Project title:	Development and demonstration of an automated selective broccoli harvester
Project number:	0002829
Project leader:	Prof Tom Duckett, University of Lincoln
Report:	Annual report, 2018
Previous report:	
Key staff:	Prof Tom Duckett
	Dr Grzegorz Cielniak
Location of project:	Lincoln, UK
Industry Representative:	Prof Simon Pearson, Lincoln Institute for Agri-Tech (LIAT), University of Lincoln
Date project commenced:	19th October 2017

DISCLAIMER

While the Agriculture and Horticulture Development Board seeks to ensure that the information contained within this document is accurate at the time of printing, no warranty is given in respect thereof and, to the maximum extent permitted by law the Agriculture and Horticulture Development Board accepts no liability for loss, damage or injury howsoever caused (including that caused by negligence) or suffered directly or indirectly in relation to information and opinions contained in or omitted from this document.

AUTHENTICATION

We declare that this work was done under our supervision according to the procedures described herein and that the report represents a true and accurate record of the results obtained.

[Name]

[Position]

[Organisation]

Signature .... Date ....

[Name]

[Position]

[Organisation]

Signature .... Date ....

Report authorised by:

[Name]

[Position]

[Organisation]

Signature .... Date ....

[Name]

[Position]

[Organisation]

Signature .... Date ....

CONTENTS

GROWER SUMMARY

Headline

The development of a fully automated harvesting system is crucial for broccoli growers who currently harvest manually and face increasing labour scarcity as well as the full impact of the National Living Wage by 2021. An automated selective harvester would address the urgent need to reduce labour and improve the growers' ability to control production costs. We seek to improve and to develop algorithms for the perception module of the proposed harvester.

Background

This project will incorporate learnings from work previously carried out by the Lincoln Centre for Autonomous Systems (L-CAS) at the University of Lincoln to accurately identify broccoli plants in the field, as well as measure the size of each plant head in order to determine whether or not it is suitable for cutting. A more robust solution to this problem will be achieved by integrating this previous work based on low-cost 3-D structured light cameras. The goal is to ensure that the automated selective harvester could be built using a competitively priced imaging system able to deliver the required levels of accuracy, reliability and scalability.

Summary

In the UK, broccoli is currently harvested by hand, usually by teams of 7, relying on the ability of workers to visually assess the size of each broccoli head against a predefined specification to estimate whether it can be cut. The ability of field workers to do this will vary and their effectiveness inevitably deteriorates over a working shift. Furthermore, the methods used to breed broccoli have caused the heads to grow at different rates. This makes them difficult to harvest. As a result, nearly 50% of the crop can be harvested economically. Similar challenges and difficulties have also been reported by growers in other countries in Europe and in America. In addition, between 2016 and 2021, the cost of manual workers in the UK will have increased by 35% due to increases in the National Living Wage. A variety of other factors ranging from political pressures to migration dynamics have made suitable workers to become increasingly scarce.

Financial Benefits

The current manual harvesting process is labour intensive, expensive and it leads to high levels of waste. The direct benefit to growers of an automated harvester is the significant reduction in increasingly expensive and hard to find field workers. There will be also significant improvements to the growers' production, as an automated selective harvester has the potential to produce larger quantities of marketable product. Additionally, cutting only the heads that meet the required size criteria will also reduce processing and packaging costs.

Action Points

Currently the harvesting process is totally dependent on the ability of the field workers to accurately assess the size of each head of broccoli against a predefined specification before making the decision to cut it or leave it in the field. Growers must be open to adapt and modify their current harvesting system in order to automate the process and potentially adapt their harvesting methods when the automated harvester evolves towards other brassica and leaf crops. An automated harvesting system is expected to be able to work in very low light conditions. This presents a valuable opportunity to growers as harvesting at night when ambient temperatures are lower improves both shelf life and product quality and could also reduce refrigeration costs. Currently the high cost of labour prevents growers to harvest either at night or round the clock.

SCIENCE SECTION

Introduction

This PhD research is concerned with robotic perception using state-of-the-art 3D sensors for real-time detection and size estimation of crops; it specially focuses on perception for robotic harvesting of *broccoli crops*.

Broccoli is a vegetable in the cabbage family that belongs to the Brassica Oleracea plant species. The interest in its cultivation has grown in recent years due to genetic improvement programmes developed in several countries, and to the healthy compounds contained in the crop that have increased its consumption [28]. The worldwide production of this vegetable family reached 25.23 million tonnes in 2016 [11].

A consequence of the methods used to breed broccoli is that the heads grow at different rates [28]. This makes them difficult to harvest. Broccoli heads are harvested when they reach maturity, i.e. when the clusters of buds are tight and of dark-green colour. Almost all broccoli is currently harvested by hand, relying on visual grading of size to estimate whether a head can be cut [34]. As a result, only around 50% of broccoli heads can be harvested economically. The remaining crops either mature too quickly and have to be left in the field, or they simply do not reach full maturity on time during the harvesting season.

Two approaches can be readily compared when considering traditional and/or automated harvesting, namely, slaughter harvesting, i.e. cutting everything in one pass, and selective harvesting, i.e. cutting individually each crop [10]. Slaughter harvesting is not a productive option as it potentially produces large quantities of unmarketable broccoli heads, whereas selective harvesting presents its own challenges as it relies on a subjective assessment by each person cutting the broccoli as to which head is ready. Additionally, labour has become increasingly scarce and more expensive due to a variety of factors ranging from political pressures to migration dynamics [3].

