A Comparative Study of 3D Plant Modeling Systems Based on Low-Cost 2D LiDAR and Kinect

Harold Murcia(✉), David Sanabria, Dehyro Méndez, and Manuel G. Forero

Universidad de Ibagué, Ibagué-Tolima 730002, Colombia
{harold.murcia,manuel.forero}@unibague.edu.co,
{2420141016,2d20191003}@estudiantesunibague.edu.co

Abstract. Morphological information of plants is an essential resource for different agricultural machine vision applications, which can be obtained from 3D models through reconstruction algorithms. Three dimensional modeling of a plant is an XYZ spatial representation used to determine its physical parameters from, for example, a point cloud. Currently two low-cost methods have gained popularity in terms of 3D object reconstructions in 360° employing rotating platforms, based on 2D LiDAR and Kinect. In this paper, these two techniques are compared by getting a 3D model of a *Dracaena braunii* specie and evaluating their performance. The results are shown in terms of their accuracy and time consumption using a Kinect V1 and a LiDAR URG-04LX-UG01, a well-performance low-cost scanning rangefinder from Hokuyo manufacturer. In terms of error calculation, the Kinect-based system probed to be more accurate than the LiDAR-based, with an error less than 20% in all plant measurements. In addition, the point cloud density reached with Kinect was approximately four times higher than with LiDAR. But, acquisition and processing time was about twice than LiDAR system.

Keywords: Low-cost · Phenotyping · LiDAR · Kinect · Point clouds · 3D modeling

1 Introduction

The importance of plant phenotyping, i.e. the determination of plant structures and morphological parameters is widely recognized among researchers from different scientific fields. Phenotyping platforms are necessary to allow the determination of plant features and the formulation of genomic models for plant breeding. An appropriate plant surface sampling, with a convenient resolution of the 3D modeling, leads to different morphological measurements such as leaf area and angle, and plant topology. The requirements for real-time responses to post-processing data are an important task in different perception fields. Therefore, in recent years, a new generation of 3D sensors has appeared, known as depth

E. Roman-Rangel et al. (Eds.): MCPR 2021, LNCS 12725, pp. 272–281, 2021.
https://doi.org/10.1007/978-3-030-77004-4_26

cameras, which main advantage is the rapid acquisition of depth images. Some of them are based on structured light emission sensors, such as the Microsoft Kinect or Asus Xtion, and others on laser scanning sensors, using what is known as Time of Flight (ToF) technology, representing a revolution in 3D imaging due to its performance offered at a low cost. ToF allows getting distances by measuring phase difference between the modulated signal emitted and received with a specific wavelength. LiDAR devices are based on ToF and its use in phenotyping applications is not entirely new. However, its application has not yet been fully explored. LiDAR-based techniques are popular in field and laboratory applications, given their wide resistance to dust, robustness to changing lighting conditions, wide measurement range, fast time response and ease of deployment.

Besides, they are suitable for 3D plant and foliage reconstruction, a fundamental factor in obtaining characteristic plant models and their monitoring over time. The use of LiDAR technology in the phenotyping process includes the measurement of density and volume in plants and crops for the estimation of different parameters such as height, biomass, leaf indices, etc. [10,16]. LiDAR sensors are frequently chosen, in a large number of applications, to provide range data, such as plant differentiation by height and pattern, automatic identification of stem and leaf organs using point clouds obtained from the 3D scanning of barley and other cereals, according to their geometric shapes and histograms. They are also used to create virtual plants and tree models on large-scale phenotyping platforms. The 3D models obtained with LiDAR sensors can also be merged with visual information to generate geometric and multispectral models that allow developing more complete phenotyping processes, and also new classification and automatic processing techniques used for pest and disease control, irrigation, fertilization and plant stress, among others.

