

MEASURING NUT VOLUMES USING THE AZURE KINECT

Xiong Peng¹, Ting On Chan^{1,2*}

¹ School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China;

² Guangdong Provincial Key Laboratory of Urbanization and Geo-Simulation, Sun Yat-sen University, Guangzhou 510275, China.

KEYWORDS: Azure Kinect, Nuts volume estimation, Geometric model, Single nut segmentation, SBF algorithm.

ABSTRACT:

Rapid and accurate nut count and volume measurement techniques are critical for nut production and its automation. In this work, a three-dimensional (3D) depth camera, the Azure Kinect, was used to measure the volume of the nuts, along with our proposed point cloud processing framework. A group of nut16rowns was first collected as point clouds using the Azure Kinect, and then the point nearest neighbor (PNN) algorithm was used to perform fast nut single segmentation. After that, noise caused by the multi-path effect was identified from the collected point clouds. Most of the noise was eliminated using the surface boundary filter (SBF) algorithm. After the preprocessing, the volumes of the nut point cloud were estimated using the least-squares ellipsoid fitting (LSEF). We selected three different nuts in the experiment: 15 each of walnuts, pecans, and macadamia nuts. Their volumes were first measured by the Azure Kinect, and then by the water displacement method (WDM) as control. The results show that the proposed measurement setup can accurately count nuts and the average nuts volume estimation accuracy is 92.1% when compared to the reference volumes. The Azure Kinect can effectively count the number of nuts and accurately estimate the volume of nuts of different types and sizes, based on our proposed framework.

1. INTRODUCTION

The global nuts market size was valued at USD 54.11 billion in 2021 and is projected to expand at a compound annual growth rate (CAGR) of 4.6% from 2022 to 2028 (INC, 2022), as the growing awareness on the health benefits among consumers is promoting the market. After the harvest of the nut, nut count and volume measurement are performed for sales. In addition, the quantitative assessment of the morphological traits of nuts is important to agricultural and botanic research (Wang and Nguang, 2007; Gharibzadeh et al., 2012), including genetics, physiology, and plant breeding. Therefore, low-cost, user-friendly, non-contact nut count and nut volume measurement systems are always desired by many individuals.

Traditionally, the grading and quality assessment of nuts are done manually (Costa et al., 2021; Solak and Altinişik, 2018), so they are inefficient and costly. Besides, nuts are small, so nuts making manual measurements of external morphological characteristics is time-consuming and laborious.

Based on the computed tomography (CT), Bernard et al. (2020) proposed a method to simultaneously measure different morphological characteristics of several walnuts, such as nut length, nut surface, and nut outline diameter. In this method, the features of different parts of the walnut were extracted and quantified by using CT images to reconstruct three-dimensional (3D) models. However, the walnuts need to be placed in a specific position. Arendse et al. (2016) used CT to characterize and quantify the internal structure of pomegranate fruit. Because the bench CT scanner sold on the market are usually designed for medical and industrial use. Their turntable is not suitable for scanning small objects. In addition, CT machines are expensive, and it is hard for them to be widely applied for nut production.

Omid et al. (2010) proposed a method to reconstruct a three-dimensional (3D) model of fruit using two cameras. Fruit volumes were calculated by dividing the 3D model into several elliptical cylinders. The volume estimated by this method is in good agreement with the actual volume measured by the water displacement method (WDM). The R-square of lemon, lime, orange, and tangerine was 0.962, 0.970, 0.985, and 0.959,

respectively. The above technology mainly uses two-dimensional (2D) images to reconstruct 3D models, but this process often requires multiple cameras to shoot the targets in different fields of view.

In this paper, we propose a point cloud processing framework that is executed along with the Azure Kinect 3D camera operation to count the nuts and measure their sizes. The paper was divided into three sections. The materials and methods section focuses on the proposed processing framework and related algorithm, the experiments section provides experimental design, and the result section shows the outcome of our method and the WDM.

2. MATERIALS AND METHODS

2.1 Proposed Measurement System Overview

As seen in Figure 1, the measurement prototype consists of a Microsoft Azure Kinect camera, a camera tripod, a stage for placing the nuts, and a laptop computer.



Figure 1. Proposed measurement setup.

