Sigma Journal of Engineering and Natural Sciences
Web page info: https://sigma.yildiz.edu.tr
DOI: 10.14744/sigma.2022.00003

Research Article

Image processing based package volume measurement system using kinect sensor

Hasan OCAK[1], Gürkan KÜÇÜKYILDIZ[2,*], Suat KARAKAYA[1]

[1]Department of Mechatronics Engineering, Kocaeli University, Kocaeli, Turkey
[2]Department of Electrical-Electronics Engineering, Usak University, Usak, Turkey

ARTICLE INFO

Article history
Received: 05 May 2020
Accepted: 28 February 2021

Key words:
Camera calibration; Hough transform; Image processing; Kinect; Package volume detection

ABSTRACT

In this study, an image processing based volume measurement system is presented. The proposed scheme provides high accuracy, low cost and ease of calibration. There are no additional requirements such as alignment of objects during measurement, positioning the objects at a certain position in the measuring field, and these advantages reduce the labor effort in real-time. In the first stage, a problem-specific calibration process was performed and then Hough transform was applied in the image processing phase. Although the proposed method has been developed primarily to measure the volume of box-type objects, it also provides an approximate volume measurement for objects of random geometry. The system tested under various external disturbances has produced satisfactory results.

Cite this article as: Hasan O, Gürkan K, Suat K. Image processing based package volume measurement system using kinect sensor. Sigma J Eng Nat Sci 2022;40(1):26–44.

INTRODUCTION

Volume measurement systems are of interest to logistics companies and postal service providers to save time and costs. The trend in this area has concentrated on non-destructive measurement systems, which are mainly based on image processing techniques [1]. Various hardwares are currently being used, such as laser-based sensors, infrared sensors, stereo systems and acoustic systems. Laser sensors can be basically a combination of point lasers or high-cost 2D laser scanners [2]. In these systems misalignment between the reference and the package can lead to incorrect calculations. Deformation of the package surface is another major problem associated with such measuring approaches. More sensors should be used to increase the accuracy. It is also possible to measure volume on a moving platform using point laser source and camera together. Such an approach is also used in 3D laser scanners and studies have been carried out using these sensors for volume

***Corresponding author.**
*E-mail address: iusta@eskisehir.edu.tr

This paper was recommended for publication in revised form by Regional Editor N. Özlem Ünverdi

Published by Yıldız Technical University Press, İstanbul, Turkey
Copyright 2021, Yıldız Technical University. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

measurement in various studies [3]. The study given in [4] has been able to provide accurate measurements up to 1 mm.

Camera based methods are also used in volume measurement applications. Image-based volume measurement approaches can be classified as stereo camera and structured light systems [5]. Stereo systems use depth data provided by a pair of calibrated cameras [6]. An example study of such approach is provided in [7]. The authors used two calibrated cameras to measure the volume and position of a box. They calculated the vanishing points of the box border to obtain the 3D model. The measurement accuracy of these systems depends on various factors, such as proper stereo calibration, image resolution, camera and lens quality. Moreover, such systems may be adversely affected by varying visible light conditions. Therefore, camera-based systems typically require a natural light-free environment where the amount of light is controlled by artificial light sources. Unfortunately, this may not always be appropriate to implement in many work places. It also has a relatively higher computational burden in real-time applications. Structured light systems are a current approach to the problem of volume measurement [8].

Another approach consists of a single camera and a rotating platform, which usually measures the volume of small volume objects. The main problem in this approach is that the objects to be measured are smaller than the common-used commercial products and that no satisfying results are obtained in real-time industrial applications [9].

Volume measurement methods are also available by processing single images obtained from different angles. The surface volume of the human body was estimated by applying the Elliptical Tube Volume method given in [10]. It is known that it is far from being an industrial method.

Another hardware used for the volume measurement is the depth sensors. Kinect, which operates on the principle of structured infrared light, offers a measurement range of 0.7 cm to 6 m for (mostly) indoor environments. The measurement error increases up to a maximum of 4 cm proportional to the increase in distance. Depth images that can be captured with high-cost infrared cameras before the development of Kinect are now easily available. Depth data can be used in many areas such as medicine [11], gaming technology [12], mapping [13] and indoor layout [14]. Academic interest in Kinect continues to grow. Kinect has been used in a variety of areas such as hand gesture recognition [15], human tracking, and avoidance of obstacles [16].

Kinect depth data processing on transparent surfaces is a compelling problem. Improvement on the performance of the accuracy level of the depth data on transparent surfaces is presented in [17]. Kinect calibration studies were performed. In a study [18], calibration is performed by a quadratic transformation. Kinect data was used as 2D laser distance data in certain conditions [19]. Kinect has disadvantages with 2D laser scanners, such as narrower viewing angles and larger output data. On the other hand, Kinect is more cost-effective than 2D laser scanners. Alternative depth cameras were calibrated and compared on the same platform [20]. The authors presented that SLR stereo was the most accurate while the SR-4000 was the least accurate. In another study, which presented the volume-intersection approach using Kinect, finding the volumes of objects of various geometry, the volumes of medium-sized household objects could be calculated with an average error of 5.2% [21].

Figure 1. Box type package volume measurement platform.

In our study, we have developed a measurement platform supported with Kinect. The sensor was mounted on a custom made measuring cage shown in Figure 1.

Although the proposed system has been developed for volume measurement of box-type packages, it also provides an approximate solution for randomly shaped objects. The proposed system is not only a cost-effective alternative to existing systems, but also provides accurate measurements of up to 1 cm at a distance of 1.5 m. Unlike stereo camera-based systems, measurement accuracy is not affected by changing light conditions unless placed directly under sunlight or infrared light source. The method presented in this study is not affected by visible light, unlike stereo camera based systems. It significantly increases the measurement capacity by providing continuous operation in both dark

Table 1. Kinect parameters

Parameter	Value
f/sx	5.9104053696870778e+02
f/sy	5.9421434211923247e+02
cx	2.4273913761751615e+02
cy	3.3930780975300314e+02

and illuminated environments. It can operate with the plug-and-play principle without the need for additional camera calibration and light control despite changing operating conditions. The proposed method is robust against disturbances in the sensor field of view.

The rest of the article is organized as follows: In section 2, the calibration routine, background subtraction and Hough transform are presented. In section 3, the results are interpreted. Finally, the conclusions are discussed in section 4.

TECHNICAL BACKGROUND

Kinect calibration for measuring on planar surface

The calibration process is basically performed to compensate for the misalignment caused by the fact that the surface of the measuring plate on which we put the object is not parallel to the kinect camera plane. There are frequently used parameters and analytical solutions in the calibration stages. Calibration process between Kinect and measuring plate is given in steps by using parameters and equations which can be examined in the literature [22-24]. A problem-specific solution was proposed during the calibration stage. Roll and pitch angular misalignments were calculated with the assumption that the height of the sensor from the ground is known without the use of tools such as calibration board in standard calibration procedures.

Intrinsic matrix depends on camera internal parameters such as focal length (f), principal point (cx, cy), pixel width (sx) and pixel height (sy). The intrinsic matrix, M, of a camera is given in equation (1).

$$\mathbf{M} = \begin{bmatrix} -\frac{f}{s_x} & 0 & -c_x \\ 0 & -\frac{f}{s_y} & -c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

The internal parameters of Kinect were obtained from Kinect Software Development Kit (SDK). The parameters are given in Table 1. Real world coordinates of a pixel can be calculated using equation (2).

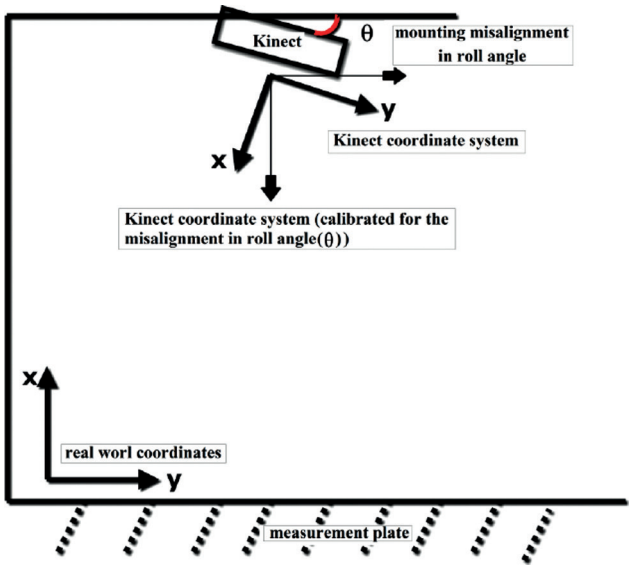


Figure 2. Misalignement of the sensor and tranformed coordinates.

$$\begin{bmatrix} \frac{x_k}{z_k} \\ \frac{y_k}{z_k} \\ 1 \end{bmatrix} = \mathbf{M}^{-1} \begin{bmatrix} i \\ j \\ 1 \end{bmatrix} \tag{2}$$

Extrinsic matrix depends on the camera placement. Because of the imperfections in the manufacturing process, the measurement and the Kinect mounting plates may not be perfectly aligned. Consequently, it results in the depth value of each point on the measuring plate being different. Therefore, it is necessary to make a coordinate transformation so that all points on the measurement plate should have the approximately equal depth value. To make this transformation, the rotation matrices of the Kinect coordinate system relative to the x and y axes (roll and pitch angles) must be defined. Rotation matrices ($\mathbf{R}_x$, $\mathbf{R}_y$) around the x and y axes are given in equation (3).

