CUCUMBER DETECTION FOR PRECISION AGRICULTURE APPLICATIONS*

ROEMI FERNÁNDEZ†, VICTOR ALIENDE, PABLO GONZALEZ-DE-SANTOS, MANUEL ARMADA

Centre for Automation and Robotics CAR CSIC-UPM
Ctra.Campo Real, km 0.2, La Poveda, Arganda del Rey 28500, Madrid, Spain

JELENA SURDILOVIC

Leibniz-Institut für Agrartechnik und Bioökonomie e.V. ATB
Max-Eyth-Alle 100, D-14469, Potsdam, Germany

DRAGOLJUB SURDILOVIC

Fraunhofer-Institut für Produktionsanlagen und Konstruktionstechnik IPK
Palcalstraße 8-9, 10587 Berlin, Germany

The objective of this research was to explore the feasibility of detecting cucumber fruits in field conditions for autonomous robotic harvesting applications. A high resolution colour camera and a time-of-flight camera are proposed as primary sensors for the design of the sensory system. The preliminary detection algorithm includes a pixel-based classifier that labels areas of interest that belong to cucumber fruits and a registration procedure that combines the results of the aforementioned classifier with the range data provided by the time-of-flight camera. The detection algorithm is extremely simple and efficient, and provides a satisfactory discrimination of the cucumbers fruits with respect to the rest of the

---

* This work was supported in part by the FP7 Project ECHORD++ (European Clearing House for Open Robotics Development, Grant Agreement Number 601116) – Experiment Call 2 – CATCH – Cucumber Gathering – Green Field Experiments, by the Spanish Ministry of Economy, Industry and Competitiveness under grant DPI2017-84253-C2-1-R, by the CSIC under the project Robótica y Sensores para los Retos Sociales (ROBSEN-PIE 20165E050) and by the ROBOCITY2030-III-CM Project (Robótica Aplicada a la Mejora de la Calidad de Vida de los Ciudadanos. Fase III; S2013/MIT-2748), funded by Programa de Actividades de I + D en la Comunidad de Madrid and cofounded by Structural Funds of the EU.

† Work partially supported by the Spanish Ministry of Economy, Industry and Competitiveness under the Ramón y Cajal Programme.

H. Montes et al. (eds.), *CLAWAR 2018: 21st International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines*,
Robotics Transforming the Future

elements of the scene. Several experimental tests have been carried out in outdoor conditions in order to evaluate and demonstrate the capabilities of the proposed approach.

1. Introduction

Over the past two decades, several studies have been conducted to provide automatic detection and measurement systems for different features of cucumbers crops. For instance, van Henten et al. [1] proposed a combination of two CCD cameras, one equipped with an 850 nm filter and the other with a 970 nm filter for detecting cucumber fruits grown in greenhouses using a high-wire cultivation system, in which every plant is attached to a wire. Experimental tests showed that the designed system was able to detect more than 95% of the fruits in a dataset of 106 cucumbers. A segmentation algorithm based on rough set theory was presented in [2] for solving the problem of cucumber identification in greenhouses. However, no quantitative results derived from the proposed algorithm were presented. In [3], a three-layer back-propagation neural network combined with a texture analysis was established to detect greenhouse cucumber fruits. The algorithm was tested on 40 cucumber plant images, and a detection rate of approximately 76% was achieved considering only backlighting conditions. A multi-template matching method was utilized by Bao et al. [4] to recognize matured Radit cucumbers grown vertically in a greenhouse. Proportional scaling and rotation operations were applied to a standard cucumber image to build a multi-template matching library, whereas the multi-template matching method was developed by using the normalized correlation coefficients algorithm.

This paper presents the first steps carried out for the design, implementation and validation of an efficient, adaptable and robust sensing system for the detection of cucumbers fruits in field conditions. The proposed solution differs from all previous approaches found in the literature, which mainly focus on the detection of cucumbers grown in greenhouses by using wire cultivation systems.

