2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
Daejeon Convention Center
October 9-14, 2016, Daejeon, Korea

Development of An Autonomous Tomato Harvesting Robot with Rotational Plucking Gripper

Hiroaki Yaguchi[1], Kotaro Nagahama[1], Takaomi Hasegawa[2] and Masayuki Inaba[1]

***Abstract*— In this paper, we present a design and development an autonomous tomato harvesting robot. We developed a harvesting robot with stereo camera which can measure depth in short range and direct sunlight and plucking gripper using the infinity rotational joint. We also evaluate the developed robot through harvesting in the tomato robot competition and the real farm. In the tomato robot competition, the robot harvested tometoes from tomato clusters and tomato trees, harvesting speed was about 80[s/fruit] and successful rate was about 60%. In the real farm we evaluated the robot with tomato trees in semi-outdoor environment to show the effectiveness and robustness under direct sunlight. According to the result of harvesting with real tomatoes, we improved the robot motion and finally harvesting speed was up to 23[s/fruit], however the gripper may grasp multiple fruits in case of very cluttered cluster and the calyx also may be broken when the stem angle is deep from the rotation axis. To avoid this situation, a grasp state estimation of the gripper and simultaneous recognition fruit and stem positions are next problems to improve the harvesting successful rate.**

I. INTRODUCTION

In this paper, we present a design and development an autonomous tomato harvesting robot. In cultivation of the tomato, mechanization of the farm itself is progressing to automate and optimize growth of fruits. However, fruit harvesting is not enough automated yet. In the mechanization, some autonomous vehicles using hot water piping as rails, e.g. fruits cargo are introduced, so to decrease initial cost, the robot have to use the infrastructure. The design of the robot is needed to fit to existing installations.

The real farm uses direct sunlight to grow tomatoes, and direct sunlight may jams optical sensors. The farm also has limited workspace because the distance between the ridges are shallow to increase the efficiency of harvest. So it is necessary to employ sensors which can work in an outdoor environment and to design the body size of the robot as small as possible.

Recently, similar harvesting robots are proposed. For example, Van et al. [1][2] developed and evaluated a cucumber harvesting robot using thermal cutter, and its harvesting successful rate was 74.4%. Hayashi et al. [3] proposed a strawberry harvesting robot using scissors and vacuum mechanism and evaluated through field test in the real farm, its harvesting rate was about 40%. Note that it is difficult to compare the harvesting rate because target crops are different.

We proposed a tomato harvesting solution using humanoid robot and scissor hand [4]. In previous work, the recognition rate is about 80%, however, harvesting success rate is about 60% and the robot needs teleoperation support of human. In case of operating the robot in the real farm, the previous robot system is not suitable because we used IR projection 3D camera only for indoor environment for recognition, however the real farm seems to be outdoors under direct sunlight, and the robot size is also larger than workspace of the real farm. In this paper, we develop a tomato harvesting robot using sensor that is capable under direct sunlight and rotational plucking gripper considering autonomous operation in the real farm. We also evaluate the robot function through harvesting experimentation in both indoor and outdoor environments.

II. DESIGN CONCEPT AND RELATED WORKS

According to the survey of Bachche [5], the basic strategy of harvesting robot is subdivided into; harvesting, recognition, and movement. Especially harvesting and recognition are the most important factor, so in this section we discuss these 2 elements.

In our approach, the design requirements of these 2 elements are defined as followings. For harvesting end effector, we do employ a plucking method to simplify recognition methods and to avoid damage to tomato trees and fruits. For recognition, we employ color camera capable under the direct sunlight.

A. END EFFECTOR

According to the survey of Blanes et al. [6], harvesting end effector is classified into the air, contact, and etc. Monta et al. [7] proposed a tomato harvesting mechanism using air vacuum and contact plucking. Bachche et al. proposed some contact cutting end effectors using scissor [8] and thermal cutter [9] Blanes's survey classifies by physical property of gripper, considering methods to detach fruit from tree, solutions are classified into plucking and cutting. For plucking, at least tomato fruit recognition is required, however, for cutting, the recognition, target is not a fruit, but the stem, so cutting method requires more information than fruit recognition. Especially in the tomato recognition, fruit recognition is much easier than stem recognition because tomato fruit has size and color feature. Moreover, to avoid