A single point in Kinect coordinate system ($\mathbf{p}_k$=[x_k y_k z_k]) can be transformed to a corresponding point ($\mathbf{p}_m$ = [x_m y_m z_m]) in the new coordinate system parallel to x-y plane of the measurement plate by equation (5).

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\theta) & -sin(\theta) \\ 0 & sin(\theta) & cos(\theta) \end{bmatrix}, \quad \mathbf{R}_y = \begin{bmatrix} cos(\phi) & 0 & sin(\phi) \\ 0 & 1 & 0 \\ -sin(\phi) & 0 & cos(\phi) \end{bmatrix} \tag{3}$$

Combined rotation matrix ($\mathbf{R} = \mathbf{R}_x\mathbf{R}_y$) is given in equation (4). $\mathbf{R}_z$ is neglected since rotation around z-axis does not affect the depth measurement.

The elimination of the roll-pitch angular misalignment is shown visually in Figure 2 for one dimension.

$$\mathbf{R} = \begin{bmatrix} \mathrm{c}(\phi) & 0 & \mathrm{s}(\phi) \\ \mathrm{s}(\theta)\mathrm{s}(\phi) & \mathrm{c}(\theta) & -\mathrm{s}(\theta)\mathrm{c}(\phi) \\ -\mathrm{c}(\theta)\mathrm{s}(\phi) & \mathrm{s}(\theta) & \mathrm{c}(\theta)\mathrm{c}(\phi) \end{bmatrix} \tag{4}$$

$$\mathbf{p}_\mathrm{m} = \mathbf{R}\mathbf{p}_\mathrm{k} \tag{5}$$

Equation (5) can be rewritten as,

$$\begin{aligned} x_m &= r_{11}x_k + r_{12}y_k + r_{13}z_k \\ y_m &= r_{21}x_k + r_{22}y_k + r_{23}z_k \\ z_m &= r_{31}x_k + r_{32}y_k + r_{33}z_k = d_m \end{aligned} \tag{6}$$

where r_{ij} is the i-th row and the j-th column element of $\mathbf{R}$. The depth value d_m must be constant for any point on the measurement plate. The third row of equation (6) can be rewritten as equation 7,

$$\mathbf{r}_3^\mathrm{T}\mathbf{p}_\mathrm{k} = d_m \tag{7}$$

where $\mathbf{r}_3$ is equal to the third row of the rotation matrix $\mathbf{R}$ and d_m is equal to the depth value of the measurement plate. The superscript T means the transpose of the matrix. Every discrete point on the measurement plate must satisfy equation (8). N points were selected on the measuring plate to find solutions for $\mathbf{r}_3$ and d_m. The set of N equations is re-expressed in matrix format as following equation.

$$\begin{bmatrix} \mathbf{p}_{\mathrm{k1}}^\mathrm{T} & -1 \\ \mathbf{p}_{\mathrm{k2}}^\mathrm{T} & -1 \\ \mathbf{p}_{\mathrm{k3}}^\mathrm{T} & -1 \\ . & . \\ . & . \\ \mathbf{p}_{\mathrm{kN}}^\mathrm{T} & -1 \end{bmatrix} \begin{bmatrix} \mathbf{r}_3^\mathrm{T} \\ d_m \end{bmatrix} = 0 \tag{8}$$

Equation 8 can be expressed as $\mathbf{Ax} = 0$ format, where;

$$\mathbf{A} = \begin{bmatrix} \mathbf{p}_{\mathrm{k1}}^\mathrm{T} & -1 \\ \mathbf{p}_{\mathrm{k2}}^\mathrm{T} & -1 \\ \mathbf{p}_{\mathrm{k3}}^\mathrm{T} & -1 \\ . & . \\ . & . \\ \mathbf{p}_{\mathrm{kN}}^\mathrm{T} & -1 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} \mathbf{r}_3^T \\ d_m \end{bmatrix} \tag{9}$$

This equation can be solved by using singular value decomposition (SVD). The solution is the null space of $\mathbf{A}$. As a result of the solution, R matrix is defined using θ and ϕ angles. The equivalents of a point in the Kinect space in a new coordinate system that is concentric but with a

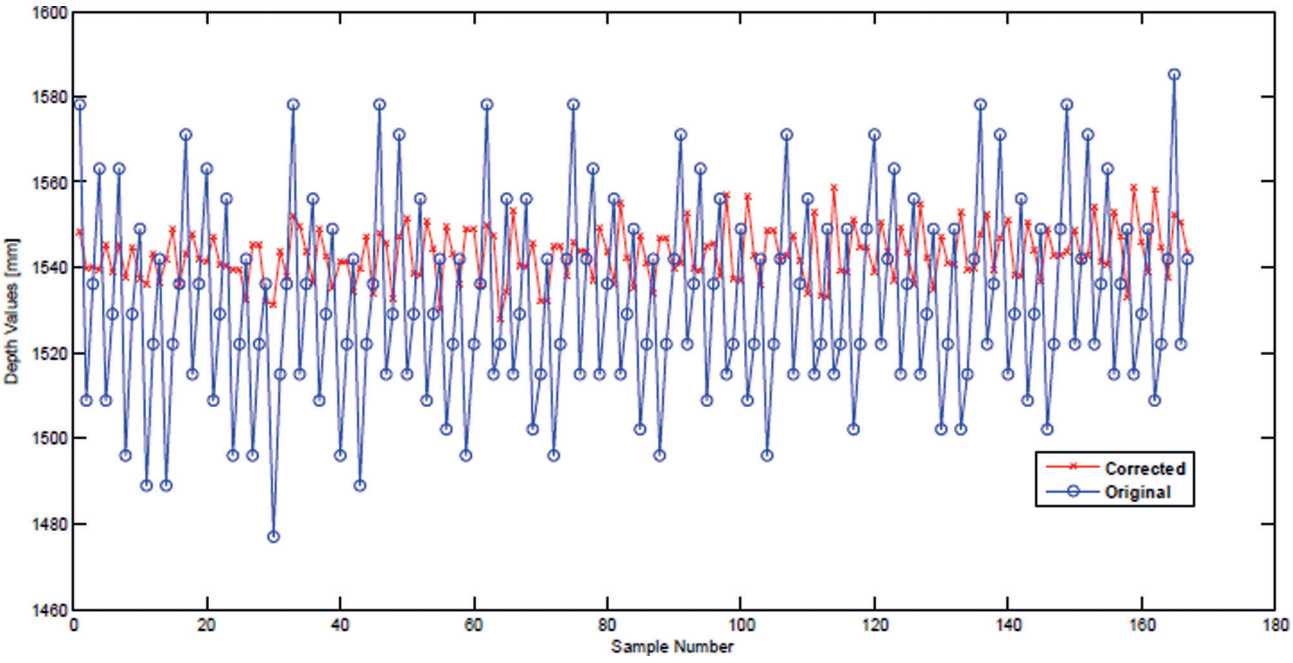


Figure 3. Original and corrected depth values.

distortion equal to θ and ϕ angles are calculated in the last step by applying equation (5).

Figure 3 depicts the depth measurements taken at various points on the measurement plate before and after a sample calibration procedure. The blue circles and the red crosses represent the original and the corrected depth values, respectively. Table 2 provides the mean, standard deviation and the range of the original and corrected measurements. It can be seen that the standard deviation of the original depth values on the plate is very high compared to the one for the corrected depth values. After the calibration procedure, depth values on the plate deviates from the mean by approximately ±2 cm, which is roughly equivalent to Kinect's measurement accuracy at 150 cm.

In the image on Figure 4.a, the depth map shows an inhomogeneous distribution on the ground and the measuring plate. This is due to the fact that the sensor is not fully parallel to the ground and this is converted to the format on the right by applying the recommended calibration process. In this way, the depth information of the upper surface of the box to be measured is evenly distributed (Figure 4.b).

Table 2. Original and calibrated image statistics

	Original Depth Image	Corrected Depth Image
Standard Deviation	22.3334	6.5437
Mean Value (mm)	1531.2	1542.9
Range (Min. – Max.) (mm)	1470 - 1585	1522.8 - 1565.1

Background Subtraction

Since a height measurement will be made with reference to the depth of the measuring plate, the average of the first 20 depth images is calculated before the initialization. Images read at 30fps for a period of 3.3 ms are subjected to averaging to minimize noise. The averaging depth image is assumed to be the background (reference) image. In the next step, the volume measurement algorithm is triggered when pixels above a certain threshold are detected on the measuring plate. Step by step definition of the background substraction is given below.

- Estimate the background depth matrix by averaging the 20 frames captured at the initial state,
- Subtract the background depth matrix from the input frame,
- Apply a treshold to the absolute difference to obtain the foreground image (height of the box on the measurement plate), (equation 10).

$$\mathbf{I}_b = \sum_{t=1}^{20} \frac{\mathbf{I}(x,y,t)}{n}, n = 20$$
$$\mathbf{I}_f(x,y,t) = \left|\mathbf{I}(x,y,t) - \mathbf{I}_b\right| > threshold \quad (10)$$

where, $\mathbf{I}_b$ is the background depth matrix, $\mathbf{I}(x, y, t)$ is input frame at time t and $\mathbf{I}_f(x, y, t)$ is the foreground image at time t.

