Improving obstacle awareness for robotic harvesting of sweet-pepper

Cornelis Wouter Bac

Improving obstacle awareness for robotic harvesting of sweet-pepper

C. Wouter Bac

Thesis committee

Promotor
Prof. Dr E.J. van Henten
Professor of Farm Technology
Wageningen University

Co-promotor
Dr J. Hemming
Researcher Computer Vision & Robotics, Wageningen UR Greenhouse Horticulture
Wageningen University and Research Centre

Other members
Prof. Dr C.J.F. ter Braak, Wageningen University
Prof. Dr P.P. Jonker, Delft University of Technology
Prof. Dr J.G. De Baerdemaeker, KU Leuven, Belgium
Dr G.W. Kootstra, Wageningen UR Food and Biobased Research

This research was conducted under the auspices of the Graduate School of
Product Ecology & Resource Conservation (PE&RC)

Improving obstacle awareness for robotic harvesting of sweet-pepper

C. Wouter Bac

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 9 January 2015
at 1:30 p.m. in the Aula.

C. Wouter Bac
Improving obstacle awareness for robotic harvesting of sweet-pepper
186 pages

PhD thesis, Wageningen University, Wageningen, NL (2015)
With references, with summaries in Dutch and English

ISBN 978-94-6257-180-8

Abstract

Obstacles are densely spaced in a sweet-pepper crop and they limit the free workspace for a robot that can detach the fruit from the plant. Previous harvesting robots mostly attempted to detach a fruit without using any information of obstacles, thereby reducing the harvest success and damaging the fruit and plant. The hypothesis evaluated in this research is that a robot capable of distinguishing between hard and soft obstacles, and capable of employing this knowledge, improves harvest success and decreases plant damages during harvesting. In line with this hypothesis, the main objective was to develop a sweet-pepper harvesting robot capable of distinguishing between hard and soft obstacles, and of employing this knowledge.

As a start, the thesis describes the crop environment of a harvesting robot, reviews all harvesting robots developed for high-value crops, and defines challenges for future development. Based on insights from this review, we explored the ability to distinguish five plant parts. A multi-spectral imaging set-up and artificial lighting were developed and pixels were classified using a decision tree classifier and a feature selection algorithm. Classification performance was found insufficient and therefore post-processing methods were employed to enhance performance and detect plant parts on a blob basis. Still, performance was found insufficient and a focussed study was conducted on stem localization. The imaging set-up and algorithm developed for stem localization were used to provide real stem locations for motion planning simulations. To address the motion planning problem, we developed a new method of selecting the grasp pose of the end-effector. The new method and the stem localization algorithm were both integrated in the harvesting robot, and we tested their contribution to performance. This research is the first to report a performance evaluation of a sweet-pepper harvesting robot tested under greenhouse conditions. The robot was able to harvest sweet-peppers in a commercial greenhouse, but at limited success rates: harvest success was 6% when the Fin Ray end-effector was mounted, and 2% when the Lip-type end-effector was mounted. After simplifying the crop, by removal of fruit clusters and occluding leaves, harvest success was 26% (Fin Ray) and 33% (Lip-Type). Hence, these properties of the crop partly caused the low performance. The cycle time per fruit was commonly 94 s, i.e. a factor of 16 too long compared with an economically feasible time of 6 s. Several recommendations were made to bridge the gap in performance. Additionally, the robot's novel functionality of stem-dependant determination of the grasp pose was evaluated to respond to the hypothesis.

Testing the effect of enabling stem-dependent determination of the grasp pose revealed that, in a simplified crop, grasp success increased from 41% to 61% for the Lip-type end-effector, and stem damage decreased from 19% to 13% for the Fin Ray end-effector. Although these effects seem large, they were not statistically significant and therefore resulted in rejection of the hypothesis. To re-evaluate significance of the effects, more samples should be tested in future work.

In conclusion, this PhD research improves the obstacle awareness for robotic harvesting of sweet-pepper by the robot's capability of perceiving and employing hard obstacles (plant stems), whereas previous harvesting robots either lumped all obstacles in one obstacle class, or did not perceive obstacles. This capability may serve as useful generic functionality for future robots.

Table of Contents

Chapter 1 – General Introduction 1

Chapter 2 – Harvesting robots for high-value crops: state-of-the-art review and challenges ahead 9

Chapter 3 – Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper 47

Chapter 4 – Pixel classification and post-processing of plant parts using multi-spectral images of sweet-pepper 77

Chapter 5 – Stem localization of sweet-pepper plants using the support wire as a visual cue 89

Chapter 6 – Analysis of a motion planning problem for fruit harvesting in a dense obstacle environment 109

Chapter 7 – Performance evaluation of a harvesting robot for sweet-pepper 131

Chapter 8 – Conclusion, General Discussion and Recommendations 157

Summary 169

Samenvatting 173

Dankwoord / Acknowledgements 177

Curriculum Vitae 180

List of Publications 181

PE&RC PhD Training Certificate 184

The Chapters 2, 3, 4, and 5 have been published, and Chapters 6 and 7 were submitted as journal articles, as indicated on the first page of each chapter. Except for a different font and addition of chapter-based numbering, the text of the manuscripts was integrally adopted and maintained throughout this thesis. Citations should be made to the original article(s).

Chapter 1 – General Introduction

C.W. Bac

1.1 Background

Harvesting of fruit is an activity performed since the existence of mankind. The act of fruit harvesting was presumably first described in the biblical narrative about Adam and Eve, who picked and ate a fruit from one of the trees present in the Garden of Eden. The estimated period of this writing ranges from 1440 to 1410 B.C (Tyndale 2008).

Nowadays, fruit is harvested manually or by a machine, depending on the type of crop that bears the fruit. Two types of crop exist: 'single-harvest' or 'multiple-harvest' crops. Single-harvest crops produce only one batch of edible product that can involve fruit, tubers or the root of a plant. These produce, hereafter referred to as fruit, are fairly equal in ripeness. As a result, all fruit can be harvested at once and plants bearing the fruit can be damaged in case of annual crops such as wheat and maize. Perennial crops grown for juice extraction (apple, orange, grape) are also single-harvest crops, but they require a more careful treatment of the plant. Their fruit are commonly harvested through shaking mechanisms (Sanders 2005) because fruit damages are not critical for juice extraction. Many machines have been developed in the 20th century to harvest single-harvest crops (Reid 2011). Multiple-harvest crops, however, produce several batches of fruit along a growing season. And, their fruit do not ripen equally. Multiple-harvest crops therefore require selective harvesting, where the ripeness level needs to be assessed, to decide which fruit are ready to harvest. Harvesting of multiple-harvest crops is still done manually throughout the world because machines are not able to perform selective harvesting. The high labour input required, increases production costs and therefore multiple-harvest crops are also referred to as 'high-value crops' (Temu and Temu 2005) or 'specialty crops' (USDA 2014a). The term 'high-value crops' is used throughout the thesis. To selectively harvest a high-value crop, advanced machines are needed, i.e. robots.

Many different definitions exist to define a 'robot'. Although people are mostly able to distinguish between a robot and a machine, fuzzy classifications will probably remain. In this research, the definition by Sevila and Baylou (1991) is used:

> *'a robot refers to a device operating on a variety of tasks and objects and includes sensors, data processing systems, and actuators for adaptive and selective actions.'*

Robots typically differ from machines in their capability of adaptive and selective actions. At the moment of writing, there is only one harvesting robot commercially available for high-value crops: the Spanish company AGROBOT commercialized a strawberry harvesting robot in 2012 (Agrobot 2014).

To advance robotics research for crops, 14 participants from different countries initiated the CROPS project (Clever Robots for crOPS, www.crops-robots.eu) funded by the European Commission (Grant Agreement no. 246252). Wageningen UR Greenhouse Horticulture led the development of a harvesting robot for sweet-pepper, which is the context of this PhD research.

Sweet-pepper (*Capsicum annuum*) is also referred to as 'bell pepper', 'capsicum', or just 'pepper'. Major producers of sweet-pepper fruit are China (14 mln tons per year), Mexico (1.7 mln tons per year) and Indonesia (1.1 mln tons per year). Production takes commonly place in the open field or in greenhouses (plastic or glass). The Netherlands annually produces 0.4 mln tons at about 1300 ha of glasshouses (USDA 2014b).

1.2 Motivations to develop harvesting robots

Harvesting robots are expected to play an important role in facing the challenge of feeding the growing world population and fulfilling the increasing need for high-quality food produced at a competitive price. Harvesting robots do not only reduce current labour costs that constitute 29% of the production costs (Jukema and Van de Meer 2009), but they also enable new functionality using sensing abilities that humans either not possess, or cannot realize at a similar accuracy, consistency and cost. One can think of:

- Early disease detection and tracking-and-tracing of fruit. If a plant or fruit is infected by a disease during growth, the disease symptoms can be observed in an early stage using, for instance, chlorophyll fluorescence imaging (Gorbe and Calatayud 2012). After harvesting, a diseased fruit can be traced back to the specific diseased plant, if the fruit or packaging material was labelled. As a result, treatment can focus on only a specific area in the greenhouse, thereby saving chemicals use.
- Measuring plant physiological properties to apply plant-specific treatment and local climate control. Currently growers lack detailed information about spatial variation in crop productivity caused by varying local climate conditions, fertilization, or the biological variation present in plants. By extracting yield data per plant, or measuring leaf and stem properties, such feed-back data can enable the grower to take appropriate measures.
- Market-driven fruit picking with instantaneous grading and packing. Markets that require fruit of a certain weight, ripeness level, shelf life, hygiene level, or appearance can be targeted. Although current harvesting robots are not yet able to measure these properties, future robots could be equipped with such ability to be able to classify, grade and pack the fruit for the appropriate market. As a result, this feature might reduce food waste because, for instance, fruit with a short shelf life can be sold to the local market, whereas other fruit can be exported.

This new functionality offered by harvesting robots is expected to contribute to the "more with less" mission for food production in the coming decades (Gomiero et al. 2011). Harvesting robots are an enabling technology and improve the "farm to fork" efficiency in terms of lower production costs, and a better crop management, food quality and food security.

An often observed complaint is that robots steal people's jobs. In Dutch greenhouse horticulture, jobs in harvesting and plant maintenance are mostly performed by workers from Eastern Europe due to the difficulty of recruiting skilled local workers. Dutch citizens consider working in greenhouses as tedious and heavy work (Van Henten 2006; Van der Bol 2012) that has to be performed under high temperatures and a high humidity. The pool of skilled local workers is declining and aging, not only in the Netherlands, but also in the USA

(He et al. 2013) and Japan (Morio et al. 2013). Furthermore, hiring foreign workers increases the risk of mistreatment, especially if they are unfamiliar with the employment regulations. In the Netherlands, for instance, there have been several incidents of mushroom and strawberry growers who mistreated foreign workers (KVW 2012). Hiring foreign workers is therefore not a desired solution and it is unclear whether workers from Eastern Europe will be willing to do the job on the long term. Introducing harvesting robots resolves these difficulties by gradually transforming physically intensive jobs into new jobs in application development, sales, or other services.

1.3 Problem description and scope

Figure 1-1 provides a view of a sweet-pepper plant with two candidate fruit for harvesting. When analysing this picture, a number of challenges can be observed for robotic harvesting. At this camera viewpoint, both fruit are occluded by leaves, the plant stem, and other plant parts. Occlusions make it difficult to detect and localize the fruit, to determine ripeness, and to determine the position and orientation for a mechanism that can detach the fruit from the plant. If, despite these occlusions, an accurate position and orientation can be determined, another challenge arises: accessing and detaching the target fruit without damaging the target fruit, a nearby fruit, or stems and leaves. Hence, leaves and other plant parts not only occlude the target fruit, but also act as an obstacle when a mechanism approaches the target fruit for harvesting. Furthermore, these obstacles are densely spaced and therefore limit the free workspace for a mechanism that can detach the fruit. The problem at the start of this PhD research was the lack of a robot able to avoid obstacles and successfully harvest a sweet-pepper fruit, without damaging the fruit or other plant parts.

Figure 1-1. View of a sweet-pepper plant with two ripe fruit attached.

Based on the problem description, a number of requirements were established for the robot (Hemming et al. 2011) in terms of economic, technical, safety and social aspects. The most relevant requirement for economic feasibility was that the maximum cost of the robot should not exceed 196 k€ if a cycle time of 6 s per fruit can be achieved for 20 h per day, with

95% of the ripe fruit successfully harvested. An alternative scenario of a cycle time of 20 s per fruit would correspond to a maximum robot cost of 59 k€. Technical requirements were that the robot should successfully harvest a fruit without damaging the crop; should not spread diseases; should work under a wide range of temperature, CO_2 and humidity levels; should physically fit in-between the crop rows etc. Regarding safety, the robot should allow for safe interaction with a human that operates the robot. Lastly, the value of harvesting robots for society should be explained for acceptance of the robot upon market introduction.

Using these requirements, an overview of eight functions to harvest a fruit was established in Figure 1-2. This PhD research focused mostly on three of these functions: localize obstacles; determine grip and cut location of the target fruit; approach target fruit (functions 3-5). The other functions were developed by colleagues of WUR Greenhouse Horticulture and partners of various other work packages within the CROPS project.

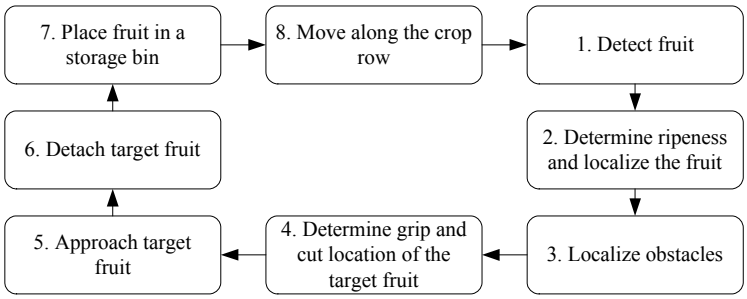


Figure 1-2. Functionality required to harvest a fruit.

The reason to focus on the obstacles involved in functions 3-5 was, on the one hand, the need to avoid obstacles for successful sweet-pepper harvesting and, on the other hand, the lack of existing work on obstacles. The literature review chapter in this thesis shows that current harvesting robots mostly attempted to detach a fruit without using information of obstacles. There were a few exceptions of harvesting robots with some ability of obstacle detection by assuming that any detected object in the environment is an obstacle, except for the target fruit (Van Henten et al. 2002; Nguyen et al. 2013). However, this approach can limit the accessibility of fruit because not every object is necessarily an obstacle in reality. To illustrate, when humans approach a fruit they sometimes push a leaf aside without causing any damage to the plant. Humans commonly treat the plant stem more carefully because a damage can seriously hamper plant growth and, consequently, future production of fruit. A leaf was therefore considered a soft obstacle, whereas the plant stem was considered a hard obstacle. A robot capable of distinguishing between hard and soft obstacles is currently lacking. If capable, a robot can take actions that may benefit from such knowledge. The hypothesis posed in this research is that a robot capable of distinguishing between hard and soft obstacles, and of employing this knowledge, improves harvest success and decreases plant damages during harvesting.

1.4 Objective

In line with the hypothesis, the main objective was *to develop a sweet-pepper harvesting robot capable of distinguishing between hard and soft obstacles, and of employing this knowledge*. Sub-objectives were the following:

- to describe the crop environment of a harvesting robot, to review all harvesting robots developed for high-value crops, and to define challenges for future development;
- to distinguish between different plant parts of a sweet-pepper plant and group them in hard and soft obstacles, using pixel-based classification;
- to convert detected pixels of plant parts into blobs using post-processing techniques;
- to localize the stem of a sweet-pepper plant;
- to develop and test a motion planning algorithm for the robot arm and end-effector that plans a motion to a target fruit while avoiding hard obstacles;
- to test the harvesting robot under greenhouse conditions and evaluate its performance, and to validate the hypothesis of this PhD research.

Coherence of these sub-objectives is discussed in the following. The literature review in the first sub-objective served as a basis to establish the main objective and sub-objectives. As an explorative study, five different plant parts were classified. Classification performance was found insufficient and therefore post-processing methods were employed to enhance performance and detect plant parts on a blob basis. Still, performance was found insufficient and a focussed study was conducted on stem localization. The imaging set-up and algorithm developed for stem localization were used to provide real stem locations for motion planning simulations. A result of analysing the motion planning problem was a new method of selecting the grasp pose of the end-effector. This new method and the stem localization algorithm were both integrated in the harvesting robot, together with other hardware and software modules. In the last sub-objective, the robot was tested under real greenhouse conditions. The contribution of stem localization and grasp pose selection to performance were determined in these tests, to validate the hypothesis of this PhD research. Furthermore, the performance of other modules was reported and discussed.

1.5 Demarcation

Automating navigation of the harvesting robot along the crop row (function 8 in Figure 1-2) is outside the scope of this thesis and was done manually. This demarcation was agreed upon within the consortium of the CROPS project because solutions already exist for this functionality in, for instance, spraying robots (Holland Green Machine BV, The Netherlands). The CROPS project and this PhD research therefore focused on the remaining functionality that is more challenging to automate. Furthermore, transportation, grading and packing of harvested fruit is outside the scope of this research.

1.6 Outline

The outline of the thesis corresponds to the sub-objectives of this research and consists of eight chapters: the general introduction (Chapter 1), a literature review (Chapter 2), five research chapters (Chapter 3-7) and a conclusion with a general discussion and recommendations (Chapter 8).

Chapter 2 describes a literature review concerning harvesting robots. This chapter characterizes the working environment, i.e. the crop environment, of a harvesting robot. Furthermore, all harvesting robots developed in the last three decades were reviewed using

quantitative performance indicators, and the chapter ends with challenges and R&D directions for future development of harvesting robots.

Chapter 3 deals with pixel-based classification of five plant parts: the stem, fruit, top of a leaf, bottom of a leaf and the petiole. A multi-spectral imaging set-up was developed and images recorded were used to extract pixel-based features for classification. The classifier, a decision tree, was trained by a novel performance measure that not only considers the classification accuracy, but also the consistency of classification among different scenes, i.e. the robustness of classification. Results of pixel-based classification were further processed in Chapter 4. A number of morphological image processing techniques and blob analyses were implemented to obtain detection performance on a blob basis.

The performance of stem detection was found insufficient and therefore a different approach was taken in Chapter 5 to localize the plant stems using the support wires as a visual cue. Stereo images were recorded using colour cameras mounted with a small baseline. An algorithm consisting of five steps was developed that includes novel components such as adaptive thresholding, use of object-based and 3D features and use of minimum expected stem distance. Performance of stem localization was evaluated under strong and moderate sunlight conditions commonly encountered in horticultural practice.

In Chapter 6 the motion planning problem of sweet-pepper harvesting is addressed. Locations of fruit and stems were recorded in the greenhouse and were combined with a 3D representation of the actual harvesting robot, to create a realistic environment for motion planning simulations. Subsequently, two novel analyses were conducted. The first analysis compared two methods of selecting the azimuth angle of the end-effector with respect to the fruit targeted. The second analysis evaluated the sensitivity of five parameters specifying the crop (stem spacing and fruit location), the robot (end-effector dimensions and robot position) and the planning algorithm on successfully finding a collision-free goal configuration and path towards the fruit. The novel method of selecting the azimuth angle was integrated in the harvesting robot.

In the last phase of the PhD project, components developed in this research and by several project partners were integrated into the harvesting robot. After integration, the robot was tested in a commercial greenhouse and performance indicators were determined and reported in Chapter 7. A protocol was established to measure the contribution of stem localization, in addition to only localizing the fruit, to a higher performance and less damages to the crop.

Lastly, Chapter 8 concludes the thesis by validating the hypothesis and by discussing Chapters 1-7 in a general discussion. Major findings are put in a broader perspective and discussed with respect to existing literature. Limitations and other relevant assumptions are debated and recommendations are made for future research.

References

Agrobot (2014). Agrobot. http://www.agrobot.com/. Accessed 5 September 2014.

Gomiero, T., Pimentel, D., & Paoletti, M. G. (2011). Is There a Need for a More Sustainable Agriculture? *Critical Reviews in Plant Sciences, 30*(1-2), 6-23.

Gorbe, E., & Calatayud, A. (2012). Applications of chlorophyll fluorescence imaging technique in horticultural research: A review. *Scientia Horticulturae, 138*, 24-35.
He, L., Zhang, Q., & Charvet, H. J. (2013). A knot-tying end-effector for robotic hop twining. *Biosystems Engineering, 114*(3), 344-350.
Hemming, J., Bac, C. W., & Van Tuijl, B. A. J. (2011). CROPS project deliverable 5.1: Report with design objectives and requirements for sweet-pepper harvesting. Wageningen, The Netherlands: Wageningen UR Greenhouse Horticulture.
Jukema, G., & Van de Meer, R. (2009). Labor costs in arable farming and greenhouse horticulture; in Dutch: Arbeidskosten in de akkerbouw en glastuinbouw. The Hague, The Netherlands: Landbouw Economisch Instituur (LEI).
KVW (2012). TV documentary about Dutch mushroom production, in Dutch. Broadcasted by KRO at 20 December 2012. http://keuringsdienstvanwaarde.kro.nl/seizoenen/2012/afleveringen/20-12-2012.
Morio, Y., Shibata, K., & Murakami, K. Agricultural worker tracking and gesture recognition system for remote assistance. In *IFAC Proceedings Volumes (IFAC-PapersOnline), 2013* (PART 1 ed., Vol. 1).
Nguyen, T. T., Kayacan, E., De Baedemaeker, J., & Saeys, W. Task and motion planning for apple harvesting robot. In *IFAC Proceedings Volumes (IFAC-PapersOnline), 2013* (PART 1 ed., Vol. 4, pp. 247-252).
Reid, J. F. (2011). The Impact of Mechanization on Agriculture. *The Bridge, 41*(3), 22-29.
Sanders, K. F. (2005). Orange harvesting systems review. *Biosystems Engineering, 90*(2), 115-125.
Sevila, F., & Baylou, P. (1991). Ch. 5: The principles of robotics in agriculture and horticulture. In *Progress in agricultural physics and engineering* (pp. 119-147).
Temu, A. E., & Temu, A. A. High Value Agricultural Products for Smallholder Markets in Sub-Saharan Africa: Trends, Opportunities and Research Priorities. In *High Value Agricultural Products Workshop, Cali, Colombia, October 2005*.
Tyndale (2008). *Holy Bible Text Edition NLT: New Living Translation*. Carol Stream, IL, USA: Tyndale House Publishers, Inc.
USDA (2014a). USDA Definition of Specialty Crop. http://www.ams.usda.gov/AMSv1.0/getfile?dDocName=STELPRDC5082113. Accessed 15 July 2014.
USDA (2014b). World bell and chili peppers production: 1990-2007. http://usda.mannlib.cornell.edu/usda/ers/BellAndChile/Table64.xls. Accessed 15 July 2014.
Van der Bol, B. (2012). Unemployed workers in West-Netherlands are unwilling; in Dutch: Westlandse werkelozen willen niet. *10 August 2012* (pp. 8-9): NRC.
Van Henten, E. J. (2006). Greenhouse Mechanization: State of the Art and Future Perspective. *Acta Horticulturae*(710), p. 55-69.
Van Henten, E. J., Hemming, J., Van Tuijl, B. A. J., Kornet, J. G., Meuleman, J., Bontsema, J., et al. (2002). An Autonomous Robot for Harvesting Cucumbers in Greenhouses. *Autonomous Robots, 13*(3), 241-258.

Chapter 2 – Harvesting robots for high-value crops: state-of-the-art review and challenges ahead

C.W. Bac[a,b], E.J. van Henten[a,b], J. Hemming[a], Y. Edan[c]

Affiliations:
[a] Wageningen UR Greenhouse Horticulture, Wageningen University and Research Centre
[b] Farm Technology Group, Wageningen University and Research Centre
[c] Department of Industrial Engineering and Management, Ben-Gurion University of the Negev

This chapter was published in *Journal of Field Robotics* (2014), Vol. 31(6), p. 888-911

Abstract

This review article analyzes state-of-the-art and future perspectives for harvesting robots in high-value crops. Objectives were to characterize the crop environment relevant for robotic harvesting; to perform a literature review on the state-of-the-art of harvesting robots using quantitative measures; and to reflect on the crop environment and literature review to formulate challenges and directions for future research and development. Harvesting robots were reviewed regarding the crop harvested in a production environment, performance indicators, design process techniques used, hardware design decisions, and algorithm characteristics. On average localization success was 85%, detachment success was 75%, harvest success was 66%, fruit damage was 5%, peduncle damage was 45%, and cycle time was 33 s. A kiwi harvesting robot reached the shortest cycle time of 1 s. Moreover, performance of harvesting robots did not improve in the last three decades and none of these 50 robots were commercialized. Four future challenges with R&D directions were identified to realize a positive trend in performance and to successfully implement harvesting robots in practice: (1) simplifying the task; (2) enhancing the robot; (3) defining requirements and measuring performance; (4) considering additional requirements for successful implementation. This review article may provide new directions for future automation projects in high-value crops.

2.1 Introduction

Harvesting is performed several times during production of a high-value crop and is a candidate operation for automation. High-value crops are generally non-staple crops such as fruit, vegetables, ornamentals, condiments, and spices (Temu and Temu 2005). A major reason for a crop to be classified as a high-value crop is high labor input required. Labor costs in Dutch greenhouse horticulture, for instance, constitute 29% of the production costs (Jukema and Van de Meer 2009). These high costs motivate automation of harvesting and, furthermore, other motivations involve social, environmental, and food quality aspects (Lewis et al. 1983; Van Henten 2006).

To automate the harvesting process, robots have been actively developed over the last 30 years (Sarig 1993; Grift et al. 2008; Li et al. 2010; Li et al. 2011). Reviews specifically targeting harvesting robots included reviews of complete systems (Sarig 1993; Li et al. 2010; Li et al. 2011), or a sub-task of the robot such as guidance and navigation (Li et al. 2009; González et al. 2009; Bechar 2010) and fruit localization (Jiménez et al. 2000a; Kapach et al. 2012). These reviews indicate robots were capable to harvest fruit autonomously, under a certain range of environmental conditions. Yet, harvesting robots are still far from mature and harvesting is still manual due to the limited performance of current robots. A gap in the literature exists regarding a better understanding of this limited performance, and regarding future challenges that can generate a positive trend in performance.

We attempted to close this gap by pursuing the following three objectives. First of all, we characterize the crop environment, which is defined as the working environment of the robot. We describe three sources of variation, to elucidate how the complexity of the crop environment forms the main bottleneck to better performance. This description also supports

terms and findings posed throughout the paper. Second, quantitative performance indicators were reviewed in only one article (Jiménez et al. 2000a), whereas they can be useful to compare robots, to determine the state-of-the-art in performance, and to identify future challenges for development. Therefore, this review tries to identify and review quantitative performance indicators for harvesting robots developed in the last 30 years. Third, we obtained new insights by completing the first and second objective, and complemented these insights with our experience, to establish future challenges and R&D directions for development of harvesting robots.

These three objectives inspired the outline of this paper. Section 2.2 characterizes the crop environment in which harvesting robots have to operate. Section 2.3 presents the method, Section 4 results, and Section 5 the discussion of the literature review. Finally, Section 2.6 presents future challenges and R&D directions written from the authors' perspective. In fact, this section can be viewed by the international robotics community as a plea to tackle the remaining challenges to develop reliable, economically viable, and useful harvesting robots, as asserted in a similar way for a review dealing with transmission line maintenance robots (Toussaint et al. 2009).

2.2 Characterizing the Crop Environment for Robotic Harvesting

A crop environment includes three sources of variation relevant for robotic harvesting. First, there is a lot of variation of objects within a crop. Harvesting is performed on objects with ill-defined positions, shapes, sizes, and colors. Objects are also hard to see and reach due to occluding branches and leaves (Section 2.1). Second, the environment in which the robot must operate provides a lot of variation. Crops are grown in different production environments, where lighting conditions can vary. Growers employ different cultivation systems to support, train and maintain their crops, which influence visibility and accessibility of fruit (Section 2.2.2). Third, many different kinds of crops are grown and technical challenges and market potential for robotic harvesting can vary among crops (Section 2.2.3). Consequently, these three sources of variation render fruit harvesting a complex task and, furthermore, define harvesting robots as a unique and challenging field of agricultural robots.

2.2.1 Variation of objects in a crop

Objects, i.e. fruit and other plant parts, in a crop vary in position, size, shape, and reflectance due to the natural variation that exists in nature. High-value crops all contain this object variation and we show a sweet-pepper crop (Figure 2-1) as an example. Positions of fruit are widely distributed in a height range of about 1 m. Currently, crop growth models cannot predict where fruit will occur and each fruit therefore must be localized. Furthermore, shapes of fruit vary (e.g., sweet peppers are cylindrical, but the width/height ratio is not constant and widths vary from 6-11 cm; tomatoes can be round or elongated and vary in size). As a result, sensory techniques must cope with this variation. Reflectance (mostly color and near-infrared) of fruit is a visual cue often used to distinguish fruit from other plant parts and strongly varies. Further challenges for fruit detection regarding position, shape, size, reflectance, and texture are discussed by Kapach et al. (2012).

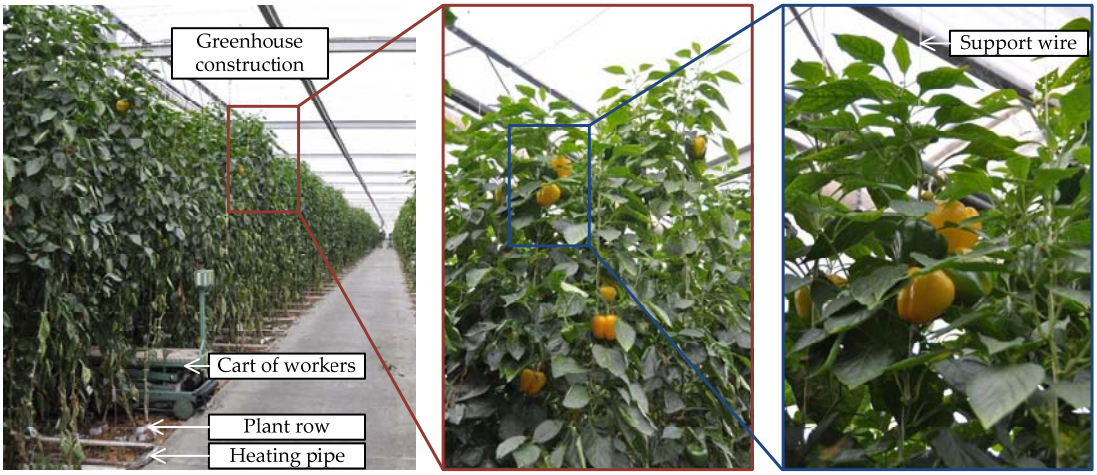


Figure 2-1. Sweet-pepper crop in a commercial greenhouse. Center photo is an enlargement of the left photo. Right photo is an enlargement of the center photo.

Color is used as a ripeness indicator for many crops and required ripeness level can vary. For example, in the beginning of the season (March-April), sweet-pepper fruit are considered ripe if >80% of the fruit surface is colored. Later (May-October), temperature rises and coloring of 50% is acceptable for a fruit to be ready for harvesting. Furthermore, these percentages can vary depending on market demand. Consequently, a robot must handle this color variation.

When a robot grasps sweet-peppers it requires a mechanism that can handle a range in sizes (width of sweet-pepper fruit vary between 6-11 cm) and shapes. A fruit's susceptibility to damage is another important factor for grasping. For instance, apples and avocado must be handled with great care, whereas kiwis can be handled with little risk of bruises.

Another aspect for harvesting is accessibility and visibility of fruit. The right photo in Figure 2-1 displays a fruit cluster consisting of two ripe fruit and two unripe fruit. Accessing these ripe fruit with a gripper is a challenge because the other three fruit must be avoided to prevent damages to these fruit. Also, other obstacles such as plant parts (leaves, stems, fruit), support wires, and construction elements need to be avoided once a gripper follows a path towards a targeted fruit. Therefore, localization of these objects is required to avoid damage to the crop, fruit or construction elements. Obstacles not only block access to the fruit, but also reduce visibility. Fruit visibility can be occluded by other fruit and leaves (

Figure 2-1) causing difficulties for fruit localization (Kapach et al. 2012).

Age of the plant, pests and diseases, and cultivation practices of growers can influence object variation as well. An old rose crop, for instance, produces less straight stems than a young rose crop and such change can decrease performance of a robot. Also, pest and diseases can influence leaf angle and color of objects. Furthermore, growers sometimes apply different fertilization regimes or climate management that cause variation of objects and their locations. The object variation described so far, already holds for one specific cultivar of a crop. But, this variation becomes even larger when considering that different cultivars exist for each crop. A cultivar exhibits slight genetic differences resulting in fruit and plant parts with differences in position, shape, size, and color. For peppers, for instance, about 25 different cultivars are grown in the Netherlands. Worldwide this number is even larger (Figure 2-2).

Figure 2-2. Pepper cultivars commercially offered by Westland Seeds. Source: www.westlandseeds.nl

In view of variation, an additional complicating factor is the short market lifetime of cultivars nowadays (about eight years). Breeding companies continuously develop new cultivars for better production or to respond to new market demands. Ultimately, a robot should be able to handle each of these cultivars.

2.2.2 Variation in the environment

A crop grows in a production environment and a cultivation system is used to guide and maintain the plant. We describe production environments and cultivation systems in this section and their impact on robotic harvesting.

High-value crops grow in four production environments: orchard, greenhouse, indoor, and open field (Figure 2-3). Large plants, such as trees, typically appear in orchards, whereas smaller plants typically appear in the other three production environments. Although few crops are grown indoor, this production environment is increasingly investigated, such as a plant factory to grow lettuce (Shimizu et al. 2011).

Figure 2-3. High-value crops appearing in four production environments: orchard hosting apples (left; source: http://fruit.cfans.umn.edu/apples/beforeyoustart/), greenhouse hosting tomatoes (center-left), indoor hosting mushrooms (center-right; source: http://www.tuinadvies.be/champignons_telen.htm), and open field hosting melons (right; source: http://angelvalleyfarm.files.wordpress.com/2012/06/melons.jpg)

We identified several factors that may influence design and performance of harvesting robots and that differ among production environments (Table 2-1).

Table 2-1. Differences among four production environments, for factors that can have an effect on harvesting robots. Factors scored from poor (--) to excellent (++).

Factors	Production environment			
	Orchard	Greenhouse	Indoor	Open Field
Wind and rain protection	--	+	++	--
Controllable lighting	--	--	++	--
Consistent plant development	+/-	+	++	-
Visibility of objects	--	--	+	+/-
Accessibility of objects	--	--	+	+/-
Ease of navigation	--	+	++	--
Suitable for stationary robots	--	+/-	++	--

A robot must handle wind and rain as additional disturbance factors for crops grown in orchards and the open field. Lighting conditions can only be controlled for indoor crops, and are a strong disturbance for image processing, which can limit harvest success. For instance, Plebe and Grasso (2001) showed that harvest success in a cloudy sky was 85%, but was only 52% under a low sun angle. Plant development is more consistent for crops grown in greenhouses and indoor because climate conditions (CO_2, temperature, humidity) can be controlled. As a result, objects in these crops will be more consistent in terms of position, shape, size and color, rendering it easier for a robot to handle these objects. Nevertheless, visibility and accessibility of objects is generally worse for orchard and greenhouse crops due to a denser canopy with stronger presence of obstacles (construction elements, other plant parts) that need to be avoided by the robot. Navigation in orchards and the open field is more challenging because robots have to rely on guidance systems, whereas in greenhouses and indoors robots can drive on rail systems. Indoor grown crops can be suitable for stationary robots if crops grow on movable benches that can be transported to the robot. Such benches are sometimes applied in greenhouses as well (Hayashi et al. 2011). Movable benches cannot be used in the open field because of little economic incentive and also not in orchards because trees are too heavy to transport. In conclusion, an indoor production environment seems most suitable for harvesting robots (Table 2-1).

Growers use different cultivation systems (also referred to as “training system”) to grow crops, such as the V-system and “Spanish” system for sweet-pepper (Jovicich et al. 2004). The cultivation system influences accessibility and visibility of fruit. For instance, in the conventional cultivation system for cucumber, fruit are more occluded than in the high-wire cultivation system (Van Henten et al. 2002). In apple orchards, a “fruit wall” cultivation system is under investigation (Saeys and Nguyen 2012) to simplify both manual harvesting and robotic harvesting (Figure 2-4). This system involves small tree spacing with tightly pruned trees.

Figure 2-4. Fruit are better visible and accessible in the "fruit wall" cultivation system (left) than in the conventional cultivation system (right). Photos courtesy of PCFruit, Belgium (http://www.pcfruit.be/Homepage/22724/pcfruit).

Pruning is performed to influence plant growth and can have a large effect on geometry of the tree. Orange trees, for instance, occur in short and wide or narrow and tall. These different geometries influence the travelling distance along the different ripe fruit to be picked. It turned out, obviously, that horizontal travelling is shorter for short and wide trees, whereas vertical travelling is shorter for narrow and tall trees. Furthermore it was shown that vertical travelling was 5% shorter than horizontal travelling for typically shaped orange trees (Edan et al. 1990). In conclusion, the cultivation system influences suitability of the crop environment for robotic harvesting.

2.2.3 Variation among crops

As indicated in the previous sections, there exists a lot of variation within one crop and also the environment adds variation, but variation becomes even more pronounced when considering all high-value crops. To demonstrate how geometry of plants differs among crops and how geometry determines the challenges for harvesting, we compare a rose crop and a tomato crop (Figure 2-5). Such comparison can be made between other crops as well.

Figure 2-5. Rose crop (left) and a tomato crop (right). The cutting position of a rose stem is located deep into the crop, at the bottom, whereas the cutting position for the tomato truss reasonably accessible, at the intersection between the main stem and the vine of the truss.

The challenge for rose harvesting is to travel over a long distance through a dense canopy to finally reach the cutting position, whereas for tomato harvesting the cutting position is easier to reach and not surrounded by many obstacles. The challenge for tomato harvesting is to accurately determine ripeness while large parts of the fruit are not visible, whereas ripeness determination is a relatively simple task for rose harvesting because, when viewed from the top, the flower bud is completely visible.

The total variation aggregates when considering all high-value crops grown. To illustrate the large number of crops, we show crops grown in greenhouses in the Netherlands (Figure 2-6), probably such distributions hold for other countries as well. These crops impose specific technical challenges and crops with similar challenges for robotic harvesting may exist. Researchers should therefore thoroughly analyze the crop of interest to design a robot that can handle the crop.

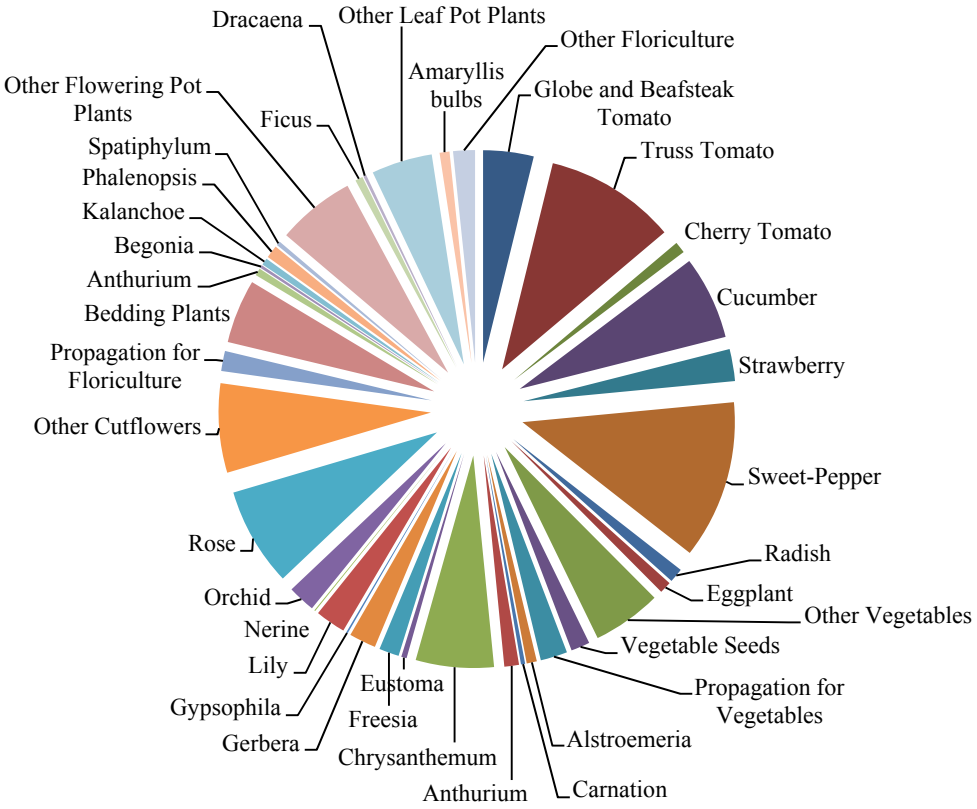


Figure 2-6. Area distribution of crops grown in greenhouses in the Netherlands in 2006. A large number of crops is grown and most cover a small area. Sweet-pepper and truss tomato are the only crops covering an area > 1000 ha. The total area of greenhouses is 10603 ha. Source: adapted from LEI & CBS (2009)

Another aspect that can strongly vary among crops is the market potential a crop offers for robotic harvesting. We discuss two factors. First, the area covered by a high-value crop is generally small (Figure 2-6). Even if a robot would be able to handle all cultivars included in this area, still few robots can be marketed for a crop. This small market potential provides little incentive for industry to develop harvesting robots, as also indicated in a review on robotic harvesters in the USA (Glancey and Kee 2005). Second, depending on the crop,

demand for harvesting can be irregularly distributed over the year. As an example, Figure 2-7 displays labor requirement for two crops: sweet-pepper and rose.

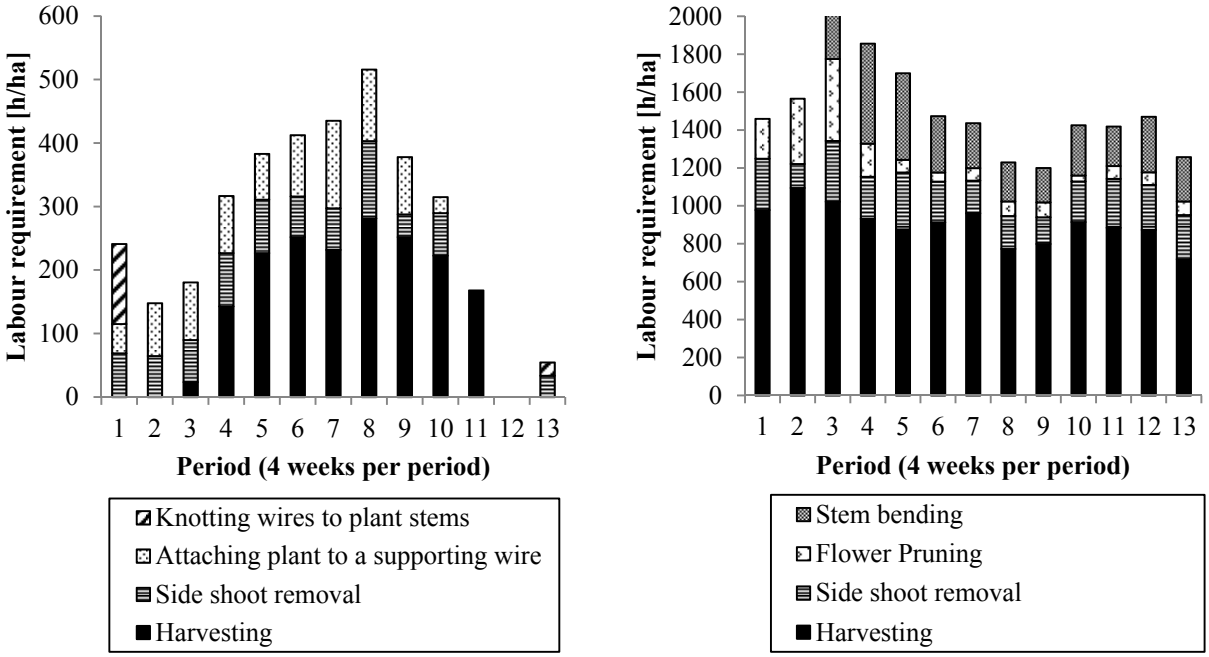


Figure 2-7. Annual human labor requirement of plant-maintenance operations performed in sweet-pepper (left) and rose (right). Labour requirement for harvesting in sweet-pepper is more irregularly distributed over the year than for rose.

In sweet-pepper cultivation, harvesting is performed in periods 3 through 11, which is 36 weeks of the year, and a robot would be out of use in other periods of the year. A rose harvester clearly has a better market potential because roses are harvested year-round and the robot is therefore better utilized. The stable demand for rose harvesting is because of continuous supplementary lighting and because rose plants grow for at least five years, whereas sweet-pepper production is restarted annually and are mainly grown without supplementary lighting.

2.3 Literature Review Method

Harvesting robots were reviewed in-line with the threefold complexity of the crop environment (Section 2.2) with a focus on performance, design, and algorithms. State-of-the-art performance was quantitatively assessed to determine how well robots handled object and environment variation (Section 2.3.1) and performance was compared among projects (Section 2.3.2). Since the robot design influences the robot's ability to deal with object and environment variation, we assessed how researchers carried out the design process (Section 2.3.3) and what design decisions they made (Section 2.3.4). For algorithms we reviewed if they were reported and on the presence of adaptive algorithms so as to enable adaptation to object and environment variation (Section 2.3.5). We defined adaptive algorithms as algorithms that change their parameters or actions based on on-line feedback data. Examples of adaptive algorithms are adaptive thresholding techniques used in image processing and

adaptive motion planning algorithms using, for instance, visual servoing or learning controllers. If the robot contained an algorithm to perform another harvest attempt after failure, we considered such an algorithm non-adaptive because actions were not adapted on-line, i.e. during the operation.

Results of the review were ordered by the production environment (orchard, open field, greenhouse or indoor) and by the particular crop on which research had been focused. In summary, the review addressed the following questions:

- Which crops have been investigated for robotic harvesting
- Which performance measures were reported and which performance measures are relevant to assess a harvesting robot
- What percentage of harvesting robots were autonomous
- Which tested conditions were reported
- What is the overall performance of robots developed so far
- How does performance compare between production environments, crops and along time?
- How did researchers carry out the design process in terms of systematic design and economic analysis
- What hardware components did researchers select
- Which algorithms were reported for the main tasks of fruit harvesting
- Which robots contained adaptive algorithms and did those perform better

A project was included in the review if and only if a complete functional system was built and reported in an English written conference paper or peer-reviewed journal article.

2.3.1 Performance Indicators

The performance indicators were selected based on measures reported in the literature. For each indicator an explanation is given for its relevance.

Performance indicators considered included categorical and continuous indicators. Two performance indicators were measured categorically: whether robots were autonomous (true/false), and whether robots were tested in the lab or the field. We defined a robot as autonomous if the robot performed tasks without human intervention once a human operator placed the robot in the field and set the hardware and algorithm parameters at the start of each field test. The difference between lab and field tests was considered important because a lab environment is usually much more structured than a field environment. As a result, higher performance can be achieved in a lab environment than in the field. Performance reported under field conditions is relevant because a harvesting robot must be eventually implemented under field conditions.

Eight performance indicators were analyzed as continuous indicators.

- Fruit localization success [%]: The number of localized ripe fruit per total number of ripe fruit in the canopy. This indicator was included because a fruit must be localized to determine fruit ripeness and to approach the fruit for detachment.
- False-positive fruit detection [%]: The number of objects falsely detected as fruit per total number of ripe fruit in the canopy. This indicator was included because false-

positive detections can cause failed pick attempts, damages to fruit or plant and increase cycle times.

- Detachment success [%]: The number of successfully harvested ripe fruit per total number of localized ripe fruit. This indicator was included because it measures performance of collision-free motion planning toward a fruit, and of interaction between the end-effector and fruit.
- Harvest success [%]: The number of successfully harvested ripe fruit per total number of ripe fruit in the canopy. This indicator measures overall performance of a harvest cycle
- Cycle time [s]: time of an average full harvest operation, including ripeness determination, localization, fruit detachment, transport of a detached fruit, and robot transport to the next fruit. This time includes time loss caused by failed attempts. This indicator is relevant to determine economic feasibility of the robot.
- Damage rate [%]: the number of damaged fruit or peduncles per total number of localized ripe fruit, caused by the robot. A peduncle is the connecting stem between the fruit and the main stem or branch. Peduncle pull of apples were considered peduncle damages. Damages to fruit or peduncle reduce market value of fruit and are therefore relevant for economic feasibility of the robot.
- Number of fruit evaluated in a test [#]: The number of fruit that were evaluated to calculate localization success, false-positive fruit detection, detachment success, harvest success, damage rate and cycle time. Such numbers are useful to evaluate the statistical significance of the reported performance indicators. As a result, field test results, if reported, with statistical significance allow for performance comparison with other research projects.
- Detachment attempt ratio [-]: The number of detachment attempts [#] divided by the number of successfully detached ripe fruit [#]. This performance measure was included to show the relevance of the detachment success reported. The relevance of reported detachment success is partly influenced by the number of attempts the robot performed. That is, more attempts, especially from different platform positions, increase detachment success.

Ripe fruit were chosen for the performance indicators because high-value crops require ripeness determination. Though incorrect from an agronomic point of view, we will use the term “fruit” throughout the article to indicate harvestable fruit, vegetables, and flowers.

The accuracy of ripeness determination was not reviewed because we did not find a common ground truth measure that holds for several crops. Also, positioning accuracy of the manipulators and end-effectors were not reviewed because accuracy requirements are not expected to be critical for harvesting robots. An end-effector positioning accuracy of ± 0.5 cm is satisfactory to grasp a fruit because of the required wide tool aperture to accommodate for varying fruit size. Such accuracy can be achieved by most mechanisms.

In this article, units were chosen such that multiplication of fruit localization success and detachment success yields harvest success. If one of these three values was missing, the missing value was calculated based on the other two, provided all units were clearly defined

in the article and a "(C)" was added after the performance indicator. Also, a "(C)" was added if performance values had to be converted to match the units of this review article (Table 2-2). Some authors, however, did not report units of the performance indicators, which complicates conversion into similar units. For instance, some authors report cycle times of a "full harvest cycle" while it is unclear whether the harvest operation also includes platform transport to the next fruit. For cases where conversion of cycle times was impossible, unit unclear "(U?)" was added after the performance indicator.

2.3.2 Comparison of Performance

To assess the effect of the crop environment, performance indicators were compared among production environments and among crops in a production environment. In addition, performance indicators were compared between decades to determine how performance advanced along time. Decades were taken as time step because the number of projects reported was too small for smaller time intervals. For comparison of performance indicators, the number of projects, average, and range were extracted for success rates, damage rate and cycle time. If data for comparison of performance were too sparse for statistical significance, we only indicated if the data showed a trend.

2.3.3 Design Process Techniques

To determine if use of design process techniques contributes to better performance, projects were reviewed on the use of systematic design methods and economic analyses. The literature indicates that these two design process techniques can greatly contribute to technical and economic feasibility of a robot (Angeles et al. 2008) and are common practice in industrial robotics implementation (Nof 2009). Systematic design assists researchers in design choices, structures the design process and stimulates creativity, whereas an economic analysis considers allowable costs of a system.

Systematic design methods reviewed include either process-based design methods or systems engineering methods. Examples of process-based methods are methodological design (Siers 2007) or engineering design (Cross 2008). An example of systems engineering is the determination of number of robot arms, multiple arm configuration, degrees-of-freedom, and horticulture workspace design (Edan and Miles 1994).

2.3.4 Hardware Design Decisions

Decisions for hardware design influence the robot performance and, eventually, influence economic feasibility. Hardware design decisions were reviewed in terms of the number of degrees-of-freedom (DOF) used in manipulators and the use of off-the-shelf or custom-made components for the robot platform (travelling device), sensors, manipulator, and end-effector. If DOF used was not reported, the DOF were extracted from figures or pictures in the article. In case we were uncertain regarding the DOF used a question mark was added. Furthermore, we reviewed if authors reported an analysis explaining the choice of the number of DOF.

Hardware components were considered off-the-shelf if it was clear that the component was purchased and used without any further modification to the component, such as an industrial manipulator. A multi-spectral camera that was assembled using off-the-shelf components was considered off-the-shelf. In other cases, hardware components were

considered custom-made. Information was extracted from figures and pictures in the article if the details were not reported in the text.

2.3.5 Algorithm Characteristics

Description of algorithms is important for reuse in future projects. If algorithms were described we determined whether these algorithms were described *partially* in text or *fully* in parameterized flow charts, equations, code (or pseudo-code). We considered a citation to or description of an algorithm to be *fully* described if parameter values were given. Details of the parameter values allow to repeat the experiment.

Presence of algorithms was evaluated for five major tasks involved in fruit harvesting: fruit localization, ripeness determination, obstacle localization, task planning (i.e. which ripe fruit to pick first), and motion planning.

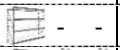
Furthermore we evaluated if any of the algorithms for these five tasks was adaptive (true/false), to determine if robots with an adaptive algorithm dealt better with object and environment variation.

2.4 Literature Review Results

Projects included in the review are projects performed in the last 30 years (Table 2-2). Information of older projects was hard to retrieve. To the best of our knowledge, 50 distinct research projects have been reported.

Table 2-2. Performance evaluation of harvesting robots reported in the period 1982 through 2012, ordered by production environment, crop, and time (present to past)

Crop			Research Location (*Prototype name*) & References	Performance indicators									Design process		Hardware design choices					Algorithm characteristics						
Production environment	Name	Photo of fruit		Autonomous	Fruit loc. success [%]	False-pos. fruit det. [%]	Detachment suc. [%]	Harvest success [%]	Damage rate [Fruit % / Peduncle %]	Cycle time [s]	Number of fruit [Loc. suc. # / False-pos # / Det. suc. # / Harv. suc. # / Cycle time # / Damages #.]	Detach. attempt ratio [-]	Test: Lab ( )/Field ( )	Economic analysis	Syst. design methods	Manipulator DOF	Custom-made ( ) / Off-the-shelf ( ) Platform	Sensors	Manipulator	End-effector	Algorithms: Full ( ) / Partial ( ) / Not reported ( ) Fruit loc.	Ripeness det.	Obst. loc.	Task planning	Mot. planning	Adaptive alg. used
Orchard	Apple		China (De-An et al. 2011)	✓	-	-	77	-	-/-	14.3	-/39/-/-/39/-	-		✗	✗	5					✓	✗	½	✗	✓	✓
	Apple		Belgium (Baeten et al. 2008)	✓	-	-	-	80 (U?)	0/30	9 (U?)	-/-/-/-/-/-	-		✗	✗	6					½	✗	✗	½	½	✗
	Apple		Hungary (*AUFO*) (Kassay and Slaughter 1993)	✗	70 (U?)	-	42 (U?)	-	1/-	2.4 (C)	∞/∞/-/∞/∞	-		✓	✗	5?					✗	✗	✗	✗	✗	✗
	Apple/ Peach		USA (Sites and Delwiche 1988)	✗	90 (U?)	20	-	-	-/-	-	145/145/-/-/-/-	-		✗	✗	-			-	-	✓	✗	✗	✗	✗	✗
	Apple		France (*MAGALI*) (Grand d'Esnon et al. 1987; Sarig 1993);	✗	-	-	-	50 (U?)	0/25	3 (U?)	-/-/-/-/-/-	-		✗	✗	4					✓	✗	✗	✗	✗	✗
	Date palm		Saudi Arabia (Aljanobi et al. 2010)	✗	-	-	-	-	-/-	-	-/-/-/-/-	-		✗	✗	4					✗	✗	✗	✗	✗	✗
	Grape		Japan (Monta 1995)	✓	-	-	-	-	-/-	-	-/-/-/-/-	-		✗	✗	5					½	✗	✗	✗	½	✓
	Kiwi		New Zealand (Flemmer et al. 2009)	✓	-	-	-	-	-/-	1 (U?)	-/-/-/-/-	-		✗	✓	3					✗	✗	✗	✗	✗	✗
	Lychee		China (Liu et al. 2011)	✓	-	-	-	-	-/-	-	-/-/-/-/-	-		✗	✗	-		-			✗	✗	✗	✗	✗	✗
	Oil Palm		Thailand (Sitthichareonchai et al. 1995)	✗	-	-	-	-	-/-	-	-/-/-/-/-	-	-	✗	✗	3	-		-	-	½	½	✗	✗	✗	✗
	Orange		USA (Lee and Rosa 2006)	✗	-	-	84 (U?)	-	-/-	-	-/917/-/-/-	-		✗	✗	3?					✗	✗	✗	✗	✗	✗
	Orange		Italy (Muscato et al. 2005)	✗	-	-	-	-	-/-	8.7	-/-/-/11/-	-		✗	✗	3					✓	✗	✗	✗	½	✓
	Orange		Italy (*OPR*) (Plebe and Grasso 2001)	✓	87	15	-	52-85 (U?)	-/-	5.5-9.7 (U?)	673/673/721/721/721 /-	-		✗	✗	5	-				✓	✗	✗	✗	½	✗
	Orange		Spain (*AGRIBOT*) (Jiménez et al. 2000b; Ceres et al. 1998)	✗	80	0	-	-	-/-	2	-/-/-/-/-/-	-		✗	✗	4					✓	✗	✗	✗	½	✗
	Orange		France and Spain (*CITRUS*) (Sarig 1993; Plá et al. 1993)	✓	75	8	75 (U?)	-	-/-	-	165/165/-/-/-/-	-	-	✗	✗	4	-		-	-	✓	½	✗	✗	✗	✗
	Orange		USA (*CPR*) (Pool and Harrell 1991; Harrell et al. 1990)	✓	-	-	69	-	7/-	3-7	-/-/154/-/154/-	-		✓	✗	3					✓	✗	✗	✗	✓	✓
	Cherry		Japan (Tanigaki et al. 2008)	✓	59	-	67	40	-/-	14 (U?)	80/-/12/12/-/-	-		✗	✗	4					½	½	½	✗	½	✗
Greenhouse	Cucumber		China (Tang et al. 2009)	✓	-	-	93 (U?)	-	-/-	-	-/-/60/-/-/-	-		✗	✗	6	-				✗	✗	✗	✗	✗	✗
	Cucumber		The Netherlands (*CUPID*) (Van Henten et al. 2002; Van Henten et al. 2003)	✓	95	18	78 (C)	74	-/-	124	106/106/179/195/195 /-	2.1		✗	✓	7					½	✓	✗	✗	✓	✗
	Cucumber		Japan (Arima and Kondo 1999)	✓	-	-	-	-	-/-	-	-/-/-/-/-/-	-		✗	✗	7					½	✗	✗	✗	✗	✗
	Eggplant		Malaysia (Wan Ishak et al. 2010)	✗	-	-	-	-	-/-	42 (U?)	-/-/-/-/-/-	-		✗	✗	-	-				✗	✗	✗	✗	✗	✗
	Eggplant		Japan (Hayashi et al. 2002)	✓	-	-	63	-	13 (C)/-	64.1	-/-/40/-/40/40	-		✗	✗	5					✓	✗	✗	✗	½	✓

Greenhouse	Gerbera		Germany (Rath and Kawollek 2009)	✓	97	-	82 (C)	80	5/-	197 (C) (U?)	111/-/-/232/-/57	-		✗	✗	6	-		✓	✗	✗	✗	✓	✗
	Rose		The Netherlands (Noordam et al. 2005)	✓	-	-	-	-	-/-	-	-/-/-/-/-/-	-	-	✗	✗	-	-		-	-	½	✗	✗	✗	✗	✓
	Strawberry		Korea (Han et al. 2012)	✓	-	-	-	-	-/-	7	-/-/-/-/-/-	-		✗	✗	4					✓	✓	✗	✗	✗	✗
	Strawberry		China (Feng et al. 2012)	✓	-	-	86	-	-/-	31.3	-/-/100/-/-/-	-		✗	✗	6					✓	✓	✗	✗	✗	✗
	Strawberry		Japan (Hayashi et al. 2010)	✓	60	12	69 (C)	41	-/-	12.5	-/1130/1130/1130/ 1130/-	1.3		✗	✗	3					½	✓	✗	✗	½	✗
	Strawberry		Japan (Guo et al. 2008)	✓	93	-	-	-	5/-	30	100/-/-/-/100/100	-		✗	✗	3	NI				✓	✓	✗	✗	✗	✗
	Sweet Pepper		Japan (Kitamura and Oka 2005)	✓	-	-	-	-	-/-	-	-/-/-/-/-/-	-		✗	✗	-					✓	✗	✗	✗	½	✓
	Tomato (truss)		Japan (Kondo et al. 2008)	✓	65	-	-	-	-/-	-	17/-/-/-/-/-	-		✗	✗	6?					✓	✗	✗	✗	✗	✗
	Tomato		USA (Ling et al. 2004)	✓	95 (U?)	-	85 (U?)	-	-/-	227	-/-/-/-/-/-	-		✗	✗	6	-				✗	✗	✗	✗	✓	✓
	Tomato		Italy (*AGROBOT*), (Jiménez et al. 2000a; Buemi et al. 1996)	✓	90	-	-	-	-/-	-	-/-/-/-/-/-	-	-	✗	✗	6	-				½	✗	✗	✗	✗	✗
	Tomato (cherry)		Japan (Kondo et al. 1996)	✗	100	-	-	70	-/-	3-5 (U?)	-/-/-/62/-/-	3.3 (C)	-	✗	✗	7					✓	✗	✗	✗	✓	✓
	Tomato		Japan (Hayashi and Sakaue 1996)	✗	-	-	-	-	-/-	41	-/-/-/-/-/-	-		✗	✗	5	-				✗	✗	✗	✗	✗	✗
	Tomato		France (Balerin et al. 1991)	✓	-	-	60	-	-/80	-	-/-/-/-/-/-	-		✗	✗	6					½	½	✗	✗	½	✗
	Tomato		Japan (Namikawa and Ogawa 1989)	✓	-	-	-	-	-/-	22 (U?)	-/-/-/-/-/-	-	-	✗	✗	3	-				✗	✗	✗	✗	✗	✗
	Tomato		Japan (Kawamura et al. 1984)	✓	-	-	-	-	-/-	-	-/-/-/-/-/-	-	-	✗	✗	5					✗	✗	✗	✗	✗	✗
Indoor	Mushroom		UK (Reed et al. 2001)	✓	90	-	84 (C)	76	-/-	6.7	-/-/2290/2506/-/-	1.3		✗	✗	3					✓	½	✗	✗	✗	✗
	Mushroom		UK (Reed 1994)	✓	84	-	68 (C)	57	-/-	-	815/-/689/815/-/-	-		✗	✗	3					✓	✗	✗	✗	✗	✗
Open Field	Asparagus (white)		Greece (Chatzimichali et al. 2009)	✓	-	-	-	-	-/-	-	-/-/-/-/-/-	-		✗	✗	3?					½	✗	✗	✗	½	✗
	Asparagus		Japan (Irie et al. 2009)	✓	-	-	-	-	-/-	13.7	-/-/-/-/-/-	1.0		✗	✗	6					✗	✗	✗	✗	½	✗
	Asparagus		USA (Clary et al. 2007)	✗	-	-	69	-	10/-	-	-/-/-/-/-/-	-		✓	✗	-					✗	✗	✗	✗	✗	✗
	Asparagus		Australia (Arndt 1997)	✓	-	-	-	-	-/-	-	-/-/-/-/-/-	-	-	✗	✗	-					✗	✗	✗	✗	✗	✗
	Asparagus		USA (Humburg and Reid 1991)	✓	86-97	-	-	-	-/-	-	-/-/-/-/-/-	-		✗	✗	-					✗	✗	✗	✗	✗	✗
	Melon		USA, Israel (Edan et al. 2000; Edan 1995)	✓	94	20	92 (C)	86	7/-	15	400/400/374/400/ 400/400	1.3 (C)		✓	✓	3					✓	✓	✗	✓	✓	✗
	Radicchio		Italy (Foglia and Reina 2006)	✗	100	-	-	-	-/-	7	6/-/-/-/-/-	-		✗	✓	2					✓	✗	NI	✗	½	✓
	Saffron		Italy (*Zaffy*) (Antonelli et al. 2011)	✓	-	-	60	-	-/-	-	-/-/-/-/-/-	-		✗	✗	3					✗	✗	✗	✗	✗	✗
	Watermelon		Japan (Sakai et al. 2008)	✓	-	-	87	-	0 (U?)/-	14	-/-/-/-/-/-	-		✗	✓	4					✗	✗	NI	✓	✓	✗
	Watermelon		Korea (Hwang and Kim 2003)	✓	-	-	-	-	-/-	15	-/-/-/-/-/-	-		✗	✓	4					✗	✗	NI	✗	½	✗
	Watermelon		Japan (Tokuda et al. 1995)	✓	-	-	-	-	-/-	-	-/-/-/-/-/-	-		✗	✗	-					✓	✗	NI	✗	✓	✗

Legend: "-" = Not reported or unable to extract from figures or tables; "(U?)" = Unit unclear; "(C)" = Calculated value; "?" = Uncertain about value; "NI" = Task not of interest for the application

Results in Figure 2-8 show how many projects were performed per crop.