Kinect devices introduced by Microsoft (version V1 and V2) as an interface to track body position for Xbox videogame consoles have gained popularity in engineering and robotics applications [12]. One of its main advantages is its low cost compared to other sensors. Kinect sensors cost about 0.1% of commercial research LiDAR systems. The Kinect projects a pattern of points using an infrared laser on the scene of interest. The target is captured by an infrared camera and aligned with the image obtained from a standard one. In this way, the Kinect produces a point cloud with XYZRGB dimensions in a distance range from 0.5 to 4 m approximately. Compared to version 1, the Kinect V2 has improved characteristics like higher video resolution of 1920×1080 pixels, data transfer rate of 30 fps, field of view of $70° \times 60°$ and better signal to noise ratio in daylight scenes [1]. Given the recent interest in this type of work and the absence of comparative technical studies between these types of sensors, the aim of this work is to present and compare two methods for 3D plant modeling from a point cloud. The following section summarizes the basics of ToF perception and briefly reviews the literature dealing with the generation of plant point clouds. Section 3 describes the elements and methods used. Finally, Sects. 4 and 5 present the experimental results and conclusions.

2 Related Work

Although the state-of-the-art mentions the generation of 3D models using multivision systems, these are not included among those that provide depth clues and require a neutral background, easily separable from the object due to its strong contrast, to facilitate better segmentation [11]. Several 3D reconstruction applications based on 2D LiDAR sensors have been developed in the literature. Most of them are installed on mobile platforms for outdoor applications that present high-quality results after a camera calibration process [2,5,8,9]. However, some equipments are used indoors for phenotyping purposes. Wang et al. [14] reported accurate results in an indoor environment with a lower-cost RP-LiDAR laser scanner integrated into a mobile proximal detection system. Thapa et al. [13] proposed a scanning system consisting of a SICK LMS511 LIDAR and a 360° rotating platform. Panjvani et al. [7] presented a LIDAR system based on a SICK LMS400 device integrated in a linear moving platform for leaf feature extraction. Unfortunately, most of them are based on expensive sensors, which are not easily accessible to many users.

Therefore, several approaches have been developed for the reconstruction of 3D models using low-cost sensors such as Kinect, in particular for plant phenotyping. Li et al. [3] introduced a method to segment leaves without occlusions from three different types of 3D image platforms: stereo cameras, a Kinect V2 sensor and a multi-vision stereo camera in a mobile phone, scanning four types of plants. The technique included the automatic estimation of morphological features such as area, length and width of the leaf. Point coverage rate between 87% and 99% and accuracy of almost 100% were obtained in all cases. McCormick et al. [5] developed a semiautomatic image acquisition and processing pipeline for shoot segmentation of a sorghum variety. For each plant, a series of 12 depth and RGB images were acquired and the resulting point clouds were processed to segmented meshes. Image-based measurements like shoot and leaf height, surface area, leaf width and angle were well correlated with manual measurements. Root-Mean-Square Difference (RMSD) coefficient of variation for the measurements ranged from 0.07 to 0.3 within the same range as real values. Yamamoto et al. [15] employed a method to extract a 3D model and evaluate volume and diameter for fruits and vegetables. Both Kinect V1 and V2 were used to create 111 3D models. Depth and color information were processed and several features of 3D models were examined using an open-source software. Both model shapes were similar to the real fruit. However, they were slightly different from each other. The Kinect V2 model had a more uneven shape because of noise from the TOF sensor. The accuracy of the fruit volume estimate and the largest diameter was 93% and 86% respectively, with Kinect V2. Liu et al. [4] identified different kind of fruits and leaves using RealSense F200 and Kinect V1 depth sensors. The RealSense F200 has a color and an infrared cameras and an infrared laser projector. 120 depth data samples were collected from one plant, placed in 64 different positions modified manually and from different angles. Results showed that little occlusion and low adhesion brought the fruit recognition rate up to 80–100%. Different species with occlusions had a lower detection rate.