At this stage, pixels with a height higher than a certain threshold in depth value are considered as a possible

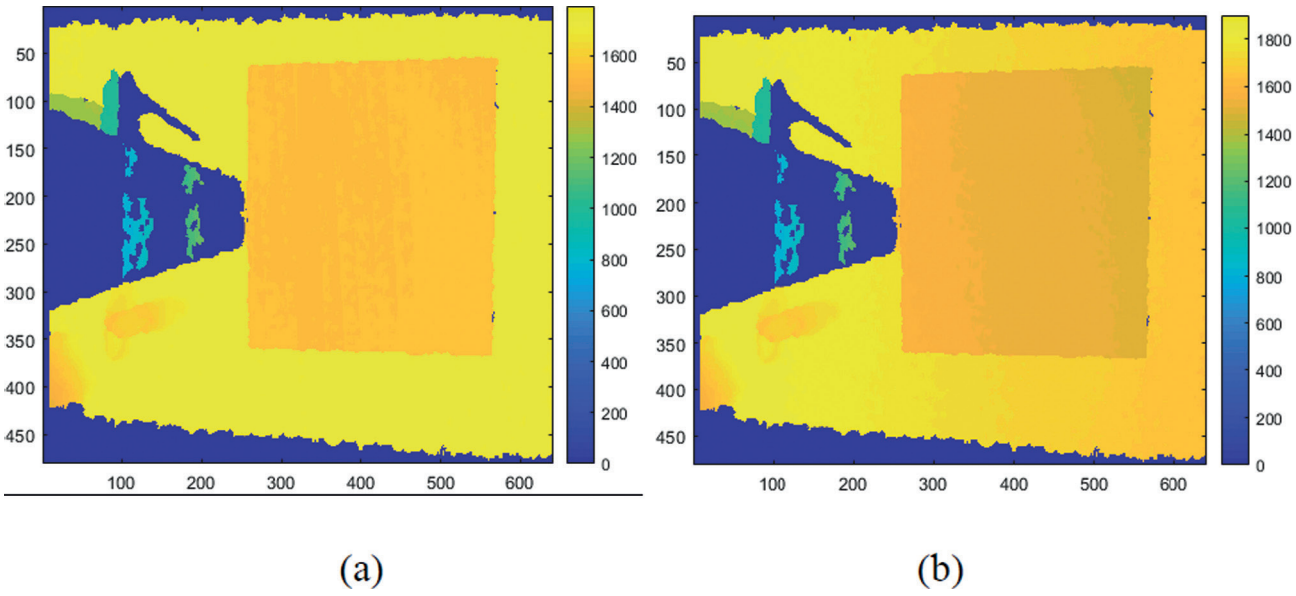


Figure 4. Original (a) and corrected (b) depth map after calibration.

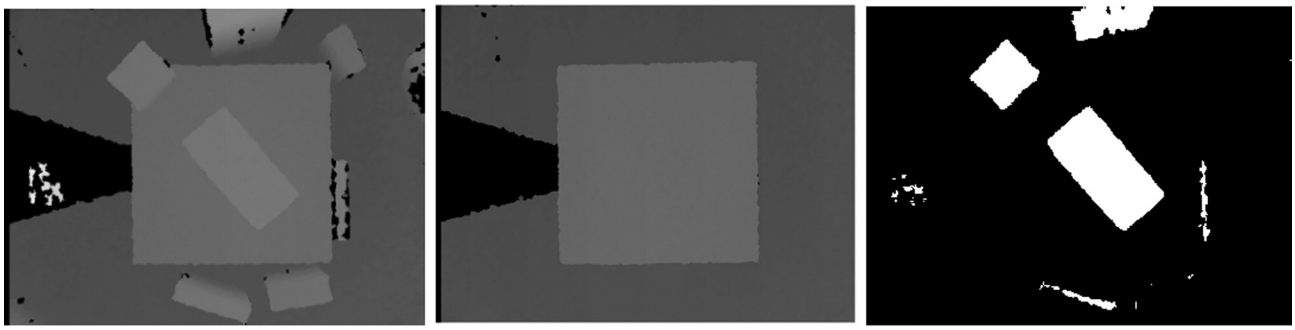


Figure 5. Sample (a), background (b) and morphological operations applied (c) images.

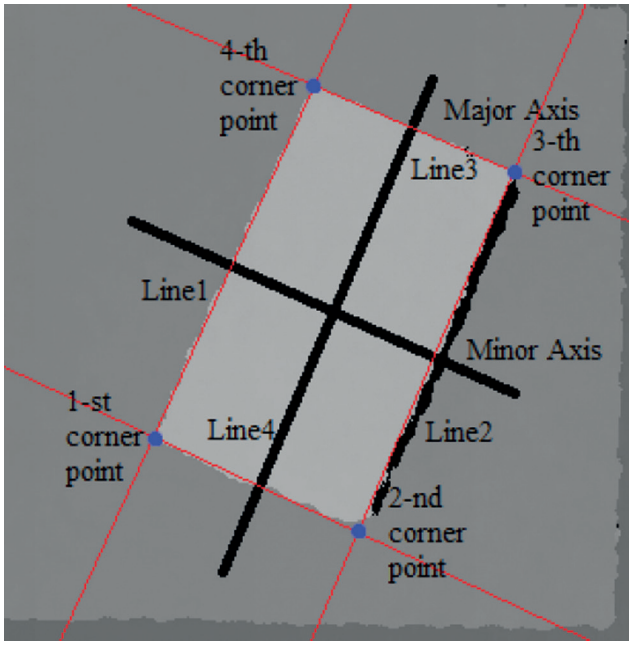


Figure 6. Hough lines and major-minor axes.

box and used to obtain a black and white (BW) image. Morphological opening was applied to the image for noise removal. Morphological closing and region filling were then applied to fill the gaps. After morphological operations, connected components algorithm was used to segment the image. The object which has the maximum pixel area overlapping with the measurement plate was assumed as package of interest and others were eliminated. A sample image (a), background image (b) and morphological operations applied image (c) are given in Figure 5.

The procedures applied for detecting the minimum bounding box on the resulting BW image are shown in the following topic.

Hough Transform

Theory of the Hough transform was presented to literature in 1972 [23]. Another major study given in [25] proved that geometric forms defined in parametric expressions, can be detected by using Hough transform. Hough transform can be described in equation (11).

$$\rho = x\cos(\theta) + y\sin(\theta) \tag{11}$$

Hough transform has robustness against gaps and noise, but its computational cost can be challenging in real time applications [24].

Pseudo code of Hough transform is given in the following statement:

```
For each pixel in the binary image{
  If the pixel is white {
    For each θ {
      Compute
        ρ = xcos(θ)+ysin(θ)
      T(ρ, θ) = T(ρ, θ ) + 1 }}}
```

The final BW image was divided into four zones with major and minor axes to reduce computational cost. Orientation angle (β), approximate width (d_1) and approximate height (d_2) of the object were obtained by connected components labeling algorithm. The line passing through the middle of the detected object and having the orientation angle β is accepted as the major axis.

Similarly, the line which passes through the center point of the detected object and has an orientation angle $\beta + 90$ was assumed to be the minor axis. The parameter ρ given in equation (10) is calculated for both axes (ρ_1, ρ_2). Hough lines and major-minor axis representation for a sample object are given in Figure 6.

Hough lines, which have θ values $\beta \pm \delta\theta$ and ρ values $\rho_1 - 0.5d_1 \pm \delta\rho$ and $\rho_1 + 0.5d_1 \pm \delta\rho$, were searched in the Hough space. Maximum scored lines above and below the major

