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

ARTICLE INFO

*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

ABSTRACT

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

an IR camera, which is a CMOS sensor. The IR camera is equipped with an 850–1100 nm bandpass filter, which captures the depth image. The RGB camera has a 400–800 nm bandpass filter. The system captures the frames at a rate of 30 frames per second and a depth image resolution of $640 \times 480$ pixels using the structured light method. The comparison with other time-of-flight systems reveals a low depth range. The Kinect for Windows v1 can be switched to the Near Mode, which provides a range of 500–3000 mm. However, this system has a minimum limit of 800 mm and a maximum limit of 4000 mm to work in the normal mode. The distance between the object and the Kinect sensor leads to a different accuracy in the measurement range. The sensor output is a high number of frames per second with redundant information overlapping in the continuous measurements. Thus, static objects lead to higher accuracy. The redundant information can be used to automatically remove outliers in the point cloud of the 3D model. This low-cost system uses the Kinect Fusion algorithms to recreate a 3D point cloud from the depth video stream. The Skanect®3D scanning software was used for data acquisition with a Kinect sensor. Point clouds were acquired from the Kinect sensor and the running software on an Intel laptop with Windows 7 and a 2 Gb Nvidia GeForce GTX graphic card with a Graphics Processing Unit (GPU) and CUDA. The Kinect was manually hand held and turned 360 degrees around a plant, while a video stream of point clouds was recorded. The software could merge the point cloud into a single point cloud by storing the relative position of the sensor in movement. The coordinate system of data from the Kinect is relative to the sensor itself. The clouds were rotated and translated in space into a common coordinate system. The overlapping areas in the clouds allow the algorithm to fuse the point clouds by common points. The hardware equipped in the laptop allows the software to process this information in real time. Every 3D model was scanned and modeled at distances of 70–100 cm away and moved along the plant.

### 2.2. Sampling measurements

In total, 30 assessments were made during two days throughout the growing period. Individual plants of cauliflower, i.e., entire fruits including leaves around them (hereinafter, samples), were collected in a commercial field in San Martín de la Vega (Madrid, Spain) in order to evaluate real field plants which can differ from those cultivated under greenhouse conditions. The samples were collected in two different states of fruit ripening. The first samples were collected in July 2014 at an early growing stage with a non-developed fruit. The plants were cut from the base with a part of the root system and properly stored in a humid and fresh atmosphere until the measurements could be taken in the next few hour to avoid the loss of turgor, reproducing the field conditions in the lab. On the second sampling day, a commercial time with a developed inflorescence was found in October 2014. Plants of various sizes and weights were harvested from the two dates, and a total of 10 and 20 plants were analyzed in the first and second samplings, respectively. The structural parameter of each plant, which are major axis along the equatorial perimeter (width and length), and the maximum height were manually measured using a tape measure. In addition, the weight was also measured using a digital precision balance before the Kinect readings.

An experiment was conducted with the usual lighting conditions in the laboratory, which are fluorescent lamps and natural daylight through the windows, but direct sunlight, which impedes correct measurement, was always avoided. Because the Kinect sensor has an error in depth images, which increases with a quadratic relationship with the distance between the sensor and the measured object, some calibration tests were performed to establish the correct scanning distance between the sensor and the plant. This error varies from a few millimeters to several centimeters when the object is located at the maximum range. Two different measurement procedures were performed according to the sampling date. On the first date, only the plants were scanned because the fruits were not developed. Plants were supported on pots during the scanning process using the root system. On the second date, when those plants had a totally developed inflorescence, the plants were scanned in two steps: before and after removing the leaves, i.e., the plants and fruits were individually scanned. The entire plants were individually scanned by moving the sensor 360° around the plant while recording a point cloud. The software merged different point clouds into one cloud using software fusion.