3 Materials and Methods

Two experiments were carried out to produce a three-dimensional point cloud of plants. The first experiment was done with a Kinect V1 sensor and the second was based on a 2D LiDAR device. The basic principle of both scanners is optical depth measurement. The three-dimensional modeling system consisted of a depth sensor placed on a fixed tripod, a rotating disk moved by a stepper motor, a drive motor, a power source, and a computer used to control it, as shown in Fig. 1a. The modeling methods were written in Python 3.7 language on an Intel Core i5 7th Gen 2.5 GHz with 8 GB of RAM to control the turntable under Ubuntu 16.04 OS and using Open3d [17] library for 3D point cloud registration process. The computer system was connected via USB to an Arduino Nano, which controlled a stepper motor through a driver V44A3967. The disk angle was estimated with an open-loop counter algorithm, which returned the position to the master system every time it moved. Geometrical information from the depth sensor (LiDAR 2D or Kinect) was acquired using a ROS (Robot Operating System) environment with a Kinetic version, once the initial and final angles and angular steps were configured. The obtained point clouds were stored in LASer format and visualized using the free access software CloudCompare. The stepper motor used in the rotating platform shown in Fig. 1b had a torque of 9.4 kg/cm, a gearbox with a ratio of 100:1 and an angular speed between 1.2 to 3.6 RPM. The motor, which was adapted to a mechanism by a worm, had the capacity to move every 0.36°. The worm wheel had 26 teeth and a transmission ratio of 36:1, which increased the torque system and angular resolution, resulting in the following final mechanical characteristics: torque of 300 kg/cm, a gearbox of 3600:1 and the ability to move the disc every 0.01°. However, the final angular speed was reduced to a range of 0.033–0.1 RPM. The turntable was modeled in SolidWorks 2018. For the experimental setup a plant of the species *Dracaena braunii* was scanned, with an approximate height of 0.8 m measured from the base of the pot to the top of the plant, in a controlled environment using white artificial light.

3.1 LiDAR Modeling

This experiment was based on a 2D LiDAR sensor URG-04LX-UG01, one of the simplest of the manufacturer Hokuyo. Basically, the laser emits an infrared beam on a rotating mirror, which changes its direction, illuminating a specific region of the scene. The reflected light is then used to determine the distance to the target. The main specifications of the scanner are shown in Table 1. The acquisition protocol and the software used were based on an earlier version developed by Murcia et al. [6]. They presented a methodology for the calibration and reconstruction process of the same sensor and motor unit. However, in this work, a new kinematic model was established for the developed platform system.

Kinematic Model. The purpose of the kinematic model is to acquire a 3D initial coordinate P_i to a reference coordinate frame P_o located in the center of

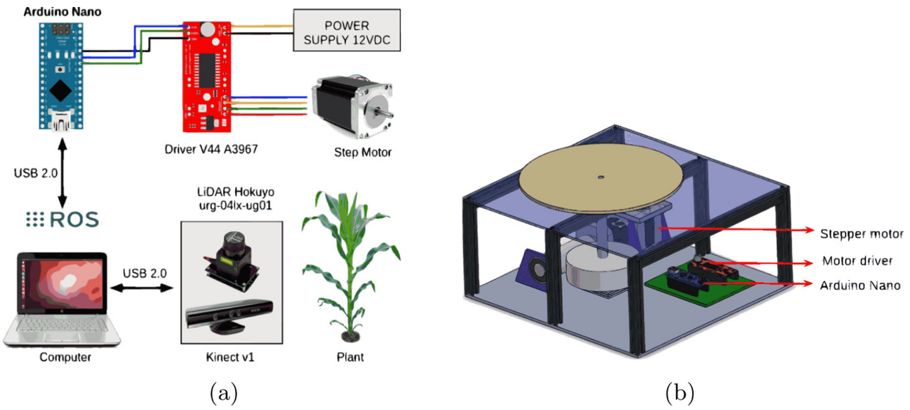


Fig. 1. Main components of the scanning system: (a) functional diagram; (b) representation of the used rotating platform and its main components.

Table 1. Main specifications of the 2D LiDAR sensor URG-04LX-UG01.