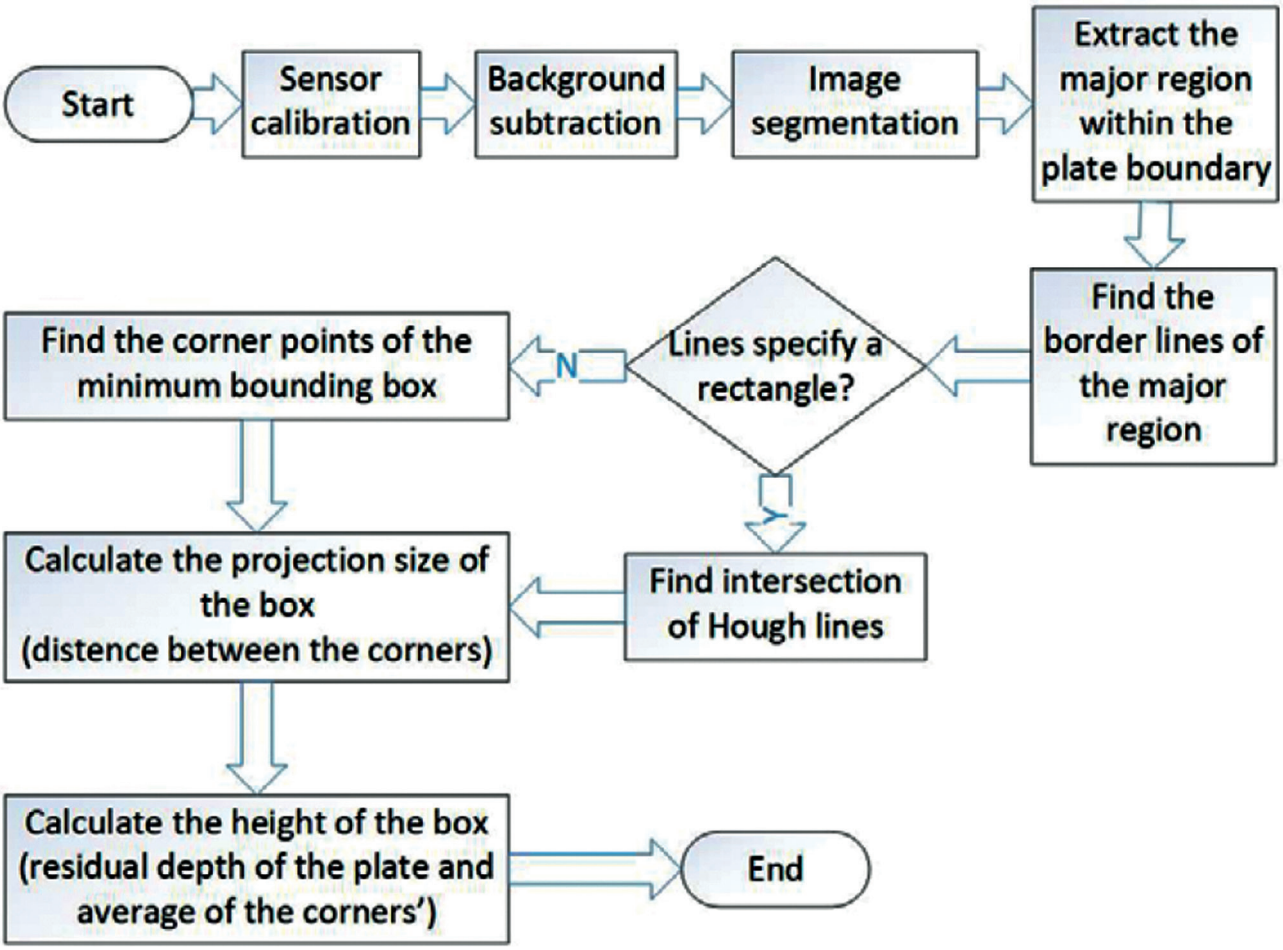


Figure 7. Flow-chart of the proposed algorithm.

axis ($Line_1$, $Line_2$) were assumed as the first and the second edge lines of the package surface. Similarly, Hough lines, which have θ values $\beta + 90 \pm \delta\theta$ and ρ values $\rho_2 - 0.5d_1 \pm \delta\rho$ and $\rho_2 + 0.5d_1 \pm \delta\rho$, were searched in the Hough space. Maximum scored lines above and below minor axis ($Line_3$, $Line_4$) were assumed as the third and the fourth edge lines of the package surface. Searching dominant lines without regard to edge information will likely cause false detections. Dividing the image into four regions and searching an independent Hough line along each edge guarantees that the selected four Hough lines cover the entire package surface. Figure 6 shows the four detected Hough lines for a sample image. Package corners were determined by finding the intersection points of the Hough lines. Real world coordinates of the corner points were calculated by equation (6). The flow chart of developed image processing algorithm is given in Figure 7.

Because of the non-perfect package geometries and deformations on the package corners, the detected corner points can possibly fall on the measurement plate. Therefore, the depth values of a corner point was calculated by averaging the depth values of the points on a circle centered on the corner point with a certain radius which overlap with the package. The average of the Euclidian distances between 1st and 2nd corner points and 3rd and 4th corner points in real world coordinates were calculated and assumed to be the width of the package. Similarly, the average of the Euclidian distances between 1st and 4th corner points and 2th and 3rd corner points in real world coordinates were calculated and assumed to be the length of the package. The package height was determined as the difference between the plate depth (which was found after the calibration step) and the average depth values of points on the package surface.

RESULTS

The proposed algorithm was executed on a Windows based PC with Matlab software. To test the robustness of the proposed volume measurement algorithm, possible disturbances must be considered. Due to operating conditions,

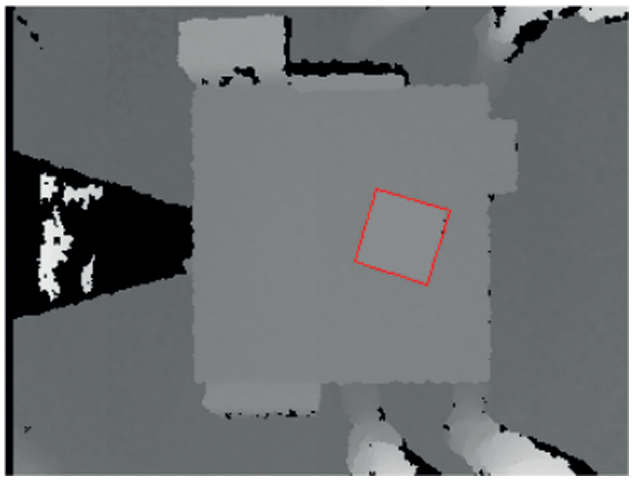


Figure 8. Measurement result for package 1.

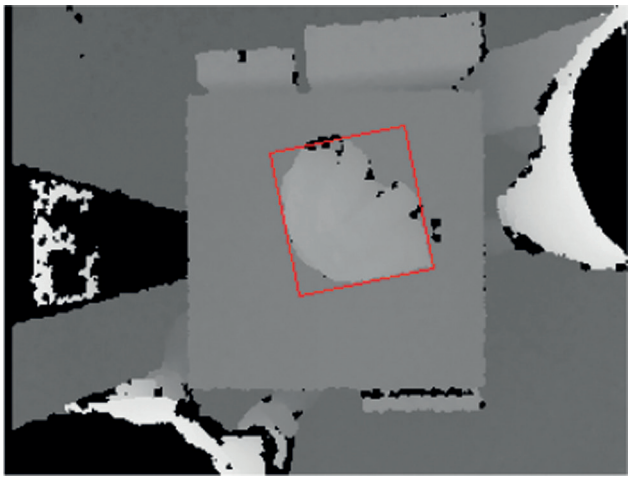


Figure 9. Measurement result for non-rectangular package 2.

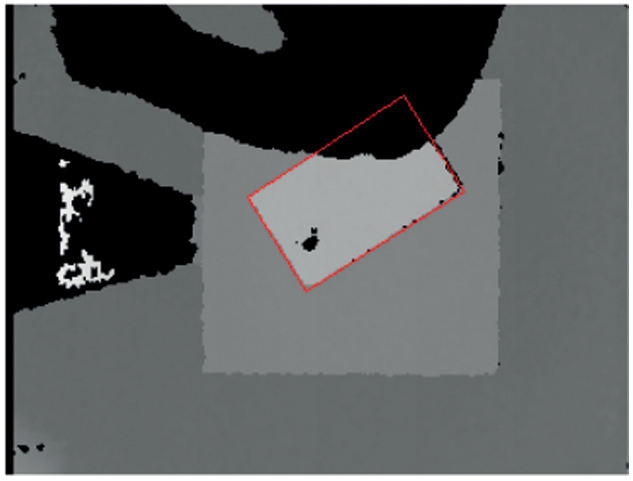


Figure 10. Measurement result for package 3.

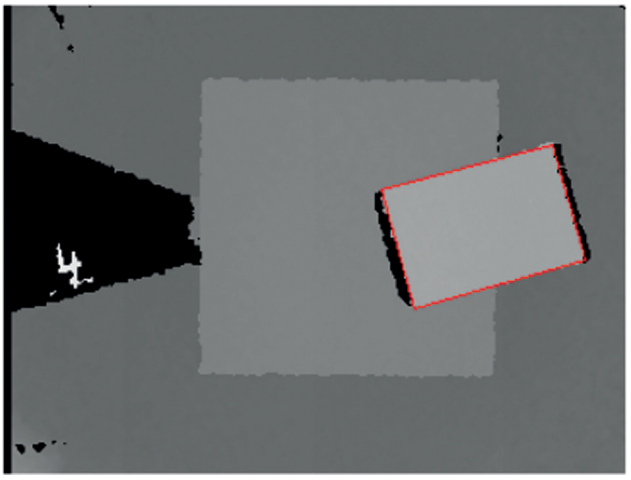


Figure 11. Measurement result for package 4.

the operator might partly obstruct Kinect's field of view while handling the packages. In addition, there might be other packages around the measuring plate either handled or waiting to be handled. These packages might completely or partly fall in Kinect's field of view.

Given the fact that Kinect's minimum range for depth measurement is 80 cm and the horizontal and vertical viewing angles are 43° and 57°, respectively, the sensor was placed at a height for measuring packages up to 60cm x 60cm x 60cm in size.

Another scenario is when part of the package is not on the measuring plate. Various tests were performed under these circumstances. The results prove that the algorithm could determine the package dimensions (width, length, height) with maximum 1 cm error in all dimensions regardless of the disturbances.

Figure 8 illustrates a sample measurement performed while other packages are present on the floor around the measurement plate. The legs of the operator are also visible in the depth image as another disturbance effect. The measurement results for the given scenario are provided in Table 2 under package number 1. The measurement errors are 0.1 cm, 0.1 cm and 0.3 cm for the width, length and the height, respectively.

If a non-rectangular object is placed on the measurement platform, then the image processing algorithm cannot detect four Hough lines. In such a case as illustrated in Figure 9, object dimensions are determined based on the minimum bounding rectangle.

