Computers and Electronics in Agriculture 122 (2016) 67–73

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Using depth cameras to extract structural parameters to assess the growth state and yield of cauliflower crops

Dionisio Andújar [a,*], Angela Ribeiro [a], César Fernández-Quintanilla [b], José Dorado [b]

[a] Center for Automation and Robotics, CSIC-UPM, Arganda del Rey 28500, Spain
[b] Institute of Agricultural Sciences, CSIC, Madrid 28006, Spain

A R T I C L E I N F O

Article history:
Received 3 June 2015
Received in revised form 4 December 2015
Accepted 12 January 2016
Available online 3 February 2016

Keywords:
Kinect
Plant structure characterization
Weight estimation
Volume estimation

A B S T R A C T

The use of robotic systems for horticultural crops is widely known. However, the use of these systems in cruciferous vegetables remains a challenge. The case of cauliflower crops is of special relevance because it is a hand-harvested crop for which the cutting time is visually chosen. This methodology leads to a yield reduction, as some inflorescences are cut before ripening because the leaves hide their real state of maturity. This work proposes the use of depth cameras instead of visual estimation. Using Kinect Fusion algorithms, depth cameras create a 3D point cloud from the depth video stream and consequently generate solid 3D models, which have been compared to the actual structural parameters of cauliflower plants. The results show good consistency among depth image models and ground truth from the actual structural parameters. In addition, the best time for individual fruit cutting could be detected using these models, which enabled the optimization of harvesting and increased yields. The accuracy of the models deviated from the ground truth by less than 2 cm in diameter/height, whereas the fruit volume estimation showed an error below 0.6% overestimation. Analysis of the structural parameters revealed a significant correlation between estimated and actual values of the volume of plants and fruit weight. These results show the potential of depth cameras to be used as a precise tool in estimating the degree of ripeness during the harvesting of cauliflower and thereby optimizing the crop profitability.

1. Introduction

Non-destructive methods for yield estimation may be a powerful tool to decide harvest time. In the case of cauliflower plants, the right time for harvest is when heads are full size, firm, compact and white. Usually, a visual assessment is conducted when leaves surrounding the cabbage heads are open, hand harvesting all the units that have the right requirements. This visual procedure frequently leads to harvest some small units that should had been cut in a later stage, reducing crop yield and commercial quality. The task of obtaining an unbiased estimation of cabbage yield before cutting the plant is a challenge, particularly considering new opportunities for automatic harvesting.

Increased knowledge in the use of sensors for plant reconstruction leads to a better understanding of the involved processes to increase yield and crop management. The use of accurate and efficient methods for plant phenotyping is crucial to obtain models that enhance crop yields. The description of plant architecture using phenotyping methodologies requires a set of methodologies and protocols to measure the plant characteristics with accuracy and precision at different scales of organization: from organs to canopies. The ongoing studies in robotic and automation processes using sensors as well as imaging and non-imaging technologies have provided a great variety of applications for plant characterization. In recent years numerous new methods have been devised for plant phenotyping (using noninvasive technologies to measure plant traits with a high precision and accuracy) (Dhondt et al., 2013; Fiorani and Schurr, 2013; Paulus et al., 2014). Cameras sensitive to the visible spectrum are an affordable solution that allows rapid characterization. Visible images are mostly based on CCD or CMOS arrays, which are sensitive to visible bands and create images in two dimensions. They are used for several purposes, such as shape description, growing processes, diseases and stress detection (Bock et al., 2010; Berge et al., 2012), crop yield estimation (Duan et al. 2011; Diago et al., 2012), phenology monitoring (Crimmins and Crimmins, 2008), plant structure characterization (Cescatti, 2007), weed shape description (Weis and Sökefeld, 2010), seedling vigor (Fiorani and Schurr, 2013), and biomass and nitrogen needs (Hunt et al., 2005). However, this type of camera offers limited information for certain physiological parameters. Visible imaging is the simplest method for phenotyping, in which

* Corresponding author. Tel.: +34 913 36 30 61; fax: +34 913 36 30 10.
E-mail address: dionisioandujar@hotmail.com (D. Andújar).

http://dx.doi.org/10.1016/j.compag.2016.01.018
0168-1699/

the images only provide physiological information. Challenges remain when the images are processed to extract phenotypic information, such as the case of leaf area and associated biomass, mainly because of the overlap of close leaves in the image. The use of fluorescence cameras provides information about the photosynthetic status, yields, stress and diseases, which are obtained from the excitation of the plant chloroplast and observation of the responses. The cameras incorporate a CCD that is sensitive to fluorescence signals by illuminating samples with visible or UV light. The fluorescence techniques can estimate the photosynthesis status to relate the effects of plant pathogens (Balachandran et al., 1997) and detect early stress responses to an effect in future yield (Chaerle et al., 2007). In plant phenotyping, fluorescence imaging provides information about physiological phenomena that are related to metabolism. On the contrary, most fluorescence imaging techniques are limited to a lab model, and the required rapid reproducibility and data analysis for larger scales are not sufficiently robust for on-field applications. NDVI sensors can distinguish green plant spectrum based on the light reflection in the red and near-infrared bands (Sui et al., 2008; Andújar et al., 2011); these active sensors use a self-illuminated light source in the red and near-infrared wavelengths. The use of NDVI sensors is mainly focused in crop nutrient management (Tremblay et al., 2009) and stress determination in some cases. These types of sensor are only used as complement for plant characterization. Thermal cameras indicate the temperature of an object's surface by measuring infrared radiation. The leaf temperature can be used as an indicator of evaporation or transpiration, with decreases indicating the response to water stress and transpiration (Munns et al., 2010). Spectral images of imaging spectroscopy can be obtained using multispectral or hyperspectral imaging cameras. The principle involves the effect of solar radiation in plants. The sensor can indirectly assess some parameters, such as spectral reflectance or near infrared measurement, to assess the yield and crop growth status (Ferrio et al., 2004). The measured parameters can be related with some other parameters, such as canopy status, water and chlorophyll content or green crop biomass.