The goal of growing fresh fruit and vegetables is to keep the quality high while minimising costs. The integration of autonomous systems technologies into agriculture operations can have a notable impact in crop production. It is therefore desirable to find a method to harvest more frequently, more quickly, more accurately, with less waste, and that reduces labour and overall operation costs [2, 3]. Thus, developing an automated method for selective harvesting would help to increase productivity and to better control production costs. This project will develop 3D imaging methods to accurately identify broccoli plants in the field, establish the precise location of each broccoli head selected for cutting, and accurately measure the size of each plant head and compare it against pre-agreed criteria in order to establish whether or not it is suitable for cutting. The developed methods will be validated on multiple real world datasets and can be integrated into a prototype robotic system for automatic selective harvesting, which can work for long periods to cut and collect broccoli heads of the preferred size.

The project builds on previous work where a set of experiments were conducted with a 3D vision system pipeline for detection, classification and tracking of mature broccoli heads [25, 26]. The proposed research will develop this framework further to realise a multi-purpose 3D perception system which will be able to robustly detect and classify the target crops in real time, resulting in 3D maps that can be used by a robotic system for physically harvesting the crop. Additionally, the project will develop complementary algorithms for carrying out quality measurements including checking whether the crop is both ready for harvest and that it meets customer specifications for size and shape.

Research objectives

This PhD thesis aims to develop and verify a framework for detection, classification, and size estimation of broccoli heads using 3D point cloud data. It also aims to demonstrate the suitability of the approach for integration into a robotic broccoli harvester in the field. The research will be guided by the following research objectives:

1. Develop and evaluate new approaches for accurately clustering 3D point clouds that are both size and point density invariant. The developed approaches will take into account the inherent features of broccoli crops such as shape or texture.
2. Develop and evaluate new approaches for detection of broccoli heads beyond the current state-of-the-art, e.g., using deep learning. These approaches will be evaluated and compared to the developed segmentation methods.
3. Develop and test an approach for 3D size estimation of broccoli heads robust to occlusions and partial data to determine if a head is ready to be harvested.
4. Evaluate the performance of the developed algorithms using a variety of real-world datasets, including comparisons for different varieties of broccoli. The goal is to assess the generalisation of the developed methods, i.e., the ability of the algorithms to be effective across a range of input datasets.
5. Develop an optimised version of the system suitable for deployment in a robotic broccoli harvester.

Progress to date

The initial PhD research direction has focused on the optimisation and evaluation of a previous study by Kusumam *et al.* [26]. The progress made so far in this PhD project includes the following:

- The entire 3D system pipeline used in the experiments reported by Kusumam *et al.* has been re-implemented in a single application written in C++ to replicate and improve their results. To this end, the filtering step that removed outliers (Statistical Outlier Removal) was cancelled as it was computationally expensive and contributed poorly to the final result. The updated pipeline layout can be seen in Appendix A, Figure 2 on page 12. Also, a Support Vector Machines (SVM) classification module based on the PCL[1] implementation of the LIBSVM library[2] was added to the system.
- A close review of the available ground truth data showed that about 10% were incorrectly labelled, and that even annotating a fraction of existing or new datasets was extremely time consuming. To solve this, an annotation tool for 3D point cloud datasets has been implemented based on PCL. This allowed to correct, in a much shorter time, all ground truth data from previous experiments and to annotate new data.
- The SVM classifier was fine-tuned using a different kernel function, various kernel parameters, and a 75-25% proportion of the datasets for training and testing. Consequently, the classification performance based on the average precision score at various discrimination threshold settings improved by around 5% for the UK and Spanish datasets available. More details of these results are listed in Appendix A.
- The average run time of the entire pipeline per image frame is now 1/s on an Intel i7 CPU, 2.2 GHz. The previous result was an average of 5-6/s on an Intel i7 CPU, 3.4 GHz.
- The system pipeline has also been tested with new plant varieties datasets collected from fields in California, USA. The results showed that the current method does not account properly for larger distances from the sensor to the crops (four times larger in the California dataset). The segmented clusters were on average a third smaller, and with higher point density than those from other datasets. Appendix A also includes

[1] http://www.pointclouds.org
[2] https://www.csie.ntu.edu.tw/~cjlin/libsvm

more details of the classification results using this dataset. A more robust approach for segmenting point clouds is currently being investigated that is both size and point density invariant [30, 9, 16]. In addition, the following drawbacks of previous work were identified:

  - About 2% of broccoli heads in the datasets were not segmented by the clustering extraction algorithm and were not even considered in the performance evaluation.
  - About 3% of broccoli heads were segmented as part of much larger clusters that included parts of other elements in the point cloud. As a result, they were classified as a negative sample, i.e. not broccoli.
  - The time taken by the entire system pipeline is still too long (one frame per second) for its integration into a robotic broccoli harvester in the field.

- Lastly, in collaboration with Earth Rover[3], a company focused on robotic applications for agriculture, a first large dataset of broccoli crops in the field was recorded at Pollybell Farms[4]. The data was collected using an Intel RealSense Depth Camera D435[5] mounted on a tractor. The tractor moved several times over rows of planted broccoli ready to be harvested. This new dataset will further contribute to evaluate the performance of the developed algorithms on different varieties of broccoli, and on a variety of datasets.

Literature review

Harvesting robots usually consist of three independent systems: a recognition system to identify and locate the product, a picking system to perform grasping and cutting operations, and a navigation system to allow the robot to move around the cultivated crop plants [3]. This section is concerned with the first system and reviews several published approaches for detection, recognition, localisation that have been part of harvesting systems of various crops.

Two major challenges in autonomous harvesting are the recognition of the crop and its detachment from the plant. Some resistant crops such as olives and almonds can be harvested using a branch shaker device [13]. However, soft and delicate fruits cannot be harvested using this method, and even produce that may undergo some degree of bruising

[3] https://earthrover.cc
[4] https://www.pollybell.co.uk
[5] https://realsense.intel.com/depth-camera

may not be suited for the fresh market [7]. One of the first and common approaches has been to detect crops using 2D images. This can be promptly perceived in the wealth of techniques based on computer vision available in the literature [21, 3, 53].