In a sample measurement depicted in Figure 10, the operator's arm is partly obstructing Kinect's field of view while handling the packages. The system successfully

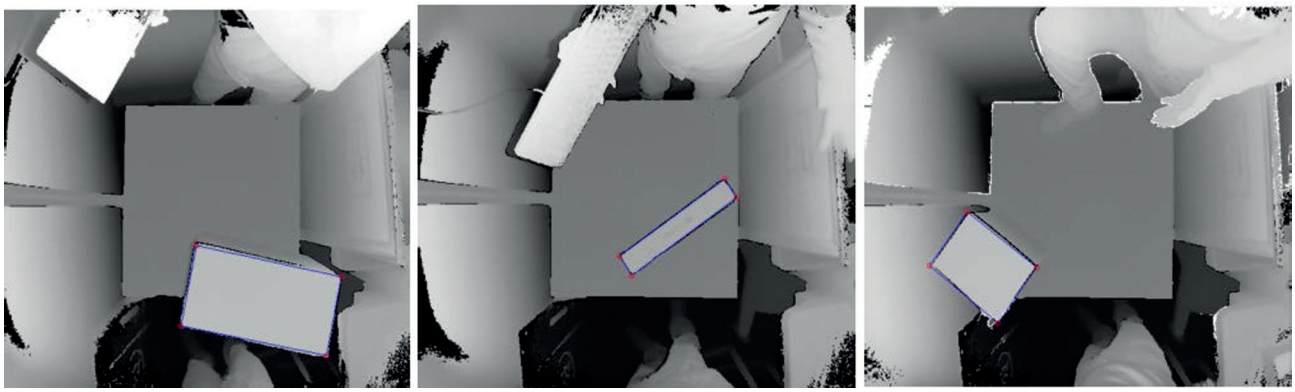


Figure 12. Left to right (a), (b), (c).

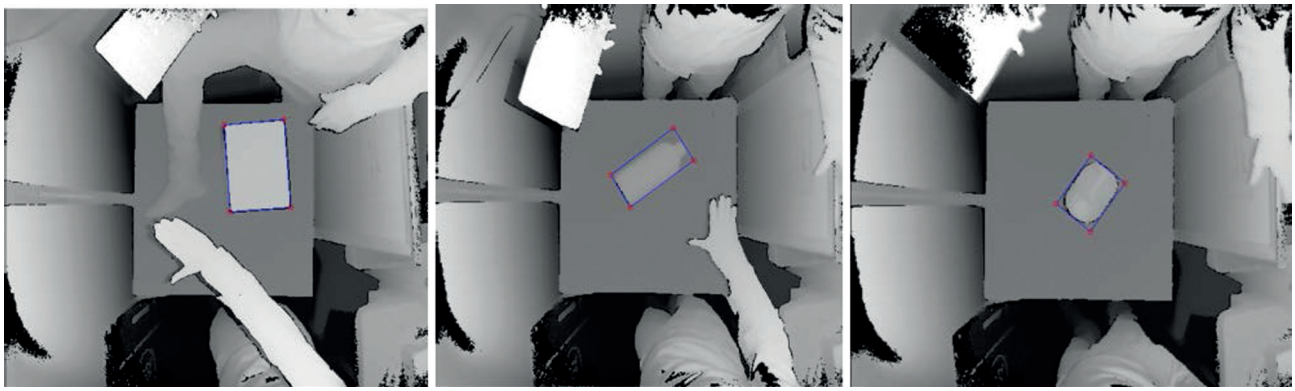


Figure 13. Left to right (a), (b), (c).

Table 3. Measurement results

Package number	Width (cm)	Length (cm)	Height (cm)	Width error (cm)	Length error (cm)	Height error (cm)
1	24,20000	29.90000	6.40000	0.10000	0.10000	-0.30000
2	24.30000	39.70000	27.90000	-0.80000	0.80000	0.40000
3	30.20000	43.10000	12.40000	0.40000	-0.90000	-0.10000
4	27.40000	40.20000	22.90000	0.10000	0.30000	-0.60000
5	8.90000	20.60000	44.00000	-0.10000	0.10000	0.20000
6	9.00000	20.70000	43.20000	0.00000	-0.20000	1.00000
7	32.50000	35.00000	23.50000	0.50000	-1.00000	0.00000
8	7.60000	18.00000	20.00000	0.40000	0.00000	-0.80000
9	20.10000	23.70000	11.20000	-0.40000	-0.20000	0.20000
10	19.20000	19.40000	7.50000	-0.20000	0.80000	-0.20000
11	13.40000	31.20000	8.90000	0.20000	0.20000	0.00000
12	13.00000	19.30000	6.30000	-0.30000	0.80000	-0.20000
13	19.90000	20.80000	7.00000	0.10000	-0.20000	0.00000
14	30.20000	27.40000	6.40000	-0.20000	0.10000	-0.10000
15	24.20000	38.00000	11.10000	-0.80000	1.00000	0.60000
Average:	20.27333	28.46667	17.24667	-0.06667	0.11333	0.00667

measured the package volume even though part of the package was missing in the depth image. The measurement results for this scenario are provided in Table 3 under package number 2. The measurement errors are 0.8 cm, 0.8 cm and 0.4 cm for the width, length and the height, respectively.

Figure 11 presents a sample measurement taken while part of the package was outside the measuring plate. Package dimensions were also effectively measured for the given case with measurement errors of 0.1 cm, 0.3 cm and 0.6 cm for the width, length and the height, respectively. The measurement results for this case are given in Table 3 under package number 4.

Figure 12(a) illustrates the case that a part of the package is outside the measuring plate and Figure 12(b) shows the case where the Kinect's field of vision is partially obstructed where Figure 12(c) shows the cases both the Kinect's field of vision is partially obstructed and the a part of the package is outside the measuring plate.

Figure 13(a) shows the case where the Kinect's field of view is obstructed by multiple external disturbances. Figure 13(b) and Figure 13(c) show minimum bounding rectangle solutions calculated for non-rectangular objects. The presented scenarios show that the system behaves robust against the effects of the operators on the real-time system.

The results of the measurements in width, height and length dimensions are given in Table 3. The average of the actual measurements and measurement error data were calculated by the operation given in equation (12). The average values of the actual measurements are calculated in equation (13) and the average errors are calculated by the operations given in equation (15).

Average error values were calculated by applying equation (15) for each dimension separately. The volume measurement with measurement error is given in equation (16) and the actual volume calculation is given in equation (17). The mean absolute percentage error value of the volume measurement system is given in equation (18).

The subscript *act* used in the equations shows the actual values and the *meas* subscript shows the measured values. The subscript *w* used in the equations means width, *h* stands for height and *l* stands for length.

The results show that the proposed system can reliably measure the volume of a box-type package with 1 cm accuracy in all dimensions even under the presence of several disturbances. The average absolute percentage error value of the volume measurement system is calculated as 0.10666. More results are given in the Supplemantary file of the paper.

In the laser range finder based system (Zhang et al., 2005) which was developed for volume measurement in the literature, the measurement error occurred at 1.80 percent. The system based on structured light technology developed in (Li et al., 2016) resulted in a 3.03 percent error. In the study using computer vision (Jadhav & Kamble, 2016) a measurement error of 0.168 was revealed. The comparison is given in Table 4.

Table 4. Comparison of volume measurement approaches

Volume measurement approach	Percentage error (%)
Laser range finder based system (Zhang et al., 2005)	1.8000
Computer vision based system (Jadhav & Kamble, 2016)	0.1680
Constructed light based system in (Li et al., 2016)	3.0300
Kinect based system	0.1066

$$f(x)=\frac{\sum_{i=1}^{n_{sample}} x_i}{n_{sample}} \tag{12}$$

$$\overline{w}_{act}=f(w_{act}),\ \overline{h}_{act}=f(h_{act}),\ \overline{l}_{act}=f(l_{act}) \tag{13}$$

$$\overline{\delta}_w=f(\delta_w),\ \overline{\delta}_h=f(\delta_h),\ \overline{\delta}_l=f(\delta_l) \tag{14}$$

$$\overline{w}_{meas}=\overline{w}_{act}+\overline{\delta}_w,\ \overline{h}_{meas}=\overline{h}_{act}+\overline{\delta}_h,\ \overline{l}_{meas}=\overline{l}_{act}+\overline{\delta}_l \tag{15}$$

$$\overline{vol}_{meas}=\overline{w}_{meas}\times\overline{h}_{meas}\times\overline{l}_{meas} \tag{16}$$

$$\overline{vol}_{act}=\overline{w}_{act}\times\overline{h}_{act}\times\overline{l}_{act} \tag{17}$$

$$\varepsilon_{vol}=100\times\frac{\left|\overline{vol}_{meas}-\overline{vol}_{act}\right|}{\overline{vol}_{act}} \tag{18}$$

CONCLUSIONS AND FUTURE WORK

In this study an image processing based volume measurement system is presented. The system consists of the Kinect sensor and a computer. Captured depth images were procesed by a developed image processing algorithm. Algorithms are designed on a suitable basis for software and hardware transition for embedded operation on a single board that reduces cost. The developed code can also be employed on a embeded computer such as Raspberry Pi. On the other hand, the proposed system could integrate with enterprise resource planning (ERP) systems via web service. The proposed system could determine only one box's dimensions per mesurement. The code can also easily be adapted to measure multiple packages simultaneously.

The manufacturing of the Kinect sensors(both Kinect One, Kinect has expired. Therefore, an commercial depth cameras (Intel, Logitech etc) could be used to gather the depth images.

ACKNOWLEDGEMENT

The prototype of this study is currently evaluated as a commercial product by Romeda Inc. and thus limited visual material is shared.

AUTHORSHIP CONTRIBUTIONS

Authors equally contributed to this work.

DATA AVAILABILITY STATEMENT

The authors confirm that the data that supports the findings of this study are available within the article. Raw data that support the finding of this study are available from the corresponding author, upon reasonable request.