### 2.3. Volume and surface measurements from point clouds

The open software Meshlab® was used for offline processing of the acquired meshes. The software is aimed to process unstructured 3D models that arise from the depth camera readings. The software has some internal filters that allow the removal of duplicated points, unreferenced vertices, non-manifold edges, and null faces (Meshlab, Visual Computing Lab-ISTI-CNR). The stored data acquired during the readings were managed and plotted to create a model with the values taken by the sensor. The software was used to process unstructured models that arose in 3D scanning. The software includes remeshing tools and filters, which enable cleaning, managing, smoothing and tools for the curvature analysis, visualization and processing the obtained meshes. From the original point cloud to the final mesh, four processes were performed sequentially. (1) After importing, the data outliers and noise were filtered and removed using filters. Individual points at 1 cm outside the grid were removed using internal filters. A human visual procedure followed the automatic process to remove the outliers that were not removed using the filters. (2) Cleaning the supporting structure. The plants and inflorescences were held in a pot that was manually deleted from the point cloud. Following, the object of interest + was isolated after removing every source of interference inside the bounding box that the target object occupied. Some filters were tested to remove undesirable parts or noise based on color selection. However, no filter was found to be useful for isolating the target object because of the high similarity in color and the continuous point cloud without significant space between the object and the structure. Thus, a manual process was the option to remove those parts that did not correspond with the studied plant or inflorescence. (3) The samples were divided into subsamples to simplify the work. This process reduces the data resolution but maintains a high level of fidelity in the point cloud data. This process enables the use of different tools to reach the best fit. However, the final models were not subsampled, and every point was considered in the mesh. Among the various methods to process the points, the adjustment to the Poisson distribution appeared to have the best results in the automated filters. The noise filtering by the Poisson distribution takes the entire area of interest and looks at the corresponding distribution of points in 3D space. Although the Poisson method worked, some points were manually removed from the data. Then, the normals were calculated on the sub-sample to determine which side of the point is facing "out" and which is "in". (4) From this point, the surface reconstruction using the Poisson formulation algorithm was run. This algorithm simultaneously considers all points and do not resort to heuristic spatial partitioning. From the point cloud, it adapts for an approximate function of the inferred solid, whose gradient best matches the input normals.

The obtained meshes were processed to extract the main parameters and compared with the actual parameters. The major axis of each sample was measured by searching along the equatorial plant perimeter using the measuring tool in Meshlab.

The maximum height of the plant was also measured using the same tool. In addition, the mesh surface and mesh volume were calculated using different tools and stored to the subsequent statistical analysis. Plant surface is calculated by the Meshlab tool which summarizes the surface of every face in the mesh. The plant and fruit volume was calculated in Meshlab by its internal tool based on computation of polyhedral Mass properties (Mirtich, 1997). The algorithm computes the moment of inertia about various axes in rigid bodies. The algorithm is based on a three-step reduction of the volume integrals to successively simpler integrals. The algorithm also minimizes the errors of poor alignment of polyhedral faces. All volume integrals of a polyhedron are computed together during a single walk over the boundary of the polyhedron.

### 2.4. System to measure the actual fruit volume and main parameters

A measuring system was constructed to measure the actual fruit volume. The system based on the Archimedes principle consisted of a container and a top structure to push the fruit into the container. Hence, the volume of each fruit was equivalent to the volume of liquid displaced, which was measured after submerging a fruit in the container that was fully filled with water. Because cauliflowers are not compact, they were plasticized to prevent water fill in the gaps within the fruit. The air inside the bag was vacuumed. The water container was fully filled with water, and a horizontal bar was located in the top of the container to reduce the effect of accumulative volume by water surface tension.

Then, the plasticized sample was submerged into the container. The fruit was pushed down with a metallic rod in the top of the system which passed through a hole in the top of the structure. The rod enabled a total immersion of the sample into the water to prevent the fruit from floating on the surface because of the lower density. The displaced water was collected with a second container under the water container (Fig. 1). The water volume of the second container was then measured using a measuring cylinder. The volume of the empty vacuum bag and the rod immersed in the water were also measured using the same technique and subtracted from the water volume.