Although all the aforementioned sensors can estimate several plant parameters, their use in cauliflower yield estimation is difficult since these sensors estimate the external characteristics of the plant based on 2D measurements and the cauliflower cabbage head should be cut when the fruit reach a commercial desirable volume. 3D modeling, particularly TLS (Terrestrial Laser Scanner), has led to increased effectiveness in plant science in recent years (Méndez et al., 2014). Its use for geometrical characterization has been widely explored. It is a relatively low-cost tool and a simple-to-operate solution. LIDAR sensors allow the scanning of any type of object by measuring the distance between the sensor and the impacted objects. The sampling creates a large spatial density of points at a notably high sampling frequency to reconstruct 2D or 3D models by displacing the sensor and storing the relative position. There are other systems that use 3D techniques, such as radar systems (Bongers, 2011), magnetic resonance and X-ray (Rudall and Rowe, 2003), ultrasonic systems (Andújar et al., 2011), hemispherical photography (Chen et al., 1991), stereovision using two or more combined cameras from motion for 3D data (Andersen et al., 2005) and depth cameras (Dal Mutto et al., 2012). Although these sensors could estimate cauliflower plant volume accurately, color information would be lost. This information may be critical for object separation between plant, ground and other object with different color in the field.

Regarding depth cameras, the low cost and high frame rate are revolutionizing the phenotyping industry. Between the two available types in the market: Time of Flight (ToF) and structured-light emission cameras, the latter is the most used in 3D modeling. There are two manufacturers in the market: Microsoft with the Kinect sensor and Asus with Xtion. Various authors have addressed the usefulness of these types of 3D cameras. Nock et al. (2013) showed the accuracy of both cameras in *Salix* branch segments of 2–13 mm. By scanning at different distances, they quantified the effect of the scanning distance and showed the possibilities for branch architecture reconstruction in woody plants. Paulus et al. (2014) compared two 3D imaging low-cost systems (David laser scanning system and the Microsoft Kinect device) to reconstruct the volumetric shape of sugar beet taproots and their leaves. The sensors created models that were similar to the reality and showed that phenotyping was possible using automated applications. Concerning agricultural robotics, Agrawal et al. (2012) developed an inexpensive robot of small-medium size that could pick and place plants and seedlings from one spot to other with an arm, which moves on rails, using a camera and a Microsoft Kinect sensor. Wang and Zhang (2013) explored the use of 3D reconstructing technology in the Kinect in a dormant tree and concluded that the system could reconstruct a 3D dormant tree that was sufficiently accurate for robotic pruning.

Although the detection was not as accurate as in TLS, the authors stated that these low-cost sensors could replace an expensive laser scanner in many plant-phenotyping scenarios. The measurement of plant size and shape was consistent with the horizontal and vertical measurements (Azzari et al., 2013). In addition, the calculated plant volumes using three-dimensional convex hulls were related to the plant biomass. Chéné et al. (2012) assessed the potential of 3D depth imaging systems for plant phenotyping with a self-develop algorithm to segment depth images of a plant from a single top view. They showed the possible applications for leaf morphology, orientation or pathogens detection. Wang and Changying (2014) estimated the onion fruit volume using the Kinect sensor. They measured the maximum diameter and volume of sweet onions. The predicted volume showed that this tool was a good non-destructive method to estimate the onion density. Other horticultural crops, such as cauliflower, have a similar problems for non-destructive yield estimation. However, in some cases, the fruit is covered and not visible. A non-destructive method for yield estimation before harvest is necessary, which is the case of cauliflower crops. New technologies to improve the agronomic management, reduce the cultivation cost, increase yields and improve the environmental quality could be required. In the case of this crop, the yield estimation is difficult and the use of sensing technologies can substantially improve the accuracy and precision of current methods based on eyeball and experience by matching yield with some subtracted information from 3D models. The estimation of plant volume and its relationship with internal fruit weight and fruit volume can increase yields. Additionally, cauliflower fruit weight and volume needs to be assessed since the volume estimation can be good indicator of turgor in relation with fruit weight. We propose an innovative approach to estimate cauliflower yield before cutting the plant using a Kinect depth camera. The goal of this paper is to explore the possibilities of the Kinect sensor to estimate the major parameters (weight, volume, distances, LAI) related to the final yield and, as a result, to define the optimum harvest time.

2. Material and methods

2.1. Sampling system

The samples were assessed with a Kinect sensor, which is identical to the one originally designed for gaming using Microsoft Xbox. The Kinect sensor is a depth camera containing a structured-light device integrated with an RGB camera, an infrared (IR) emitter and an IR depth sensor. The IR depth sensor includes