CONFLICT OF INTEREST

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ETHICS

There are no ethical issues with the publication of this manuscript.

REFERENCES

[1] Balogh Z, Magdin M, Monlar G. Motion detection and face recognition using raspberry pi, as a part of, the internet of things. Acta Polytech Hungarica 2019;16:167–185. [CrossRef]

[2] Yamada T, Ito T, Ohya A. Detection of road surface damage using mobile robot equipped with 2D laser scanner. Proceedings of IEEE/SICE International Symposium on System Integration (SII). Kobe, Japan 2013. [CrossRef]

[3] Wei G, Shalei S, Bo Z, Shuo S, Faquan L, Xuewu C. Multiwavelength canopy Lidar for remote sensing of vegetation, design and system performance. ISPRS J. Photogramm. Remote Sens 2012;69:1–9. [CrossRef]

[4] Zhang X, Morris J, Klett R. Volume measurement using a laser scanner. Communication and Information Technology Research (CITR) Technical Report. Computer Science Department, The University of Auckland, 2005.

[5] Li J, Liu G, Liu Y. A dynamic volume measurement system with structured light vision. Proceedings of 31st Youth Academic Annual Conference of Chinese Association of Automation. Wuhan, China, 2016. [CrossRef]

[6] Yang R, Cheng S, Yang W, Chen Y. Robust and accurate surface measurement using structured light. IEEE Trans Instrum Measurement 2008;57:1275–1280. [CrossRef]

[7] Kurka P, Rudek M. Three dimensional volume and position recovering using a virtual reference box. IEEE Trans Image Processing 2007;1:573–576. [CrossRef]

[8] Mao J, Lou X, Li W, Zhao Y. Binocular stereo vision based online structured light. Optical Technique 2016;42:10–15.

[9] Jadhav T, Kamble S. Volume measurement of object using computer vision. IEEE International Conference On Recent Trends In Electronics Information Communication Technology. Bangalore, India, 2016. [CrossRef]

[10] Supranata, TH, Davin PS, Jeremy DK, Pratiwi AE, Wulandari M. Body weight measurement using image processing based on body surface area and elliptical tube volume. International Conference on Information Technology and Electrical Engineering (ICITEE). Bali, Indonesia, 2018. [CrossRef]

[11] Kitsunezaki N, Adachi, E., Masuda, T., Mizusawa, J.: Kinect applications for the physical rehabilitation. IEEE International Symposium on Medical Measurements and Applications Proceedings (MeMeA). Gatineau, QC, Canada, 2013. [CrossRef]

[12] Duarte N, Postolache O, Scharcanski J. KSGphysio-Kinect serious game for physiotherapy. International Conference and Exposition on Electrical and Power Engineering (EPE). Lasi, Romania, 2014. [CrossRef]

[13] Khoshelham K, Elberink SO. Role of dimensionality reduction in segment-based classification of damaged building roofs in airborne Laser scanning data. Proceedings of the 4th GEOBIA. Rio de Janeiro, Brazil, 2012.

[14] Joydeep B, Manuela V. Depth camera based indoor mobile robot localization and navigation. Int J Rob Res 2013;2:1679–1694.

[15] Yi L. Hand gesture recognition using Kinect. IEEE International Conference on Computer Sciences and Applications. Beijing, China, 2012. [CrossRef]

[16] DeSouza GN, Kak AC. Vision for mobile robot navigation: A survey. IEEE Trans Pattern Anal Mach Intell 2002;24:237–267. [CrossRef]

[17] Saygili G, Maaten L, Hendriks E. Hybrid kinect depth map refinement for transparent objects. ieee international conference on pattern recognition. Stockholm, Sweden, 2014. [CrossRef]

[18] Essmaeel K, Gallo L, Damiani E, Pietro, G, Dipanda A. Temporal denoising of Kinect depth data. IEEE International Conference on Signal Image

Technology and Internet Based Systems. Naples, Italy, 2012. [CrossRef]

[19] Zug S, Penzlin F, Dietrich A, Tran T, Albert S. Are laser scanners replaceable by Kinect sensors in robotic applications. IEEE International Symposium on Robotic and Sensors Environments (ROSE). Magdeburg, Germany, 2012. [CrossRef]

[20] Smisek J, Jancosek M, Pajdla T. 3D with Kinect. IEEE International Conference on Computer Vision Workshops (ICCV Workshops). Barcelona, Spain, 2011. [CrossRef]

[21] Dellen B, Rojas, I. Volume measurement with a consumer depth camera based on structured infrared light. Catalan Conference on Artificial Intelligence. Catalonia, Spain, 2013.

[22] Pagliari D, Pinto L. Calibration of kinect for Xbox one and comparison between the two generations of microsoft sensors. Sensors 2015;15:27569–27589. [CrossRef]

[23] Richard OD, Peter EH. Use of the Hough transformation to detect lines and curves in pictures. J Commun ACM 1972;15:11–15. [CrossRef]

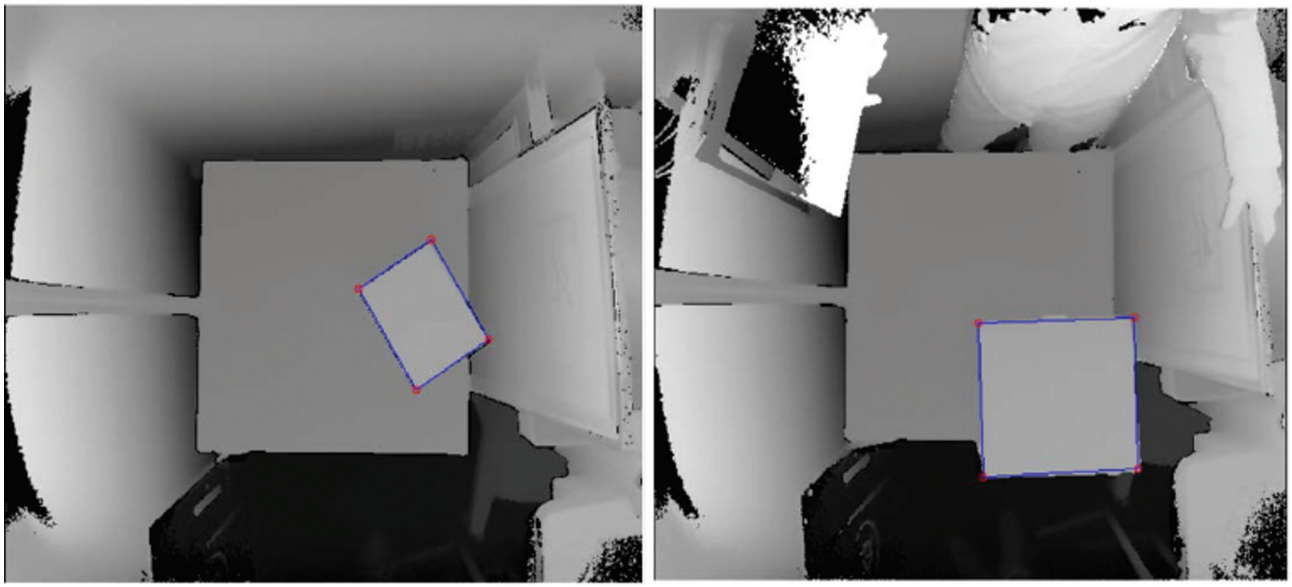
[24] Kucukyildiz G, Ocak H. Development and optimization of a DSP-based real-time lane detection algorithm on a mobile platform. Turk J Elect Eng Comput Sci 2014;22:1484–1500. [CrossRef]

[25] Ballard DH. Generalizing Hough transform to detect arbitrary shapes. J Pattern Recognit 1981;13:111–122. [CrossRef]

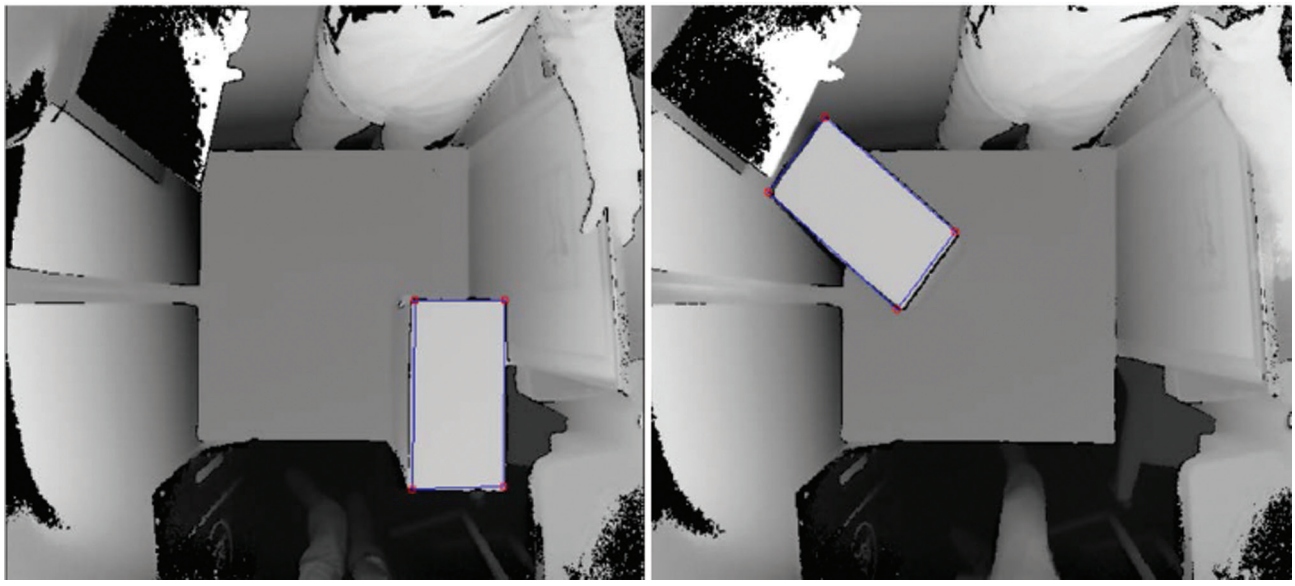
Supplemantary Table 1. Measurement results