In the following subsections, we focus on the Azure Kinect functioning and the proposed processing framework, which are the components of the measurement system.

2.2 Microsoft Azure Kinect

The Microsoft Azure Kinect represents the latest generation of Microsoft's Kinect depth camera (Kurillo et al., 2022). In Figure 2, the Azure Kinect hardware includes a color camera, a ToF depth sensor, the gyroscope and accelerometer, and a circular array of seven microphones. In addition, the software supports Windows and Linux platforms. The Azure Kinect supports a variety of different operating modes. The maximum field of view of the depth camera is $120^\circ \times 120^\circ$ WFOV mode and $75^\circ \times 65^\circ$ NFOV mode. The operating range in WFOV mode is 0.25 to 2.21 m (0.25 to 2.88 m), while the operating range in NFOV mode is 0.50 to 3.86 m (0.50 to 5.46 m). If the measurement range of the depth camera is 1-2 m, the range error is less than 2 mm. In addition, the Azure Kinect supports many applications. For 18 rows 18 biomedical engineering (Yoshimoto and Shinya, 2022; Shamim et al., 2022), human posture control (Antico et al., 2021), and poultry product measurement (Chan et al., 2018), etc, are built with the Kinect.



Figure 2. Microsoft Azure Kinect

According to previous research (Tölgýessy et al., 2021), the Azure Kinect camera has a lower standard deviation for the measurement at pixel level when using NFOV Binned mode compared to the other three operation modes (NFOV Unbinned, WFOV Binned, WFOV Unbinned), so we adopt NFOV binned mode in our experiment (Tölgýessy et al., 2021). Azure Kinect depth data does not carry RGB information. The depth data and RGB data need to be registered first to generate the colorized

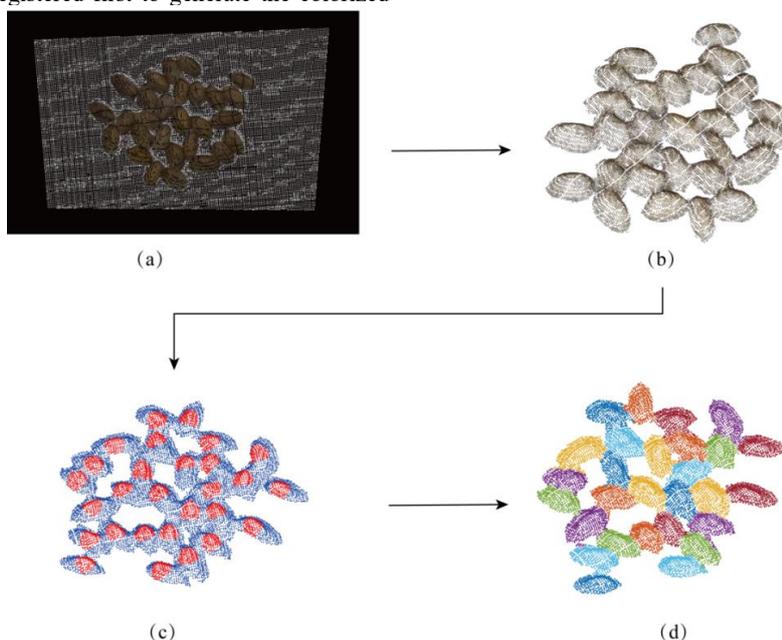


Figure 3. Proposed single nut segmentation algorithm. (a) Raw point cloud; (b) Background point cloud removal; (c) The highest

point clouds.

2.3 Proposed Processing Framework

2.3.1 The Point Nearest Neighbor (PNN) Algorithm

In the processing framework, a novel segmentation algorithm for the nut point cloud is proposed, which is built upon the segmentation algorithm for point clouds of individual tree crowns (Li et al., 2012). The algorithm segmented nuts into individual point cloud by considering the relative spacing between nuts. In general, the spacing between the top of the two nuts will be larger than that between the bottom (Li et al., 2012). The algorithm has the following steps: Step (1), extract multiple nut point clouds from the white stage point cloud based on the color threshold method in Figure 3(a) and (b); Step (2), set the highest value within the threshold radius as the vertex of a single nut point cloud, as shown in Figure 3(c). In Figure 4, Points A B are the highest and hence they are treated as the top of Nut #1 and Nut #2. All the nut point cloud vertices will be extracted in this Step; Step (3), points will be correctly assigned to their corresponding nut clusters by comparing the distance between the points and the classified points. In Figure 4, when $d_{AC} < d_{BC}$ is satisfied, Point C is correctly assigned to Nut #1. After that, Point D will be classified as Nut #2 by comparing to Points B and C. Finally, all the nut are correctly segmented as individual nuts, as shown in Figure 3(d).