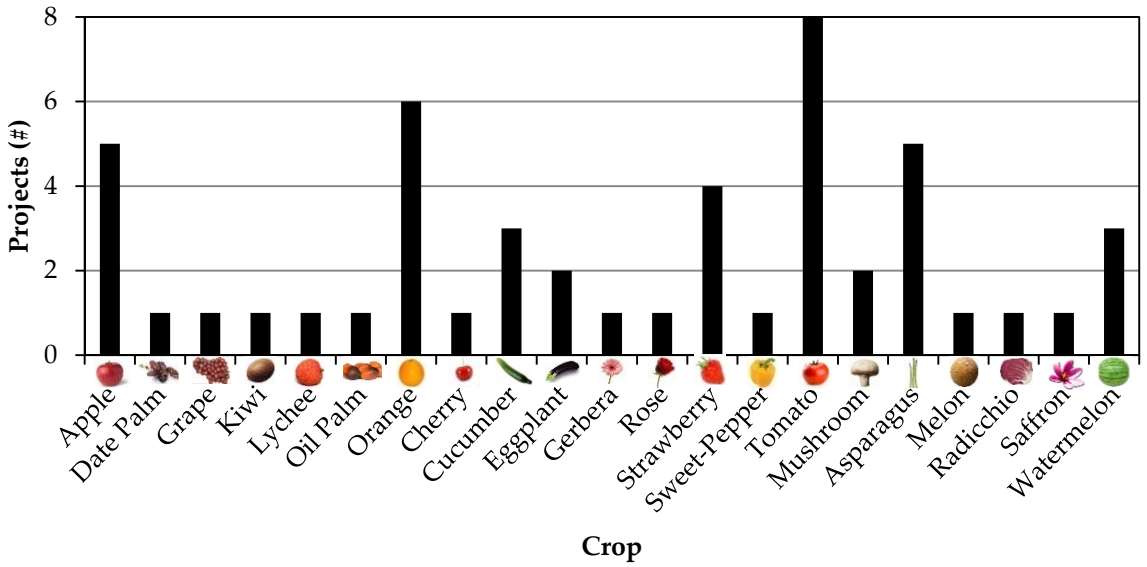


Figure 2-8. Number of distinct projects per crop. Most projects were aimed at tomato, orange, apple, or asparagus harvesting

Projects were aimed at many different crops (Figure 2-8). Only three projects were aimed at ornamental harvesting (rose, saffron, and gerbera) and all others focused on fruit or vegetable harvesting. Half of the projects – 48% (24/50) – were aimed at tomato, orange, apple, or asparagus harvesting.

Results concerning performance indicators (Sections 2.4.1 and 2.4.2), design process techniques (Section 2.4.3), hardware design decisions (Section 2.4.4), and algorithm characteristics (section 2.4.5) are reported in the following sections.

2.4.1 Performance Indicators

One or several quantitative performance indicators were reported for 76% (38/50) of the projects. But, few instances were reported for several of the individual indicators: 19 for localization success, seven for false-positive fruit detection, 20 for detachment success, 11 for harvest success, 10 for fruit damage, three for peduncle damage, and 28 for cycle time.

Most of the projects concerned autonomous robots – 74% (37/50). Only few authors – 12% (6/50) – reported the number of attempts the robot made to harvest a fruit. The average (σ) number of attempts was 1.7 (0.8) attempts per successfully detached ripe fruit. Most performance tests were done in the field – 68% (34/50), a few in the lab – 16% (8/50), or the location of tests was not reported – 16% (8/50).

The average values and range (minimum-maximum) of localization success, detachment success, harvest success, fruit damage, and peduncle damage are in Figure 2-9.

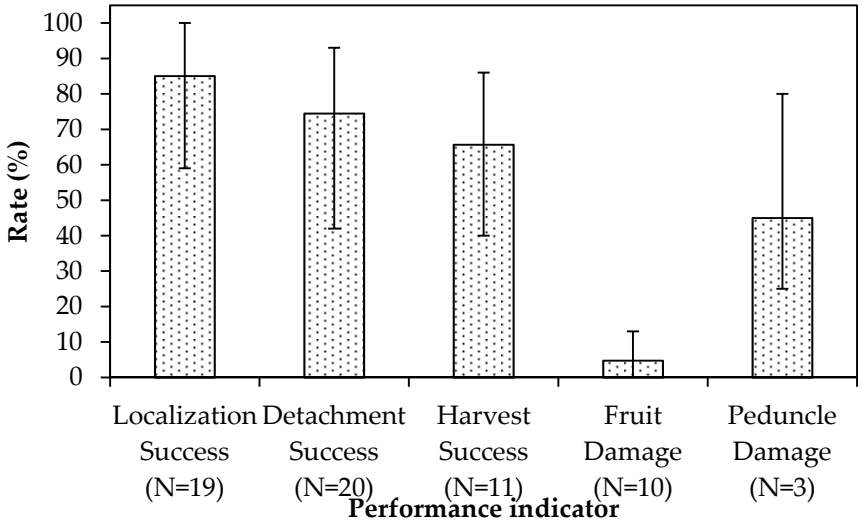


Figure 2-9. Averages and range of reported quantitative performance indicators: localization success, detachment success, harvest success, fruit damage, and peduncle damage. N represents the number of distinct projects

Localization success (85%; 59-100%) was, on average, slightly higher than detachment success (75%; 42-93%). Overall harvest success was 66% (40-86%). Fruit damage was 5% (25-80%) of the localized ripe fruit. Peduncle damage was 45% (25-80%) of the localized ripe fruit. Cycle time showed a large range of 1 to 227 s with an average of 33 s (N=28).

2.4.2 Comparison of Performance

Data were too sparse to reveal statistically significant differences among performance indicators for the four production environments (Figure 2-10). Also, comparison of crops within a production environment revealed no statistically significant differences and data were too sparse for visualization or analysis.

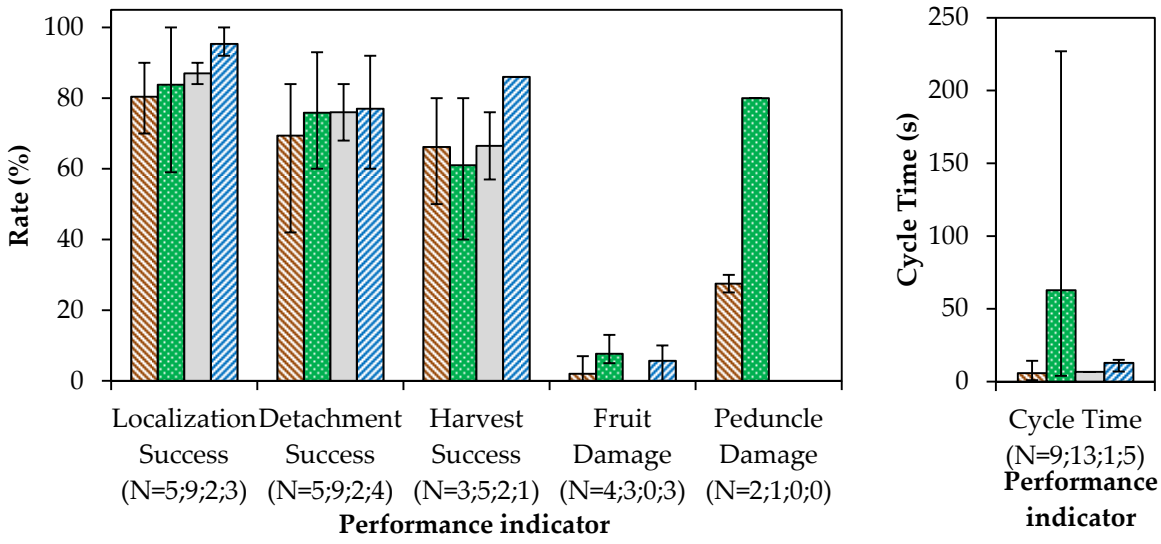


Figure 2-10. Performance indicators for four production environments: orchard (; 16 projects), greenhouse (; 21 projects), indoor (; 2 projects), and open field (; 11 projects). Averages and range of localization success, detachment success, harvest success, fruit damage, and peduncle damage (left). Average and range of cycle time (right). N represents instances per performance indicator per production environment.

Analyses of averages and range of performance indicators for three decades (Figure 2-11) reveal that localization success, harvest success, fruit damage, peduncle damage and cycle time did not improve along time. Only detachment success shows an improving trend.

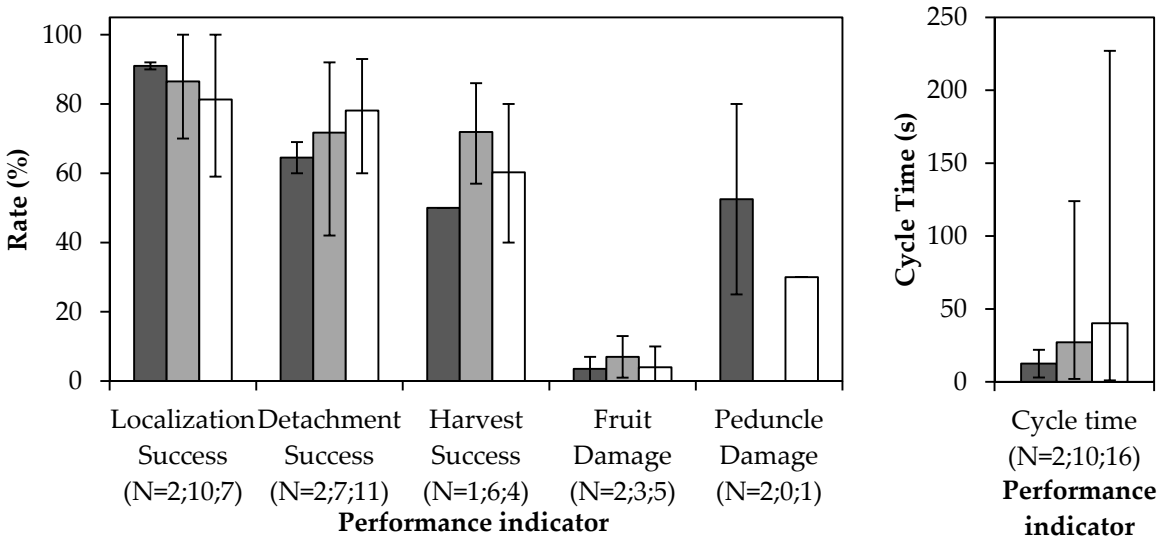


Figure 2-11. Performance indicators for three decades: 1984-1992 (■; 7 projects), 1993-2002 (■; 17 projects), and 2003-2012 (□; 26 projects). Averages and range of localization success, detachment success, harvest success, fruit damage, and peduncle damage (left). Average and range of cycle time (right). N represents instances per performance indicator per decade.

2.4.3 Design Process Techniques

Only 12% (6/50) of the authors reported systematic design methods. Also, only 8% (4/50) of the authors performed an economic analysis. Unfortunately, performance indicators were missing for projects in which systematic design was used and the same crop (e.g. tomato) was harvested. Similarly, performance indicators were missing for projects in which an economic analysis was performed. We were therefore unable to determine if use of systematic design or an economic analysis contributes to better performance.

2.4.4 Hardware Design Decisions

It turned out that 82% (41/50) of the authors reported the degrees-of-freedom of the manipulator (Figure 2-12). In most of the projects 3-DOF manipulators were used. These manipulators were mostly Cartesian manipulators and some were 3-DOF anthropomorphic arms. None of the authors reported an analysis explaining the choice of the number of DOF for their application.

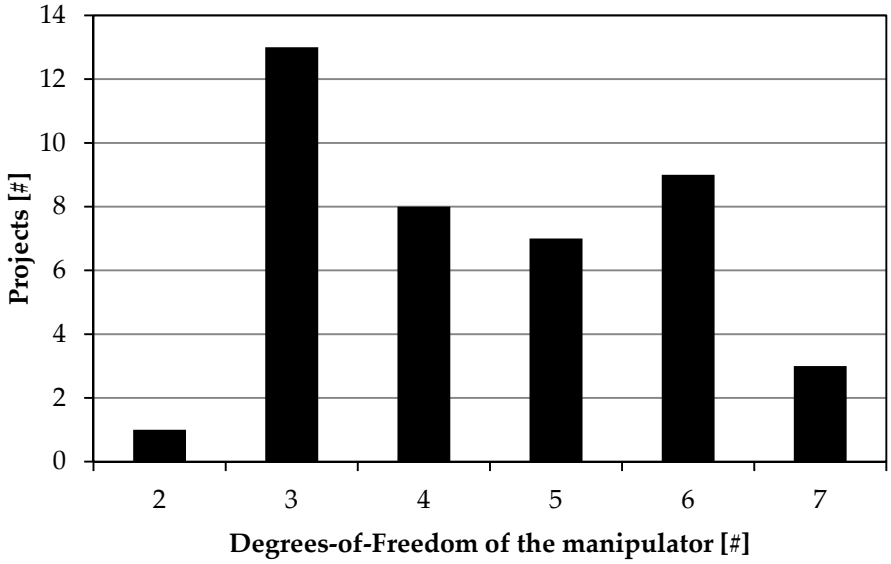


Figure 2-12. Degrees-of-freedom of the manipulators that were used in the projects. Most manipulators were 3-DOF Cartesian manipulators

The use of custom-made or off-the-shelf hardware components is shown in Figure 2-13.

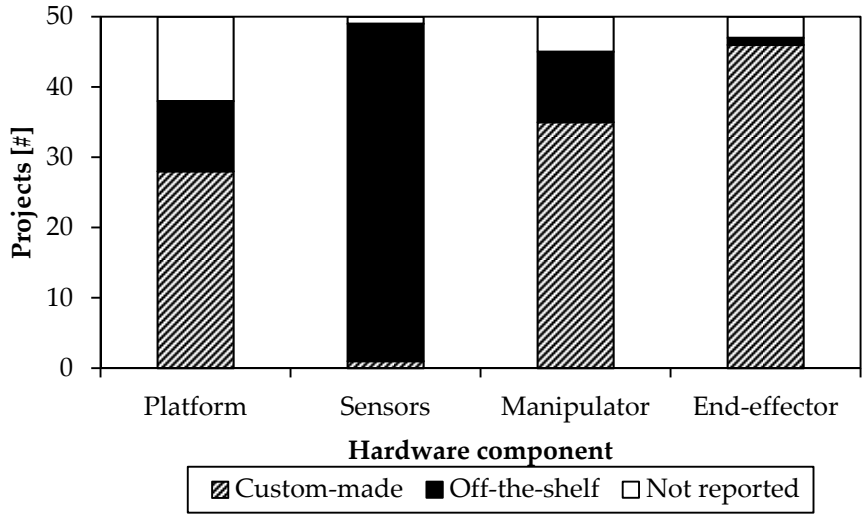


Figure 2-13. Use of custom-made or off-the-shelf hardware components in the projects. Sensors were off-the-shelf, the platform, manipulator and end-effector were mostly custom-made

Most platforms were custom-made – 74% (28/38). Sensors were almost always off-the-shelf – 98% (48/49) – and, for fruit localization, included mostly color and multispectral cameras. More details on the type of sensors used in agricultural robots can be found in other reviews (Jiménez et al. 2000a; Li et al. 2011; Kapach et al. 2012). Manipulators were mostly custom-made – 78% (35/45). Almost all end-effectors were custom-made – 98% (46/47).

2.4.5 Algorithm Characteristics

Analysis of whether algorithms were reported for the five major tasks performed during fruit harvesting, i.e. fruit localization, ripeness detection, obstacle localization, task planning, and motion planning, is shown in Figure 2-14.

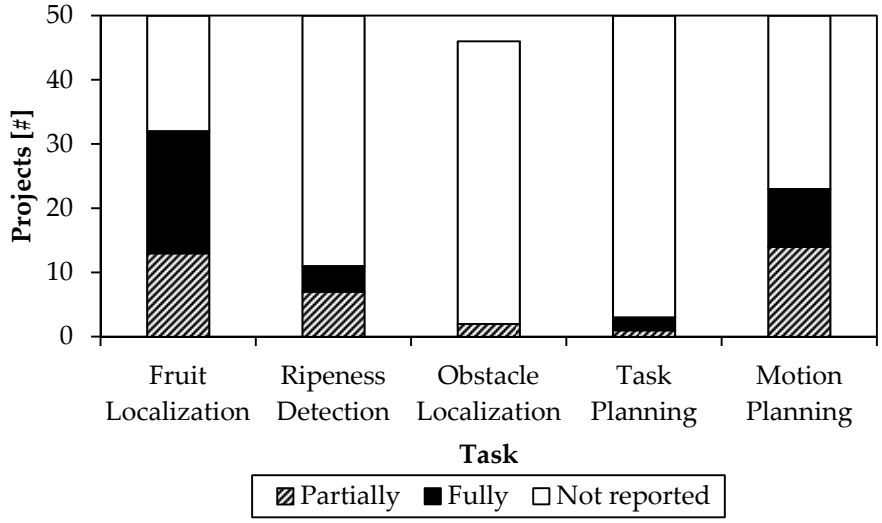


Figure 2-14. Reported algorithms for five major tasks, performed during fruit harvesting. If algorithms were reported, they were mostly partially reported and few authors fully reported algorithms

Figure 2-14 clearly shows that algorithms for fruit localization were reasonably reported – 64% (32/50). However, few authors reported which algorithms they used for ripeness detection – 22% (11/50), obstacle localization – 4% (2/46), task planning – 6% (3/50), and motion planning – 46% (23/50). Although some authors even distributed the description of algorithms over several articles, none of the authors reported algorithms for all five tasks.

Adaptive algorithms were used in 20% (10/50) of the projects. Whether the use of adaptive algorithms resulted in better performance is hard to conclude because the number of projects performed on a crop was too small or because performance indicators were missing. The only minor observation that can be made is that fruit localization success was highest for the two tomato harvesting projects where adaptive algorithms were used.

2.5 Discussion

Results reveal that performance of harvesting robots did not improve over the last three decades. We expected at least to observe a positive trend in performance because of two reasons. First, R&D advances in the field of sensors, mechatronics, computing power, computer vision, artificial intelligence and other robot domains might have contributed better hardware and software. Second, the number of projects increased from 7 to 17 and 26 projects during the three decades investigated (Figure 10) and we expected a learning effect from previous projects. The absence of increasing performance may partly elucidate why, to the best of our knowledge, none of the 50 harvesting robots reached a commercial stage.

When discussing how far harvesting robots are away from the required performance for implementation in practice, sparse data of average performance per crop (Table 2) complicate drawing conclusions. Exceptions include cucumber and orange harvesting. For cucumber harvesting, a cycle time of 10 s was proven to be economically feasible (Van Henten et al. (2002). The cycle time achieved was a factor of 12 too long (124 s) and clearly shows a gap must be bridged. For orange harvesting, comparing cycle time was possible for only one

project: 3 s required (Harrell 1987) vs. 3-7 s achieved (Harrell et al. 1990), i.e. a factor of about two too long. Although this gap is smaller, all performance indicators are required for a more conclusive analysis.

Despite the absence of increasing performance, we observed best practices (Brannan et al. 2008) that researchers may employ in future projects. In the following we discuss hardware design, reporting the design process, performing economic analysis, and testing the robot in the field. Sensing is already discussed by others (Kapach et al. 2012; Jiménez et al. 2000a). Regarding hardware, the design of the 3-DOF manipulator and end-effector combination of the kiwi harvesting robot seems outstanding, which is also clear from the low cycle time of 1 s. The manipulator was custom-made and low-cost (5% of an off-the-shelf anthropomorphic arm), thereby improving economic feasibility. The manipulator did not contain motors at each joint – enabled by applying four-bar mechanisms – and therefore prevented leaves and branches getting stuck behind motors. Furthermore, a pipe was mounted on the gripper for quick disposal of fruit in a storage bin (Flemmer et al. 2009). The melon harvesting robot (Edan et al. 2000) included the most complete description of the design process and performance indicators. The most thorough economic analysis was reported for orange harvesting (Harrell 1987) and considered 19 variables affecting economic feasibility. The best practice for field tests was conducted for the strawberry harvesting robot (Hayashi et al. 2010). Performance was evaluated for five different classes of fruit positions, providing a better understanding of the influence of the crop environment on harvest success, and many fruit (>1000) were evaluated. In addition, the effect of the gripping modes (suction on/off) and causes of unsuccessful harvesting attempts were quantitatively assessed.

To understand why performance is still limited (harvest success of 66% and a cycle time of 33 s), we attempt to extract bottlenecks limiting performance (ordered on descending priority):

1. The threefold variation in the crop environment renders fruit harvesting a hard task.
2. It seems few attempts were made to simplify the task, and therefore the requirements to be met remained very challenging.
3. Robot designs were probably not optimized with respect to the requirements because requirements were hardly defined and because design process techniques were hardly used.
4. Since requirements, testing conditions, and performance indicators were hardly reported, it was hard to determine which techniques, devices, or algorithms are successful as-is, or need improvement.
5. The limited description of hardware and software complicated determining best practices and therefore enforced researchers to “reinvent the wheel” (Brannan et al. 2008). As a result, researchers may not always have implemented state-of-the-art techniques, devices, and algorithms.

To address these bottlenecks, we propose future challenges in Section 2.6.

2.6 Future Challenges and R&D Directions

To address bottlenecks and realize a positive trend in performance, we propose three challenges with R&D directions (Sections 2.6.1-2.6.3). The first two challenges “simplifying

the task" and "enhancing the robot" provide solutions to improve performance. The third challenge "defining requirements and measuring performance" is a prerequisite to successfully implement the first and second challenges. The first challenge addresses bottlenecks 1 and 2, the second challenge addresses bottlenecks 1 and 3, and the third challenge addresses bottlenecks 4 and 5. Apart from a better performance, we believe more requirements must be met to successfully implement harvesting robots in practice, as discussed in the fourth challenge (Section 2.6.4).

Relevance of the challenge is explained at the start of a section and, subsequently, R&D directions are given. The identified challenges and R&D directions follow from the analysis of the variation in a crop environment (Section 2.2) and the literature review (Sections 2.4 and 2.5). Yet, they are additionally based on the authors' 25 years of experience with development of harvesting robots.

2.6.1 Simplifying the Task

Simplifying the task, by modifying the crop environment, helps to improve performance and should be investigated. This challenge relates to workspace design, a common practice when introducing robots into manufacturing systems (Nof 2007), which was also applied when introducing the milking robot (Halachmi et al. 2000). We propose several modifications based on the crop environment described in Section 2.2. Yet, probably more modifications can be identified in future.

2.6.1.1 Modified Cultivation Systems

To simplify the crop environment, the cultivation system can be modified. An example is the implementation of the high-wire cultivation system for cucumber. The high-wire system reduced occlusion drastically (Van Henten et al. 2002). For rose harvesting, plants were transported to a stationary robot by a movable cultivation system. This stationary robot allowed to implement parallel actuation mechanisms, which are known to be fast and accurate (Noordam et al. 2005).

Plant locations can be adapted in the row. For melon harvesting it has been proven that by alternating the plants along the row, robot speed can be increased (Edan and Miles 1993).

Trees or plants can be pruned to obtain a suitable plant geometry for robotic harvesting. The literature indicated that differences in tree geometry can influence the cycle times and travelling distance along ripe oranges in a tree (Edan et al. 1990). A tree can be pruned such that travelling distances between fruit are short and, hence, the cycle time is reduced.

Yet, growers might be reluctant to accept modifications because of the technical and financial risks of unproven cultivation systems. The researcher should therefore involve growers along development of a modified cultivation system (Section 2.6.4.4).

2.6.1.2 Cultivar Selection and Breeding

Selection of cultivars can be aimed at increasing the free workspace of the robot. An example is the cucumber harvester that performed much better on cultivars with a long peduncle. Investigation of convenient cultivar characteristics was suggested for cherry tomato harvesting as well (Kondo et al. 1996).

An even more interesting option would be to include requirements for robotic harvesting in the process of breeding a new cultivar. For instance, Dutch breeding companies indicated to us they are able to breed sweet-pepper cultivars that generate fruit at distributed positions, whereas currently fruit appear in clusters that are difficult to access and view. Plant breeders and roboticists seem to worked separately so far and collaboration may be beneficial for both disciplines, as also indicated in another review (Houle et al. 2010).

2.6.1.3 Supportive Mechanisms

Mechanisms that temporarily mitigate occlusion or temporarily increase the free workspace can be used. Examples are canopy volume reduction for oranges by a mechanism (Lee and Rosa 2006), leaf occlusion reduction for melon harvesting by an air blower (Edan et al. 2000), or mechanisms that pull or push neighboring plants aside.

2.6.1.4 Alternative Cultivation Practices

Additional operations such as fruit or flower thinning in sweet-pepper or apple cultivation can be applied to avoid fruit clusters. Fruit clusters typically cause occlusion of ripe fruit and therefore complicate fruit localization and detachment. Performing additional operations, however, increases costs and a trade-off should be made with revenues.

An option to reduce the peak demand for harvesting capacity is by model-based climate control. Different climate management can smooth yield peaks and as a result the required number of robots, i.e. the required investment, on a site might reduce. Such climate management techniques have been proposed for sweet-pepper cultivation (Buwalda et al. 2006; Van Henten et al. 2006a). Shifting plant dates over the season can also spread yield peaks, as is done in melon cultivation (Edan et al. 1992a). Shifting plant dates are also applied in cauliflower and broccoli cultivation.

2.6.2 Enhancing the Robot

In addition to the advantages of a simplified task, the robot itself can be enhanced to realize a positive trend in performance. Several options are discussed to improve the robot's design and ability to deal with the complex crop environment (Section 2.2). These options adhere to challenges discussed in the EU Strategic Research Agenda (euRobotics 2013) and Roadmap for US Robotics (CCC 2009).

2.6.2.1 Sensing, World Modeling, and Reasoning

Sensing is required to detect and localize the objects of interest. World modeling and reasoning manipulate, combine and enhance sensor output to improve actuation. In addition to sensors already used (RGB, infrared, and multi-spectral cameras, and laser range finders), an R&D challenge is to advance algorithms for novel sensors (e.g., hyperspectral, time-of-flight, light-field, laser triangulation, structured light, and chlorophyll fluorescence cameras). Additionally, multiple sensors can be fused to enhance performance of object detection and localization (Armada et al. 2014). Specific additional challenges for computer vision are described by Kapach et al. (2012). Given the large number of articles dealing with fruit detection (Jiménez et al. 2000a), it seems that researchers so far focused mainly on developing improved sensing.

Yet, better performance may be achieved by combining sensing, world modeling, reasoning and actuation. Such combinations may improve actions and awareness of the robot, i.e. better cognition (euRobotics 2013). Some examples of promising R&D directions employing such combinations are detailed in the following. Active perception (Bajcsy 1988) can be used to move the camera until visibility of the fruit improves. Visual servoing is already used for orange harvesting (Mehta and Burks 2014) and involves updating object locations while an end-effector – with a camera mounted – moves toward a fruit. In order to plan appropriate actions, sensor data from multiple sensor types can be fused into a world model containing 3D crop models (Weiss and Biber 2011) and other objects and behaviour in the environment. These crop models and other objects can be based on a-priori maps of a crop environment, similar as Google Street View that is used as a-priori map for the Google driverless car. Such maps can be obtained with aid of image parsing (Yao et al. 2010) or probabilistic approaches (Thrun et al. 2006; Hiremath et al. 2014). Other recent work regarding world modeling and reasoning include symbolic planning and control methods (Belta et al. 2007), fuzzy temporal logic (Lu et al. 2012), multi-agent systems, dynamic Bayesian networks, and Markov logic networks (Lavee et al. 2009).

2.6.2.2 Adaptation and Learning

To better deal with the complex crop environment, a robot can be equipped with adaptation and learning capabilities. A learning system not only adapts to changing conditions, but also learns and improves to better handle changes over time, and preferably learns on-line within a short time (Hagras et al. 2002). We therefore consider a learning system an advanced way to tackle the complex sensing and manipulation problems involved in fruit harvesting.

The literature indicates the potential of adaptive algorithms both for sensing (Edan et al. 2000; Nieuwenhuizen et al. 2010) and manipulation (Grift et al. 2008). As an example of the effect of an adaptive classifier, Nieuwenhuizen et al. (2010) detect potato plants that occur as weed in a sugar beet crop. Under changing natural light conditions, the classification accuracy improved from 34.9% (non-adaptive) to 67.7% (adaptive). Learning has been scarcely applied for agricultural robots, and we found examples regarding outdoor navigation (Hagras et al. 2002) and ground detection (Reina and Milella 2012; Reina et al. 2012). An R&D direction is to explore learning capabilities and another review (Ye et al. 2012) discusses possible learning techniques.

2.6.2.3 Human-robot Collaboration

Given the moderate harvest success (66%) and complex crop environment (Section 2.2), human skills may still be needed to enhance robot performance. By taking advantage of human perception skills and the robot's accuracy and consistency, a combined human-robot system may result in improved performance and lower costs compared with an autonomous robot. Limited work has been done regarding human-robot collaboration for agricultural robots (Murakami et al. 2008; Berenstein et al. 2012) but results indicate the potential of this direction due to the increase in performance, i.e. faster and higher detection rates (Bechar and Edan 2003). Furthermore, it can simplify the robot providing reduced costs and increased robustness.

2.6.2.4 Specialized Hardware

Developing specialized hardware can enhance robot performance and might even reduce cost. For instance, an increasing number of degrees-of-freedom (DOF) of a manipulator complicates motion planning, decreases manipulator speed and reliability, and increases costs. Despite these disadvantages, authors used manipulators with two to seven DOF, without further explanation of this particular design choice. Researchers prioritized use and design of custom-made platforms (74%), manipulators (78%) and end-effectors (98%), but more analysis tools can be used to optimize a custom-made design or to assess available off-the-shelf components.

The literature provides several tools for optimization of a manipulator (Jian et al. 2007; Li et al. 2008; Van Henten et al. 2009; Sivaraman and Burks 2006; Sakai et al. 2008). Systems engineering and modeling tools can help to address design issues regarding manipulation: type of arm (e.g., articulated, Cartesian), number of arms, required speed, and minimum required degrees-of-freedom (Edan and Miles 1994). To enhance and compare end-effectors, developers can use systems engineering or employ modeling techniques such as Finite Element Modeling (FEM) (Cardenas-Weber 1991; Edan et al. 1992b).

2.6.2.5 Exploring Alternative Robot Designs

To best fit the requirements, alternatives to the standard robot design, consisting of a travelling device with a single manipulator and an end-effector, should be explored. Robot designs such as snake robots, flying robots, spider robots and climbing robots were never applied in high-value crops. Such designs can be investigated to determine whether they fit the requirements better than the standard robot design. The suggestion for such robot designs is not based on pure speculation; an attempt has been undertaken to develop a climbing robot for tomato (Figure 2-15). The example indicates that development of climbing robots is a possible direction to pursue.

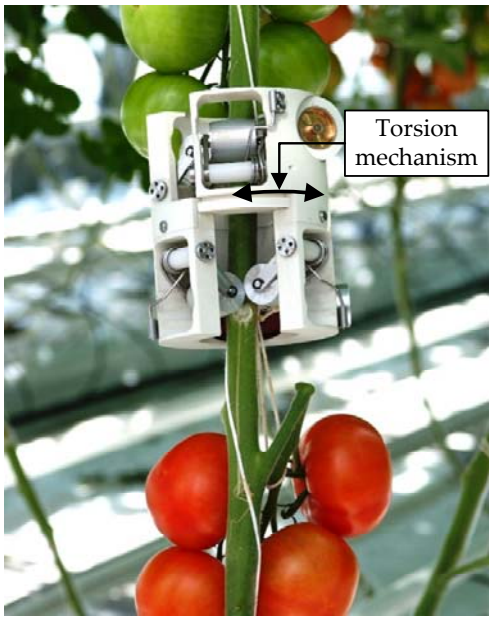


Figure 2-15. Functional model of a climbing robot for tomato. The torsion ability of the frame allows for peduncle or leaf avoidance during movement along the stem. Photograph courtesy of Bart van Tuijl, Wageningen UR

The choice for a certain robot design can be a rather difficult process and multiple disciplines should be involved to develop a successful design (Section 2.6.4).

2.6.3 Defining Requirements and Measuring Performance

Requirements must be defined to identify goals for performance indicators and, eventually, to measure if these goals are met. Performance indicators used (Section 2.3.1) can serve as benchmarks for development of harvesting robots in future projects. Benchmarking and performance comparison is important to foster research and to enable implementation in practice (Madhavan et al. 2009; Bonsignorio et al. 2009).

2.6.3.1 Defining Requirements

Requirements must be defined at the start of a project to define robot capabilities, to identify goals for performance indicators, and to determine research focus. Examples of requirements are: “the robot must be able to operate within a temperature range of 5 to 40° C”, "the robot must detect 90-95% of the targets with less than 10% false detections", "cycle time must be less than 6 s per fruit", “system cost must not exceed 100 k€”. Only a few authors reported requirements (Tillett 1993; Van Henten et al. 2002; Hwang and Kim 2003; Belforte et al. 2006; Foglia and Reina 2006).Reporting requirements therefore deserves attention in future projects.

Requirements must be defined for non-technical aspects as well, such as: economic, aesthetic, social, safety, ethical and legal aspects (Section 2.6.4). These aspects are inspired by the 15 “aspects of experience”, established by the Dutch philosopher Dooyeweerd (Clouser 2010).

Setting technical and economic requirements is a complicated interplay between variables such as cycle time, success rates, and system costs because these strongly interact. To illustrate, we demonstrate two economically feasible scenarios for sweet-pepper harvesting using an economics simulation model (Pekkeriet 2011). In the first scenario, maximum cost of the robot is 196 k€ if a cycle time of 6 s per fruit with 95% harvest success can be achieved for 20 hours per day; In the second scenario, maximum cost of the robot is 59 k€ if a cycle time of 20 s per fruit with 95% harvest success can be achieved for 20 hours per day. As a result, design choices strongly depend on the scenario selected.

2.6.3.2 Reporting Performance Indicators

Performance indicators must be reported to verify if requirements are met and to enable performance comparison. There was only one project in which all quantitative performance indicators were reported (Edan et al. 2000), whereas all performance indicators should be reported to enable statistically significant comparison of, for instance, the effect of production environment and crop. Especially false-positive detections and damage rate were hardly reported and therefore deserve special attention.

Additionally, authors should explicitly state what results were obtained, under which conditions (field/lab; weather conditions; autonomous operation; crop environment), and which hardware and software components were used. The results of the review clearly showed that this data was hardly reported, or only partially reported for 46% (33/71) of the reported algorithms. Once these data are reported, performance can be compared with respect

to requirements defined. If requirements are not met, developers might either adapt the requirements or try to improve robot performance.

2.6.3.3 Reporting Test Samples

Test samples should be reported to determine reliability of the performance indicator. In total 57% (56/98) of the reported performance indicators were supported by the number of test samples evaluated. It is unclear what number of test samples the other authors evaluated to calculate the reported performance indicators.

Units were lacking in 21% (21/98) of the reported performance indicators. The unit should be reported to compare performance with other projects. Standardization of units can be useful for comparison and units used in this review (Section 2.3.1) can serve as a starting point.

2.6.3.4 Testing the Robot Under a Broad Range of Conditions

The robot must be tested under a broad range of climate conditions and crop environments to demonstrate if required robot capabilities are met. The number of test samples evaluated varied from 11 to 2506. Hence, given the complexity of harvesting, one can doubt if reliability of a performance indicator based on 11 fruit samples can represent the robot performance for a common harvest day, in which a human picks up to several thousands of fruit, in a complex crop environment. Moreover, these test samples were probably evaluated on a specific hour on a day and did not include the weather variations along the day. Due to the large variation along and between seasons it is important to conduct long term tests. Evaluation methods should be developed to define the experimental setup and statistical analysis to ensure the performance indicators are not specific for the evaluated test cases (Cohen et al. 2006).

2.6.3.5 Defining Additional Performance Indicators

Additional performance indicators are required to validate requirements that are not covered by performance indicators in Section 2.3.1. For instance, performance indicators are yet to be defined to validate to what extent a robot damages the plant. Plant damage reduces vitality of the crop and consequently yield can drop or wounds might act as access points for diseases. Also, it is a fact from practice, that irrespective of quantifiable losses, farmers do not appreciate nor accept significant damage to the crop. In the literature, only Pool and Harrell (1991) reported that in 57 of the 154 pick attempts (37%) damage was caused to leaves or branches of the orange tree. Damages to the plant might be hard to quantify since a plant is capable of self-healing and a plant does not suffer from all damages.

Additional indicators can be reported for obstacle localization (Bac et al. 2013). Although obstacle localization received little attention (only 4% of the projects reported algorithms) it is required to avoid collisions that can damage the plant or construction elements.

2.6.4 Considering Additional Requirements for Successful Implementation

Simplifying the task and enhancing the robot can improve performance, but they are not sufficient for successful implementation of a harvesting robot in practice. Disciplines other than robotics must be involved in the development of the robot (Burks et al. 2005), to address

the following requirements: the robot must be technically capable to perform the task, economically feasible, safe, match the logistics processes, and must be accepted by growers and society. We believe all requirements must be met for successful implementation. Although most research focused on technical capability, as discussed previously (Sections 2.6.2-2.6.3), R&D directions for the other four requirements are discussed in the following sections.

To successfully involve all disciplines, reflexive design methods can be used. Such methods require all stakeholders (growers, engineers, scientists, economists) to be involved and assure needs of stakeholders are reflected in project objectives, requirements, functions, and working principles. Examples of reflexive design methods are reflexive interactive design (Bos et al. 2009), methodological design (Siers 2007) and engineering design (Cross 2008). The use of such design methods turned out to be effective in previous research (Van Henten et al. 2006b; Bakker 2009; Nieuwenhuizen 2009).

2.6.4.1 Economics

Incorporating economics is not only important to set performance requirements (Section 2.6.3.1), but can also provide economic incentives that enable implementation in practice. To analyze economics, existing tools can be used as a starting point (Harrell 1987; Edan et al. 1992a; Sarig 1993; Tillett 1993; Pedersen et al. 2006; Clary et al. 2007; Nof 2009). Economic analyses performed so far (four projects in Table 2-2) dealt with only performance requirements, but examples of additional economic incentives are human-robot co-working or adding features apart from the harvesting task.

An emerging R&D direction to investigate is human-robot co-working because 100% replacement of human labor does not seem technically feasible with current harvest success (66%) achieved. Japanese researchers currently investigate co-working for strawberry cultivation. Growers indicated to accept a harvest success rate of 60% if such rate improves economic feasibility by skipping complex harvest cases. Humans can harvest the remaining complex harvest cases (Hayashi, pers. comm.).

Examples of added features are tracking-and-tracing of food or better ripeness determination. Robots can log for each fruit where it grew and, in case of a disease outbreak, growers can quickly localize its origin and apply targeted crop protection treatments. A more equalized ripeness of harvested fruit will result in a longer shelf life of the fruit and probably a higher market value. As a result, higher prices can be generated if robots can determine ripeness better than humans, which may lower performance requirements of the robot. Chlorophyll fluorescence imaging is a possible candidate to measure ripeness accurately (Polder et al. 2004; Bron et al. 2004).

2.6.4.2 Logistics

Logistics must be analyzed to successfully integrate a robot in logistics of operations and harvested fruit. Current logistics are already complex (Van 't Ooster et al. 2014) and become even more complex if future robots can perform more tasks than harvesting. So far most robots were developed for only harvesting, with a few exceptions (Monta 1995; Van Henten et al. 2002; Van Henten et al. 2006b). Examples of tasks that can be added apart from

harvesting are chlorophyll fluorescence sensors for disease detection (Gorbe and Calatayud 2012) or sorting and quality assessment of fruit directly after harvesting (Qiao et al. 2005). Developing such a multi-operational robot might even save development efforts because a platform and manipulator has to be developed only once.

2.6.4.3 Safety

Safe operation of robots is necessary to avoid harming humans and is critical for actual implementation (Pedersen et al. 2008). If humans (Section 2.6.2.3) or other robots can appear in the robot workspace, safety becomes even more important. Safety seems an unexplored field for harvesting robots. The only work known dealing with safe manipulators is an article by Vermeulen & Wisse (2008). R&D directions for safety are provided in the Multi-Annual Roadmap for Robotics in Europe (euRobotics 2014)

2.6.4.4 Acceptance by Growers and Society

Involving growers in the development is not only important for reflections on the design, but also to investigate acceptance to alternative cultivation systems or logistics, and to identify social aspects. A clear example for which non-technical aspects were relevant for implementation, is the milking robot. An incentive to purchase a robot was the flexible work schedule enabled by the release from fixed milking hours (Mathijs 2004).

Acceptance by society is relevant for successful implementation as well and seems an unexplored field for harvesting robots. Society may have reservations to such innovations and, in addition, employees may reject co-working with a robot (Section 2.6.4.1).

2.7 Conclusion

The first contribution of this review involves identifying and elucidating three sources of variation in a crop environment that must be considered in development of a harvesting robot: variation in objects, environment and crops. Consequently, this variation renders fruit harvesting a threefold complex task, and is the main bottleneck to better performance. Approaches to overcome these unique variations are proposed.

The second contribution is a quantitative review of the literature on developed harvesting robots, which revealed that 50 distinct projects were performed in the last 30 years. The quantitative method resulted in the following findings. On average localization success was 85%, detachment success was 75 %, harvest success was 66%, fruit damage was 5%, peduncle damage was 45%, and cycle time was 33 s. Unfortunately, data were too sparse to provide averages per crop, which would have been more meaningful. Nevertheless, these results indicate the need for continued R&D to improve performance. Systematic design methods (12%) and economic analyses (8%) were scarcely used. Most manipulators were designed with two to seven degrees-of-freedom (DOF), where three DOF were common. Platforms (74%), manipulators (78%), and end-effectors (98%) were mostly custom-made, whereas sensors were taken off-the-shelf (98%). Algorithms were poorly reported for the five main robot harvesting tasks: fruit localization (64%), ripeness detection (22%), obstacle localization (4%), task planning (6%) and motion planning (46%). Moreover, if algorithms were reported they were mostly partially reported. Adaptive algorithms were used in 20% of the projects. The kiwi harvesting robot developed in New Zealand reached the shortest cycle

time (1 s) probably because a low-cost and specialized manipulator and end-effector were developed and because fruit were easily accessible and hard to bruise. Cycle time was a factor of 12 too long for cucumber and a factor of about two for orange, compared with required cycle time. Furthermore, the comparisons conducted on data of the review indicate performance did not improve over the last three decades. Data should be re-analyzed once more projects are available to supply statistically significant trends. We identified best practices that researchers can employ in future projects, and five bottlenecks that limit performance of current harvesting robots.

Our final contribution involves four future challenges combined with R&D directions to improve performance and implement harvesting robots in practice. The first two challenges – simplifying the task and enhancing the robot – provide solutions to improve performance. As a prerequisite, the third challenge must be implemented: defining requirements and measuring performance. The list of performance indicators provided in this paper could be used as benchmarks in future development of harvesting robots, and should be obtained from extensive field tests that include a wide range of conditions. The fourth challenge lists additional requirements that must be considered to successfully implement harvesting robots in practice. Robots will be needed more and more to optimize plant maintenance and, as a result, to contribute to production of high-quality food for an increasing world population.

Acknowledgements

We thank rose growers Thijs and Arie van den Berg and a sweet-pepper grower, who preferred to remain anonymous, for the labor registration data they made available to us. This research was funded by the European Commission in the 7th Framework Programme (CROPS GA no. 246252) and by the Dutch horticultural product board (PT no. 14555). This research was partially supported by Ben-Gurion University of the Negev, Paul Ivanier Center for Robotics Research and Production Management, and the Rabbi W. Gunther Plaut Chair in Manufacturing Engineering.

References

Aljanobi, A. A., Al-hamed, S. A., & Al-Suhaibani, S. A. (Robotic System). A Setup Of Mobile Robotic Unit For Fruit Harvesting. In *IEEE 19th International Workshop on Robotics in Alpe-Adria-Danube Region (RAAD), Budapest, Hungary, June 2010* (pp. 105-108).

Angeles, J., Park, F. C., Siciliano, B., & Khatib, O. (2008). Ch. 10: Performance Evaluation and Design Criteria. In B. Siciliano, & O. Khatib (Eds.), *Springer Handbook of Robotics* (pp. 229-244): Springer Berlin Heidelberg.

Antonelli, M. G., Auriti, L., Beomonte Zobel, P., & Raparelli, T. (2011). Development of a new harvesting module for saffron flower detachment. *Romanian Review Precision Mechanics, Optics and Mechatronics*(39), 163-168.

Arima, S., & Kondo, N. (1999). Cucumber Harvesting Robot and Plant Training System. *Journal of Robotics and Mechatronics, 11*(3), 208-212.

Armada, M. A., Sanfeliu, A., Ferre, M., Fernández, R., Salinas, C., Montes, H., et al. Validation of a Multisensory System for Fruit Harvesting Robots in Lab Conditions. In *ROBOT2013: First Iberian Robotics Conference, 2014* (Vol. 252, pp. 495-504).

Arndt, G., Rudziejewski, R., Stewart, V.A. (1997). On the future of automated selective asparagus harvesting technology. *Computers and Electronics in Agriculture, 16*, 137-145.

Bac, C. W., Hemming, J., & Van Henten, E. J. (2013). Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper. *Computers and Electronics in Agriculture, 96*, 148-162.

Baeten, J., Donné, K., Boedrij, S., Beckers, W., & Claesen, E. (2008). Autonomous fruit picking machine: A robotic apple harvester. *Springer Tracts in Advanced Robotics, 42*, 531-539.

Bajcsy, R. (1988). Active perception. *Proceedings of the IEEE, 76*(8), 996-1005.

Bakker, T. (2009). *An autonomous robot for weed control : design, navigation and control*. PhD. Thesis, Wageningen University, Wageningen.

Balerin, S., Bourly, A., & Sévila, F. Mobile robotics applied to fruit harvesting: the case of greenhouse tomatoes. In *Automated Agriculture for the 21st Century 1991 Symposium, Chicago, USA, December 1991* (pp. 236-244).

Bechar, A. (2010). Robotics in horticultural field production. *Stewart Postharvest Review, 6*(3), 1-11.

Bechar, A., & Edan, Y. (2003). Human-robot collaboration for improved target recognition of agricultural robots. *Industrial Robot, 30*(5), 432-436.

Belforte, G., Deboli, R., Gay, P., Piccarolo, P., & Ricauda Aimonino, D. (2006). Robot Design and Testing for Greenhouse Applications. *Biosystems Engineering, 95*(3), 309-321.

Belta, C., Bicchi, A., Egerstedt, M., Frazzoli, E., Klavins, E., & Pappas, G. J. (2007). Symbolic planning and control of robot motion [Grand Challenges of Robotics]. *IEEE Robotics & Automation Magazine, 14*(1), 61-70.

Berenstein, R., Ben Halevi, I., & Edan, Y. A remote interface for a human-robot cooperative sprayer. In *Proceedings of the 11th International Conference on Precision Agriculture, Indianapolis, IN, July 2012*.

Bonsignorio, F., Hallam, J., & del Pobil, A. Defining the requisites of a replicable robotics experiment. In *IROS 2009 Workshop on Performance Evaluation and Benchmarking for Next Intelligent Robots and Systems, St. Louis, MO, 11-15 October 2009*.

Bos, A. P., Groot Koerkamp, P. W. G., Gosselink, J. M. J., & Bokma, S. (2009). Reflexive interactive design and its application in a project on sustainable dairy husbandry systems. *Outlook on Agriculture, 38*(2), 137-145.

Brannan, T., Durose, C., John, P., & Wolman, H. (2008). Assessing best practice as a means of innovation. *Local Government Studies, 34*(1), 23-38.

Bron, U. I., Ribeiro, R. V., Azzolini, M., Jacomino, A. P., & Machado, E. C. (2004). Chlorophyll fluorescence as a tool to evaluate the ripening of Golden' papaya fruit. *Postharvest Biology and Technology, 33*(2), 163-173.

Buemi, F., Massa, M., Sandini, G., & Costi, G. (1996). The AGROBOT project. *Advances in Space Research, 18*(1-2), 185-189.

Burks, T., Villegas, F., Hannan, M., Flood, S., Sivaraman, B., Subramanian, V., et al. (2005). Engineering and horticultural aspects of robotic fruit harvesting: Opportunities and constraints. *HortTechnology, 15*(1), 79-87.

Buwalda, F., Van Henten, E. J., De Gelder, A., Bontsema, J., & Hemming, J. (2006). Toward an optimal control strategy for sweet pepper cultivation - 1. A dynamic crop model. *Acta Horticulturae* (Vol. 718, pp. 367-374).

Cardenas-Weber, M. C., Stroshine, R.L., Haghighi, K., Edan, Y. (1991). Melon material properties and finite element analysis of melon compression with application to robot gripping. *Transactions of the ASAE, 34*(3), 920-929.

CCC (2009). *A Roadmap for US Robotics: From Internet to Robotics*. Computing Community Consortium study on Robotics, Snobird, UT.

Ceres, R., Pons, J. L., Jiménez, A. R., Martín, J. M., & Calderón, L. (1998). Design and implementation of an aided fruit-harvesting robot (Agribot). *Industrial Robot, 25*(5), 337-346.

Chatzimichali, A. P., Georgilas, I. P., & Tourassis, V. D. (Robotic System). Design of an advanced prototype robot for white asparagus harvesting. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics, July 2009* (pp. 887-892).

Clary, C. D., Ball, T., Ward, E., Fuchs, S., Durfey, J. E., Cavalieri, R. P., et al. (2007). Performance and economic analysis of a selective asparagus harvester. *Applied Engineering in Agriculture, 23*(5), 571-577.

Clouser, R. (2010). A brief sketch of the philosophy of Herman Dooyeweerd. *Axiomathes, 20*(1), 3-17.

Cohen, O., Edan, Y., & Schechtman, E. (2006). Statistical evaluation method for comparing grid map based sensor fusion algorithms. *International Journal of Robotics Research, 25*(2), 117-133.

Cross, N. (2008). *Engineering Design Methods* (4th Revised Edition ed.): John Wiley and Sons Ltd.

De-An, Z., Jidong, L., Wei, J., Ying, Z., & Yu, C. (2011). Design and control of an apple harvesting robot. *Biosystems Engineering, 110*(2), 112-122.

Edan, Y. (1995). Design of an autonomous agricultural robot. *Applied Intelligence, 5*(1), 41-50.

Edan, Y., Benady, M., & Miles, G. E. Economic analysis of robotic melon harvesting. In *The ASAE 1992 International Winter Meeting, Nashville, TN, 1992a* (Vol. ASAE paper no. 92-1512).

Edan, Y., Flash, T., Shmulevich, I., Sarig, Y., & Peiper, U. M. (1990). An algorithm defining the motions of a citrus picking robot. *Journal of Agricultural Engineering Research, 46*(C), 259-273.

Edan, Y., Haghighi, K., Stroshine, R., & Cardenas-Weber, M. (1992b). Robot gripper analysis: finite element modeling and optimization. *Applied Engineering in Agriculture, 8*(4), 563-570.

Edan, Y., & Miles, G. E. (1993). Design of an agricultural robot for harvesting melons. *Transactions of the ASAE, 36*(2), 593-603.

Edan, Y., & Miles, G. E. (1994). Systems engineering of agricultural robot design. *IEEE Transactions on Systems, Man and Cybernetics, 24*(8), 1259-1265.

Edan, Y., Rogozin, D., Flash, T., & Miles, G. E. (2000). Robotic melon harvesting. *IEEE Transactions on Robotics and Automation, 16*(6), 831-835.

euRobotics (2013). Strategic Research Agenda for Robotics in Europe. Draft 0v42.

euRobotics (2014). Robotics 2020 Multi-Annual Roadmap. Initial Release B 15/01/2014.

Feng, Q., Zheng, W., Qiu, Q., Jiang, K., & Guo, R. (Robotic System). Study on strawberry robotic harvesting system. In *IEEE International Conference on Computer Science and Automation Engineering (CSAE), 25-27 May 2012* (Vol. 1, pp. 320-324).

Flemmer, C., Flemmer, R., & Scarfe, A. Robotics in Horticulture. In *Proceedings of the Australia and New Zealand Avocado Growers Conference, Cairns, Australia, July 2009*.

Foglia, M. M., & Reina, G. (2006). Agricultural robot for radicchio harvesting. *Journal of Field Robotics, 23*(6-7), 363-377.

Glancey, J. L., & Kee, W. E. (2005). Engineering aspects of production and harvest mechanization for fresh and processed vegetables. *HortTechnology, 15*(1), 76-79.

González, R., Rodríguez, F., Sánchez-Hermosilla, J., & Donaire, J. G. (2009). Navigation techniques for mobile robots in greenhouses. *Applied Engineering in Agriculture, 25*(2), 153-165.

Gorbe, E., & Calatayud, A. (2012). Applications of chlorophyll fluorescence imaging technique in horticultural research: A review. *Scientia Horticulturae, 138*, 24-35.

Grand d'Esnon, A., Pellenc, R., Rabatel, G., Journeau, A., & Aldon, M. J. Magali: a self propelled robot to pick apples. In *ASAE Annual International Meeting, Baltimore, MD, 1987*.

Grift, T. E., Zhang, Q., Kondo, N., & Ting, K. C. (2008). A Review of Automation and Robotics for the Bio-industry. *Journal of Biomechatronics Engineering, Vol. 1*(1), 37-54.

Guo, F., Cao, Q., & Masateru, N. (2008). Fruit detachment and classification method for strawberry harvesting robot. *International Journal of Advanced Robotic Systems, 5*(1), 41-48.

Hagras, H., Colley, M., Callaghan, V., & Carr-West, M. (2002). Online Learning and Adaptation of Autonomous Mobile Robots for Sustainable Agriculture. *Autonomous Robots, 13*(1), 37-52.

Halachmi, I., Metz, J. H. M., Maltz, E., Dijkhuizen, A. A., & Speelman, L. (2000). Designing the optimal robotic milking barn, part 1: Quantifying facility usage. *Journal of Agricultural Engineering Research, 76*(1), 37-49.

Han, K.-S., Kim, S.-C., Lee, Y.-B., Kim, S.-C., Im, D.-H., Choi, H.-K., et al. (2012). Strawberry Harvesting Robot for Bench-type Cultivation. *Journal of Biosystems Engineering, 37*(1), 65-74.

Harrell, R. (1987). Economic analysis of robotic citrus harvesting in Florida. *Transactions of the American Society of Agricultural Engineers, 30*(2), 298-304.

Harrell, R. C., Adsit, P. D., Munilla, R. D., & Slaughter, D. C. (1990). Robotic picking of citrus. *Robotica, 8*(4), 269-278.

Hayashi, S., Ganno, K., Ishii, Y., & Tanaka, I. (2002). Robotic harvesting system for eggplants. *Japan Agricultural Research Quarterly, 36*(3), 163-168.

Hayashi, S., Saito, S., Iwasaki, Y., Yamamoto, S., Nagoya, T., & Kano, K. (2011). Development of circulating-type movable bench system for strawberry cultivation. *Japan Agricultural Research Quarterly, 45*(3), 285-293.

Hayashi, S., & Sakaue, O. Tomato Harvesting by Robotic System. In *ASAE Annual International Meeting, Phoenix, AZ, 1996*.

Hayashi, S., Shigematsu, K., Yamamoto, S., Kobayashi, K., Kohno, Y., Kamata, J., et al. (2010). Evaluation of a strawberry-harvesting robot in a field test. *Biosystems Engineering, 105*(2), 160-171.

Hiremath, S. A., van der Heijden, G. W. A. M., van Evert, F. K., Stein, A., & Ter Braak, C. J. F. (2014). Laser range finder model for autonomous navigation of a robot in a maize field using a particle filter. *Computers and Electronics in Agriculture, 100*, 41-50.

Houle, D., Govindaraju, D. R., & Omholt, S. (2010). Phenomics: The next challenge. *Nature Reviews Genetics, 11*(12), 855-866.

Humburg, D. S., & Reid, J. F. (1991). Field performance of machine vision for the selective harvest of asparagus. *SAE (Society of Automotive Engineers) Transactions, 100*(Sect 2), 81-92.

Hwang, H., & Kim, S. (Robotic System). Development of multi-functional tele-operative modular robotic system for greenhouse watermelon. In *Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2003), Piscataway, NJ, 2003* (Vol. vol.2, pp. 1344-1349).

Irie, N., Taguchi, N., Horie, T., & Ishimatsu, T. Asparagus harvesting robot coordinated with 3-D vision sensor. In *Proceedings of the IEEE International Conference on Industrial Technology, Gippsland, Australia, 2009* (pp. 1-6).

Jian, S., Xueyan, S., Tiezhong, Z., Bin, Z., & Liming, X. (2007). Design optimisation and simulation of structure parameters of an eggplant picking robot. *New Zealand Journal of Agricultural Research, 50*(5), 959-964.

Jiménez, A. R., Ceres, R., & Pons, J. L. (2000a). A survey of computer vision methods for locating fruit on trees. *Transactions of the American Society of Agricultural Engineers, 43*(6), 1911-1920.

Jiménez, A. R., Ceres, R., & Pons, J. L. (2000b). A vision system based on a laser range-finder applied to robotic fruit harvesting. *Machine Vision and Applications, 11*(6), 321-329.

Jovicich, E., Cnatliffe, D. J., Sargent, S. A., & Osborne, L. S. (2004). Production of Greenhouse-Grown Peppers in Florida. Gainesville, FL: University of Florida, IFAS Extension.

Jukema, G., & Van de Meer, R. (2009). Labor costs in arable farming and greenhouse horticulture; in Dutch: Arbeidskosten in de akkerbouw en glastuinbouw. The Hague, The Netherlands: Landbouw Economisch Instituur (LEI).

Kapach, K., Barnea, E., Mairon, R., Edan, Y., & Ben-Shahar, O. (2012). Computer Vision for Fruit Harvesting Robots - State of the Art and Challenges Ahead. *Int. J. of Computational Vision and Robotics, 3*(1/2), 4-34.

Kassay, L., & Slaughter, D. C. Electrical, Mechanical and Electronic Problems With the Hungarian Apple Harvester Robot During the 1992 Fall Field Tests. In *ASAE International Summer Meeting, Spokane, WA, June 1993* (pp. 1-16).

Kawamura, N., Namikawa, K., & Fujiura, T. Study on Agricultural Robot (II). In *Proceedings of the 7th Vehicle Automation Symporium JAACE, 1984* (pp. p. 35-38).

Kitamura, S., & Oka, K. (Vision). Recognition and cutting system of sweet pepper for picking robot in greenhouse horticulture. In *IEEE International Conference on Mechatronics and Automation, Niagra Falls, ON, July 2005* (Vol. 4, pp. 1807-1812).

Kondo, N., Nishitsuji, Y., Ling, P. P., & Ting, K. C. (1996). Visual feedback guided robotic cherry tomato harvesting. *Transactions of the ASAE, 39*(6), 2331-2338.

Kondo, N., Yamamoto, K., Yata, K., & Kurita, M. (Robotic System). A Machine Vision for Tomato Cluster Harvesting Robot. In *ASAE Annual International Meeting, Providence, RI, June 2008*.

Lavee, G., Rivlin, E., & Rudzsky, M. (2009). Understanding video events: A survey of methods for automatic interpretation of semantic occurrences in video. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, 39*(5), 489-504.

Lee, B. S. H., & Rosa, U. A. (2006). Development of a canopy volume reduction technique for easy assessment and harvesting of valencia citrus fruits. *Transactions of the ASABE, 49*(6), 1695-1703.

LEI, & CBS (2009). Statistics of Agriculture and Horticulture in the Netherlands; in Dutch: Land- en tuinbouwcijfers 2009. (pp. 23). The Hague, The Netherlands.

Lewis, A., Watts, P. L., & Nagpal, B. K. (1983). Investment Analysis for Robotic Applications. *Technical Paper - Society of Manufacturing Engineers. MS*.

Li, B., Vigneault, C., & Wang, N. (2010). Research development of fruit and vegetable harvesting robots in China. *Stewart Postharvest Review, 6*(3), 1-8.

Li, M., Imou, K., Wakabayashi, K., & Yokoyama, S. (2009). Review of Research on Agricultural Vehicle Autonomous Guidance. *International Journal of Agricultural and Biological Engineering, 2*(3), 1-16.

Li, P., Lee, S.-h., & Hsu, H.-Y. (2011). Review on fruit harvesting method for potential use of automatic fruit harvesting systems. *Procedia Engineering, 23*, 351-366.

Li, Z., Liu, J., Li, P., & Li, W. (Systematic Design). Analysis of workspace and kinematics for a tomato harvesting robot. In *International Conference on Intelligent Computation Technology and Automation, ICICTA 2008, Changsha, Hunan, China, October 2008* (Vol. 1, pp. 823-827).

Ling, P. P., Ehasani, R., Ting, K. C., Chi, Y., Ramalingam, N., Klingman, M. H., et al. (Robotic System). Sensing and End-Effector for a Robotic Tomato Harvester. In *ASAE Annual International Meeting, Ottawa, ON, August 2004*.

Liu, T. H., Zeng, X. R., & Ke, Z. H. (Robotic System). Design and prototyping a harvester for litchi picking. In *4th International Conference on Intelligent Computation Technology and Automation, ICICTA 2011, Shenzhen, China, March 2011* (Vol. 2, pp. 39-42).

Lu, Z., Augusto, J., Liu, J., Wang, H., & Aztiria, A. (2012). A system to reason about uncertain and dynamic environments. *International Journal on Artificial Intelligence Tools, 21*(5).

Madhavan, R., Lakaemper, R., & Kalmár-Nagy, T. Benchmarking and standardization of intelligent robotic systems. In *2009 International Conference on Advanced Robotics, ICAR 2009, 2009*.

Mathijs, E. Socio-economic impacts of automatic milking. In *The International Symposium on Automatic Milking, Lelystad, The Netherlands, March 2004*.

Mehta, S. S., & Burks, T. F. (2014). Vision-based control of robotic manipulator for citrus harvesting. *Computers and Electronics in Agriculture, 102*(0), 146-158.

Monta, M., Kondo, N., Shibano, Y. Agricultural robot in grape production system. In *The IEEE International Conference on Robotics and Automation, IEEE International Conference on Robotics and Automation, May 1995* (Vol. 3, pp. 2054-2509).

Murakami, N., Ito, A., Will, J. D., Steffen, M., Inoue, K., Kita, K., et al. (2008). Development of a teleoperation system for agricultural vehicles. *Computers and Electronics in Agriculture, 63*(1), 81-88.

Muscato, G., Prestifilippo, M., Abbate, N., & Rizzuto, I. (2005). A prototype of an orange picking robot: Past history, the new robot and experimental results. *Industrial Robot, 32*(2), 128-138.

Namikawa, K., & Ogawa, Y. (1989). Study on the hand of fruits harvesting robot. *Kansai Branch Report of the Japanese Society of Agricultural Machinery, 65*, 38-41.

Nieuwenhuizen, A. T. (2009). *Automated detection and control of volunteer potato plants*. PhD. Thesis, Wageningen University, Wageningen.

Nieuwenhuizen, A. T., Hofstee, J. W., & van Henten, E. J. (2010). Adaptive detection of volunteer potato plants in sugar beet fields. *Precision Agriculture, 11*(5), 433-447.

Nof, S. Y. (2007). Robot Ergonomics: Optimizing Robot Work. In S. Y. Nof (Ed.), *Handbook of Industrial Robotics*. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Nof, S. Y. (2009). *Springer Handbook of Automation* (Springer Handbook of Automation): Springer Berlin Heidelberg.

Noordam, J. C., Hemming, J., van Heerde, C., Golbach, F., van Soest, R., & Wekking, E. (2005). Automated Rose Cutting in Greenhouses with 3D Vision and Robotics: Analysis of 3D Vision Techniques for Stem Detection. *Acta Hort. (ISHS), 691*, 885-892.