Feature	Hokuyo URG-04LX-UG01
Measurement distance	20 to 5600 mm
Resolution	1 mm
Scan angle	240°
Angular resolution	0.36° (360°/1024 steps)
Accuracy	± 30 mm (For distance above < 10000 mm)
Scanning time	100 ms/scan
Power source	5 VDC ± 5% (USB Bus Power)

the disk for relative LiDAR detection. The final 3D point cloud in P_o is represented in a XYZ space as a function of LiDAR horizontal angle θ and radial distance γ, as well as the disk angle β and system constant parameters, which represent the distances between the sensor and the target. The transformation matrix $T_p = R_z * T$ in 3D space was obtained using homogeneous coordinates by means of a $[4 \times 4]$ dimensional matrix. Where T is a translation transformation with three parameters t_x, t_y and t_z, which represent the respective translations of P_o along X, Y, Z axes and Rz is a rotation matrix around Z axis (yaw). Each one of these matrices is described below. Thus, the final XYZ reconstruction of the studied plant was calculated with a three-frame transformation based on T_p as shown below:

$$T = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}, R_z = \begin{bmatrix} cos(\gamma) & -sin(\gamma) & 0 & 0 \\ sin(\gamma) & cos(\gamma) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1}$$

$$P_o = T * R_z * P_i \Rightarrow P_i = \begin{bmatrix} r * cos(\theta) \\ 0 \\ r * sin(\theta) \\ 1 \end{bmatrix} \quad (2)$$

P_i is represented as an input XYZ matrix with dimensions [4xm], where m is the number of samples or points. Where r is a range vector of m samples obtained from a LiDAR ROS node, which represent the measurement of each LiDAR angle θ_i.

3.2 Kinect Modeling

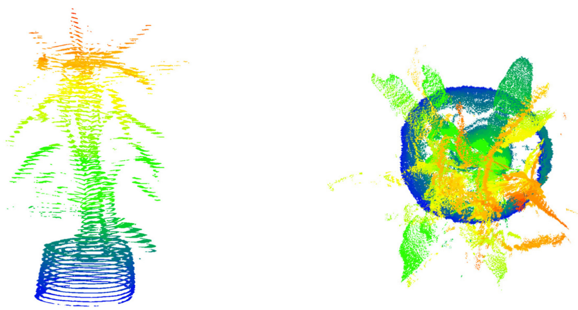
The main technical characteristics of the Kinect sensor are the resolutions of its color and depth cameras: 320×240 pixels and 640×480 pixels respectively. The TOF camera has a depth range of 0.5 to 3.5 m. It was mounted at a height of 1.5 m with a horizontal view at an angle of less than 27°. The camera's horizontal and vertical fields of view are 57° and 43° respectively and the sampling rate is 30 fps. The Kinect sensor was mounted on the same tripod used by the LiDAR platform. The data capture process took approximately 1 h and 40 min and 361 point clouds were obtained.

Point Cloud pre-processing. Initially, the algorithm pre-processed the data in four stages: Down-sampling, statistical removal of points, outliers removal, and estimation of normal vectors. Voxel downsampling function uses a regular voxel mesh to create a uniformly reduced point cloud. The algorithm works in two steps: Points are grouped into voxels. Each occupied voxel generates an exact point by averaging all the points within it. The statistical outlier removal function is a filter that eliminates points that are farther away from their neighbors compared to the point cloud average. A function of the standard deviation of those average distances is established as a threshold level. The lower this number, the more aggressive the filter will be. The radius outlier removal function is a filter that eliminates points that have few neighbors on a given sphere around them, by setting the minimum number of points on the sphere and the radius of the sphere that will be used to count the neighbors. Finally, the function to estimate normals finds adjacent points and calculates the principal axis of the adjacent points using analysis of covariance.

ICP Algorithm. Once the point clouds were pre-processed, they were merged into a single one, also unifying the framework of reference. For this purpose, the point-to-plane ICP (Iterative Closest Points) algorithm was used, which minimizes the Root-Mean-Square-Error RMSE between the transformed point clouds. After fusion, new outliers may appear. To remove them, a point optimization method was performed based on the neighboring nodes and the voxel size. This step also reduces possible false alignments between edges and, avoid duplication and excess points. The ICP algorithm requires a number of parameters to be tuned. The setting values found are: The voxel size for sample reduction should be 0.001 or less to improve performance. The threshold of the recording edge should be 0.006 or less, the search radius equal to 0.5 cm and maximum of neighbors equal to 100. The number of iterations of the fixed ICP in 5 for greater efficiency. Although the point cloud range is not required for processing, it is included to determine the fitting error between the point clouds.