The manuscript is organized as follows. Section 2 describes the sensory rig that has been designed and manufactured for the acquisition of images. Section 3 presents the algorithm proposed for the detection of cucumbers fruits in outdoor conditions. In Section 4, results of this study are discussed. Finally, major conclusions and lines of future extensions are summarised in Section 5.

2. Description of the sensory system

The proposed multisensory system consists of an AVT Prosilica GC2450 high resolution CCD colour camera and a Mesa SwissRanger SR-400011 TOF (Time-Of-Flight) 3D camera. The 5-megapixel GC2450 has a frame rate of up to 15 fps at 2448 × 2050 pixels resolution. Meanwhile, the TOF camera provides a depth map and an amplitude image at the resolution of 176 × 144 pixels with 16 bit floating-point precision and a maximum frame rate of 54fps, as well as *x*, *y* and *z* coordinates to each pixel in the depth map. The detection range (radial distances) of this device goes from 0.1 m to 5.0 m, and its field of view is 69° (h) × 56° (v). The proposed multisensory system is shown in Figure 1.

Intrinsic and extrinsic calibration parameters of both cameras (RGB and TOF) were estimated by using the Matlab camera calibration toolbox [5]. A distance measurement calibration was also carried out in Matlab (http://www.mathworks.com/products/matlab/) for the TOF camera by following the method proposed in [6].

Figure 1. Proposed multisensory system.

3. Detection algorithm

Taking into consideration the configuration described in the previous Section, two complementary procedures are proposed for the detection algorithm: a pixel-based classification procedure that labels areas of interest that are candidates for belonging to cucumber fruits and a registration procedure that combines the results of the aforementioned classifier with the data provided by the TOF camera for the 3D reconstruction of the desired regions. These procedures are described below.

The first procedure, based on Support Vector Machines (SVMs), is capable of labelling each pixel of the image into three classes that are: cucumber, leaves and background. SVM is a supervised learning method utilized for classifying set of samples into two disjoint classes, which are separated by a hyperplane defined on the basis of the information contained in a training set [7-10]. In the case at hand, three SVMs are utilized sequentially, each one for detecting a class against the rest. Therefore, after the first SVM is applied, pixels identified as belonging to cucumber class are labelled and a mask is generated in such a way that only the remaining pixels are considered for the following SVMs. This step is then repeated for the leaves class and the background class. The SVM classifiers are trained by selecting a random subset of samples from the RGB images and manually labelling the regions of interest from these images into the three semantic classes mentioned above.

Once regions of interest have been detected in the scene, it is necessary to locate them spatially. The TOF camera included in the proposed multisensory system provides amplitude, depth and confidence data simultaneously for each pixel of the captured image. The amplitude represents the greyscale information, the depth is the distance value calculated within the camera and the confidence is the strength of the reflected signal, which means the quality of the depth measurements. Although TOF data is fundamental for localisation purposes, it is still necessary to automatically match this information with the classification map obtained from the previous step in a common reference frame [11]. For accomplishing this procedure it should be taken into account that TOF images and resulting classification maps come from sensors that exhibit different field of view and different pixel array size. Thus, data will only depict the same content partially, and the pixel correspondence will not be direct. To overcome this problem, the random sample consensus (RANSAC) algorithm is adopted for the multisensory registration [11, 12]. RANSAC is one of the most robust algorithms for model fitting to data containing a significant percentage of errors [13]. This iterative method estimates parameters of a mathematical model from a set of observed data which contains outliers. As the multisensory system has been designed in an enclosure that prevents the relative movements between the different elements that compose it, the idea is to use the RANSAC method to find the rotation and translation (R, T) that enable the transformation of the TOF data into the reference frame of the classification map. In this way, the transformation given by (R, T) may be applied to any image acquired with the TOF camera, obtaining quickly and efficiently the registered data and it won’t be necessary to recalculate this transformation as long as the multisensory rig is not modified.

4. Experimental results

In order to validate the proposed approach, several experimental campaigns for acquiring training and ground truth data were carried out in the summer of 2017.

The experimental field was planted with Liszt cucumbers from the seed producer Ryk Zwaan GmBH.