[1] Hiroaki Yaguchi, Kotaro Nagahama and Masayuki Inaba are with Department of Creative Informatices, Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan {h-yaguchi, nagahama, inaba}@jsk.t.u-tokyo.ac.jp

[2] Takaomi Hasegawa is with DP-Robotics section, DENSO CORPORATION, 1-1 Syowa-cho, Kariya, Aichi, Japan takaomi_hasegawa@denso.co.jp

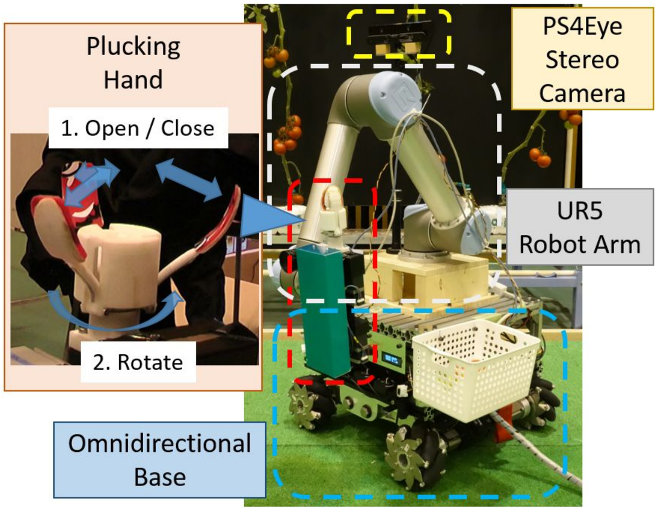


Fig. 1: Tomato Robot

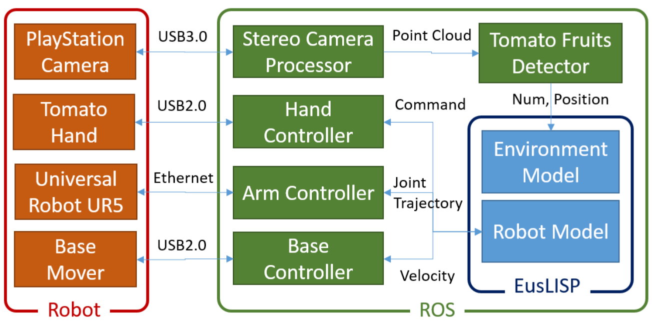


Fig. 2: System Diagram of Tomato Robot

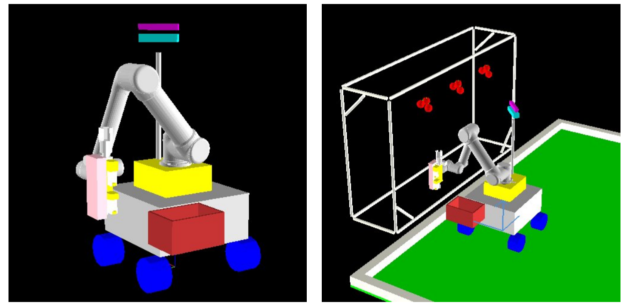
unexpected damages to tomato, we consider plucking is safer than cutting method using edge blade.

B. RECOGNITION

Bachche [5] classified sensor method for recognition into color camera and multispectral camera. We focus into color camera in this paper. Bachche et al. [10] proposed a green sweet peppers recognition using stereo camera and evaluate various color space models. Kondo et al. [11] proposed a recognition method of tomato cluster using stereo camera and illuminator with PL filter. Jiang et al. [12] proposed a tomato fruit recognition method using color detection and stereo camera. Guo et al. [13] proposed a method to detect fruit and stem of strawberry using color features and shape symmetricity of fruit. In these related work recognition method was employed considering the nature of target fruit, for example, color feature is effective for both tomato and strawberry. On the other hand, 3D measurement information is required to reach end effector for harvesting, many research employ stereo camera because it is constructed by 2 color camera and measurement algorithm is also simple.