2.3.2 The Surface Boundary Filter (SBF) Algorithm

The surface boundary filter (SBF) algorithm (Li et al., 2019) is used to remove the nut boundary noise point cloud, which is similar to the 2D image erosion. It is a 3D surface boundary point extraction algorithm based on principal component analysis (PCA). This algorithm was originally used to extract the edge points of overlapping areas of plant leaves (Li et al., 2019). The noise caused by the multipath effect can be removed by this algorithm.

point of a single Pecan; (d) Segmentation of individual Pecan.

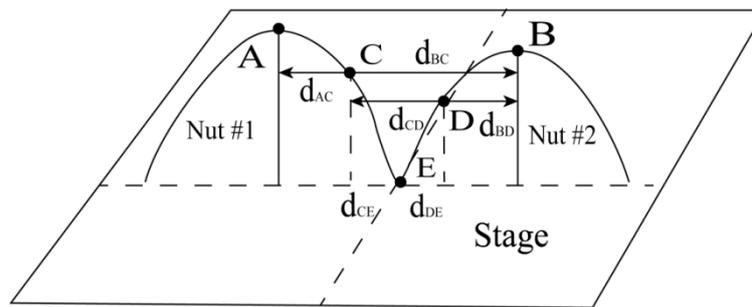


Figure 4. The spacing between points on two nuts.

2.3.3 Least Squares Ellipsoid Fitting

As seen in Figure 5, although different types of nuts vary in size and shape, their outlines are similar to ellipsoids. Therefore, we considered using the ellipsoid to fit the point clouds captured by the camera. An ellipsoid is defined in terms of center coordinates (x_c, y_c, z_c) , semi-axes parameters (a, b, c) , and rotation angles $(\theta_1, \theta_2, \theta_3)$ about the x,y,z-axes. These rotation angles are treated as the orientation of the ellipsoid (Turner et al., 1999). The ellipsoid model can be expressed as:

$$\mathbf{R}_1 = \begin{bmatrix} \cos\theta_1 & \sin\theta_1 & 0 \\ -\sin\theta_1 & \cos\theta_1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{R}_2 = \begin{bmatrix} \cos\theta_2 & 0 & \sin\theta_2 \\ 0 & 1 & 0 \\ -\sin\theta_2 & 0 & \cos\theta_2 \end{bmatrix}, \text{ and } \mathbf{R}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_3 & \sin\theta_3 \\ 0 & -\sin\theta_3 & \cos\theta_3 \end{bmatrix} \quad (2)$$

$\mathbf{R}_1, \mathbf{R}_2$ and \mathbf{R}_3 are the rotation matrices about the X, Y, and Z-axes, respectively. (x, y, z) is the 3D coordinates of the point cloud captured by the camera. We adopt the Gauss-Newton least-squares algorithm (Gill et al., 1981) to determine estimate ellipsoid parameters, $\hat{E} = [\hat{x}_c, \hat{y}_c, \hat{z}_c, \hat{a}, \hat{b}, \hat{c}, \hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3]^T$ using.

As seen in Figure 6, S is the section area of the ellipsoid perpendicular to the major semi-axis. once the ellipsoid parameters a, b , and c are estimated, volume (V) of the ellipsoid can be computed as follows:

$$V = \int_{-a}^a S(x) dx = \frac{3}{4} \pi abc \quad (3)$$

where

$$S(x) = \pi bc \left(1 - \frac{x^2}{a^2}\right) \quad (4)$$

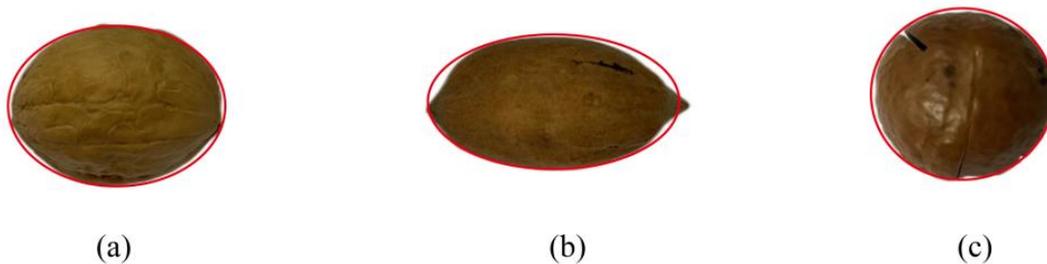


Figure 5. Different types of nut shapes and the ellipse (Red). (a) Walnut; (b) Pecan; (c) Macadamia nut.

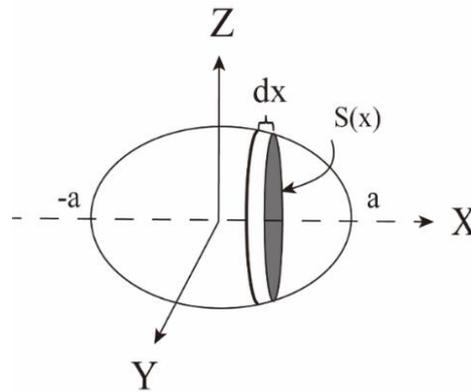


Figure 6. Schematic diagram of the volume of an ellipsoid.

3. EXPERIMENTS

3.1 Three experiments were performed: two preliminary experiments and one main experiment. The preliminary experiment was performed to test the camera temporal stability and the optimal capturing distance, and the main experiment was performed to estimate the volume of three different nuts according to the proposed algorithm framework. Temporal Stability of the Azure Kinect

Before the nut volume measurement, the temporal stability of the Azure Kinect was first examined (Chan and Lichti, 2015; Kurillo et al., 2022; Tölgyessy et al., 2021). As seen in Figure 6(a), the Azure Kinect was used to capture a spherical target (with a radius of 7.25 cm) for 6 hours inside a room with controlled environmental conditions (e.g., regulated temperature). The center of the spherical target was estimated by using the sphere fitting for the analysis. The Kinect was placed approximately 1.8

m away from the spherical target.

3.2 Capturing Distance Optimization

In general, the range errors of the point cloud captured by the Azure Kinect increase with distances. However, if the Azure Kinect is too close to the target, the top point cloud of the target will be destroyed due to the scattering effect. Therefore, we quantified the distance from the Azure Kinect to the target, which ensure the collection of a complete point cloud with the highest accuracy. A table tennis ball (with a radius of 20 mm) was captured with the Azure Kinect at camera-to-ball distances of 0.4 m, 0.5 m, 0.6 m, and 0.7 m, as illustrated in Figure 6(b). The point clouds were then fitted to spherical model with the least-squares. This experiment aims at optimizing the distance from the camera to a small target (Chan et al., 2018).



(a)



(b)

Figure 7. Preliminary experiments. (a) Measurement setup for temporal stability analysis; (b) The capturing distance optimization

3.3 Volume Measurement of Nuts

Three kinds of nuts were selected in this experiment: walnut, pecan, and macadamia nut. For each kind of nuts, we placed 15

individuals on the stage. Multiple nuts are captured at three different camera-to-nut distances of 0.4 m, 0.5 m, and 0.6 m. As seen in Figure 8, as a control experiment, the volume of each nut was measured using the WDM. The error of the WDM was assumed to be ± 1 ml.



Figure 8. Set up for the nut with the WDM.

4. RESULTS AND DISCUSSIONS

4.1 Analysis of the Temporal Stability of the Azure Kinect

The center of the spherical target was estimated by using the sphere fitting. The result shows that the X and Y coordinates of the target center tend to be stable over time. From Figure 9, it is known that the X coordinate is increased by only 0.3 mm from the initial coordinate and the Y coordinate has increased by 0.45 mm within 6 hours of the camera working. During that six-hours, the Z coordinate of the target center of the sphere fluctuated between 778.03 mm and 778.3 mm (in a range of ~ 0.3 mm). Therefore, the Y coordinate of the camera is relatively the most erroneous. As a result, the Azure Kinect generates highly accurate measurement with high temporal stability. Since the

errors are less than a half of a millimeter over six hours, no warm-up time is strictly needed for most applications.