Based on the number of research publications, Bechar *et al.* [6] found that some crops have been more intensively investigated than others. One of the largest research efforts has been invested in apple orchards [31]. Tabb *et al.* [43] developed a method for locating apples based on a background modelling process that mixes the Gaussian distributions of pixels in several image sequences. The algorithm correctly identified around 90% of both red and yellow apples at a rate of 14-16 frames per second. Similarly, Wachs *et al.* [48] discussed a vision system that uses both infra-red and colour images to detect apples. The system combines two independent subsystems; one based only on standard image processing operations, and the other based on a Haar wavelets detector. A voting scheme is then used to combine both systems and get the final detection results. Ji *et al.* [20] proposed a method for detecting apples using colour images acquired by a CCD camera. The method uses a segmentation algorithm based on region growing and colour feature extraction. Then a support vector machine classifier determines whether the segmented region is an apple or not. The system was tested on a harvesting robot under natural conditions and showed a detection success rate of 89%.

The identification of mature citrus fruits has also been extensively investigated in autonomous harvesting [27]. Hannan *et al.* [18] introduced an algorithm for the recognition of orange fruits. Firstly, a colour image of the crops was segmented based on a global thresholding approach that uses the proportion of red and green colours in the image. Then a perimeter circle detection procedure locates the oranges at different distances even when they are occluded. Results showed a success rate of 93% and a false detection rate of 4% in images with different amounts of illumination. Okamoto *et al.* [33] developed a vision system for detecting green fruit on citrus trees. The method extracted the background from each image and then used a template matching process to find circular objects on an edge image. A success rate of 91% was achieved on complete fruits, while 80% of occluded citrus fruits were correctly detected. Similarly, Kurtulmus *et al.* [24] presented an algorithm to detect immature green citrus fruits using colour images. The algorithm works by scanning the citrus tree image with a square region of three different sizes. Each region was tested by three different reference templates, two based on the well-known *eigenface* approach dubbed here *eigenfruit*: an eigenfruit intensity-based component, an eigenfruit saturation-based component, and third template based on circular Gabor texture analysis. A majority voting approach was then used for merging the results and to determine the region that is a fruit. Despite being heavily occluded by leaves of similar colour, 75% of the fruits were successfully detected. A more robust approach for detecting multiple fruits (oranges, grapes, and lychees -a red oval berry-like fruit-) of different colour under sunny and cloudy days was introduced by Wang *et al.* [49]. The method involved the application of a two-dimensional discrete wavelet transform to normalise the illumination, an image enhancement algorithm to highlight the fruit objects, and a segmentation process based on k-means clustering. Even on cloudy days, the average success rates for the three fruits were 92%, 94% and 87%, respectively.

Another crop that has seen extensive research interest is tomato, as a considerable amount is produced in greenhouses under more controlled conditions than those found in conventional large-scale farms [53]. Kondo *et al.* [23] proposed a vision system for greenhouse tomatoes. It consisted of two cameras and four lighting lamps to acquire images from three viewing angles during the night to avoid the influence of sun light. After acquisition, the RGB images were converted into HSI images, then the image was binarised and a polygonal image was created based on a H-S plane analysis classification. To detect the region of a tomato fruit, the polygonal image was examined with templates containing physical properties (lengths, diameters, and angles) of stems, peduncles, and tomatoes. Success identification rates when these three elements were cleanly isolated from other plant parts was 100%. However, success rate of images with occluded tomato clusters, stem, or peduncle was 65%. This was relevant when the system was integrated and tested on a robotic harvester [22]. Zhao *et al.* [54] described a tomato recognition algorithm based on multiple features merged at pixel level using a wavelet transformation. These features were the a* and I components extracted from the L*a*b* and YIQ colour spaces, respectively. The final segmentation result was processed by an adaptive threshold algorithm, and a set of morphology operations to reduce noise that segmented the target tomato from the background. The detection tests showed that 93% of tomatoes were successfully recognised by the system. The system was tested on a robotic platform that included a dual-arm frame of 3-DoF manipulators [55].

An assortment of other vegetables and fruits have as well been extensively investigated for autonomous harvesting. These include sweet peppers [4, 1, 47], cucumbers [35, 46], strawberries [12, 36, 52], melon and watermelon [29], and eggplants [19, 17].

Using only computer vision techniques may be insufficient to estimate the crop location. Therefore, many approaches have used 3D sensors to integrate depth perception. Silwal *et al.* [42] developed an image processing method consisting of circular Hough transformation and blob analysis to iteratively identify visible and partially visible apples in colour images. The images were acquired by a single CCD camera and depth information was added by using a 3D camera. Experimental testing in clustered and complex tree canopy structures reached a detection accuracy greater than 90%. Additionally, the system was integrated into a robotic apple harvester and tested in a commercial apple orchard [41]. Barnea *et al.* [5] combined RGB and range data to analyse shape features of objects captured in both a 2D image and a 3D space. In particular, they combined 3D surface normal features, 3D plane-reflective symmetry, and image plane highlights from elliptic surface points to detect sweet peppers, regardless of their colour. Xiang *et al.* [51] presented a method for recognising clustered tomatoes using stereo images acquired thorough two CCD cameras. Clusters were divided into two types based on the depth difference between the front and back regions using Otsu's method, a technique to perform clustering-based image thresholding. The method fused depth map segmentation and edge curvature analysis for overlapping tomatoes, whereas edge curvature analysis sufficed for non-overlapping fruits. Many other methods have also used a similar approach of combining RGB and depth images to recognise and locate different crops [44, 14, 40, 32, 38]. Similarly, numerous research efforts have also been reported using spectral and thermal cameras [15], as well as laser range finder sensors [21, 50] for detecting and localising of a variety of crops.