Pedersen, S. M., Fountas, S., & Blackmore, S. (2008). Ch. 21: Agricultural Robots – Applications and Economic Perspectives In Y. Takahashi (Ed.), *Service Robot Applications*: InTech.

Pedersen, S. M., Fountas, S., Have, H., & Blackmore, B. S. (2006). Agricultural robots - System analysis and economic feasibility. *Precision Agriculture, 7*(4), 295-308.

Pekkeriet, E. J. (2011). CROPS project deliverable 12.1: Economic viability for each application. Wageningen, The Netherlands: Wageningen UR Greenhouse Horticulture.

Plá, F., Juste, F., & Ferri, F. (1993). Feature extraction of spherical objects in image analysis: an application to robotic citrus harvesting. *Computers and Electronics in Agriculture, 8*, 57-72.

Plebe, A., & Grasso, G. (2001). Localization of spherical fruits for robotic harvesting. *Machine Vision and Applications, 13*(2), 70-79.

Polder, G., Van der Heijden, G. W. A. M., Van der Voet, H., & Young, I. T. (2004). Measuring surface distribution of carotenes and chlorophyll in ripening tomatoes using imaging spectrometry. *Postharvest Biology and Technology, 34*(2), 117-129.

Pool, T. A., & Harrell, R. C. (1991). An end-effector for robotic removal of citrus from the tree. *Transactions of the ASAE, 34*(3), 373-378.

Qiao, J., Sasao, A., Shibusawa, S., Kondo, N., & Morimoto, E. (2005). Mapping yield and quality using the mobile fruit grading robot. *Biosystems Engineering, 90*(2), 135-142.

Rath, T., & Kawollek, M. (2009). Robotic harvesting of Gerbera Jamesonii based on detection and three-dimensional modeling of cut flower pedicels. *Computers and Electronics in Agriculture, 66*(1), 85-92.

Reed, J. N., Miles, S. J., Butler, J., Baldwin, M., & Noble, R. (2001). Automatic mushroom harvester development. *Journal of Agricultural Engineering Research, 78*(1), 15-23.

Reed, J. N., Tillet, R.D. (1994). Initial experiments in robotic mushroom harvesting. *Mechatronics, 4*(3), 265-279.

Reina, G., & Milella, A. (2012). Towards autonomous agriculture: Automatic ground detection using trinocular stereovision. *Sensors, 12*(9), 12405-12423.

Reina, G., Milella, A., & Underwood, J. (2012). Self-learning classification of radar features for scene understanding. *Robotics and Autonomous Systems, 60*(11), 1377-1388.

Saeys, W., & Nguyen, T. T. (2012). CROPS project deliverable 6.3: Report on the plantation and pruning methods to have fruit readily suited for automated harvesting. Leuven, Belgium: KU Leuven.

Sakai, S., Iida, M., Osuka, K., & Umeda, M. (2008). Design and control of a heavy material handling manipulator for agricultural robots. *Autonomous Robots, 25*(3), 189-204.

Sarig, Y. (1993). Robotics of Fruit Harvesting: A State-of-the-art Review. *Journal of Agricultural Engineering Research, 54*(4), 265-280.

Shimizu, H., Saito, Y., Nakashima, H., Miyasaka, J., & Ohdoi, K. Light environment optimization for lettuce growth in plant factory. In *IFAC Proceedings Volumes (IFAC-PapersOnline), 2011* (PART 1 ed., Vol. 18, pp. 605-609).

Siers, F. J. (2007). *Methodological Design; in Dutch: Methodisch Ontwerpen*. Groningen, The Netherlands: Wolters-Noordhoff.

Sites, P. W., & Delwiche, M. J. (1988). Computer Vision to Locate Fruit on a Tree. *Transactions of the ASAE, 31*(1), 257-263, 272.

Sittichareonchai, A., Khaorapapong, T., & Limsiloratana, S. Oil Palm Harvesting Robot: Automatic Manipulator Control. In *International Symposium on Automation and Robotics in Bioproduction and Processing, ARBIP95, Kobe, Japan, November 1995* (pp. 65-70).

Sivaraman, B., & Burks, T. F. (2006). Geometric performance indices for analysis and synthesis of manipulators for robotic harvesting. *Transactions of the ASABE, 49*(5), 1589-1597.

Tang, X., Zhang, T., Liu, L., Xiao, D., & Chen, Y. (Robotic System). A new robot system for harvesting cucumber. In *American Society of Agricultural and Biological Engineers Annual International Meeting 2009, Reno,NV, June 2009* (Vol. 6, pp. 3873-3885).

Tanigaki, K., Fujiura, T., Akase, A., & Imagawa, J. (2008). Cherry-harvesting robot. *Computers and Electronics in Agriculture, 63*(1), 65-72.

Temu, A. E., & Temu, A. A. High Value Agricultural Products for Smallholder Markets in Sub-Saharan Africa: Trends, Opportunities and Research Priorities. In *High Value Agricultural Products Workshop, Cali, Colombia, October 2005*.

Thrun, S., Burgard, W., & Fox, D. (2006). *Probabilistic Robotics*. Cambridge, MA: MIT Press.

Tillett, N. D. (1993). Robotic manipulators in horticulture: a review. *Journal of Agricultural Engineering Research, 55*(2), 89-105.

Tokuda, M., K., N., Suguri, M., Umeda, M., & Iida, M. Development of watermelon harvesting robot. In *International Symposium on Automation and Robotics in Bioproduction and Processing, ARBIP95, Kobe, Japan, November 1995* (Vol. 2, pp. 9-24).

Toussaint, K., Pouliot, N., & Montambault, S. (2009). Transmission line maintenance robots capable of crossing obstacles: State-of-the-art review and challenges ahead. *Journal of Field Robotics, 26*(5), 477-499.

Van 't Ooster, A., Bontsema, J., Van Henten, E. J., & Hemming, S. (2014). Simulation of harvest operations in a static rose cultivation system. *Biosystems Engineering, 120*, 34-46.

Van Henten, E. J. (2006). Greenhouse Mechanization: State of the Art and Future Perspective. *Acta Horticulturae*(710), p. 55-69.

Van Henten, E. J., Buwalda, F., De Zwart, H. F., De Gelder, A., Hemming, J., & Bontsema, J. (2006a). Toward an optimal control strategy for sweet pepper cultivation - 2. Optimization of the yield pattern and energy efficiency. *Acta Horticulturae* (Vol. 718, pp. 391-398).

Van Henten, E. J., Hemming, J., Van Tuijl, B. A. J., Kornet, J. G., Meuleman, J., Bontsema, J., et al. (2002). An Autonomous Robot for Harvesting Cucumbers in Greenhouses. *Autonomous Robots, 13*(3), 241-258.

Van Henten, E. J., Van't Slot, D. A., Hol, C. W. J., & Van Willigenburg, L. G. (2009). Optimal manipulator design for a cucumber harvesting robot. *Computers and Electronics in Agriculture, 65*(2), 247-257.

Van Henten, E. J., Van Tuijl, B. A. J., Hemming, J., Kornet, J. G., Bontsema, J., & Van Os, E. A. (2003). Field Test of an Autonomous Cucumber Picking Robot. *Biosystems Engineering, 86*(3), 305-313.

Van Henten, E. J., Van Tuijl, B. A. J., Hoogakker, G. J., Van Der Weerd, M. J., Hemming, J., Kornet, J. G., et al. (2006b). An Autonomous Robot for De-leafing Cucumber Plants grown in a High-wire Cultivation System. *Biosystems Engineering, 94*(3), 317-323.

Vermeulen, M. M. A., & Wisse, M. (2008). Maximum allowable manipulator mass based on cycle time, impact safety and pinching safety. *Industrial Robot, 35*(5), 410-420.

Wan Ishak, W. I., Kit, W. H., & Awal, M. A. (2010). Design and Development of Eggplant Harvester for Gantry System. *Pertanika Journal of Science and Technology, 18*(2), 231-242.

Weiss, U., & Biber, P. (2011). Plant detection and mapping for agricultural robots using a 3D LIDAR sensor. *Robotics and Autonomous Systems, 59*(5), 265-273.

Yao, B. Z., Yang, X., Liang, L., Mun Wai, L., & Song-Chun, Z. (2010). I2T: Image Parsing to Text Description. *Proceedings of the IEEE, 98*(8), 1485-1508.

Ye, J., Dobson, S., & McKeever, S. (2012). Situation identification techniques in pervasive computing: A review. *Pervasive and Mobile Computing, 8*(1), 36-66.

Chapter 3 – Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper

C.W. Bac[a,b], J. Hemming[a], E.J. van Henten[a,b]

Affiliations:
[a] Wageningen UR Greenhouse Horticulture, Wageningen University and Research Centre
[b] Farm Technology Group, Wageningen University and Research Centre

This chapter was published in *Computers and Electronics in Agriculture* (2013), Vol. 96, p. 148-162

Abstract

Sweet-pepper plant parts should be distinguished to construct an obstacle map to plan collision-free motion for a harvesting manipulator. Objectives were to segment vegetation from the background; to segment non-vegetation objects; to construct a classifier robust to variation among scenes; and to classify vegetation primarily into soft (top of a leaf, bottom of leaf and petiole) and hard obstacles (stem and fruit) and secondarily into five plant parts: stem, top of a leaf, bottom of a leaf, fruit and petiole. A multi-spectral system with artificial lighting was developed to mitigate disturbances caused by natural lighting conditions. The background was successfully segmented from vegetation using a threshold in a near-infrared wavelength (>900 nm). Non-vegetation objects occurring in the scene, including drippers, pots, sticks, construction elements and support wires, were removed using a threshold in the blue wavelength (447 nm). Vegetation was classified, using a Classification and Regression Trees (CART) classifier trained with 46 pixel-based features. The Normalized Difference Index features were the strongest as selected by a Sequential Floating Forward Selection algorithm. A new robust-and-balanced accuracy performance measure P_{Rob} was introduced for CART pruning and feature selection. Use of P_{Rob} rendered the classifier more robust to variation among scenes because standard deviation among scenes reduced 59% for hard obstacles and 43 % for soft obstacles compared with balanced accuracy. Two approaches were derived to classify vegetation: Approach A was based on hard vs. soft obstacle classification and Approach B was based on separability of classes. Approach A (P_{Rob} = 58.9) performed slightly better than Approach B (P_{Rob} = 56.1). For Approach A, mean true-positive detection rate (standard deviation) among scenes was 59.2 (7.1)% for hard obstacles, 91.5 (4.0)% for soft obstacles, 40.0 (12.4)% for stems, 78.7 (16.0)% for top of a leaf, 68.5 (11.4)% for bottom of a leaf, 54.5 (9.9)% for fruit and 49.5 (13.6)% for petiole. These results are insufficient to construct an accurate obstacle map and suggestions for improvements are described. Nevertheless, this is the first study that reports quantitative performance for classification of several plant parts under varying lighting conditions.

3.1 Introduction

This research is part of the EU funded CROPS project, 'Clever Robots for Crops', in which a sweet-pepper harvesting robot will be developed (Hemming et al. 2011). A harvesting robot includes a manipulator and end-effector for which collision-free motions should be planned to be able to reach a target (fruit or peduncle). The motion planner therefore requires locations of obstacles. But, before obstacle locations can be determined, obstacles must be detected. Consequently, obstacle detection is the focus of this research.

We decided to separate obstacles in hard obstacles and soft obstacles because the dense obstacle map requires the manipulator to push some obstacles aside to reach the target. Soft obstacles (leaves and petioles) can be pushed aside because damage to these plant parts would hardly harm the physiological status of the plant. Collisions with hard obstacles (stems, fruits, support wires, construction elements) are, however, critical. For instance, a damage to the stem may strongly limit future fruit set. Similarly, a small scratch in the fruit skin will reduce the market value of a fruit. Support wires are twisted around the stem to guide plant growth

upward. A broken support wire would cause the plant to drop on the floor and such an event would severely damage the plant. A collision with a greenhouse construction element can cause damage to the greenhouse construction, manipulator or end-effector.

Moreover, leaves were classified in the top side of a leaf (adaxial) and the bottom side of a leaf (abaxial) because this information may be required if either an upward or downward leaf motion is desired. Generally, leaves will move upward when pushed from the bottom side and move downward when pushed from the top side. In some motion planning cases it can be useful to know whether a leaf will move upward or downward when pushed. For instance, a leaf on top of a ripe fruit should be moved upward by the gripper to avoid leaf damage during fruit detachment.

A low-cost sensing solution, multi-spectral imaging, was selected to improve economic feasibility of the harvesting robot. Alternative sensors used in related applications, such as LIDAR for detection of canopy structure in apple trees (Fleck et al. 2004) and X-ray for rose stem detection (Noordam et al. 2005), were considered to be too expensive. We used multi-spectral imaging instead of colour imaging to use the spectral information present in both visible and near-infrared wavelengths. Two near-infrared wavelengths were selected: one within the red edge (696-736 nm) (Filella and Penuelas 1994) and one within a water absorption band (900-1000 nm) (Center 2008). Combinations of a wavelength in visible and near-infrared spectrum turned out to be useful features in this research.

Little research has been performed regarding obstacle detection for fruit harvesting (Bac et al. 2014) and quantitative performance was not reported. We reviewed studies performed under varying lighting conditions using either multi-spectral imaging or colour imaging. Two studies describe classification of cucumber plant parts into leaves, stems and fruits: a study regarding a cucumber leaf picking robot using two near-infrared wavelengths (Van Henten et al. 2006) and a multi-spectral imaging study in which several wavelengths and sensors were compared (Noble and Li 2012). Lu et al. (2011) detected branches of citrus using multi-spectral imaging. Stems of lychee were detected using colour imaging (Deng et al. 2011). Unfortunately, these studies lack quantitative classification performance. The article most closely related to the work presented here is classification of grape foliage into leaves, branches and fruits (green or coloured) using RGB cameras with a Support Vector Machine classifier. For green grapes, true-positive rate was 91.9% with a false-positive rate of 2.7%. Performance for branches and leaves was not reported (Dey et al. 2012). To the best of our knowledge, quantitative performance of plant part classification under varying lighting conditions was only reported for fruit detection (Jiménez et al. 2000), whereas this study reports quantitative performance for five plant parts.

A Classification And Regression Trees (CART) classifier was used for classification (Breiman et al. 1984). CART is a variant of a decision tree classifier and performs similar to variant C4.5 (Unay et al. 2006). Decision trees were hardly used in related classification studies, whereas decision trees can reach classification accuracies similar to other classifiers (Kavdır and Guyer 2004) and fit the requirements listed in this article.

Two literature reviews served as inspiration for the approach taken. Firstly, McCarthy et al (2010) discuss how disturbances by natural lighting can be addressed upon design of the vision system. Based on that, we assembled intense lighting to mitigate natural lighting

disturbances. Moreover, a new classification performance measure, robust-and-balanced accuracy P_{Rob}, is introduced that considers both balanced accuracy and robust classification performance among scenes. Robustness is defined as the ability of a system to resist change without adapting its initial stable configuration (Wieland and Wallenburg 2012). We show the advantage of P_{Rob} compared with balanced accuracy as performance measure. Secondly, Kapach et al. (2012) show that few studies use visual cues other than colour, i.e. shape, size and especially texture, for agricultural applications. We therefore evaluated spectral and texture features to classify pixels.

Objectives were to 1) detect plant vegetation in a crop row; 2) segment non-vegetation objects; 3) prune a decision tree and select features such that the classifier is robust to variation among scenes; 4) classify vegetation primarily into hard and soft obstacles and, secondarily, into stems, top of leaves, bottom of leaves, green fruits and petioles.

The experimental set-up, experiments and new performance measure presented in this article may contribute tools to perform plant part classification under varying lighting conditions.

3.2 Image processing sequence

Figure 3-1 presents an overview of this article. The image processing sequence to classify vegetation is described in Section 3.2. Algorithms (performance measures, classifier, feature selection algorithm and features) for vegetation classification are described in Section 3.3. These algorithms were used in Experiment 1 and 2, described in Section 3.4. Experiment 1 relates to the third objective and Experiment 2 relates to the fourth objective.

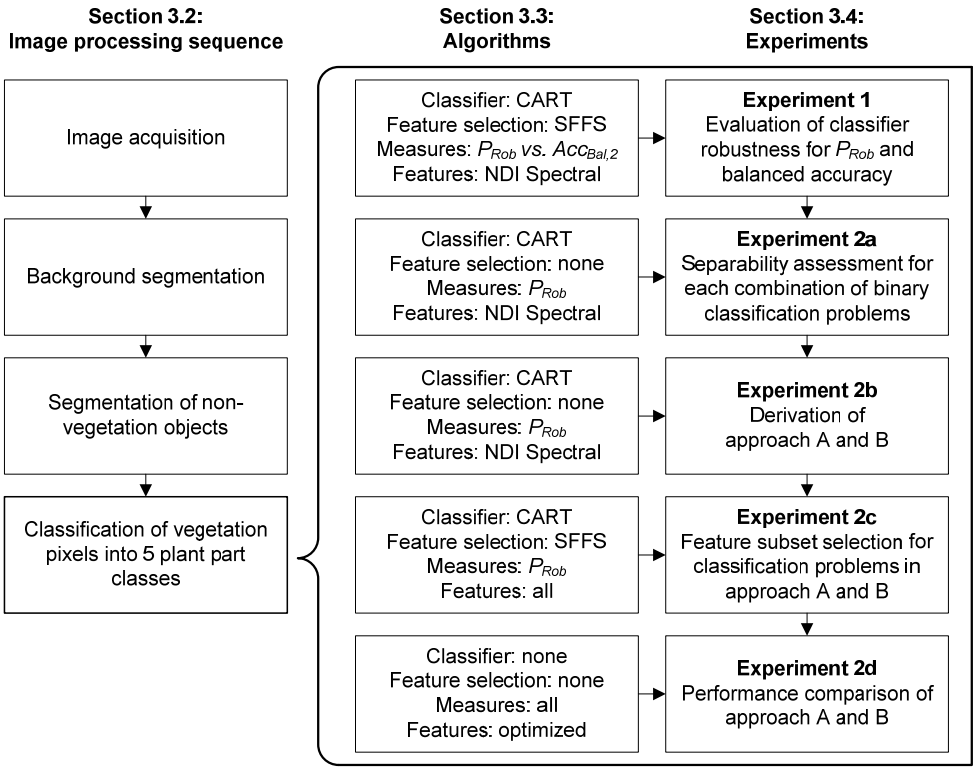


Figure 3-1. Overview of the image processing sequence and of algorithms used for each experiment.

The five-class classification problem, in Experiment 2, was split into four binary classification problems because Kavdir and Guyer (2004) demonstrated that accuracy for binary classification problems was greater than for multi-class classification problems.

3.2.1 Image acquisition

The multispectral imaging set-up developed consisted of a monochrome camera, housing around the camera and a filter wheel (Section 3.2.1.1) holding six filters (Section 3.2.1.2). The scene was illuminated by continuous lighting (Section 3.2.1.3). An overview of the experimental set-up is in Figure 3-2.

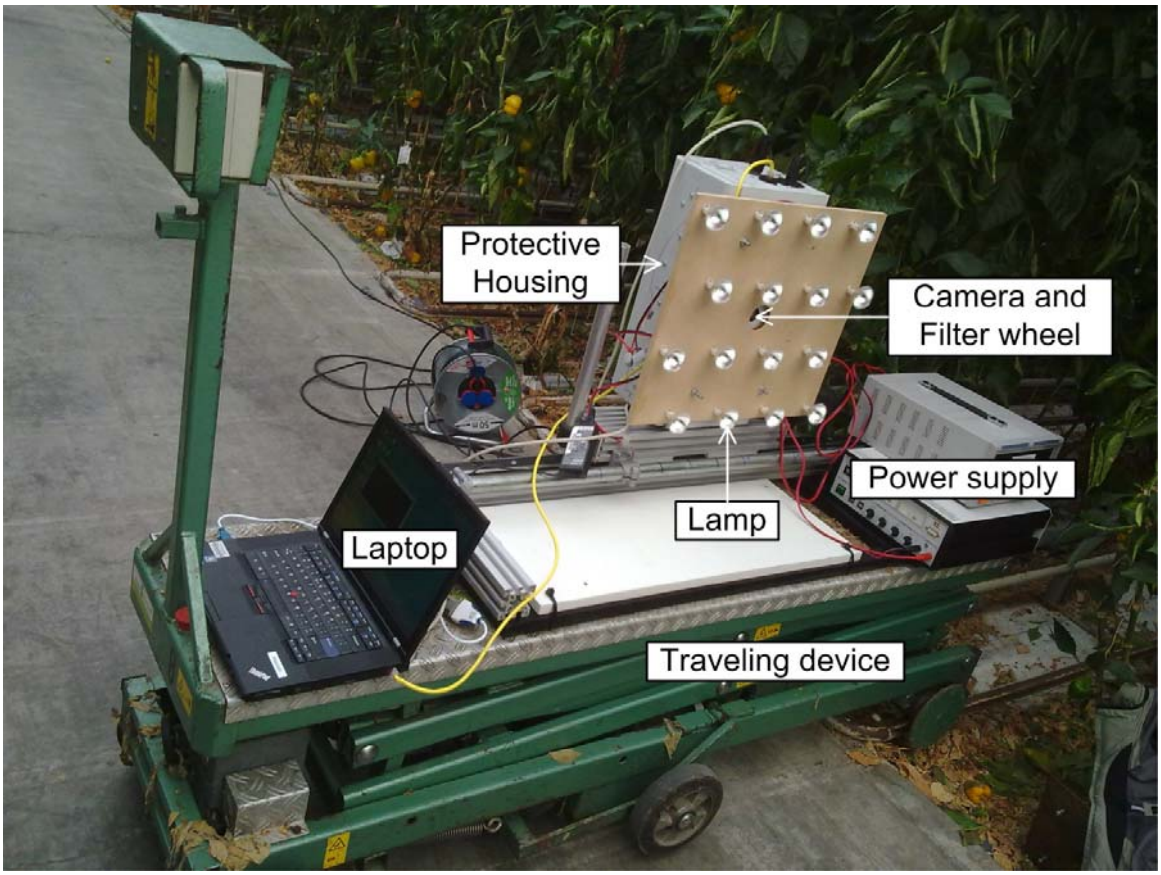


Figure 3-2. Overview of the experimental set-up in a greenhouse.

3.2.1.1 Camera and filter wheel

The monochrome camera used was a 5 megapixel camera with a 2/3" CCD (Manta G504B; Allied Vision Technologies GmbH, Germany). A low-distortion lens with 8 mm focal length (LM8JCM; Kowa GmbH, Germany) was mounted on the camera. Additionally, a filter wheel (Stock no. NT56-658; Edmund Optics Ltd., UK), holding six Ø25 mm filters, was placed in front of the lens. Exposure time was set such that less than 5 % of the pixels were overexposed. After determining this time for each filter, exposure time was set constant during recording.

We assembled a motorized filter wheel because an off-the-shelf motorized filter wheel was too thick (>15 mm) and would occlude the edges of the image. The assembled filter

wheel comprised a manually rotatable filter wheel (thickness 6.35 mm) with a stepper motor (QMot QSH4218; Trinamic Motion Control GmbH & Co. KG, Germany) and a belt. Recording of six images, including wheel rotation, took about 2 s. Misalignment of the image was 0-3 pixels vertically and 0-1 pixel horizontally, with respect to the first image taken in a wheel rotation. We consider this misalignment to be of negligible influence to the results because labelled ground-truth regions were well within the edge of plant parts.

Image processing was performed on a computer with an Intel Core i5 CPU 2.4 GHz Quad core processor with 4 GB memory.

3.2.1.2 Wavelength selection

Six wavelengths were selected to distinguish between plant parts. Selection of these wavelengths was based on unpublished work (Fransen 2011) in which stem, peduncle, fruit, top of a leaf and bottom of a leaf were recorded with a hyperspectral camera – Specim ImSpector V10E (Spectral Imaging Ltd., Finland). Hence, six filters were used (Edmund Optics Ltd., UK):

- 447 nm – bandwidth 60 nm (Stock no. NT48-074)
- 562 nm – bandwidth 40 nm (Stock no. NT48-085)
- 624 nm – bandwidth 40 nm (Stock no. NT48-087)
- 692 nm – bandwidth 40 nm (Stock no. NT48-148)
- 716 nm – bandwidth 40 nm (Stock no. NT67-039)
- >900 nm – longpass (Stock no. NT66-237)

Note that measured bandwidth of the >900 nm longpass filter is limited to 1000 nm due to limited sensitivity of the CCD chip in the camera.

3.2.1.3 Artificial lighting

For illumination of the scene, 16 halogen lamps (12VDC, 75 W, OSRAM GmbH, Germany) were used. Lamps were placed 5 cm in front of the camera. Each lamp was equipped with a dichroic reflector (Ø 51 mm) for equal light distribution. During measurements in a dark room it turned out a lamp emitted light horizontally at an angle of 36° and vertically at an angle of 43°. Based on these measured light emission angles, 16 lamps were distributed over four rows. The horizontal distance between lamps, in a row, was 13 cm. The vertical distance between rows was 13.6 cm. Each row was horizontally shifted (6.5 cm) with respect to the previous row. As a result, equal lighting distribution was ensured for a camera-canopy distance of 25 until 55 cm (Figure 3-3). At distances closer than 25 cm, the scene was not equally illuminated. At distances farther than 55 cm, lighting was less intense at edges of the image than at the centre because light spots overlapped at the centre, whereas light spots did not overlap at edges.

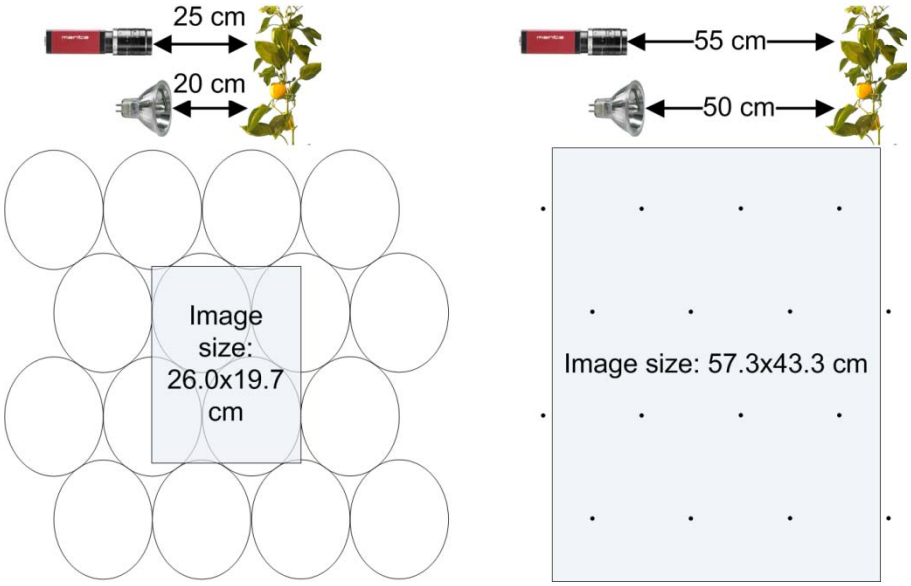


Figure 3-3. Ovals display lighting distribution, at a camera-canopy distance of 25 cm (left). The gray rectangle displays the image size at a camera-canopy distance of 55 cm (right).

3.2.2 Environment and objects

Plants were of the red sweet-pepper cultivar 'Viper' and were cultivated in the V-system (Jovicich et al. 2004). Plants (height = 1.5 m) were grown in an experimental greenhouse in Wageningen, The Netherlands. Images were recorded from 11:00 am until 11:30 am under a clear sunny sky. Outdoor solar irradiance was measured during image recording and varied between 374 and 435 W/m2. Plant(stem)-camera distance was on average 80 cm and varied in a range of 63 cm to 109 cm among scenes. Camera-plant(leaf) distance was mostly at least 25 cm and the lighting (Section 3.2.1.3) was therefore fit for this distance.

We classified objects, which occurred in the scene, for later use in a motion planning problem (Table 3-1). Only green unripe fruits occurred in recorded scenes.

Table 3-1. Objects that occurred in recorded scenes. Each object was classified for motion planning by an image processing operation.

Object type	Classified for motion planning as	Operation used for classification	Abbre-viation
Objects with distance >1.2 m	Background	Threshold (>900 nm)	Back
Support wire	Hard obstacle	Threshold (447 nm) and area-based segmentation	SW
Stick, dripper and pot	Hard obstacle	Threshold (447 nm) and area-based segmentation	Pot
Construction elements	Hard obstacle	Threshold (447 nm) and area-based segmentation	Const
Stem	Hard obstacle	Pixel-based classification	Stem
Green Fruit	Hard obstacle	Pixel-based classification	Fruit
Petiole	Soft obstacle	Pixel-based classification	Pet
Top of a leaf	Soft obstacle	Pixel-based classification	TL
Bottom of a leaf	Soft obstacle	Pixel-based classification	BL

An example of objects occurring in a scene are displayed in Figure 3-4.

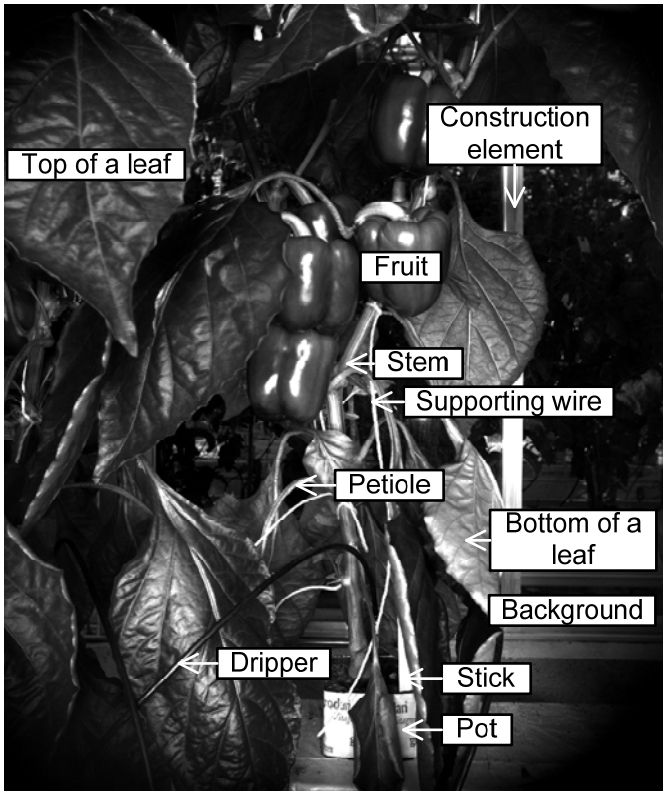


Figure 3-4. Overview of objects that occurred in an arbitrary scene.

Construction elements in the scene were aluminium window frames. Other construction elements such as support poles did not occur in the recorded scenes.

3.2.3 Background segmentation

After image acquisition, the background was segmented using a gray-value threshold, which was empirically determined in HALCON© 10.0.1 (MVTec GmbH, Germany). The threshold was applied on the >900 nm image because the background hardly reflected in this spectral range due to low solar irradiance in the range of 925-975 nm (Center 2008). In other words, the scene was mainly illuminated by artificial lighting and therefore hardly any light reflected from distances farther than 1.2 m. As a result, the background occurred dark and was easily segmented. Remaining holes in the background were filled with a morphological 'fill up' operation.

3.2.4 Segmentation of non-vegetation objects

Objects other than vegetation, i.e. support wire, construction elements, stick, pot and dripper, were removed by a gray-value threshold, which was empirically determined in HALCON©. A blue wavelength (447 nm) was selected to segment non-vegetation by this threshold because green vegetation hardly reflected, whereas non-vegetation objects strongly reflected. Segmented regions varied in size and an area-based threshold was used to assign small regions to the background and larger regions to hard obstacles. Larger regions were enlarged with a dilation (3x3 mask) operation, to include object edges.

3.3 Algorithms

Performance measures, the classifier, the feature selection algorithm and features used for classification of vegetation pixels are described in the following sections.

3.3.1 Performance measures

Only binary classification problems were performed in this research. Standard performance measures for these problems are described in Section 3.3.1.1 and a new measure is introduced in Section 3.3.1.2. After four binary classification problems were performed, five classes remained and five-class performance measures were used to assess overall performance (Section 3.3.1.3).

3.3.1.1 Two classes: standard measures

A confusion matrix (Table 3-2) was used to assess classification performance. Labels 'Object I' and 'Object II' were substituted for combinations of plant part classes 'TL', 'BL', 'Fruit', 'Stem', 'Pet' throughout this article. In addition, label 'Object I' was substituted for 'Hard', which indicates union of two hard obstacle classes (Stem ∪ Fruit), and label 'Object II' was substituted for 'Soft', which indicates union of three soft obstacle classes (TL ∪ BL ∪ Pet).

Table 3-2. Confusion Matrix

		Actual Class	
		Object I	Object II
Classified Class	Object I	TP_I	FP_I
	Object II	FP_{II}	TP_{II}

Where: TP_I (-) true-positive detection of object I
TP_{II} (-) true-positive detection of object II
FP_I (-) false-positive detection of object I
FP_{II} (-) false-positive detection of object II

The confusion matrix was used to evaluate classification performance in terms of: total classification accuracy, true-positive detections for each class, false-positive detections for each class and balanced accuracy (Equations (3-1)-(3-6)). The number '2' in the symbols refers to a binary problem.

$$Acc2_{Tot} = \frac{100 \cdot TP_I + TP_{II}}{TP_I + TP_{II} + FP_I + FP_{II}} \quad (\%) \qquad (3\text{-}1)$$

$$TPR2(I) = \frac{100 \cdot TP_I}{TP_I + FP_{II}} \quad (\%) \qquad (3\text{-}2)$$

$$TPR2(II) = \frac{100 \cdot TP_{II}}{TP_{II} + FP_I} \quad (\%) \qquad (3\text{-}3)$$

$$FPR2(I) = \frac{100 \cdot FP_I}{FP_I + TP_{II}} \quad (\%) \qquad (3\text{-}4)$$

$$FPR2(II) = \frac{100 \cdot FP_{II}}{FP_{II} + TP_{I}} \quad (\%) \quad (3\text{-}5)$$

$$Acc2_{Bal} = 0.5 \cdot \left(TPR2(I) + TPR2(II)\right) \quad (\%) \quad (3\text{-}6)$$

Where: $Acc2_{Tot}$ (%) total classification accuracy
$TPR2(I)$ (%) true-positive detection rate of object I
$TPR2(II)$ (%) true-positive detection rate of object II
$FPR2(I)$ (%) false-positive detection rate of object I
$FPR2(II)$ (%) false-positive detection rate of object II
$Acc2_{Bal}$ (%) balanced accuracy

In the literature, researchers mostly use total classification accuracy as a performance measure to evaluate algorithms. A drawback of this measure is that each class needs to be equal in sample size, otherwise classification accuracy will approximate true-positive detection rate of the class containing most samples. A workaround would be to calculate mean true-positive rate for two classes, i.e. balanced accuracy. Still then, such a performance measure does not consider classification accuracy variation among scenes. Classification accuracy variation among scenes can be large due to three reasons: different lighting conditions, different camera-object distances and the varying shape, colour, size, texture and position of objects. As robustness to these three disturbances is currently not assessed, we introduce a new performance measure.

3.3.1.2 Two classes: new measure

We established a new performance measure *robust-and-balanced accuracy* P_{Rob}, which compensates for class size differences and considers classification robustness among scenes. This new measure was inspired by the *d'* measure, commonly used in signal detection theory (Wickens 2001) and was applied to distinguish between orange fruits and leaves (Bulanon et al. 2010). In contrast to the *d'* measure, P_{Rob} considers variation among scenes (Equation (3-7)).

$$P_{Rob} = \frac{Rob_{Mit} \cdot 0.5 \cdot \left(M_{TPR2(I)} + M_{TPR2(II)}\right)}{0.5 \cdot \left(SD_{TPR2(I)} + SD_{TPR2(II)}\right) + Rob_{Mit}} \quad (-) \quad (3\text{-}7)$$

Where: P_{Rob} (-) robust-and-balanced accuracy
$M_{TPR2(I)}$ (%) mean of multiple scenes with true-positive detection rate of object I
$M_{TPR2(II)}$ (%) mean of multiple scenes with true-positive detection rate of object II
$SD_{TPR2(I)}$ (%) standard deviation of multiple scenes[1] with true-positive detection rate of object I

[1] Note that standard deviation $SD_{FPR2(I)}$ is equal to $SD_{TPR2(II)}$, and similarly $SD_{FPR2(I)}$ is equal to $SD_{TPR2(II)}$

$SD_{TPR2(II)}$	(%) standard deviation of multiple scenes with true-positive detection rate of object II
Rob_{Mit}	(%) robustness mitigation factor, must be >0. Value was set to 20%

P_{Rob} is a division of the mean TPR among scenes by its sample standard deviation (SD). An increasing SD will punish P_{Rob}, whereas an increasing mean TPR will reward P_{Rob}. We use P_{Rob} to determine the pruning level of the classifier and to select features.

SD is a measure to quantify the effect of disturbances on consistent classification accuracy among scenes. By incorporating SD, the classifier becomes more robust to disturbances. Sometimes, additional robustness of the classifier may come at a cost of accuracy loss. To vary the weight of robustness vs. accuracy, a robustness mitigation factor Rob_{Mit} was added. As such, P_{Rob} has similarity with the coefficient of variation, but weight on mean (accuracy) vs. SD (robustness) can be adjusted through Rob_{Mit}. The user must set Rob_{Mit} *a priori* based on required robustness in his application. A user can calculate Rob_{Mit} if allowable average standard deviations are known for which performance P_{Rob} should drop by 50% because, by definition, P_{Rob} drops by 50% when $Rob_{Mit} = 0.5 \cdot (SD_{TPR2(I)} + SD_{TPR2(I)})$. In this application, the exact required robustness of the classifier was unknown. But, Rob_{Mit} was empirically set to 20 to assure a stronger punishment of P_{Rob} by an increase of $0.5 \cdot (SD_{TPR2(I)} + SD_{TPR2(I)})$ than a similar decrease of $0.5 \cdot (M_{TPR2(I)} + M_{TPR2(I)})$.

3.3.1.3 Five classes

To assess combined performance of four binary problems, true- and false-positive detection rates, total accuracy and balanced accuracy were calculated for a 5 by 5 confusion matrix. Similar to measures for two classes (Section 3.3.1.1), total accuracy for five classes $Acc5_{Tot}$ is the division of the summed confusion matrix diagonal by the sum of all confusion matrix elements. Mean true-positive detection rates, taken over multiple scenes, of the stem ($M_{TPR5(Stem)}$), top of leaf ($M_{TPR5(TL)}$), bottom of a leaf ($M_{TPR5(BL)}$), fruit ($M_{TPR5(Fruit)}$) and petiole ($M_{TPR5(Pet)}$) are calculated as the ratio between correctly classified pixels of a class and the total actual pixels of that class. Balanced accuracy $Acc5_{Bal}$ is the average of these five mean true-positive detection rates. Each class includes four false-positive detection rates, i.e. 20 values in total.

3.3.2 Classifier

Three requirements resulted in the choice for a decision tree classifier: 1) the classifier must consume little computational load because the application requires an overall harvesting cycle time of 6 s per fruit; 2) the classifier must handle data which are non-normally distributed because Q-Q plots of the gray-value distributions per class and per image revealed that data were non-normally distributed. 3) the classifier should be easily transferable to another programming language. Non-parametric classifiers, i.e. decision tree, artificial neural network and k-nearest neighbour, do not assume a mathematical distribution and are preferred. Linker et al. (2012) also selected a non-parametric classifier (nearest neighbor) for similar reasons. A decision tree requires less computational load than an artificial neural network and decision trees can be easily implemented in any programming language by a set of if-then-else

statements. Hence, a Classification And Regression Trees (CART) classifier (Breiman et al. 1984) was implemented in MATLAB® 2007b.

Moisen (2008) explains basics of CART and discusses which properties should be set for CART training. The Gini Diversity Index (GDI) is a commonly used splitting criterion and was used in this research as well because other splitting criteria (twoing and deviance reduction) performed equally or worse. After training, the tree was pruned to avoid overfitting (Section 3.3.2.1). Other relevant parameters for decision tree training are initial probabilities and weights (Section 3.3.2.2) and equalization of the training samples in a class (Section 3.3.2.3).

3.3.2.1 Pruning

As Moisen (2008) points out, pruning can be performed by n-fold cross-validation, the 1-SE rule, or maximizing performance on an independent test set. We optimized pruning on an independent test set. To evaluate effects of the new performance measure, P_{Rob} and $Acc2_{Bal}$ were each used to prune the decision tree (Equation (3-8)).

$$prune_{opt} = \max(P_{Rob} \text{ or } Acc2_{Bal}) \quad (-) \qquad (3\text{-}8)$$

Where $prune_{opt}$ (-) is the optimal prune level, which determines tree length. Note that performance is always reported for testing data, i.e. the independent test set, throughout the article.

3.3.2.2 Initial probabilities and weights

Initial probabilities set the likeliness an arbitrary pixel belongs to a certain class. We set initial probabilities to equal values because it turned out that for other probabilities large classes were favoured over small classes and, as a result, performance P_{Rob} decreased.

Weights are another option to influence decision tree training. A weight $w_{i,j}$, for $i \neq j$, is the cost of pixel of class i to be misclassified as class j. In a brute force simulation of all possible weight combinations it turned out that weights, other than equal weights, reduced total classification accuracy. We therefore set weights to an equal value ($w_{i,j} = 1$).

3.3.2.3 Equalization of training set size per class

CART has the tendency to favour classes with most training samples (Cieslak and Chawla 2008). Training sets were therefore reduced to the size of the smallest training class such that training sets were balanced.

3.3.3 Feature selection algorithm

As Blum and Langley (1997) indicated, two critical issues in machine learning are training data selection and feature selection. Training samples were abundantly available and therefore efforts were put on feature selection only. A suboptimal Sequential Forward Floating Selection (SFFS) feature selection method was chosen. This 'floating search method' is, compared with the optimal branch-and-bound method, less computationally intensive, able to handle problems that do not require the monotonicity condition and approaches optimality closely (Jain 1997). The best feature set is selected based on classification performance in terms of P_{Rob} or $Acc2_{Bal}$. The SFFS method was also used for classification of phalaenopsis stem and root (Huang and Lee 2008) and for apple stem and calyx detection (Unay et al.

2006). An improved version of the SFFS method has been developed as well (Nakariyakul and Casasent 2009), but its computational load is at least four times greater than the SFFS method. Given that feature selection with SFFS already lasted three days per binary problem, we used the SFFS method for feature selection.

We also tested Principal Component Analysis (PCA), but results of PCA are not reported because the first four components (98 % of correlation) performed worse than feature selection applied to features in Section 3.3.4.

3.3.4 Pixel-based features

An overview of features used is given in the following sections. In total, 46 features (6+6+15+15+2+2) were calculated for each pixel. Except for entropy, features were calculated in MATLAB®.

3.3.4.1 Raw data

Merely the reflectance at six wavelengths, i.e. six intensity values per pixel.

3.3.4.2 Entropy

An entropy operation (mask size = 3x3) was applied on the raw data using HALCON© and yielded six values per pixel. Entropy is a method to calculate texture within the specified mask. The relative frequencies of gray-values within the mask is used as an input to calculate entropy of the centre pixel in the mask.

3.3.4.3 Normalized Difference Index (NDI)

The Normalized Difference Index (NDI) is a measure which describes the normalized difference between two different wavelength channels (Davies 2009). An advantage of this index is that lighting influences are reduced through a division by two channels. NDI is commonly applied to raw data but here we tested its effect on entropy as well. In total 15 possible combinations of two wavelengths from a set of six wavelengths exist. That is, applied for raw data and entropy, 30 values were calculated per pixel.

3.3.4.4 Spectral Angle Mapper (SAM)

The spectral angle mapper is a commonly used feature for hyperspectral or multi-spectral analysis. It measures similarity of an image pixel spectrum with respect to a reference spectrum (Yang et al. 2008). One spectral angle was calculated per class, i.e. each pixel yielded two values.

3.3.4.5 Mahalanobis distance

The Mahalanobis distance algorithm is an improvement of the commonly known Euclidean distance algorithm because correlations of wavelengths are considered (DeVries 2005). An average intensity vector was calculated from labelled ground-truth regions in the training image. As for the SAM, each pixel yielded two values: one value per class.

3.4 Experiments

Ground truth labelling, assigning training and testing data and Experiments 1 and 2 are explained in the following sections.

3.4.1 Ground truth labelling

In total, 12 scenes were recorded using the experimental set-up, one plant per scene. Ground truth labelling was manually performed, using HALCON©, for the five plant parts to be classified. A result of labelling is in Figure 3-5.

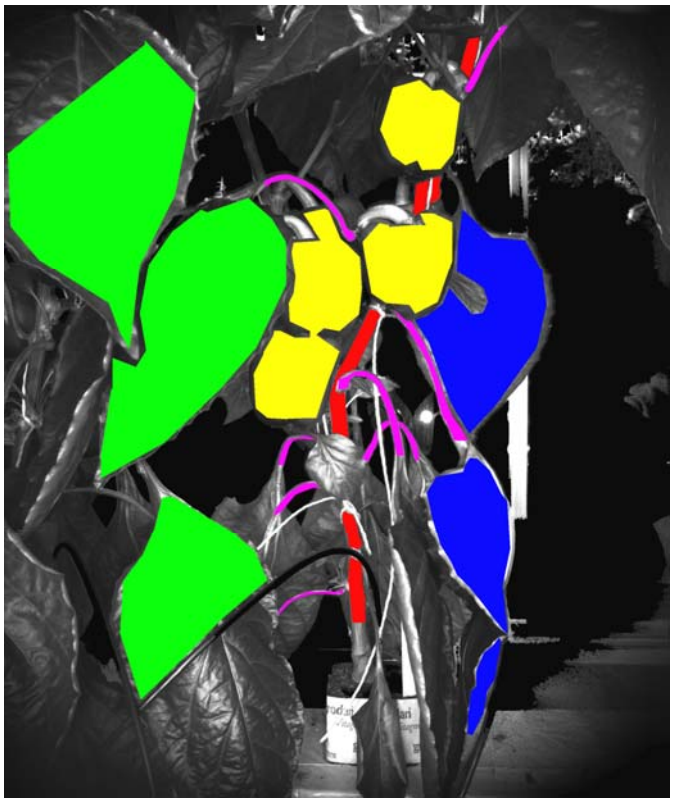


Figure 3-5. Example of manually drawn ground truth labels for five classes: stem (red), top of a leaf (green), bottom of a leaf (blue), green fruit (yellow) and petiole (purple).

In total, $14.6 \cdot 10^6$ pixels were labelled and this number comprised 29.4% of the vegetation present in the 12 scenes. The number of labelled pixels varied among scenes because visibility of plant parts also varied among scenes. Most of labelled pixels were leaves – TL (54.6%), BL (22.4%) – followed by fruits (15.6%), stems (3.7%) and petioles (3.7%). Vegetation pixels which were not labelled (70.6%) were mostly leaves and these were not labelled because we assumed labelled leaves already represented the majority of the leaf variation occurring in the scene. In addition, labelling all samples would increase the computational load dramatically, especially during feature selection.

3.4.2 Training and testing data

We evaluated the number of scenes required for training by evaluating all possible combinations of selecting this number of scenes from the population of 12 scenes (Table 3). We performed hard vs. soft obstacle classification using 15 NDI spectral features as input. Only NDI spectral features were used because these features were strong and therefore sufficient to assess the size of the training set.

Table 3-3. Effect of different combinations and number of training scenes on classification performance

Scenes used for training vs. testing	Combinations evaluated (#)	Mean (SD) P_{Rob} (-) among combinations	Performance increase of mean P_{Rob} (-)
1 vs. 11	12	44.4 (7.7)	
2 vs. 10	66	49.6 (5.2)	5.2
3 vs. 9	220	52.3 (4.1)	2.7
4 vs. 8	495	54.7 (3.4)	2.4

We selected two scenes for training and ten for testing because performance increase was greater (5.2) than the increase to three training scenes (2.7) or four training scenes (2.4). Performance increase for three or four training scenes may have been partially due to the decreasing number of testing scenes. Regarding the combination selected out of 66 possible combinations, a combination that performed on average (P_{Rob} = 49.4) was selected to perform Experiment 1 and 2. These two scenes comprised $2.9 \cdot 10^6$ labelled pixels; ten scenes for testing comprised $11.7 \cdot 10^6$ labelled pixels.

3.4.3 Experiment 1: Evaluation of classifier robustness

The effect of the new measure P_{Rob} vs. $Acc2_{Bal}$ was separately evaluated in terms of accuracy and consistent classification among scenes, i.e. robustness. Each measure was used for decision tree pruning and for selection of the optimal feature subset, using 15 NDI spectral features. The best measure was P_{Rob} and was therefore used for pruning and feature selection in Experiment 2.

3.4.4 Experiment 2a: Separability assessment for each combination of binary classification problems

As mentioned earlier, the five plant part classes were classified by an approach of four binary classification problems. To define such an approach, the separability of all possible binary combinations were evaluated. In total, 15 combinations exist to divide five classes into a binary problem. Hence, unions of 2 and 3 classes (10 combinations) or 1 and 4 classes (5 combinations) were taken.

Similar to the evaluation of training data set sizes (Section 3.4.2), a fixed set of 15 NDI spectral features was used to calculate the separability for each of the 15 combinations. Feature selection was therefore only performed in Experiment 2c to avoid a computation time of 45 days (3 days*15 combinations). And, we doubt whether feature selection would lead to a different best separable combination.

3.4.5 Experiment 2b: Derivation of Approach A and B

There are many approaches possible to reduce a 5-class classification problem into four binary classification problems, based on results from Experiment 2a. We investigated two different approaches. Firstly, Approach A was inspired by the application. Here, hard (Stem ∪ Fruit) and soft (TL ∪ BL ∪ Pet) obstacles were classified first. Secondly, Approach B was inspired by the separability of classes. Here, the best separable classes, identified in Experiment 2a, were classified first. Once both approaches were defined, their four binary classification problems were optimized using feature selection (Section 3.4.6).

3.4.6 Experiment 2c: Feature subset selection

Feature selection was performed for each binary classification problem of Approach A and B, using the feature selection algorithm and pixel-based features described in Section 3.3. To reduce the computational load of feature selection for a set of 46 features, feature selection was first performed per type of feature (Section 3.3.4). Secondly, these selected sets (7) were clustered in a 'Cluster' set for which feature selection was applied.

3.4.7 Experiment 2d: Performance comparison of Approach A and B

Approach A and B were compared in terms of robust-and-balanced accuracy of hard and soft obstacle classification and in terms of 5-class performance measures. Hard and soft obstacle classification performance is more critical for motion planning than the classification performance for five classes. The best approach was therefore chosen based on robust-and-balanced accuracy of hard and soft obstacle classification.

3.5 Results

3.5.1 Background segmentation

To segment the background from other objects in the scene, a threshold was set to ≤ 27 (>900 nm image; 8 bit). Subsequently holes were filled (Figure 3-6). Calculation took 7 ms.

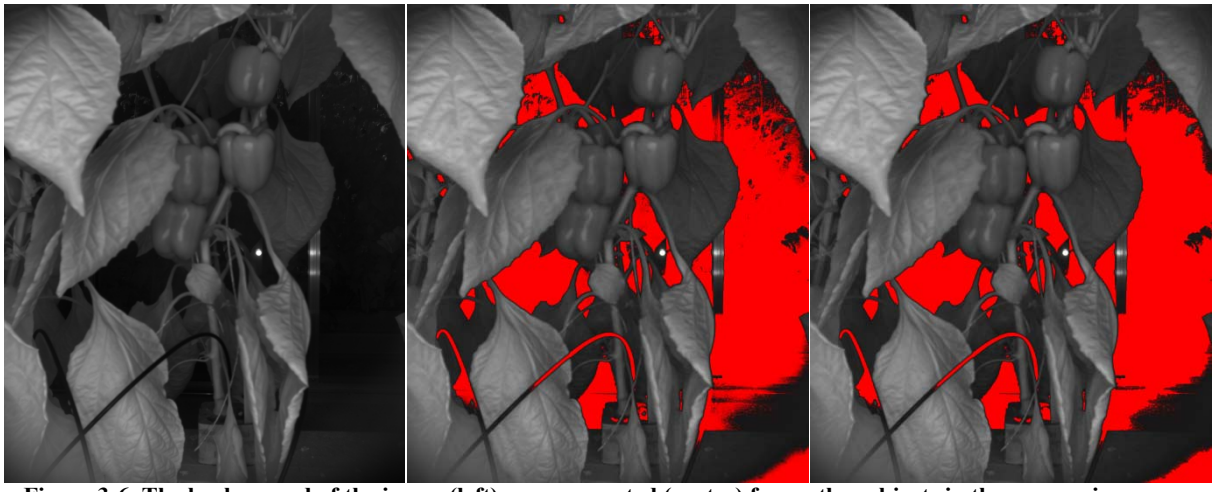


Figure 3-6. The background of the image (left) was segmented (centre) from other objects in the scene using a gray value threshold (≤ 27). Subsequently holes were filled (right). Unfortunately, part of the dripper is classified as background whereas it should have been classified as a hard obstacle.

The vegetation area decreased by 2.7% if the gray-value threshold was set to 30, i.e. an increase of 11%. Hence, the threshold value has an effect on classifiable vegetation area and had to be selected carefully.

3.5.2 Segmentation of non-vegetation objects

To segment all non-vegetation objects, a threshold was set to >139 (447 nm image; 8 bit). Segmented regions with an area larger than or equal to 300 pixels included construction elements, stick, pot, and supporting wires. These hard obstacles were therefore removed before classification of vegetation. Segmented regions with an area smaller than 300 pixels were assigned to the background class (Figure 3-7). Calculation took 6 ms.

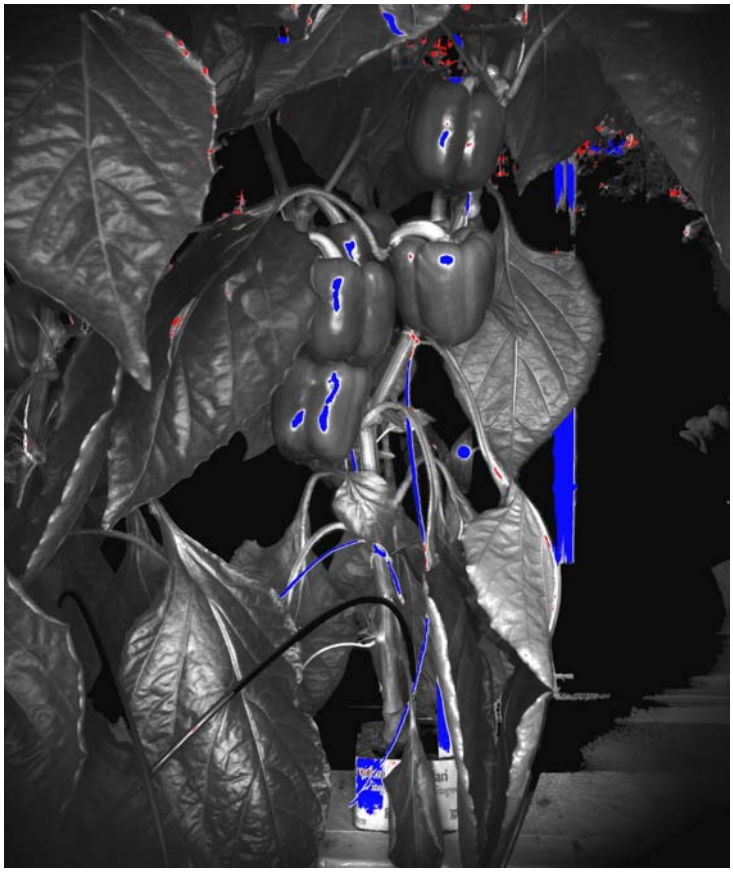


Figure 3-7. Segmented regions, mostly non-vegetation objects, with an area smaller than 300 pixels (red) and segmented regions with an area larger than or equal to 300 pixels (blue).

Some vegetation pixels, about 1% of total vegetation area, were also part of the segmented regions (Figure 3-7) and were therefore not classified. These vegetation pixels were mainly fruits and because fruits are hard obstacles it is not problematic that segmented regions included some vegetation pixels. These fruit segments were overexposed due to direct light reflection on the glossy fruit surface.

The pot, support wire, construction elements and stick were removed by this threshold operation. Drippers were not detected due to their black colour and use of white drippers may be a solution to detect drippers as a hard obstacle. Nevertheless drippers were classified as background (Figure 3-6) and a scene remained that contained only the five plant parts to be classified. All non-vegetation objects were removed from the scene.

3.5.3 Experiment 1: Evaluation of classifier robustness

Due to the use of the new performance measure P_{Rob}, we observed that feature selection resulted in different optimal feature sets compared with feature selection based on only balanced accuracy as measure. As an example, Table 3-4 shows that accuracy is slightly greater for the balanced accuracy performance measure (77.1) than for the robust-and-balanced accuracy performance measure (75.4). But, for balanced accuracy, the standard deviation on the true-positive-rates is about two times greater. Hence, the classifier is more robust to variations among scenes when P_{Rob} was used instead of balanced accuracy.

Table 3-4. Comparison between balanced accuracy and robust-and-balanced accuracy as performance measure for decision tree pruning and feature selection. Feature selection was applied to NDI spectral features for classification problem A1: hard vs. soft obstacle classification.

	Performance measure used	
	Balanced accuracy: $Acc2_{Bal}$	Robust-and-balanced accuracy: P_{Rob}
Features (NDI spectral) in the pruned decision tree; ordered on occurrence.	562&900; 692&716; 692&900; 562&716; 624&692; 562&624	447&624; 624&900; 692&716; 562&624
Tree length before pruning (nodes)	9427	9287
Tree length after pruning (nodes)	67	9
Balanced accuracy $Acc2_{Bal}$ (%)	77.1	75.4
Robust-and-balanced accuracy P_{Rob} (-)	47.7	58.9
$Acc2_{Tot}$ (SD) (%)	84.2 (6.0)	84.5 (5.0)
$M_{TPR2(hard)}$ ($SD_{TPR2(hard)}$) (%)	66.5 (17.2)	59.2 (7.1)
$M_{TPR2(soft)}$ ($SD_{TPR2(soft)}$) (%)	87.4 (7.0)	91.5 (4.0)

Table 3-4 shows that, for P_{Rob}, two features (447&624, 624&900) were selected that were not selected when $Acc2_{Bal}$ was taken as performance measure. One may conclude these two features are less sensitive to influences of lighting variations and plant-camera distance variations.

3.5.4 Experiment 2a: Separability assessment for each combination of binary classification problems

Figure 3-8 demonstrates the separability of the 15 combinations possible.

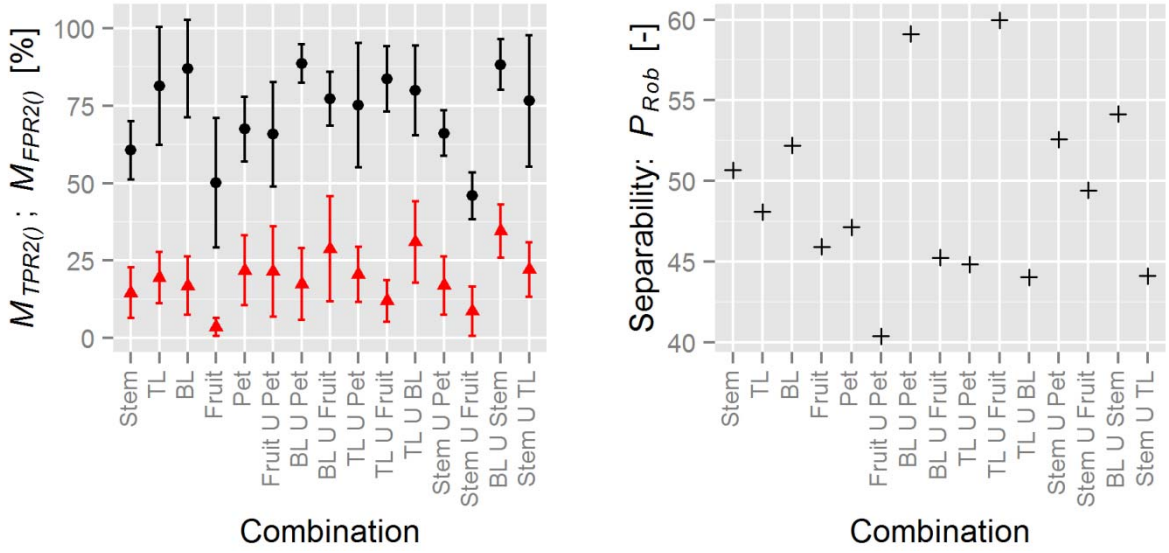


Figure 3-8. Separability for 15 combinations, each combination represents a binary classification problem. Classes in each combination were taken vs. all remaining classes. Mean and SD (N=10) of true-positive detection rates $M_{TPR2()}$ (●) and false-positive detection rates $M_{FPR2()}$ (▲) (left) were used as input to calculate the separability performance measure P_{Rob} (+) (right). Combination 'TL ∪ Fruit' (P_{Rob} = 60.0) is best separable.

Union of top of a leaf (TL) and fruit is the best separable combination (P_{Rob} = 60.0). This combination was therefore classified first in Approach B. Separabiliy of hard (Stem ∪ Fruit)

vs. soft (TL ∪ BL ∪ Pet) obstacles (P_{Rob} = 49.4) is average compared with the other 15 combinations.

3.5.5 Experiment 2b: Derivation of Approach A and B

An overview of Approach A is shown in Figure 3-9. This approach was based on relevance for the application, i.e. hard vs. soft obstacle classification first.

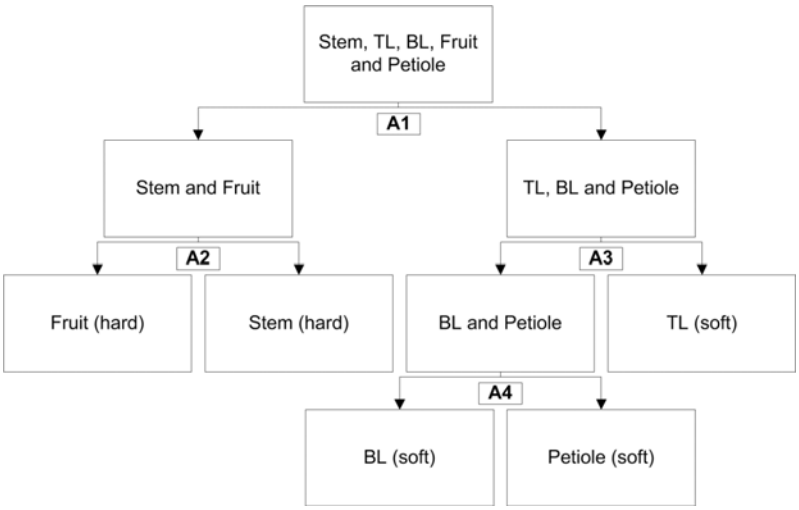


Figure 3-9. Approach A to reduce a five-class classification problem into four binary problems. The classification sequence is based on the relevance to the application, i.e. hard vs. soft obstacle classification first.

Three different options were available to divide three classes for classification problem A3. In a classification where the 15 NDI spectral features were fed to the classifier, it turned out that TL and union of BL and Pet (P_{Rob} = 57.7) resulted in better separability than BL and union of TL and Pet (P_{Rob} = 49.1) or Pet and union of BL and TL (P_{Rob} = 45.8). Consequently, TL and union of BL and Pet were taken as classes in classification problem A3 and finally BL and Pet were classified in problem A4.

An overview of Approach B is shown in Figure 3-10. This approach was based on separability of classes, i.e. the easiest classification problem first.

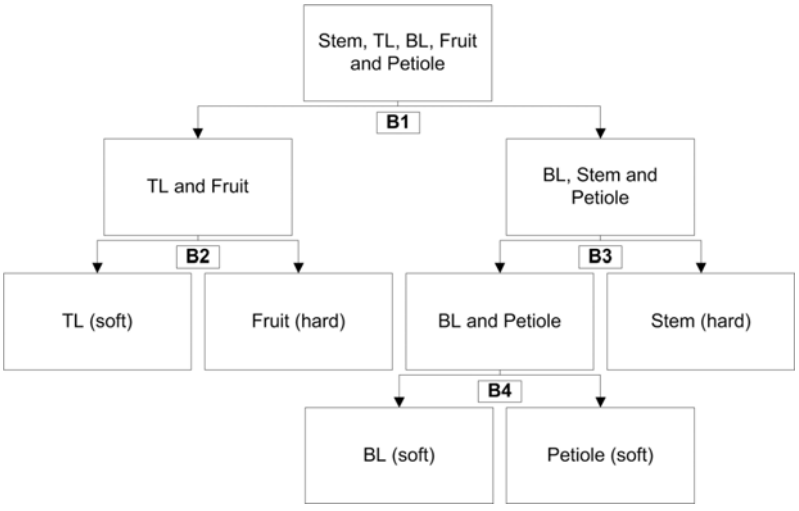


Figure 3-10. Approach B to reduce a five-class classification problem into four binary problems. The classification sequence is based on separability of the classes, i.e. best separable first.