Package number	Width (cm)	Length (cm)	Height (cm)	Width error (cm)	Length error (cm)	Height error (cm)
1	20,8	26,2	17,2	0.20000	-0.10000	0.30000
2	33,5	37,2	18,8	0.00000	-0.30000	0.30000
3	38,2	33,6	17,6	0.40000	0.10000	-0.90000
4	34,2	38	17,5	-0.10000	0.30000	0.20000
5	22,9	45,5	35,2	-0.10000	0.00000	0.20000
6	22	34	9,3	0.00000	-0.50000	0.30000
7	7	42,1	24,5	-0.50000	0.10000	0.50000
8	8,2	24,1	41,7	0.70000	0.10000	-0.30000
9	23,9	37,5	17,8	-0.10000	-0.20000	0.80000
10	14,5	19,2	21	0.50000	0.20000	0.00000
11	42,3	12,6	25,4	-0.20000	0.10000	0.40000
12	35,3	33,5	20,2	0.30000	-0.50000	0.20000
13	27,4	20,4	15,2	-0.10000	0.40000	-0.30000
14	26,5	35,3	30,3	0.50000	0.30000	-0.20000
15	8,4	43,6	33	-0.10000	-0.40000	0.00000
16	15,6	18,3	20,1	0.10000	-0.20000	0.10000
17	22,7	21,3	23,5	0.70000	-0.20000	0.00000
18	32,2	33,5	16,4	0.20000	1.00000	-0.40000
19	21,3	35,2	17,6	-0.20000	0.20000	0.60000
20	41,6	9,3	33,4	-0.90000	0.30000	-0.60000
21	45,8	10,1	35,2	0.80000	0.10000	0.20000
22	9,1	10,6	10	-0.90000	0.60000	0.00000
23	26,2	21,6	25,4	0.20000	0.60000	0.40000
24	38,4	33,3	30,3	0.40000	-0.20000	-0.20000
25	21,1	20,9	22,8	-0.20000	0.20000	0.60000
26	34,2	30,5	29,7	-0.90000	0.30000	-0.60000
27	45,6	8,5	30,3	0.80000	0.10000	0.20000
28	30,9	30,7	34,2	0.90000	0.70000	0.20000
29	7,6	9,9	10,1	-0.40000	-0.10000	0.10000
30	21,2	19,3	23,8	0.20000	-0.20000	-0.20000
31	27,8	20,1	25,8	0.80000	0.10000	0.30000
32	30,5	9	42,2	0.00000	0.00000	-0.30000
33	25,5	8,8	22,2	0.50000	0.70000	0.20000
34	19,7	10,2	20,3	-0.30000	0.70000	0.30000
35	10,5	15,9	14,2	0.0000	-0.10000	-0.30000
Averages:				0.0914	0.1200	0.0600

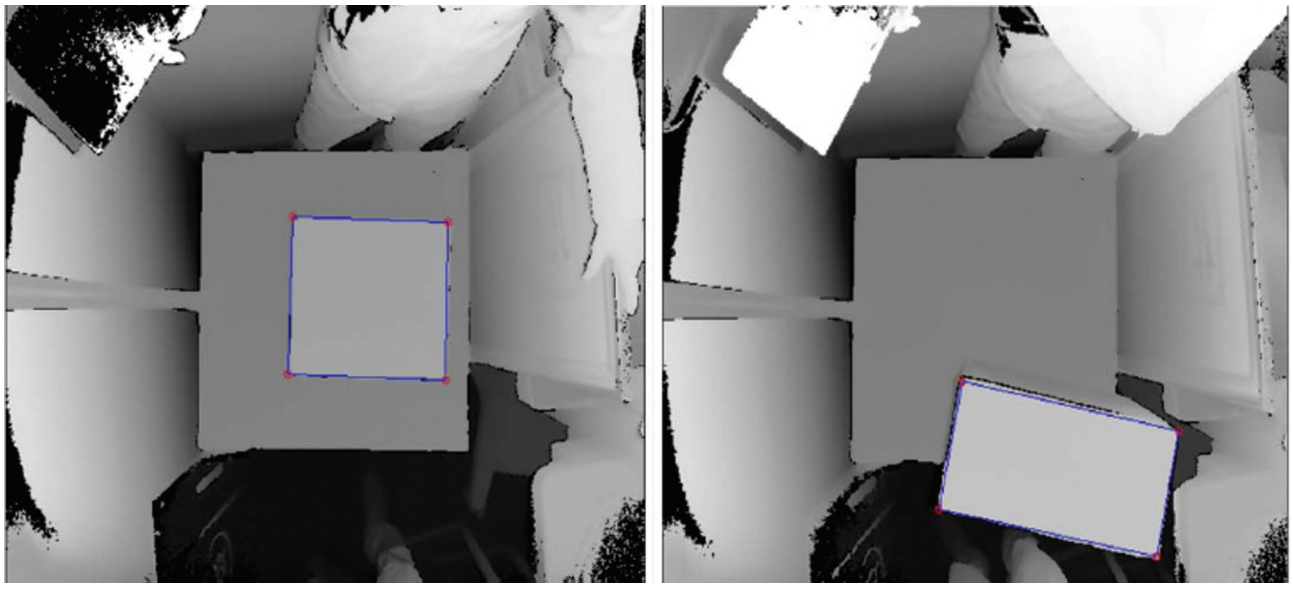
Supplemantary Depth Measurement Figures



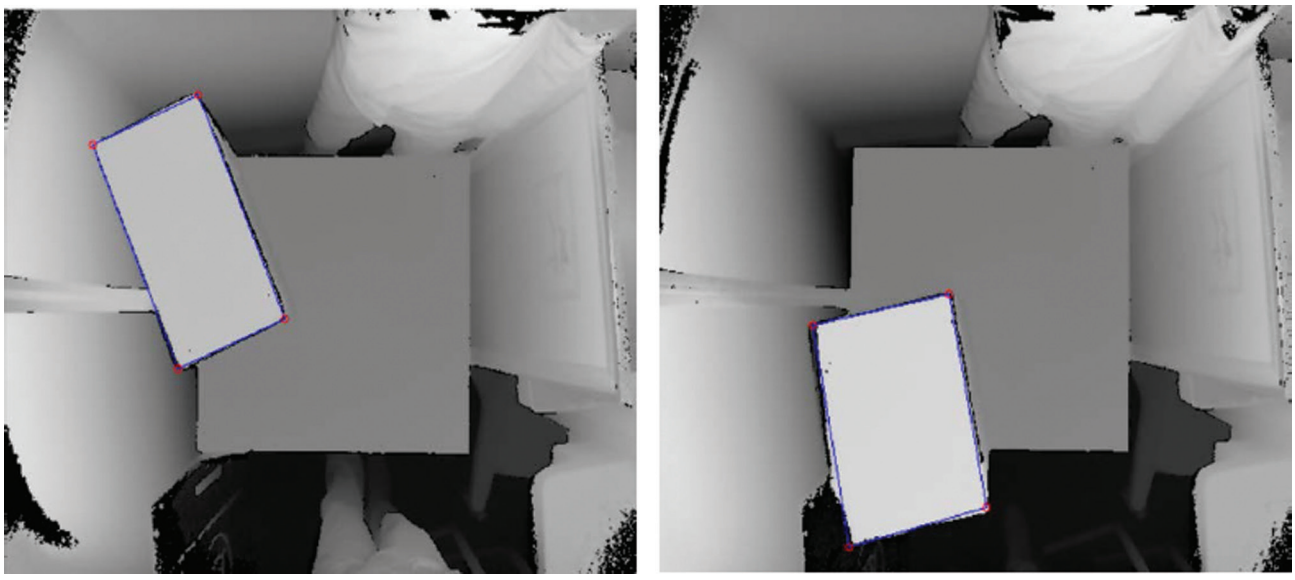
Supplemantary Figure 1. Sample Measurements 1.



Supplemantary Figure 2. Sample Measurements 2.