For the particular case of broccoli, some approaches have used RGB images to separate the broccoli head from the soil and other plant parts. Ramirez [37] developed an algorithm to locate the broccoli head within an image of an entire broccoli plant. To locate the head, first the method finds the leaf stems using a threshold, a Canny edge detector, and a Hough transform to extract geometric features that approximate lines that can be fit to the stems. Then the broccoli head can be located based on contrast texture analysis at the intersection of the stems. The method also determined the maturity of the crop using statistical texture analysis. Tu *et al.* [45] published results of a method to grade broccoli heads. The goal was to assess the quality decay of the harvested crop based on a set of colour and shape parameters. The system determined the area and roundness as the shape parameters and extracted the colour features using standard vision techniques. The resultant quality of the broccoli head was then decided by a neural network classifier. More recently, Blok *et al.* [8] presented a method for detecting and sizing broccoli heads based on computer vision techniques. The method segmented an image based on texture and colour of the broccoli head buds. Firstly, the contrast of the image was enhanced to emphasise high frequency areas, followed by a series of filters and several morphological operations to fine-tune the image. Then, pixel connectivity was used to generate connected green-coloured components. Lastly, a shape-based feature selection on the connected area was conducted to separate small non-connecting components from the foreground. The segmented heads were sized using circle templates, and the mean image processing time took a little less than 300 ms. The system was part of a prototype harvesting device attached to a modified tractor and was tested in cultivated broccoli fields reaching an accuracy of 94%.

PhD timeline			Months																		
Name	**Start**	**End**	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36
Reimplement pipeline	Oct-17	Aug-18																		
3D point cloud segmentation	Sep-18	Dec-18																		
Deep learning	Jan-19	Apr-19																		
Crop ready for harvesting	May-19	Oct-19																		
Test final version	Nov-19	Apr-20																		
Submit PhD thesis	May-20	Oct-20																		

Figure 1: PhD timeline.

Research plan

This section lists the tasks to be completed and the target dates by which specific milestones will be achieved. Figure 1 shows a timeline of both the finished and remaining tasks.

- Develop and evaluate new approaches for clustering point clouds that are both size and point density invariant (month 14).
- Develop and evaluate new approaches beyond the current state-of-the-art, e.g., deep learning (month 18).
- Develop and test the framework for estimating when broccoli is ready for harvesting (month 24).
- Prepare final versions including testing in broccoli harvesting scenario (month 30).
- Submit PhD thesis (month 36).

Knowledge and Technology Transfer

Activity 1. in collaboration with Earth Rover (https://earthrover.cc) a large dataset of broccoli crops in the field was collected at Pollybell Farms (https://www.pollybell.co.uk). The data was collected using an Intel RealSense Depth Camera D435 mounted on a tractor. The tractor moved several times over rows of planted broccoli ready to be harvested. The new dataset will further contribute to evaluate the performance of the developed algorithms on different varieties of broccoli, and on a variety of datasets.

Activity 2. A poster presentation with the first results of a method for accurately segmenting broccoli heads was presented at the "Smart Industry 4 Workshop 2019" in the Nottingham Trent University (http://smartindustry4.uk).

Activity 3. Introductory talk: "Robotics Research at the University of Lincoln", Isaac Newton Building, University of Lincoln, 12th February 2019, for the Lincolnshire Robotics Forum, to representatives from local industry with interests in robotics, followed by discussions and lab tour, sponsored by NatWest Bank, Business Lincolnshire and ERDF.

References

[1] C. W. Bac, J. Hemming, and E. J. Van Henten. Stem localization of sweet-pepper plants using the support wire as a visual cue. Computers and electronics in agriculture, 105:111–120, 2014.

[2] C. Wouter Bac, Eldert J. van Henten, Jochen Hemming, and Yael Edan. Harvesting robots for high-value crops: State-of-the-art review and challenges ahead. Journal of Field Robotics, 31(6):888–911, 2014.

[3] Shivaji Bachche. Deliberation on design strategies of automatic harvesting systems: A survey. Robotics, 4(2):194–222, 2015.

[4] Shivaji Bachche and Koichi Oka. Distinction of green sweet peppers by using various colour 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, 7(2):178–196, 2013.

[5] Ehud Barnea, Rotem Mairon, and Ohad Ben-Shahar. Colour-agnostic shape-based 3D fruit detection for crop harvesting robots. Biosystems Engineering, 146:57–70, 2016.

[6] Avital Bechar and Clément Vigneault. Agricultural robots for field operations: Concepts and components. Biosystems Engineering, 149:94–111, 2016.

[7] Avital Bechar and Clément Vigneault. Agricultural robots for field operations. Part 2: Operations and systems. Biosystems Engineering, 153:110–128, 2017.

[8] Pieter M. Blok, Ruud Barth, and Wim van den Berg. Machine vision for a selective broccoli harvesting robot. IFAC-Papers Online, 49(16):66–71, 2016. 5th IFAC Conference on Sensing, Control and Automation Technologies for Agriculture AGRICONTROL 2016.

[9] Robert Cupec, Emmanuel Karlo Nyarko, and Ratko Grbic. Survey of state-of-the-art point cloud segmentation methods. Technical report, University of Osijek, 2016.

[10] Tom Duckett, Simon Pearson, Simon Blackmore, Bruce Grieve, et al. Agricultural robotics: The future of robotic agriculture. UK-RAS Network White Papers. arXiv preprint arXiv:1806.06762, 2018.