III. TOMATO HARVESTING ROBOT

In this section we introduce the tomato harvesting robot. Fig. 1 shows the tomato harvesting robot. The robot constructed by; Universal Robot UR5[14] for single arm,



(a) Robot Model (b) Scene Model

Fig. 3: EusLISP model based simulator

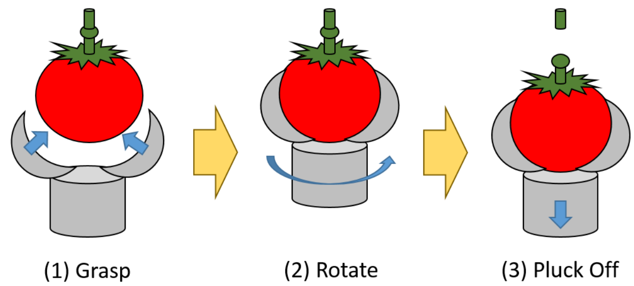


Fig. 4: Image of Plucking

rotational plucking gripper for end effector, PlayStation Camera[15] for recognition, and omnidirectional mover.

We employ the ROS [16] framework for the robot operation, and also use EusLISP [17] and ROSEus (ROS - EusLISP interface) as the geometrical model-beased development environment, as shown in fig. 2. Figs. 3a - 3b show the geometrical robot model and the scene model of the tomato robot competition respectively. Harvesting procedure is simulated by these robot and environment models, and motion of simulated robot model can send to the real robot immediately via EusLISP interpreter.

A. ROTATIONAL PLUCKING GRIPPER

In this paper, we employ a plucking mechanism using gripper and infinity rotational joint. Comparing with scissor hand, scissor requires estimation accuracy of a stem because a stem must put into a gap of blades correctly, on the other hand, plucking hand is robust for estimation error because it can harvest when a fruit is grasped regardless of stem direction.

Red rectangle in fig. 1 shows a picture of the hand. The hand has 2 degree of freedom, fingers open / close and wrist infinity rotation. Fig. 4 also shows sequence of harvesting. In harvesting, firstly the hand grasps a tomato fruit using fingers, nextly whole hand unit rotates. Then a pedicel (bold part of stem as shown in figure) is broken and the hand can pluck off a tomato fruit with its calyx. Finger open / close mechanism has spring and grasp force does not exceed over spring tension for close direction to avoid breaking tomato fruit. Because of this mechanism grasp state can not be estimated from motor load or finger distance, so the hand

TABLE I: Comparison between PlayStation Camera and Xtion PRO LIVE

Camera	PlayStation Camera	Xtion PRO LIVE
Resolution	640×400[px $\times$ px]	640×480[px $\times$ px]
Frame Rate	60[fps] (640×400)	30[fps] (640×480)
Focal Length	2.5[mm]	n/a
Baseline Length	80[mm]	n/a
Pixel Size	6[μ m] (640×400)	n/a
Field of View	83[°]	70[°]
Distance of Use	0.3-∞ [m]	0.8-3.5 [m]

requires to rotate specific time to pluck off fruit.

B. 3D STEREO CAMERA

In previous work we employed an IR pattern projection 3D camera Xtion PRO LIVE[18]. Problems of the IR pattern projection based 3D measurement like Xtion are; it can not use under brighter illumination than projector e.g. direct sunshine, and it has shallow measurement range. In case of operating in real farm, the effect of the sunlight is not negligible. In terms of the measurement range, the short range limit of xtion is 0.8 [m] and the reachable distance of UR5 is less than 0.85 [m], then its margin is only 50 [mm]. It is also farther than the supposed distance to the tomato tree 0.5 [m] expected from the distance between ridges. So in this paper we employ a stereo camera PlayStation Camera[15] for recognition. Stereo camera can measure distance from disparity between left and right images via simple block matching, it depends on only camera image and there is no limitation according to IR projection. Center column of table I shows a optical specification of the PlayStation Camera. According to the comparison with Xtion PRO LIVE as shown in table I, the PlayStation Camera can use under direct sunlight and has enough measurement range especially near side.

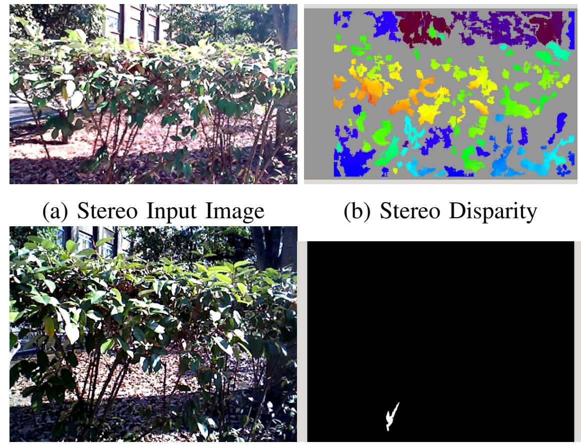
Fig. 5 shows the comparison of distance measurement under direct sunlight between stereo and IR projection camera. In this experiment illuminance of the sunny side is about 75000 [lx] and the shadow side is about 3500[lx]. Results of distance measurement are shown as colored region in Fig. 5b and Fig. 5d for each camera, note that the unmeasured region is colored as black. Compared with IR projection camera can not measure almost all of viewing, stereo camera can measure for a most part of viewing.