The Leaf Area (LA) of every plant was calculated from the images. All leaves were placed on a white surface, and digital images were taken using a Nikon D70 (Nikon Corporation, Tokyo, Japan) camera, which was fitted with a 50 mm Nikkor lens. It incorporated a 6.1 megapixel DX Format CCD image sensor that produces 3008 × 2000-pixel images. The RGB images were transformed to binary images. The excess green index (ExG = 2G-R-B) was applied (Meyer et al., 1998). This process creates a grey level image, where green objects (plants) are highlighted, and other objects and the background appear in dark. Then, a threshold of 8%, which corresponds to the grey level of 20 = 255 × 0.08, was applied to the ExG to separate the objects pixel-wise into foreground (plants) and background in a binary image. From this image, the pixel values that appear in black indicate the leaf area, and the pixels off the range, i.e., white pixels, correspond to the background. A standard 100 $cm^2$ black square was placed beside the leaves to be measured as a reference surface to calculate the LA.

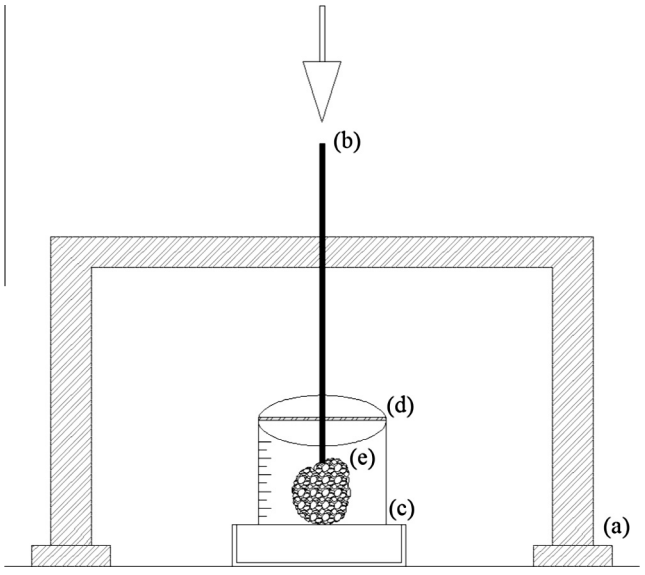


**Fig. 1.** Schematic design of the water displacement method to measure the fruit volume: (a) top structure to push the fruit into the container; (b) metallic rod to push the fruit into the water container; (c) and (d) water containers; (e) vacuumed plastic bag that contains the cauliflower fruit.

### 2.5. Statistical analyses

The recorded data in Kinect were validated with the actual values using regression methods. Prior to the regression analysis, a correlation analysis was conducted to make an initial examination of bivariate relationships among the variables. Pearson's correlation coefficient was used to analyze simple linear relationships between the actual values (ground truth values) and those measured with the sensor.

## 3. Results and discussion

The measured parameters that were obtained using the system showed a high correlation for Kinect models and vegetation structural parameters. Indeed, a good consistency in linear correlation equations between the actual parameters of plant and fruit and the corresponding estimated values from the models was found. For the plants in an earlier growing stage, the maximum length showed a strong correlation ($R^2$ = 0.998) and an RMSE of 2.31 cm. The maximum plant height and width were overestimated with RMSEs of 1.19 cm and 1.34 cm, respectively, and good correlation values ($R^2$ = 0.991 and $R^2$ = 0.998, respectively). These parameters remained almost constant with no significant differences between dates. Actual values of plant height and length had an average value of 63.81 cm and 70.1 cm for the first sampling date and 69.44 and 69.62 cm for the second date respectively. In general, for developed plants, the results showed that the fruit weight and volume could be well estimated measuring the entire plant volume (significant correlations at $P < 0.01$) using the sensor before harvesting. This value indicates the potential of depth cameras for yield estimation before harvesting. Other structural parameters were also well estimated. The maximum length of the plant showed a strong correlation ($R^2$ = 0.998) between the estimated value using the Kinect model and the actual value, which was assessed with a tape measure. The comparison of the model and ground truth values showed a root mean squared error (RMSE) of 2.25 cm. The real values were always overestimated by an average of 2.44% of the maximum length. Similarly, the maximum height and width were overestimated with RMSEs of 2.19 cm and 1.12 cm, respectively, which corresponded to a 2.7% and 1.31% average overestimation, respectively. Nock et al. (2013) showed similar results in the branch detection of *Salix* plants, which ranged from 2 to 13 mm in diameter, using an Asus Xtion and Kinect camera to quantify the limits and accuracy of detection. They found that the branches below 6.5 mm in diameter were not detectable by the camera, and there was a similar accuracy in the range of 0.5–1.25 m from the target object. Similar to our study, the errors were always related to the overestimation of the measurements. Thus, the authors suggested the necessity for improvements in