[11] FAOSTAT. Production quantities of cauliflowers and broccoli by country, 2016. Food and Agricultural Organization of the United Nations, Statistics Division. http://www.fao.org/faostat/en/#data/QC/visualize. Retrieved 18 September 2018.

[12] Guo Feng, Cao Qixin, Cui Yongjie, and Nagata Masateru. Fruit location and stem detection method for strawberry harvesting robot. Transactions of the Chinese Society of Agricultural Engineering, 10, 2008.

[13] L. Ferguson, U. A. Rosa, S. Castro-Garcia, S. M. Lee, J. X. Guinard, J. Burns, W. H. Krueger, N. V. O'connell, and K. Glozer. Mechanical harvesting of California table and oil olives. Advances in Horticultural Science, pages 53–63, 2010.

[14] Jingyao Gai, Lie Tang, and Brian Steward. Plant recognition through the fusion of 2D and 3D images for robotic weeding. In 2015 ASABE Annual International Meeting, page 1. American Society of Agricultural and Biological Engineers, 2015.

[15] A. Gongal, Suraj Amatya, Manoj Karkee, Q. Zhang, and Karen Lewis. Sensors and systems for fruit detection and localization: A review. Computers and Electronics in Agriculture, 116:8–19, 2015.

[16] Eleonora Grilli, Fabio Menna, and Fabio Remondino. A review of point clouds segmentation and classification algorithms. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W3:339–344, February 2017.

[17] Song Han, Sun Xueyan, Zhang Tiezhong, Zhang Bin, and Xu Liming. Design optimisation and simulation of structure parameters of an eggplant picking robot. New Zealand Journal of Agricultural Research, 50(5):959–964, 2007.

[18] M. W. Hannan, T. F. Burks, and Duke M. Bulanon. A machine vision algorithm combining adaptive segmentation and shape analysis for orange fruit detection. Agricultural Engineering International: The CIGR e-journal, XI (Manuscript 1281), December 2009.

[19] Shigehiko Hayashi, Katsunobu Ganno, Yukitsugu Ishii, and Itsuo Tanaka. Robotic harvesting system for eggplants. Japan Agricultural Research Quarterly: JARQ, 36(3):163–168, 2002.

[20] Wei Ji, Dean Zhao, Fengyi Cheng, Bo Xu, Ying Zhang, and Jinjing Wang. Automatic recognition vision system guided for apple harvesting robot. Computers & Electrical Engineering, 38(5):1186–1195, 2012.

[21] A. R. Jimenez, R. Ceres, and J. L. Pons. A survey of computer vision methods for locating fruit on trees. Transactions of the ASAE, 43(6):1911, 2000.

[22] Naoshi Kondo, Koichi Tanihara, Tomowo Shiigi, Hiroshi Shimizu, Mitsutaka Kurita, Mitsuyoshi Tsutsumi, Vui Kiong Chong, and Shigemune Taniwaki. Path planning of

tomato cluster harvesting robot for realizing low vibration and speedy transportation. Engineering in Agriculture, Environment and Food, 2(3):108–115, 2009.

[23] Naoshi Kondo, Kazuya Yamamoto, Hiroshi Shimizu, Koki Yata, Mitsutaka Kurita, Tomoo Shiigi, Mitsuji Monta, and Takahisa Nishizu. A machine vision system for tomato cluster harvesting robot. Engineering in Agriculture, Environment and Food, 2(2):60–65, 2009.

[24] Ferhat Kurtulmus, Won Suk Lee, and Ali Vardar. Green citrus detection using 'eigenfruit', colour and circular Gabor texture features under natural outdoor conditions. Computers and Electronics in Agriculture, 78(2):140–149, 2011.

[25] Keerthy Kusumam, Tomáš Krajník, Simon Pearson, Grzegorz Cielniak, and Tom Duckett. Can you pick a broccoli? 3D-vision based detection and localisation of broccoli heads in the field. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Korea, 2016. IEEE.

[26] Keerthy Kusumam, Tomáš Krajník, Simon Pearson, Tom Duckett, and Grzegorz Cielniak. 3D-vision based detection, localization, and sizing of broccoli heads in the field. Journal of Field Robotics, 34(8):1505–1518, 2017.

[27] Peilin Li, Sang-heon Lee, and Hung-Yao Hsu. Review on fruit harvesting method for potential use of automatic fruit harvesting systems. Procedia Engineering, 23:351– 366, 2011.

[28] Lorenzo Maggioni, Roland von Bothmer, Gert Poulsen, and Ferdinando Branca. Origin and domestication of cole crops (Brassica oleracea L.): linguistic and literary considerations. Economic botany, 64(2):109–123, 2010.

[29] Moshe Mann, Boaz Zion, Itzhak Shmulevich, and Dror Rubinstein. Determination of robotic melon harvesting efficiency: a probabilistic approach. International Journal of Production Research, 54(11):3216–3228, 2016.

[30] Anh Nguyen and Bac Le. 3D point cloud segmentation: A survey. In 6th IEEE Conference on Robotics, Automation and Mechatronics (RAM), pages 225–230, Nov 2013.

[31] Tien Thanh Nguyen, Erdal Kayacan, Josse De Baedemaeker, and Wouter Saeys. Task and motion planning for apple harvesting robot. IFAC Proceedings Volumes, 46(18):247–252, 2013.

[32] Tien Thanh Nguyen, Koenraad Vandevoorde, Niels Wouters, Erdal Kayacan, Josse G. De Baerdemaeker, and Wouter Saeys. Detection of red and bicoloured apples on tree with an RGB-D camera. Biosystems Engineering, 146:33–44, 2016.