4 Results

Figure 2 shows a frontal and top view of the resulting 3D point cloud. 3D representation had a density of 558639 points and the whole acquisition and processing took about 1.25 h, with no color information. Figures 3a and 3b show a frontal view of the resulting 3D point cloud. Unlike the reconstruction done by LiDAR, the Kinect includes dimensions of color information (Red, Green and Blue, RGB) in the acquired data. The final 3D reconstruction had 2230527 points. The registration algorithm took 45 min meanwhile the complete procedure took about 2.5 h.



(a) (b)

Fig. 2. Reconstructed 3D plant point cloud with Hokuyo URG-04LX-UG01 using a color scale based on altitude: a) side view of point cloud, b) top view of point cloud.

4.1 3D Modeling Comparison

Both methods were tested in the same laboratory conditions using the same rotating platform and acquisition software. The comparative results are summarized in Table 2. As was expected, experiment carried out with Kinect presented a higher point density and processing time regarding the procedure with LiDAR. A second comparison was performed to determine an error estimation in four measurements of the plant called A, B, C, D as shown in Fig. 3c according to Eq. 3.

(a) (b) (c)

Fig. 3. (a) side view of reconstructed 3D point cloud with Kinect using a color scale based on altitude. (b) side view of reconstructed 3D point cloud with Kinect using color information. (c) Illustration of manual plant measurements: (A) Height of the first stem. (B) Diameter of the second stem, (C) Diameter of the first stem (D) Pot diameter.

Table 2. Comparative features of generated point clouds.

Feature	Kinect V1	URG-04LX-UG01
Number of points	2230527	558639
Acquisition time [s]	6000	4530
Processing time [s]	2735	4
Color	Yes	No

$$error[\%] = \left| \frac{X - W}{W} \right| \tag{3}$$

where X is the software measurement from 3D point cloud using point picking tool in CloudCompare and W is the real measurement obtained in the laboratory, called reference. Table 3 shows the measurement obtained from each point cloud and the error calculated in each case.

As can be observed on Table 2, the acquisition and processing time was almost double in the Kinect-based system compared to the LiDAR one. This was mainly due to the fact that the Kinect takes longer to acquire the data, in addition to simultaneously obtaining the color information. In this case, the number of points in the acquired point clouds was exactly four times higher than that obtained with LiDAR. The acquisition rates measured in points per second for the Kinect and LiDAR systems were 255 points/s and 123 points/s respectively. Thus, the Kinect, despite requiring more time, was more than twice as fast as LiDAR in acquiring relevant plant data. According to the results in Table 3, the Kinect-based system had the highest accuracy, as the error found in

Table 3. Error calculation from point clouds and manual measurements.

Measurements		Kinect V1		URG-04LX-UG01	
Letter	Ref. [mm]	meas. [mm]	Error[%]	meas. [mm]	Error [%]
A	280	308.54	10.19	246.83	23.86
B	14	16.79	19.92	23.09	64.92
C	15	15.55	3.66	22.99	53.26
D	131	117.3	10.45	210.13	60.40

the plant parameter measurements was lower than with LiDAR. In both cases, the largest error occurred with the B parameter. This error was due to the characteristic occlusions of the plant architecture making it difficult to correctly acquire the points in that particular part of the plant.

5 Conclusions

In this paper, two low-cost three-dimensional modeling of the plant based on rotating platform and depth sensors are presented. Two point cloud reconstruction experiments based on 2D LiDAR and Kinect V1 were performed. The Kinect procedure presented better results, in terms of point cloud density and rendering quality, due to the color information used. However, acquisition and processing times were longer than those of 2D LiDAR. Results using Hokuyo URG-04LX-UG01 showed a point cloud acquisition without color or intensity information that required less time. Employing the open source software CloudCompare, the 3D point clouds obtained with LiDAR were filtered to reduce noise. On the other hand, the lowest error was achieved with the Kinect sensor, so it turned out to be the most suitable system for accuracy. In addition, the LiDAR stage required a kinematic model, so a characterization of the data was necessary to determine the model parameters, which could introduce distortions in the point clouds if they were far from the real values. The acquisition data and reconstruction codes, as well as the LAS files obtained, are available online at https://github.com/HaroldMurcia/plant_reconstruction.git. As future work, the improvement of the 3D reconstruction results using the LiDAR system by introducing a new sensor with better technical characteristics, such as divergence angle, resolution, signal-to-noise ratio and fusion techniques between digital cameras and LiDAR, is proposed.