Figure 2(a, b) shows the RGB image for one scene, together with the classification map resulting after applying the first part of the detection algorithm. In the classification map, yellow, green and white colours are utilised to visualize pixels classified as cucumbers, leaves and background, respectively. In addition, the blobs extracted from the classification map are displayed on Figure 3. For each blob, the centroid and the area are calculated.

(a) (b)

Figure 2. (a) Original RGB image. (b) Resulting classification map.

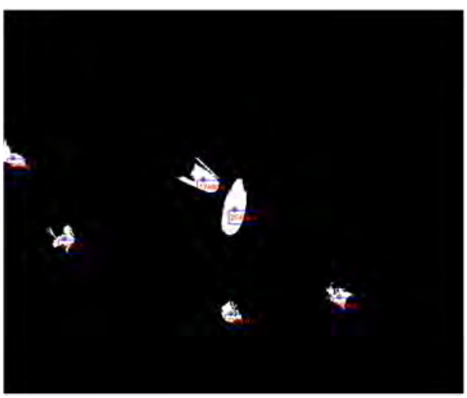


Figure 3. Resulting blobs provided by the first part of the detection algorithm.

Computation of the registration procedure required a significant collection of ground truth correspondences (about 300 pairs of points), trying to cover the entire field of view of the cameras [11, 12]. Registration procedure was tested indoors and outdoors, in order to evaluate possible interferences due to natural light (direct sun light), as well as the response to objects exhibiting different geometries. For this purpose, several landmarks were located in a scene with known objects, in order to measure the deviations. Table 1 summarizes not only the minimum and maximum errors but also the mean errors in each axis for indoor conditions. Table 2 gathers the obtained results for outdoor conditions. Figure 4 shows two close-up views of a registered region of interest extracted from the scene displayed on Figure 2(a).

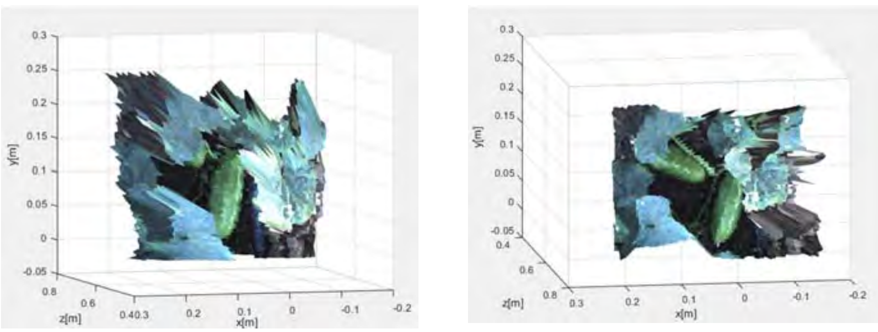


Figure 4. Close-up views of a registered region of interest extracted from the scene displayed on Figure 2(a).

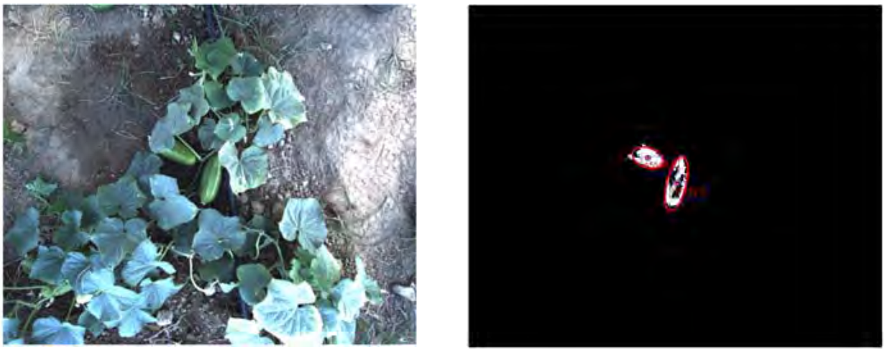
Table 1. Position errors from the TOF registered data in indoor conditions.