Fig. 6 also shows the comparison of measurement distance of dummy tomato about 50[cm] apart between stereo and IR projection camera. As shown in Fig. 6c and Fig. 6a, 2 cameras has almost same viewing, however, IR projection camera can not measure tomato fruits in center of the image as shown in Fig. 6d. On the other hand, as shown in Fig. 6b, stereo camera can measure distance to tomato fruits.

IV. TOMATO HARVESTING SOLUTION

A. TOMATO FRUIT DETECTION

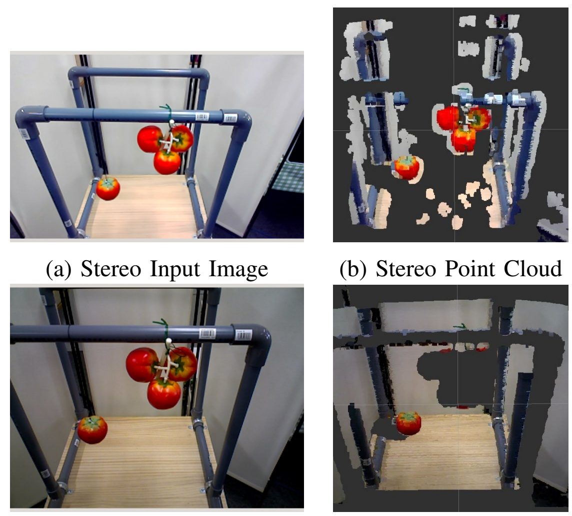
In this paper, we improve recognition method proposed in previous work [4]. Fig. 7 shows the algorithm. In tomato recognition, firstly color extraction is applied to 3D colored point cloud from stereo camera to detect tomato candidate point cloud which has similar color to tomato fruit. We use HSI color space in the color extraction and threshold of tomato candidates are; $-13 < H < 13$[°], $80 < S < 255$, and $10 < I < 255$. Secondly, euclidean distance clustering is applied to the candidate point cloud to determine tomato clusters, and Z-sorting is also applied to determine the order of each clusters. The nearest tomato cluster is decided as target cluster. Lastly, sphere fitting using RANSAC is applied to target cluster point cloud, while rest points after rejecting inliers of previous fitting are enough left, sphere fitting is applied to rest points iteratively to detect all tomato fruits in the cluster. In RANSAC phase, 10000 iterations are tried to fit sphere. Whole procedure of tomato fruit detection runs about 5[Hz], takes about 0.2[s].



(a) Stereo Input Image (b) Stereo Disparity

(c) IR Proj. Input Image (d) IR Proj. Depth Image

Fig. 5: Comparison Stereo and IR Projection 3D Camera In Outdoor



(a) Stereo Input Image (b) Stereo Point Cloud

(c) IR Proj. Input Image (d) IR Proj. Point Cloud

Fig. 6: Comparison Stereo and IR Projection 3D Camera, Target objects are in 50 cm front of camera.