4.2 Accuracy Analysis versus Capturing Distance

As illustrated in Figure 10, Kinect suffers from the multi-path effect, this work adopts the SBF algorithm to eliminate the edge noise of the table tennis point clouds. After preprocessing, the estimation radii of the spheres are 19.3 mm, 19.2 mm, 18.9 mm, and 18.6 mm at 0.4 m, 0.5 m, 0.6 m, and 0.7 m, respectively. Overall, the volume of the table tennis ball (with known size) can be estimated accurately at a camera-to-ball distance between approximately 40 cm to 60 cm, in which 40 cm is recommended in the setup to obtain the highest radii accuracy.

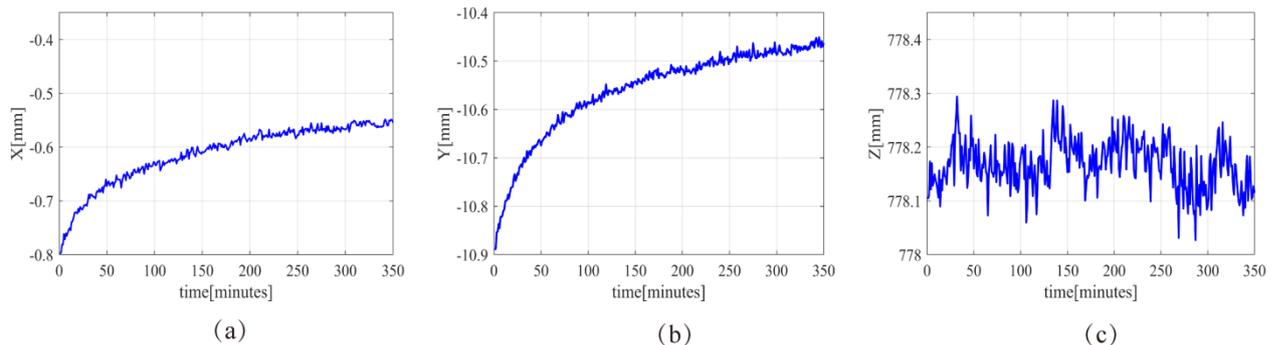


Figure 9. Three-dimensional coordinates of the center of the table tennis ball point cloud using the sphere fitting (a) X coordinate; (b) Y coordinate; (c) Z coordinate.

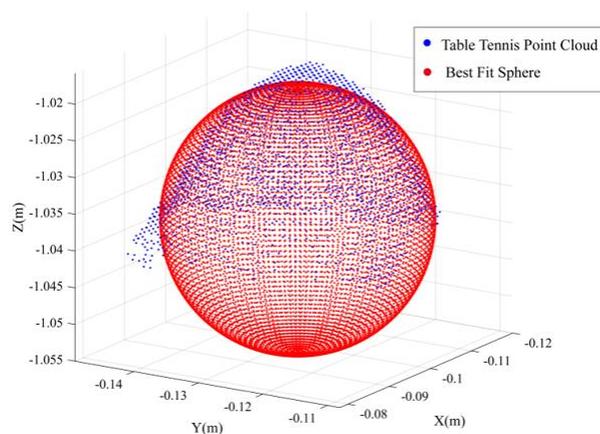


Figure 10. Best fit sphere (Red) to the table tennis ball point cloud (Blue).