[33] Hiroshi Okamoto and Won Suk Lee. Machine vision for green citrus detection in tree images. Environmental Control in Biology, 48(2):93–99, 2010.

[34] M. D. Orzolek, W. J. Lamont, L. F. Kime Jr., and J. K. Harper. Broccoli production. In Agricultural Alternatives series, Agricultural Alternatives series. Penn State Cooperative Extension, 2012.

[35] T. Ota, J. Bontsema, S. Hayashi, K. Kubota, E. J. Van Henten, E. A. Van Os, and K. Ajiki. Development of a cucumber leaf picking device for greenhouse production. Biosystems engineering, 98(4):381–390, 2007.

[36] Peter Rajendra, Naoshi Kondo, Kazunori Ninomiya, Junzo Kamata, Mitsutaka Kurita, Tomowo Shiigi, Shigehiko Hayashi, Hirotaka Yoshida, and Yasushi Kohno. Machine vision algorithm for robots to harvest strawberries in tabletop culture greenhouses. Engineering in Agriculture, Environment and Food, 2(1):24–30, 2009.

[37] Rachael Angela Ramirez. Computer vision based analysis of broccoli for application in a selective autonomous harvester. MA thesis, Virginia Polytechnic Institute and State University, July 2006.

[38] Inkyu Sa, Chris Lehnert, Andrew English, Chris McCool, Feras Dayoub, Ben Upcroft, and Tristan Perez. Peduncle detection of sweet pepper for autonomous crop harvestingCombined colour and 3D information. IEEE Robotics and Automation Letters, 2(2):765–772, 2017.

[39] Takaya Saito and Marc Rehmsmeier. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS ONE, 10(3):e0118432, 2015.

[40] Yongsheng Si, Gang Liu, and Juan Feng. Location of apples in trees using stereoscopic vision. Computers and Electronics in Agriculture, 112:68–74, 2015.

[41] Abhisesh Silwal, Joseph R. Davidson, Manoj Karkee, Changki Mo, Qin Zhang, and Karen Lewis. Design, integration, and field evaluation of a robotic apple harvester. Journal of Field Robotics, 34(6):1140–1159, 2017.

[42] Abhisesh Silwal, Manoj Karkee, and Qin Zhang. A hierarchical approach to apple identification for robotic harvesting. Transactions of the ASABE, 59(5):1079–1086, 2016.

[43] Amy L. Tabb, Donald L. Peterson, and Johnny Park. Segmentation of apple fruit from video via background modelling. In ASABE Annual International Meeting, page 1. American Society of Agricultural and Biological Engineers, 2006.

[44] Kanae Tanigaki, Tateshi Fujiura, Akira Akase, and Junichi Imagawa. Cherry harvesting robot. Computers and Electronics in Agriculture, 63(1):65–72, 2008.

[45] Kang Tu, Ke Ren, Leiqing Pan, and Hongwen Li. A study of broccoli grading system based on machine vision and neural networks. In International Conference on Mechatronics and Automation, pages 2332–2336. IEEE, 2007.

[46] E. J. Van Henten, D. A. Van't Slot, C. W. J. Hol, and L. G. Van Willigenburg. Optimal manipulator design for a cucumber harvesting robot. Computers and electronics in agriculture, 65(2):247–257, 2009.

[47] Efi Vitzrabin and Yael Edan. Changing task objectives for improved sweet pepper detection for robotic harvesting. IEEE Robotics and Automation Letters, 1(1):578– 584, 2016.

[48] Juan P. Wachs, H. I. Stern, T. Burks, and V. Alchanatis. Low and high-level visual feature-based apple detection from multi-modal images. Precision Agriculture, 11(6):717–735, 2010.

[49] Chenglin Wang, Yunchao Tang, Xiangjun Zou, Weiming SiTu, and Wenxian Feng. A robust fruit image segmentation algorithm against varying illumination for vision system of fruit harvesting robot. Optik-International Journal for Light and Electron Optics, 131:626–631, 2017.

[50] Ulrich Weiss and Peter Biber. Plant detection and mapping for agricultural robots using a 3D LIDAR sensor. Robotics and autonomous systems, 59(5):265–273, 2011.

[51] Rong Xiang, Huanyu Jiang, and Yibin Ying. Recognition of clustered tomatoes based on binocular stereo vision. Computers and Electronics in Agriculture, 106:75– 90, 2014.

[52] Yongwei Xu, Kenji Imou, Yutaka Kaizu, and Kiyotaka Saga. Two-stage approach for detecting slightly overlapping strawberries using HOG descriptor. Biosystems engineering, 115(2):144–153, 2013.

[53] Yuanshen Zhao, Liang Gong, Yixiang Huang, and Chengliang Liu. A review of key techniques of vision-based control for harvesting robot. Computers and Electronics in Agriculture, 127:311–323, 2016.

[54] Yuanshen Zhao, Liang Gong, Yixiang Huang, and Chengliang Liu. Robust tomato recognition for robotic harvesting using feature images fusion. Sensors, 16(2):173, 2016.

[55] Yuanshen Zhao, Liang Gong, Chengliang Liu, and Yixiang Huang. Dual-arm robot design and testing for harvesting tomato in greenhouse. IFAC-Papers Online, 49(16):161–165, 2016.

Appendices

Appendix A

A Preliminary broccoli head identification results

This appendix outlines the results of preliminary experiments conducted on broccoli head identification. The initial stage of this PhD research has focused on the evaluation and optimisation of a previous work by Kusumam *et al.* [26]. In their paper, they showed a method for detecting and locating mature broccoli heads in cluttered outdoor field conditions based on depth images acquired by a RGBD sensor. The paper evaluates a combination of Viewpoint Feature Histograms (VFH), Support Vector Machines (SVM) classifier, and a temporal filter to track the detected heads. Their results showed a precision rate of 95.2% and 84.5% on datasets collected from fields in the UK and Spain, respectively.