Three different options were available to divide three classes for problem B3. In a classification where the 15 NDI spectral features were fed to the classifier, it turned out that stem and union of BL and Pet (P_{Rob} = 53.3) resulted in better separability than Pet and union of Stem and BL (P_{Rob} = 41.3) or BL and union of Stem and Pet (P_{Rob} = 35.0). Consequently, Stem and union of BL and Pet were taken as classes in problem B3 and finally BL and Pet were classified in problem B4.
Section 3.5.6 describes the result of feature selection on each of the four binary problems, for Approach A (A1, A2, A3 and A4) and for Approach B (B1, B2, B3 and B4).

3.5.6 Experiment 2c: Feature subset selection

To assess feature strength and the effectiveness of feature selection, we show detailed results of classification problem A1. Such detailed results are shown only for A1 to avoid repetition of similar findings in other classification problems.

Figure 3-11 demonstrates the result of feature selection for features listed in Section 3.3.4. Details of feature selection and performance are in Table 3-5. Only selected raw spectral (3), raw entropy (4), NDI spectral (4) and NDI entropy (3) were combined in the 'Cluster' set (Figure 3-11; Table 3-5). SAM and Mahalanobis distance were not added to the 'Cluster' feature set because, due to their poor performance, deteriorated performance of the 'Cluster' set. Furthermore, they were not investigated in successive classification problems.

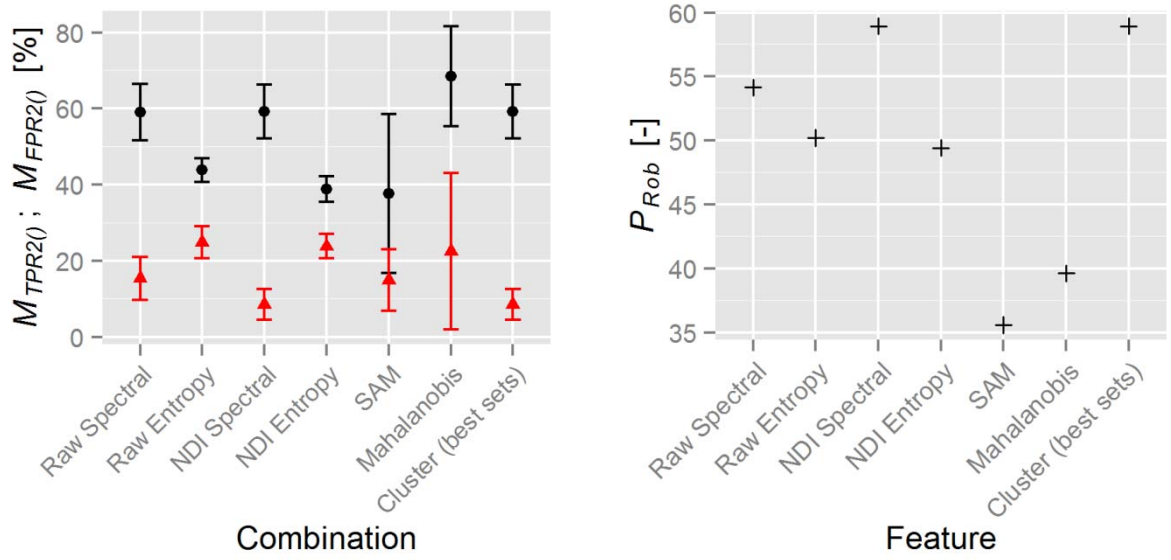


Figure 3-11. Performance per feature after feature selection, using the SFFS algorithm. Results are for classification problem A1: hard vs. soft obstacles. Mean and SD (N=10) of true-positive detection rates MTPR2() (●) and false-positive detection rates MFPR2() (▲) per feature (left) were used to calculate the performance measure PRob (+) per feature (right).

Fusion of the best four selected sets did not result in better performance. The NDI spectral features were again selected from the 'Cluster' set (P_{Rob} = 58.9). Despite the suggested potential of texture features (entropy) in the literature (Kapach et al. 2012), the best set only contained NDI spectral features.

Table 3-5. Classification performance results and classifier characteristics per feature type, for problem A1 (hard vs. soft obstacles).

	Raw Spectral	Raw Entropy	NDI Spectral	NDI Entropy	SAM	Mahalano-bis Distance	Cluster (best sets)
Features before selection (#)	6	6	15	15	2	2	14
Feature subset length after selection (-) / (features)	3 / 447; 624; 716	4 / 447; 562; 624; 692	4 / 447&624; 624&900; 692&716; 562&624;	3 / 447&562; 624&716; 624&900	2 / SAM (hard); SAM (soft)	1 / Mahalano-bis (hard)	4 / NDI Spectral: 447&624; 624&900; 692&716; 562&624;
Tree length before pruning (nodes)	4469	13247	9287	32995	39645	60823	9287
Tree length after pruning (nodes)	105	27	9	19	3	3	9
P_{Rob} (-)	54.1	50.2	58.9	49.4	35.6	39.6	58.9
$Acc_{Tot,2}$ (SD)	79.0 (4.8)	68.4 (6.0)	84.5 (5.0)	68.2 (5.5)	75.7 (10.0)	75.6 (15.0)	84.5 (5.0)
$M_{TPR\,2(hard)}$ ($SD_{TPR\,2(hard)}$)(%)	59.1 (7.4)	43.8 (3.1)	59.2 (7.1)	38.8 (3.4)	37.7 (20.9)	68.5 (13.1)	59.2 (7.1)
$M_{TPR\,2(soft)}$ ($SD_{TPR\,2(soft)}$)(%)	84.6 (5.7)	75.1 (4.2)	91.5 (4.0)	76.2 (3.2)	85.1 (8.1)	77.5 (20.5)	91.5 (4.0)

In general, feature selection improved classification results for all feature types. As an example, Figure 3-12 (left) demonstrates how the feature subset strongly influences classification performance.

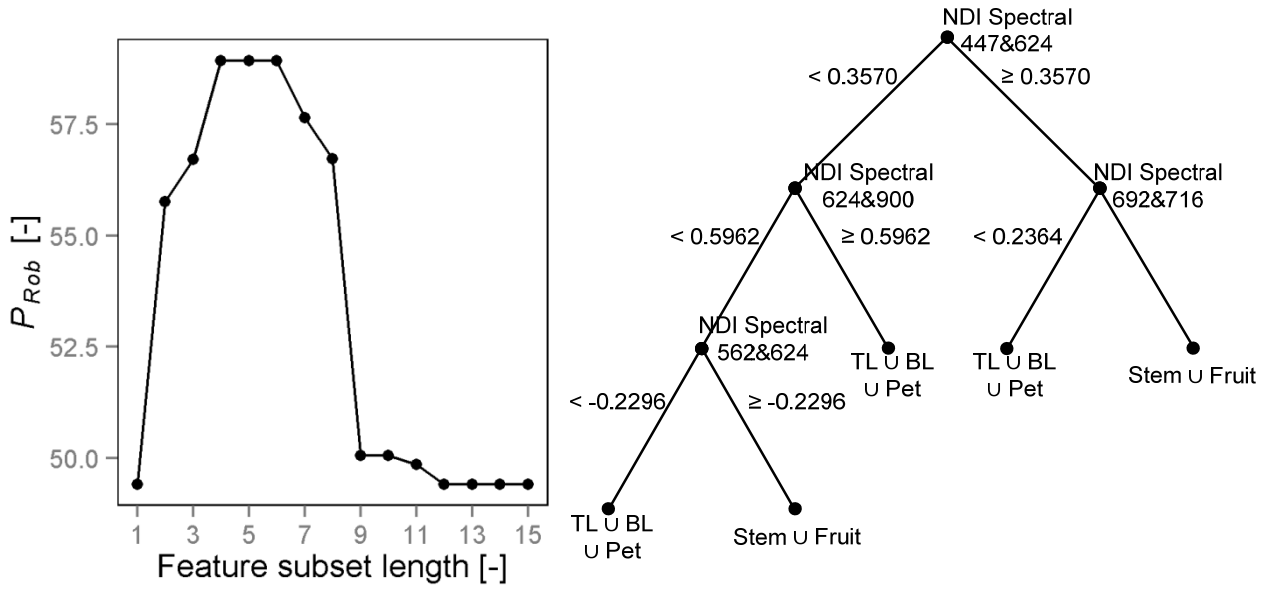


Figure 3-12. Classification performance of optimized feature subset lengths by SFFS, for NDI spectral features (left). Results are for classification problem A1: hard vs. soft obstacles. The decision tree for the optimal feature subset with length = 4 (right).

Clearly, the feature subset is optimal at feature subset length = 4. For lengths 5 and 6, CART is still able to find the optimized tree. Once the feature subset becomes longer than 6 features, performance decreases. This process of decreasing performance is caused by the suboptimal nature of classifiers and aggravates with high-dimensional feature sets, an effect which is often observed in machine learning (Jain 1997).

3.5.6.1 Performance of A1, A2, A3 and A4

Figure 3-11 already shows how means and standard deviations relate to the performance measure P_{Rob}. Therefore, only the performance P_{Rob} was shown in Figure 3-13, for the most useful features per classification problem.

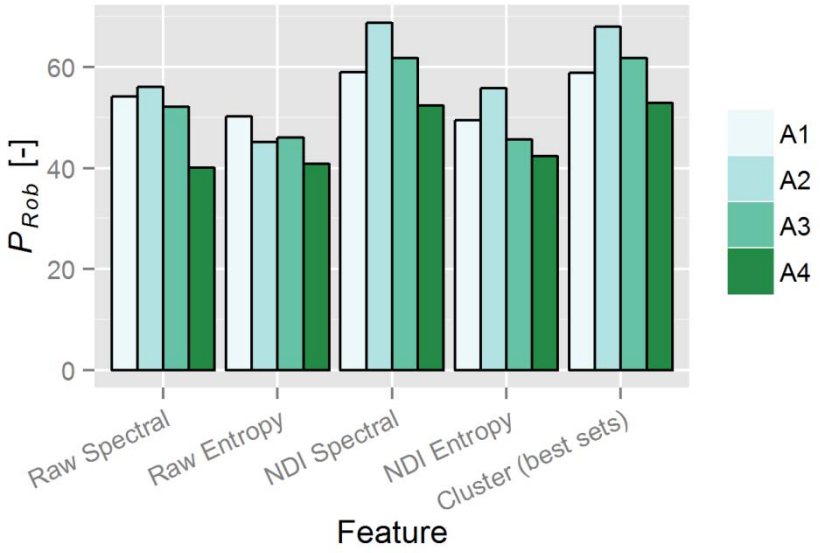


Figure 3-13. Performance of classification problems A1 (Stem ∪ Fruit vs. TL ∪ BL ∪ Pet), A2 (Stem vs. Fruit), A3 (TL vs. BL ∪ Pet) and A4 (BL vs. Pet) per feature. NDI Spectral was the strongest feature in A1, A2 and A3. The combination of the best sets was the strongest feature in A4. A2 was the easiest classification problem.

Classification problem A2 (stem vs. fruit) was clearly the easiest among the classification problems. Apparently the stem and fruit are easier to separate than other classes. Especially problem A4 is hard (BL vs. Pet).

In general, NDI spectral features were the strongest features in all classification problems. Features are given in parenthesis and are ordered on occurrence in the pruned tree: A1 comprised 4 NDI spectral features (447&624, 624&900, 692&716, 562&624); A2 comprised 4 NDI spectral features (562&624, 692&716, 447&692, 447&716); A3 comprised 5 NDI spectral features (692&900, 447&562, 624&716, 447&900 , 562&716; A4 comprised 4 NDI spectral features (692&716, 562&900, 562&624, 624&900) and 1 raw spectral feature (624). All six wavelengths were used. Texture features were never selected and one can conclude that texture features did not contribute relevant information in Approach A.

3.5.6.2 Performance of B1, B2, B3 and B4

The performance P_{Rob}, for the most useful features per classification problem, is shown in Figure 3-14.

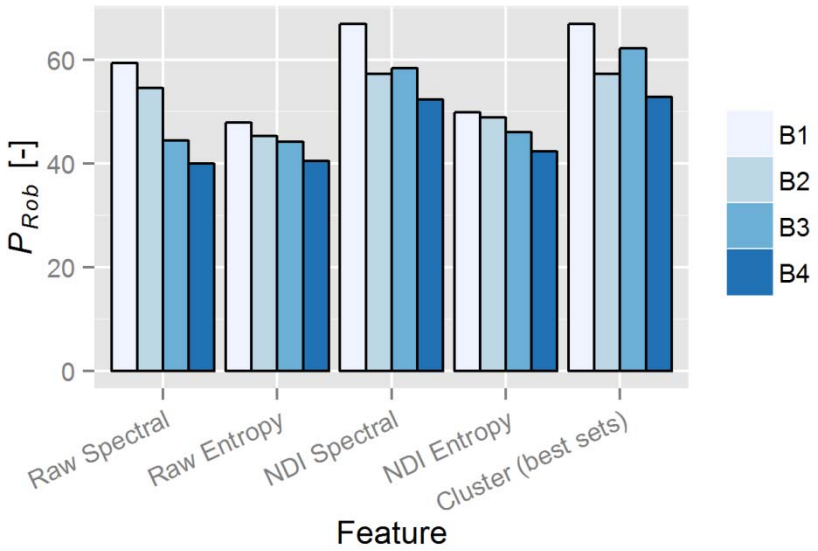


Figure 3-14. Performance of classification problems B1 (TL ∪ Fruit vs. Stem ∪ BL ∪ Pet), B2 (TL vs. Fruit), B3 (Stem vs. BL ∪ Pet) and B4 (BL vs. Pet) per feature. NDI Spectral was the strongest feature in B1 and B2. The combination of the best sets was the strongest feature in B3 and B4. B1 was the easiest classification problem.

Classification problem B1 (TL ∪ Fruit vs. TL ∪ BL ∪ Pet) was clearly the easiest among the classification problems.

In general, NDI spectral features were the strongest features in all classification problems. Features are given in parenthesis and are ordered on occurrence in the pruned tree: B1 comprised 4 NDI spectral features (562&900, 624&716, 447&624, 562&624); B2 comprised 3 NDI spectral features (447&624, 447&900, 692&716); B3 comprised 4 NDI spectral features (692&716, 716&900, 562&624, 447&900) and three raw spectral features (716, 900, 692); B4 comprised 4 NDI spectral features (692&716, 562&900, 562&624, 624&900) and 1 raw spectral feature (624). All six wavelengths were used. Texture features were not selected in Approach B and one can conclude that texture features did not contribute relevant information in Approach B.

3.5.7 Experiment 2d: Comparison of Approach A and B

The choice for the best approach was based on performance of hard vs. soft obstacle classification (Section 3.5.7.1). Furthermore, Section 3.5.7.2 describes the classification performance among the five plant part classes.

3.5.7.1 Performance of hard vs. soft obstacle

The performance of hard (Stem ∪ Fruit) vs. soft (TL ∪ BL ∪ Pet) obstacle classification are in Table 3-6.

Table 3-6. Classification performance of hard vs. soft obstacle classification by Approach A and B.

Performance measure	Approach A	Approach B
$M_{TPR2(hard)}$ ($SD_{TPR2(hard)}$) (%)	59.2 (7.1)	61.6 (8.2)
$M_{TPR2(soft)}$ ($SD_{TPR2(soft)}$) (%)	91.5 (4.0)	89.4 (5.6)
Robust-and-balanced accuracy P_{Rob} (-)	58.9	56.1
$Acc2_{Tot}$ (SD) (%)	84.5 (5.0)	83.1 (6.3)

Robust-and-balanced accuracy was greater for Approach A (P_{Rob} = 58.9) than for Approach B (P_{Rob} = 56.1), which is caused by lower standard deviations. Hence, Approach A resulted in a slightly better performance.

3.5.7.2 Performance of five plant part classes

Figure 3-15 displays classification performance per class, for both approaches.

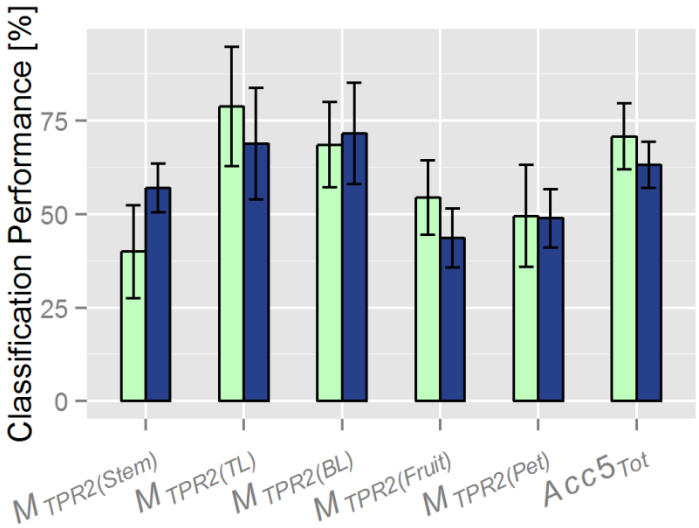


Figure 3-15. Classification performance of Stem, top of a leaf (TL), bottom of a leaf (BL), Fruit and Petiole (Pet) for Approach A ( ) and B ( ). Performance shown in terms of mean and SD (N=10) of true-positive detection rate $M_{TPR5()}$ and total accuracy $Acc5_{Tot}$. Total accuracy for Approach A (70.8%) is greater than for Approach B (63.1%).

Approach A and Approach B do not seem to differ strongly in performance. For Approach A, total accuracy (70.8) is greater than Approach B (63.1%) and one may conclude that Approach A is better. But, TPRs differ strongly per class. For instance, $M_{TPR5(Stem)}$ is greater in Approach B (57.0%) than Approach A (40.0%). The average of mean TPRs, $Acc5_{Bal}$, is similar for Approach A (58.2%) and Approach B (58.0%).

Moreover, Figure 3-15 clearly shows low TPRs for stem (<59%), fruit (<55%) and petiole (< 50%). Such low detection rates are not problematic as long as FPRs are also low. Figure 3-16 therefore displays how labelled ground truth regions were incorrectly classified.

Figure 3-16. Labelled ground truth (GT) regions were classified into 5 classes by Approach A (left) and Approach B (right). In both approaches, the highest mean (N=10) detection rate of a class occurred in the corresponding labelled class. For instance, stem detection was highest in GT Stem.

For both approaches, the highest detection rate of a class occurred in the corresponding labelled class. This finding indicates that post-processing can be promising because false-positive detections might be removed by morphological image processing such as opening or erosion. In conclusion, Figure 3-16 shows in which plant parts false-positive detections occur and, as a result, a strategy for post-processing can be established. For instance, many stem and petiole pixels were falsely assigned to TL and BL classes and one might consider post-processing techniques that convert false TL and BL detections into true stem and petiole detections.

3.5.8 Example of a classified scene

Results of classification by Approach A are in Figure 3-17 and Figure 3-18.

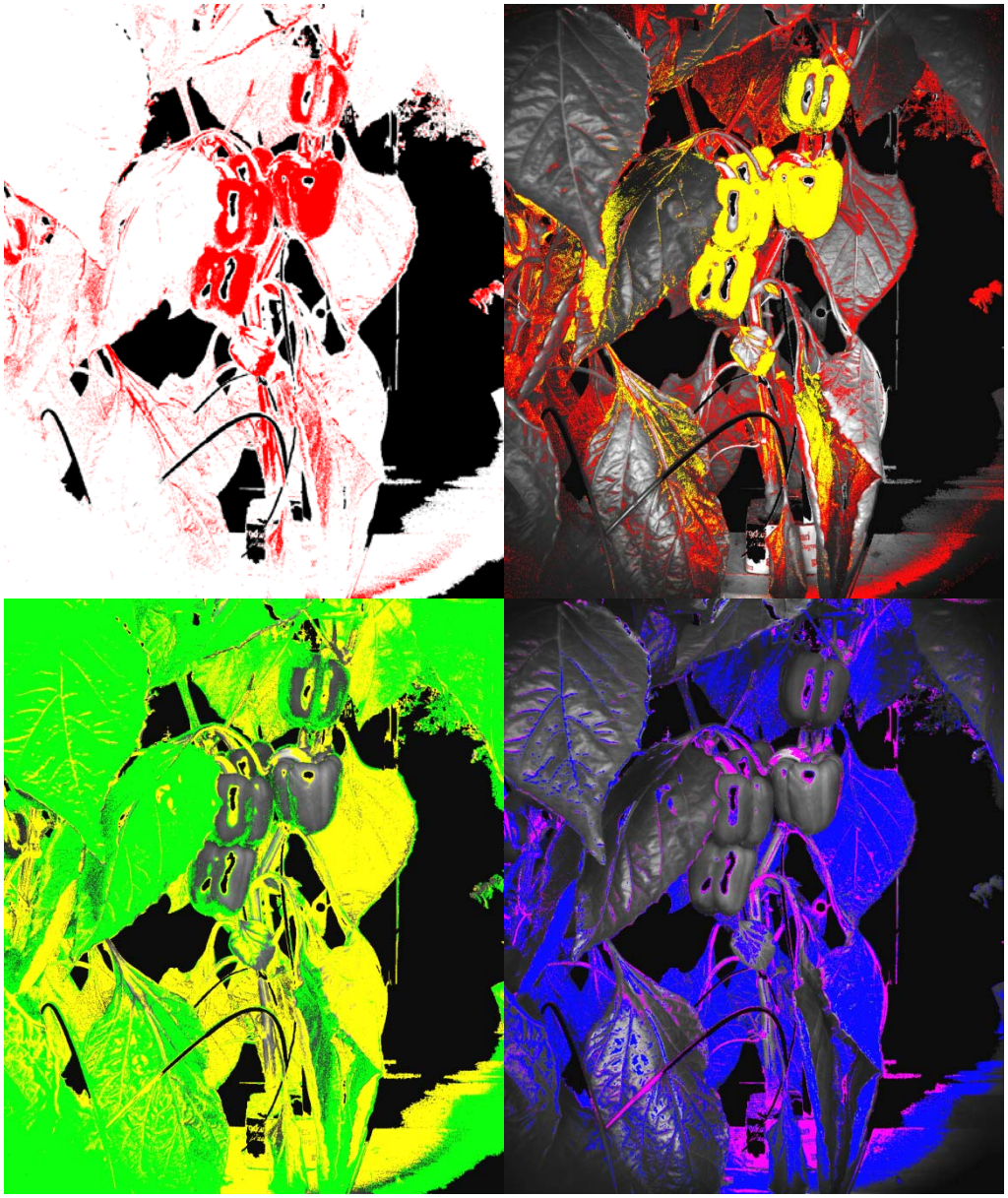


Figure 3-17. Result of pixel-based classification by Approach A. A1 (top-left) represents hard (red) vs. soft obstacles (white). A2 (top-right) represents stem (red) vs. fruit (yellow). A3 (bottom-left) represents top of a leaf (green) vs. bottom of a leaf ∪ petiole (orange). A4 (bottom-right) represents bottom of a leaf (blue) vs. petiole (magenta).

Total execution time for one scene in HALCON© was 2.3 s: 1.5 s for calculation of 12 NDI spectral features, 0.23 s for A1, 0.14 s for A2, 0.18 s for A3 and 0.29 s for A4. Figure 3-18 demonstrates that many false-positive detections occur in the scene.

Figure 3-18. Full classification of vegetation into five classes: stem (red), top of a leaf (green), bottom of a leaf (blue), fruit (yellow) and petiole (magenta). Black parts are either background or segmented non-vegetation objects.

3.6 Discussion

Classification performance was too low to build a reliable obstacle map for motion planning. Mean true-positive detection rate remains limited to 59.1% for hard obstacles and 91.5% for soft obstacles. As a result, 8.5% of the soft obstacles are falsely classified as hard obstacle and these false detections block a collision-free path for the harvesting manipulator. We were not able to remove these false detections, by a morphological 'opening' operation (Bac et al. 2013), without removing less than 50% of true-positive detections. An option to investigate in future work is to fuse Approach A and B in a judges-based approach.

Classification performance of related work conducted under varying lighting conditions is unknown and we only found performance of a study conducted indoor, under controlled lighting conditions. Humphries and Simonton (1993) performed pixel-based classification of geranium plant parts and report true-positive detection rates of 85% for leaves, 74% for stems and 21% for petioles. These slightly greater detection rates are probably due to controlled lighting conditions, and perhaps more constant camera-object distances. Humphries and Simonton (1993) improved detection rates to 97% for leaves, 95% for stems and 93% for petioles after addition of object-based features. Such an increase after addition of object-based features is also demonstrated by others (Wang et al., 2004) and addition of object-based features is therefore a task for future work.

Two causes may elucidate why classification performance was low. Firstly, the strong camera-object distance variation among scenes and within scenes resulted in illumination differences and, as a result, misclassifications. For instance, leaves located in front of a stem mainly received artificial illumination, whereas leaves behind a stem were further away and received a mixture of artificial illumination and natural lighting. Secondly, natural lighting varied (374-435 W/m2) and, in addition, light incidence on objects varied because light penetration is influenced by the varying shape and size of plant parts and of non-vegetation objects. Effects of both issues on classification accuracy was not studied and is an objective for future work. Two solutions can be investigated: use of a reference card to compensate for temporal variation in lighting conditions (Ting et al. 2012) and use of distance information to compensate for spatial variation in lighting conditions.

3.7 Conclusion

Unfortunately, classification performance was too low to build a reliable obstacle map for motion planning. Nevertheless, sweet-pepper plant vegetation was successfully segmented from the background using a threshold in a near-infrared wavelength (>900 nm). Subsequently non-vegetation objects, which included drippers, pots, sticks, construction elements and support wires, were removed by a threshold in the blue wavelength (447 nm). Remaining plant vegetation was classified into five classes using a CART decision tree with 46 features. For classification of hard (Stem ∪ Fruit) vs. soft (TL ∪ BL ∪ Pet) obstacles, NDI spectral features (P_{Rob} = 58.9) are stronger than raw spectral (P_{Rob} = 54.1), raw entropy (P_{Rob} = 50.2), NDI entropy (P_{Rob} = 49.4), SAM (P_{Rob} = 35.6), Mahalanobis distance (P_{Rob} = 39.6) or the clustered sets (P_{Rob} = 58.9).

Two approaches were derived to reduce the 5-class classification problem into four consecutive binary problems: Approach A was based on the problem nature and Approach B was based on separability of classes. Approach A is slightly better than Approach B because, on hard vs. soft obstacle classification, Approach A (P_{Rob} = 58.9) is equally accurate but more robust than Approach B (P_{Rob} = 56.1).

For Approach A, mean true-positive detection rate (standard deviation) among scenes was 59.2 (7.1)% for hard obstacles, 91.5 (4.0)% for soft obstacles, 40.0 (12.4)% for stems, 78.7 (16.0)% for top of a leaf, 68.5 (11.4)% for bottom of a leaf, 54.5 (9.9)% for fruit and 49.5 (13.6)% for petiole. These performance values are low and post-processing techniques will be investigated in future research. This research also shows benefits of the new performance measure P_{Rob}, which considers both balanced accuracy and classification variation among scenes. Use of P_{Rob} rendered the classifier more robust to variation among scenes because standard deviation among scenes reduced 59% for hard obstacles and 43 % for soft obstacles compared with balanced accuracy. This improvement is achieved, interestingly, because features probably less sensitive to lighting variations and varying plant-camera distances are selected.

Acknowledgements

We thank Yael Edan and Efi Vitzrabin for their useful comments and suggestion to improve this manuscript. We thank Anja Dieleman who provided the experimental greenhouse in

Wageningen. Bart van Tuijl is acknowledged for his contributions to the experimental set-up. This research was funded by the European Commission in the 7th Framework Programme (CROPS GA no. 246252) and by the Dutch horticultural product board (PT no. 14555).

References

Bac, C. W., Hemming, J., & Van Henten, E. J. (2013). *Pixel classification and post-processing of plant parts using multi-spectral images of sweet-pepper.* Paper presented at the IFAC Biorobotics Conference, Sakai, Japan,

Bac, C. W., Van Henten, E. J., Hemming, J., & Edan, Y. (2014). Harvesting robots for high-value crops: state-of-the-art review and challenges ahead. *Journal of Field Robotics, DOI: 10.1002/rob.21525*.

Blum, A. L., & Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial Intelligence, 97*(1-2), 245-271.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Boca Raton, FL: Chapman & Hall/CRC press.

Bulanon, D. M., Burks, T. F., & Alchanatis, V. (2010). A multispectral imaging analysis for enhancing citrus fruit detection. *Environmental Control in Biology, 48*(2), 81-91.

Center, S. E. (2008). *Solar Radiation Hand Book*. India: MNRE, Indian Metrological Department.

Cieslak, D. A., & Chawla, N. V. (2008). Learning decision trees for unbalanced data. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 5211 LNAI, pp. 241-256).

Davies, E. R. (2009). The application of machine vision to food and agriculture: A review. *Imaging Science Journal, 57*(4), 197-217.

Deng, J., Li, J., & Zou, X. (Vision). Extraction of litchi stem based on computer vision under natural scene. In *Proceedings - International Conference on Computer Distributed Control and Intelligent Environmental Monitoring, CDCIEM 2011, 2011* (pp. 832-835).

DeVries, R. J. (Vision). Spatial modelling using the Mahalanobis statistic: Two examples from the discipline of plant geography. In *MODSIM05 - International Congress on Modelling and Simulation: Advances and Applications for Management and Decision Making, Proceedings, 2005* (pp. 1368-1374).

Dey, D., Mummert, L., & Sukthankar, R. (Vision). Classification of plant structures from uncalibrated image sequences. In *Proceedings of IEEE Workshop on Applications of Computer Vision, 2012* (pp. 329-336).

Filella, I., & Penuelas, J. (1994). The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *International Journal of Remote Sensing, 15*(7), 1459-1470.

Fleck, S., Van Der Zande, D., Schmidt, M., & Coppin, P. (2004). Reconstructions of tree structure from laser-scans and their use to predict physiological properties and processes in canopies. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36*(Part 8/W2), 119-123.

Fransen, T. S. (2011). Determination of specific reflectance curves for different parts of the sweet-pepper plant. Wageningen, The Netherlands: BSc Thesis, Farm Technology Group, Wageningen University, Open-access via: http://edepot.wur.nl/176380.

Hemming, J., Bac, C. W., & Van Tuijl, B. A. J. (2011). CROPS project deliverable 5.1: Report with design objectives and requirements for sweet-pepper harvesting. Wageningen, The Netherlands: Wageningen UR Greenhouse Horticulture.

Huang, Y. J., & Lee, F. F. (2008). Classification of Phalaenopsis plantlet parts and identification of suitable grasping point for automatic transplanting using machine vision. *Applied Engineering in Agriculture, 24*(1), 89-99.

Humphries, S., & Simonton, W. (1993). Identification of Plant Parts Using Color and Geometric Image Data. *Transaction of the ASABE, 36*(5), 1493-1500.

Jain, A. (1997). Feature selection: evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*(2), 153-158.

Jiménez, A. R., Ceres, R., & Pons, J. L. (2000). A survey of computer vision methods for locating fruit on trees. *Transactions of the American Society of Agricultural Engineers, 43*(6), 1911-1920.

Jovicich, E., Cnatliffe, D. J., Sargent, S. A., & Osborne, L. S. (2004). Production of Greenhouse-Grown Peppers in Florida. Gainesville, FL: University of Florida, IFAS Extension.

Kapach, K., Barnea, E., Mairon, R., Edan, Y., & Ben-Shahar, O. (2012). Computer Vision for Fruit Harvesting Robots - State of the Art and Challenges Ahead. *Int. J. of Computational Vision and Robotics, 3*(1/2), 4-34.

Kavdır, İ., & Guyer, D. E. (2004). Comparison of Artificial Neural Networks and Statistical Classifiers in Apple Sorting using Textural Features. *Biosystems Engineering, 89*(3), 331-344.

Linker, R., Cohen, O., & Naor, A. (2012). Determination of the number of green apples in RGB images recorded in orchards. *Computers and Electronics in Agriculture, 81*, 45-57.

Lu, Q., Tang, M., & Cai, J. (Vision). Obstacle recognition using multi-spectral imaging for citrus picking robot. In *Proceedings - PACCS 2011: 2011 3rd Pacific-Asia Conference on Circuits, Communications and System, Wuhan, China, 17-18 July 2011 2011* (pp. 1-5).

McCarthy, C. L., Hancock, N. H., & Raine, S. R. (2010). Applied machine vision of plants: A review with implications for field deployment in automated farming operations. *Intelligent Service Robotics, 3*(4), 209-217.

Moisen, G. G. (2008). Classification and Regression Trees. In *Encyclopedia of Ecology* (pp. 582-588). Oxford: Academic Press.

Nakariyakul, S., & Casasent, D. P. (2009). An improvement on floating search algorithms for feature subset selection. *Pattern recognition, 42*(9), 1932-1940.

Noble, S., & Li, D. (2012). *Segmentation of Greenhouse Cucumber Plants in Multi-spectral Imagery.* Paper presented at the International Conference of Agricultural Engineering, CIGR-Ageng, Valencia, Spain, 8-12 July

Noordam, J. C., Hemming, J., van Heerde, C., Golbach, F., van Soest, R., & Wekking, E. (2005). Automated Rose Cutting in Greenhouses with 3D Vision and Robotics: Analysis of 3D Vision Techniques for Stem Detection. *Acta Hort. (ISHS), 691*, 885-892.

Ting, Y., Kondo, N., & Wei, L. (2012). Sunlight Fluctuation Compensation for Tomato Flower Detection Using Web Camera. *Procedia Engineering, 29*(0), 4343-4347.

Unay, D., Gosselin, B., & Debeir, O. Apple stem and calyx recognition by decision trees. In *Proceedings of the 6th IASTED International Conference on Visualization, Imaging, and Image Processing, VIIP 2006, 2006* (pp. 549-552).

Van Henten, E. J., Van Tuijl, B. A. J., Hoogakker, G. J., Van Der Weerd, M. J., Hemming, J., Kornet, J. G., et al. (2006). An Autonomous Robot for De-leafing Cucumber Plants grown in a High-wire Cultivation System. *Biosystems Engineering, 94*(3), 317-323.

Wickens, T. D. (2001). Ch 2. The equal-variance Gaussian model. In *Elementary Signal Detection Theory*. USA: Oxford University Press.

Wieland, A., & Wallenburg, C. M. (2012). Dealing with supply chain risks: Linking risk management practices and strategies to performance. *International Journal of Physical Distribution and Logistics Management, 42*(10), 887-905.

Yang, C., Everitt, J. H., & Bradford, J. M. (2008). Yield estimation from hyperspectral imagery using spectral angle mapper (SAM). *Transactions of the ASABE, 51*(2), 729-737.

Chapter 4 – Pixel classification and post-processing of plant parts using multi-spectral images of sweet-pepper

C.W. Bac[a,b], J. Hemming[a], E.J. van Henten[a,b]

Affiliations:
[a] Wageningen UR Greenhouse Horticulture, Wageningen University and Research Centre,
[b] Farm Technology Group, Wageningen University and Research Centre

This chapter was published in proceedings of the *IFAC Bio-Robotics Conference* (2013), 27-29 March, Sakai, Japan, Vol 1., p. 150-155

Abstract

As part of the development of a sweet-pepper harvesting robot, obstacles should be detected. Objectives were to classify sweet-pepper vegetation into five plant parts: stem, top of a leaf (TL), bottom of a leaf (BL), fruit and petiole (Pet); and to improve classification results by post-processing. A multi-spectral imaging set-up with artificial lighting was developed to acquire images of sweet-pepper plants. The background was segmented from the vegetation and vegetation was classified into five plant parts, through a sequence of four two-class classification problems. True-positive detection rate/scaled false-positive rate achieved, on a pixel basis, were 40.0/179% for stem, 78.7/59.2% for top of a leaf (TL), 68.5/54.8% for bottom of a leaf (BL), 54.5/17.2% for fruit and 49.5/176.0% for petiole (Pet), before post-processing. The opening operations applied were unable to remove false stem detections to an acceptable rate. Also, many false detections of TL (>10%), BL (14%) and Pet (>15%) remained after post-processing, but these false detections are not critical for the application because these three plant parts are soft obstacles. Furthermore, results indicate that TL and BL can be distinguished. Green fruits were post-processed using a sequence of fill-up, opening and area-based segmentation. Several area-based thresholds were tested and the most effective threshold resulted in a true-positive detection rate, on a blob basis, of 56.7 % and a scaled false-positive detection rate of 6.7 % for green fruits (N=60). Such fruit detection rates are a reasonable starting point to detect obstacles for sweet-pepper harvesting. But, additional work is required to complement the obstacle map into a complete representation of the environment.

4.1 Introduction

This research is part of the EU funded CROPS project, 'Clever Robots for Crops', in which a sweet-pepper harvesting robot will be developed. The manipulator of this harvesting robot should approach a target (fruit or peduncle) while avoiding obstacles. These obstacles should be detected, and eventually localized. Obstacle detection is the scope of this article and obstacles comprise supporting wires, construction elements and plant parts (stem, leaf, fruit and petiole). We separated obstacles in hard obstacles and soft obstacles. Hard obstacles (stems, fruits, supporting wires, construction elements) should be avoided by a manipulator or end-effector, whereas soft obstacles (leaves and petioles) can be touched or pushed aside. In addition, the top of a leaf and the bottom of a leaf were discerned to be able to control the motion of a pushed leaf. Pushing a top of a leaf, namely, will usually result in downward motion of the leaf and pushing a bottom of a leaf will result in an upward motion of the leaf.

Vision-based plant part classification studies under natural lighting conditions are scarce. Two studies describe classification of cucumber plant parts into leafs, stems and fruits: a study on a cucumber leaf picking robot using two near-infrared wavelengths (Van Henten et al., 2006) and a multi-spectral imaging study in which several wavelengths and sensors are compared (Noble and Li, 2012). Unfortunately, both studies lack classification performance values. The article most closely related to the work presented here is classification of grape foliage into leaves, branches and fruits (green or coloured) using RGB cameras under natural

lighting conditions. For green grapes, the true-positive rate was 91.9 % with a false-positive rate of 2.7 % (Dey et al., 2012).

Objectives were to 1) classify sweet-pepper vegetation into five plant parts: stem, top of a leaf (TL), bottom of a leaf (BL), fruit and petiole (Pet); 2) improve classification results by post-processing techniques.

4.2 Materials and Methods

A multi-spectral imaging set-up was used to acquire images of sweet-pepper plants (Figure 4-1). The crop cultivar was "Viper" and only unripe green fruits occurred in the scenes. A 5 megapixel monochrome camera was used in combination with a filter wheel containing filters with the following wavelengths (bandwidth): 447 (60) nm, 562 (40) nm, 624 (40) nm, 692 (40) nm, 716 (40) nm, 950 (100) nm. For each scene, a set of six 8-bit images was acquired with a resolution of 2082 by 2493 pixels. In total, 12 scenes were recorded under outdoor lighting conditions with additional artificial lighting (Figure 4-1).

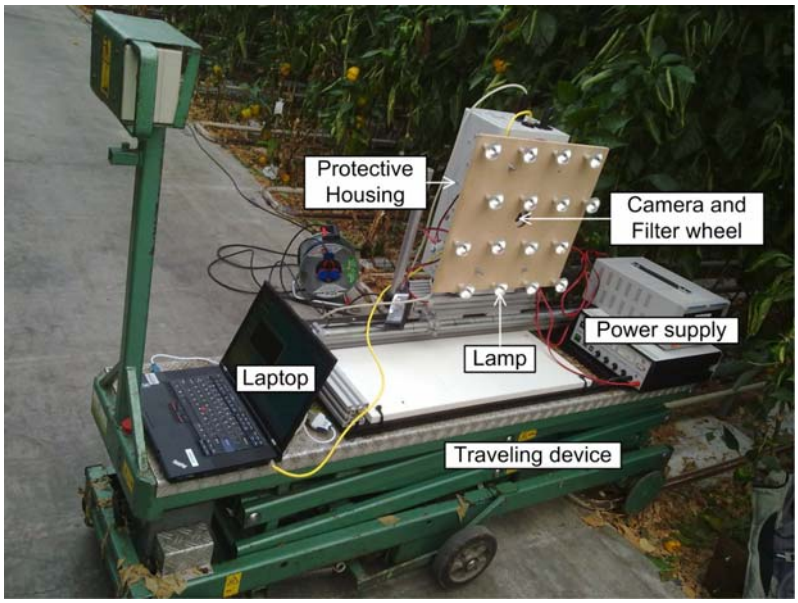


Figure 4-1. Overview of the experimental set-up in the greenhouse.

Images were recorded from 11:00 am until 11:30 am under a clear sunny sky. Outdoor solar irradiance was measured during image recording and varied between 374 and 435 W/m^2. Camera-stem distance was on average 80 cm and varied in a range of 63 cm to 109 cm among scenes.

Ground truth data was obtained through manual labelling of pixels in recorded images. In total, $14.6 \cdot 10^6$ pixels were labelled and this number comprised 29.4% of the vegetation present in the 12 scenes. Labelled pixels were mostly leaves – TL (54.6%), BL (22.4%). Other labelled pixels were fruits (15.6%), stems (3.7%) and petioles (3.7%). A large part of the vegetation pixels, mostly leaves, was not labelled (70.6%) because we assumed labelled leaves already represented the majority of the leaf variation occurring in the scene. In addition, labelling all samples would increase the computational load dramatically during training of the classifier. An example of a labelled scene is in Figure 4-2.

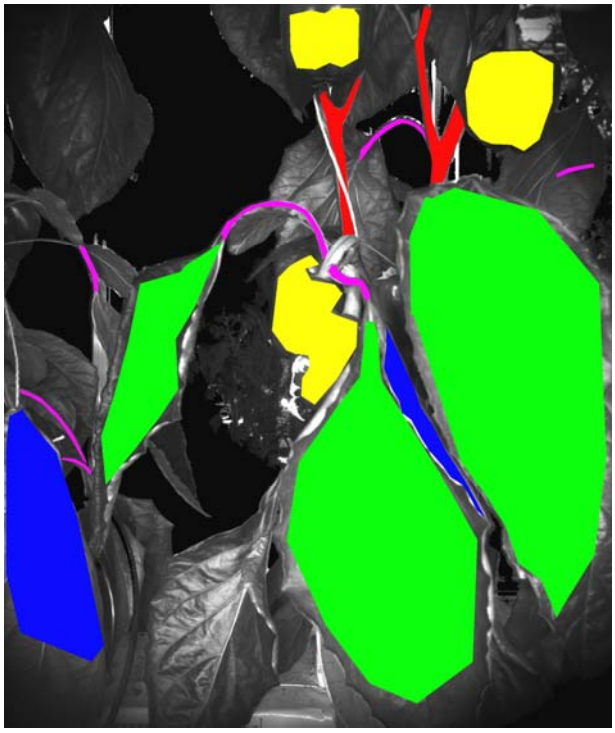


Figure 4-2. Labelled image comprising five classes: stem (red), top of a leaf (green), bottom of a leaf (blue), green fruit (yellow) and petiole (purple).

4.2.1 Performance assessment

Performance was assessed as a binary classification problem. Hence, we compared the detection rates for one class versus the union of the other four classes. As a result, we calculated five 2 by 2 sized confusion matrixes. The elements in each matrix describe true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN) detected pixels (Bradley 1997). Based on these elements, a true-positive detection rate TPR (Equation (4-1)) and a scaled false-positive detection rate SFPR (Equation (4-2)) were calculated.

$$TPR = \frac{100 \cdot TP}{TP + FN} \quad (\%) \qquad (4\text{-}1)$$

$$SFPR = \frac{100 \cdot FP}{TP + FN} \quad (\%) \qquad (4\text{-}2)$$

Note that this measure SFPR is identical to how authors of previous and recent fruit detection literature refer to 'false-positive rate' (Jiménez et al. 2000; Bulanon et al. 2010; Linker et al. 2012). Such a definition of false-positive rate is, however, confusing because in other research disciplines false-positive rate is calculated as FP/(FP+TN) (Mackinnon 2000; Gu et al. 2009). Hence, we use scaled false-positive rate instead of false-positive rate to avoid ambiguous definitions of false-positive rate.

Similar to authors of fruit detection literature, we consider scaled false-positive rate to be a more useful measure to report than false-positive rate because false-positive detections are expressed in terms of the class to be detected. False detection rate is therefore not biased by unbalanced class sizes, as is the case with false-positive rate. A drawback of scaled false-positive rate is, however, that rates can exceed 100 %.

4.2.2 Pixel classification of vegetation

Pixels were classified using Classification And Regression Trees (CART) in MATLAB® 2007b in combination with the Sequential Floating Forward Selection (SFFS) feature selection algorithm (Pudil et al. 1994). Vegetation classification was performed on a computer with an Intel Core i5 CPU 2.4 GHz Quad core processor including 4 GB memory.

The first step in the image processing sequence was to remove the background to obtain remaining pixels of interest. To remove the background, a threshold operation was applied on the 900 nm image (gray-value threshold: ≤ 27) and holes in the background region were filled by a fill-up operation. Subsequently, overexposed regions, mostly construction elements and supporting wires, were removed by a threshold (>139) on the 447 nm image. As a result, only vegetation remained. The vegetation was classified into five plant parts: stem, top of a leaf, bottom of a leaf, green fruits and petiole. Pixel-based features were used, i.e. Normalized Difference Index (NDI) (Davies 2009) and raw gray-values. In total 15 NDI features were calculated from the six wavelengths. Consequently, 15+6 = 21 features were used.

Classification of the five plant parts was split into four two-class classification problems instead of one five-class classification problems because this approach resulted in greater accuracy in previous research (Kavdır and Guyer 2004). Figure 4-3 shows which plant part classes were separated in each classification problem.

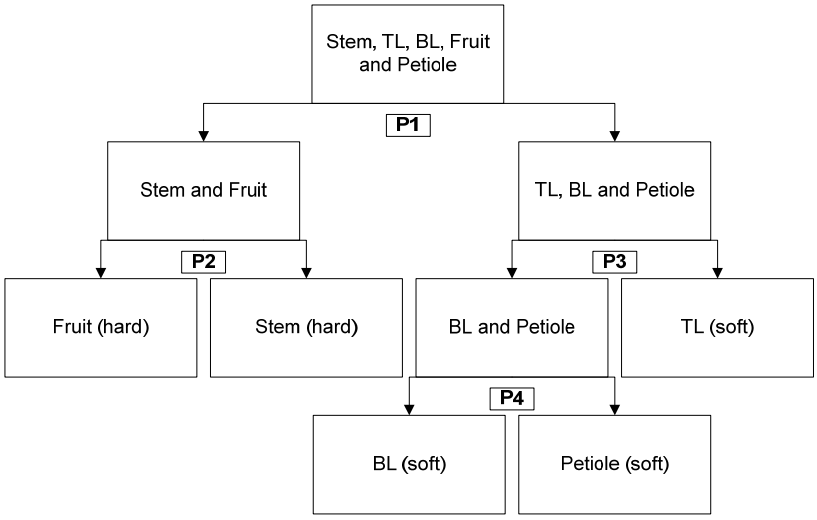


Figure 4-3. Approach taken to reduce a five-class classification problem into four two-class problems: P1, P2, P3 and P4.

The decision trees for each two-class problem were trained and pruned in MATLAB®. Two scenes were used for training and ten scenes were used for testing. Before training, class sample sizes were balanced such that the classifier would equally favour both classes. After pruning, trees were exported to Halcon (MVTEC 2012). As a result, both classification and post-processing techniques were applied in Halcon.

4.2.3 Post-processing applied to each class

For post-processing, morphological image processing was first applied to improve results of pixel classification. We applied an opening operation, with rectangular mask sizes of 3x3, 5x5 and 7x7 pixels, to each of the five plant parts. Subsequently, results were compared with

unprocessed data, in terms of true-positive detection rate and scaled false-positive detection rate. Note that only labelled regions in the image were assessed for performance of classification and post-processing on a pixel basis.

4.2.4 Green fruit detection

Detection of the fruits, turned out to be more successful than the other plant parts and therefore additional post-processing was applied to improve fruit detection. We applied a sequence of fill-up, opening (circular mask with a radius 2.5 of pixels), connection and area-based segmentation. The circular mask was chosen because fruits mimic a circular shape more than a rectangular shape. The value for area-based segmentation was tested in a range of 1000 to 13000 pixels, with steps of 2000 pixels, to determine the effect on true-positive fruit detection rate and scaled false-positive detection rate.

Performance of fruit detection was not only determined on a pixel basis, but also on a blob (or region) basis. To compare results of blob analysis to the literature, the full image was classified instead of only labelled regions. Separation of fruit clusters into individual fruits is a challenging task in fruit detection (Linker et al. 2012) and was not performed in this research. To calculate the number of individual fruits detected we manually counted the number of fruits present in a cluster, before and after detection. If, for instance, a blob covered three fruits, we counted such a blob as three successfully detected fruits. We counted a fruit as successfully detected if at least some part (>0 %) of the visible fruit surface was detected. In total, 60 distinct fruits were visible (partially or completely) in the ten test scenes.

4.3 Results

4.3.1 Pixel classification of vegetation

NDI features turned out to be the strongest features to classify plant parts. True-positive detections with NDI features were about 4-6 % greater than for raw spectral features. The result of plant part classification, using NDI features, is shown in Figure 4-4.

Figure 4-4. Classification of vegetation into five classes: stem (red), top of a leaf (green), bottom of a leaf (blue), fruit (yellow) and petiole (magenta). Black parts are either background regions or overexposed regions.

Figure 4-4 shows that performance of classification based on only pixel information is limited because many false-positive detections occur in the scene. At pixel level, average true-positive detection rates (standard deviations) among scenes (N=10) are: 59.2 (7.1)% for hard obstacles and 91.5 (4.0)% for soft obstacles. Furthermore, detection rates per class are: 40.0 (12.4)% for stem, 78.7 (16.0)% for top of a leaf, 68.5 (11.4)% for bottom of a leaf, 54.5 (9.9)% for fruit and 49.5 (13.6)% for petiole.

Total execution time for one scene in Halcon was 2.3 s: 1.5 s for calculation of 12 NDI features and 0.8 s for decision tree classification. Calculation time of post-processing methods reported in the following sections was in the order of 1-20 ms and therefore negligible to the time required for classification.

4.3.2 Post-processing applied to each class

To assess the effect of post-processing on true-positive detection rates and scaled false-positive detection rates, results are shown for three post-processing operations and for unprocessed data (Figure 4-5).

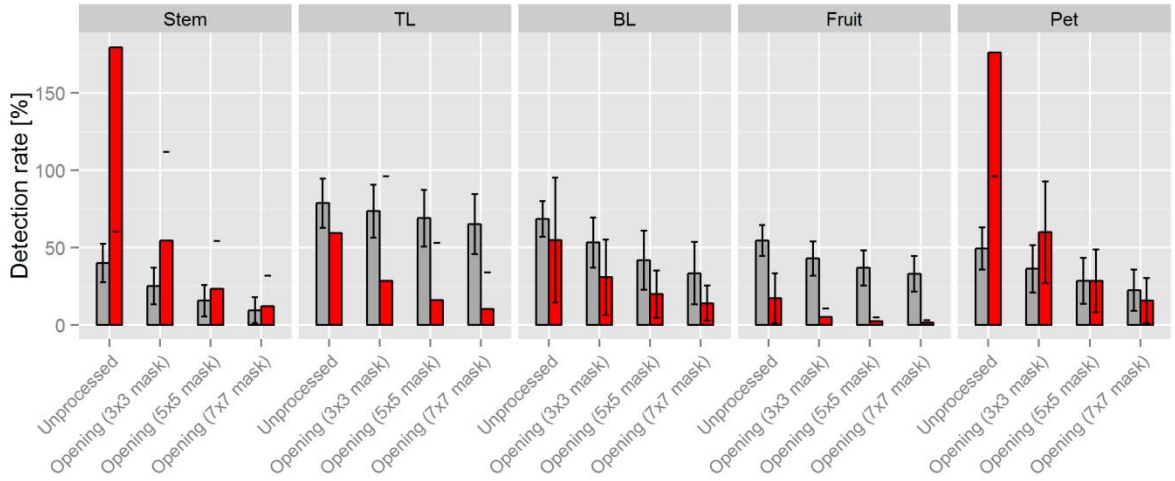


Figure 4-5. Mean (N=10 scenes) and SD of true-positive detection rate (▭) and scaled false-positive detection rate (▭), on a pixel basis, for unprocessed pixels and for three post-processing operations: opening with mask sizes 3x3, 5x5 or 7x7 pixels. These results are shown for each of the five plant part classes: Stem, Top of a Leaf (TL), Bottom of a Leaf (BL), Fruit and Petiole (Pet). Post-processing of fruits resulted in the lowest scaled false-positive rate compared with other classes: 5.2 % for a 3x3 mask, 2.5 % for a 5x5 mask and 1.4 % for a 7x7 mask.

Post-processing improved ratio between true-positive detection rate and scaled false-positive detection rate for all classes. Yet this ratio does not exceed a value of one for the stem and petiole, which indicates that stem and petiole detection was difficult. Furthermore, standard deviation (SD) of detection rate among scenes does not decrease significantly, except for false-positive fruit detections. Apparently, these opening operations do not decrease the variability of detection among scenes.

Figure 4-6 demonstrates that many false-positive stem detections remain after post-processing. Such false stem detections are unacceptable because the motion planning algorithm of the robot arm considers these false 'hard obstacle' detections as forbidden areas during calculation of a collision-free path, whereas in reality a path can be planned through these soft obstacles (e.g. leaf).

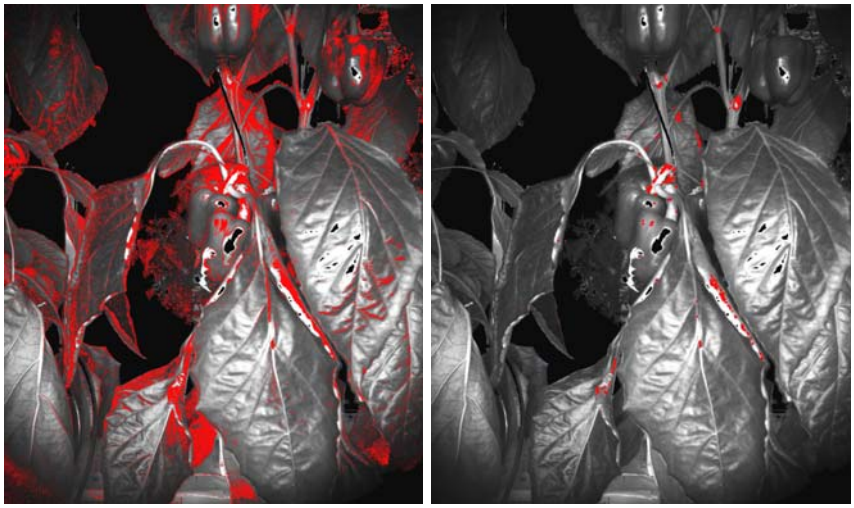


Figure 4-6. Stem detections after classification and unprocessed (left). Stem detections after an opening operation (7x7) mask (right). Despite the opening operation, many false detections occur on the leaf and fruit, and many true detections disappear.

Although ratio between true-positive detection rate and scaled false-positive detection rate increased after post-processing (Figure 4-5), scaled false-detection rates remain greater than 10 % for TL and greater than 14 % for BL. An example of an opening operation applied to TL is in Figure 4-7.

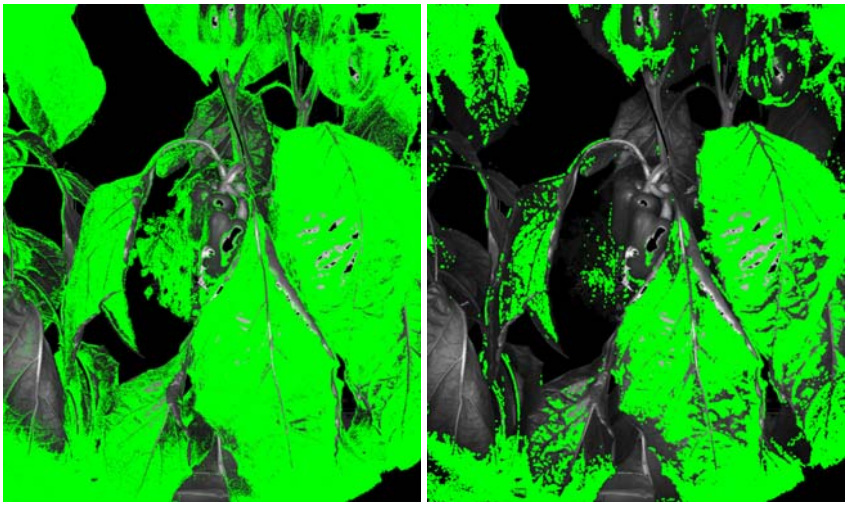


Figure 4-7. Top of a Leaf (TL) detections after classification and unprocessed (left). TL detections after an opening operation (7x7) mask (right). Despite the opening operation, many false detections occur on the stem, bottom of a leaf, fruit and petiole.

Figure 4-7 (right) demonstrates that few false TL detections remain after an opening operation (in bottom left part of image). Such remaining false detection may be removed through area-based segmentation. Similarly, few false BL detections occurred in the TL (not shown). Hence, these results indicate that TL and BL can be distinguished.

4.3.3 Green fruit detection

In contrast to detection of other plant parts, a scaled false-positive detection rate of <5.2%, on a pixel basis, was achieved for fruit detection (Figure 4-5). The remaining blobs were further

processed and Figure 4-8 demonstrates that many false-positive blob detections were removed, whereas most true-positive fruit detections remained.

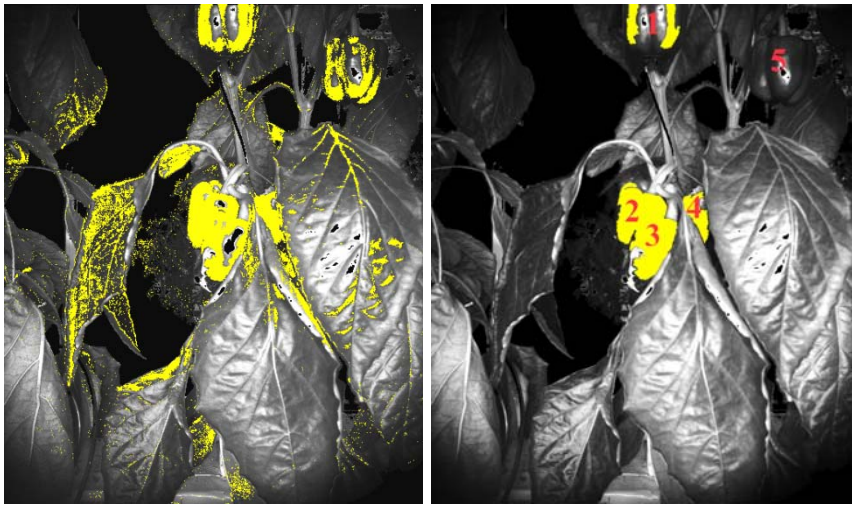


Figure 4-8. Classified fruit pixels (left) and the result after fill-up, opening, connection and area-based segmentation (right). The fruit visible in the top right (no. 5) is unfortunately not detected after a sequence of fill-up, opening and area-based (>5000 pixels) segmentation. Hence, four (no. 1-4) of the five fruits (no. 1-5) were detected.

Results of different thresholds, for area of blobs, are in Table 4-1.

Table 4-1. True-positive detection rate (TPR) and scaled false-positive detection rate (SFPR) of green fruits, on a blob basis, for different area-based thresholds.

Area-based threshold (>pixels)	True-positive detection rate (%)	Scaled false-positive detection rate (%)	Ratio of TPR/SFPR (-)
1000	88.3	126.7	0.70
3000	83.3	55.0	1.51
5000	78.3	30.0	2.61
7000	68.3	18.3	3.73
9000	68.3	13.3	5.14
11000	61.7	11.7	5.27
13000	61.7	10.0	6.17
15000	56.7	8.3	6.83
17000	56.7	6.7	8.46
19000	55.0	6.7	8.20
21000	51.7	6.7	7.72

We chose the threshold with greatest ratio of true-positive detection rate vs. scaled false-positive detection rate because here the post-processing approach is most effective in removing false detections while preserving true fruit detections. Greatest ratio (8.46) of true-positive vs. false-positive detections is achieved at an area-based threshold of 17000 pixels: TPR = 56.7% and SFPR = 6.7%.

4.4 Discussion

Accurate hard obstacle detection is more critical for the application than accurate soft obstacle detection because false hard obstacle detection limit the free workspace of a robot manipulator, whereas false soft obstacle detections do not. A motion without hitting either soft or hard obstacles is ideal, but, if required, a motion can be planned through soft obstacles. A motion through hard obstacles is, however, unacceptable because a damage to the stem affects growth of the plant and a damage to the fruit causes yield loss.

Detection rate achieved for soft obstacles (top of a leaf, bottom of a leaf and petiole) is limited, but acceptable for this application.

Regarding hard obstacles, stem detection rate is too small for a useful obstacle map because, after an opening operation (7x7 mask), a TPR of 9.5% and a SFPR of 12.2% was achieved. For green fruit detection, a TPR of 56.7% and a SFPR of 6.7% was achieved. This performance is far from perfect, but probably a reasonable starting point to find a collision-free path to a ripe fruit. However, with such fruit detection rates, the obstacle map will not be complete and additional detection is therefore required. Such detection may be obtained from additional sensors on the end-effector. Also, the manipulator should be able to adapt its path during motion if the sensor detects additional obstacles, or if a false-positive obstacle detection is re-detected as a true negative detection.

Fruit detection rate is worse than related work regarding green apple detection. Linker et al. (2012) achieved a TPR of 88% and a SFPR of 25% under intense natural lighting and a TPR of 95% and a SFPR of 4% under diffuse natural lighting. The methods used in this article are rather basic compared with methods used for apple detection, which may elucidate why those authors achieved a better performance. However, for green citrus detection, Kurtulmus et al. (2011) achieved a TPR of 75% and a SFPR of 27%, which is comparable to results achieved in this research.

The artificial lighting used mitigates the effect of outdoor lighting variations. Nevertheless, classification performance is rather limited and, in addition to outdoor lighting variations, varying plant-camera distances may elucidate why classification performance was rather limited. In related work, geranium cuttings were classified based on RGB images recorded indoor and, seemingly, camera-object distances were more constant (Humphries and Simonton 1993). These authors achieved a TPR of 85% for leaf, 21% for petiole and 74% for stem, on a pixel basis. They did not report false-positive detection rate, which renders it hard to compare results with this research. Nevertheless, their true-positive rates are slightly higher than results reported here, except for petiole classification. Yet, our study probably benefitted from the near-infrared wavelengths used and the performance gap may have been larger if we would have recorded RGB images combined with such varying lighting conditions and varying camera-object distances. In summary, one can observe that pixel-based classification is limited and addition of object-based features can improve performance because, after addition of object-based features, Humphries and Simonton (1993) achieved a TPR of 97% for leaf, 95% for petiole and 94% for stem.

4.5 Concluding remarks

This study is one of the first multi-spectral imaging studies, under varying lighting conditions, in which detection rates are reported. True-positive detection rate/scaled false-positive rate achieved, on a pixel basis, are: 40.0/179% for stem, 78.7/59.2% for top of a leaf (TL), 68.5/54.8% for bottom of a leaf (BL), 54.5/17.2% for fruit and 49.5/176.0% for petiole (Pet). The opening operations applied were unable to remove false-positive stem detections to an acceptable rate. An improved stem detection algorithm is therefore a task for future work. Also, many false detections of TL, BL and Pet remained after post-processing, but these false detections are not critical for the application because these three plant parts are soft obstacles. Furthermore, results indicate that TL and BL can be distinguished.

Green fruits were post-processed using a sequence of fill-up, opening and area-based segmentation. Several area-based thresholds were tested and the most effective threshold resulted in a true-positive detection rate of 56.7 % and a scaled false-positive detection rate of 6.7 % for green fruits (N=60). Such fruit detection rates are a reasonable starting point for an obstacle map in an application regarding sweet-pepper harvesting. But, additional sensor information and detection is required to complement the obstacle map into a complete representation of the environment.