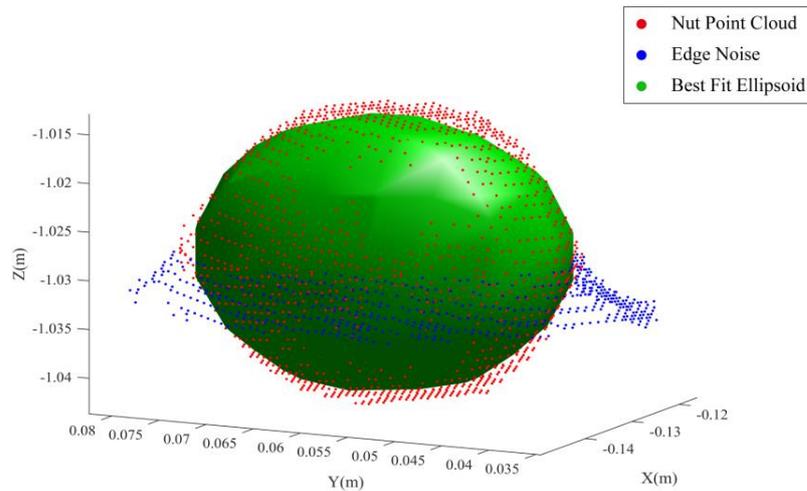


Figure 11. Best fit ellipsoid to the pecan point cloud: denoised nut point cloud (Red), edge noise (Blue), and the best fit ellipsoid model (Green).

Table 1. Comparison of the estimated volume of nuts with SBF algorithm and without SBF algorithm

Nuts	Label	Without SBF/Vol (ml)	With SBF/Vol (ml)	WDM Vol (ml)
Walnut	W1	26.7	22.1	21.5
	W2	35.9	27.7	28.6
Pecan	P1	19.2	12.2	12.5
	P2	25.9	15.7	14.8
Macadamia nut	M1	16.5	9.2	9.5
	M2	19.8	11.5	11.2

4.3 Accuracy Analysis of Estimating Volumes of Nuts with the least-squares fitting

As seen in Figure 11, the point cloud of a nut is fitted with the ellipsoid model (green) with the least-squares method, as well as the blue point cloud is the noise point on the edge, and the denoised nut point cloud (red). Table 1 shows the estimated volumes. As can be seen, the volume contributed by the SBF algorithm is significantly higher than that without removing edge noise points. In addition, the lower half of the point cloud is simulated assuming the nuts are symmetric to enhance the geometry of the observations for the fitting. This method can fully realize the measurement of the number of nuts in Table 2. The comparison results between the proposed method and the

WDM for different types of nuts are shown in Figure 12 (Huynh et al., 2020). The R-square of walnut at 0.4 m, 0.5 m, and 0.6 m height were 0.905, 0.9, and 0.9, respectively. The R-square of Pecan at 0.4 m, 0.5 m, and 0.6 m height were 0.911, 0.902, and 0.860, respectively. The R-square of Macadamia nut at 0.4 m, 0.5 m, and 0.6 m height were 0.925, 0.916, and 0.879, respectively. The R^2 can be interpreted as the proportion of variance accounted for with the actually measured variance in the fit. The higher the R^2 , the closer the estimated volume is to the measured volume by the WDM. Furthermore, the lower the height, the higher the estimation accuracy. It can be seen that up to approximately an average of 92.1% accuracy was achieved.