A.1 3D vision system pipeline

The entire 3D system pipeline by Kusumam *et al.* was re-implemented in a single application written in C++ to replicate and improve their results. To achieve this goal, one of the initial filtering steps, called Statistical Outlier Removal, was removed as it was computationally expensive and contributed poorly to the final result. A layout of the updated pipeline can be seen in Figure 2.

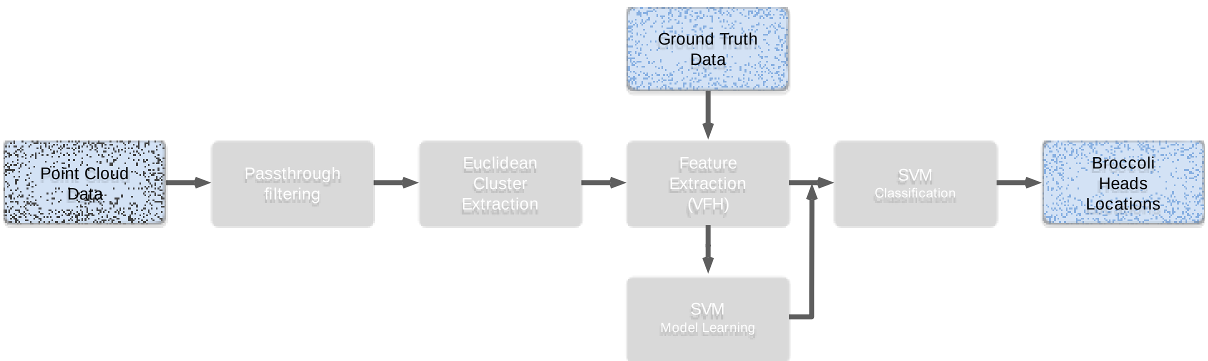

Figure 2. Updated broccoli detection pipeline. Each 3D point cloud data frame is first processed by passthrough (depth) filtering and cluster extraction routines. Then VFH features vectors are extracted and matched with ground truth data to either learn a model or to predict the broccoli heads locations.

The experiments reported in this appendix were performed using the same datasets as reported in Kusumam's paper, and included an additional dataset collected from fields in California, USA. Table 1 shows the main characteristics of the datasets along with the various parameters used by the Euclidean Cluster Extraction algorithm. Similarly, Table 2 lists the distribution of some datasets elements after been processed by the pipeline, as well as the number of descriptors used to train and test the classifier.

	Frames	Min	Max	Tolerance	Clusters	Avg cluster size
Surfleet run 2	300	500	10,000	0.005	12,984	2,092
Surfleet run 4	300	500	10,000	0.005	12,956	2,153
Fuente Alamo	300	300	10,000	0.004	17,182	1,627
California	300	70	6,000	0.0048	60,657	349

Table 1. Euclidean cluster extraction (ECE) parameters. The method is very sensitive to the variation of the tolerance parameter as even a variation of 0.0001 has a notable impact on the number of clusters produced and on their sizes. Since to the distance from the sensor to the crops was larger for the California dataset, more clusters are produced with a smaller number of points each.

	Frames	AvgSqrD	Descriptors	Positive		Negative	
Surfleet run 2	300	0.29	12,984	1,028	7.9%	11,929	92.1%
Surfleet run 4	300	0.29	12,956	1,039	8.0%	11,946	92.0%
Fuente Alamo	300	0.34	17,182	960	5.6%	16,223	94.4%
California	300	0.87	60,657	2,830	4.7%	57,827	95.3%

Table 2. Datasets distribution. The average square distance (AvgSqrD) is measured taking the 20 nearest points to the origin at coordinates (0,0,0). Also, note that all datasets for these experiments are highly unbalanced, i.e. the sample size in the positive and negative data classes are unevenly distributed.

A.2 Classification results

We evaluated the 3D system pipeline for detecting broccoli heads using VFH feature descriptors. For each cluster generated by the pipeline a VFH feature descriptor was produced. If the cluster was broccoli, according to ground truth data, it was labelled as a positive sample; otherwise it was labelled negative. These descriptors, in turn, constitute the set of training and testing samples to be classified. Note that both sets are highly unbalanced, i.e. they exhibit a significant difference in the number of positive and negative samples. In this case, the negatives notably outnumber the positives as large portions of each point cloud frame are from leafs, soil or other elements. All classification results were performed using a Support Vector Machine (SVM) classifier.

An SVM is a machine learning algorithm used for both classification or regression problems. The algorithm produces a model that is a representation of training samples as points in a space divided into classes by a separation gap. Testing samples are then mapped into that space and predicted to be part of a class based on which side of the gap they fall. An SVM has shown to be efficient even in cases where the data is not linearly separable. A non-linear classification can be achieved by using kernels, i.e. functions that map samples into a different feature space. Other parameters can also be tuned to improve the classifier behaviour. The regularization parameter (often named C parameter) controls the balance between classification accuracy (generalisation) and model simplicity. For large values of C, the classifier will look for a smaller (and often harder to find) gap to separate classes. On the other hand, small values of C will cause the classifier to choose a larger separating gap, even if it misclassifies some samples.

Finally, the gamma parameter defines the influence of a single training sample in finding the separation gap. A low gamma value causes that points far away from a viable separation gap are considered in its calculation. Conversely, a high gamma value increases the influence of closer points. All classification results were performed using the following parameters: RBF (radial basis function) kernel, C = 2.0, and gamma = 0.0078125. These parameters were determined by k-fold cross validation (k = 10) based on a grid search. The final classification output of the system pipeline is a set of clusters representing the broccoli heads as shown in Figure 3.