Acknowledgements

We thank Anja Dieleman for providing and maintaining the sweet-pepper plants in the experimental greenhouse in Wageningen. This research was funded by the European Commission in the 7th Framework Programme (CROPS GA no. 246252) and by the Dutch horticultural product board (PT no. 14555).

References

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern recognition, 30*(7), 1145-1159.

Bulanon, D. M., Burks, T. F., & Alchanatis, V. (2010). A multispectral imaging analysis for enhancing citrus fruit detection. *Environmental Control in Biology, 48*(2), 81-91.

Davies, E. R. (2009). The application of machine vision to food and agriculture: A review. *Imaging Science Journal, 57*(4), 197-217.

Gu, Q., Zhu, L., & Cai, Z. (2009). Evaluation measures of the classification performance of imbalanced data sets. *Communications in Computer and Information Science* (Vol. 51, pp. 461-471).

Humphries, S., & Simonton, W. (1993). Identification of Plant Parts Using Color and Geometric Image Data. *Transaction of the ASABE, 36*(5), 1493-1500.

Jiménez, A. R., Ceres, R., & Pons, J. L. (2000). A survey of computer vision methods for locating fruit on trees. *Transactions of the American Society of Agricultural Engineers, 43*(6), 1911-1920.

Kavdır, İ., & Guyer, D. E. (2004). Comparison of Artificial Neural Networks and Statistical Classifiers in Apple Sorting using Textural Features. *Biosystems Engineering, 89*(3), 331-344.

Kurtulmus, F., Lee, W. S., & Vardar, A. (2011). Green citrus detection using 'eigenfruit', color and circular Gabor texture features under natural outdoor conditions. *Computers and Electronics in Agriculture, 78*(2), 140-149.

Linker, R., Cohen, O., & Naor, A. (2012). Determination of the number of green apples in RGB images recorded in orchards. *Computers and Electronics in Agriculture, 81*, 45-57.

Mackinnon, A. (2000). A spreadsheet for the calculation of comprehensive statistics for the assessment of diagnostic tests and inter-rater agreement. *Computers in Biology and Medicine, 30*(3), 127-134.

MVTEC (2012). HALCON image processing library developed by Machine Vision Technologies (MVTEC). HALCON version: 11.0.0.1. www.mvtec.com.

Pudil, P., Novovičová, J., & Kittler, J. (1994). Floating search methods in feature selection. *Pattern Recognition Letters, 15*(11), 1119-1125.

Chapter 5 – Stem localization of sweet-pepper plants using the support wire as a visual cue

C.W. Bac[a,b], J. Hemming[a], E.J. van Henten[a,b]

Affiliations:
[a] Wageningen UR Greenhouse Horticulture, Wageningen University and Research Centre,
[b] Farm Technology Group, Wageningen University and Research Centre

This chapter was published in *Computers and Electronics in Agriculture* (2014), Vol. 105, p. 111-120

Abstract

A robot arm should avoid collisions with the plant stem when it approaches a candidate sweet-pepper for harvesting. This study therefore aims at stem localization, a topic so far only studied under controlled lighting conditions. Objectives were to develop an algorithm capable of stem localization, using detection of the support wire that is twisted around the stem; to quantitatively evaluate performance of wire detection and stem localization under varying lighting conditions; to determine depth accuracy of stereo-vision under lab and greenhouse conditions. A single colour camera was mounted on a pneumatic slide to record image pairs with a small baseline of 1 cm. Artificial lighting was developed to mitigate disturbances caused by natural lighting conditions. An algorithm consisting of five steps was developed and includes novel components such as adaptive thresholding, use of support wires as a visual cue, use of object-based and 3D features and use of minimum expected stem distance. Wire detection rates (true-positive/scaled false-positive) were more favourable under moderate irradiance (94/5%) than under strong irradiance (74/26%). Error of stem localization was measured, in the horizontal plane, by Euclidean distance. Error was smaller for interpolated segments (0.8 cm), where a support wire was detected, than for extrapolated segments (1.5 cm), where a support wire was not detected. Error increased under strong irradiance. Accuracy of the stereo-vision system (±0.4 cm) met the requirements (±1 cm) in the lab, but not in the greenhouse (±4.5 cm) due to plant movement during recording. The algorithm is probably capable to construct a useful collision map for robotic harvesting, if the issue of inaccurate stereo-vision can be resolved by directions proposed for future work. This is the first study regarding stem localization under varying lighting conditions, and can be useful for future applications in crops that grow along a support wire.

5.1 Introduction

This research is part of a project in which a robot is developed to harvest sweet-pepper in a greenhouse (Hemming et al. 2011). The manipulator and end-effector of this harvesting robot should avoid obstacles during motion towards a target (fruit or peduncle). The motion planner requires locations of these obstacles. In our prior research, plant stems were obstacles more difficult to detect than fruit, leaves or petioles (Bac et al. 2013a; Bac et al. 2013b). This work therefore focuses on stem localization.

A low-cost sensor, a Red, Green, Blue (RGB) camera, was selected to fit economic feasibility requirements for the harvesting robot (Pekkeriet 2011). Alternative sensors, such as LIDAR for detection of canopy structure in apple trees (Fleck et al. 2004) were considered to be too expensive. X-ray scanners were used for rose stem detection (Noordam et al. 2005), but are rather expensive and require the object to be placed between a source and receiver, which is a complicated configuration in a greenhouse environment. In our previous work, multi-spectral imaging was used (Bac et al. 2013b). But, we selected an RGB camera because the algorithm, described in this work, relies little on spectral features and uses mostly object-based features (size and shape). Such features can be extracted from RGB images as good as from multi-spectral images.

Stem detection and localization was studied under controlled lighting conditions (Paproki et al. 2012). Yet, we reviewed studies pertaining to our work that include experiments conducted under varying lighting conditions, and employ either multi-spectral imaging or colour imaging. Two studies describe classification of cucumber plant parts into leaves, stems and fruit: a study regarding a cucumber leaf picking robot (Van Henten et al. 2006) and a multi-spectral imaging study (Noble and Li 2012). Lu et al. (2011) detected branches of citrus using multi-spectral imaging. Stems of lychee were detected using colour imaging (Deng et al. 2011). Branches and trunk of apple trees were detected using colour imaging (Jidong et al. 2012). Two articles describe classification of grape foliage into several plants parts: a study using RGB (Dey et al. 2012), and a study using multi-spectral imaging (Fernández et al. 2013). Our previous work also dealt with detection of several plants parts (Bac et al. 2013b). Although work exists regarding stem detection or fruit localization (Bac et al. 2014), to the best of our knowledge, only one article exists in which stem localization was briefly described as part of a leaf picking robot (Van Henten et al. 2006).

To localize the stem, we used stereo-vision. Accuracy of stereo matching has been thoroughly investigated (Scharstein and Szeliski 2002), but depth accuracy of stereo-vision seems mostly qualitatively described for applications in a crop (Song et al. 2011; Van Der Heijden et al. 2012). To fill this gap, this study quantified depth accuracy and validated if the required accuracy (±1 cm) can be achieved, to localize obstacles for robotic harvesting (Hemming et al. 2011).

The approach included novel elements in terms of the baseline and algorithm. A small baseline of 1 cm was taken to improve matching score of stereo-vision and to decrease occlusion of the stem. Delon & Rougé (2007) note that few studies applied a small baseline so far and describe the advantages. Yet, a disadvantage is the difficulty to record images simultaneously. Regarding the algorithm, support wires were used as a visual cue to localize the plant stem because wires are twisted around the stem and can be distinguished from the vegetation. Support wires therefore approximate the location of the stem. Furthermore, the algorithm developed employed adaptive thresholding, object-based and 3D features, and filtering by minimum expected stem distance, to better handle varying lighting conditions.

Objectives were to (1) develop an algorithm capable of stem localization using detection of the support wire; (2) quantitatively evaluate performance of wire detection and stem localization under varying lighting conditions; (3) determine depth accuracy of stereo-vision under lab and greenhouse conditions.

This is the first study regarding stem localization and can be useful for future applications, to localize plant stems under varying lighting conditions. The algorithm and experimental set-up may not only be useful to localize obstacles for collision-free harvesting in sweet-pepper, but also in other crops that grow along a support wire, such as tomato, cucumber or egg-plant. The algorithm may furthermore fit for tasks other than harvesting, such as leaf picking, side shoot removal or plant phenotyping.

5.2 Image acquisition

Images of plants were recorded using an experimental set-up shown in Figure 5-1. Plants were of the red sweet-pepper cultivar ‘Waltz’ and were cultivated in the V-system (Jovicich et al.

2004). A total of 151 stems were recorded in 38 scenes. Solar irradiance was measured and ranged from 140 to 880 W/m^2 for these recordings.

Figure 5-1. Experimental set-up comprising a cart with a height-adjustable imaging set-up on top (left). Detailed view of the imaging set-up during recording in a row (right).

5.2.1 Camera and pneumatic slide

For stereo-vision, a camera was mounted on a pneumatic slide (Mini slide SLT; Festo AG & Co. KG, Germany). After recording the left image, the pneumatic slide was shifted to record the right image. Shifting took 0.4 s. The camera used was a 5 megapixel camera with a 2/3" CCD (Prossilica GC2450C; Allied Vision Technologies GmbH, Germany). A low-distortion lens with 5 mm focal length (LM5JC10M; Kowa GmbH, Germany) was mounted on the camera. A digital laser rangefinder was used (PLR 50; Bosch GmbH, Germany) to validate calculated depth values of the stereo-vision system.

5.2.2 Artificial lighting

For illumination of the scene, 30 halogen lamps (230VAC, 50 W) were used. Six rows of lamps illuminated the vertical range of the image (2448 pixels). Distance between the rows was 15 cm. Each row consisted of five lamps to cover the horizontal range of the image (2050 pixels). Distance between lamps in a row was 14 cm. Similar to previous research (Bac et al. 2013b), each row was horizontally shifted (7 cm) with respect to the previous row to improve equal light distribution.

Lamps were equipped with a dichroic reflector (Ø 51 mm), to reduce strong reflections in the centre of the image. Two types of reflectors were used: a reflector causing a beam angle of 25° (GU10/50/Clear Prolite; Ritelite Ltd., UK) and a reflector causing a beam angle of 50° (HI-Spot ES50; Sylvania Europe Ltd., UK). Lamps (N=18) emitting a beam angle of 25° were positioned at the edge of the lighting set-up, whereas lamps (N=12) emitting a beam angle of 50° were positioned in the centre of the lighting set-up. As a result, light was more diffuse in the centre of the image than at the edge.

5.3 Algorithm

The algorithm to localize stems (Figure 5-2) was developed in the image processing library HALCON© 11.0.1 (MVTec Software GmbH, Germany). Images were processed on a computer with an Intel Core i5 CPU 2.4 GHz Quad core processor with 4 GB of memory.

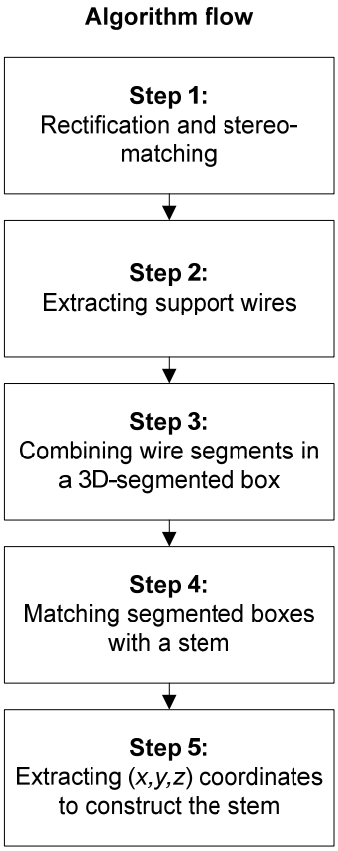


Figure 5-2. Algorithm flow to localize the stem of sweet-pepper plant using stereo-RGB images

The algorithm starts with rectification and stereo-matching in Step 1, to obtain 3D information from the scene. In Step 2, support wires, which are twisted around the stem, are detected. After Step 2, it is yet unclear to which stem a detected wire belongs and some leaf edges are falsely detected as support wires. Steps 3 and 4 solve these issues by matching wire segments with a real stem and by removing false detections. In Step 5, (*x,y,z*) coordinates are extracted from the support wires to construct a sequence of coordinates that represent the stem.

5.3.1 Step 1: Rectification and Stereo-matching

Images were rectified using intrinsic camera parameters. Intrinsic camera parameters were determined by a calibration procedure (Steger et al. 2007) in HALCON that uses a calibration plate (30x30 cm) ordered from MVTec Software GmbH (Germany). The procedure estimated the centre pixel of the image, focal length, width and height of a pixel, width and height of the image, and one parameter ($\kappa = 278$) to compensate for lens distortion. Another calibration procedure (Steger et al. 2007) in HALCON was used to determine extrinsic camera parameters, i.e. the relative pose between left and right image. As a result, images satisfied the

epipolar constraint (Brown et al. 2003), with a maximum measured error of 0.093 pixel. Baseline, i.e. linear displacement between left and right image, was 9.87 mm.

After rectification, we applied a local stereo-matching approach using the Normalized Cross-Correlation (NCC) as block-matching method (Hannah 1974; Brown et al. 2003). The matching method performed linear sub-pixel interpolation along the epipolar line. We used the green channel of the colour image to match vegetation pixels reflecting in this channel. A window of size 15x15 pixels was taken, which is about 1/3rd of the width of a stem, because smaller windows resulted in 'salt and pepper noise' and larger windows resulted in 'foreground fattening' (Hu and Mordohai 2012). Minimum and maximum disparities were set corresponding to a depth range of 0.3-0.8 m because plant stems always appeared within this range. To further reduce the disparity range searched, two levels of image pyramids were used (Scharstein and Szeliski 2002). As a result, computation time was saved. A left-right consistency check (Hu and Mordohai 2012) was performed to filter out erroneous matches. To match objects of interest (stem and support wire) and remove other objects, a threshold was set on correlation score (>0.95). This value did not cause a false removal in the images tested. Output of Step 1 was a disparity map.

5.3.2 Step 2: Extracting support wires

Support wires were extracted in six sub-steps: adaptive thresholding, intersection with stereo-matches, morphological image processing, thresholding on orientation, segmentation on roundness and removal of regions intersecting with the sky. We discuss each sub-step in the following paragraphs

In Step 2a, wires – assumed as bright objects in the scene – were detected by a threshold on the R, G and B channel. Thresholds were adapted to handle illumination changes in the images. Maximum threshold was fixed at a grey value of 255, whereas minimum threshold $Threshold_{min}$ was determined (Equation (3-6)) for each channel. Equation (3-6) ensured that thresholds were taken around the mean grey value of a channel and, in addition, shift from this mean was small for dark images and larger for bright images.

$$Threshold_{\min} = M_{match} + (255 - M_{match}) \cdot c \quad (-) \qquad (5\text{-}1)$$

Where: M_{match} (-) is the mean grey value within stereo-matched region for either R, G or B channel; c (-) is a constant to control threshold shift from M_{match}. Constant c was experimentally determined such that all wire segments were detected and few other objects were detected. c was 0.1 for the red channel, 0.1 for the green channel and 0.05 for the blue channel. These values were sufficient for images tested, but perturbations of 0.05 led to a worse detection performance.

In Step 2b, an intersection of four regions was taken to remove non-wire regions. These four regions comprised the three regions obtained in step 2a and the region where a stereo-match occurred (Step 1).

In Step 2c, morphological top hat filtering was applied on the output region of Step 2b, to extract rectangular shapes that represent a wire. The rectangular element applied (width = 15 pixels and height = 2 pixels) assured that wire segments remained. As a final step, small regions (<30 pixels) were removed because these never included wire segments.

In Step 2d, more undesired regions were removed through a threshold on orientation of the region in the image (Haralick and Shapiro 1992). Wires were twisted around the stem and therefore approximately vertically oriented. An orientation of -1.57 rad referred to an upright orientation and regions that were reasonably upright (orientation between -2.1 and -0.9 rad or between 0.9 to 2.1 rad) were kept, other regions were removed. These values were empirically determined.

In Step 2e, regions were filtered on 'roundness'. Roundness (Haralick and Shapiro 1992) is a feature in HALCON and its values range between 0 (non-round regions) and 1 (round regions). Support wires are non-round regions with roundness between 0 and 0.47, values were empirically determined. Other regions (>0.47) were removed.

In Step 2f, regions intersecting with the sky were removed. These regions mostly occurred at the edge between a leaf and the sky and were falsely detected as a wire. To obtain the sky as region, an empirically determined threshold of >200 was applied on the blue channel of the image.

5.3.3 Step 3: Combining wires in a 3D-segmented box

In this step, 3D features are used to combine wires in a box. Step 3 starts with intersecting detected wires (Step 2) with the disparity map (Step 1) to extract (*x,y,z*) coordinates by triangulation in Equations (5-2)-(5-4) (Rodriguez and Aggarwal 1990).

$$x = \frac{b \cdot x_L}{d} \quad \text{(m)} \qquad (5\text{-}2)$$

$$y = \frac{b \cdot y_L}{d} \quad \text{(m)} \qquad (5\text{-}3)$$

$$z = \frac{f \cdot b}{d} \quad \text{(m)} \qquad (5\text{-}4)$$

Where: x is the direction of the plant row; y is the height-direction; z is depth; b (m) is the baseline; x_L (m) and y_L (m) represent the coordinate system of the left image; f (m) is focal length; d (m) is disparity.

The function *'auto_threshold'* (Steger et al. 2007), in HALCON, was applied to determine thresholds for histograms of the *x*-direction and *z*-direction. This function solved the issue of strongly differing histograms among scenes, which caused fixed thresholds to fail. The function smoothed the histogram using a Gaussian filter with a manually adjustable standard deviation. Subsequently, thresholds were extracted at minima of the smooth histogram, to generate regions. A higher standard deviation results in stronger smoothing and fewer regions. Sometimes smoothing was too strong to obtain regions. Therefore we ran the function in a loop starting at a standard deviation of 0.04 m and decreased it until *'auto_threshold'* returned regions. As a result, Step 3 generated boxes for matching with a stem in Step 4.

5.3.4 Step 4: Matching segmented boxes with a stem

Step 4 uses expected stem distance as indicator to delete boxes probably not corresponding to a stem, i.e. to remove false detections (Figure 5-3). Expected object distance has been used earlier to detect sugar beet in a row (Bontsema et al. 1991).

```
% Inputs and Outputs
Boxes: array of regions that cover a box
x_coord: array of mean x-coordinates (m) for each box
Area: array of areas (pixels) for each box
% Parameters
min_displ: threshold for min expected stem displacement (m)
% Local variables
Indices: array of indices of array Boxes that should be kept
FP: boolean to indicate if there are any false-positive stem
detections
x_displ: displacement between two x-coordinates

FP:=true
while(FP)
   Count number of elements N in Boxes
   Generate Indices with values 0, 1,...N-1
   FP:=false
   for i:=0 to N-2 by 1
      x_displ:=x_coord[i]-x_coord[i-1]
      if (x_displ<min_displ)
         FP:=1
         if (Area[i]>Area[i-1])
            Delete element i-1 from Indices
         else
            Delete element i from Indices
         endif
         Create new array of Boxes with Indices of previous
         array of Boxes. Same for 'Area' and 'x_coord'.
         Break execution of for loop
      endif
   endfor
endwhile
```

Figure 5-3. Pseudo-code to match segmented boxes with a stem, using minimum expected stem distance.

Three properties were used for each box: the region that covers all pixels in the box, area (pixels) of this region, and mean *x*-coordinate (m) of all pixels in the box. Three corresponding arrays were constructed, where number of array elements corresponds to the number of boxes. Elements were sorted based on mean *x*-coordinate, i.e. from small to large values, and x-displacement was computed between subsequent elements. If x-displacement was smaller than expected minimum distance of the stem, the element with the smallest number of pixels was removed. For minimum expected distance of the stem, we took a value of 10 cm because stems were mostly 10-30 cm apart.

5.3.5 Step 5: Extracting (x,y,z) coordinates to construct the stem

In Step 5, the algorithm extracts (*x,y,z*) coordinates, from pixels in a box, and connects the coordinates to represent a stem. More precisely, four operations were implemented. Firstly, pixels in a box were connected, to obtain blobs, and (*x,y,z*) coordinates were determined. Secondly, for each blob, the pixel with minimum *y*-coordinate and the pixel with maximum *y*-coordinate were determined. Thirdly, these two pixels were connected by a line to represent an interpolated wire segment. Fourthly, interpolated wire segments were connected by a line to represent extrapolated wire segments. In addition, extrapolated wire segments were added at the smallest and largest *y*-coordinate of pixels in a box. The smallest *y*-coordinate was

connected to y=0 m. Similarly, the greatest y-coordinate was connected to y=5 m, which represents greenhouse height. Consequently, a localized stem consisted of interpolated wire segments and extrapolated wire segments.

5.4 Experiments

Four experiments were performed in this research. Firstly, performance of wire detection was visually and quantitatively assessed, to validate the algorithm (Section 5.4.1). Secondly, we determined the error between localized wire segments and a labelled stem, to validate accuracy of stem localization (Section 5.4.2). Thirdly, precision and accuracy of stereo-vision were determined, under lab conditions, to test if required accuracy (±1 cm) for obstacle localization (Hemming et al. 2011) can be achieved by stereo-vision (Section 5.4.3). Fourthly, accuracy was determined under greenhouse conditions as well, for comparison with lab conditions (Section 5.4.4).

5.4.1 Wire detection

Wire detection was qualitatively assessed by visualizing the output, for each of the five steps in the algorithm. Furthermore, wire detection was quantitatively assessed using the four elements of a confusion matrix: true-positive (TP), false-positive (FP), false-negative (FN) and true-negative (TN). A wire was counted as detected (TP) if at least one of its interpolated segments was actually on a real stem. If only an extrapolated wire segment, or no segment at all, intersected with the real stem, it was counted as a false detection (FP). If the algorithm did not detect any wire segment on the stem, it was counted as a missed detection (FN). Finally, a correctly removed detection (TN) was counted if a box was successfully removed in Step 4 of the algorithm (Section 5.3.4).

True-Positive Rate (TPR), (TP/(TP+FN), and Scaled False-Positive Rate (SFPR), (FP/(TP+FN), were calculated to demonstrate performance of wire detection on 151 stems that were visible in 38 scenes. These scenes were split into two subsets based on measured outdoor solar irradiance, to test the effect of irradiance on detection rate. A subset of 28 scenes, recorded under moderate irradiance of 140-300 W/m^2, was compared with a subset of 10 scenes, recorded under strong irradiance of 300-880 W/m^2. Average grey-value of stem pixels, recorded in the green channel, ranged between 14-26 for scenes recorded under moderate irradiance and between 27-49 for scenes recorded under strong irradiance.

5.4.2 Error of stem localization

A stem location was approximated by connecting (x,y,z) coordinates extracted from wire segments (Section 5.3.5). The error of this approximation was determined by comparing localized wire segments with labelled pixels on the actual stem. We labelled visible parts of the stem by drawing a line through the vertical axis of the stem (Figure 5-4). Subsequently, (x,y,z) coordinates of localized wire segments were compared with (x,y,z) coordinates of the labelled stem, obtained by intersecting the labelled stem with the disparity map and applying Equations (5-2)-(5-4). Thus, this experiment involved a stereo-to-stereo comparison.

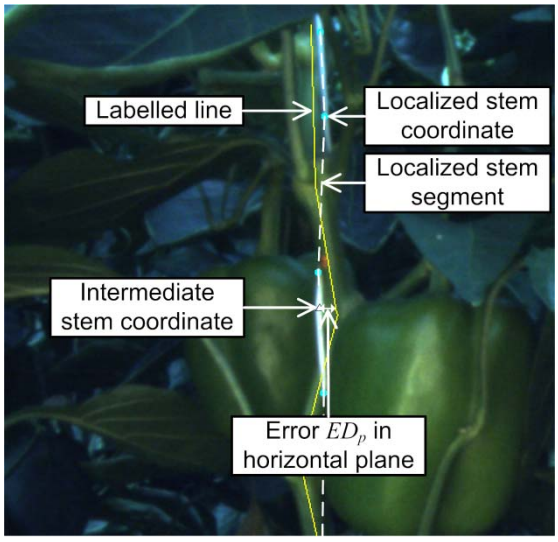


Figure 5-4. The labelled line (stem) was used to calculate error ED_p relative to a localized wire segment. Only visible parts of the stem were labelled.

To compare more coordinates than the two coordinates extracted from a localized wire segment (Section 5.3.5), intermediate (*x,y,z*) coordinates (Figure 4) were generated through interpolation. We interpolated these coordinates such that y-coordinates were aligned with *y*-coordinates of the labelled stem. As a result, we were able to calculate errors for a large number of coordinates per wire segment. To express error in the horizontal plane, errors in *x*- and *z*-direction were used to calculate Euclidean distance (Equation (5-5)), for each labelled pixel (Figure 4).

$$ED_p = \sqrt{(Lab_x - Loc_x)^2 + (Lab_z - Loc_z)^2} \quad \text{(cm)} \qquad (5\text{-}5)$$

Where: ED_p (cm) Euclidean distance between a labelled and localized pixel
Lab_x (cm) *x*-coordinate of a labelled pixel
Loc_x (cm) *x*-coordinate of a localized pixel
Lab_z (cm) *z*-coordinate of a labelled pixel
Loc_z (cm) *z*-coordinate of a localized pixel

Finally, Euclidean distance of a wire segment, to the stem, was calculated by averaging ED_p of all pixels in the wire segment.

We calculated mean and sample Standard Deviation (SD) for wire segments grouped by two variables. The first variable, type of wire segment, validated the hypothesis that Euclidean distance is smaller for interpolated segments than for extrapolated segments. We expected larger errors for extrapolated segments because of a greater distance to the stem. The second variable, irradiance, validated if Euclidean distance was greater for wires recorded under strong irradiance (300-880 W/m^2) than for wires recorded under moderate irradiance (140-300 W/m^2). For strong irradiance, 30 wires were analysed that were composed of 446 segments: 208 interpolated and 238 extrapolated. Due to complete occlusion of some of these segments, we calculated Euclidean distance for only 101 interpolated and only 130 extrapolated wire segments. For moderate irradiance, 41 wires were analysed that were composed of 415 segments: 187 interpolated and 228 extrapolated. Due to complete occlusion

of some of these segments, we calculated Euclidean distance for only 126 interpolated and only 138 extrapolated wire segments.

5.4.3 Accuracy of stereo-vision under lab conditions

Precision was tested to evaluate consistency of stereo-matching. Subsequently, accuracy was tested to evaluate both consistency of stereo-matching and accuracy of calibration (Section 5.3.1).

Precision was tested by recording 100 image pairs of a constant scene (Figure 5-5), at a distance of 54 cm. To compare precision for two different objects, a black dot of a calibration plate and a support wire segment of a sweet-pepper plant were detected, using a threshold in a manually drawn region-of-interest around these objects. Subsequently, a mean (x,y,z) coordinate was calculated over matched pixels in the object detected, using matching settings and triangulation equations of the algorithm (Section 5.3). Finally, we calculated precision, separately for x, y and z-direction, by taking population standard deviation (σ) over these 100 mean coordinates.

Whereas precision was calculated in three directions, accuracy was tested in only the z-direction (Equation (5-6)) because literature indicated most inaccuracy can be expected in the z-direction (Van Henten et al. 2002) and this statement was supported by our results of precision.

$$Accuracy = \frac{\sum_{i=0}^{N} | z(Stereo)_i - z(GT)_i |}{N} \quad \text{(m)} \qquad (5\text{-}6)$$

Where: *z(Stereo)* (m) depth measurement *i* of the stereo-vision system
z(GT) (m) depth measurement *i* of the ground-truth measurement device

We recorded the same scene as for precision (Figure 5-5), but camera-dot distance was varied in a range of 30-120 cm with steps of 1.2 cm. Hence, 76 image pairs were recorded.

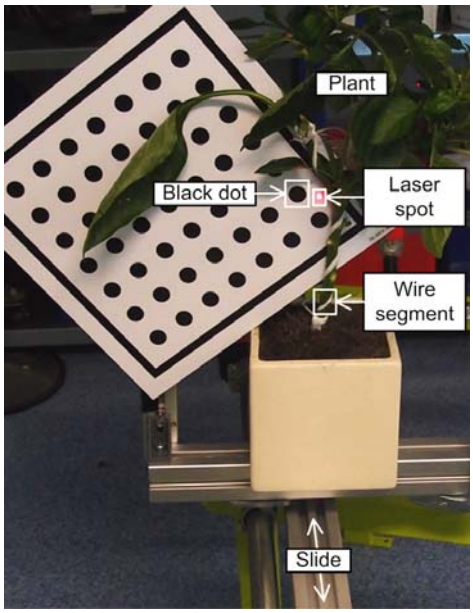


Figure 5-5. Set-up used for accuracy validation of stereo-vision. Distance between camera and black dot was varied using the slide. The laser measurement served as ground truth of distance. The stereo-vision system determined distance of both the black dot and the wire segment indicated.

To measure ground-truth in depth direction, a laser rangefinder was pointed at the calibration plate. Specs of the laser rangefinder report an accuracy of ±1 mm for depths less than 1 m. In addition, inaccuracy of ground truth slightly aggravated with distance due to misalignment. To align camera and laser rangefinder, we rotated the laser rangefinder until (*x,y*) coordinate of the localised black dot was constant for the depth range investigated. Similarly, we aligned slide and laser rangefinder by translating the slide until the laser spot appeared at the same location at the calibration plate for the depth range investigated. Such alignment is not perfectly accurate and we therefore assume a maximum offset of 3 cm in (*x,y*) direction at a distance of 1 m, which corresponds to an error of 0.45 mm in depth. Summing the errors (misalignment and laser range finder) led to an accuracy of ±1.1 mm at a distance of 0.3 m and ±1.5 mm at a distance of 1.2 m. Hence, we considered laser measurement a proper ground-truth because its accuracy (±1.5 mm) is greater than theoretical depth resolution of stereo-vision (10.1 mm) calculated by Equation (5-7) (Rodriguez and Aggarwal 1990).

$$\Delta z = \frac{-z^2 \cdot p \cdot s}{b \cdot f + z \cdot p \cdot s} \quad \text{(m)} \qquad (5\text{-}7)$$

Δz (m) is the depth resolution; p pixel size was 3.45·10-6 m; s precision of sub-pixel interpolation was assumed 1/10; b (9.87·10-3 m) is the baseline; f (0.005 m) is the focal length; z is the depth taken at 1.2 m.

To determine the optical camera centre, we aligned the front side of the laser rangefinder and the front side of the lens. Subsequently, mean offset between depth measurements (N=76) of laser rangefinder and stereo-vision were calculated for the black dot. This offset represented the distance between the optical camera centre and the front of the lens. In addition, we added 8 mm to depth measurements of the wire to compensate for the measured depth offset between wire and calibration plate (Figure 5-5).

Precision of sub-pixel interpolation s was determined for the average and maximum error measured, using Equation (5-8) (Brown et al. 2003). To be able to assess s, we first checked if mismatches of one or more pixels occurred.

$$s = \frac{b \cdot f \cdot E_z}{p \cdot z^2 - p \cdot z \cdot E_z} \quad \text{(-)} \qquad (5\text{-}8)$$

E_z (m) is the error measured in the depth direction.

5.4.4 Accuracy of stereo-vision under greenhouse conditions

For depth validation in the greenhouse, the laser beam was pointed at a stem and an image pair was recorded. Subsequently, the depth measurement of the laser rangefinder was read from the display and reported. There was a time lag of about 20 s between recording of the image pair and reading the depth measurement. In this time period, stems sometimes moved due to a breeze in the greenhouse and therefore we expect the error of the ground-truth to be about ±1 cm. In addition, we quantified the effect of stem movement on accuracy of stereo-vision.

To compare distance of stereo-vision with laser distance, the pixel containing the laser spot was manually selected. The laser spot did not disturb matching because the red spot was not visible in the green channel used for stereo-matching. In case no disparity was found on

the laser spot, the nearest matched pixel was taken. In total, 91 stems were measured: 19 spots appeared at a wire, 72 spots appeared at a stem. In lab tests we did not find any difference in distance measurements between spots appearing at the wire and spots appearing at a stem.

5.5 Results

The following sub-sections correspond to experiments described in sub-sections of Section 5.4.

5.5.1 Wire detection

Output for each of the five steps in the algorithm (Section 5.3) is visualized in Figure 5-6. Regarding performance of wire detection for scenes recorded under moderate irradiance (113 stems), true-positive rate was 94% and scaled false-positive rate was 5%. For scenes recorded under strong irradiance (38 stems), true-positive rate was 74% and scaled false-positive rate was 26%. Hence, strong irradiance deteriorated performance. In addition, under strong irradiance, constructed stems sometimes contained a combination of correct detections of wire segments and of incorrect detections of leaf edges and fruit detected as wire segment. As a result, constructed stems contained bends that increased error of stem localization (Section 5.5.2). Yet, this effect was hardly observed in images recorded under moderate irradiance.

5.5.2 Error of stem localization

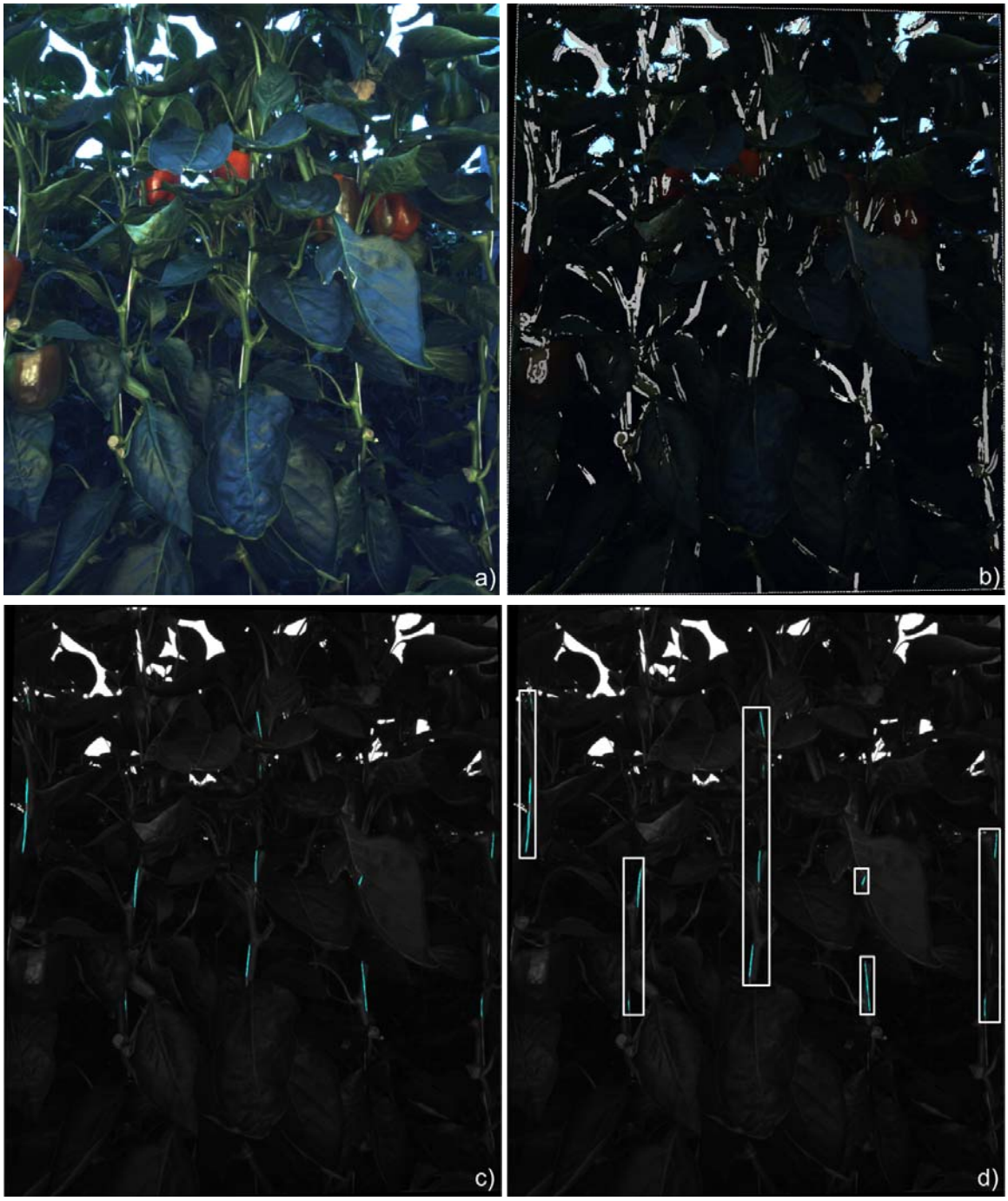
Euclidean distance in the horizontal plane was smaller for interpolated wire segments (0.8 cm; 2.9 cm) than for extrapolated wire segments (1.5 cm; 3.9 cm), both under moderate and strong irradiance (Table 5-1). Furthermore, for interpolated wire segments, Euclidean distance was almost four times greater under strong irradiance (2.9 cm) than under moderate irradiance (0.8 cm). Also, for extrapolated wire segments, Euclidean distance was two times greater under strong irradiance (3.9 cm) than under moderate irradiance (1.5 cm).

5.5.3 Accuracy of stereo-vision under lab conditions

Precision of stereo-vision for the black dot (N=100) detected was 1.2 mm in *x*-direction, 0.1 mm in *y*-direction and 2.7 mm in *z*-direction. For the detected wire segment (N=100), precision was similar: 1.0 mm in *x*-direction, 0.5 mm in *y*-direction and 2.5 mm in *z*-direction. Since precision in *x*-direction is relatively high (<1.2 mm), we conclude that shifts of the pneumatic slide were accurate and did not cause an inconsistent baseline among scenes.

Error of stereo-vision for a distance of 0.3-1.2 m is in Figure 5-7. The optical camera centre was 9.1 mm behind the front surface of the lens. For the range in which stems occur (0.3-0.8 m), depth measurements of the black dot (±0.2 cm) were slightly more accurate than the wire segment (±0.4 cm). For a longer range (>0.8 m), accuracy decreased for both black dot (±0.3 cm) and wire segment (±0.6 cm). Consequently, under lab conditions, accuracy of wire localization (±0.4 cm) fits the requirement (±1 cm) (Hemming et al. 2011).

A mismatch did probably not occur because a mismatch of one pixel would cause an error of 0.6 cm at a depth of 30 cm and an error of 9.5 cm at a depth of 120 cm. Such large errors were not observed (Figure 5-7). Therefore we were able to assess precision of sub-pixel interpolation: 1/11 for an average error (±0.4 cm) and 1/5 for the maximum error observed (1.1 cm).


a)
b)
c)
d)

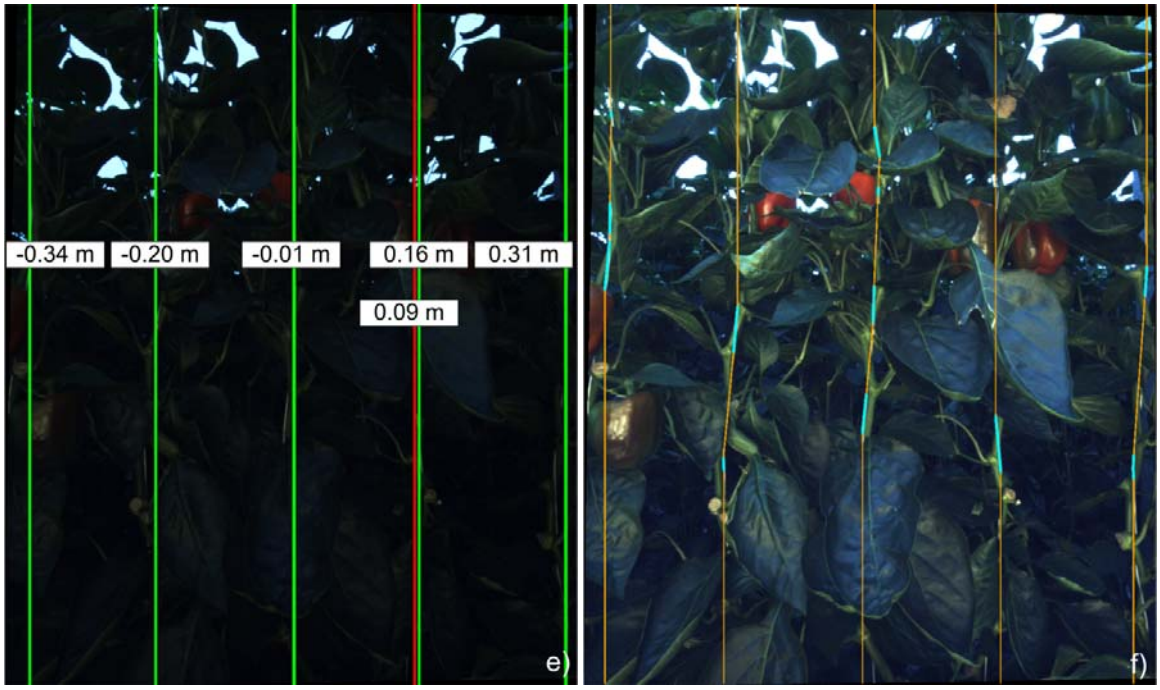


Figure 5-6. Results of stem localization algorithm that consists of Steps 1-5. a) left frame of recorded RGB frame (histogram was scaled for better appearance). b) disparity image, overlayed on the rectified left frame (Step 1). c) Extracted support wires, indicated in cyan (Step 2). d) Wires combined in six 3D-segmented boxes (Step 3), where regions belonging to a box are indicated by a white rectangle. e) Mean x-value (m) of pixels in each box after matching (vertical bars are for visualization of x-value). The red bar (x=0.09 m) indicates a correctly rejected box. The other five green bars are correct matches of a box with a stem (Step 4). f) Constructed stems that consist of interpolated segments, indicated in cyan, and extrapolated segments, indicated in orange (Step 5; histogram was scaled for better appearance).

Table 5-1. Mean (SD) Euclidean distance (cm), in the horizontal plane, of wire segments to the actual stem. Data is shown for scenes recorded under moderate or strong irradiance

Moderate irradiance (140-300 W/m^2)		Strong irradiance (300-880 W/m^2)	
Interpolated wire segments (N=126)	Extrapolated wire segments (N=138)	Interpolated wire segments (N=101)	Extrapolated wire segments (N=130)
0.8 (1.2)	1.5 (1.6)	2.9 (4.6)	3.9 (4.3)

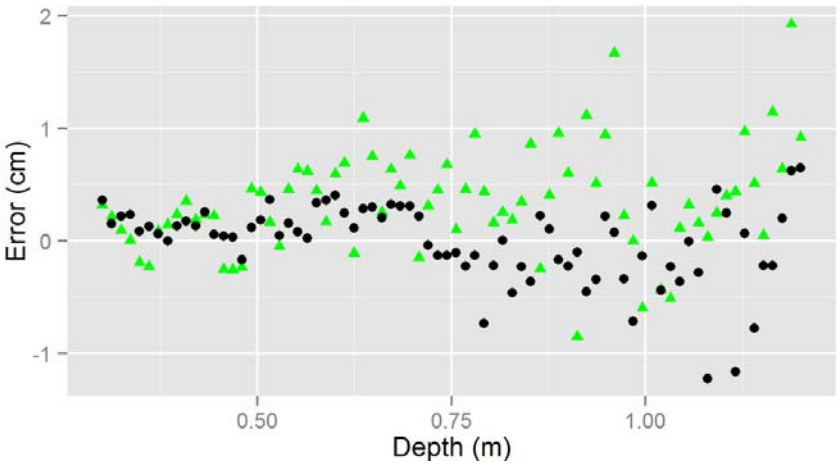


Figure 5-7. Error of depth for the stereo-vision system, measured under lab conditions. Objects localized were a black dot of a calibration plate (●) and a wire segment (▲). The error increases with depth.

5.5.4 Accuracy of stereo-vision under greenhouse conditions

Accuracy of stereo-vision was ±4.5 cm and was based on 91 measurements. This accuracy deviates from the accuracy (±1 cm) required for harvesting (Hemming et al. 2011). Analysis of a scene with a large error (-8.7 cm) shows that the pixel was correctly matched (Figure 5-8).

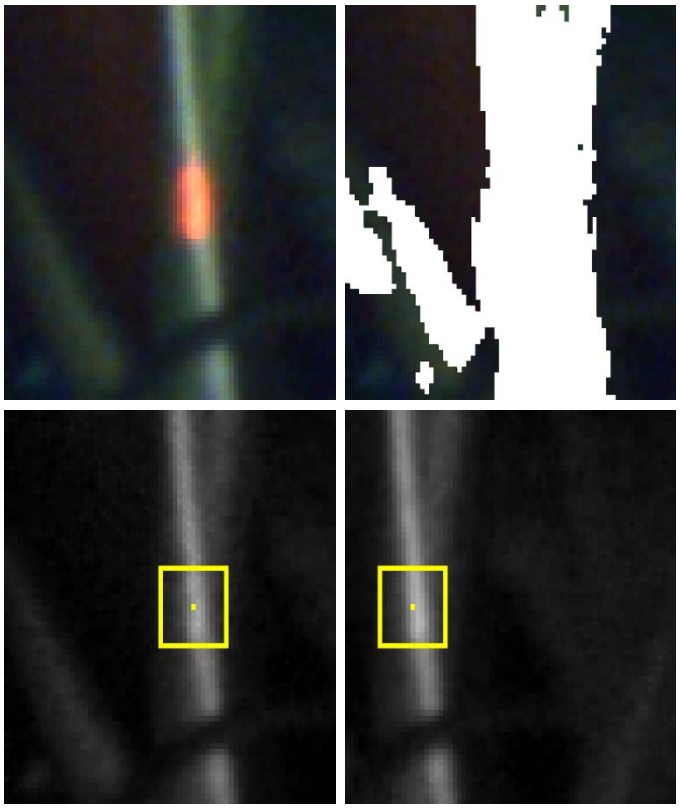


Figure 5-8. An example of a stereo-match on the laser spot. RGB image with the laser spot visible on a wire (top-left). The disparity map (top-right) obtained after stereo-matching of the left frame (bottom-left) and right frame (bottom-right). The mask (15x15 pixels) is displayed in the left and right frame. Also note that the laser spot was not visible in the green channel used for stereo-matching.

Whereas we did not observe a mismatch of pixels (Figure 5-8), movement of the plant or cart during recording might explain the large error (- 8.7 cm). Probably about one cm of this error can be explained by inaccuracy of stereo-vision or ground-truth. The remaining error corresponds to a disparity offset of three pixels at a laser-stem distance of 66.5 cm. If a plant would have moved horizontally by three pixels, speed of plant movement would have been 3.0 mm/s. We calculated this speed by dividing real-world width of three pixels (1.2 mm), at a depth of 66.5 cm, by the duration of the slide movement (0.4 s). Whereas we did not measure speed of plant movement, a speed of 3.0 mm/s seems likely to have occurred because of the breeze we observed in the greenhouse.

5.6 Discussion

We discuss each experiment in the following paragraphs.

Wire detection yielded a True-Positive Rate (TPR) of 94% and a Scaled False-Positive Rate (SFPR) of 5% under moderate irradiance. To avoid the performance drop under strong irradiance (TPR of 74%; SFPR of 26%), exposure time of the camera can be adapted and algorithm parameters can be optimized in future work. We were unable to compare performance with related work regarding stem detection (Bac et al. 2013b; Fernández et al. 2013) because these studies report performance on a pixel basis, whereas this study reports performance on a blob basis. Therefore we compared performance with the state-of-the-art in green fruit detection. For green apple detection, Linker et al. (2012) achieved similar performance: a TPR of 95% and a SFPR of 4% under diffuse natural lighting and a TPR of 88% and a SFPR of 25% under intense natural lighting.

Errors of stem localization under strong irradiance (2.9 cm) were mainly due to leaf edges and fruit falsely being detected as wire. Adapting exposure time may decrease these errors because, in darker images, leaf edges and fruit appeared darker than wires and were therefore not detected as wire. To compensate for errors, an additional danger zone can be taken into account during planning of collision-free motions. Planning such motions may improve harvest success and decrease damages to the plant compared with current harvesting robots that mostly do not consider obstacles (Bac et al. 2014). However, adding danger zones may reduce the chance to find a collision-free motion and might increase calculation time needed for motion planning.

Error of stereo-vision (±0.4 cm) under lab conditions can be explained by inaccurate calibration, inaccuracy of ground-truth (±1.1-1.6 mm) and varying level of sub-pixel interpolation (±2.6 mm) due to illumination change (Brown et al. 2003). Sub-pixel interpolation achieved was about 1/11 and this value fits with Brown et al. (2003) who indicate it is difficult to obtain sub-pixel interpolation better than 1/10.

Comparing accuracy of stereo-vision under greenhouse conditions (±4.6 cm), with the literature, was hard because little work exists. In a related study sweet-pepper fruit were localized under greenhouse conditions using stereo-vision and time-of-flight images (Song et al. 2011; Van Der Heijden et al. 2012). Although these authors do not report accuracy, they indicate pixel values were not reliable for matching due to changes of perspective, lighting and noise. Such changes may partly explain why accuracy was much worse under greenhouse conditions (±4.6 cm) than under lab conditions (±0.4 cm). Yet, the main reason of poor accuracy is probably movement of the plant or cart during recording. Due to the poor accuracy achieved, we were unable to separately assess the effect of perspective, lighting and noise because of the strong effect of plant movement on error of depth measurements. For the same reason we were unable to assess precision of sub-pixel interpolation. Therefore simultaneous acquisition of the left and right image is critical for accurate localization in future work, as also indicated by others (Biskup et al. 2007). Simultaneous acquisition is only possible using a wider baseline than the one chosen in this research (1 cm) because the width of the current camera restricts minimum baseline to 3.5 cm. A drawback of such wider baseline is less similar images (more occlusion, more noise, larger geometrical deformations) and causes less accurate disparity measurements due to a more difficult matching process (Delon and Rougé 2007). In addition, performance of the wire detection algorithm might drop because support wires can become more occluded by leaves. Therefore we suggest to use two

baselines: a small baseline of 1 cm to detect stems and a baseline of 3.5 cm to accurately localize as many of the detected pixels as possible. Alternatively a solution with a mirror and a beam splitter can be used (Pachidis and Lygouras 2005), or a combination of RGB and time-of-flight images.

5.7 Conclusion

The algorithm developed is capable of stem localization using the support wire as a visual cue and using stereo-images with a small baseline (1 cm). Novel components of the algorithm include adaptive thresholding, use of support wires as a visual cue, use of object-based and 3D features and use of minimum expected stem distance.

Wires were detected with a true-positive rate (TPR) of 94% and a scaled false-positive rate (SFPR) of 5%, under moderate irradiance. Although related work that includes detection performance does not exist, this performance is comparable with state-of-the-art performance of green fruit detection. The algorithm, however, suffers from strong irradiance because detection rate dropped to a TPR 74% and an SFPR of 26%. Also, error of stem localization suffers from strong irradiance because error was two times greater for interpolated wire segments and four times greater for extrapolated wire segments. Adapting exposure time for irradiance is therefore a task for future work.

The algorithm is probably capable to construct a useful collision map for robotic harvesting, if the issue of inaccurate localization can be resolved in future work. Accuracy of stereo-vision meets the required accuracy (±1 cm) under lab conditions (±0.4 cm), but not under greenhouse conditions (±4.5 cm). Plant movement seems the major cause of this decreasing accuracy and simultaneous acquisition of the left and right frame is needed to avoid errors caused by plant movement. A possible direction for future work is to investigate double-baseline stereo, a combination with time-of-flight images, or a solution with mirrors and a beam splitter.

Acknowledgements

We thank Yael Edan and Bart van Tuijl for their contributions to the experimental set-up and experiments. We are grateful to growers Ted and Robert Vollebregt and Cees and Rolf Vijverberg who allowed us to perform experiments in their greenhouses. This research was funded by the European Commission in the 7th Framework Programme (CROPS GA no. 246252) and by the Dutch horticultural product board (PT no. 14555).

References

Bac, C. W., Hemming, J., & Van Henten, E. J. (2013a). *Pixel classification and post-processing of plant parts using multi-spectral images of sweet-pepper.* Paper presented at the IFAC Biorobotics Conference, Sakai, Japan,

Bac, C. W., Hemming, J., & Van Henten, E. J. (2013b). Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper. *Computers and Electronics in Agriculture, 96*, 148-162.

Bac, C. W., Van Henten, E. J., Hemming, J., & Edan, Y. (2014). Harvesting robots for high-value crops: state-of-the-art review and challenges ahead. *Journal of Field Robotics, DOI: 10.1002/rob.21525*.

Biskup, B., Scharr, H., Schurr, U., & Rascher, U. (2007). A stereo imaging system for measuring structural parameters of plant canopies. *Plant, Cell and Environment, 30*(10), 1299-1308.

Bontsema, J., Grift, T. E., & Pleijsier, K. Mechanical weed control in sugar beet growing: the detection of a plant in a row. In Y. Hashimoto, & W. Day (Eds.), *Mathematical and control applications in agriculture and horticulture, Matsuyama, Japan, 30 Sep-3 Oct 1991* (pp. 207-212).

Brown, M. Z., Burschka, D., & Hager, G. D. (2003). Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 25*(8), 993-1008.

Delon, J., & Rougé, B. (2007). Small baseline stereovision. *Journal of Mathematical Imaging and Vision, 28*(3), 209-223.

Deng, J., Li, J., & Zou, X. (Vision). Extraction of litchi stem based on computer vision under natural scene. In *Proceedings - International Conference on Computer Distributed Control and Intelligent Environmental Monitoring, CDCIEM 2011, 2011* (pp. 832-835).

Dey, D., Mummert, L., & Sukthankar, R. (Vision). Classification of plant structures from uncalibrated image sequences. In *Proceedings of IEEE Workshop on Applications of Computer Vision, 2012* (pp. 329-336).

Fernández, R., Montes, H., Salinas, C., Sarria, J., & Armada, M. (2013). Combination of RGB and multispectral imagery for discrimination of Cabernet Sauvignon grapevine elements. *Sensors (Switzerland), 13*(6), 7838-7859.

Fleck, S., Van Der Zande, D., Schmidt, M., & Coppin, P. (2004). Reconstructions of tree structure from laser-scans and their use to predict physiological properties and processes in canopies. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36*(Part 8/W2), 119-123.

Hannah, M. J. (1974). *Computer Matching of Areas in Stereo Images*. PhD Thesis, Stanford University, CA, USA.

Haralick, R. M., & Shapiro, L. G. (1992). *Computer and Robot Vision* (Vol. 1). Boston, USA: Addison-Wesley.

Hemming, J., Bac, C. W., & Van Tuijl, B. A. J. (2011). CROPS project deliverable 5.1: Report with design objectives and requirements for sweet-pepper harvesting. Wageningen, The Netherlands: Wageningen UR Greenhouse Horticulture.

Hu, X., & Mordohai, P. (2012). A quantitative evaluation of confidence measures for stereo vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 34*(11), 2121-2133.

Jidong, L., De-An, Z., Wei, J., Yu, C., & Ying, Z. (Vision). Research on trunk and branch recognition method of apple harvesting robot. In *Measurement, Information and Control (MIC), 2012 International Conference on, 18-20 May 2012 2012* (Vol. 1, pp. 474-478).

Jovicich, E., Cnatliffe, D. J., Sargent, S. A., & Osborne, L. S. (2004). Production of Greenhouse-Grown Peppers in Florida. Gainesville, FL: University of Florida, IFAS Extension.

Linker, R., Cohen, O., & Naor, A. (2012). Determination of the number of green apples in RGB images recorded in orchards. *Computers and Electronics in Agriculture, 81*, 45-57.

Lu, Q., Tang, M., & Cai, J. (Vision). Obstacle recognition using multi-spectral imaging for citrus picking robot. In *Proceedings - PACCS 2011: 2011 3rd Pacific-Asia*

Conference on Circuits, Communications and System, Wuhan, China, 17-18 July 2011 2011 (pp. 1-5).

Noble, S., & Li, D. (2012). *Segmentation of Greenhouse Cucumber Plants in Multi-spectral Imagery*. Paper presented at the International Conference of Agricultural Engineering, CIGR-Ageng, Valencia, Spain, 8-12 July

Noordam, J. C., Hemming, J., van Heerde, C., Golbach, F., van Soest, R., & Wekking, E. (2005). Automated Rose Cutting in Greenhouses with 3D Vision and Robotics: Analysis of 3D Vision Techniques for Stem Detection. *Acta Hort. (ISHS), 691*, 885-892.

Pachidis, T. P., & Lygouras, J. N. (2005). Pseudo-stereo vision system: A detailed study. *Journal of Intelligent and Robotic Systems: Theory and Applications, 42*(2), 135-167.

Paproki, A., Sirault, X., Berry, S., Furbank, R., & Fripp, J. (2012). A novel mesh processing based technique for 3D plant analysis. *BMC Plant Biology, 12*(1), 63.

Pekkeriet, E. J. (2011). CROPS project deliverable 12.1: Economic viability for each application. Wageningen, The Netherlands: Wageningen UR Greenhouse Horticulture.

Rodriguez, J. J., & Aggarwal, J. K. (1990). Stochastic analysis of stereo quantization error. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 12*(5), 467-470.

Scharstein, D., & Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision, 47*(1-3), 7-42.

Song, Y., Glasbey, C. A., Van Der Heijden, G. W. A. M., Polder, G., & Dieleman, J. A. (2011). Combining stereo and time-of-flight images with application to automatic plant phenotyping. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6688 LNCS, pp. 467-478).

Steger, C., Ulrich, M., & Wiedemann, C. (2007). Ch. 3: Machine Vision Algorithms. In *Machine Vision Algorithms and Applications*. Weinheim, Germany: Wiley-VCH.

Van Der Heijden, G., Song, Y., Horgan, G., Polder, G., Dieleman, A., Bink, M., et al. (2012). SPICY: Towards automated phenotyping of large pepper plants in the greenhouse. *Functional Plant Biology, 39*(11), 870-877.

Van Henten, E. J., Hemming, J., Van Tuijl, B. A. J., Kornet, J. G., Meuleman, J., Bontsema, J., et al. (2002). An Autonomous Robot for Harvesting Cucumbers in Greenhouses. *Autonomous Robots, 13*(3), 241-258.

Van Henten, E. J., Van Tuijl, B. A. J., Hoogakker, G. J., Van Der Weerd, M. J., Hemming, J., Kornet, J. G., et al. (2006). An Autonomous Robot for De-leafing Cucumber Plants grown in a High-wire Cultivation System. *Biosystems Engineering, 94*(3), 317-323.

Chapter 6 – Analysis of a motion planning problem for fruit harvesting in a dense obstacle environment

C.W. Bac[a,b], T. Roorda[b], R. Reshef[c], S. Berman[c], J. Hemming[a], E.J. van Henten[b]

Affiliations:
[a] Wageningen UR Greenhouse Horticulture, Wageningen University and Research Centre,
[b] Farm Technology Group, Wageningen University and Research Centre
[c] Department of Industrial Engineering and Management, Ben-Gurion University of the Negev

This chapter was submitted for publication in *Biosystems Engineering*

Abstract

Robotic fruit harvesting requires a collision-free motion, of the manipulator and end-effector, to a target fruit in an obstacle-dense crop environment. We conducted a novel twofold analysis of a case study of an actual sweet-pepper harvesting robot based on data of fruit (N=158) and stem locations collected from a greenhouse. The first analysis compared two methods of selecting the azimuth angle of the end-effector. Our new 'constrained-azimuth' method avoided risky paths and achieved a motion planning success similar to the 'full-azimuth' method. In the second analysis, we conducted a sensitivity analysis for five parameters specifying the crop (stem spacing and fruit location), the robot (end-effector dimensions and robot position) and the planning algorithm, to evaluate their effect on successfully finding a collision-free goal configuration and path. Reducing end-effector dimensions and widening stem spacing are promising research directions because they both significantly improved goal configuration success, from 63% to 84%. Yet, the fruit location at the stem is the strongest influencing parameter and therefore provides an incentive to train or breed plants that develop more fruit at the front side of the plant stem. The two analyses may serve as useful tools to study motion planning problems in a dense obstacle environment.

6.1 Introduction

Robotic fruit harvesting is motivated by several factors, including a need to reduce the production costs (Lewis et al. 1983) of high-value crops such as tomato, cucumber or sweet-pepper. Harvesting these crops is challenging given the complex working environment for the robot that comprises varying poses, sizes, shapes and colours of fruit and other objects (Bac et al. 2014b). Furthermore, a target fruit can be surrounded by densely spaced obstacles, i.e. plant parts, support wires, and fruit clusters. The manipulator and end-effector need to avoid these obstacles to successfully approach a target fruit and prevent damages to the plant or nearby fruit. To address this motion planning problem of an obstacle-dense crop environment, we conduct a novel twofold analysis of a case study of sweet-pepper harvesting. To provide a realistic case, we collected measurements from a crop grown in a greenhouse and performed simulations with the actual harvesting robot developed for the crop.

The first analysis focuses on properly selecting the azimuth angle of the end-effector with respect to the target fruit. Preferably, the azimuth angle should be selected such that the end-effector is positioned in front of the fruit, with the stem located behind the fruit (Figure 6-1), because preliminary field tests of the end-effector mechanism revealed that such a pose increases the probability of successful fruit detachment while reducing the probability of damaging the fruit or stem. However, accessing this preferred pose is sometimes problematic for fruit located at the left, right, or backside of the stem, for two reasons. Firstly, in the few cases a path exists to the preferred pose, there is a high probability of damaging a plant part while executing the path. Secondly, planning a path to the preferred pose may result in a long planning time, even for the efficient bi-directional Rapidly exploring Random Trees (bi-RRT) planner (Lavalle 2006) deployed in this research. To avoid these issues and to quickly find a successful path, we select the azimuth angle to approach the fruit from a constrained range of 120° (taken orthogonal to the direction of the aisle), i.e. the 'constrained-azimuth' method,

conversely to selecting the azimuth angle from the full range of 360°, i.e. the 'full-azimuth' method. We compare these methods and demonstrate how this new 'constrained-azimuth' method avoids needless path planning without reducing motion planning success.

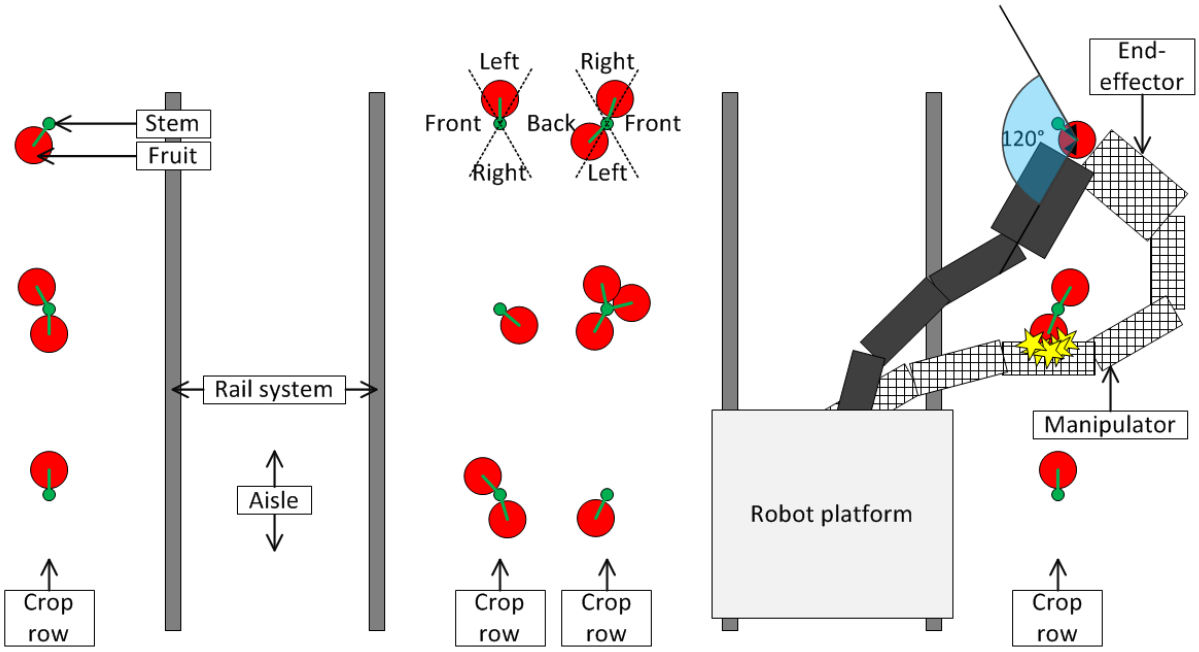


Figure 6-1. Top view of four crop rows, along two aisles, in the greenhouse. Two end-effector poses at a backside fruit are shown. In the preferred pose (⊟), the end-effector is positioned in front of a fruit and the stem behind the fruit. However, this pose leads to a collision (indicated by stars) between the manipulator and a nearby fruit. Therefore, a collision-free pose (■) is selected from a constrained range of the azimuth angle (120°) taken orthogonal to the direction of the aisle, instead of from the full range of 360°.