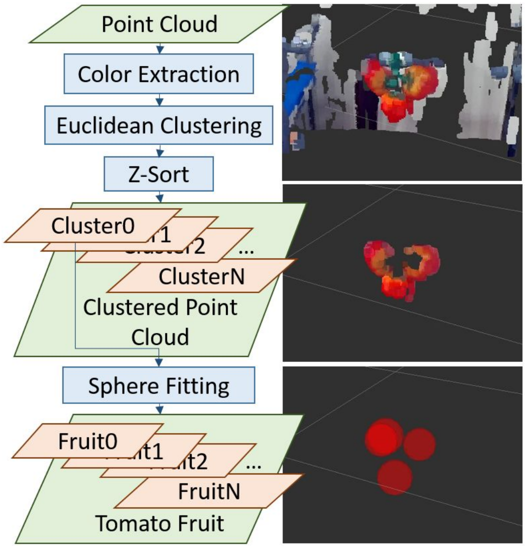


Fig. 7: Tomato Fruit Detection

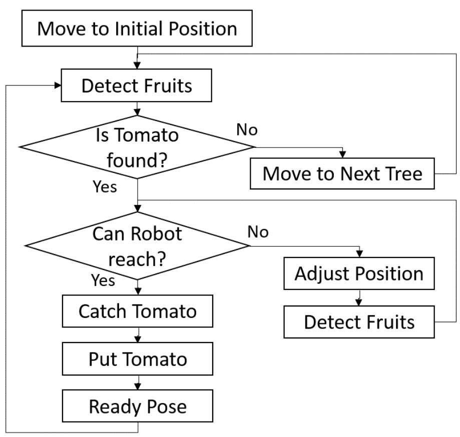


Fig. 8: Tomato Harvesting Algorithm

B. HARVESTING ALGORITHM

Tomato harvesting algorithm is shown in Fig. 8. First of all the robot moves to initial position, in case of the competition it means the front of the ridge, and in case of the real farm it means on a rail. Next tomato fruits are detected and harvest while tomatoes are detected. The robot position is calculated from the nearest tomato position, then arm pose is calculated using inverse kinematics algorithm of EusLISP. If the robot can not reach to the fruit, the robot adjusts its position. Note that adjustment distance is decided heuristically. When there are no tomato which can be harvested by the robot. Then the robot moves to the next tomato tree.

V. EXPERIMENTATION

A. TOMATO ROBOT COMPETITION

We evaluated the robot system in the 2nd tomato robot competition. In this competition, the robot have to run to harvesting spot, and there are 3 stages; Stage 1: approaching to one tomato fruit (harvesting is not necessary), Stage 2: harvesting from multiple tomato fruits clusters, Stage 3: harvesting from real tomato trees. Competition time limit is 10 minutes including moving in the field of each stage.

Fig. 9 shows images of the harvesting tomato fruit from tomato clusters hanged from frame in the competition stage 2. In this case the harvesting takes about 85[s] per fruit, about 12[s] is spent to move hand, about 30[s] is spent to rotate hand to pluck off fruit. and about 18[s] is spent to put tomato fruit into basket. Note that the recognition process takes about 0.2[s] and this is enough faster than robot motion. In this stage the robot tried to harvest 3 times and succeeded 2 times, however both succeeded 2 times the calyx of tomato fruit was broken. The failure case is misgrasping by moving hand into the incorrect position.

Fig. 10 also shows images the practice of stage 3. In this stage the robot tried to harvest 5 times and succeeded 3 times. The calyx of tomato is also broken for all succeeded cases. The one of failure cases is misgrasping and the other case is multiple grasping.

B. HARVESTING IN GREENHOUSE

We also evaluate the robot in the greenhouse. Fig. 11 shows images of harvesting. The robot move in the shallow corridor and under the direct sunlight in the greenhouse, against in the wide field and under the artificial illumination in the competition. As shown in images, the robot harvested tomato fruit successfully from clutter scene with tomato vines under the direct sunlight. Note that a calyx of fruit was not broken because tomato breed of the greenhouse has larger fruit and bolder stem than the competition, therefore a calyx is also stronger.

C. EVALUATION AND DISCUSSION OF HARVESTING RESULTS

In this section, we evaluate our system through results of harvesting in previous section, by considering about harvesting time. We also discuss about problems and future solutions by considering failure cases.

1) HARVESTING TIME: Harvesting time table (table II) indicates time ratio of whole sequence of robot motion. According to this, following 2 factors spent most part of harvesting; Hand rotation for plucking and Arm motion for each process. Especially in the competition stage 2, tomato clusters were fixed loosely to the frame by chain of plastic bands and it could rotate itself easily, so pluck-off harvesting spent many time. Comparing with this, in the real farm harvesting, about a half of time is spent to pluck off fruit. The difference of plucking time means the total revolution of hands to pluck pedicel off, depends on the tomato breed. We try to optimize motions of robot arm, hand open/close and rotation speed supposing to the tomato breed in the real farm, finally the robot can harvest 23[s] per fruit. On the other hand, harvesting rate of the proposed system was $5/8 = 62.2\%$. Note that this successful rate is inaccurate because there is not enough harvesting trials, however, according to analyzing failure cases, at least the system needs more information from recognition methods. So the harvesting time will increase for recognition to improve harvesting rate.