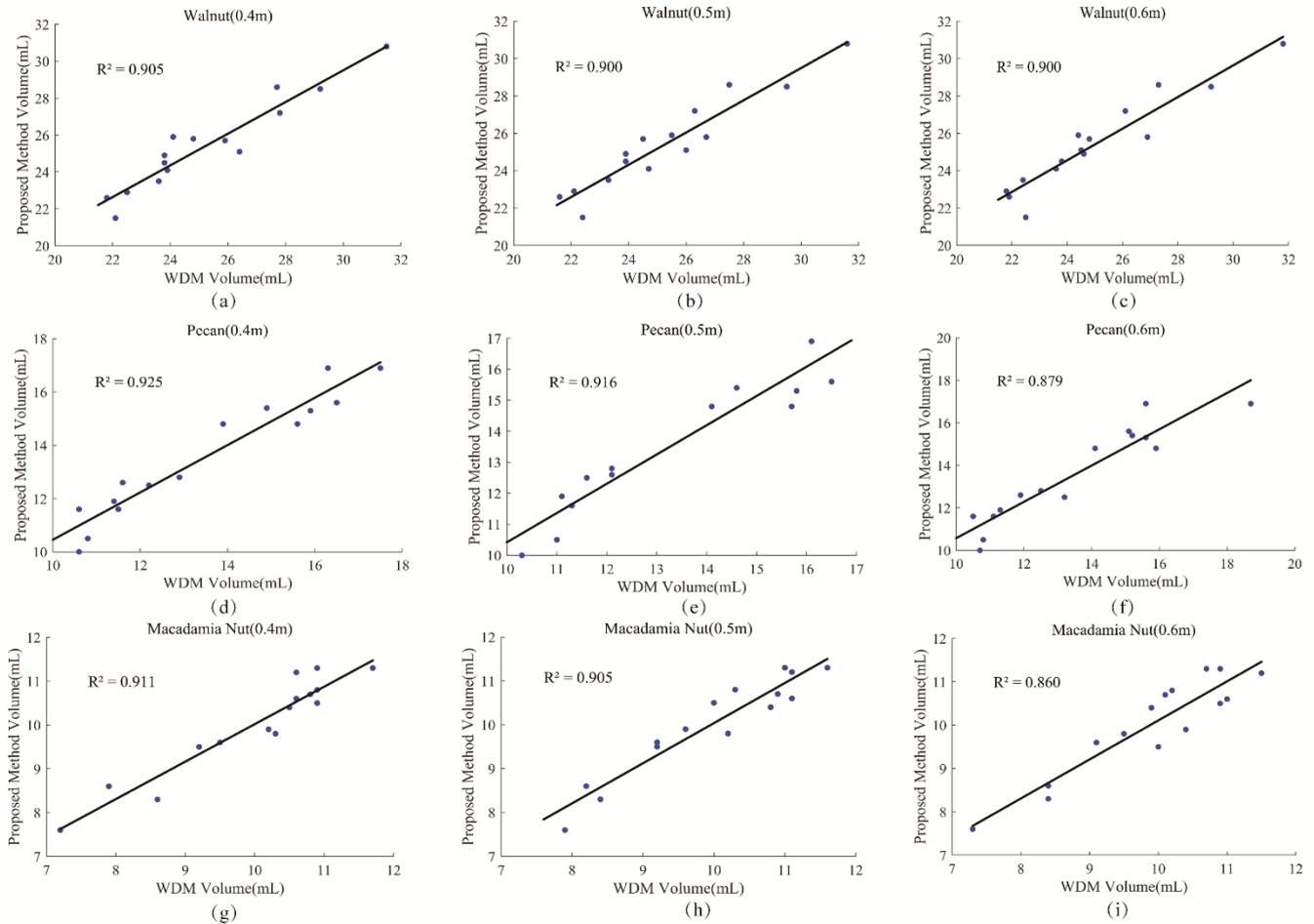


Figure 12. Comparison of estimated and measured volume of various nuts with the proposed method and the WDM at different heights. (a-c) Comparison of estimated and measured volume of Walnuts at the heights of 0.4 m, 0.5 m, and 0.6 m; (d-f) Comparison of estimated and measured volume of Pecans at the heights of 0.4 m, 0.5 m, and 0.6 m; (g-i) Comparison of estimated and measured volume of Macadamia nuts at the heights of 0.4 m, 0.5 m, and 0.6 m.

Table 2. Summary of the volume accuracy of the nuts.

Nuts	Total	Detected	Height (m)	R ²	Accuracy (%)
Walnut	48	48	0.4	0.905	95.1
			0.5	0.900	94.2
			0.6	0.900	93.6
Pecan	28	28	0.4	0.925	92.2
			0.5	0.916	91.5
			0.6	0.879	90.3
Macadamia nut	32	32	0.4	0.911	95.2
			0.5	0.902	95.1
			0.6	0.860	92.1

5. CONCLUSIONS

In this study, we proposed a cost-effective and efficient point cloud processing framework for counting the number of nuts and estimating the volume of a single nut, based on the Azure Kinect 3D camera. In the proposed algorithm processing framework, multiple nut point clouds were first segmented individually by PNN algorithm. Next, the noise was eliminated by SBF algorithm, and then the other halves of the point clouds collected by the camera were simulated. The volume was further estimated by the least-squares method. The proposed framework has a high degree of automation, which can be easily developed

into a real-time nut count and volume measurement system. The WDM in this study is used to verify the effectiveness of the method. In addition, we also considered the optimal camera-to-nut distance to achieve the highest accuracy. Experimental results showed the average volume estimation accuracy of this method was 92.1%. The framework is shown to be able to enhance the efficiency and accuracy of nut evaluation and eventually benefit the nut industries.