accuracy for certain tasks with the development of new software. The characterization of vegetation structures with depth cameras shows an overestimation of the actual plant values because these sensors cannot capture the end details. However, low-cost 3D imaging devices are highly reliable for plant phenotyping with the potential to be implemented in automated application procedures (Paulus et al., 2014). A similar case was studied with onions, where the diameter was estimated using its color images and depth images. The results of this work showed an underestimation of the onion size (i.e., maximum diameter) by using color images for the measurements. The method was compared with the calculation of the same dimensions using depth images, in which the latter showed more precise measurements than the color images, demonstrating that the depth cameras offer greater possibilities (Wang and Changying, 2014). However, the method showed a slight overestimation of the diameter of 2 mm. The 3D modeling method is achievable using depth cameras, and the plant reconstruction with Kinect is highly accurate (Fig. 2). Fig. 2a and b show a similar appearance, which is a indicator that LAI remains almost invariable during the fruit development. Fig. 2c shows a cauliflower cabbage head fully developed after removing leaves. The models tend to slightly overestimate the actual values because the sensors do not detect small details at the end of the leaves.

The Leaf Area measurements showed different values and evolution depending on the sampling date. Apparently, the LA grows to fruit development and subsequently stops. There are no significant correlations between the fruit volume or weight and the LAI. The LA remains almost constant during this period and shows no significance with some of the parameters. The LA index was correlated with the plant weight ($R^2$ = 0.65) and plant area, which were measured using the Kinect sensor ($R^2$ = 0.89). The LA evolution was previously reported to be affected by temperature and irradiance during vegetative growth for up to 60 days from the beginning of the experiment time, from which it becomes constant (Olesen and Grevsen, 1997). Similarly, the increase in number of leaves stops at 60 days at the end of the vegetative growing period, and the dry-matter production rate of leaves decreases (Kage and Stützel, 1999). Although the Kinect sensor could relate the LA with the plant area, this index does not prove to be a good indicator of the final yield.

The correlation between the estimated volume of the cauliflower cabbage head and the actual values of volume and weight of the entire plant was properly analyzed (Fig. 3). The actual fruit volume and the estimated volume using the sensor were consistent with an overestimation of the volume with an RMSE of 19.57 $cm^3$, which corresponds to an average error of 0.58%. Thus, the volume estimation by Kinect is a highly accurate methodology. Wang and Changying (2014) predicted the volume of onions with an accuracy of 96.3% and showed an RMSE of 18.5 $cm^3$. The cauliflower fruits have lower error possibly because of the higher volume, which minimizes the error during the measurements. The potential of this sensor during harvest is to automatically classify fruits based on the size and weight because the fruit volume is directly related to the weight. In addition, the regression curve between actual and estimated fruit weight value ($R^2$ = 0.87) confirms this relationship.

Regarding the relation of structural characteristics with features that were detected from the 3D model information, the results indicate the importance of an accurate algorithm that can detect the best matching between the measured parameters and the growth stage and yield. In this case, the plant volume can be easily calculated using the depth information. The entire plant volume that was calculated using the polyhedral mass property algorithm showed important information related to the fruit volume and fruit weight. By establishing this relationship, the best matching between the yield and the 3D model is useful to detect which plants should be harvested. This effect is of crucial importance because the leaves completely occlude the fruit at harvest time, and many fruits are usually cut before ripeness. The common practice is not a selective harvest, and some fruits are discarded after cut, which results in consequent economic loss. The results showed a significant relationship between the entire plant volume that was estimated using 3D models and the actual fruit volume and fruit weight. The linear regression models were adjusted (Fig. 4). The model's precision is high enough to predict volume and weight of the cauliflower cabbage head by measuring plant volume before harvesting. The actual fruit volume and the plant volume that was estimated using the sensor were highly correlated, which shows that Kinect is a promising tool to assess the fruit stage. Significant correlations between Kinect volume readings and the cited variables were observed at $P < 0.01$.