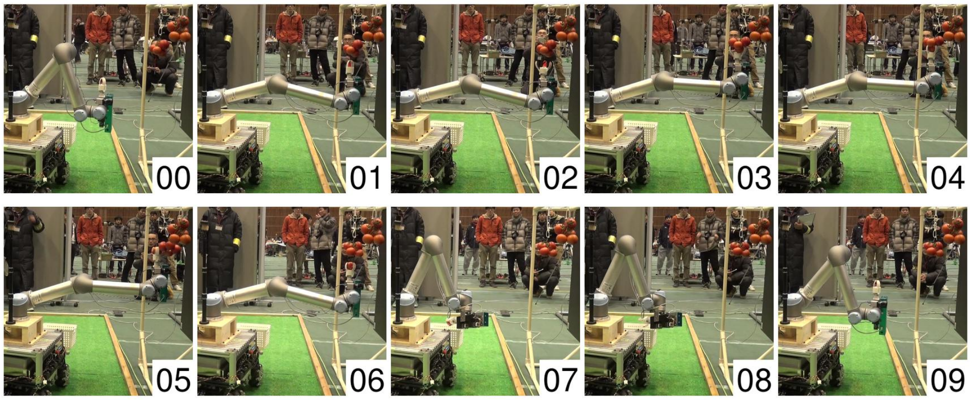


Fig. 9: Harvesting in Tomato Robot Competition Stage 2

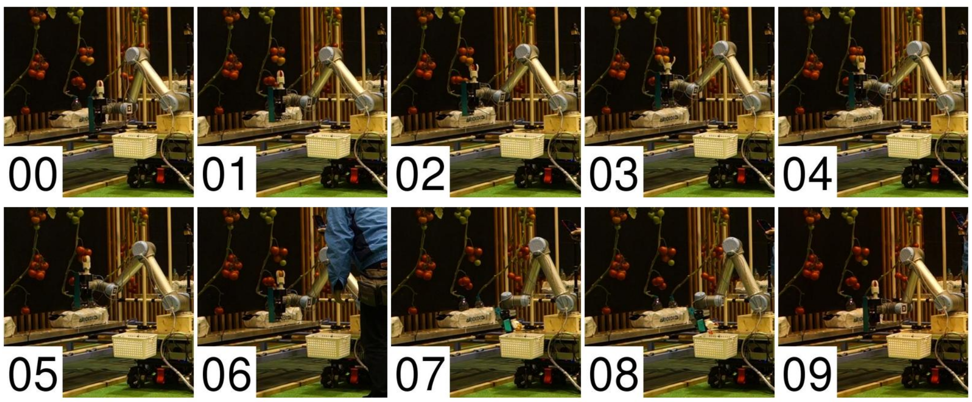


Fig. 10: Harvesting in Tomato Robot Competition Stage 3

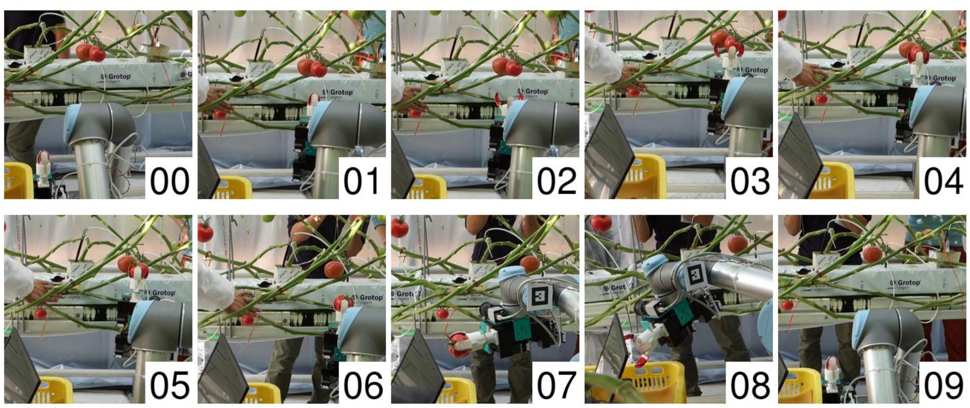


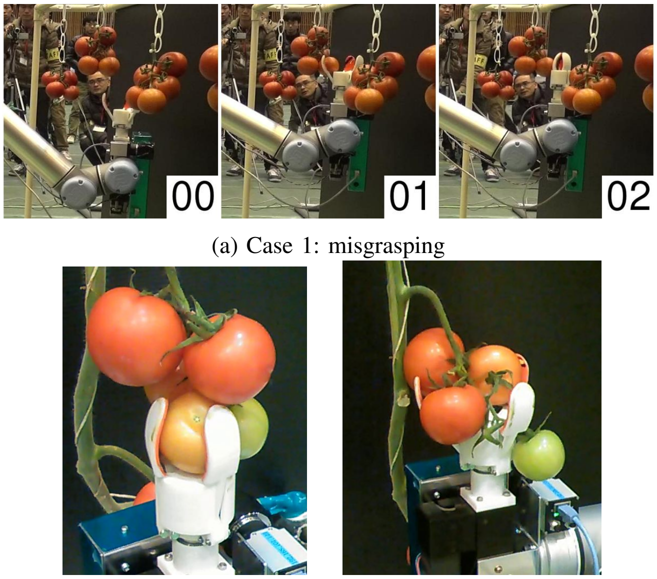
Fig. 11: Harvesting in Real Farm

TABLE II: Evaluation of Harvesting Time

Item	Comp. [s]	Real Farm [s]	Improved [s]
Fruit Recognition	3	3	3
Approaching	12	12	4
Putting	18	19	8
Hand Open/Close	20	20	4
Hand Rotation	31	15	4
Total	85	69	23

2) FAILURE CASES: Analyzing failure cases of harvesting, there are 3 type of failure; a) Tomato is not grasped into fingers, b) Tomato fruit is harvested successfully but calyx is broken, and c) Tomato is not harvested because the hand grasps multiple fruits.

In the first case a), the reason is also classified as following 2 subcases; the robot reached to a void space because the tomato fruit detector returns incorrect position, or the tomato fruits is moved outside of fingers by pushing another part of tomato tree or large error of recognition. In first subcase, it depends on the recognition algorithm, especially recognition will easily fails with cluttered tomato clusters, because the



(a) Case 1: misgrasping

(b) Case 2: breaking pedicel (c) Case 3: grasping multiple tomatoes

Fig. 12: Failure Case

recognition process depends on the area of tomato fluits in viewing and in cluttered fruits scene behind tomatoes may have small area in viewing. To avoid this, multiple view direction measurement is effective. In second subcase, grasp state check will detect misgrasp and cancel motion to shorten the retry time.