Acknowledgment. This work was funded by the OMICAS program: Optimización Multiescala In-silico de Cultivos Agrícolas Sostenibles (Infraestructura y validación en Arroz y Caña de Azúcar), anchored at the Pontificia Universidad Javeriana in Cali and funded within the Colombian Scientific Ecosystem by The World Bank, the Colombian Ministry of Science, Technology and Innovation, the Colombian Ministry of Education, the Colombian Ministry of Industry and Tourism, and ICETEX, under grant FP44842-217-2018 and OMICAS Award ID: 792-61187.

References

1. Corti, A., Giancola, S., Mainetti, G., Sala, R.: A metrological characterization of the Kinect V2 time-of-flight camera. Robot. Autonomous Syst. **75**, 584–594 (2016)
2. Guo, Q., et al.: Crop 3D—a LiDAR based platform for 3D high-throughput crop phenotyping. Sci. China Life Sci. **61**, 328–339 (2018)
3. Li, D., et al.: An overlapping-free leaf segmentation method for plant point clouds. IEEE Access **7**, 129054–129070 (2019)
4. Liu, J., et al.: Experiments and analysis of close-shot identification of on-branch citrus fruit with realsense. Sensors **18**(5) (2018)
5. McCormick, R.F., Truong, S.K., Mullet, J.E.: 3D sorghum reconstructions from depth images identify QTL regulating shoot architecture. Plant Physiol. **172**(2), 823–834 (2016)
6. Murcia, H.F., Monroy, M.F., Mora, L.F.: 3D scene reconstruction based on a 2D moving LiDAR. In: Florez, H., Diaz, C., Chavarriaga, J. (eds.) ICAI 2018. CCIS, vol. 942, pp. 295–308. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01535-0_22
7. Panjvani, K., Dinh, A.V., Wahid, K.A.: LiDARPheno - a low-cost LiDAR-based 3D scanning system for leaf morphological trait extraction. Front. Plant Sci. (2019)
8. Paulus, S.: Measuring crops in 3D: using geometry for plant phenotyping (2019)
9. Rosell-Polo, J.R., et al.: Advances in structured light sensors applications in precision agriculture and livestock farming. Adv. Agron. **133**, 71–112 (2015)
10. Saeys, W., et al.: Estimation of the crop density of small grains using LiDAR sensors. Biosyst. Eng. (2009)
11. Santos, T.T., Koenigkan, L.V., Barbedo, J.G.A., Rodrigues, G.C.: 3D plant modeling: Localization, mapping and segmentation for plant phenotyping using a single hand-held camera. In: Agapito, L., Bronstein, M., Rother, C. (eds.) ECCV 2014. LNCS, vol. 8928. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-319-16220-1_18
12. Spoliansky, R., et al.: Development of automatic body condition scoring using a low-cost 3-dimensional kinect camera. J. Dairy Sci. (2016)
13. Thapa, S., Zhu, F., Walia, H., Yu, H., Ge, Y.: A novel LiDAR-Based instrument for high-throughput, 3D measurement of morphological traits in maize and sorghum. Sensors (Switzerland) (2018)
14. Wang, H., Lin, Y., Wang, Z., Yao, Y., Zhang, Y., Wu, L.: Validation of a low-cost 2D laser scanner in development of a more-affordable mobile terrestrial proximal sensing system for 3D plant structure phenotyping in indoor environment. Comput. Electron. Agric. **140**, 180–189 (2017)
15. Yamamoto, S., et al.: 3D reconstruction of apple fruits using consumer-grade RGB-depth sensor. Eng. Agric. Environ. Food **11**(4), 159–168 (2018)
16. Zhang, L., Grift, T.E.: A LIDAR-based crop height measurement system for Miscanthus giganteus. Comput. Electron. Agric. **85**, 70–76 (2012)
17. Zhou, Q.Y., Park, J., Koltun, V.: Open3D: A modern library for 3D data processing. arXiv:1801.09847 (2018)