In the second analysis, we use the 'constrained-azimuth' method and conduct a sensitivity analysis for five parameters specifying the crop (stem spacing and fruit location), the robot (end-effector dimensions and robot position) and the planning algorithm. Sensitivity analysis enables to identify the parameters that are most critical for success, thereby providing directions for future research. Repeated-measures logistic regression was used to evaluate if changes in parameter values significantly influenced the success of finding a collision-free goal configuration of the manipulator and end-effector, and of planning a collision-free path that is smooth.

These two analyses intend to contribute beyond existing literature that deals with motion planning or sensitivity analysis of motion planning. Existing studies of fruit harvesting range from simple manipulator control and kinematics to advanced motion planning, and address various crops: melon (Edan et al. 2000), watermelon (Sakai et al. 2008), orange (Hannan and Burks 2004; Sivaraman and Burks 2006), apple (Guo et al. 2010), tomato (Liang and Wang 2010; Kondo et al. 2007; Kondo et al. 1996) and cucumber (Van Henten et al. 2003a; Van Willigenburg et al. 2004; Van Henten et al. 2010). The current study focuses not only on solving the motion planning problem, but also analyses the azimuth angle of the end-effector and parameter sensitivity. In other robot application fields, sensitivity analysis focussed on improving trajectories (de Luca et al. 1991), comparing robot designs (Tannous et al. 2014), or positioning errors of the end-effector (Zhang et al. 2012). Whereas previous studies

concentrated only on robot parameters, the contribution of our sensitivity analysis is to also consider parameters of the planning algorithm, and of the crop, i.e. the working environment of the harvesting robot.

The 'constrained-azimuth' method may serve as a useful generic method to approach fruit in crops with a similar plant architecture, such as tomato, cucumber, eggplant, or cherry. Furthermore, the sensitivity analysis demonstrates a strong effect of fruit location on motion planning success. This finding is an incentive for future research on modifying the crop environment, besides improving the robot design and planning algorithm. The methodology applied may be useful to study motion planning in other applications with a dense obstacle environment.

6.2 Components of the motion planning problem

The robot has been developed as part of the research project CROPS 'Clever Robots for Crops' (CROPS 2014). Section 6.2.1 details this robot, Section 6.2.2 describes the crop row used and Section 6.2.3 describes the algorithms used to address the motion planning problem.

6.2.1 Robot

The robot consisted of a platform, a manipulator and an end-effector (Figure 6-2). The platform developed (Jentjens Machinetechniek B.V., The Netherlands) uses the heating pipes of the greenhouse as a rail system to travel along the crop row.

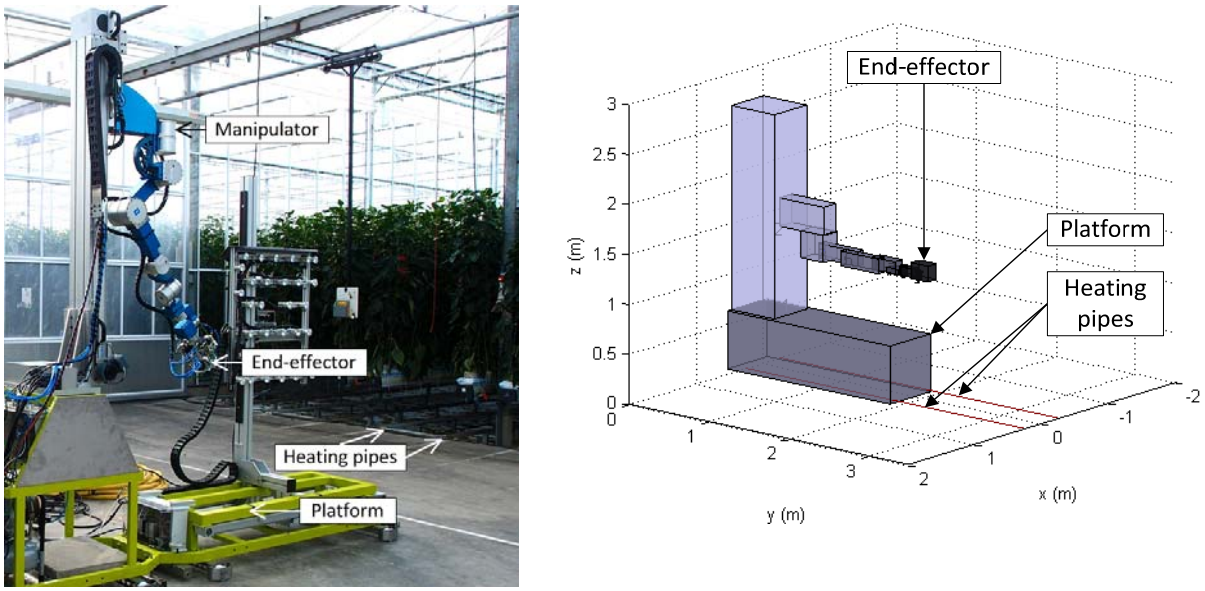


Figure 6-2. Photo of the platform, manipulator and end-effector (left), and a model representation of the manipulator based on bounding boxes (right). The robot platform travels in the y-direction (along the crop row) and uses the heating pipes as a rail system.

The manipulator (Technical University Munich, Germany) comprises nine degrees-of-freedom (DOF). The first joint is prismatic, whereas the other joints are rotational. We derived transformations between the links (Table 6-1), using the modified Denavit-Hartenberg convention (Craig 2004), where a_{i-1} refers to the link length (m) of link i, α_{i-1} refers to the

link twist (rad) of link i, d_i refers to the link offset (m) of link i, and θ_i refers to the joint angle (rad) at axis i.

Table 6-1 Denavit-Hartenberg parameters describing transformations between links of the manipulator.

Link i	a_{i-1} (m)	α_{i-1} (rad)	d_i (m)	θ_i (rad)
1	0.115	0	0.796	$\pi/2 + 0.2352$
2	0.480	0	-0.200	-0.2352
3	0	$\pi/2$	0	0
4	0.350	π	0	0
5	0.350	π	0	0
6	0.180	$-\pi/2$	0	$-\pi/2$
7	0	$-\pi/2$	0.150	0
8	0	$\pi/2$	0	0
9	0	$-\pi/2$	0.065	0

Manipulators used for fruit harvesting typically consist of two to seven DOF (Bac et al. 2014b). An advantage of using many DOFs is a better target-reachability in cluttered environments. Yet, the disadvantage is that common motion planning algorithms have difficulty in finding a collision-free path within a reasonable time. Furthermore, additional DOFs decrease speed (the stiffness reduces, which imposes lower motor torques to avoid vibrations) and reliability, and increase costs of the manipulator.

The end-effector grasps the fruit using a suction cup and subsequently cuts the peduncle, which is the connecting stem between the fruit and main stem. The bounding box of the end-effector has dimensions of 170 mm (width), 260 mm (length), and 110 mm (height). More details are described in the patent (Tuijl and Wais 2014).

6.2.2 Crop row

6.2.2.1 Measurements in the greenhouse

A total of 60 plant stems, and 165 attached fruit, were recorded from a crop row (sweet-pepper cultivar: Waltz) grown in a commercial greenhouse in the Netherlands. We measured and mapped only hard obstacles (fruit, stem, support wire, robot) that may cause damage to the crop or robot. Soft obstacles (leaves, petioles) can be pushed aside (Bac et al. 2013) and were therefore disregarded. Stems and support wires were localized by stereo-vision (Bac et al. 2014a). The issue of inaccurate localization was resolved by using two cameras, resulting in a localization accuracy of ±4 mm in the depth direction and ±0.5 mm in width and height.

The following fruit parameters were manually measured (measurement resolution is indicated in parenthesis):

- Height of the fruit centre relative to the robot platform (10 mm) (Figure 6-3);
- Azimuth angle φ_{Fruit} with respect to the stem (10 degrees), to determine if the fruit is located at the front, left or right side of the stem (Figure 6-3);
- Fruit width (1 mm);
- Fruit height (1 mm);
- Offset, if present, between the fruit surface and stem (10 mm) (Figure 6-3).

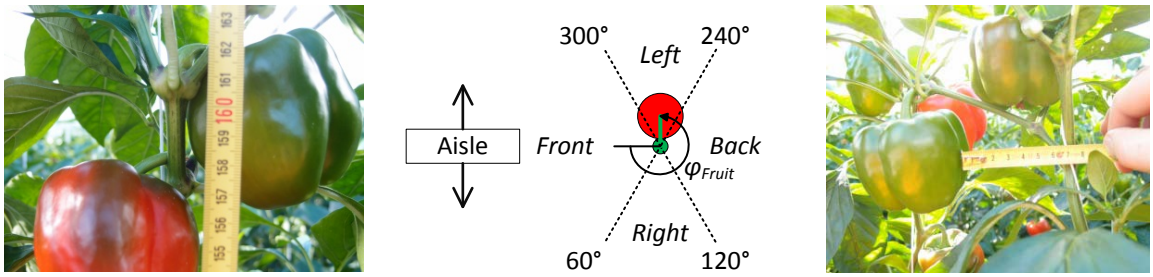


Figure 6-3. Manual measurements: height of the fruit centre relative to the robot platform (left), azimuth angle with respect to the stem (centre), and offset between fruit surface and stem (right).

During preliminary testing in the greenhouse, it was found that the end-effector can successfully detach a fruit if a collision-free azimuth angle is selected for the end-effector, whereas the skew and elevation angle of the fruit hardly influenced detachment success. Therefore, the elevation and skew angle were not measured and fruit were assumed to be vertically oriented, i.e. orthogonal to the floor. Correspondingly, all end-effector poses were taken in the horizontal plane and only the azimuth angle of the end-effector was varied. In addition, the position and orientation of the peduncle was not measured because the end-effector does not require this information for a successfully cutting the peduncle. These assumptions concerning the peduncle, and the skew and elevation angle may have to be reconsidered if a different type of end-effector is used in future work.

6.2.2.2 3D mapping

The measurements recorded in the greenhouse were used to generate a 3D virtual environment, using Matlab® 2012a (Mathworks Inc., USA). The 3D environment consisted of the robot, plant stems, attached fruit, and a support wire that was twisted around each stem (Figure 6-4).

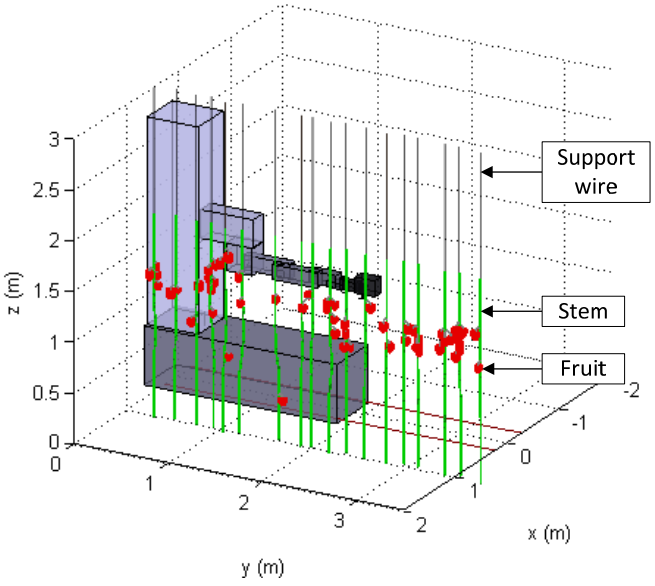


Figure 6-4. 3D virtual environment consisting of the manipulator, end-effector, plant stems, attached fruit, and a support wire that was twisted around each stem.

The workspace of the manipulator (Table 6-1) was limited to a height range of about one meter. Out of the total 165 fruit, seven were hanging below this height range (e.g. the lowest fruit in Figure 6-4). Therefore, only the 158 reachable fruit were analysed.

6.2.3 Algorithms

A sequence of algorithms (Figure 6-5) was implemented in Matlab® that includes three steps: calculation of the start and goal configuration (Section 6.2.3.2), path planning (Section 6.2.3.3) and path smoothing (Section 6.2.3.4). Each step uses a collision detection module (Section 6.2.3.1). Algorithms were run on a computer with an Intel Core i5 CPU 2.4 GHz processor with 4 GB of memory and Windows 7 as the operating system.

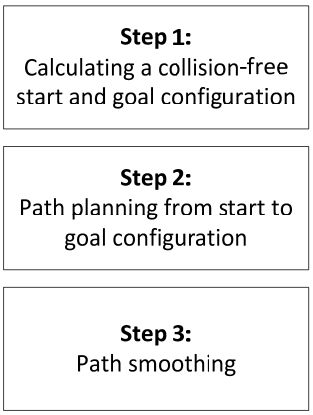


Figure 6-5. Sequence of algorithms for motion planning

6.2.3.1 Collision detection

The manipulator and end-effector were checked for self-collisions and for collisions with wires, stem segments or fruit. Self-collision detection was implemented using checks between spheres positioned at planes, corners and edges of the bounding box of links and at the end-effector. To reduce the number of checks, spheres were combined in an enclosing sphere (Figure 6-6), i.e. an octrees implementation (Meagher 1982). Link pairs 1-3, 5-9, 6-9, and 7-9 never collided and were therefore excluded from collision detection. Wires, stem segments and fruit were represented by a cylinder to resemble the shape of the object in reality. Therefore, a collision check with the manipulator or end-effector involved a sphere-cylinder check, which is more time consuming than a sphere-sphere check (Kodam et al. 2010).

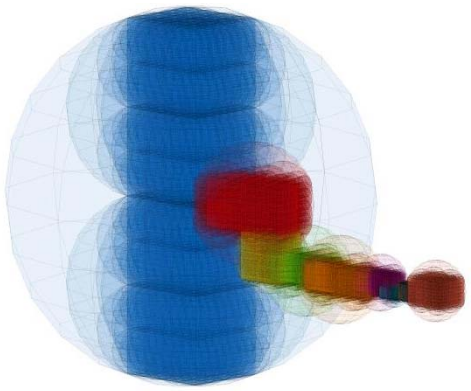


Figure 6-6. Bounding box representation of the manipulator with octrees of spheres used for collision detection. Colour coding is used to distinguish between the links.

6.2.3.2 Step 1: calculating a start and goal configuration

To calculate a goal configuration, an end-effector pose was selected with respect to the fruit and subsequently an inverse kinematics solution was calculated for this pose. Since end-effector poses were taken in the horizontal plane, only the azimuth angle of the end-effector φ_{EE} was varied (the pose is orthogonal to the aisle at $\varphi_{EE} = 0°$). The preferred angle of φ_{EE} is a pose where the end-effector is in line with the fruit and stem, as visualized in Figure 6-7. However, the preferred pose was not always reachable, due to presence of obstacles, and therefore a φ_{EE} was selected that deviates from this preferred pose, at an angle φ_{Dev} (Equation (3-6)). Preliminary field tests of the end-effector mechanism revealed that an increasing φ_{Dev} increases the probability of unsuccessful fruit detachment and of damaging the fruit or stem. Therefore, the pose with the smallest φ_{Dev} was selected, for path planning, from poses resulting in a collision-free goal configuration.

$$\varphi_{Dev} = \left|\varphi_{Fruit} - \varphi_{EE}\right| \quad (°) \qquad (6\text{-}1)$$

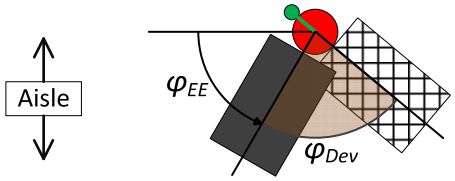


Figure 6-7. Azimuth angle φ_{EE} of the end-effector with respect to the fruit. Depending on φ_{EE}, the selected end-effector pose (■) may deviate from the preferred pose (□) by φ_{Dev}.

Inverse kinematics were calculated using the Robotics, Vision & Control toolbox (Corke 2011). The inverse kinematics solution was furthermore checked for collisions. The goal configuration was calculated for each fruit, whereas a fixed start configuration (Figure 6-4) was taken that was collision-free for all fruit.

6.2.3.3 Step 2: path planning

Most path planning algorithms were not suitable for this nine DOF manipulator because planning time increases exponentially with the number of DOFs (Choset et al. 2005). However, the sampling-based planner implemented, a balanced bi-RRT (Kuffner and Lavalle 2000; Lavalle 2006), is less affected by the number of DOFs. This planner randomly samples the configuration space using a user-defined sampling resolution ε. If the sample is collision-free, it is added to a growing tree. After a number of iterations, the trees growing from the start and goal configuration may connect and a path is found. In preliminary analyses, paths were mostly found within 200 iterations and, in an exceptional case, 800 iterations were needed. The maximum number of iterations was therefore set to 1500, which corresponds to a planning time of about half an hour. Such a long time is acceptable for objectives of this research, but the cycle time of the real robot should not exceed 6 s to be economically feasible (Pekkeriet 2011). Although reducing planning time is outside the scope of this research, a possible direction would be to implement the algorithm in C++ code and to use a fast collision detection algorithm (Sucan and Chitta 2014).

6.2.3.4 Step 3: path smoothing

Paths found by RRT-based planning algorithms tend to be tortuous due to the random nature of such sampling algorithms. Therefore, a path-smoothing algorithm was implemented following the path-planning phase. The implemented algorithm is an adaptation of a classical heuristic based on the divide and conquer concept, which consists of decomposing the original path and omitting some of its nodes, iteratively, in order to shorten it (Carpin and Pillonetto 2005). Additionally we incorporated a "memory matrix" where calculated sub-path options are retained for subsequent use. These modifications facilitate achieving shorter processing times. To ensure the smoothened path is indeed collision-free, consecutive vertices were linearly interpolated using ten configurations. A path was determined as collision-free if all interpolated configurations were collision-free.

6.3 Performance measures

Performance was measured in terms of success rate, planning time, and path quality. Success rate was comprised of the following:

- Suc_{Goal} - Goal configuration success (%), calculated by:

$$Suc_{Goal} = \frac{100 \cdot [\text{no. of fruits reached}]}{[\text{total no. of fruits on plants}]} \quad (\%) \qquad (6\text{-}2)$$

  This measure expresses the ratio of fruits for which a collision-free goal configuration was found.

- Suc_{Path} - Path planning success (%), calculated by:

$$Suc_{Path} = \frac{100 \cdot [\text{no. of fruits for which a collision-free path is planned}]}{[\text{no. of fruits reachable}]} \quad (\%) \qquad (6\text{-}3)$$

  A path was considered collision-free if it was collision-free before and after path smoothing.

- Suc_{Mot} - Motion planning success (%), calculated by:

$$Suc_{Mot} = \frac{Suc_{Goal} \cdot Suc_{Path}}{100} \quad (\%) \qquad (6\text{-}4)$$

Planning time refers to the wall-clock time required to plan a path. Path quality was quantified by the Joint Angles Index of Curvature (JAIC) (Equation (6-5)). JAIC, measured in the configuration space, is based on the city-block distance metric capturing the overall sum of the joint angle paths. As JAIC measures how much the joints rotate, it reflects power consumption by joint actuators. JAIC is normalized by the sum of joint angle paths for the "line-of-sight" path (the shortest path from start point to goal point), thus JAIC ranges from 1 (the shortest path) to zero. The normalization facilitates comparing path quality among groups of fruit.

$$JAIC = \frac{\sum_{j=1}^{J} \left| x_{1,j} - x_{V,j} \right|}{\sum_{i=2}^{V} \sum_{j=1}^{J} \left| x_{i-1,j} - x_{i,j} \right|} \quad (-) \qquad (6\text{-}5)$$

Where: $x_{i,j}$ (deg.) is the i^{th} vertex's position in the j^{th} dimension of the configuration space (the j^{th} joint's angle value); V the total number of vertices in the path; and J the total number of joints, which is nine in the current implementation.

6.4 Parameters

The influence of five parameters, on success or planning time, was investigated in this research (Table 6-2): Average stem spacing (S_{Stem}), fruit location at a side of the stem (S_{Fruit}), dimensions of the end-effector (D_{EE}), Robot position (P_{Robot}) and sampling resolution (ε). Few values were selected per parameter to simplify analyses in Section 6.5.

Table 6-2. Parameters analysed in this research

Parameter	Description	Values tested
S_{Stem}	Change (%) in average stem spacing (m)	[0 (narrow), 50 (wide)]
S_{Fruit}	Fruit location at a side of the stem	[Front, Left, Right, Back]
D_{EE}	Change (%) in dimensions (width, length, height) of the end-effector	[-25 (small), 0 (big)]
P_{Robot}	Position of robot (m) along the rail, relative to the fruit	[0, 0.1, ...0.7]
ε	Sampling resolution of RRT algorithm (-)	[0.03,0.05,0.07]

Average stem spacing (S_{Stem}) was expected to influence Suc_{Goal} and Suc_{Path} because it modifies the density of the obstacle map. The parameter value depends on the cultivation system used for sweet-pepper production. Plants were planted precisely 0.4 m apart and pruned such that two stems developed, in an approximately upward direction. Stem spacing therefore strongly varied from 0.1-0.3 m, among harvest cases, whereas average stem spacing was precisely known (0.2 m). We tested an increase of 50%, to 0.3 m, to realise a strong change in density of the obstacle map.

Fruit location at a side of the stem (S_{Fruit}) was expected to influence Suc_{Goal} and Suc_{Path} because, for instance, backside fruit are further away from the robot and obstacles are more likely to obstruct a path, than for front-side fruit. Fruit were grouped in four sides based on the measured azimuth angle at the fruit φ_{Fruit} (Figure 6-3): front ($300° < \varphi_{Fruit} \leq 60°$), right ($60° < \varphi_{Fruit} \leq 120°$), back ($120° < \varphi_{Fruit} \leq 240°$), or left side ($240° < \varphi_{Fruit} \leq 300°$).

Changing the dimensions of the end-effector (D_{EE}) was expected to influence goal configuration success (Suc_{Goal}) and path planning success (Suc_{Path}) because D_{EE} influences the size of the collision-free workspace of the manipulator. Dimensions were precisely known

and do not vary among harvest cases. A 25% decrease was investigated because such dimensions are feasible after another design iteration.

Robot position (P_{Robot}) influences dexterity of the manipulator (Klein and Blaho 1987), i.e. the range of motion toward a fruit. In addition, variation in position and orientation of the fruit influence the robot position required to reach a fruit. Therefore, P_{Robot} was expected to influence Suc_{Goal} and Suc_{Path}. For each fruit, eight robot positions were tested. The first position was always taken at an offset of 1.5 m between robot and target fruit. Subsequently, steps of 0.1 m were made until the last position (Figure 6-8). Investigating this range of robot positions assured that fruit were within the workspace, whereas for a wider range of robot positions, fruit were outside the workspace. To choose a robot position for path planning, we selected the one with most collision-free azimuth angles of the end-effector (φ_{EE}). If several robot positions resulted in the same number of angles, the one closest to the middle position (P_{Robot} = 0.3 or 0.4) was selected because dexterity was greatest at these positions.

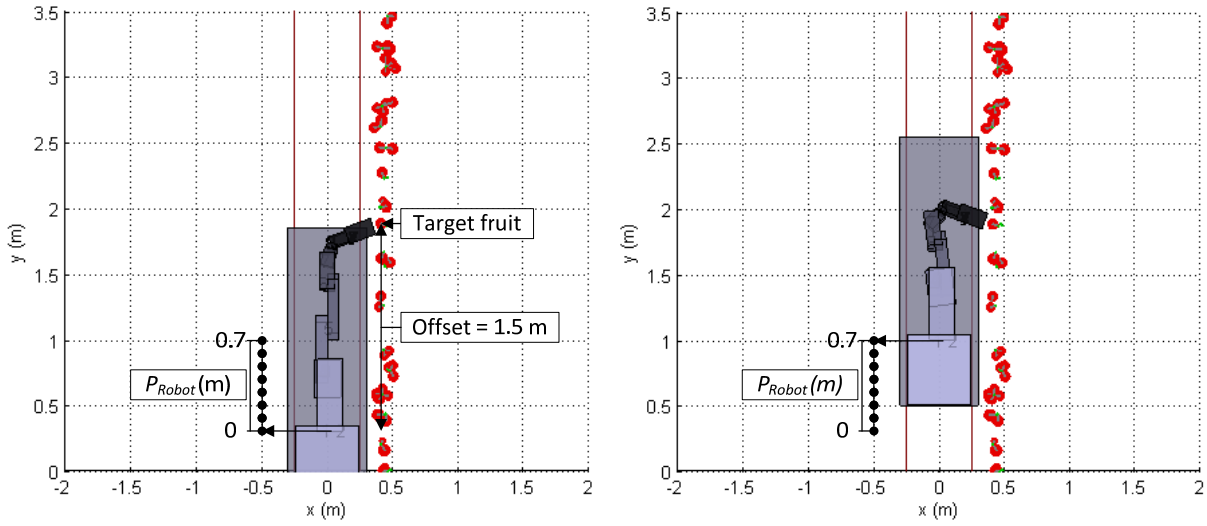


Figure 6-8. Visualization of robot position (P_{Robot}), with values ranging from of 0 m (left) to 0.7 m (right) in steps of 0.1 m. Initial robot position (P_{Robot} = 0) was put at a fixed offset of 1.5 m, between the robot and target fruit.

Sampling resolution (ε) is a parameter in the bi-RRT algorithm and determines the probability of finding a path. A smaller resolution increases this probability at the cost of additional iterations. A small resolution is typically required for fruit surrounded by many obstacles. At each fruit, the resolution was initially set to 0.07. If a path was not found, or if the path contained a collision after path smoothing, a resolution of 0.05 and 0.03 were subsequently attempted. Smaller resolutions were not tested because the probability of existence of a solution would become very low and the maximum number of iterations would easily be exceeded.

6.5 Analyses

Two analyses were performed by simulation. In the first analysis, all three steps of the algorithm sequence were performed and goal configuration success, path planning success and motion planning success were compared between the constrained-azimuth method and the full-azimuth method (Section 6.5.1). The constrained-azimuth method turned out most

favourable and was therefore used for the second analysis that involved a sensitivity analysis. We tested the effect of stem spacing, fruit location and end-effector dimensions on goal configuration success (Section 6.5.2). Subsequently, the goal configurations obtained were used to plan a path an smooth the path, where we tested sensitivity of stem spacing, end-effector dimensions and the sampling resolution (ε) on path planning success (Section 6.5.3).

6.5.1 Constrained-azimuth method vs. full-azimuth method

As mentioned in the Introduction, a constrained-azimuth method may offer advantages in terms of planning time and avoiding a motion that risks collisions. To validate this expectation, we compared this method with the full-azimuth method and determined motion planning success. We evaluated goal configurations in simulation and subsequently path planning was performed. For the constrained-azimuth method, 13 azimuth angles φ_{EE} (-60, -50, ... 60°) were simulated, whereas all 36 azimuth angles φ_{EE} (0, 10, ... 350°) were simulated for the full-azimuth method. The configuration with the smallest φ_{Dev} was selected as goal configuration for path planning.

The performance measures (Section 6.3) were measured for both methods. In addition, we counted fruit that were either reachable by only the full-azimuth method, or reachable by both methods, but where the full-azimuth method resulted in a smaller φ_{Dev}.

6.5.2 Sensitivity analysis of finding a goal configuration

This sensitivity analysis tested if a successful goal configuration (Suc_{Goal}) existed for a fruit, regardless whether a path existed toward that goal configuration. In Section 6.5.2.1, we assessed the effect of stem spacing, fruit location and dimensions of the end-effector on Suc_{Goal}. Section 6.5.2.2 describes the effect of distinct robot positions on Suc_{Goal}.

6.5.2.1 Stem spacing, fruit location and dimensions of the end-effector

S_{Stem} and D_{EE} each involved two levels and therefore four combinations of these two parameters were tested on goal configuration success (Suc_{Goal}). Subsequently, results were grouped by S_{Fruit}: all sides (N=158), the front side (N=46), left (N=31) and right side (N=32) combined, and the backside of the stem (N=48).

To determine if S_{Stem}, S_{Fruit}, and D_{EE} had a statistically significant effect on Suc_{Goal}, we performed a repeated measures logistic regression (Field 2009). The effect of robot position was disregarded, to simplify the regression analysis. Hence, a fruit was successfully reached, if at least one of the eight robot positions resulted in a successful goal configuration. Logistic regression is a multiple regression with a categorical output variable and predictor variables that are continuous or categorical. The values of the predictor variables determine probability of membership to a category: success or failure. Predictor variables were pair-wise compared, to validate the following hypotheses:

- Suc_{Goal} is significantly greater for fruit at the front side than any of the other three sides, i.e. three pair-wise comparisons;
- Dimensions of the end-effector (D_{EE}) have a significant effect on Suc_{Goal}, i.e. a single pair-wise comparison;
- Stem spacing (S_{Stem}) has a significant effect on Suc_{Goal}, i.e. a single pair-wise comparison;

- There are significant parameter interaction effects between D_{EE}, S_{Stem} and side of fruit location.

The output was presented in a table, with an odds ratio and a probability value to determine significance.

The side predictor involved multiple pair-wise comparisons on a single set of data. To reduce the chances of obtaining false-positive results in such a statistical model, we performed post-hoc analysis adjusted using Bonferroni correction (Field 2009). Furthermore, we applied the Quasi likelihood Independence Criterion (QIC) to test the contribution of each predictor to the statistical model (Field 2009).

6.5.2.2 Distinct robot positions

In previous analyses all robot positions were tested, whereas in this analysis the effect of distinct robot positions (P_{Robot}) on goal configuration success (Suc_{Goal}) was tested, for combinations of end-effector dimensions and stem spacing: S_{Stem} = narrow & D_{EE} = big, S_{Stem} = narrow & D_{EE} = small, S_{Stem} = wide & D_{EE} = big and S_{Stem} = wide & D_{EE} = small. This analysis enabled to determine if goal configuration success was sensitive to distinct robot positions.

We validated if it was required to optimize the robot position in order to reach the fruit, by comparing Suc_{Goal} between a fixed robot position (P_{Robot} = 0.3) vs. any of the eight positions. This test was conducted for one parameter combination: S_{Stem} = narrow and D_{EE} = big.

To evaluate if robot position influences the expected success of detachment, φ_{Dev} was measured at distinct robot positions. At each position, a mean φ_{Dev} was taken over the fruit belonging to one of the twelve possible parameter combinations of S_{Stem}, D_{EE} and S_{Fruit}.

6.5.3 Sensitivity analysis of path planning and path smoothing

Goal configurations obtained in Section 6.5.2 were used in this analysis, to determine if a collision-free and smooth path can be planned from the start configuration to a goal configuration. We tested the influence of D_{EE} and S_{Stem} (same combinations as in Section 6.5.2.2) on path planning success Suc_{Path} and path quality measure $JAIC$. Furthermore, only one planning attempt was performed for each fruit due the probabilistic completeness property of the bi-RRT planner. Probabilistic completeness means that, if a solution exists, the probability that the planner solves the problem goes to 1 as the planning time goes to infinity (Berenson and Srinivasa 2010). Probabilistic completeness is closely approximated, but not fully met, because a maximum number of iterations (1500) is set. This maximum is about double the maximum number of iterations observed to find a path (800). Hence, we consider the probability of finding a solution beyond 1500 iterations to be extremely low.

6.6 Results

6.6.1 Constrained-azimuth method vs. full-azimuth method

Goal configuration success (Suc_{Goal}) of the full-azimuth method (66%) was slightly greater than for the constrained-azimuth method (63%) (Table 6-3). Five fruit were reachable by only the full-azimuth method, but for only one fruit a path was found. Eight fruit were reachable by both methods where the full-azimuth method resulted in a smaller φ_{Dev}, but for only one fruit

a path was found. As a result, motion planning success (Suc_{Mot}) of the constrained-azimuth method (63%) was similar to the full-azimuth method (64%).

Table 6-3. Comparison of success rates between the constrained-azimuth method and the full-azimuth method.

	Constrained-azimuth	Full-azimuth
Suc_{Goal} (%)	63	66
Suc_{Path} (%)	100	96
Suc_{Mot} (%)	63	64

6.6.2 Sensitivity analysis of finding a goal configuration

6.6.2.1 Stem spacing and dimensions of the end-effector

Figure 6-9 displays the effect of stem spacing S_{Stem} and D_{EE} on goal configuration success (Suc_{Goal}), grouped by fruit location at the stem. Fruit located at the front side (Suc_{Goal}=93%) of the stem were better reachable than left+right side (Suc_{Goal}=59%) and backside (Suc_{Goal}=41%). In addition, smaller end-effector dimensions (D_{EE}) and wide stem spacing (S_{Stem}) both improve goal configuration success (Suc_{Goal}).

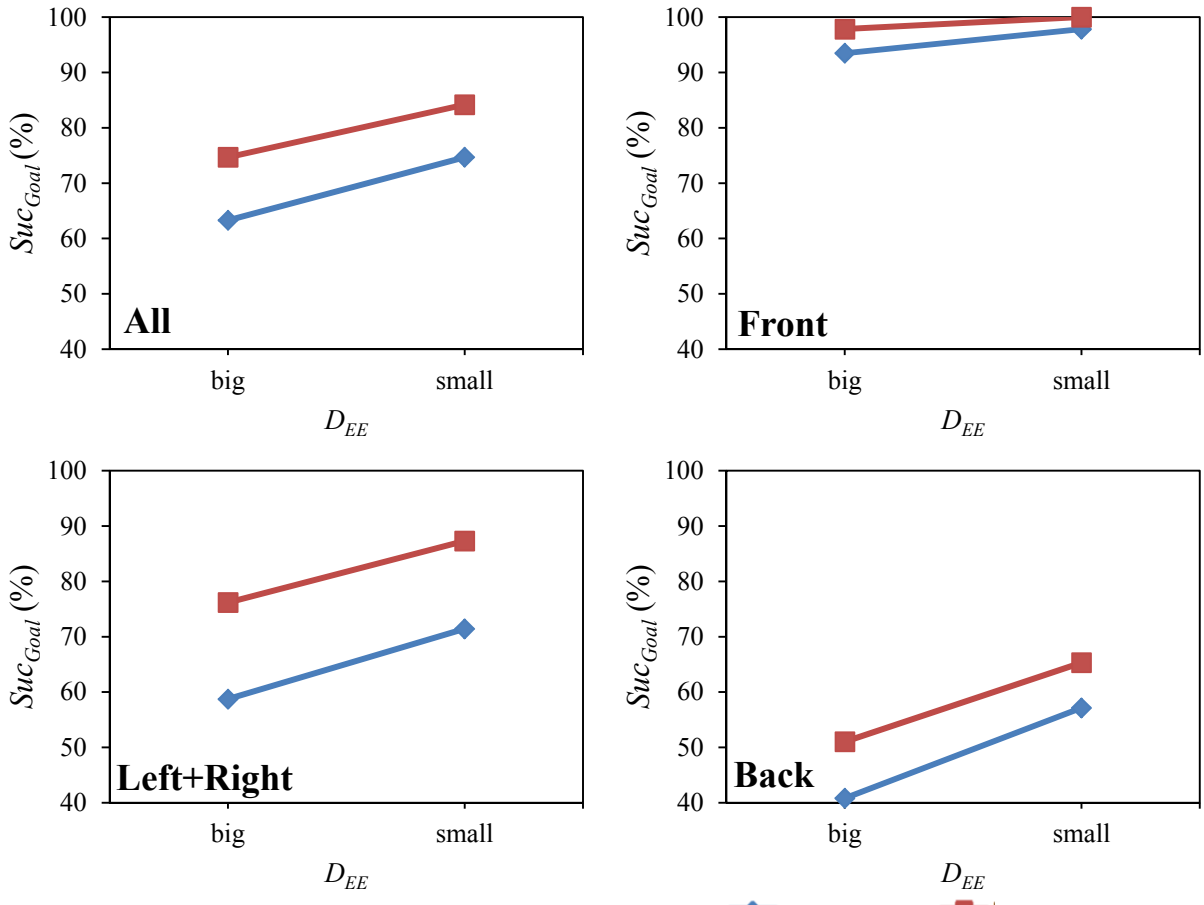


Figure 6-9. Effect of end-effector dimensions (D_{EE}) and stem spacing (S_{Stem} = narrow; S_{Stem} = wide) on goal configuration success (Suc_{Goal}). Graphs are grouped by fruit location: all 158 fruit (top-left), 46 front-side fruit (top-right), 63 left+right-side fruit (bottom-left), and 48 back-side fruit (bottom-right).

In the repeated measures logistic regression of Suc_{Goal}, interactions were found to have no significant influence on Suc_{Goal}, and therefore were all removed from the model. The main effects were all found significant and results are presented in Table 6-4. All three predictors have a significant effect ($p \leq 0.01$) on goal configuration success. The odds of reaching a fruit, using smaller end-effecter dimensions (-25%), are about twice (1/0.52) as high than using the big end-effector dimensions. In addition, the odds of reaching a fruit, when stems are additionally spaced apart (+50%), is about twice as high (1/0.52) than with narrow stem spacing. Yet, the strongest influencing predictor was the side of fruit location. The odds of reaching fruit at the front side is significantly higher than any other side (right = 1/0.099, left = 1/0.112, back = 1/0.046).

Table 6-4. Results of repeated measures logistic regression for goal configuration success. Predictors are fruit location (three pair-wise comparisons of the side), stem spacing (S_{Stem}), and dimensions of the end-effector (D_{EE}).

		95% Confidence interval for Odds Ratio		
	Beta (Std. Error)	Lower	Odds Ratio	Upper
Constant	3.91 (0.63)**	14.62	49.82	169.77
S_{Fruit} = right vs. front	-2.31 (0.71)*	0.025	0.099	0.394
S_{Fruit} = left vs. front	-2.19 (0.71)*	0.028	0.112	0.450
S_{Fruit} = back vs. front	- 3.09 (0.67)**	0.012	0.046	0.169
S_{Stem} = narrow vs. wide	-0.65 (0.11)**	0.416	0.520	0.650
D_{EE} = big vs. small	-0.65 (0.11)**	0.420	0.520	0.645

Note: **p<0.001, *p<0.01

The Bonferroni corrected post-hoc analysis of side revealed that only front is significantly different from all other sides (right and left: p<0.01, back: p<0.001). Although the influence of the side predictor S_{Fruit} is much higher than that of D_{EE} or S_{Stem}, including D_{EE} and S_{Stem} predictors improves the model according the Quasi likelihood Independence model Criterion (QIC) (610 for all three predictors, 629 for just S_{Fruit}). Hence, this finding explains that all three predictors should be considered in the statistical model.

6.6.2.2 Distinct robot positions

Figure 6-10 displays the effect of distinct robot positions on goal configuration success (Suc_{Goal}). The difference between minimum and maximum goal configuration success is greater for S_{Stem} = narrow & D_{EE} = big (14%) than for S_{Stem} = wide & D_{EE} = small (7%). This finding indicates that robot position becomes more critical for decreasing stem spacing and increasing end-effector dimensions. Moreover, for the best position (P_{Robot} = 0.3 m), goal configuration success was 61% (Figure 6-10). When testing all positions, Suc_{Goal} further increased to 63% (not visible in Figure 6-10). The three fruit (2%) that were not reachable from P_{Robot} = 0.3 m, were accessible from P_{Robot} = 0.4 or 0.5 m. Therefore, one can conclude that fruit were always accessible from robot positions where the manipulator had a high dexterity (P_{Robot} = 0.2-0.5 m).

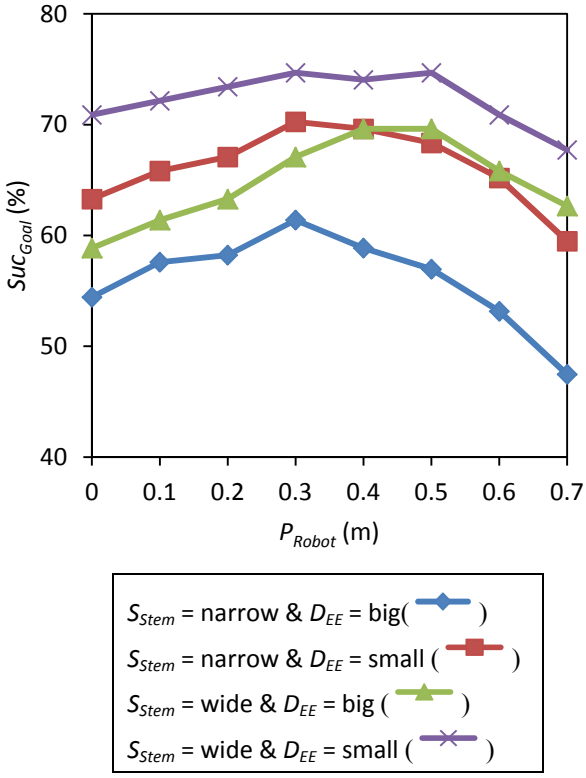


Figure 6-10. Effect of robot position (P_{Robot}) on goal configuration success (Suc_{Goal}), grouped by four combinations of S_{Stem} and D_{EE}. All 158 fruit were evaluated to determine each value.

The effect of distinct robot positions on φ_{Dev} is presented in Figure 6-11. Parameter combinations of S_{Stem} and D_{EE} do not strongly influence φ_{Dev}, whereas fruit location has a strong influence.

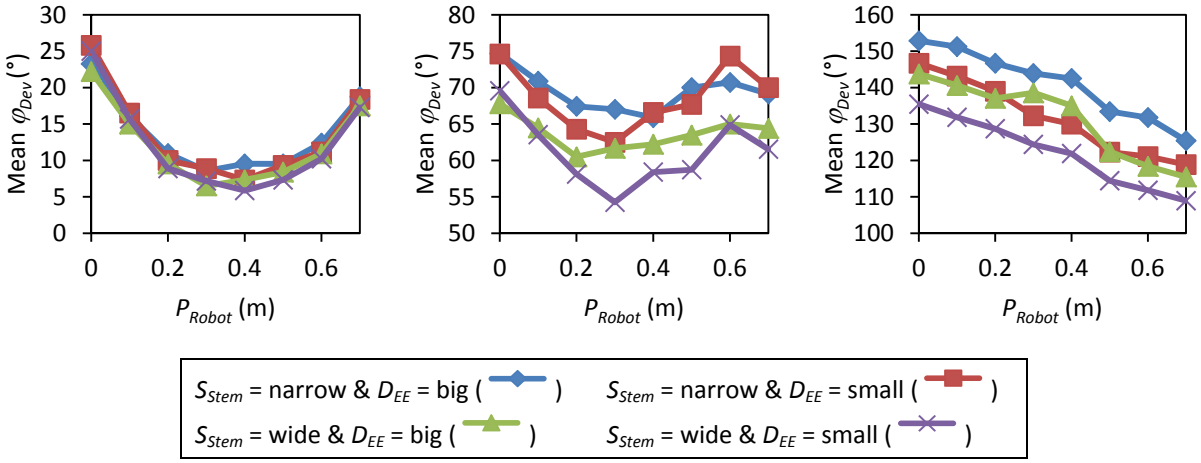


Figure 6-11. Effect of robot position (P_{Robot}) on mean φ_{Dev} , for fruit locations at the front-side (left), left+right-side (centre) and backside (right) of the stem. Lines in each graph indicate parameter combinations of S_{Stem} and D_{EE}.

Fruit locations at the front-side of the stem can be reached with a small mean φ_{Dev} of <30° that is desirable for successful detachment. At the left and right side, the deviation angle φ_{Dev} deteriorated with values ranging between 55° and 75°. For backside fruit, mean φ_{Dev}

drastically worsened with values greater than 110°. Moreover, for front, left and right side fruit, mean φ_{Dev} is smaller for robot positions ranging between 0.2 and 0.5 m, than for "extreme" robot positions. This finding supports earlier results shown in
Figure 6-10.

6.6.3 Sensitivity analysis of path planning and path smoothing

Results of path planning and smoothing reveal no parameter influences on Suc_{Path} (Table 6-5) and path planning is generally successful (>97%). No notable difference in path quality was observed, in all four combinations.

Table 6-5. Results of path planning and path smoothing

Combination	Path planning				Path Quality	
	Suc_{Path} in (%)	Number of fruit evaluated (#)	Proportion of paths found per ε: 0.07, 0.05, 0.03	Mean Planning time in s (SD)	Mean JAIC before smoothing	Mean JAIC after smoothing
S_{Stem} = narrow & D_{EE} = big	100	100	90%,7%,3%	56 (96)	0.73	0.91
S_{Stem} = narrow & D_{EE} = small	99	118	85%, 12%, 3%	63 (136)	0.74	0.93
S_{Stem} = wide & D_{EE} = big	97	118	95%, 5%, 0%	86 (161)	0.74	0.91
S_{Stem} = wide & D_{EE} = small	98	133	95%, 4%, 1%	97 (227)	0.74	0.94

Sampling resolution (ε) required was 0.07 for at least 85% of the paths. For the remaining paths a finer resolution was required. Planning time, for combined parameter combinations, was on average (range) 59 (3-1494) s for ε=0.07 with N=423, 137 (23-607) s for ε=0.05 with N=27, and 401 (42-901) s for ε=0.03 with N=8. Consequently, a smaller resolution increases planning time and planning time strongly varies.

6.7 Discussion

The new constrained-azimuth method resulted in successfully planned motions for a total of 100 fruit (63%), and only one additional fruit was reached by the full-azimuth method. The constrained-azimuth method simplifies the challenging planning problems by selecting a different goal configuration. Furthermore, the constrained-azimuth method may result in less damages to the crop because the full-azimuth method may collide with an unseen plant part when approaching a fruit from the side or backside of the stem. Other crops with a similar plant architecture may benefit from this method as well and validating this hypothesis is therefore a task for future work. Lastly, the value of 120°, to constrain the azimuth angle, was empirically determined and not optimized in this study. Optimizing this value is therefore a task for future work.

Fruit location strongly determined goal configuration success. Especially backside fruit were hard to reach and, if reached, the end-effector pose resulted in a large deviation angle (φ_{Dev}>110°) that will probably lead to unsuccessful fruit detachment. How strongly

detachment success will decrease, at such an increasing deviation angle, is currently unknown and should be determined in future field tests. In related work, fruit location also played a dominant role. For strawberry harvesting, harvest success was only 8% for fruit located behind other fruit, yet 72% for fully accessible fruit (Hayashi et al. 2010). To address the issue of backside fruit or complex harvest cases, three directions can be pursued. Firstly, the end-effector design can be adapted to widen the allowable range of φ_{Dev}. Secondly, adaptations in genetics, the cultivation system, or cultivation practices can be implemented to generate more fruit at the front side of the plant stem (Bac et al. 2014b). Thirdly, human-robot co-working can be applied, where the robot only harvests easy harvest cases, whereas humans harvest the remaining complex cases.

Although the end-effector interacts more with the plant than the manipulator, the dimensions of manipulator links and the kinematic structure may also influence motion planning success. Different lengths and dimensions of the links may influence the reachable workspace of the manipulator and optimizing the robot position may therefore become more relevant. Optimizing the robot position increased goal configuration success by only 2% for the manipulator used in this research. Furthermore, the number of degrees-of-freedom (DOF) influences the manoeuvrability around obstacles. In this application, planned motions were relatively simple and fewer DOFs may not have a strong influence on the result. However, in other applications, DOFs could play an important role. When reusing the methodology in future work, these manipulator aspects should be considered.

An assumption was that no contact or movement of an obstacle is allowed. In practice, however, workers commonly touch or move fruit or stems aside to better access fruit. Relaxing the hard obstacle constraint and implementing danger zones (Sent and Overmars 2001) may improve motion planning success and is therefore a task for future work.

Another assumption was that soft obstacles, i.e. leaves and petioles, could be pushed aside and were therefore disregarded in the analyses. If a soft obstacle would be pushed aside, the fruit position might change and lead to unsuccessful detachment. Furthermore, leaves can get stuck in the end-effector, which may lead to contamination and improper functioning of the end-effector. Hence, the influence of soft obstacles on motion planning success deserves further study in future work.

A collision-free path was not found for 37% of the sweet-pepper fruit, whereas for cucumber the motion planner was responsible for 4.9% of the total number of failed attempts during a field test (Van Henten et al. 2003b). Although robot designs differed, one may conclude that sweet-pepper is more challenging to harvest than cucumber, probably due to the dense obstacle spacing in a sweet-pepper crop.

The success rates reported may differ from success rates of a future field test with the actual robot because additional sources of error may lower the motion planning success. These errors include hardware related errors, imprecise positioning of the end-effector, and path obstructions by objects not included in the obstacle map (Van Henten et al. 2003b). Such errors were not investigated in this research, but are relevant upon implementation of the robot in a greenhouse. A comparison of success rates between this study and an actual field test is therefore a task for future work. Nevertheless, this study provides more insight in which parameters cause the suboptimal motion planning success of 63%.

6.8 Conclusion

Motion planning success was 63% (N=158) for the constrained-azimuth method and 64% for the full-azimuth method. Approaching fruit from only the aisle side is successful because motion planning success hardly decreases, whereas this method avoids long planning time and avoids executing paths with a potential risk for colliding and damaging a plant.

The sensitivity analysis revealed that reducing end-effector dimensions and widening stem spacing both significantly improve the goal configuration success, from 63% to 84%. Yet, the strongest influencing factor on goal configuration success is the side of fruit location. Fruit located at the front of the stem are significantly better reachable than any other side. Hence, this new finding provides an incentive to train or breed plants in such a way that more fruit develop at the front side of the plant. Furthermore, goal configuration success was 61% for a fixed robot position and 63% for an optimized robot position. Robot position therefore hardly affects goal configuration success in this study. However, robot position influences deviation in azimuth angle between end-effector and fruit. The effect of this deviation on success of fruit detachment by the end-effector is yet to be tested in future field tests.

Path planning success was in the range of 98-100% for the four parameter combinations of end-effector dimensions and stem spacing. Hence, paths are almost flawlessly found, probably because challenging planning problems were simplified by applying the constrained-azimuth method to determine a goal configuration. Furthermore, more than 85% of the paths were found by a sampling resolution of 0.07. Smaller sampling resolutions increased planning time.

Acknowledgements

This research was funded by the European Commission in the 7th Framework Programme (CROPS GA no. 246252) and by the Dutch horticultural product board (PT no. 14555). We are grateful to growers Cees and Rolf Vijverberg who allowed us to record measurements in their greenhouse.

References

Bac, C. W., Hemming, J., & Van Henten, E. J. (2013). Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper. *Computers and Electronics in Agriculture, 96*, 148-162.

Bac, C. W., Hemming, J., & van Henten, E. J. (2014a). Stem localization of sweet-pepper plants using the support wire as a visual cue. *Computers and Electronics in Agriculture, 105*, 111-120.

Bac, C. W., Van Henten, E. J., Hemming, J., & Edan, Y. (2014b). Harvesting robots for high-value crops: state-of-the-art review and challenges ahead. *Journal of Field Robotics, DOI: 10.1002/rob.21525*.

Berenson, D., & Srinivasa, S. S. Probabilistically complete planning with end-effector pose constraints. In *Proceedings - IEEE International Conference on Robotics and Automation, 2010* (pp. 2724-2730).

Carpin, S., & Pillonetto, G. (2005). Motion planning using adaptive random walks. *IEEE Transactions on Robotics, 21*(1), 129-136.

Choset, H., Lynch, K. M., Hutchinson, S., Kantor, G., Burgard, W., Kavraki, L. E., et al. (2005). *Principles of Robot Motion*. Cambridge, MA: MIT Press.

Corke, P. I. (2011). *Robotics, Vision & Control*. Germany, Berlin: Springer.

Craig, J. J. (2004). *Introduction to Robotics* (3rd ed.). Upper Saddle River, NJ: Pearson Education Inc.

CROPS (2014). CROPS: Clever Robots for Crops. http://www.crops-robots.eu/. Accessed 1-9-2014.

de Luca, A., Lanari, L., & Oriolo, G. (1991). A sensitivity approach to optimal spline robot trajectories. *Automatica, 27*(3), 535-539.

Edan, Y., Rogozin, D., Flash, T., & Miles, G. E. (2000). Robotic melon harvesting. *IEEE Transactions on Robotics and Automation, 16*(6), 831-835.

Field, A. (2009). Ch. 8: Logistic Regression. In *Discovering statistics using SPSS* (3rd ed.): Sage Publication Ltd. London.

Guo, J., Zhao, D. A., Ji, W., & Xia, W. (Motion Planning). Design and control of the open apple-picking-robot manipulator. In *Proceedings - 2010 3rd IEEE International Conference on Computer Science and Information Technology, ICCSIT 2010, 2010* (Vol. 2, pp. 5-8).

Hannan, M. W., & Burks, T. F. Current developments in automated citrus harvesting. In *ASAE Annual International Meeting 2004, 2004* (pp. 4187-4196).

Hayashi, S., Shigematsu, K., Yamamoto, S., Kobayashi, K., Kohno, Y., Kamata, J., et al. (2010). Evaluation of a strawberry-harvesting robot in a field test. *Biosystems Engineering, 105*(2), 160-171.

Klein, C. A., & Blaho, B. E. (1987). DEXTERITY MEASURES FOR THE DESIGN AND CONTROL OF KINEMATICALLY REDUNDANT MANIPULATORS. *International Journal of Robotics Research, 6*(2), 72-83.

Kodam, M., Bharadwaj, R., Curtis, J., Hancock, B., & Wassgren, C. (2010). Cylindrical object contact detection for use in discrete element method simulations. Part I – Contact detection algorithms. *Chemical Engineering Science, 65*(22), 5852-5862.

Kondo, N., Nishitsuji, Y., Ling, P. P., & Ting, K. C. (1996). Visual feedback guided robotic cherry tomato harvesting. *Transactions of the ASAE, 39*(6), 2331-2338.

Kondo, N., Taniwaki, S., Tanihara, K., Yata, K., Monta, M., Kurita, M., et al. (2007). End-Effector and Manipulator Control for Tomato Cluster Harvesting Robot.

Kuffner, J. J., & Lavalle, S. M. (Motion Planning). RRT-connect: an efficient approach to single-query path planning. In *Proceedings - IEEE International Conference on Robotics and Automation, 2000* (Vol. 2, pp. 995-1001).

Lavalle, S. M. (2006). *Planning Algorithms*. Cambridge, UK: Cambridge University Press.

Lewis, A., Watts, P. L., & Nagpal, B. K. (1983). Investment Analysis for Robotic Applications. *Technical Paper - Society of Manufacturing Engineers. MS*.

Liang, X., & Wang, Y. (2010). Analysis on the workspace of a tomato harvesting manipulator with 7 DOF. *Applied Mechanics and Materials* (Vol. 37-38, pp. 155-161).

Meagher, D. (1982). Geometric modeling using octree encoding. *Computer Graphics and Image Processing, 19*(2), 129-147.

Pekkeriet, E. J. (2011). CROPS project deliverable 12.1: Economic viability for each application. Wageningen, The Netherlands: Wageningen UR Greenhouse Horticulture.

Sakai, S., Iida, M., Osuka, K., & Umeda, M. (2008). Design and control of a heavy material handling manipulator for agricultural robots. *Autonomous Robots, 25*(3), 189-204.

Sent, D., & Overmars, M. H. (Motion Planning). Motion planning in environments with dangerzones. In *Proceedings - IEEE International Conference on Robotics and Automation, 2001* (Vol. 2, pp. 1488-1493).

Sivaraman, B., & Burks, T. F. (2006). Geometric performance indices for analysis and synthesis of manipulators for robotic harvesting. *Transactions of the ASABE, 49*(5), 1589-1597.

Sucan, I. A., & Chitta, S. (2014). MoveIt. http://moveit.ros.org. Accessed 3 March 2014.

Tannous, M., Caro, S., & Goldsztejn, A. (2014). Sensitivity analysis of parallel manipulators using an interval linearization method. *Mechanism and Machine Theory, 71*(0), 93-114.

Tuijl, B. A. J., & Wais, E. (2014). *Harvesting Device*. Patent pending, Dutch Patent Office, Netherlands.

Van Henten, E. J., Hemming, J., Van Tuijl, B. A. J., Kornet, J. G., & Bontsema, J. (2003a). Collision-free Motion Planning for a Cucumber Picking Robot. *Biosystems Engineering, 86*(2), 135-144.

Van Henten, E. J., Schenk, E. J., van Willigenburg, L. G., Meuleman, J., & Barreiro, P. (2010). Collision-free inverse kinematics of the redundant seven-link manipulator used in a cucumber picking robot. *Biosystems Engineering, 106*(2), 112-124.

Van Henten, E. J., Van Tuijl, B. A. J., Hemming, J., Kornet, J. G., Bontsema, J., & Van Os, E. A. (2003b). Field Test of an Autonomous Cucumber Picking Robot. *Biosystems Engineering, 86*(3), 305-313.

Van Willigenburg, L. G., Hol, C. W. J., & Van Henten, E. J. (2004). On-line near minimum-time path planning and control of an industrial robot for picking fruits. *Computers and Electronics in Agriculture, 44*(3), 223-237.

Zhang, X., Ba, P., & Mu, L. (2012). Position error sensitivity analysis for polishing robot. *Advanced Materials Research* (Vol. 500, pp. 326-330).

Chapter 7 – Performance evaluation of a harvesting robot for sweet-pepper

C.W. Bac[a,b], J. Hemming[a], B.A.J. van Tuijl[a], R. Barth[a], E. Wais[b], E.J. van Henten[b]

Affiliations:
[a] Wageningen UR Greenhouse Horticulture, Wageningen University and Research Centre,
[b] Farm Technology Group, Wageningen University and Research Centre

This chapter was submitted for publication in *Autonomous Robots*

Abstract

This paper evaluates a robot developed for autonomous harvesting of sweet-peppers in a greenhouse. Objectives were to evaluate the robot for two types of end-effectors (Fin Ray; Lip-type) and crop conditions (unmodified; simplified), and to evaluate the performance contribution of stem-dependent determination of the grasp pose. We describe and discuss performance of hardware and software components developed for fruit harvesting in a complex environment that includes lighting variation, occlusions and densely spaced obstacles. Generally, the robot had difficulty in successfully picking sweet-peppers. Harvest success was 6% (Fin Ray) and 2% (Lip-type). After simplifying the crop, harvest success improved to 26% (Fin Ray) and 33% (Lip-type). Enabling stem-dependent determination of the grasp pose decreased stem damage from 19% to 13% (Fin Ray) and increased grasp success from 41% to 61% (Lip-type). The robot's novel capability of perceiving the stem of a plant may serve as useful functionality for future robots.

7.1 Introduction

Robotic harvesting of sweet-pepper is motivated by several factors, including a need to reduce the costs of fruit production (Lewis et al. 1983; Van Henten 2006). Whereas harvesting robots have been developed for tomato in eight earlier projects (Bac et al. 2014b), only one harvesting robot for sweet-pepper has been developed before, in a project at Kochi University, Japan (Kitamura and Oka 2005; Bachche and Oka 2013). A major reason for this difference is presumably the complexity of sweet-pepper harvesting. In contrast to tomato, leaves are not picked and often cover the fruit, rendering it difficult to detect and localize fruit. Furthermore, fruit sometimes develop in clusters, which makes it hard to localize, access, and detach fruit without damaging. Hence, automating sweet-pepper harvesting provides a scientific challenge where multiple disciplines are involved: hardware design, computer vision, control algorithms and electronics, software architecture and, last but not least, horticulture. The current paper describes the components and performance evaluation of an autonomous robot capable of sweet-pepper harvesting in a greenhouse.

An economic analysis of sweet-pepper harvesting in the Netherlands showed that a cycle time of 6 s should be achieved, if a catalogue price of 100k€ is assumed for a robot that harvests 50% of the sweet-peppers and can be used for about 120 hours per week over a period of 35 weeks per year in which fruit can be harvested (Pekkeriet 2011). We based these numbers on a payback period of three years. Under these assumptions, the market potential would be about 460 robots for the area (1300 ha) of sweet-pepper greenhouses in the Netherlands.

Previous work provides an overview of harvesting robots developed in the last three decades (Bac et al. 2014b). These 50 robots reviewed typically consisted of a movable platform mounting an imaging system, a manipulator and an end-effector. For some of these robots, a comprehensive field test was conducted, including use of performance indicators, failure analysis and a discussion on the testing conditions. Developers of a melon harvesting robot were among the first to report the performance of the robot using several indicators (Edan 1995; Edan et al. 2000). A robot for mushroom harvesting was unique regarding the

large number of 2506 test samples evaluated (Reed et al. 2001). Plebe & Grasso (2001) conducted a field test of orange harvesting and were novel in testing the effect of different weather conditions on performance. A cucumber harvesting robot (Van Henten et al. 2002; Van Henten et al. 2003) was tested and the causes of failed picking attempts were categorized. For strawberry harvesting, the effect of fruit location on harvest success was analysed for the first time (Hayashi et al. 2010).

We attempt to contribute to this literature by pursuing two objectives. Firstly, to provide the first performance evaluation of a harvesting robot for sweet-pepper. The robot was evaluated in two crop environments, thereby contributing to previous evaluations, and for two types of end-effectors developed. Furthermore, four new performance indicators are introduced. Secondly, to evaluate the performance contribution of stem-dependent determination of the grasp pose. The robot is novel in the ability to distinguish and localize the stem of the plant. Using the stem location, we calculate a grasp pose for the end-effector, to reduce damaging plant parts and to increase detachment success.

The current paper may provide tools and performance indicators to develop and test an autonomous robot for complex environments that include lighting variations, occlusions and densely spaced obstacles. The robot's novel capability of perceiving the stem, and calculating a grasp pose, may be reused in future work.

7.2 Crop

The harvesting experiments were conducted in a sweet-pepper crop (*Capsicum annum*; cultivar 'Waltz') grown in a commercial greenhouse in the Netherlands. The greenhouse layout consisted of pairs of crop rows with aisles in-between pairs to move the robot along a crop row. Each crop row consisted of plants positioned on a hanging gutter that was attached to the roof of the greenhouse. Plants were cultivated in the V-system (Jovicich et al. 2004) in which the stems grew along a vertical support wire. Due to plant maintenance and the unstructured growth of a plant, the stem spacing ranged between 10-30 cm and was on average 20 cm. Two heating pipes were present in the aisle and were used as a rail system for the robot, to travel along the plants (Figure 7-1). Crop rows were spaced 113 cm at the aisle side and 47 cm at the backside. The practical width of the aisle was about 80 cm because fruit and leaves attached to the plants further limited the free space.

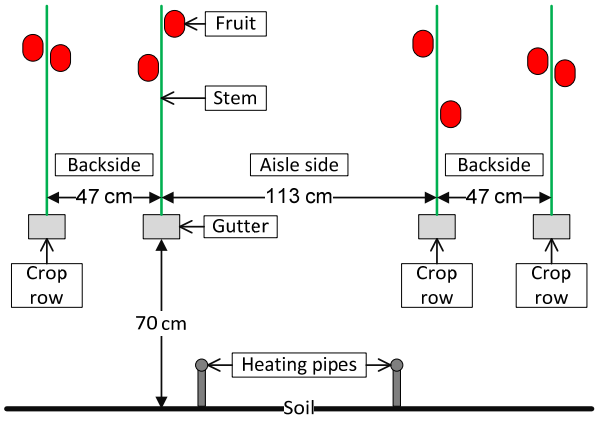


Figure 7-1. Schematic side view of the crop

7.3 Robot components

The robot was developed as part of the research project CROPS 'Clever Robots for Crops' (CROPS 2014). Hardware components, the task sequence and algorithms are described in the following sections.

7.3.1 Hardware

Figure 7-2 provides an overview of the harvesting robot. The platform (Jentjens Machinetechniek BV, The Netherlands) consisted of two connectable modules: a manipulator module and a sensing module.