In the second case b), there is also 2 subcases; the gap of tomato stem and rotation axes of the hand is too large because approach angle is decided only from position of the fruit, or strength of the pedicel is depend to tomato breed itself. Especially in contrast, in the real farm calyx of all tomatoes were not broken, however in the competition calyx of all tomatoes were broken. Tomato stem recognition may avoid these situation or decrease break probability to decide approach angle to same as the stem.

In the third case c), the hand may grasps multiple fruits in case of dense tomato cluster. The reason of this situation seems to the hand may push another tomato in a cluster and the target fruit move to unexpected position. The solution of this is the grasp state estimation as same as case a), and tomato fruit tracking is also effective while grasping to detect failure.

VI. CONCLUSIONS

In this paper, we propose a design and development of an autonomous tomato harvesting robot. The robot is composed of a rotational plucking gripper, a 6DoF single arm, a stereo camera capable under the direct sunlight, and a omnidirectional mover. We also proposed an autonomous harvesting algorithm and evaluate through harvesting experimentation using real tomatoes in the tomato robot competition and the real farm. Finally the robot can harvest tomatoes about 23 [s] per fruit, however, its harvesting successful rate is roughly about 60%. In the real harvesting process, many kinds of disturbance are caused for each procedure, both recognition and harvesting motion. Especially, grasp state estimation to avoid misgrasp or multiple grasp and tomato recognition not only position, but how fruit grows from the tree are necessary towards more correct harvesting.