However, the system has certain limitations. Since the study was carried out by moving the sensor 360° around the plant, the use on field conditions needs the implementation of two or three sensor, storing the relative position of the subsequent sensors in the space, to obtain a complete model on the go. It is either not immune to high-lighting environments and should be used at certain hours during the day under outdoor conditions. High natural illumination impedes the collection of a sufficient number of points and an accurate resolution in the point cloud because of IR interference, which affects the depth data information. It is a similar case to the water scenarios, in which the reflected IR from the surface of the glass impedes the readings. For sunlight, direct sunlight will create a high level of interference and lead to noisy readings in the depth data. There are IR filters that can filter and reflect the external IR radiation. Using the near mode will properly detect close-up objects under more general lighting conditions. Low-light illumination is also a problem because low resolution is obtained, and the color detection is not sufficiently accurate. In the case of low illumination, the use of artificial light enables the Kinect sensor to capture the frames. The new model of Kinect for Windows v2 has new active-IR capabilities, which allows work in almost any lighting condition. This improvement will allow the use of the studied methodology in outdoors conditions under any

**Fig. 2.** Examples of 3D models obtained from Kinect readings: (a) sample from the first date; (b) sample from the second date; (c) fruit model.

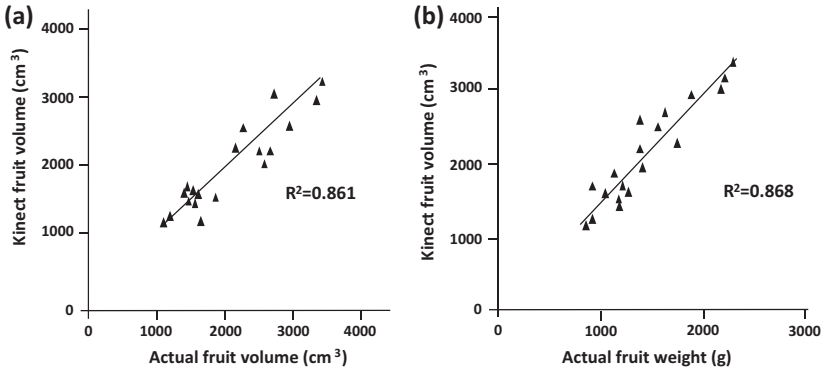


**Fig. 3.** Relationships between estimated and actual fruit volumes (a) and fruit weights (b).

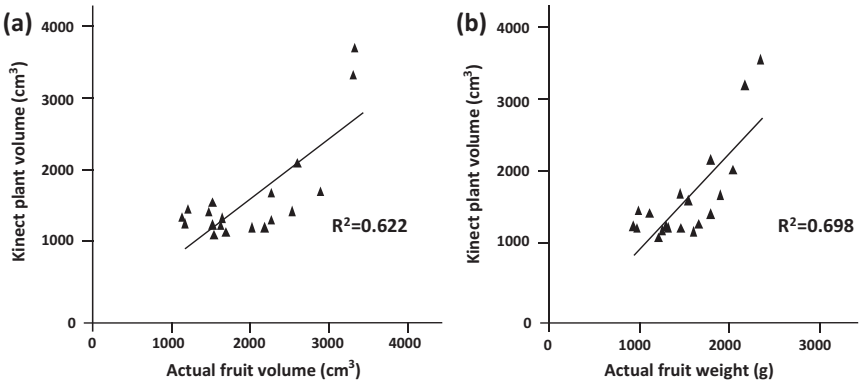


**Fig. 4.** Simple linear regression between plant volume estimated with Kinect sensor and actual fruit volume (a) and fruit weight (b).