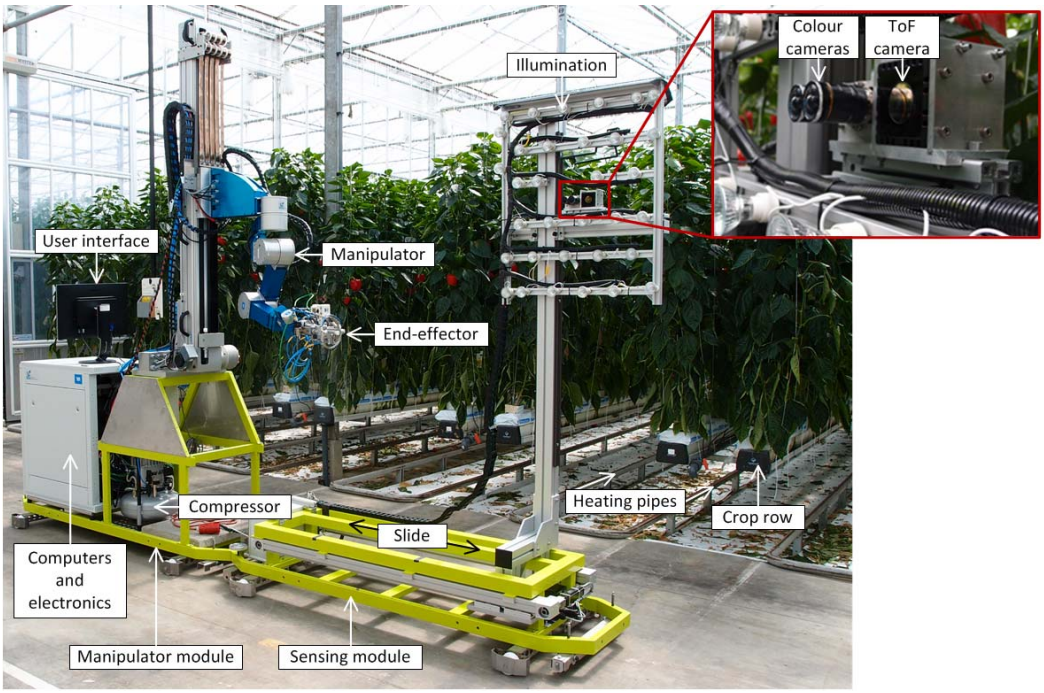


Figure 7-2. Overview of the harvesting robot.

The manipulator module comprised a manipulator, end-effector, air compressor for the pneumatics of the end-effectors, and a cabinet containing computers and electronics. The manipulator (Institute of Applied Mechanics, Technical University Munich, Germany) comprised nine degrees-of-freedom. Additional details on the kinematic structure, motors and electronics are described in earlier work (Schuetz et al. 2014; Bac et al. 2015). Figure 7-3 visualizes the two end-effectors that were developed. The Fin Ray end-effector (Festo AG & Co. KG, Germany) used four fingers that employ the 'Fin Ray principle' to adjust the fingers to the curvature of the fruit (Gauchel and Saller 2012). Scissors were mounted on top of the fingers to cut the peduncle after a fruit was grasped. A position sensor detected if the fruit was grasped or not. The lip-type end-effector (Wageningen UR Greenhouse Horticulture, The Netherlands) used a suction cup to grasp the fruit and a vacuum sensor to detect grasp success. After grasping, the two lips were closed to cut the peduncle. A unique patented feature (Tuijl and Wais 2014) of the end-effector is that both lips move independently from each other. If the peduncle obstructed either one of the two lips, the other lip continued to move. As a result, the two lips met at the peduncle and therefore no information on the pose

(position and orientation) of the peduncle was required for cutting. Furthermore, this feature reduces the likelihood to cut into the fruit or stem. The knife spanned the full length of the upper lip and cut into a slit present in the lower lip. In addition, both end-effectors contained LED lighting and two cameras: a colour camera (VRmMS-12, VRMagic GmbH, Germany) with a resolution of 754x480 pixels, and a mini Time-of-Flight (ToF) camera (CamBoard nano, PMD technologies GmbH, Germany) with a resolution of 160x120 pixels.

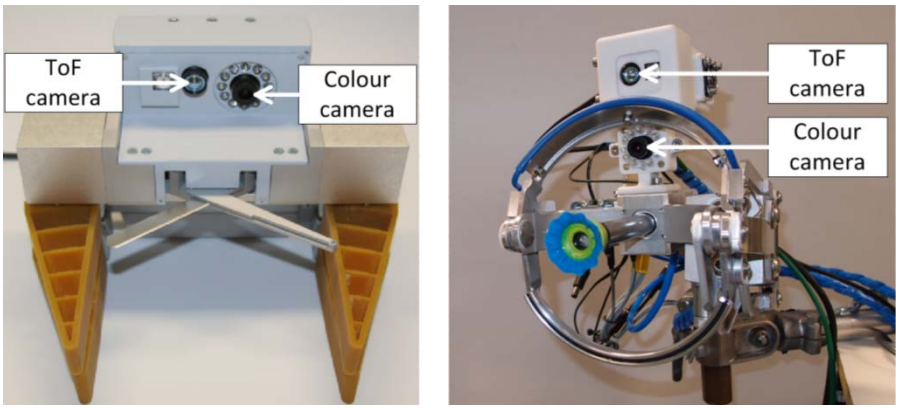


Figure 7-3. Fin Ray end-effector (left) and Lip-type end-effector (right).

The sensing module carried a sensor rig that was actuated by a motor for horizontal movement along a slide. Sensors mounted on this rig included two colour cameras for stereo-vision and a ToF camera (SR4000, Mesa Imaging AG, Switzerland). The colour cameras (Prossilica GC2450C; Allied Vision Technologies GmbH, Germany) had a resolution of 5 megapixel and low-distortion lenses with 5 mm focal length (LM5JC10M; Kowa GmbH, Germany) were mounted on the cameras. The halogen lighting set-up used to illuminate the scene was detailed in earlier work (Bac et al. 2014a).

Overall hardware costs of the robot were about 128k€: 90k€ for the manipulator, 20k€ for sensors and illumination, 10k€ for the end-effector, and 8k€ for the platform. These prices hold for 2012-2014, i.e. the years in which materials were purchased. Yet, if these parts will be taken into production, the costs will likely reduce.

7.3.2 Task sequence

Figure 7-4 displays the task sequence performed by the harvesting robot. The robot was fully autonomous except for moving the robot along the aisle and between aisles. Automating these tasks was outside the scope because commercial robots already perform these tasks, e.g. in spraying robots (Micothon BV, The Netherlands). Therefore, we manually positioned the robot in the aisle. Displacements were done in small steps of 20 cm, to increase fruit localization success by using several viewpoints (Hemming et al. 2014a). At each robot position, we initiated the task sequence to attempt harvesting the ripe sweet-peppers present in the workspace. A fruit was defined as ripe if at least 50% of the fruit surface was coloured.

The task sequence diagram was grouped in three high-level tasks: fruit localization, stem-dependent determination of the grasp pose, and fruit detachment. For each high-level task, we detail its low-level tasks. To localize the ripe fruit, the manipulator first moved aside to allow the sensor rig to slide into the manipulator workspace and acquire RGB and ToF images.

Subsequently, the RGB and ToF images were registered and the algorithm for fruit localization (Section 7.3.3.1) detected the ripe fruit and determined the fruit locations. The second high-level task is novel compared with earlier harvesting robots (Bac et al. 2014b) and involved stem localization to generate a list of best possible grasp poses (Section 7.3.3.2) for each fruit localized. The third high-level task, fruit detachment, was repeated until each localized fruit was attempted for detachment. A task planner ordered the sequence of fruit based on their height coordinate, from high to low. Subsequently, the manipulator moved to the 'central waypoint', in the centre of the workspace, because the motion planner (Section 7.3.3.3) was sometimes unable to find a collision-free path to the mini camera viewpoint, in front of the fruit. This mini camera viewpoint was taken 30 cm in front of the fruit, at the same orientation as the cameras on the sensor module, to increase the likelihood of re-detecting the target fruit. In case the mini cameras were used, images were acquired and processed (Section 7.3.3.1) and, if a fruit was detected and localized, the target position was updated. A comparable implementation was used in a strawberry harvesting robot that took a first image to localize all mature fruit and took a second closer image to refine the target position (Hayashi et al. 2010).

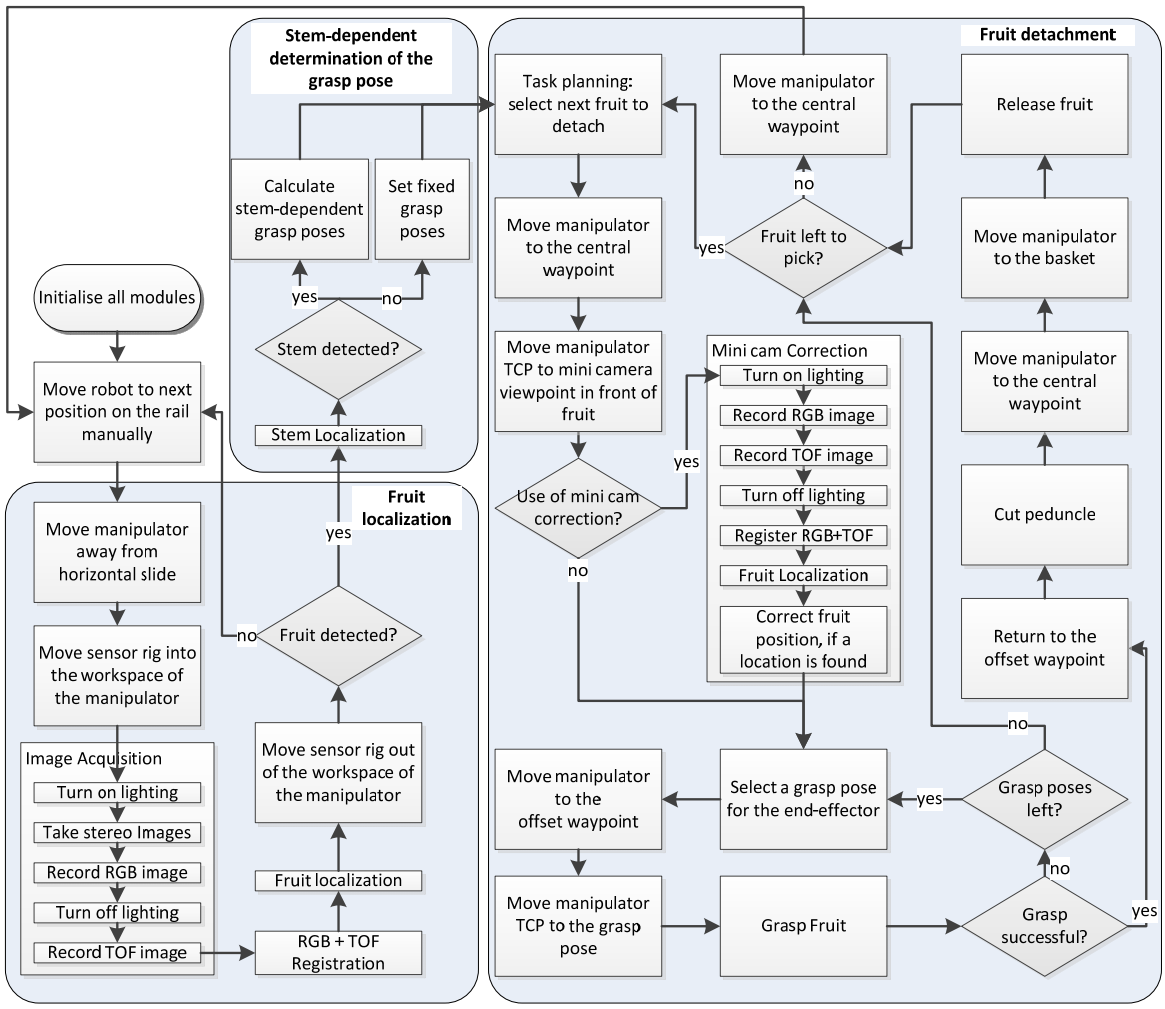


Figure 7-4. Task sequence diagram

Hereafter, the software cycled through the list of grasp poses (Section 7.3.3.2) until either a fruit was grasped, or until no poses were remaining in the list. The manipulator moved to each grasp pose, but stopped at an 'offset waypoint' at 10 cm from the grasp pose and at an identical orientation. A constant orientation of the end-effector in close vicinity to the fruit was required for a successful grasp. When grasping succeeded, the manipulator returned to the offset waypoint to pull the fruit slightly away from the stem, to avoid cutting into the stem. Thereafter, the cutting mechanism was activated. To conclude fruit detachment, the manipulator moved to the basket, via the central waypoint, and then released the fruit.

The tasks described were written in C++ and Python code and implemented using the Groovy distribution of Robot Operating System (ROS) installed on Ubuntu 12.04. ROS serves as a communication layer between software and hardware components of the robot. Our experiences with ROS are discussed by Barth et al. (2014). In addition, we used a generic software framework that included an error manager, a user interface and a finite state machine, to handle the task sequence (Hellström and Ringdahl 2013). The state machine also monitored the execution time of the following sub-tasks: (1) move arm outside sensor rig workspace; (2) move sensors in/out of workspace; (3) fruit and stem localization, and calculation of grasp poses; (4) select next fruit to detach; (5) move to the central waypoint; (6) move to mini camera viewpoint and perform mini camera correction; (7) perform grasp attempt, including grasp pose selection, moving to the pose via the offset waypoint, grasping the fruit, and cutting the fruit in case of a successful grasp; (8) move to the central waypoint; (9) move to basket.

7.3.3 Algorithms

7.3.3.1 Fruit localization and ripeness determination

The algorithm was used on images recorded by the sensor rig and by mini cameras on the end-effector. The colour images were used for fruit detection and ripeness determination, and ToF images were used for fruit localization. Three consecutive ToF images were averaged to compensate for random noise and improve the precision of the 3D measurements. The colour image was pixel-wise registered to the ToF image, using the Direct Linear Transformation (DLT) algorithm (Shapiro 1978). Fruit detection basically involved a red blob detection. To compensate for varying lighting conditions, the Normalized Difference Index (NDI) was calculated for the green G and red R channel of the colour image (Equation (7-1)).

$$NDI = \frac{G + R}{R - G} \quad (-) \qquad (7\text{-}1)$$

Hereafter, a threshold of $NDI>0$ was applied and pixel components were connected and filled to obtain blobs. Blobs were considered as ripe fruit if the area was greater than 10000 pixels (sensor rig images) or 20000 pixels (mini camera images). Both thresholds were empirically determined. If several blobs were detected in the mini camera images, the biggest blob was selected as the target fruit. Subsequently, three iterations of an erosion operator (circular structuring element with a radius of 14.5 pixel) were applied to the binary image, to compensate for inaccuracies in image registration and for noisy ToF data. Lastly, the corresponding (x,y,z) coordinates of pixels in this inner blob were averaged, to represent the fruit location.

7.3.3.2 Stem-dependent determination of the grasp pose

Use of a stem-dependent grasp pose was implemented to reduce collisions with neighbouring fruit, the stem, or other plant parts when approaching the target fruit. The algorithm involved stem localization and, together with the fruit location, a grasp pose was determined for the end-effector. Figure 7-5 illustrates this process.

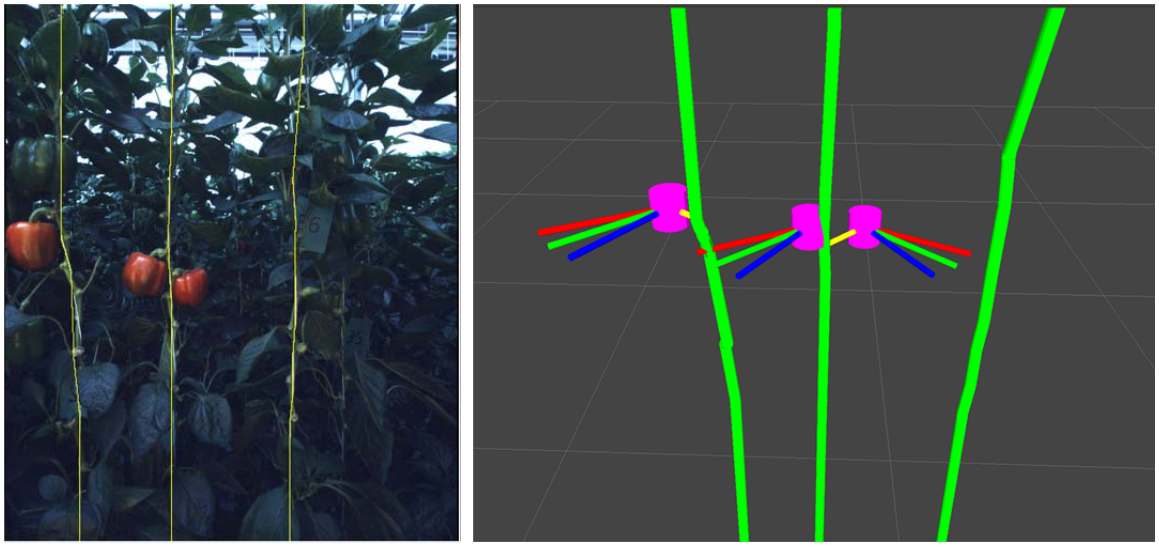


Figure 7-5. Stems were localized (yellow lines) in an image where also three fruit were present (left). The right figure displays a mapping of the stems (green) and fruit (pink) that were used to match each fruit to its 'parent' stem (indicated by yellow connecting cylinders), and to calculate the three best grasp poses (indicated by a red, green and blue cylinder) for each fruit.

Stems were localized using an algorithm described in previous work (Bac et al. 2014a) that employed stereo-vision. The previous article reported the accuracy, precision and detection performance, where stem localization was found inaccurate (±4.5 cm) because of sequential image acquisition. We resolved this issue by simultaneous image acquisition using two cameras. As a result, the localization accuracy measured was ±4 mm in the depth direction and ±0.5 mm in width and height.

The grasp pose was calculated in the horizontal plane, i.e. the algorithm only calculated the azimuth angle of the end-effector with respect to the fruit centre. Calculating the elevation and skew angle of the grasp pose would have been beneficial for a more successful grasp and cut, but requires the orientation of the fruit about its centre. We did not succeed developing an algorithm able to reliably estimate the fruit orientation. To calculate the azimuth angle, the following steps were implemented that complement earlier work (Bac et al. 2015).

1. The fruit was matched with its 'parent' stem by taking the shortest Euclidean distance between detected stem segments and the target fruit (Figure 7-5). Maximum Euclidean distance was set to 20 cm to avoid a match with a neighbouring stem in case the parent stem was not detected. Since the Euclidean distance was calculated in the horizontal plane, the height of the fruit location and each stem segment was aligned using the procedure described in Section 4.2 of previous work (Bac et al. 2014a).
2. The azimuth angle of the fruit with respect to the stem was calculated.
3. The azimuth angle of the end-effector with respect to the fruit was calculated for nine pre-selected azimuth angles in a range of 120° at the aisle side of the fruit: [-60, -45, -

30, -15, 0, 15, 30, 45, 60]°. This step corresponds with the 'constrained-angle' method described in Bac et al. (2015).

4. Out of these nine pre-selected angles, the algorithm selected the three angles resulting in the smallest deviation between the azimuth angle of the fruit and of the end-effector. As a result, the algorithm generated an ordered list of three grasp poses that corresponded with the three best azimuth angles (Figure 7-5). In case a parent stem was not detected, three fixed grasp poses were taken: 0°, -15°, 15°.
5. The positions of the grasp poses were converted from the camera coordinate frame into the coordinate frame of the manipulator. In preliminary tests, we manually adjusted the transformation matrix between these coordinate frames, until a satisfactory hand-eye calibration was realized.

7.3.3.3 Motion planning

The motion planner calculated a straight path through the workspace, i.e. point-to-point planning. To realize this motion, inverse kinematics were solved using an algorithm capable of handling the redundant degrees-of-freedom (Baur et al. 2012). Furthermore, the motion was checked for self-collisions (Baur et al. 2013). To compensate for the difficulty to plan a straight path through the workspace, the central waypoint was implemented (Section 7.3.2).

7.4 Methods and Experiments

The robot was tested in the greenhouse in June 2014. We followed a protocol that included measuring ground truth (Section 7.4.1), simplifying the crop (Section 7.4.2), recording of performance indicators (Section 7.4.3) and categorizing failures (Section 7.4.4). These recordings were used to conduct two experiments: to determine the overall performance of the robot (Section 7.4.5) and to determine the performance contribution of stem-dependent determination of the grasp pose (Section 7.4.6).

7.4.1 Ground truth measurements

The first step in measuring ground truth was to consecutively number plant stems using green labels (Figure 7-6). At each of these stems we counted the ripe fruit present within the height range reachable by the manipulator. The following parameters were measured for ripe fruit (measurement resolution is indicated in parenthesis):

- Height of the fruit centre relative to the robot platform (1 cm), to correctly correspond a detected fruit with its ground truth;
- Azimuth angle with respect to the stem (10 degrees), to serve as an explanatory parameter in case a fruit was not localized or not successfully detached;
- Fruit is part of a cluster of fruit (true/false), to serve as an explanatory parameter if access to a peduncle was obstructed;
- Occlusion (10%) of the fruit when viewed orthogonal to the crop row and at the same fruit height, to serve as an explanatory parameter in case a fruit was not localized, or in case the end-effector was misplaced.

7.4.2 Crop simplifications

Presence of fruit clusters and leaves might negatively influence the accuracy of fruit locations and the free space to access a fruit. Therefore, we compared performance for an unmodified crop and for a simplified crop, in which we removed fruit clusters and leaves. Figure 7-6

demonstrates how removal of leaves and fruit clusters reduced occlusion and improved accessibility of fruit.

Figure 7-6. Crop before (left) and after (right) removal of leaves and the fruit cluster present within the white circle (left). Green labels indicate the stem number.

7.4.3 Performance indicators

In addition to performance indicators discussed in a review article on harvesting robots (Bac et al. 2014b), four new indicators were added and explained: grasp success, cut success, stem damage rate and leaf damage rate.

- Fruit localization success (%): The number of localized ripe fruit per total number of ripe fruit in the canopy.
- False-positive fruit detection (%): The number of objects falsely detected as fruit per total number of ripe fruit in the canopy.
- Grasp success (%): The number of ripe fruit successfully grasped per total number of localized ripe fruit. This indicator was added to assess the grasping mechanism of the end-effector.
- Cut success (%): The number of fruit successfully cut per total number of fruit successfully grasped. This indicator was added to assess the cutting mechanism of the end-effector.
- Detachment success (%): The number of successfully harvested ripe fruit per total number of localized ripe fruit. This rate corresponds with a multiplication of grasp success and cut success.
- Harvest success (%): The number of successfully harvested ripe fruit per total number of ripe fruit in the canopy.
- Cycle time (s): time of an average full harvest operation, including ripeness determination, localization, fruit detachment, transport of a detached fruit, but without manual displacement to the next robot position.
- Fruit damage rate (%): the number of damaged fruit per total number of localized ripe fruit. A fruit was considered as damaged if the robot cut into the fruit surface or

caused a bruise. Measuring fruit damage is highly relevant for economic feasibility because a grower cannot market a damaged fruit.

- Stem damage rate (%): the number of damaged stems per total number of labelled stems. A stem was considered damaged if the robot cut more than 0.5 cm into the stem, or if the robot pulled a stem apart during the harvest process. Stem damages are the most serious damages that can occur because the production of fruit will strongly reduce or even stop.
- Leaf damage rate (-): the number of damaged leaves per total number of labelled stems. A leaf was considered damaged if the robot cut more than 2 cm into the leaf surface, or if the robot detached a leaf during the harvest process. Leaf damages are the least serious damages that can occur because the production of fruit is generally hardly affected if a few leaves of the plant are damaged.
- Average platform attempts (-): Average number of platform positions used per successfully detached ripe fruit. This indicator provides an additional measure for detachment success.
- Average grasp attempts (-): Average number of grasp poses used per successfully detached ripe fruit. This additional indicator measures detachment success.
- Average positioning error (cm): Average positioning error of the end-effector tool centre point (TCP) with respect to the fruit centre. This error can originate from inaccurate fruit localization and imperfect positioning of the joints of the manipulator. Figure 7-7 displays how the error was measured.

Figure 7-7. Measuring the positioning error of the end-effector in depth (left), width (centre) and height (right). White arrows indicate the error measured with respect to the cut point defined between the scissors.

7.4.4 Failure categories

To obtain an understanding of the most critical steps in the harvest procedure, we categorized failures and detailed possible causes of these failures:

1. Incorrect fruit localization, due to partial detection of the fruit surface or a merged detection of two fruit;
2. Incorrect grasp pose due to an undetected or inaccurately localized stem or fruit;
3. Misplaced end-effector due to: inaccurate image registration, inaccurate camera calibration, inaccurate hand-eye calibration, or an inaccurate manipulator;
4. Plant part obstruction to the target fruit by fruit clusters, stems or undetected objects;
5. Manipulator error due to inability to calculate a motion;
6. Grasp failure of the fruit due to misplacement, or weak grasp or suction force;
7. Cut failure due to a partial or missed cut of the peduncle;

8. Fruit transportation failure after detachment, due to collisions with other objects, or due to an unintended release of the fruit;
9. Difficult fruit to harvest due to a fruit location at the backside of the stem or a non-regular fruit orientation, i.e. perpendicular to the stem (Figure 7-8).

Figure 7-8. Example of a non-regular orientation of a fruit, i.e. perpendicular to the stem.

There were cause-effect relationships between these failure categories. For instance, a misplaced end-effector often caused a grasp or cut failure. And, a plant part obstruction also caused a grasp failure.

One or several failure categories were assigned to each unsuccessful picking attempt. If several grasp poses were attempted at a robot position, we once assigned failures because preliminary tests revealed that causes of failure were mostly the same for each pose.

7.4.5 Experiment 1: robot performance

This experiment tested the performance of the robot using performance indicators reported in Section 7.4.3. In case of unsuccessful fruit detachment, we assigned failure categories (Section 7.4.4). The experiments were done under four different testing conditions: the two end-effectors were tested in the unmodified crop and in the simplified crop. These conditions enabled to compare the two end-effectors and to determine the influence of occlusions and presence of fruit clusters on performance. Altogether a total of 176 labelled fruit were attempted. Table 7-1 provides an overview of testing conditions, including weather conditions. The mean and sample standard deviation (SD) of the solar irradiance was measured during tests, to serve as an explanatory factor for performance differences between tests.

Table 7-1. Overview of tests conducted in the greenhouse

End-effector	Crop conditions	Dates of testing	Fruit labelled	Stems labelled	Weather conditions	Mean (SD) irradiance during tests, in W/m^2
Fin Ray	Unmodified	13 and 16 June	47	46	Moderate and intense sunlight	581(113)
Fin Ray	Simplified	27 and 30 June	47	69	Cloudy	489(219)
Lip-type	Unmodified	20 and 26 June	43	78	Moderate and intense sunlight	600(260)
Lip-type	Simplified	24 and 25 June	39	69	Moderate and intense sunlight	569(174)

The execution time was computed per sub-task and for several fruit, to determine the time variation between fruit and to determine the most time-consuming sub-tasks. A mean and SD was calculated for each sub-task and for the sum of sub-tasks, i.e. the total cycle time.

In a preliminary experiment, we tested the contribution of the mini cameras to a more accurate fruit location. In many cases, the mini-ToF camera did not improve the location estimates of the fruit compared with the cameras on the sensor rig. Noisy data may have been one of the causes. Therefore, the mini cameras were turned off during the experiments.

7.4.6 Experiment 2: stem-dependent determination of the grasp pose

This experiment measured the effect of stem-dependent determination of the grasp pose on performance, to indicate the contribution of this novel functionality. In fact, performance was compared for two modes that were sequentially tested: disabling and enabling stem localization. In the disabled mode, the algorithm in Section 7.3.3.2 computed three fixed grasp poses with an azimuth angle of 0, -15, and 15 degrees. At each robot position, we first grasped the fruit using the three fixed grasp poses. Thereafter, we enabled stem localization and cutting and reinitiated the task sequence. The advantage of this approach was that the same subjects (fruit) were investigated from an identical robot position and within an identical crop environment. The effect of sunlight variation was minimized because the time lag between both modes was less than three minutes. Performance indicators compared were grasp success and stem damage. Since cutting was turned off in the disabled mode, we observed the scene and estimated the stem damage. If a stem was present in-between the scissors or lips of the end-effector, stem damage was reported (Figure 7-9).

Figure 7-9. Photo of a grasp pose calculated for a target fruit when stem detection was disabled. Upon cutting, the mechanism would damage the stem present between the scissors.

The effects of enabling the mode on stem damage rate and grasp success rate were tested for significance, using a paired-samples McNemar test (SPSS Statistics, version 19). Such test fits an experimental design that meets three assumptions (Laerd 2014). Firstly, the dependent variable should be categorical with two categories, i.e. failure and success for this experiment. And, the independent variable should consist of two related groups, i.e. the disabled and enabled mode. Secondly, the two groups of the dependent variable should be mutually exclusive, i.e. a fruit cannot be successfully and unsuccessfully grasped at the same time.

Thirdly, the subjects should be a random sample from the population. Although the first two assumptions were met, the last assumption was partially met. We tested the robot on consecutive plants in a crop row, but we did not investigate if these samples represented the crop conditions for the entire greenhouse.

7.5 Results

Results of the two experiments are described in the following sections. The supplementary material encloses two videos of a successful harvest cycle: for the Fin Ray and for the Lip-type end-effector.

7.5.1 Experiment 1: robot performance

Figure 7-10 provides the success rates and damage rates of tests conducted. We explain results of the unmodified and of the simplified crop. Note that the type of end-effector is unrelated to fruit localization success.

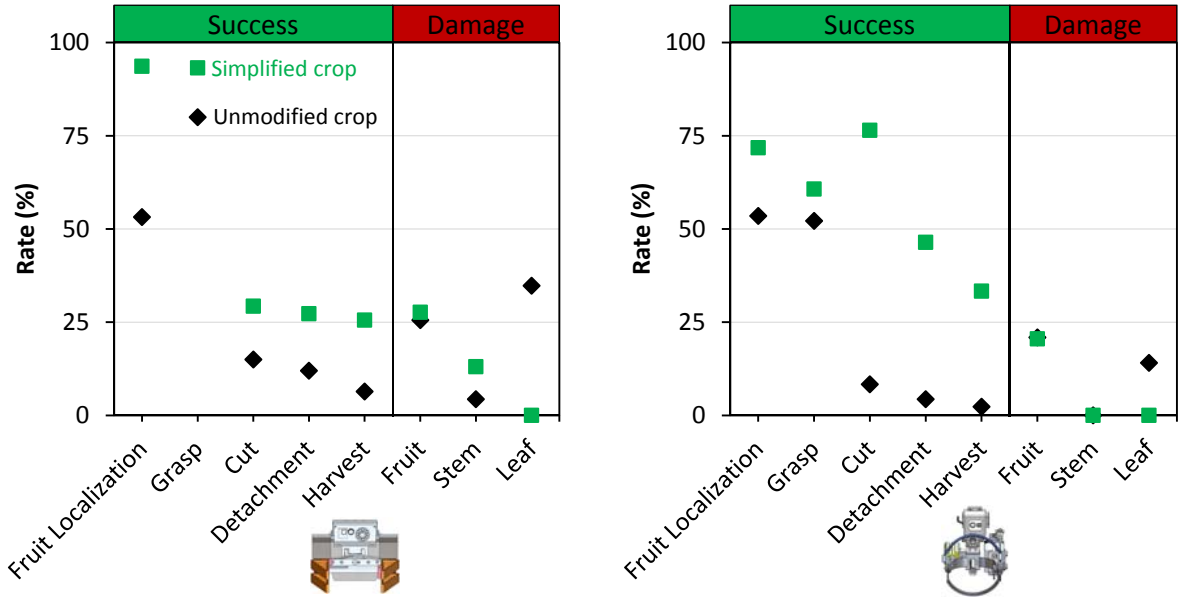


Figure 7-10. Success rates and damage rates of tests conducted using the Fin Ray end-effector (left) in the unmodified crop (◆) and simplified crop (■). Similarly, a test was conducted with the Lip-type end-effector (right) mounted on the robot.

In the unmodified crop, the Fin Ray end-effector achieved a higher grasp success (80%) than the Lip-type end-effector (52%). Cut success was poor (<15%) for both end-effectors and therefore detachment success and harvest success were also poor. Fruit damage rate of the Fin Ray end-effector (26%) was about similar to the Lip-type end-effector (21%). However, the Fin Ray end-effector caused stem damage (4%), whereas the Lip-type end-effector did not cause any stem damage. In addition, the Fin Ray end-effector caused more leaf damage (35%) than the Lip-type end-effector.

After simplifying the crop, all success rates increased. The most remarkable finding is that cut success of the Lip-type end-effector increased by a factor of eight, whereas a factor of two for the Fin Ray end-effector. Therefore, the performance of the Lip-type end-effector

increased in terms of detachment success (from 4% to 46%) and harvest success (from 2% to 33%). The leap in performance was smaller for the Fin Ray end-effector: detachment success increased from 12 to 27% and harvest success from 6% to 26%. Simplifying the crop did not result in lower damage rates. Fruit damage remained similar, whereas the stem damage for Fin Ray end-effector even increased from 4 to 13%. Leaf damage was 0%, obviously, due to absence of the leaves.

Other performance indicators are listed in Table 7-2. In the simplified crop, more fruit were visible and therefore more fruit were detected with more false detections. The simpler conditions also led to fewer platform positions used for successfully detached fruit. Yet, attempting fruit from different platform positions was clearly useful because on average 1.4 to 2.0 positions were used. Additional grasp attempts, however, were hardly useful because on average 1.0 to 1.2 attempts were used, out of the three attempts programmed in the task sequence (Section 7.3.2). We observed that positioning errors were greater if the fruit did not appear in the centre of the image. The closer a fruit appeared to the edge of the image, the greater the positioning error was.

Table 7-2. Performance indicators other than success rates and damage rates.

End-effector	Fin Ray 		Lip-type 	
Crop conditions	Unmodified	Simplified	Unmodified	Simplified
False-positive fruit detection (%)	9	17	7	13
Number of successfully detached fruit (#)	3	12	1	13
Average platform attempts (-)	2	1.4	2	1.4
Average grasp attempts (-)	1	1.2	1	1.2
Number of positioning errors observed (#)	37	27	35	16
Average positioning error (cm) in depth	1.2	0.4	2.5	1.5
Average positioning error (cm) in width	2.2	1.2	3.5	1.4
Average positioning error (cm) in height	1.4	2.8	1.8	2.2

Table 7-3 provides execution times per sub-task. The total time was 106 seconds, if the fruit was detached within one grasp pose, and if one fruit was detected in the scene. If multiple fruits were detected, the cycle time per fruit was shorter because sub-tasks 1-3 were executed only once. Hence, the cycle time was 89 s if two fruit were detected, and 84 s if three fruit were detected. On average, 1.6 fruit were detected in a scene, which corresponds to a cycle time of 94 s. If the first grasp attempt failed, the cycle time increased with 22 s per additional grasp attempt. The variation in grasp poses selected for fruit, explains the time variation (2.4 s) in sub-task 7.

Table 7-3. Mean (N=16) and SD of the execution time per sub-task, in case one fruit was detected

Sub-task	Mean (SD) execution time in seconds
1. Move manipulator away from the horizontal slide	10.7 (0.5)
2. Move sensors into and out of workspace	9.2 (0.5)
3. Fruit and stem localization, and calculation of grasp poses	12.9 (0.5)
4 Select next fruit to detach	0.0 (0.0)
5 Move to the central waypoint	10.5 (0.6)
6 Move to mini camera viewpoint and perform mini camera correction	11.9 (0.0)
7 Perform grasp attempt	32.2 (2.4)
8 Move to the central waypoint	10.2 (0.4)
9 Move to basket	8.2 (0.0)
Total time	105.8 (3.2)

Table 7-4. Result of failure analysis for four different testing conditions

End-effector	Fin Ray 		Lip-type 	
Crop conditions	Unmodified	Simplified	Unmodified	Simplified
Unsuccessful attempts	52	54	50	38
Total failures	117	136	144	82
Average number of failures per unsuccessful attempt	2.3	2.5	2.9	2.2
Contribution (%) of each failure category to total failures				
1. Incorrect fruit localization	19	18	15	9
2. Incorrect grasp pose	2	3	4	2
3. Misplaced end-effector	32	20	24	20
4. Plant part obstruction	3	7	13	12
5. Manipulator error	3	3	1	6
6. Grasp failure	4	10	24	27
7. Cut failure	31	26	10	15
8. Fruit transportation failure	0	0	0	0
9. Difficult fruit	7	13	9	10

Table 7-4 provides results of a failure analysis on unsuccessful attempts. The four testing conditions reveal a similar number of failures per unsuccessful attempt. When further analysing the underlying failure categories, one can observe frequent misplacement of the end-effector (20-32% of the failures) that is partially explained by category 1: incorrect fruit localization. Incorrect determination of the grasp pose (category 2) mostly occurred during intense sunlight. In case category 3 occurred, i.e. plant parts obstructing access to a target fruit, grasping failed. Grasp failures also occurred when difficult fruit were approached (category 9). If a difficult fruit was grasped, a cut failure was often observed. Grasp failures were more often observed for the Lip-type end-effector (24%) than the Fin Ray end-effector (4%). However, cut failures were more often observed for the Fin Ray end-effector (31%) than for the Lip-type end-effector (10%). Fruit transportation failures were never observed.

To demonstrate these failures, Figure 7-11 displays typical examples of incorrect fruit localization and of a cut failure, when the lip-type end-effector was used. The knife not only failed to fully cut the peduncle, but also cut into a neighbouring fruit. Incorrect grasp poses are further discussed in the next section.

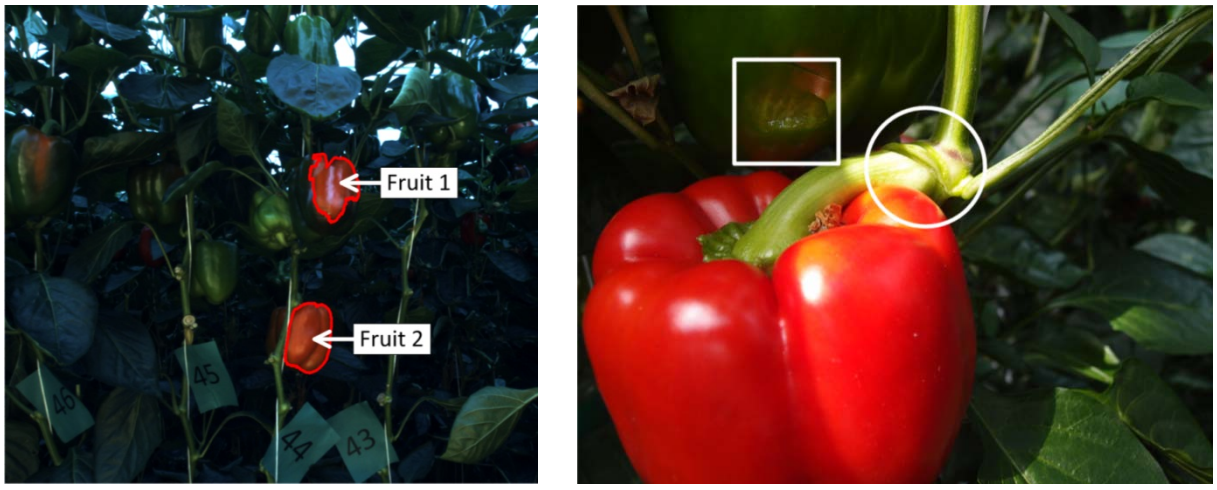


Figure 7-11. Typical example of erroneous fruit localization (left). Fruit 1's centre is erroneously localized in the centre of the coloured surface, instead of the complete fruit surface. A stem occludes Fruit 2 and causes incorrect fruit localization at the right side of the fruit. The right picture displays the result after the lip-type end-effector cut into a neighbouring fruit in a cluster (indicated by a square) and failed to cut through the peduncle (indicated by a circle).

7.5.2 Experiment 2: stem-dependent determination of the grasp pose

Figure 7-12 illustrates how stem-dependent determination of the grasp pose contributes to a more successful cut. These type of situations typically occurred for fruit located at the sides or backside of the stem. For fruit located at the front side, both modes of grasp pose calculation resulted in about the same pose.

Figure 7-12. Comparison of a grasp pose calculated when stem localization was disabled (left) and enabled (right). Due to the non-ideal pose in the left figure, the end-effector pushed against the stem and did not grasp the fruit. In the right figure, the end-effector successfully grasped and cut the fruit.

Results in Figure 7-13 demonstrate the effect of stem-dependent determination of the grasp pose on grasp success and stem damage. For the Fin Ray end-effector, no notable change in grasp success was observed after enabling. However, stem damage decreased from 7% to 4% (out of 46 stems) in the unmodified crop and from 19% to 13% (69 stems) in the simplified crop. For the Lip-type end-effector, stem damages never occurred. Grasp success increased from 48% to 52% (22 fruit) in the unmodified crop and from 41% to 61% (28 fruit) in the simplified crop. Hence, the benefits of perceiving the plant stem were: less stem damages when the Fin Ray end-effector was used, and a higher grasp success when the Lip-type end-effector was used. Yet, due to the low number of samples, the effects were not significant (p>0.05).

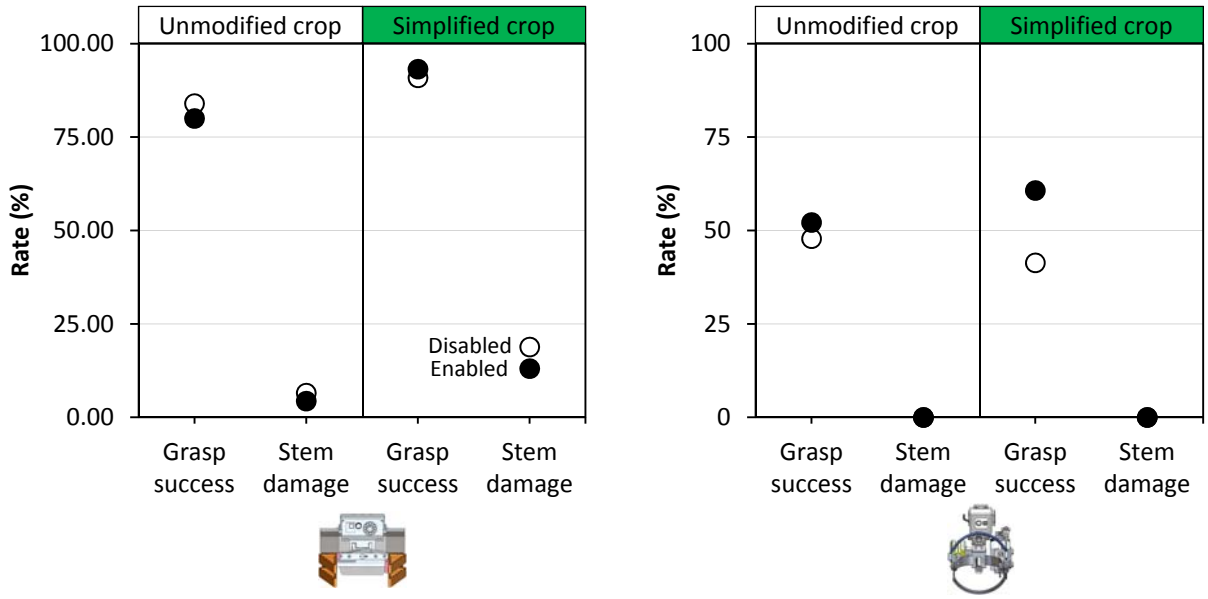


Figure 7-13. Grasp success rate and stem damage rate when stem-dependent determination of the grasp pose was disabled (○) or enabled (●). Results are shown for the Fin Ray (left) and Lip-type (right) end-effector tested in an unmodified and simplified crop.

7.6 Discussion

The harvest success rate of 6% (Fin Ray) and 2% (Lip-type) shows that the robot had great difficulty in successfully picking sweet-peppers. In the following sections we discuss lessons learned from the experiments by reflecting on the crop, sensing, calibration, manipulation, cycle times and stem-dependent determination of the grasp pose.

7.6.1 Crop

Presence of fruit clusters and occluding leaves reduced harvest success by 19% (Fin Ray) and 31% (Lip-type), and were a major reason for the poor harvest success. In addition, fruit located at the backside of the stem were generally not harvested, even in the simplified crop. Unlike sweet-pepper, in crops like tomato or cucumber the distance between fruit and main stem is commonly wider, and occluding leaves are picked as part of plant maintenance. These differences may partly explain the higher harvest success reported in related work: 50% for tomato (Kondo et al. 2010), 74% for cucumber (Van Henten et al. 2003), and 41% for strawberry (Hayashi et al. 2010).

The observation that fruit located at the backside were difficult to harvest, raises two possible interpretations. On the one hand, one might assert the robot did not fit the requirements because it was hardly able to harvest backside fruit. Yet, on the other hand, one might assert the cultivation system and configuration of aisles and crop rows was not suitable for this robot. For instance, if an aisle would be present at both sides of the crop, fruit could be approached from both sides of a crop row, thereby eliminating the need to pick backside fruit. Hence, a trade-off exists between adapting the robot or the crop.

The cultivar of the crop was 'Waltz'. In previous work, a fruit visibility analysis was conducted for cultivar 'Stayer' that produced yellow sweet-pepper fruit (Hemming et al. 2014a). Although 'Waltz' contained many fruit clusters, 'Stayer' seemed to produce even more fruit clusters. This comparison shows that presence of fruit clusters can vary among cultivars. Similarly, peduncle length varies among crops and can influence the cut success. A task for future work is, therefore, to find or breed cultivars that are suitable for robotic harvesting.

7.6.2 Fruit localization

After simplifying the crop, fruit localization success increased more for the Fin Ray end-effector (40%) than for the Lip-type end-effector (18%). This unexpected difference may be partially due to a lower irradiance (489 vs. 569 W/m^2). In addition, differing crop conditions may have caused this effect.

Estimating the fruit orientation failed in this research, due to noisy depth estimates and occlusions. Yet, such information would be useful to better grasp non-regularly oriented fruit. To the best of our knowledge, literature on estimating the fruit orientation is yet to appear, in contrast to other applications such as for head pose estimation of humans (Murphy-Chutorian and Trivedi 2009).

Although the mini cameras were not used in this research, their position corrections could have been a valuable asset to compensate for misplacement of the end-effector. To effectively use the cameras in future research, the issue of poor depth estimates should be addressed.

7.6.3 ToF camera

Depth estimates of the ToF were typically not accurate and we discuss possible causes. The specifications report that the camera should be used under indoor illumination and that the depth measurements were calibrated for a range of 0.8-5 m. Since the camera was used under sunlight and in a depth range of 0.6-0.9, the potential depth accuracy of +/- 1 cm may not have been realized. Due to the difficulty to measure time intervals in the order of several picoseconds, the depth accuracy of time-of-flight cameras might remain limited in future.

The precision is another factor to consider, specifications report 6-9 mm at a working distance of two meter. Such precision is worse than the precision measured for stereo-vision (3 mm) (Bac et al. 2014a). Although we attempted to improve the precision by averaging three images, we did not investigate effectiveness of this approach.

Another explanation for poor depth estimates is the coarse resolution of 176 x 144 pixels. Pixels may have covered plant parts with a different depth. For instance, the light may have partially reflected on a leaf in front of a fruit and on the fruit itself.

In addition, we observed greater errors under intense sunlight. On the one hand, poor fruit detection in the colour image may explain these errors. On the other hand, sunlight may have interfered with the infrared light projected by the camera. A more intense projector may overcome this issue.

7.6.4 Calibration

Results show positioning errors in the range of 0.5-3.5 cm. A difficulty in tracing back the source of errors was that three sources of error can cause misplacement: motion error, sensing error and an imperfect map of the environment (Kurniawati et al. 2012). We observed a sensing error related to rectification of the image. In case the fruit was further away from the centre of the image, large errors were observed in width and height. These errors consistently occurred and may have been caused by improper image rectification of the ToF image. The calibration procedure followed, to rectify lens distortions, should be further investigated to resolve this issue.

For registration of the colour image and ToF image, an h-matrix was calculated to project both images. Values in this matrix were fine-tuned for a specific depth, rendering inaccuracies at other depths. Hence, a drawback of this type of registration is the limited depth range in which registration is still accurate.

We were unsure if any positioning error occurred. The precision of the manipulator was briefly tested (about 1-2 mm) and did not seem to have caused a positioning error. However, the accuracy of the manipulator was not tested. The accuracy may have varied, depending on the target position in the workspace. For instance, in an outstretched configuration, we observed more vibrations of the end-effector. Inaccuracies in the manipulator may have played a role, and a detailed analysis should be conducted to reveal its influence on misplacement of the end-effector.

We manually calibrated the coordinate frames of the sensors and manipulator by iteratively adjusting for offsets. The accuracy of this calibration method was possibly not optimal and, therefore, a proper hand-eye calibration procedure should be used in future work. Motion of the sensor rig along the slide may have introduced additional inaccuracies between

recordings, despite the stable and controlled motion. Using a stationary camera would resolve this effect.

7.6.5 Manipulator and end-effectors

The robot design influenced the ability to harvest certain fruit. The manipulator and end-effector were too big to penetrate the crop and successfully grasp a fruit at the backside of the plant. Earlier work showed that a smaller end-effector helped to access fruit better (Bac et al. 2015).

A number of issues explain unsuccessful grasp and cut attempts of the Lip-type end-effector. The suction force was generally too low and caused many of the failed grasp attempts. We attempted to increase the force by replacing the vacuum generator and shortening the plastic tubes. Regardless of these improvements, more suction force is desired in future work. Similarly, the cut force was too low, explaining why some peduncles were not fully cut. However, stem damages did not occur because the fruit fell from the suction cup, if a lip hit the stem upon closing. Fruit present below the target fruit sometimes blocked motion of the lower lip, and resulted in an unsuccessful cut, whereas the Fin Ray end-effector only cut at the top of a fruit and did not suffer from this issue.

To avoid a cut in the fruit or to cut the peduncle properly, positioning of the Fin Ray end-effector in height was more critical than of the Lip-type end-effector. Positioning in depth was also critical. A positioning too far, in depth, led to a cut in the plant stem, whereas a positioning too close led to a cut failure. Although the scissor was actuated pneumatically at a fixed force, leverage of this mechanism caused the cutting force to depend on positioning of the peduncle along the scissor. Hence, either the design should change or a force-varying actuation principle should be selected. Furthermore, the knife sometimes bended, especially after a cut in a tough object such as a stem. The most important drawback of the Fin Ray end-effector, compared with the Lip-type, was the severe stem damage of 4-13%. Consequently, the grasping mechanism of the Fin Ray end-effector and the cutting mechanism of the Lip-type end-effector are most promising for future development of an end-effector (Hemming et al. 2014b).

7.6.6 Cycle time

The cycle time achieved (94 s) was a factor of 16 too long compared with the economically feasible time of 6 s (Pekkeriet 2011). Motion planning and execution were especially time-consuming because of a low speed and because of the waypoints required. Using a stationary camera would eliminate the time required to move the platform into (4.6 s) and out of the workspace (4.6 s). Possibly, the manipulator motion required to provide space for the horizontal slide (10.7 s) could be eliminated or reduced as well. The time required to move to the basket (8.2 s) could be reduced if the fruit could be disposed in a pipe attached to the end-effector, as was done in kiwi harvesting (Scarfe et al. 2009). Furthermore, both stereo-vision and colour-ToF registration were used to estimate depth. Both methods took several seconds, and in future perhaps only one method for depth estimation could be employed. The long execution time of a grasp attempt (32.2 s) was caused by the waypoints and by a low speed, whereas the motor controllers and mechanics allowed for a higher speed. Hence, execution

time could be reduced by a faster motion and by deploying a motion planner that plans in the configuration space (Kuffner and Lavalle 2000), to avoid using waypoints. Recommended changes mentioned in previous sections (crop, software, robot design) may also further reduce the cycle time.

7.6.7 Stem-dependent determination of the grasp pose

Results indicated that use of stem information reduced stem damage for the Fin Ray end-effector and increased the grasp success for the Lip-type end-effector. Results were based on few samples, which caused the statistical test to prove no significant differences in performance. However, if the increase in grasp success (41% to 61%) would have been based on double the number of samples (52 fruit), the effect would be strongly significant (p=0.007). Although we cannot guarantee significance in case of a repetition on 52 fruit, significance is imaginable because the experimental design of paired samples reduces the effect of confounding variables (Kinney 2002) present in the crop environment.

A future repetition should attempt to fully meet the requirements of the McNemar test. That is, collecting more samples at random places in the greenhouse, and alternating the disabled and enabled mode. In fact, much more than 52 fruit should be investigated to better represent the variation present of all fruit (>10,000) in the greenhouse. To proof significance under the varying crop conditions along a season (Bac et al. 2014b), future research should evaluate the robot several times along the season, to generalize the performance. Clearly, such testing would take a considerable effort.

Although stem perception enhanced the robot's performance, compared with only fruit localization, the effects might have been greater if stems were always correctly detected and localized. About 3% of the failures were caused by incorrectly calculated grasp poses, due to poor stem localization. Possible improvements are discussed in previous work (Bac et al. 2014a). In addition, perception of additional plant parts could be beneficial for higher success rates. For instance, pose estimation of leaves could be used to adjust the end-effector pose until an occlusion-free view on the fruit is realized. As a result, fruit localization will become more accurate.

7.7 Conclusion

This paper describes the first performance evaluation of a sweet-pepper harvesting robot tested in a greenhouse. The robot developed was able to harvest sweet-peppers in a commercial greenhouse, but at limited success rates: harvest success was 6% when the Fin Ray end-effector was mounted, and 2% when the Lip-type end-effector was mounted. The performance evaluation against 13 performance indicators, nine failure categories, and several testing conditions enabled to identify the key problems. After simplifying the crop, by removal of fruit clusters and occluding leaves, harvest success was 26% (Fin Ray) and 33% (Lip-Type). Hence, these properties of the crop partly caused the low performance. Other major causes were: misplacement of the end-effector (32% of the failures), incorrect fruit localization (19% of the failures), difficult fruit orientations or difficult fruit locations at the backside of the stem (7% of the failures), and grasp and cut failures. In the simplified crop, grasp success of the Fin Ray end-effector (93%) was higher than the Lip-type end-effector

(61%), whereas cut success of the Lip-type end-effector (76%) was higher than the Fin Ray end-effector (29%). A future development should therefore combine the strengths of each end-effector.

The robot was capable of stem-dependent determination of a grasp pose for the end-effector. Enabling this capability decreased the stem damage rate from 19% to 13%, for the Fin Ray end-effector. In addition, grasp success increased from 41% to 61%, for the Lip-type end-effector tested in the simplified crop. Yet, due to the low number of samples, the effects were not significant (p>0.05). More samples need to be evaluated to proof significance. A task for future work is to perceive additional plant parts, such as leaves, to further enhance the performance.

The cycle time per fruit was commonly 94 s, i.e. a factor of 16 too long compared with an economically feasible time of 6 s. Several recommendations were made to bridge this gap in performance.

Acknowledgements

This research was funded by the European Commission in the 7th Framework Programme (CROPS GA no. 246252) and by the Dutch horticultural product board (PT no. 14555). We are grateful to growers Cees and Rolf Vijverberg who offered us their greenhouse to record measurements. We especially acknowledge the CROPS project partners with whom the hardware and software components were collaboratively developed.

References

Bac, C. W., Hemming, J., & van Henten, E. J. (2014a). Stem localization of sweet-pepper plants using the support wire as a visual cue. *Computers and Electronics in Agriculture, 105*, 111-120.

Bac, C. W., Roorda, T., Reshef, R., Berman, S., Hemming, J., & van Henten, E. J. (2015). Analysis of a motion planning problem for fruit harvesting in a dense obstacle environment. *Manuscript submitted for publication*.

Bac, C. W., Van Henten, E. J., Hemming, J., & Edan, Y. (2014b). Harvesting robots for high-value crops: state-of-the-art review and challenges ahead. *Journal of Field Robotics, DOI: 10.1002/rob.21525*.

Bachche, S., & Oka, K. (2013). Design, modeling and performance testing of end-effector for sweet pepper harvesting robot hand. *Journal of Robotics and Mechatronics, 25*(4), 705-717.

Barth, R., Baur, J., Buschmann, T., Edan, Y., Hellström, T., Nguyen, T., et al. Using ROS for Agricultural Robotics - Design Considerations and Experiences. In *Second RHEA International Conference on Robotics and associated High-technologies and Equipment for Agriculture, Madrid, Spain, 21-23 May 2014* (pp. 509-518).

Baur, J., Pfaff, J., Schuetz, C., & Ulbrich, H. Dynamic modeling and realization of an agricultural manipulator. In *Proceedings of the XV International Symposium on Dynamic Problems of Mechanics, Buzios, Brazil, 17-22 February 2013* (pp. 1-7).

Baur, J., Pfaff, J., Ulbrich, H., & Villgrattner, T. Design and development of a redundant modular multipurpose agricultural manipulator. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), Kaohsiung, Taiwan, 11-14 July 2012* (pp. 823-830).

CROPS (2014). CROPS: Clever Robots for Crops. http://www.crops-robots.eu/. Accessed 1-9-2014.
Edan, Y. (1995). Design of an autonomous agricultural robot. *Applied Intelligence, 5*(1), 41-50.
Edan, Y., Rogozin, D., Flash, T., & Miles, G. E. (2000). Robotic melon harvesting. *IEEE Transactions on Robotics and Automation, 16*(6), 831-835.
Gauchel, W., & Saller, S. Adaptive gripper jaws for high-value crops harvesting. In *8th International Fluid Power Conference, Dresden, Germany, 26-28 March 2012*.
Hayashi, S., Shigematsu, K., Yamamoto, S., Kobayashi, K., Kohno, Y., Kamata, J., et al. (2010). Evaluation of a strawberry-harvesting robot in a field test. *Biosystems Engineering, 105*(2), 160-171.
Hellström, T., & Ringdahl, O. (2013). A software framework for agricultural and forestry robots. *Industrial Robot, 40*(1), 20-26.
Hemming, J., Ruizendaal, J., Hofstee, J., & van Henten, E. (2014a). Fruit Detectability Analysis for Different Camera Positions in Sweet-Pepper. *Sensors, 14*(4), 6032-6044.
Hemming, J., Tuijl, B. A. J., Gauchel, W., & Wais, E. (2014b). Field test of different end-effectors for robotic harvesting of sweet-pepper. *Acta Horticulturae* (Vol. To appear).
Jovicich, E., Cnatliffe, D. J., Sargent, S. A., & Osborne, L. S. (2004). Production of Greenhouse-Grown Peppers in Florida. Gainesville, FL: University of Florida, IFAS Extension.
Kinney, J. J. (2002). Ch. 3: Estimation and Hypothesis testing. In *Statistics for Science and Engineering*. Boston, USA: Addison-Wesley
Kitamura, S., & Oka, K. (Vision). Recognition and cutting system of sweet pepper for picking robot in greenhouse horticulture. In *IEEE International Conference on Mechatronics and Automation, Niagra Falls, ON, July 2005* (Vol. 4, pp. 1807-1812).
Kondo, N., Yata, K., Iida, M., Shiigi, T., Monta, M., Kurita, M., et al. (2010). Development of an end-effector for a tomato cluster harvesting robot. *Engineering in Agriculture, Environment and Food, 3*(1), 20-24.
Kuffner, J. J., & Lavalle, S. M. (Motion Planning). RRT-connect: an efficient approach to single-query path planning. In *Proceedings - IEEE International Conference on Robotics and Automation, 2000* (Vol. 2, pp. 995-1001).
Kurniawati, H., Bandyopadhyay, T., & Patrikalakis, N. (2012). Global motion planning under uncertain motion, sensing, and environment map. *Autonomous Robots, 33*(3), 255-272.
Laerd (2014). McNemar's test using SPSS. https://statistics.laerd.com/spss-tutorials/mcnemars-test-using-spss-statistics.php. Accessed 2 September 2014.
Lewis, A., Watts, P. L., & Nagpal, B. K. (1983). Investment Analysis for Robotic Applications. *Technical Paper - Society of Manufacturing Engineers. MS*.
Murphy-Chutorian, E., & Trivedi, M. M. (2009). Head pose estimation in computer vision: A survey. *31*(4), 607-626.
Pekkeriet, E. J. (2011). CROPS project deliverable 12.1: Economic viability for each application. Wageningen, The Netherlands: Wageningen UR Greenhouse Horticulture.
Plebe, A., & Grasso, G. (2001). Localization of spherical fruits for robotic harvesting. *Machine Vision and Applications, 13*(2), 70-79.
Reed, J. N., Miles, S. J., Butler, J., Baldwin, M., & Noble, R. (2001). Automatic mushroom harvester development. *Journal of Agricultural Engineering Research, 78*(1), 15-23.

Scarfe, A. J., Flemmer, R. C., Bakker, H. H., & Flemmer, C. L. Development of an autonomous kiwifruit picking robot. In *4th International Conference on Autonomous Robots and Agents, 2009. , 10-12 Feb. 2009 2009* (pp. 380-384).

Schuetz, C., Pfaff, J., Baur, J., Buschmann, T., & Ulbrich, H. A Modular Robot System for Agricultural Applications. In *International Conference of Agricultural Engineering AgEng, Zurich, Switzerland, 6-10 July 2014*.

Shapiro, R. (1978). Direct linear transformation method for three-dimensional cinematography. *Research Quarterly, 49*(2), 197-205.

Tuijl, B. A. J., & Wais, E. (2014). *Harvesting Device*. Patent pending, Dutch Patent Office, Netherlands.

Van Henten, E. J. (2006). Greenhouse Mechanization: State of the Art and Future Perspective. *Acta Horticulturae*(710), p. 55-69.

Van Henten, E. J., Hemming, J., Van Tuijl, B. A. J., Kornet, J. G., Meuleman, J., Bontsema, J., et al. (2002). An Autonomous Robot for Harvesting Cucumbers in Greenhouses. *Autonomous Robots, 13*(3), 241-258.

Van Henten, E. J., Van Tuijl, B. A. J., Hemming, J., Kornet, J. G., Bontsema, J., & Van Os, E. A. (2003). Field Test of an Autonomous Cucumber Picking Robot. *Biosystems Engineering, 86*(3), 305-313.

Chapter 8 – Conclusion, General Discussion and Recommendations

C.W. Bac

8.1 Conclusion

The hypothesis, main objective and sub-objectives are listed in Chapter 1. This section lists the conclusions and the contributions to the field, for the six sub-objectives that correspond to Chapters 2-7. In addition, this section responds to the hypothesis and main objective. Section 8.2 provides a general discussion and recommendations for future research.

The following was concluded from the sub-objectives:

- The literature review in Chapter 2 contributes in three ways. Firstly, we characterized the crop environment and described three sources of variation to elucidate how the complexity of the crop environment forms the main bottleneck to better robot performance. Secondly, the 50 harvesting robots developed over the past three decades were quantitatively reviewed. This review revealed that performance did not improve in the past three decades. We discussed causes for this finding and listed five bottlenecks to a better performance. Lastly, to address these bottlenecks, four future challenges with R&D directions were detailed for development of harvesting robots.
- Chapter 3 classified pixels into hard or soft obstacles and, additionally, into five plant parts. Mean true-positive detection rate (standard deviation) among scenes was 59.2 (7.1)% for hard obstacles, 91.5 (4.0)% for soft obstacles, 40.0 (12.4)% for stems, 78.7 (16.0)% for top of a leaf, 68.5 (11.4)% for bottom of a leaf, 54.5 (9.9)% for fruit and 49.5 (13.6)% for petiole. These results were found insufficient to construct a reliable obstacle map. Yet, this study is the first to report quantitative performance for classification of several plant parts under varying lighting. A new robust-and-balanced accuracy performance measure P_{Rob} was introduced to prune the decision tree classifier and to select features. Use of P_{Rob} rendered the classifier more robust to variation among scenes because the standard deviation among scenes reduced 59% for hard obstacles and 43% for soft obstacles, compared with use of balanced accuracy as a performance measure.
- Chapter 4, on post-processing of classified pixels, introduces an alternative naming of false-positive rate, i.e. scaled false-positive rate, to overcome ambiguous use of this measure. Furthermore, the chapter reveals that stems were not properly detected after post-processing, whereas 57% of the green fruit were correctly detected, with a scaled false-positive rate of 7%.
- Chapter 5 studied localization of plant stems and depth accuracy of stereo-vision. An algorithm was developed that included novel components, such as adaptive thresholding, use of support wires as a visual cue, use of object-based and 3D features and use of minimum expected stem distance. Wire detection rates (true-positive/scaled false-positive) were more favourable under moderate irradiance (94/5%) than under strong irradiance (74/26%). Accuracy of the stereo-vision system (±0.4 cm) met the requirements in the lab, but not in the greenhouse (±4.5 cm) due to plant movement during recording. Later, this accuracy issue was resolved by simultaneous image acquisition. The algorithm and imaging set-up were integrated in the final robot. This chapter is the first study regarding stem localization of a plant under varying lighting conditions. The comparison of depth accuracy provides an additional contribution.