REFERENCES

[1] E. J. van Henten, J. Hemming, B. Van Tuijl, J. Kornet, J. Meuleman, J. Bontsema, and E. Van Os, "An autonomous robot for harvesting cucumbers in greenhouses," *Autonomous Robots*, vol. 13, no. 3, pp. 241–258, 2002.

[2] E. Van Henten, B. v. Van Tuijl, J. Hemming, J. Kornet, J. Bontsema, and E. Van Os, "Field test of an autonomous cucumber picking robot," *Biosystems engineering*, vol. 86, no. 3, pp. 305–313, 2003.

[3] S. Hayashi, K. Shigematsu, S. Yamamoto, K. Kobayashi, Y. Kohno, J. Kamata, and M. Kurita, "Evaluation of a strawberry-harvesting robot in a field test," *Biosystems Engineering*, vol. 105, no. 2, pp. 160 – 171, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1537511009002797

[4] X. Chen, C. K. Chand, Y. Tanaka, K. Nagahama, H. Yaguchi, K. Okada, and M. Inaba, "Reasoning-based vision recognition for agricultural humanoid robot toward tomato harvesting," in *Proceedings of The 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2015, pp. 6487–6494.

[5] S. Bachche, "Deliberation on design strategies of automatic harvesting systems: A survey," *Robotics*, vol. 4, no. 2, p. 194, 2015. [Online]. Available: http://www.mdpi.com/2218-6581/4/2/194

[6] C. Blanes, M. Mellado, C. Ortiz, and A. Valera, "Review. technologies for robot grippers in pick and place operations for fresh fruits and vegetables," *Spanish Journal of Agricultural Research*, vol. 9, no. 4, pp. 1130–1141, 2011. [Online]. Available: http://revistas.inia.es/index.php/sjar/article/view/2437

[7] M. Monta, N. Kondo, and K. Ting, "End-effectors for tomato harvesting robot," *Artificial Intelligence Review*, vol. 12, no. 1, pp. 11–25, Feb. 1998.

[8] S. Bachche and K. Oka, "Design, modeling and performance testing of end-effector for sweet pepper harvesting robot hand," *J. of Robotics and Mechatronics*, vol. 25, no. 4, pp. 705–717, 2013.

[9] S. BACHCHE and K. OKA, "Performance testing of thermal cutting systems for sweet pepper harvesting robot in greenhouse horticulture," *Journal of System Design and Dynamics*, vol. 7, no. 1, pp. 36–51, 2013.

[10] S. Bachche and K. Oka, "Distinction of green sweet peppers by using various color space models and computation of 3 dimensional location coordinates of recognized green sweet peppers based on parallel stereovision system," *Journal of System Design and Dynamics*, vol. 7, no. 2, pp. 178–196, 2013.

[11] N. KONDO, K. YAMAMOTO, H. SHIMIZU, K. YATA, M. KURITA, T. SHIIGI, M. MONTA, and T. NISHIZU, "A machine vision system for tomato cluster harvesting robot," vol. 2, no. 2. Asian Agricultural and Biological Engineering Association, 2009, pp. 60–65. [Online]. Available: http://ci.nii.ac.jp/naid/130000108286/

[12] H. Jiang, Y. Peng, and Y. Ying, "Measurement of 3-d locations of ripe tomato by binocular stereo vision for tomato harvesting," in *2008 Providence, Rhode Island, June 29–July 2, 2008*. American Society of Agricultural and Biological Engineers, 2008, p. 1.

[13] G. Feng, C. Qixin, N. Masateru, *et al.*, "Fruit detachment and classification method for strawberry harvesting robot," *International Journal of Advanced Robotic Systems*, vol. 5, no. 1, pp. 41–48, 2008.

[14] UNIVERSAL ROBOTS, "Ur5 robot," http://www.universal-robots.com/products/ur5-robot/.

[15] PS4 Developer wiki, "Playstation 4 camera," http://www.ps3devwiki.com/ps4/PlayStation_4_Camera.

[16] Willow Garage, "ROS.org," http://www.ros.org/.

[17] T. Matsui and M. Inaba, "Euslisp: An object-based implementation of lisp," *J. Inf. Process.*, vol. 13, no. 3, pp. 327–338, July 1991. [Online]. Available: http://dl.acm.org/citation.cfm?id=110766.110773

[18] ASUS, "Xtion pro live," https://www.asus.com/3D-Sensor/Xtion_PRO_LIVE/specifications/.