ACKNOWLEDGEMENTS

This research is supported by the Science and Technology Program of Guangzhou (No. 202201011686), China

REFERENCES:

- Antico, M. et al., 2021. Postural control assessment via Microsoft Azure Kinect DK: An evaluation study. *Computer Methods and Programs in Biomedicine*, 209: 106324.
- Arendse, E., Fawole, O.A., Magwaza, L.S. and Opara, U.L., 2016. Non-destructive characterization and volume estimation of pomegranate fruit external and internal morphological fractions using X-ray computed tomography. *Journal of Food Engineering*, 186: 42-49.
- Bernard, A., Hamdy, S., Le Corre, L., Dirlwanger, E. and Lheureux, F., 2020. 3D characterization of walnut morphological traits using X-ray computed tomography. *Plant Methods*, 16(1).
- Chan, T. et al., 2018. An Egg Volume Measurement System Based on the Microsoft Kinect. *Sensors*, 18(8): 2454.
- Chan, T. and Lichti, D., 2015. Automatic In Situ Calibration of a Spinning Beam LiDAR System in Static and Kinematic Modes. *Remote Sensing*, 7(8): 10480-10500.
- Costa, L. et al., 2021. Measuring pecan nut growth utilizing machine vision and deep learning for the better understanding of the fruit growth curve. *Computers and Electronics in Agriculture*, 181: 105964.
- Gharibzahedi, S.M.T., Mousavi, S.M., Hamed, M. and Khodaiyan, F., 2012. Comparative analysis of new Persian walnut cultivars: nut/kernel geometrical, gravimetric, frictional and mechanical attributes and kernel chemical composition. *Scientia Horticulturae*, 135: 202-209.
- Gill, P.E., Murray, W. and Wright, M.H., 1981. *Practical optimization*.
- Huynh, T., Tran, L. and Dao, S., 2020. Real-Time Size and Mass Estimation of Slender Axi-Symmetric Fruit/Vegetable Using a Single Top View Image. *Sensors*, 20(18): 5406.
- INC, 2022. *NUTS & DRIED FRUITS STATISTICAL YEARBOOK*.
- Kurillo, G., Hemingway, E., Cheng, M. and Cheng, L., 2022. Evaluating the Accuracy of the Azure Kinect and Kinect v2. *Sensors*, 22(7): 2469.
- Li, D. et al., 2019. An Overlapping-Free Leaf Segmentation Method for Plant Point Clouds. *IEEE Access*, 7: 129054-129070.
- Li, W., Guo, Q., Jakubowski, M.K. and Kelly, M., 2012. A New Method for Segmenting Individual Trees from the Lidar Point Cloud. *Photogrammetric Engineering & Remote Sensing*, 78(1): 75-84.
- Omid, M., Khojastehnazhand, M. and Tabatabaefar, A., 2010. Estimating volume and mass of citrus fruits by image processing technique. *Journal of Food Engineering*, 100(2): 315-321.
- Shamim, M. et al., 2022. Proteo-Molecular Investigation of Cultivated Rice, Wild Rice, and Barley Provides Clues of Defense Responses against *Rhizoctonia solani* Infection. *Bioengineering*, 9(10): 589.
- Solak, S. and Altinişik, U., 2018. A new method for classifying nuts using image processing and k-means++ clustering. *Journal of Food Process Engineering*, 41(7): e12859.
- Tölggyessy, M., Dekan, M., Chovanec, =. and Hubinský, P., 2021. Evaluation of the Azure Kinect and Its Comparison to Kinect V1 and Kinect V2. *Sensors*, 21(2): 413.
- Turner, D.A., Anderson, I.J., Mason, J.C. and Cox, M.G., 1999. An Algorithm for Fitting an Ellipsoid to Data.
- Wang, T.Y. and Nguang, S.K., 2007. Low cost sensor for volume and surface area computation of axi-symmetric agricultural products. *Journal of Food Engineering*, 79(3): 870-877.
- Yoshimoto, K. and Shinya, M., 2022. Use of the Azure Kinect to measure foot clearance during obstacle crossing: A validation study. *PLOS ONE*, 17(3): e0265215.