lighting condition. The new model includes a high-definition (HD) color camera, and it uses the principle of Time of Flight (ToF) instead of structured light. The new sensor includes the sensor's ability to recognize features that improve the accuracy for 3D characterization. Although the improvements in the new sensor look important, the use of the first-generation Kinect remains valid for indoor applications or low illumination conditions. Paulus et al. (2014) compared the Microsoft Kinect Device with a laser scanner in measuring the volumetric shape of sugar beet taproots, the leaves of sugar beets and the shape of wheat ears. They concluded that this sensor could correctly measure volumetric objects. However, it did not precisely measure smaller and more complex structures. Azzari et al. (2013) compared depth images from a Kinect sensor with the manual measurements of plant structure and size of two plants. Although the Kinect showed some limitations for ecological observation due to lighting conditions and range, the consistency between the two methods was acceptable. Chéné et al. (2012) showed that 3D depth imaging from Kinect could identify the leaves with a high rate of success. Many applications have arisen using Kinect sensors, from robotics to human applications. Their use in agricultural tasks can also focus on yield estimation and detect the best time for harvesting. The results of this study show the potential of Kinect readings to detect the best time to collect fruits. Using an inexpensive sensor, the plants that should be cut can be automatically selected, and the immature ones in the field are left for a second cut round to increase the yield and net returns for the farmer.

## Acknowledgments

The Spanish Ministry of Economy and Competitiveness has provided support for this research via projects: AGL2011-25243 and AGL2014-52465-C4-3-R.

## References

Agrawal, D., Long, G.A., Tanke, N., Kohanbash, D., Kantor, G., 2012. Autonomous Robot for Small-Scale NFT Systems. In: proceedings: American Society of Agricultural and Biological Engineers (ASABE) Annual Meeting, Dallas, Texas.

Andersen, H.J., Reng, L., Kirk, K., 2005. Geometric plant properties by relaxed stereo vision using simulated annealing. Comput. Electron. Agr. 49, 219–232.

Andújar, D., Escolà, A., Dorado, J., Fernández-Quintanilla, C., 2011a. Weed discrimination using ultrasonic sensors. Weed Res. 51, 543–547.

Andújar, D., Ribeiro, A., Fernández-Quintanilla, C., Dorado, J., 2011b. Accuracy and feasibility of optoelectronic sensors for weed mapping in wide row crops. Sensors 11, 2304–2318.

Azzari, G., Goulden, M.L., Rusu, R.B., 2013. Rapid characterization of vegetation structure with a microsoft kinect sensor. Sensors 13, 2384–2398.

Balachandran, S., Hurry, V., Kelley, S., Osmond, C., Robinson, S., Rohozinski, J., Seaton, G., Sims, D., 1997. Concepts of plant biotic stress. Some insights into the stress physiology of virus-infected plants, from the perspective of photosynthesis. Physiol. Plantarum. 100, 203–213.

Berge, T.W., Goldberg, S., Kaspersen, K., Netland, J., 2012. Towards machine vision based site-specific weed management in cereals. Comput. Electron. Agr. 81, 79–86.

Bock, C.H., Poole, G.H., Parker, P.E., Gottwald, T.R., 2010. Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. Crit. Rev. Plant Sci. 29, 59–107.

Bongers, F., 2011. Methods to assess tropical rain forest canopy structure: an overview. Plant Ecol. 153, 263–277.

Chaerle, L., Leiononen, I., Jones, H.G., Van Der Straeten, D., 2007. Monitoring and screening plant populations with combined thermal and chlorophyll fluorescence imaging. J. Exp. Bot. 58 (4), 773–784.

Chen, J.M., Black, T.A., Adams, R.S., 1991. Evaluation of hemispherical photography for determining plant area index and geometry of a forest stand. Agr. For. Meteorol. 56, 129–143.

Chéné, Y., Rousseau, D., Lucidarme, P., Bertheloot, J., Caffier, V., Morel, P., Belin, E., Chapeau-Blondeau, F., 2012. On the use of depth camera for 3D phenotyping of entire plants. Comput. Electron. Agr. 82, 122–127.

Crimmins, M.A., Crimmins, T.M., 2008. Monitoring plant phenology using digital repeat photography. Environ. Manage. 41, 949–958.

Cescatti, A., 2007. Indirect estimates of canopy gap fraction based on the linear conversion of hemispherical photographs – methodology and comparison with standard thresholding techniques. Agr. For. Meteorol. 143, 1–12.

Dal Mutto, C., Zanuttigh, P., Cortelazzo, G.M., 2012. Time-of-Flight Cameras and Microsoft Kinect™. Series: Springer Briefs in Electrical and Computer Engineering. Springer, New York, USA.