Axis	Minimum error [cm]	Maximum error [cm]	Mean absolute error [cm]
X	0.2	1.3	0.8
Y	0.3	1.6	1.1
Z	0.1	2.2	1.0

Table 2. Position errors from the TOF registered data in outdoor conditions.

Axis	Minimum error [cm]	Maximum error [cm]	Mean absolute error [cm]
X	0.5	1.9	1.4
Y	1.0	2.3	1.7
Z	0.4	3.0	1.5

Assuming that cucumbers are always lying on the ground, it is possible to use resulting registered data for improving detection algorithms. The idea is to discard all the blobs that are located above a predefined height (false positive detections). Figure 5 shows results of this procedure for the scene displayed previously on Figures 1-3. In this way, only blobs that correspond with the cucumbers are preserved.



(a) (b)

Figure 5. (a) Original RGB image. (b) Detection result after discarding blobs located above a predefined height.

5. Conclusions

This paper proposes a multisensory approach for the detection of cucumbers in natural scenarios. The solution includes a high resolution colour camera, a TOF camera, and a detection algorithm composed of a pixel classifier based on SVMs and a registration procedure that combines the resulting classification map with the data provided by the TOF camera, in such a way that the range data can be associated to the image pixels identified as cucumbers.

Preliminary experimental results demonstrate the feasibility of the proposed approach and highlight the advantages of the solution. The simplicity of the sensory rig in comparison with other multispectral approaches decreases the total cost of the system and makes the future harvesting robot in which it is incorporated more competitive in the market.

Future work will be directed to reduce the detection and localization errors in order to achieve a more robust application.

References

1. van Henten, E.J., et al., *An Autonomous Robot for Harvesting Cucumbers in Greenhouses.* Autonomous Robots, 2002. **13**(3): p. 241-258.
2. Qing-Hua, Y., et al., *Cucumber image segmentation algorithm based on rough set theory.* New Zealand Journal of Agricultural Research, 2007. **50**(5): p. 989-996.
3. Zhang, L., et al., *Recognition of greenhouse cucumber fruit using computer vision.* New Zealand Journal of Agricultural Research, 2007. **50**(5): p. 1293-1298.
4. Bao, G., et al., *Multi-template matching algorithm for cucumber recognition in natural environment.* Computers and Electronics in Agriculture, 2016. **127**: p. 754-762.
5. Bouguet, J.-Y., *Camera calibration toolbox for MATLAB.* 2010.
6. Chiabrando, F., et al., *Sensors for 3D Imaging: Metric Evaluation and Calibration of a CCD/CMOS Time-of-Flight Camera.* Sensors, 2009. **9**(12): p. 10080.
7. Fernández, R., H. Montes, and C. Salinas, *VIS-NIR, SWIR and LWIR Imagery for Estimation of Ground Bearing Capacity.* Sensors, 2015. **15**(6): p. 13994.
8. Fernández, R., et al., *Combination of RGB and Multispectral Imagery for Discrimination of Cabernet Sauvignon Grapevine Elements.* Sensors, 2013. **13**(6): p. 7838.
9. Fernández, R., et al., *Multisensory System for Fruit Harvesting Robots. Experimental Testing in Natural Scenarios and with Different Kinds of Crops.* Sensors, 2014. **14**(12): p. 23885.
10. Mucherino, A., P.J. Papajorgji, and P.M. Pardalos, *Data Mining in Agriculture*. 2009: Springer New York.
11. Salinas, C., et al., *A New Approach for Combining Time-of-Flight and RGB Cameras Based on Depth-Dependent Planar Projective Transformations.* Sensors, 2015. **15**(9): p. 24615.
12. Salinas, C., *A non-feature based method for automatic image registration relying on depth-dependent planar projective transformations. Ph.D. Thesis,* in *Departamento de Arquitectura de Computadores y Automática*. 2015, Universidad Complutense de Madrid.
13. Fischler, M.A. and R.C. Bolles, *Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography.* Commun. ACM, 1981. **24**(6): p. 381-395.