to a PCIF and the center was calculated in relation to that of the PCIF with a higher value. When all the PCIFs were

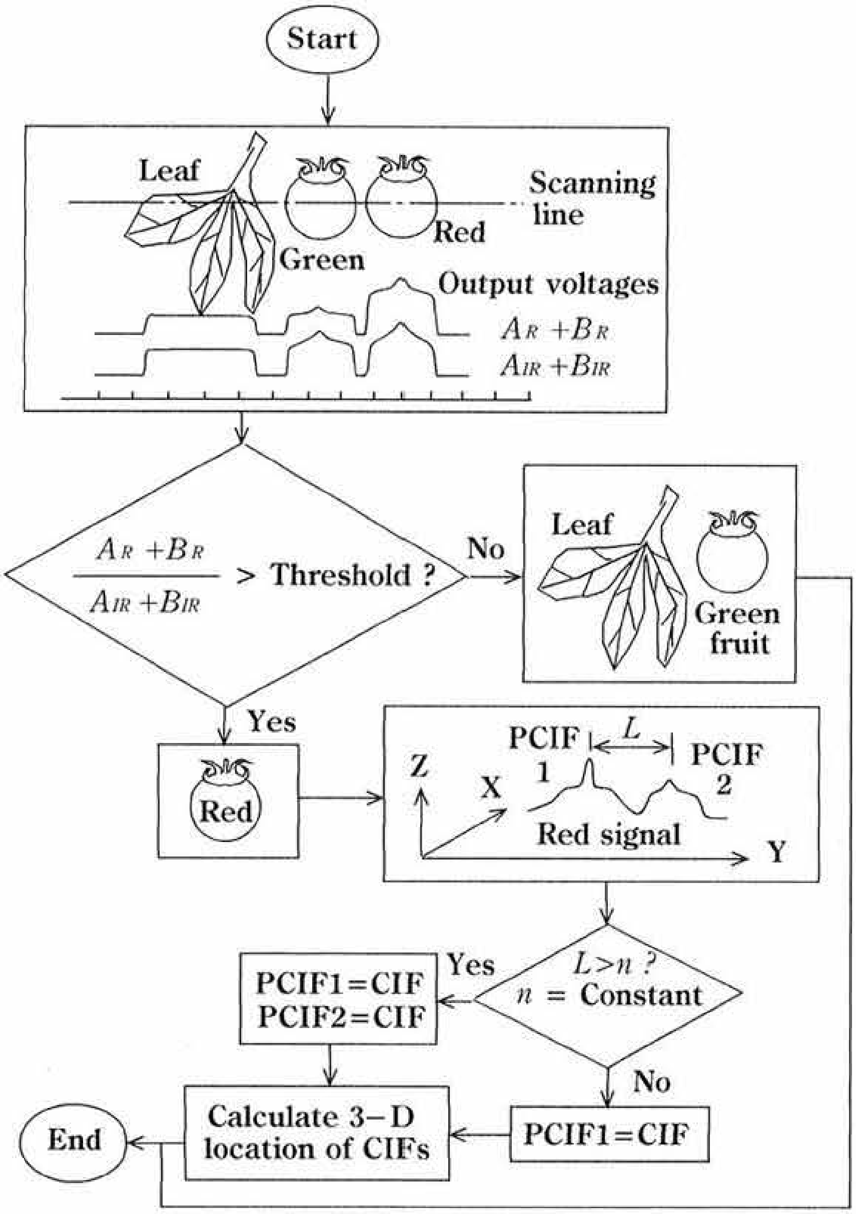

Fig. 10. Flow chart of signal processing

separated by a distance of more than n, they were considered to correspond to the center point of individual fruit (CIF). The coordinate of each CIF was calculated in pixel units and the distance from the sensor to each CIF was calculated by using the average distance of 9 adjacent pixels. The threshold value of the separating distance (n) of PCIF was set up according to the distance from the sensor to the center point of the group and was adjusted based on a field experiment. The 3-D location of each CIF in the unit of length was determined and memorized for harvesting purposes. The number of CIFs within the fruit image was considered to correspond to the number of individual fruits.

The presentation of obstacles was recognized within an area of 50 pixels horizontally and 20 pixels vertically around the CIF. Pixels of the obstacles corresponded to the pixels of the non-mature fruit with a distance from the vision sensor less than the distance of the mature fruit plus 25 mm. For obstacle avoidance, the mature fruit that was located within the area of the obstacle around the targeted fruit was also considered to be the obstacle of the targeted fruit. As the robot has 3 diverging directions (from left side, front side, and right side), the locations of the obstacles were analyzed to fulfill that requirement. If there was an obstacle only on the left side of the targeted fruit, the leftward information was memorized. If there was an obstacle only on the right side of the targeted fruit, the rightward information was memorized. If there were obstacles on both sides of the targeted fruit, the forward information was memorized.

Fig. 9 shows that the red laser beam (A_R + B_R) was well reflected by mature tomato fruits and white polypropylene rope (left side). The infrared laser beam (A_{IR} + B_{IR}) was well reflected by fruits, leaves, stems, and rope. By using the threshold value between red and yellow tomatoes, the binary image of the mature fruits was obtained. The 2-D locations of the individual fruit were recognized using the mirror reflectance characteristic phenomenon as mentioned previously. The recognition results were denoted by X sign in the binary image. The third dimension (distance) of the fruit was calculated using the average distance of 9 pixels that were located around the center point (CIF).

The left side of Plate 2 shows the results of measurements of the tomato fruit. The pixels were classified into 3 colors. The red part shows the mature fruits, the green part the objects that were measured except for the mature fruits, the black part shows the part where the distance was not calculated due to the following reason. Since the distance was generally calculated by the ratio of the A_{IR} signal to (A_{IR} + B_{IR}) signal, the results of measurements were not accurate when the values of A_{IR} and B_{IR} were small. The small values of these infrared signals showed that there was no object near the sensor in the direction of the laser beam. In such a case, since the values of the infrared signals were small, the distance was not calculated because it was estimated that there was no object in the direction of the beam. These pixels appeared in black color in Plate 2. The image on the right of Plate 2 corresponds to the extracted fruits and obstacles. The image of the rope was extracted using the shape characteristic. The location of the obstacle relative to the individual fruit was screened within the area of the obstacle as mentioned previously and was memorized for harvesting purposes.

The experimental results showed that the scanning time for 1 image was 2 seconds and the processing time for 1 image was 1 second.

Conclusions

Based on the above description, the conclusions are as follows: (1) The 3-D vision sensor can be used to recognize mature cherry tomato fruits and determine the 3-D location. (2) The MRC phenomenon can be used to extract the individual fruit from a group. (3) The shape of the crop could be recognized so that the presentation of the obstacle could be analyzed and its location relative to the individual fruit could be determined 3-dimensionally. (4) The scanning time for 1 image was 2 seconds and the processing time for 1 image 1 second.

References

1) Fujiura, T., Yamashita, J. & Kondo, N. (1992): Agricultural robots (1). Vision sensing system. ASAE Paper, #923517.
2) Kondo, N. & Kawamura, N. (1985): Methods of detecting fruits by visual sensor attached to manipulator. *J. Jpn. Soc. Agric. Mach.*, **47**(1), 60–65.

(Received for publication, March 31, 1997)