Dhondt, S., Wuyts, N., Inzé, D., 2013. Cell to whole-plant phenotyping: the best is yet to come. Trends Plant Sci. 8, 1–12.
Diago, M.P., Correa, C., Millán, B., Barreiro, P., Valero, C., Tardaguila, J., 2012. Grapevine yield and leaf area estimation using supervised classification methodology on RGB images taken under field conditions. Sensors 12, 16988–17006.
Duan, L., Yang, W., Huang, C., Liu, Q., 2011. A novel machine-vision-based facility for the automatic evaluation of yield-related traits in rice. Plant Meth. 7, 44–57.
Ferrio, J., Bertran, E., Nachit, M., Catala, J., Araus, J., 2004. Estimation of grain yield by near-infrared reflectance spectroscopy in durum wheat. Euphytica 137, 373–380.
Fiorani, F., Schurr, U., 2013. Future scenarios for plant phenotyping annual. Rev. Plant Biol. 64, 267-29.
Hunt Jr., E.R., Cavigelli, M., Daughtry, C.S.T., Mc Murtrey, J.E., Walthall, C.L., 2005. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. Precis. Agric. 6, 359–378.
Kage, H., Stützel, H., 1999. A simple empirical model for predicting development and dry matter partitioning in cauliflower (*Brassica oleracea* L. botrytis). Sci. Hortic. 80 (1–2), 19–38.
Méndez, V., Rosell-Polo, J.R., Sanz, R., Escolà, A., Catalán, H., 2014. Deciduous tree reconstruction algorithm based on cylinder fitting from mobile terrestrial laser scanned point clouds. Biosyst. Eng. 124, 78–88.
Meyer, G.E., Mehta, T., Kocher, M., Mortensen, D., Samal, A., 1998. Textural imaging and discriminant analysis for distinguishing weeds for spot spraying. T. ASABE 41, 1189–1197.
Mirtich, B., 1997. Fast and Accurate Computation of Polyhedral Mass Properties. On-line paper. <http://www.cs.berkeley.edu/~jfc/mirtich/massProps.html>.
Munns, R., James, R.A., Sirault, X.R., Furbank, R.T., Jones, H.G., 2010. New phenotyping methods for screening wheat and barley for beneficial responses to water deficit. J. Exp. Bot. 61, 3499–3507.
Nock, C.A., Taugourdeau, O., Delagrange, S., Messier, C., 2013. Assessing the potential of low-cost 3D cameras for the rapid measurement of plant woody structure. Sensors 13, 16216–16233.
Olesen, J.E., Grevsen, K., 1997. Effects of temperature and irradiance on vegetative growth of cauliflower (*Brassica oleracea* L. botrytis) and broccoli (*Brassica oleracea* L. italica). J. Exp. Bot. 48, 1591–1598.
Paulus, S., Behmann, J., Mahlein, A.K., Plümer, L., Kuhlmann, H., 2014. Low-cost 3D systems: suitable tools for plant phenoyping. Sensors 14, 3001–3018.
Rudall, P.J., Rowe, T.B., 2003. Three-dimensional analysis of plant structure using high-resolution X-ray computed tomography. Trends Plant Sci. 8, 2–6.
Sui, R., Thomasson, J.A., Hanks, J., Wooten, J., 2008. Ground-based sensing system for weed mapping in cotton. Comput. Electron. Agr. 60, 31–38.
Tremblay, N., Wang, Z., Ma, B.L., Belec, C., Vigneault, P.A., 2009. Comparison of crop data measured by two commercial sensors for variable-rate nitrogen application. Precis. Agric. 10, 145–161.
Wang, W., Changying, L., 2014. Size estimation of sweet onions using consumer-grade RGB-depth sensor. J. Food Eng. 142, 153–162.
Wang, Q., Zhang, Q., 2013. Three-Dimensional Reconstruction of a Dormant Tree Using RGB-D Cameras. In: proceedings ASABE Annual International Meeting, Kansas City, Missouri.
Weis, M., Sökefeld, M., 2010. Detection and identification of weeds. In: Precision Crop Protection-the Challenge and Use of Heterogeneity. Springer Verlag, Dordrecht, The Netherlands, pp. 119–134, 1.