(a) (b) (c)

Figure 3. A sample of the broccoli heads detection results. The output of the broccoli detection pipeline on selected examples of frames from the (a) UK, (b) Spanish, and (c) California datasets. The segments shown in red are the positive broccoli head detections produced by the system.

The performance of the classifier is evaluated using precision-recall curves (PRC), as they provide a more accurate interpretation of a classifier performance on unbalanced samples [39]. Precision represents a ratio of true positive detections to the total number of positive detections (true and false), whereas recall is a ratio of true positive detections to the total number of both true positive and false negative detections. The precision and recall values are computed over a range of discrimination threshold values of the classifier. Figure 4 (a,b,c) shows the performance evaluation using PRC plots on a selection of examples. Similarly, Table 3 shows the average precision results of the classifier using the modified pipeline for the different training-testing combinations of the datasets.

Training/ *Testing*	*Surfleet run 2*	*Surfleet run 4*	*Fuente Alamo*	*California*
Surfleet run 2	98.23%	98.31%	90.04%	17.98%
Surfleet run 4	96.75%	99.40%	88.56%	26.70%
Fuente Alamo	96.23%	97.26%	95.50%	22.39%
California	65.60%	73.19%	52.76%	67.28%

Table 3. The value shown is the average precision score (APS) at various discrimination threshold settings. Each precision score is computed as TP/(TP+FP). When a dataset is trained and tested with itself, a proportion of 75/25% is used. In any other case, 100% of each dataset is used for training and for testing.

Since the performance of the pipeline decreased significantly when it was tested with the California dataset, its performance was also evaluated when the training and testing samples were balanced. To this end, all positive samples from each set were combined with an equal number of negative samples taken at random. Table 4 shows the classification results using these balanced samples. Figure 4 (d,e,f) shows the performance evaluation using PRC plots on a selection of examples.

The results showed that the current method does not segment properly when the distance from the sensor to the crops is bigger. The segmented broccoli heads were, on average, a third smaller, and with higher point density than those from other datasets. A more robust approach for segmenting point clouds is currently being investigated that is both size and point density invariant [30, 9, 16].

Training/ *Testing*	*Surfleet run 2*	*Surfleet run 4*	*Fuente Alamo*	*California*
Surfleet run 2	99.79%	99.64%	97.67%	80.02%
Surfleet run 4	99.10%	99.73%	97.02%	82.31%
Fuente Alamo	99.24%	99.43%	99.56%	79.60%
California	89.08%	93.40%	85.42%	91.83%

Table 4. APS at various discrimination threshold settings. All datasets for these experiments are balanced, i.e. the number of both positive and negative samples is the same. Training and testing sets were split in similar fashion to those listed in Table 3.

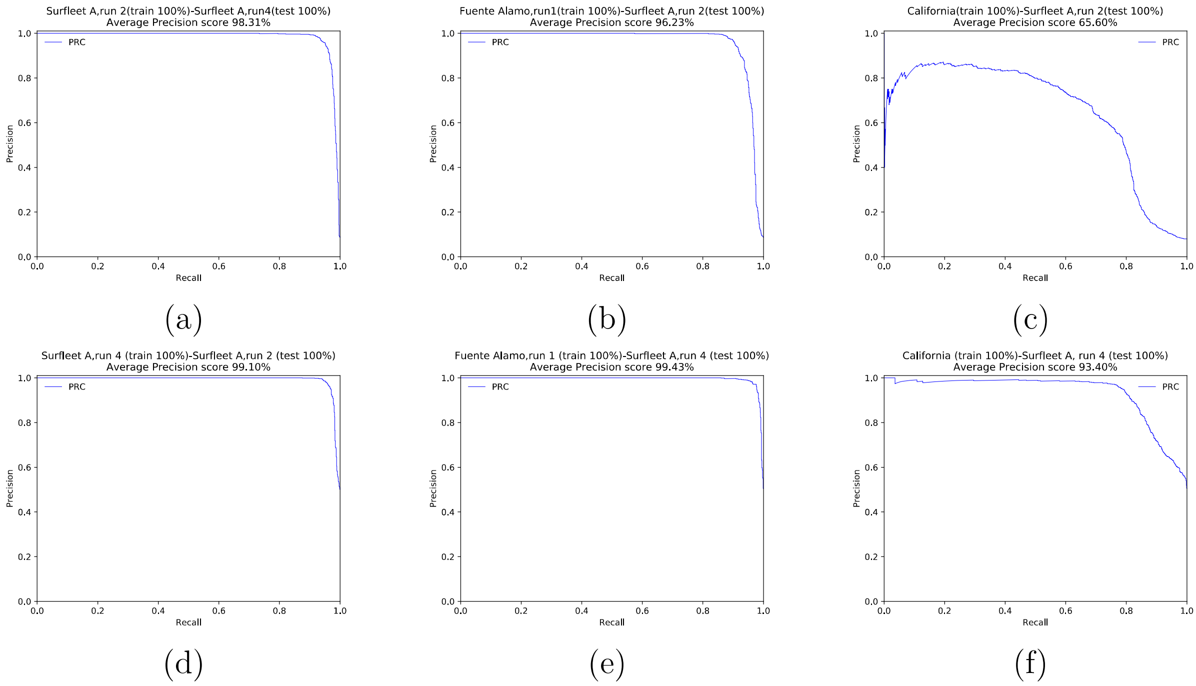

Figure 4. Performance evaluation of the SVM classification on three training-testing unbalanced datasets combinations (a), (b) and (c), and on three training-testing balanced datasets combinations: (d), (e) and (f).