- Chapter 6 describes two analyses of the motion planning problem. The first analysis compared two methods of selecting the azimuth angle of the end-effector with respect to the target fruit. The new 'constrained-azimuth' method avoided risky paths and achieved a motion planning success similar to the 'full-azimuth' method. The 'constrained-azimuth' method was integrated in the robot. The second analysis involved a sensitivity analysis for five parameters specifying the crop (stem spacing and fruit location), the robot (end-effector dimensions and robot position) and the planning algorithm, to evaluate their effect on successfully finding a collision-free goal configuration and path. The fruit location at the stem was the strongest influencing parameter. The method of analysing may serve as a generic approach to study motion planning problems in a dense obstacle environment.
- Chapter 7 describes the first performance evaluation of a sweet-pepper harvesting robot tested in a greenhouse. The robot was able to harvest sweet-peppers in a commercial greenhouse, but at a limited performance: the cycle time was 94 s per fruit, harvest success was 6% when the Fin Ray end-effector was mounted, and 2% when the Lip-type end-effector was mounted. The performance evaluation against 13 performance indicators (including four new indicators), nine failure categories, and several testing conditions enabled to identify the key problems. After simplifying the crop, by removal of fruit clusters and occluding leaves, harvest success was 26% (Fin Ray) and 33% (Lip-Type). The performance evaluation was novel in considering two different crop conditions. Additionally, the robot's novel functionality of stem-dependant determination of the grasp pose was evaluated to respond to the hypothesis.

The hypothesis of this research was that *"a robot capable of distinguishing between hard and soft obstacles, and of employing this knowledge, improves harvest success and decreases plant damages during harvesting"*. Testing the effect of enabling stem-dependent determination of the grasp pose revealed that, in the simplified crop, grasp success increased from 41% to 61% for the Lip-type end-effector, and stem damage decreased from 19% to 13% for the Fin Ray end-effector. In the unmodified crop, grasp success increased from 48% to 52% (Lip-type), and stem damage decreased from 7% to 4% (Fin Ray). Although these effects seem large, they were not statistically significant, due to the low number of samples. Consequently, the hypothesis does not hold and more samples should be tested to evaluate significance. Also, the robot requires testing and evaluation under the full range of environmental conditions expected along a growing season, to generalize the hypothesis. To strengthen the proof for the hypothesis, more performance indicators could be evaluated.

In conclusion, this PhD research improves the obstacle awareness for robotic harvesting of sweet-pepper by the robot's capability of perceiving and employing hard obstacles (plant stems), whereas previous harvesting robots either lumped all obstacles in one obstacle class, or did not perceive obstacles at all. This capability may serve as useful generic functionality for future robots. Given this conclusion, the main objective *"to develop a sweet-pepper harvesting robot capable of distinguishing between hard and soft obstacles, and of employing this knowledge"* was fulfilled. Yet, the capability of perceiving and employing soft obstacles was not integrated in the final robot and needs further improvement.

8.2 General Discussion and Recommendations

This section intends to provide a bird's-eye-view on the research performed. To avoid repeating discussions in previous chapters, this section only elaborates on topics not yet, or limitedly, discussed before. In addition, recommendations are made for future work.

8.2.1 Obstacles and cognition

Chapter 3 and 4 explored the possibility to distinguish plant parts. The performance was limited and for most plant parts, the number of correctly detected pixels was not greater than the number of incorrectly detected pixels. Integrating this algorithm would presumably have worsened the robot performance. Thereafter, it was decided to prioritize improving stem perception over leaf perception because stem damages are generally more severe than leaf damages, and knowledge of the stem was useful to calculate a better grasp pose. After integrating stem perception in the final robot, there were no resources remaining in the project to improve perception of leaves. As a consequence, occluding leaves sometimes caused an incorrect fruit location or an obstruction for the end-effector (Chapter 7). Clearly, perceiving leaves could have delivered valuable knowledge for algorithms that suffered from occluding leaves, such as for fruit localization, grasp pose determination and motion planning.

Although the stem was perceived, other hard obstacles such as unripe fruit, and construction elements, were not perceived and not considered to determine a grasp pose. Furthermore, the robot localized ripe fruit, but their locations were only used as a target, and were not considered for motion planning or to calculate a collision-free grasp pose. Future work should, therefore, address perception and employing knowledge of these hard obstacles.

In addition to classification in hard and soft obstacles, more semantic knowledge can be obtained from obstacles. One can think of the strength, sensitivity to damages, growth along time, orientation change etc. For instance, information on the strength of the stem would enable the robot to decide upon the level of interaction with the stem. The robot could avoid damage-prone stems, and push stronger stems aside. Hence, obtaining such knowledge would enhance the robot's cognition of the working environment. Tools to improve cognition are detailed in the first future challenge of Chapter 2 and include development of world models containing 3D crop models (Weiss and Biber 2011) combined with learning and reasoning frameworks (Thrun et al. 2006; Belta et al. 2007), to link perception with appropriate actions.

8.2.2 Imaging

Chapter 3 and 4 concern pixel-based classification of plant parts using multi-spectral imaging, whereas Chapter 5 concerns stem localization using colour imaging. Both technologies offer advantages and disadvantages. Multi-spectral imaging potentially offers more reflectance data from interesting wavelengths such as at water absorption bands (900-950 nm, 1350-1450 nm, 1800-1950 nm) or the red edge (696-736 nm). However, recording of the images takes longer than colour imaging if a filter wheel is implemented. Furthermore, the scene should remain unchanged during recording and the images need to be aligned. Although multi-spectral cameras with a prism can acquire the images simultaneously, these cameras are generally several times more expensive and bigger than colour cameras. Hence, a trade-off exists

between the additional performance one can expect from multi-spectral imaging, vs. additional costs.

In the final tests in the greenhouse (Chapter 7) we used colour imaging described in Chapter 5 because the algorithm for stem localization hardly used any spectral features, in contrast to the algorithms in Chapter 3 and 4. Instead, the stem localization algorithm successfully used object-based (size, shape, orientation etc.) and 3D features extracted from the stereo images. Furthermore, the support wire was a visual cue to find the stem. Using the additional spectral features, offered by a multi-spectral camera, would therefore not have added much information. Given the experiences using spectral and object-based features, plant parts can be detected via multiple research directions, or 'multiple ways lead to Rome'. In future work, one can record videos and extract features from changes between frames, to distinguish plant parts. This approach was successfully used to distinguish water in videos (Mettes et al. 2014) and was also explored to detect leaves (Hemming et al. 2005). Leaves might reveal a different motion pattern than stems or fruit. Furthermore, features could be adaptively selected depending on the lighting conditions (Nieuwenhuizen et al. 2010).

In preliminary tests, we also assessed a structured-light camera: the Microsoft Kinect. However, the light pattern strongly suffered from disturbances by sunlight. To the best of my knowledge, there is currently no camera available that projects a pattern which is visible under direct sunlight. Some alternative sensors can better cope with direct sunlight and could be explored in future work, i.e. laser triangulation and laser scanners (Sick AG, Germany), light-field cameras (Raytrix GmbH, Germany), or high-resolution time-of-flight cameras (Odos Imaging Ltd, Scotland).

8.2.3 Ground truth

Labelling of ground-truth in the research chapters 2-7 was done partly by co-workers and mainly by myself. In other words, labelling was limited by a validation of additional annotators. The literature indicates that labelling of objects in images can differ among annotators. For instance, six annotators labelled on average 269 green apples in images, with a standard deviation of 49 green apples (Linker et al. 2012). Hence, there was only a partial agreement among these annotators. Possible methods to measure the inter-annotator agreement are discussed in a review article (Banerjee et al. 1999). Despite the additional resources needed, future research should employ several annotators to improve reliability of ground truth.

8.2.4 Differing practices between application domains of computer vision

If one compares computer vision literature for applications in crops, with general literature in computer vision, a difference in methodology can be observed. General computer vision literature is characterized by a tradition of performance benchmarking on several datasets. To illustrate, Table 8-1 is copied from a publication (Weinzaepfel et al. 2013) showing how the authors' method (DeepFlow) outperforms the state-of-the-art. Yet, the authors could have supported their contribution by conducting a statistical test to validate significance of the improvement.

Table 8-1. Example of benchmarking in the field of computer vision. The DeepFlow algorithm is compared against four earlier methods, using two performance indicators where a lower value indicates better performance. Methods were applied on the same dataset. Source: (Weinzaepfel et al. 2013)

Method	End Point Error	Runtime
Deepflow	**7.212**	**19**
S2D-Matching	7.872	~2000
MDP-Flow2	8.445	709
LDOF	9.116	30
Classic+NL	9.153	301

The practice of benchmarking of algorithms on a dataset saves the time required to develop an imaging system and record proper data. Datasets of plants in greenhouses do not yet exist and making these available would enable computer vision experts to test their algorithms for agricultural applications and furthermore improves knowledge transfer between these fields of science. On the other hand, Chapters 3-5 show that development of an imaging setup and data recording are a field of research in itself with many opportunities to improve detection performance. For instance, recording of underexposed images greatly improved performance of the stem localization algorithm. Hence, making datasets available seems a worthy direction to investigate, but developing imaging set-ups, and recording of proper data, remain of interest as well.

A topic discussed in general computer vision literature is to what extend an algorithm trained for say dataset 'A' performs well on another dataset 'B' (Torralba and Efros 2011). In other words, the challenge of cross-dataset generalization. The article by Torralba and Efros (2011) clearly shows there is a challenge for generalization because all algorithms trained on a 'native' dataset dropped in performance if not re-trained on the other dataset. This result indicates that algorithms were over-learned/trained on the objects of a specific dataset. A challenge for the future is therefore to generalize the performance of object detection algorithms in crops.

8.2.5 Motion planning

Chapter 6 analyses the motion planning problem of sweet-pepper harvesting. This chapter asserted that the Lip-type end-effector did not require the orientation of the fruit and successfully detached a fruit, if the target pose was taken in the horizontal plane and with a proper azimuth angle. However, the greenhouse tests in Chapter 7 revealed that the end-effector failed to grip fruit that were irregularly orientated, i.e. perpendicular to the stem. Although only a few fruit were that extremely oriented, the assumption made in Chapter 6 was not justified. Consequently, knowing the orientation of the fruit would be useful to optimize the grasp pose of the Lip-type end-effector.

The lip-type end-effector was designed to approach the fruit in the horizontal plane. Yet, another promising option is to design an end-effector that approaches fruit from the bottom side. A recent study developed and tested such end-effector and revealed that the percentage of fully visible strawberries was 75% when viewed from the bottom side and 40% when viewed from the aisle side (Yamamoto et al. 2014). The harvest success (67%) was greater

than in an earlier harvesting robot that approached fruit from the aisle side (41%). Hence, approaching fruit from the bottom side was successful for strawberry and could be evaluated for sweet-pepper or other crops as well.

The motion planner, implemented in Chapter 7, planned a path in the workspace. This planner did not meet the completeness criterion, i.e. a path was not always found if it existed. An algorithm that plans in the configuration space commonly meets the completeness criterion, such as the bi-RRT algorithm deployed in Chapter 6. Such planner is recommended to improve the motion planning success in future work.

Chapter 6 showed that there is more to motion planning than the planning algorithm. Other parameters also played a role to address the planning problem: the spacing of obstacles, the location of the target fruit, dimensions of hardware components, and the kinematic structure of the manipulator. Whereas current motion planning literature mostly seems to focus on better algorithms (Al-Bluwi et al. 2012; Tang et al. 2012), we showed that many other parameters should be considered as well. In fact, similar to the previous discussion on computer vision, hardware, software and characteristics of the application should be simultaneously considered, as stated in the future challenges of Chapter 2.

8.2.6 Harvesting vs. other plant maintenance operations

Although Chapter 2 discussed harvesting robots, many other operations are performed to maintain a plant during growth as well: attaching plants to a support wire or stick, side shoot removal, weeding or plant thinning, fruit or flower thinning, leaf picking, lowering plants, crop protection/spraying (Van Henten et al. 2013). These Plant Maintenance Operations (PMOs) are currently done manually, where harvesting, leaf picking, and pruning generally form the majority of labour input. Chapter 2 describes that a total of 50 projects were performed for harvesting, but only three projects were performed for other PMOs. Given the comparable labour input for leaf picking or side shoot removal, it is surprising that researchers mostly focused on harvesting. The complexity of for instance leaf picking seems similar to harvesting. And, the fact that picked leaves can be dropped and damaged is advantageous for automation. In sweet-pepper cultivation, attaching the plant to a support wire was identified as a candidate task for automation (Van Tuijl and Hemming 2013). Hence, investigating PMOs other than harvesting seems an interesting research direction. Furthermore, a robot developed for a particular PMO might be easily transferable to another PMO (Van Henten et al. 2002; Van Henten et al. 2006), or could be developed for several PMOs (Monta 1995). As a result, generic robotic solutions can be developed, thereby saving the effort to start from scratch for each PMO.

8.2.7 Technology Readiness Level

An indicator used to assess maturity of a technology is the Technology Readiness Level (TRL), with levels ranging from 1 (basic principles observed) to 9 (actual system is proven and ready for market introduction) (euRobotics 2014). A TRL of 4 corresponds with a validation of the technology under lab conditions, against established benchmarks. A level of 5 corresponds with a validation under field conditions. The sweet-pepper harvesting robot in this research, and the harvesting robots reviewed in Chapter 2, are somewhere at level 4 or 5.

Harvesting robots were mostly able to pick fruit under a set of limited environmental conditions at a specific moment in the season and along the day. Although for some robots performance indicators were reported for these conditions, performance benchmarking is hardly ever done and is necessary to determine the TRL. Hence, to reach a TRL of 5 and higher, robots should be benchmarked and tested under the complete range of environmental conditions expected in practice. Given the average harvest success rate of 66% for previous harvesting robots, business cases that employ human-robot co-working solutions are promising to advance the TRL to a level of 9 (Chapter 2).

A problem of the TRL definition itself is the lack of an explicit measure for the performance level achieved vs. the level required. For instance, the cycle time reported in Chapter 7 (94 s) was a factor of 16 too long compared with an economically feasible time of 6 s. This factor clearly indicates the gap to be bridged and would be a useful addition to the TRL definition.

8.2.8 Modifying the crop environment

Chapter 2 details and explains the future challenge of simplifying the task by modifications to the crop environment. This section reflects on the application of sweet-pepper harvesting with respect to this future challenge.

After visiting several greenhouses, we noticed that some sweet-pepper cultivars contained more fruit clusters than others. A plant breeding specialist (Rijk Zwaan BV, The Netherlands) later informed us that the genes responsible for fruit clusters were about to be identified and, as a result, the company could possibly reduce clusters in future cultivars. Given the difficulties we experienced when harvesting clustered fruit, such new cultivars are promising to simplify the task and improve performance.

The sweet-pepper crop was cultivated in the V-system (Jovicich et al. 2004), where plants grow in a relatively upward direction. A fellow Japanese researcher described a promising alternative Y-system in which sweet-pepper plants grow under a slight angle (Bachche 2013). Although these two systems are not yet compared, one can expect a better fruit visibility and accessibility in the Y-system because similar adaptations in cucumber (Van Henten et al. 2002) yielded a better robot performance.

Chapter 6 revealed that a wider spacing of plants in a crop row significantly improved the success of finding a goal configuration for the end-effector. Implementing a wider stem spacing therefore seems appropriate. However, these adaptations have strong economic consequences as well because the yield per m^2 of greenhouse surface will reduce. As mentioned in the fourth future challenge in Chapter 2, one should therefore involve all stakeholders (growers, engineers, scientists, and economist) and use reflexive design methods, to tackle design trade-offs for an alternative cultivation system.

8.2.9 Reflections of growers

Several meetings were held to report progress to an advisory board consisting of three sweet-pepper growers and a representative from the Dutch horticultural product board. In the last meeting in June 2014, the robot was demonstrated in the greenhouse and the performance was disseminated. Growers were on the one hand positive about the robot's ability to function

under unmodified crop conditions. On the other hand, they were well aware of the limited performance achieved, and they acknowledged the influence of crop conditions on performance. We discussed possibilities to modify the cultivation system by placing aisles in between each crop row instead of between pairs of crop rows. As a result, fruit could be approached from both sides, to avoid harvesting of fruit located at the backside of the stem. Although they acknowledged the benefits, the growers expected high investments costs that would render the robot economically infeasible. This meeting confirmed the relevance to consider not only technical, but also economic requirements when adapting the cultivation system, as also detailed in the third future challenge of Chapter 2.

8.2.10 Design process and project reflections

Previous sections clearly promote involvement of stakeholders and use of design methods. This section therefore reflects on the design process within the CROPS project. CROPS included multi-disciplinary expertise, i.e. computer vision experts, mechanical engineers, agricultural engineers, industrial engineers, software engineers, economists and plant scientists; altogether a group of about 35 people from industry and academia, and from ten different countries. Developing the applications, with such a diverse and large group of people located far away from each other, raised a challenge to communicate and effectively employ the available expertise. This challenge was only partially tackled, the following discusses causes and implications.

The project team responsible for the sweet-pepper application, attempted to use systematic design methods, but after completing the requirements (Hemming et al. 2011) the team was hesitant to continue using these methods to also select the most promising robot design. Although the methods structured thoughts and ideas, the team could have gained more from fully using them. A complicating factor to implement the methods was that not all partners were familiar with systematic design methods. In addition, due to the complex application, it took time to inform and convince project partners of the real problems of fruit harvesting that are detailed in Chapter 2.

As a consequence, a number of hardware components could have been designed better. The manipulator contained nine degrees-of-freedom, but the need for this redundancy was never proven in an analysis. Yet, the design team involved had to provide one manipulator for both sweet-pepper harvesting and apple harvesting, which complicated things considerably. Developing the hardware and motion planning for this complex manipulator took a large effort, whereas an off-the-shelf industrial manipulator might have fit the requirements as well; thereby clearing resources. As a result, these efforts could have been spent on developing a better motion planning algorithm. In addition, the sensor rig had to move out of the workspace to provide space for the manipulator. However, if the sensor rig and manipulator would have been jointly designed, presumably a stationary camera could have been used to reduce the cycle time. Similarly, the two end-effectors were developed by separate project partners with little collaboration. Chapter 7 indicates that a merger of both end-effectors would likely have improved the detachment success.

The matrix structure of workpackages in the CROPS project complicated the clarity of responsibilities. Therefore, literature suggests to use simpler structures (Hall 2013). Also,

testing and integration was planned in the end stage of the project, which revealed many issues that otherwise could have been addressed earlier. The literature suggests test-driven development to overcome this issue (Janzen and Saiedian 2005). Nevertheless, the objective to develop and evaluate the CROPS applications was joyfully delivered upon.

Consequently, in addition to research skills, design and management skills, shared interests, project commitment, organization structure, knowledge transfer, and leadership are relevant to render such projects successful. For more insights, the reader is referred to a study on effective collaboration in large multi-disciplinary research projects (Corley et al. 2006).

8.3 References

Al-Bluwi, I., Siméon, T., & Cortés, J. (2012). Motion planning algorithms for molecular simulations: A survey. *Computer Science Review, 6*(4), 125-143.

Bachche, S. (2013). *Automatic Harvesting for Sweet Peppers in Greenhouse Horticulture*. PhD Thesis, Kochi University, Japan.

Banerjee, M., Capozzoli, M., McSweeney, L., & Sinha, D. (1999). Beyond kappa: A review of interrater agreement measures. *27*(1), 3-23.

Belta, C., Bicchi, A., Egerstedt, M., Frazzoli, E., Klavins, E., & Pappas, G. J. (2007). Symbolic planning and control of robot motion [Grand Challenges of Robotics]. *IEEE Robotics & Automation Magazine, 14*(1), 61-70.

Corley, E. A., Boardman, P. C., & Bozeman, B. (2006). Design and the management of multi-institutional research collaborations: Theoretical implications from two case studies. *Research Policy, 35*(7), 975-993.

euRobotics (2014). Robotics 2020 Multi-Annual Roadmap. Initial Release B 15/01/2014.

Hall, K. (2013). Is teamwork the problem or the solution? *Human Resource Management International Digest, 21*(6), 33-36.

Hemming, J., Bac, C. W., & Van Tuijl, B. A. J. (2011). CROPS project deliverable 5.1: Report with design objectives and requirements for sweet-pepper harvesting. Wageningen, The Netherlands: Wageningen UR Greenhouse Horticulture.

Hemming, J., Van Henten, E. J., Van Tuijl, B. A. J., & Bontsema, J. (2005). A leaf detection method using image sequences and leaf movement. *Acta Horticulturae* (Vol. 691, pp. 765-772).

Janzen, D., & Saiedian, H. (2005). Test-driven development: Concepts, taxonomy, and future direction. *38*(9), 43-50.

Jovicich, E., Cnatliffe, D. J., Sargent, S. A., & Osborne, L. S. (2004). Production of Greenhouse-Grown Peppers in Florida. Gainesville, FL: University of Florida, IFAS Extension.

Linker, R., Cohen, O., & Naor, A. (2012). Determination of the number of green apples in RGB images recorded in orchards. *Computers and Electronics in Agriculture, 81*, 45-57.

Mettes, P., Tan, R. T., & Veltkamp, R. (2014). *On the segmentation and classification of water in videos*: Department of Information and Computing Sciences, Utrecht University.

Monta, M., Kondo, N., Shibano, Y. Agricultural robot in grape production system. In *The IEEE International Conference on Robotics and Automation, IEEE International Conference on Robotics and Automation, May 1995* (Vol. 3, pp. 2054-2509).

Nieuwenhuizen, A. T., Hofstee, J. W., & van Henten, E. J. (2010). Adaptive detection of volunteer potato plants in sugar beet fields. *Precision Agriculture, 11*(5), 433-447.

Tang, S. H., Khaksar, W., Ismail, N. B., & Ariffin, M. K. A. (2012). A review on robot motion planning approaches. *Pertanika Journal of Science and Technology, 20*(1), 15-29.

Thrun, S., Burgard, W., & Fox, D. (2006). *Probabilistic Robotics*. Cambridge, MA: MIT Press.

Torralba, A., & Efros, A. A. (Vision). Unbiased look at dataset bias. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2011* (pp. 1521-1528).

Van Henten, E. J., Bac, C. W., Hemming, J., & Edan, Y. Robotics in protected cultivation. In *IFAC Proceedings Volumes (IFAC-PapersOnline), 2013* (PART 1 ed., Vol. 4, pp. 170-177).

Van Henten, E. J., Hemming, J., Van Tuijl, B. A. J., Kornet, J. G., Meuleman, J., Bontsema, J., et al. (2002). An Autonomous Robot for Harvesting Cucumbers in Greenhouses. *Autonomous Robots, 13*(3), 241-258.

Van Henten, E. J., Van Tuijl, B. A. J., Hoogakker, G. J., Van Der Weerd, M. J., Hemming, J., Kornet, J. G., et al. (2006). An Autonomous Robot for De-leafing Cucumber Plants grown in a High-wire Cultivation System. *Biosystems Engineering, 94*(3), 317-323.

Van Tuijl, B. A. J., & Hemming, J. (2013). CROPS project deliverable 5.4: Report containing description and evaluation of generic concepts other than harvesting (sweet-pepper). Wageningen, The Netherlands: Wageningen UR Greenhouse Horticulture.

Weinzaepfel, P., Revaud, J., Harchaoui, Z., & Schmid, C. DeepFlow: Large Displacement Optical Flow with Deep Matching. In *IEEE International Conference on Computer Vision (ICCV), , 1-8 Dec. 2013 2013* (pp. 1385-1392).

Weiss, U., & Biber, P. (2011). Plant detection and mapping for agricultural robots using a 3D LIDAR sensor. *Robotics and Autonomous Systems, 59*(5), 265-273.

Yamamoto, S., Hayashi, S., Yoshida, H., & Kobayashi, K. (2014). Development of a Stationary Robotic Strawberry Harvester with a Picking Mechanism that Approaches the Target Fruit from Below. *Japan Agricultural Research Quarterly: JARQ, 48*(3), 261-269.

Summary

Harvesting of fruit is an activity performed since the existence of mankind. Nowadays, fruit is harvested manually or by a machine in case of a single harvest of the fruit, such as for juice extraction. However, fruit that require selective harvesting are still picked manually, which is costly and hard to recruit labour for. Introducing harvesting robots resolves these difficulties by gradually transforming physically intensive jobs into new jobs in application development, sales, or other services. Harvesting robots do not only reduce current labour costs that constitute 29% of the production costs, but they also enable new functionality using sensing abilities that humans either not possess, or cannot realize at a similar accuracy, consistency and cost.

This thesis describes the development of a harvesting robot for sweet-pepper. More specifically, it focusses on better handling obstacles when approaching a fruit for harvesting. Obstacles include nearby fruit, stems of the plant, leaves, or construction elements. In a sweet-pepper crop, obstacles are densely spaced and therefore limit the free workspace for a mechanism that can detach the fruit from the plant. Previous harvesting robots mostly attempted to detach a fruit without using information of obstacles. There were a few exceptions of harvesting robots with some ability of obstacle detection by assuming that any detected object in the environment is an obstacle, except for the target fruit. However, this approach can limit the accessibility of fruit because not every object is necessarily an obstacle in reality. To illustrate, when humans approach a fruit they sometimes push a leaf aside without causing any damage to the plant. Humans commonly treat the plant stem more carefully because a damage can seriously hamper plant growth and, consequently, future production of fruit. A leaf was therefore considered a soft obstacle, whereas the plant stem was considered a hard obstacle. A robot capable of distinguishing between hard and soft obstacles is currently lacking. If capable, a robot can take actions that may benefit from such knowledge. The hypothesis posed in this research is that a robot capable of distinguishing between hard and soft obstacles, and of employing this knowledge, improves harvest success and decreases plant damages during harvesting.

In line with the hypothesis, the main objective was to develop a sweet-pepper harvesting robot capable of distinguishing between hard and soft obstacles, and of employing this knowledge. Sub-objectives were the following: 1) to describe the crop environment of a harvesting robot, to review all harvesting robots developed for high-value crops, and to define challenges for future development; 2) to distinguish between different plant parts of a sweet-pepper plant and group them in hard and soft obstacles, using pixel-based classification; 3) to convert detected pixels of plant parts into blobs using post-processing techniques; 4) to localize the stem of a sweet-pepper plant; 5) to develop and test a motion planning algorithm for the robot arm and end-effector that plans a motion to a target fruit while avoiding hard obstacles; 6) to test the harvesting robot under greenhouse conditions and evaluate its performance, and to validate the hypothesis of this PhD research. These sub-objectives correspond with Chapters 2-7 in the thesis.

The literature review in Chapter 2 contributes in three ways. Firstly, we characterize the crop environment and describe three sources of variation to elucidate how the complexity of the crop environment forms the main bottleneck to better robot performance. Secondly, the 50 harvesting robots developed over the past three decades were quantitatively reviewed. This review revealed that performance did not improve in the past three decades. We discussed causes for this finding and listed five bottlenecks to a better performance. Lastly, to address these bottlenecks, four future challenges with R&D directions were detailed for development of harvesting robots.

In Chapter 3, a multi-spectral imaging set-up was developed to classify pixels into hard or soft obstacles and, additionally, into five plant parts. Vegetation was classified, using a Classification and Regression Trees (CART) classifier trained with 46 pixel-based features. A feature selection algorithm was implemented and selected the normalized difference index as the strongest feature. Mean true-positive detection rate (standard deviation) among scenes was 59.2 (7.1)% for hard obstacles, 91.5 (4.0)% for soft obstacles, 40.0 (12.4)% for stems, 78.7 (16.0)% for top of a leaf, 68.5 (11.4)% for bottom of a leaf, 54.5 (9.9)% for fruit and 49.5 (13.6)% for petiole. These results were found insufficient to construct a reliable obstacle map. Yet, this study is the first to report quantitative performance for classification of several plant parts under varying lighting. A new performance measure was introduced to realize a robust classification performance among scenes. This measure was used to prune the decision tree classifier and to select features. As a result, the classifier became more robust to variation among scenes because the standard deviation among scenes reduced 59% for hard obstacles and 43% for soft obstacles, compared with use of balanced accuracy as a performance measure.

Chapter 4, on post-processing of classified pixels, introduces an alternative naming of false-positive rate, i.e. scaled false-positive rate, to overcome ambiguous use of this measure. Post-processing operations applied were unable to remove false stem detections to an acceptable rate. Also, many false detections of the top of a leaf (>10%), bottom of a leaf (14%) and petiole (>15%) remained after post-processing, but these false detections were not critical for the application because these three plant parts are soft obstacles. Furthermore, results indicated that the top of a leaf can be distinghuished from the bottom. Stems were not properly detected after post-processing, whereas 57% of the green fruit were correctly detected, with a scaled false-positive rate of 7%.

To distinguish the stem better, Chapter 5 studied localization of stems and the depth accuracy of the stereo-vision system used for localization. A single colour camera was mounted on a pneumatic slide to record image pairs with a small baseline of 1 cm. Artificial lighting was developed to mitigate disturbances caused by natural lighting conditions. The algorithm developed, included novel components such as adaptive thresholding, use of support wires as a visual cue, use of object-based and 3D features and use of minimum expected stem distance. Wire detection rates (true-positive/scaled false-positive) were more favourable under moderate irradiance (94/5%) than under strong irradiance (74/26%). Accuracy of the stereo-vision system (±0.4 cm) met the requirements in the lab, but not in the greenhouse (±4.5 cm) due to plant movement during recording. Later, this accuracy issue was resolved by simultaneous image acquisition and the algorithm and imaging set-up were

integrated in the final robot. This chapter is the first study regarding stem localization of a plant under varying lighting conditions. The comparison of depth accuracy provides an additional contribution.

Besides localizing the objects in the crop, one should also plan a motion for a robot arm (manipulator) and gripper (end-effector) that grasps the fruit. Chapter 6 describes two analyses of the motion planning problem. As a basis for the analyses, data of fruit and stem locations were collected from a commercially used greenhouse. The first analysis compared two methods of selecting the azimuth angle of the end-effector with respect to the target fruit. The new 'constrained-azimuth' method avoided risky paths and achieved a motion planning success similar to the 'full-azimuth' method. The 'constrained-azimuth' method was integrated in the robot. The second analysis involved a sensitivity analysis for five parameters specifying the crop (stem spacing and fruit location), the robot (end-effector dimensions and robot position) and the planning algorithm, to evaluate their effect on successfully finding a collision-free goal configuration and path. The fruit location at the stem was the strongest influencing parameter. The two analyses may serve as useful tools to study motion planning problems in a dense obstacle environment.

Chapter 7 describes the test results of Europe's first sweet-pepper harvesting robot. The robot was able to harvest sweet-peppers in a commercial greenhouse, but at limited success rates: harvest success was 6% when the Fin Ray end-effector was mounted, and 2% when the Lip-type end-effector was mounted. The performance evaluation against 13 performance indicators (including four new indicators), nine failure categories, and several testing conditions enabled to identify the key problems. After simplifying the crop, by removal of fruit clusters and occluding leaves, harvest success was 26% (Fin Ray) and 33% (Lip-Type). Hence, these properties of the crop partly caused the low performance. Other major causes were: misplacement of the end-effector, incorrect fruit localization, difficult fruit orientations or difficult fruit locations at the backside of the stem, and grasp and cut failures. The cycle time per fruit was commonly 94 s, i.e. a factor of 16 too long compared with an economically feasible time of 6 s. Several recommendations were made to address these problems and to bridge the gap in performance. Additionally, the robot's novel functionality of stem-dependant determination of the grasp pose was evaluated to respond to the hypothesis.

The hypothesis of this research was that *"a robot capable of distinguishing between hard and soft obstacles, and of employing this knowledge, improves harvest success and decreases plant damages during harvesting"*. Testing the effect of enabling stem-dependent determination of the grasp pose revealed that, in a simplified crop, grasp success increased from 41% to 61% for the Lip-type end-effector, and stem damage decreased from 19% to 13% for the Fin Ray end-effector. In the unmodified crop, grasp success increased from 48% to 52% (Lip-type), and stem damage decreased from 7% to 4% (Fin Ray). Although these effects seem large, they were not statistically significant, due to the low number of samples. Consequently, the hypothesis does not hold and more samples should be tested to evaluate significance. Also, the robot requires testing and evaluation under the full range of environmental conditions expected along a growing season, to generalize the hypothesis. To strengthen the proof for the hypothesis, more performance indicators could be evaluated.

In conclusion, this PhD research improves the obstacle awareness for robotic harvesting of sweet-pepper by the robot's capability of perceiving and employing hard obstacles (plant stems), whereas previous harvesting robots either lumped all obstacles in one obstacle class, or did not perceive obstacles. This capability may serve as useful generic functionality for future robots. Given this conclusion, the main objective *"to develop a sweet-pepper harvesting robot capable of distinguishing between hard and soft obstacles, and of employing this knowledge"* was fulfilled. Yet, the capability of perceiving and employing soft obstacles was not integrated in the final robot and needs further improvement.

Chapter 8 concludes the thesis and attempts to provide a bird's-eye-view on this research in a general discussion with recommendations. The discussion on hard and soft obstacles addresses how to go beyond this classification by extracting more semantic knowledge from plant parts. A robot can obtain such knowledge through learning and reasoning, thereby enhancing cognition of the working environment. Technical topics covered include imaging, ground truth labelling, data recording, and motion planning. Furthermore, harvesting as an operation is compared with other plant-maintenance operations, with reflections on the technology readiness level. The chapter suggests possible modifications to the crop environment and provides reflections by growers. Lastly, the chapter reflects on the design process and the overall CROPS project.

Samenvatting

Al sinds het begin van de mensheid worden er vruchten geoogst. Voor lange tijd was dit handwerk, maar sinds de industriële revolutie worden ook machines ingezet. Momenteel is machinaal oogsten alleen rendabel in gewassen waarin alle vruchten op hetzelfde moment geoogst kunnen worden, zoals bijvoorbeeld bij sinaasappels die bestemd zijn voor sap-productie. Veel vruchten worden echter selectief geoogst en dit is handwerk. Arbeidskosten vormen 29% van de productiekosten en het is moeilijk om vaardige werknemers te vinden. De introductie van oogstrobots lost deze moeilijkheden op door fysiek intensieve banen geleidelijk te vervangen door nieuwe banen in innovatie, sales, onderhoud en andere diensten. Oogstrobots reduceren niet alleen de arbeidskosten, maar ze maken ook nieuwe toepassingen mogelijk (bijvoorbeeld ziektedetectie) door gebruik te maken van sensorische bekwaamheden die mensen ofwel niet bezitten, of niet kunnen realiseren met een zelfde nauwkeurigheid, consistentie en kosten.

Dit proefschrift beschrijft de ontwikkeling van een paprika-oogstrobot. De focus ligt op een betere omgang met obstakels tijdens het benaderen van een oogstbare vrucht. Voorbeelden van obstakels zijn o.a.: naburige vruchten, stengels van de plant, bladeren of constructie-elementen van de kas. In een paprikagewas zijn obstakels dicht tegen elkaar gepositioneerd en dat beperkt de vrije werkruimte voor de robot. Eerder ontwikkelde oogstrobots probeerden de vrucht te oogsten zonder gebruik van informatie over obstakels. Slechts enkele oogstrobots waren in staat om obstakels enigszins te detecteren door aan te nemen dat elk gedetecteerd object in de omgeving een obstakel is, behalve de te oogsten vrucht. Deze aanpak maakt het soms onmogelijk om een route voor de robotarm naar de vrucht te berekenen, omdat niet ieder object noodzakelijkerwijs een obstakel is. Mensen benaderen een vrucht door de bladeren opzij te duwen zonder de plant te beschadigen. Daarbij gaan ze voorzichtig om met de stengel van de plant omdat een beschadiging de groei kan beperken, met oogstderving als gevolg. Dit onderzoek beschouwt een blad daarom als een zacht obstakel en een stengel als een hard obstakel. Als een robot dit onderscheid kan maken, kan hij deze kennis benutten om betere acties uit te voeren. De hypothese van dit onderzoek is daarom dat een robot die in staat is harde en zachte obstakels te onderscheiden, en deze kennis weet te benutten, een hoger oogstsucces zal realiseren en minder plant beschadigingen zal veroorzaken.

Het hoofddoel van dit onderzoek hangt nauw samen met deze hypothese, namelijk: het ontwikkelen van een paprika-oogstrobot die in staat is harde en zachte obstakels te onderscheiden en die gebruik kan maken van deze kennis, tijdens het oogsten. Dit hoofddoel is opgesplitst in de volgende sub-doelen: 1) het beschrijven van de omgeving waarin de oogstrobot moet functioneren (gewasomgeving), het schrijven van een literatuuroverzicht van alle oogstrobots die ontwikkeld zijn, en het definiëren van uitdagingen voor onderzoek en ontwikkeling in de toekomst; 2) het onderscheiden van verschillende onderdelen van een paprika plant en het groeperen daarvan in harde en zachte obstakels doormiddel van pixel-gebaseerde classificatie; 3) het omzetten van gedetecteerde pixels van plantonderdelen in blobs doormiddel van nabewerkingstechnieken; 4) het lokaliseren van de stengel van een

paprika plant; 5) het ontwikkelen en testen van een bewegingsplanning algoritme voor de robotarm en grijper, zodanig dat de vrucht bereikt wordt en harde obstakels ontweken worden; 6) het testen van de robot onder kascondities en het evalueren van performance, en het valideren van de hypothese van dit onderzoek. Deze sub-doelen corresponderen met de hoofdstukken 2-7 in het proefschrift.

Het literatuuroverzicht in hoofdstuk 2 is op drie manieren belangrijk voor het onderzoeksveld. Ten eerste, we karakteriseren de gewasomgeving en beschrijven drie bronnen van variatie om te verklaren hoe de complexiteit van de gewasomgeving de voornaamste beperking is voor een betere performance van de robot. Ten tweede, we beschrijven een kwantitatieve recensie van de 50 oogstrobots die ontwikkeld zijn gedurende de laatste 30 jaar. Deze recensie onthulde dat de performance niet is verbeterd in de laatste 30 jaar. We hebben de oorzaken bediscussieerd voor deze vondst en vervolgens vijf beperkingen geformuleerd voor een betere performance. Als laatste, om deze beperkingen aan te pakken, hebben we vier uitdagingen geformuleerd voor de ontwikkeling van oogstrobots.

In hoofdstuk 3 is een multi-spectrale beeldopname opstelling ontwikkeld om pixels te classificeren in harde of zachte obstakels en, vervolgens, in vijf plantonderdelen: de vrucht, de stengel, de bovenzijde en onderzijde van een blad en de bladsteel. De vegetatie is geclassificeerd met een Classification and Regression Trees (CART; in deze samenvatting wordt de Engelse terminologie gebruikt) classificator die getraind werd aan de hand van 46 pixel-gebaseerde features. Een feature selectie algoritme selecteerde de Normalized Difference Index (NDI) als de sterkste feature. De gemiddelde true-positive detectie ratio over de beeldopnames was 59% voor harde obstakels, 92% voor zachte obstakels, 40% voor stengels, 78% voor de bovenzijde van een blad, 69% voor de onderzijde van een blad, 55% voor de vrucht en 50% voor de bladsteel. Deze percentages waren te laag om de obstakels betrouwbaar in kaart te brengen. Desondanks zijn de bevindingen waardevol omdat dit de eerste studie is die kwantitatieve performance rapporteert voor classificatie van verschillende plantonderdelen onder variërende lichtcondities. Daarnaast is een nieuwe performance maat geïntroduceerd om een robuuste classificatie te realiseren over de beeldopnames. Deze maat is gebruikt om de decision tree classificator te snoeien en om features te selecteren. Dientengevolge werd de classificator robuuster tegen variatie tussen beeldopnames; in vergelijking met gebruik van balanced accuracy als performance maat, nam de standaard deviatie met 59% af voor harde obstakels en met 43% voor zachte obstakels.

Hoofdstuk 4 beschrijft de nabewerking van geclassificeerde pixels en introduceert een alternatieve benaming voor de false-positive ratio, namelijk de geschaalde false-positive ratio, om dubbelzinnig gebruik van deze maat te voorkomen. De gebruikte nabewerkingsoperaties waren niet in staat om onjuiste stengeldetecties te verminderen tot een acceptabele ratio. Verder bleven er na nabewerking ook vele onjuiste detecties achter van de bovenzijde van het blad (>10%), de onderzijde van het blad (14%) en van de bladsteel (>15%), maar deze onjuiste detecties waren niet kritisch voor de applicatie omdat het zachte obstakels betrof. De resultaten lieten daarnaast zien dat de bovenzijde van het blad onderscheiden kan worden van de onderzijde van het blad. Alhoewel stengels niet behoorlijk gedetecteerd waren na nabewerking, werden 57% van de groene vruchten correct gedetecteerd met een geschaalde false-positive ratio van 7%.

Om de stengel beter te onderscheiden, richt hoofdstuk 5 zich op lokalisatie van de stengels en op de dieptenauwkeurigheid van stereo-vision. De kleurencamera was gepositioneerd op een pneumatische slee om beeldparen op te nemen met een smalle baseline van 1 cm. Er is een belichtingsopstelling ontwikkeld om de verstoringen door daglicht te verminderen. Het ontwikkelde algoritme bevatte vernieuwende elementen zoals gebruik van een adaptieve instelling van de drempelwaarde, gebruik van steundraden als visuele hint, gebruik van object-gebaseerde en 3D features en gebruik van de minimum verwachte stengel afstand. Detectie ratio's van steundraden (true-positive/geschaald false-positive) waren gunstiger onder matig zonlicht (94/5%) dan onder intens zonlicht (74/26%). De nauwkeurigheid van het stereo-vision systeem (±0.4 cm) voldeed aan de eisen onder labcondities, maar niet in de kas (±4.5 cm) doordat de planten bewogen tijdens opname van het linker en rechter beeld. Later is dit nauwkeurigheidsissue opgelost door simultane beeldopname. Het algoritme en de beeldopname opstelling zijn vervolgens geïntegreerd in de uiteindelijke robot. Dit hoofdstuk is de eerste studie over stengel-lokalisatie van een plant onder variërende lichtcondities. De analyse van de dieptenauwkeurigheid levert een additionele bijdrage aan de literatuur.

Naast de lokalisatie van objecten in het gewas moet er ook een bewegingsroute geplant worden voor de robot arm (ook wel manipulator genoemd) en grijper (end-effector) die de vrucht kan pakken en afsnijden. Hoofdstuk 6 beschrijft twee analyses van het bewegingsplanning probleem. Als basis voor de analyse zijn data van stengel en vruchtlocaties verzameld in een kas. De eerste analyse vergeleek twee methodes van selectie van de horizontale hoek van de grijper ten opzichte van de te oogsten vrucht. De nieuw geïntroduceerde 'constrained-azimuth' methode voorkwam risicovolle paden en bereikte een bewegingsplanning succes dat gelijk was aan de 'full-azimuth' methode. De 'constrained-azimuth' methode is geïntegreerd in de robot. De tweede analyse betrof een gevoeligheidsanalyse voor vijf belangrijke parameters ten aanzien van het gewas (stengelafstand en vruchtlocatie), de robot (grijper dimensies en robot positie) en het bewegingsplanning algoritme, om hun effect te evalueren op het succesvol vinden van een botsingsvrij pad naar de vrucht. De vruchtlocatie aan de stengel was de meest invloedrijke parameter. Deze twee analyses kunnen dienen als nuttig gereedschap om bewegingsplanning problemen te analyseren in omgevingen met veel obstakels.

Hoofdstuk 7 beschrijft testresultaten van Europa's eerste paprika-oogstrobot. De robot was in staat paprika's te oogsten in een praktijkkas, maar met beperkte succesratio's: het oogstsucces was 6% toen de Fin Ray grijper gemonteerd was, en 2% toen de Lip-type grijper gemonteerd was. De performance evaluatie ten opzicht van 13 performance indicatoren (inclusief vier nieuwe indicatoren), negen faal categorieën en verscheidene testcondities maakte het mogelijk om de belangrijkste problemen te identificeren. Na versimpeling van het gewas, door verwijdering van vruchtclusters en bedekkende bladeren, was het oogstsucces 26% (Fin Ray) en 33% (Lip-type). Dus deze eigenschappen van het gewas verklaren ten dele waarom de performance laag was. Andere belangrijke oorzaken waren o.a.: misplaatsing van de grijper, incorrecte vruchtlokalisatie, lastige vruchtoriëntaties of lastige vruchtlocaties aan de achterzijde van de stengel, en grijp en snij falen. De cyclustijd per vrucht was 94 s, een factor 16 te lang in vergelijking met een economisch rendabele tijd van 6 s. We hebben

diverse aanbevelingen gedaan om deze problemen op te lossen en het gat in performance te dichten. Verder is de robot's nieuwe functionaliteit, stengel-afhankelijk bepalen van de grijphoek, geëvalueerd om de hypothese te kunnen beantwoorden.

De hypothese van dit onderzoek was dat *"een robot die in staat is harde en zachte obstakels te onderscheiden, en deze kennis weet te benutten, een hoger oogstsucces zal realiseren en minder plant beschadigingen zal veroorzaken"*. Het testen van het effect van stengel-afhankelijk bepalen van de grijphoek onthulde dat, in een versimpeld gewas, het grijpsucces toenam van 41% tot 61% voor de Lip-type grijper, en stengelbeschadiging afnam van 19% tot 13% voor de Fin Ray grijper. In het ongewijzigde gewas nam het grijpsucces toe van 48% tot 52% (Lip-type) en stengelbeschadiging nam af van 7% tot 4% (Fin Ray). Alhoewel dit grote effecten lijken, betrof het een kleine steekproef en daardoor waren deze effecten niet statistisch significant. Derhalve geldt de hypothese niet en zou een grotere steekproef gedaan moeten worden om significantie te her-evalueren. Daarnaast vereist de robot een evaluatie onder de volledige range van omgevingscondities die te verwachten zijn gedurende een groeiseizoen, om zo de hypothese te generaliseren. Om het bewijs voor de hypothese te verzwaren, zouden ook meerdere performance indicatoren gemeten kunnen worden.

De conclusie van dit promotieonderzoek is dat het bewustzijn van de robot ten aanzien van obstakels is verbeterd doordat de robot in staat is harde obstakels (stengels) waar te nemen en deze kennis weet te benutten, in tegenstelling tot eerder ontwikkelde oogstrobots. Gegeven deze conclusie is het hoofddoel van dit onderzoek vervuld: *"het ontwikkelen van een paprika-oogstrobot die in staat is harde en zachte obstakels te onderscheiden en gebruik kan maken van deze kennis, tijdens het oogsten"*. Alhoewel de robot harde obstakels kon waarnemen, vereist de waarneming van zachte obstakels nog verdere ontwikkeling om de performance naar een acceptabel niveau te brengen. Het bewustzijn van obstakels kan als generieke functionaliteit ingezet worden in toekomstige robots.

Hoofdstuk 8 beschrijft bovenstaande conclusie over dit proefschrift en geeft een holistische blik op dit onderzoek doormiddel van een algemene discussie met aanbevelingen. Harde en zachte obstakels worden als eerst bediscussieerd en een aanbeveling is om meer semantische kennis te ontlenen aan plantonderdelen om zo obstakels verder te classificeren dan de twee klassen in dit onderzoek. Een robot kan dergelijke kennis verkrijgen door leren en redeneren, daardoor vergroot de robot zijn cognitie van de werkomgeving. Andere technische onderwerpen die in dit hoofdstuk aan bod komen zijn: beeldverwerking, ground truth labellen, data opname en bewegingsplanning. Daarnaast wordt oogsten vergeleken met ander plant-onderhoud operaties, met een reflectie op de technology readiness level. Het hoofdstuk bediscussieert een aantal aanpassingen aan de gewasomgeving en beschrijft reflecties van de kwekers hierop. Als laatste reflecteert het hoofdstuk op het ontwerpproces en het CROPS project als geheel.

Dankwoord / Acknowledgements

Ongeveer vier jaar geleden vroeg Eldert of ik geïnteresseerd was in een vacature gericht op de ontwikkeling van een plukrobot voor paprika's, als onderdeel van het Clever Robots for crOPS (CROPS) project. Ik was direct enthousiast over dit onderwerp, vooral nadat Jochen uitlegde wat het project inhield: verschillende applicaties, deelnemers uit verschillende landen, en mogelijkheden voor reizen en internationale uitwisseling. Daarnaast bood het project een uitzicht op spin-off activiteiten aangezien oogstrobots destijds nog niet op de markt waren. Deze combinatie van een interessant onderwerp, een prachtig project, en de mogelijkheden voor spin-off, klonken erg aantrekkelijk en ik besloot te solliciteren. Een keuze waar ik geen spijt van heb gekregen.

In het eerste jaar moest ik wennen aan de kritische vragen van Eldert en Jochen. Het schrijven van het onderzoeksvoorstel verliep dan ook moeizaam. De tekst kwam er wel, maar de kwaliteit en precisie lieten op zich wachten. Daarna begonnen de experimenten en het schrijven van het overzichtsartikel, een fase die weer nieuwe energie bracht. Het schrijven van het overzichtsartikel was een intensief proces waarin Eldert en Yael Edan veel input hebben geleverd. We hebben vele discussies gevoerd en soms wisten we geen van allen meer hoe we de reviewers een beter verhaal konden aanbieden (het artikel is zes keer ingestuurd naar hetzelfde journal en door minstens zes mensen gereviewed). Ik was heel blij dat het artikel uiteindelijk geaccepteerd werd en mogelijk is dit het beste hoofdstuk van het proefschrift (hoofdstuk 2). Na deze leerervaring vielen de andere hoofdstukken vlot op hun plek. In het laatste jaar heb ik de robot mogen testen in de kas, dit was een leuke en leerzame afsluiting van dit onderzoek!

My special thanks go to my promotor Eldert, co-promotor Jochen, and co-author and host Yael Edan. Eldert you have been of a great support throughout the thesis! Our discussions can be characterized as constructive, interesting, harsh, yet fun. I learnt to think critically and to put thoughts on paper. I am also grateful for your support on the grant proposals written, you supported me both mentally and practically. I learnt a lot from you and I also appreciate the informal conversations we had, for instance during the meeting in Bled. Jochen, I learnt much from your expertise in machine vision: designing a lighting set-up, selecting cameras, developing algorithms etc. Your support along the tests, conducted in the greenhouse, was crucial to get it all done. In addition, we often discussed the collaboration with project partners of the CROPS project, from which I learnt a lot. Yael, although you were never officially an advisor, I often felt you were :). You welcomed me in Beer Sheva with a very warm hospitality. The four-month work visit in 2011 was a great pleasure. The visit not only boosted my research, I also learned a lot from you about Israel, its culture, its history, its beautiful sites etc. The desert tours you organised were great fun. Two years later, you came to Wageningen for a sabbatical year in which we completed the literature review paper and worked on the robot. Thank you for everything and I am looking forward to meet with you again! I acknowledge co-authors Sigal Berman and Roi Reshef for the constructive collaboration on the motion planning paper (Chapter 6). We mostly met over Skype and met in person for just a few hours. Nevertheless, we got the paper done and I gained a lot from our

joint work. I also acknowledge Jörg Baur for hosting me during the work visit in Munich. It was a period in which I got acquainted with the Bavarian culture.

Ceng, Kurdish buddy, in Beer Sheva I got the privilege to live with you. You explained me a lot about the relations in the Middle East and how you ended up in Beer Sheva. We had many interesting debates about science, politics, and religion. These areas provided infinite fuel for discussion. The assertion that warfare will end if enemies attempt to love one another, instead of striking back, raised a long debate on Saturday morning. Afterwards we went to a local congregation and, surprisingly, we heard a sermon about exactly this topic. I enjoyed our conversations along the dinners, coffee breaks and drinks we had. For instance, you explained in great detail how to behave on a battlefield when things get serious. Perhaps one day this knowledge might be applicable, but hopefully I won't get involved in such a situation. Lastly, based on your experience in Middle Eastern studies, you kept recommending me to use two arms on the robot, instead of one. If your solution will prove to be effective in a future harvesting robot, I hereby wish to acknowledge you for the idea :). Danny, Efi, Ron, Guillaume, Roy, Bo, and Sasa, thanks for all the joy during and after working hours. It was great fun hanging out with you.

Mijn dank gaat uit naar de collega's van WUR Glastuinbouw en de Farm Technology Group. Bart, met je gouden handen heb je vele oplossingen gevonden en je was een bron van rust in de hectiek rondom het testen van de robot. Jan, je was altijd erg positief betrokken bij mijn onderzoek en onbewust heb ik veel van jouw adviezen overgenomen. Je nuchtere Groningse blik heb ik erg gewaardeerd. Ruud, jij hebt een belangrijke rol gespeeld in de software integratie. Jouw reflectie op de meningen van Jochen en mijzelf brachten vaak nieuwe inzichten. Daarnaast delen we een passie voor start-ups. Erik, hoe je het doet is me een raadsel, maar je haalt de projecten altijd vlotjes binnen. Wandelen kun je ook. Tijdens de laatste projectmeeting in Umea hebben we zo ongeveer alle straten wel gezien en over van alles gesproken, een mooie ervaring vond ik. Verder wil ik de andere collega's bedanken voor jullie adviezen tijdens presentaties en daarbuiten. Bastiaan dank voor al je review werk, je stond altijd klaar om dit proefschrift kritisch te bekijken. Peter en Dennis ook jullie dank voor het meedenken over de tekst. Joris, je was nog maar net begonnen en toen bracht je de verlossende inzichten over world modeling, reasoning en cognition, zeer waardevolle info. Verder hebben we samen nog een EU voorstel geschreven. Balen dat het niet gefinancierd werd, maar desondanks vond ik het een plezierige samenwerking. Bert, Liansun, Hyun, Phuong en Ellen ook jullie waren fijne collega promovendi, ik ga jullie en de gezelligheid tijdens koffiepauzes missen. Tot slot wil ik PE&RC bedanken voor de vele support en trainingsmogelijkheden die jullie geboden hebben. De PE&RC dagen waren een mooie gelegenheid om even niet aan dit onderzoek te denken. Ditzelfde geldt voor de VBW in Wageningen, erg leuk om met elkaar zo'n week met activiteiten voor kinderen te organiseren.

Graag wil ik de studenten bedanken die bijgedragen hebben aan het CROPS project. Bob van der Helm en David Koolen, jullie waren de eerste studenten die ik heb begeleid. Jullie hebben creatieve oplossingen gevonden om paprika-planten een tijdje in leven te houden om experimenten aan de rijping te kunnen doen. Thomas Fransen, in samenwerking met Gerrit Polder heb je de verschillende plantcomponenten doorgemeten, met waardevolle resultaten voor het CROPS project als gevolg. Daarna heeft Mans Vossebeld ook diverse andere

sensoren onderzocht. Dirk de Hoog, jij bent begonnen met het vraagstuk van het aansturen van de ‘complexe’ robot arm. Je hebt de mogelijkheden en beperkingen in kaart gebracht. Vervolgens heeft Tim Roorda een besturingsalgoritme geschreven dat nu ook in Hoofdstuk 6 is terug te vinden. Cora van Lieshout heeft nog een additionele analyse op de bereikbaarheid van de vruchten gedaan. Het was fijn samenwerken met jullie, you rock it! Jos Ruizendaal en Stijn Leijdekkers, jullie zijn door anderen begeleid en hebben veel bijgedragen. Heel erg bedankt!

Vrienden en familie, jullie waren een grote steun. De uitjes van Rectores de la vita hadden niets met dit onderzoek te maken en waren daarmee erg ontspannend. Hetzelfde geldt voor de gezelligheid met Juriaan, Laurens en vele anderen uit Moerkapelle. Jaargenoten van BAT05, het waren mooie momenten samen en we gaan er vast nog velen beleven. Pa en ma hartelijk dank dat jullie er altijd voor me zijn. Broers, schoonfamilie en pleegzussen, ook jullie waren zeer betrokken. Het was prachtig om de kinderen te zien genieten van een demo van de robot. Derreck, enorm bedankt voor je hulp bij de business challenge en de valorisation grant. Je hebt me veel geleerd over ondernemen en misschien maak ik nog eens gebruik van je busje uit Polen. Tijdens dit onderzoek ben jij, Johanna, op mijn pad gekomen. Dank voor al je liefde in het laatste jaar. Bovenal wil ik God bedanken voor Zijn leiding en zegeningen in mijn leven.

Curriculum Vitae

Cornelis Wouter Bac was born on 30 March 1987 in Moerkapelle, the Netherlands, where he grew up at an arable farm. He attended secondary school at the Driestar College in Gouda. After receiving his VWO diploma in 2005, he started a BSc programme in Agrotechnology at Wageningen University. In 2008, he completed this programme and received the CLAAS Innovation award for his thesis. He continued his MSc in Agricultural and Bioresource Engineering at Wageningen University, which he completed in 2010 (*cum laude*). He conducted his minor MSc thesis on laser-based localization of indoor robots, at the University of Illinois at Urbana-Champaign, USA. His major thesis involved modelling of oxygen competition and ion resistance in a microbial fuel cell for electricity generation from plants. The thesis was conducted at the Systems and Control Group of Wageningen University and was awarded with the WUF-KLV award. Wouter went to Technion, Israel Institute of Technology, for an internship on modelling of a fogging system in naturally ventilated greenhouses. In September 2010, he started his PhD research that focussed on a better awareness of obstacles, as part of the development of a harvesting robot for sweet-pepper. His research was mainly conducted at Wageningen UR Greenhouse Horticulture and at the Farm Technology Group of Wageningen University. Part of this research was conducted at the Technical University Munich, Germany, to study aspects of motion planning, and at Ben-Gurion University of the Negev, Israel, to study image processing methods. Since October 2014 he started working as an entrepreneur within the start-up company Fruitomation.

List of Publications

Peer-reviewed publications in journals

Bac, C. W., Hemming, J., Van Tuijl, B. A. J., Barth, R., Wais, E., & Van Henten, E. J. (2015). Performance evaluation of a harvesting robot for sweet-pepper. Submitted for publication in *Autonomous Robots*.

Bac, C. W., Roorda, T., Reshef, R., Berman, S., Hemming, J., & Van Henten, E. J. (2015). Analysis of a motion planning problem for fruit harvesting in a dense obstacle environment. Submitted for publication in *Biosystems Engineering*.

Bac, C. W., Van Henten, E. J., Hemming, J., & Edan, Y. (2014). Harvesting robots for high-value crops: state-of-the-art review and challenges ahead. *Journal of Field Robotics*, 31(6), 888-911.

Bac, C. W., Hemming, J., & Van Henten, E. J. (2014). Stem localization of sweet-pepper plants using the support wire as a visual cue. *Computers and Electronics in Agriculture*, 105, 111-120.

Bac, C. W., Hemming, J., & Van Henten, E. J. (2013). Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper. *Computers and Electronics in Agriculture*, 96, 148-162.

Bac, C. W., Grift, T. E., & Menezes, G. (2011). Development of a tabletop guidance system for educational robots. *Applied Engineering in Agriculture*, 27(5), 829-838.

Conference Papers

Hemming, J., Bac, C.W., Van Tuijl, B., Barth, R., Bontsema, J., Pekkeriet, E., & van Henten, E. (2014). A robot for harvesting sweet-pepper in greenhouses. Paper presented at the International Conference of Agricultural Engineering, 6-10 July 2014, Zürich, Switzerland. URL: http://www.geyseco.es/geystiona/adjs/comunicaciones/304/C01140001.pdf

Bac, C. W., Hemming, J., & Van Henten, E. J. (2013). Pixel classification and post-processing of plant parts using multi-spectral images of sweet-pepper. Paper presented at the IFAC Biorobotics Conference, 27-29 March, Sakai, Japan. IFAC Proceedings Vol. 1 (PART 1), p. 150-155.

Van Henten, E. J., Bac, C. W., Hemming, J., & Edan, Y. (2013). Robotics in protected cultivation. Paper presented at the Agricontrol Conference, Espoo, Finland. IFAC Proceedings Vol. 4 (PART 1), p. 170-177.

Bac, C.W., Grift, T.E., Bhalerao, K., Menezes, G. (2009). A laser based tabletop localization system for educational robots. Paper presented at the ASABE Annual International Meeting. ASABE 2009, Vol. 1, p. 180-192.

Poster presentations

Bac, C. W., Hemming, J., & Van Henten, E. J. (2014). Detecting plant parts of sweet-pepper using pixel classification and post-processing. Poster presented at the Netherlands Conference on Computer Vision (NCCV), 24-25 April, Ermelo, The Netherlands. URL: http://edepot.wur.nl/318917

Bac, C. W., Hemming, J., & Van Henten, E. J. (2012). Robust pixel-based classification of sweet-pepper plant parts using multi-spectral imaging. Poster presented at the EPPN Plant Phenotyping Workshop, 13-14 September, Wageningen, The Netherlands. URL: http://library.wur.nl/WebQuery/wurpubs/427974

PE&RC PhD Training Certificate

With the training and education activities listed below the PhD candidate has complied with the requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of literature (6 ECTS)

- Harvesting robots for high-value crops: state-of-the-art review and challenges ahead

Writing of project proposal (4.5 ECTS)

- Development of a robotic system for sweet-pepper production

Post-graduate courses (7.5 ECTS)

- Discrete chaotic dynamics and its application; Ben-Gurion University of the Negev, Israel (2011)
- Computational methods in intelligent systems; Ben-Gurion University of the Negev, Israel (2011)
- Introduction to R for statistical analysis; PE&RC (2012)

Laboratory training and working visits (4.5 ECTS)

- Motion planning algorithms for a manipulator; Technical University Munich, Germany (2011)
- Adaptive image processing algorithms; Ben-Gurion University of the Negev, Israel (2011/2012)
- Advances in strawberry harvesting; IAM BRAIN Institute, Japan (2013)
- Advances in tomato harvesting; Kyoto university, Japan (2013)

Invited review of (unpublished) journal manuscript (2 ECTS)

- Computers and Electronics in Agriculture: location of apples for robotic harvesting based on machine vision (2012)
- International Journal of Agricultural and Biological Engineering: design of structured-light vision system for tomato harvesting robot (2014)

Competence strengthening / skills courses (4.5 ECTS)

- PhD Competence assessment; WGS (2010)
- Scientific publishing; WGS (2011)
- Scientific writing; In'to Languages (2011)
- Services workshop for entrepreneurs; StartLife (2011)
- Effective behaviour in your professional surroundings; WGS (2011)
- Voice matters – voice and presentation skills training; WGS (2012)
- Communicating with the media and the general public; WGS (2012)
- Techniques for writing and presenting a scientific paper; WGS (2012)
- Writing grant proposals; In'to Languages (2013)
- PSG Business challenge; Plant Sciences Group (2013)

PE&RC Annual meetings, seminars and the PE&RC weekend (2.4 ECTS)

- PE&RC Weekend (2010)
- PE&RC Day (2010-2013)
- PE&RC Weekend – last year's edition (2013)

Discussion groups / local seminars / other scientific meetings (7.5 ECTS)

- Maths & Statts, later re-named into Modelling & Statistics Network (MSN); chairman (2010-2013)
- EU Meetings of the "Clever Robots for crOPS" (CROPS) project (2010-2014)
- Organiser of the PE&RC days (2011/2012)
- PE&RC PhD Council (PPC); member and vice-chairman (2011-2013)

International symposia, workshops and conferences (5.1 ECTS)

- EPPN Plant Phenotyping Symposium; poster presentation; Wageningen, the Netherlands (2011)
- IFAC Biorobotics Conference; presentation of a conference paper; Sakai, Japan (2012)
- Webinar presentation at IEEE Technical Committee on Agricultural Robotics (2013)
- Netherlands Conference on Computer Vision; poster presentation; Ermelo, the Netherlands (2013)

Lecturing / Supervision of practical's / tutorials (3 ECTS)

- Sensor technology; supervision of practical (2011-2013)
- Sensor technology; lecturing (2012-2013)

Supervision of MSc students

- Design evaluation of an sweet-pepper harvesting manipulator using experiments and simulation
- Assessment of three 3D sensors for localization of sweet-pepper plant parts
- Motion planning for a redundant manipulator to harvest sweet-pepper
- Minimum space needed for a 9-DOF manipulator to harvest sweet-pepper on the backside of the stem

Colophon

This research was funded by the European Commission in the 7th Framework Programme (CROPS GA no. 246252) and by the Dutch horticultural product board (PT no. 14555).

Cover and layout design: Wouter Bac
Printed by: GVO drukkers & vormgevers B.V. | Ponsen & Looijen, Ede