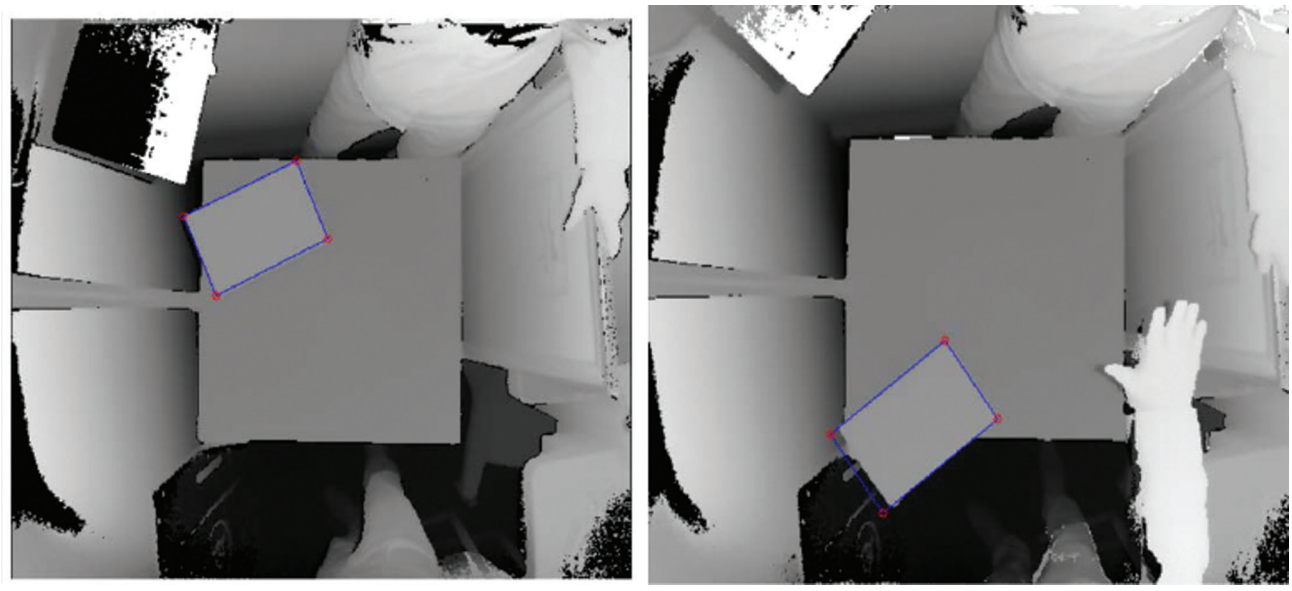


Supplemantary Figure 3. Sample Measurements 3.

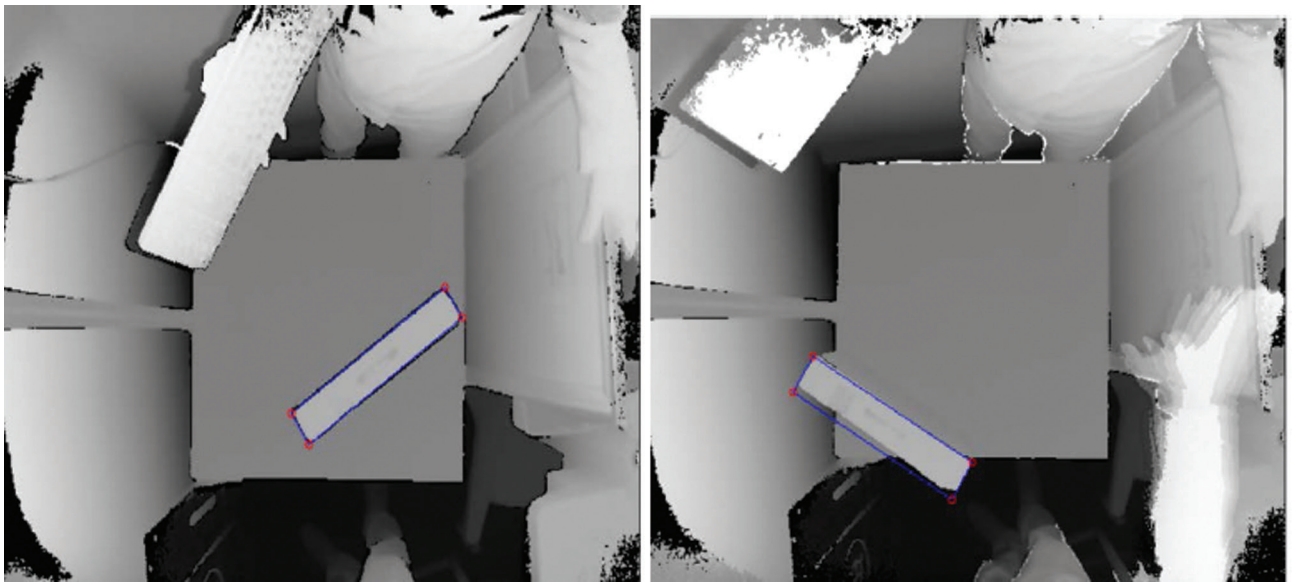


Supplemantary Figure 4. Sample Measurements 4.

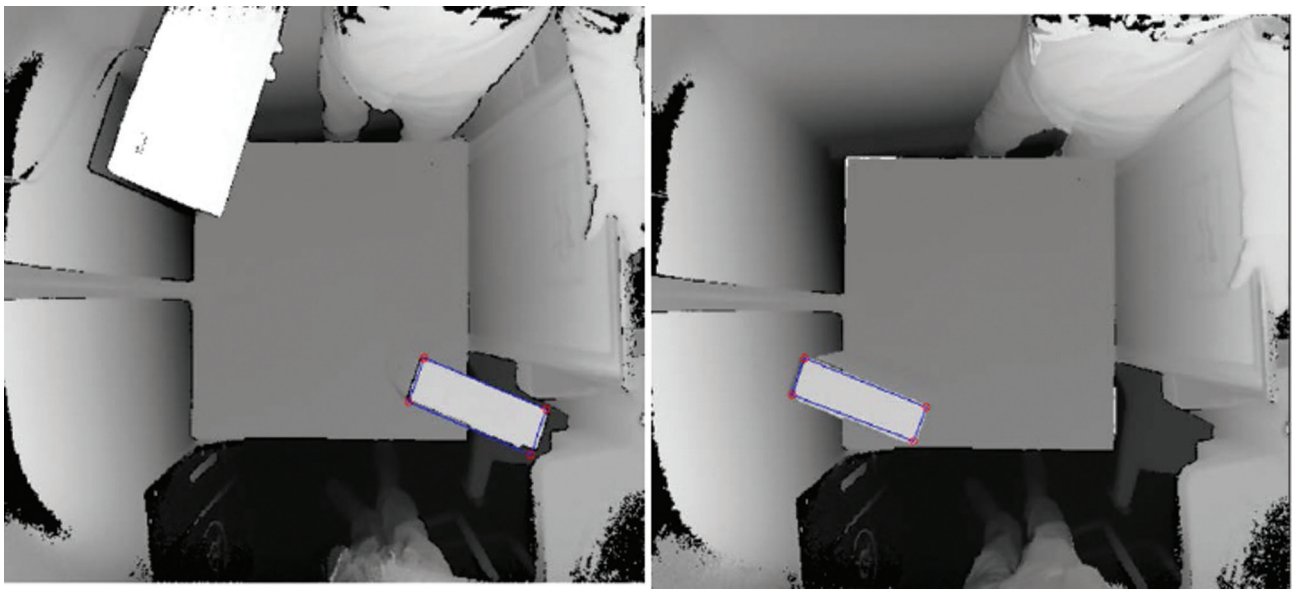
Supplemantary Figure 5. Sample Measurements 5.



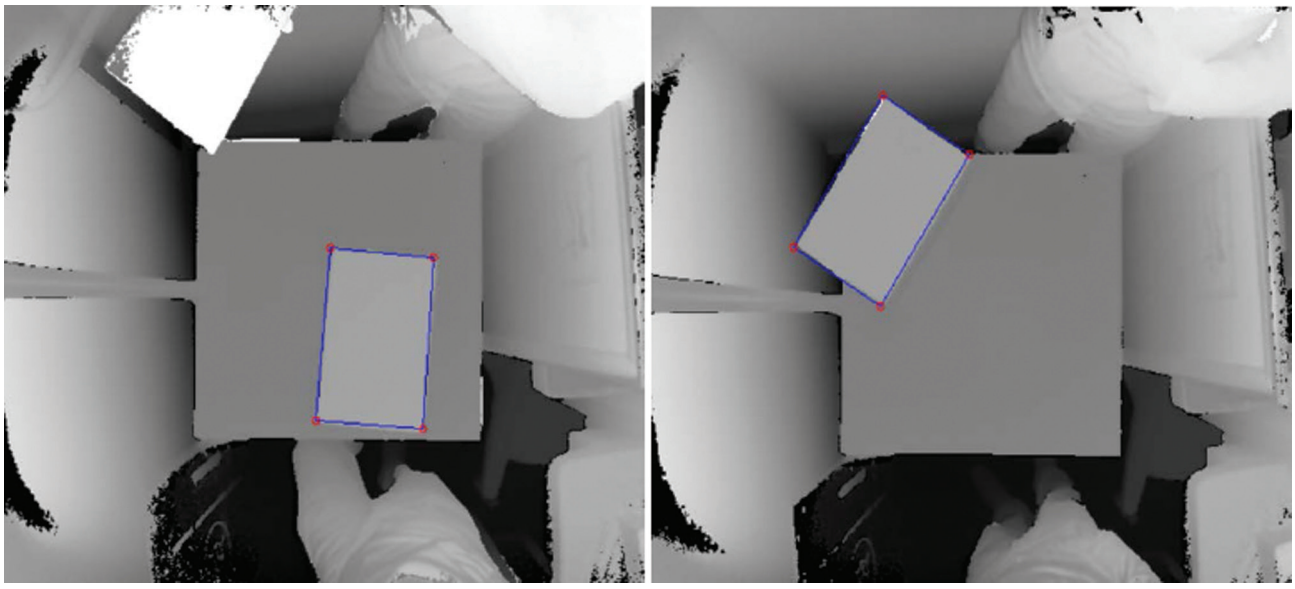
Supplemantary Figure 6. Sample Measurements 6.



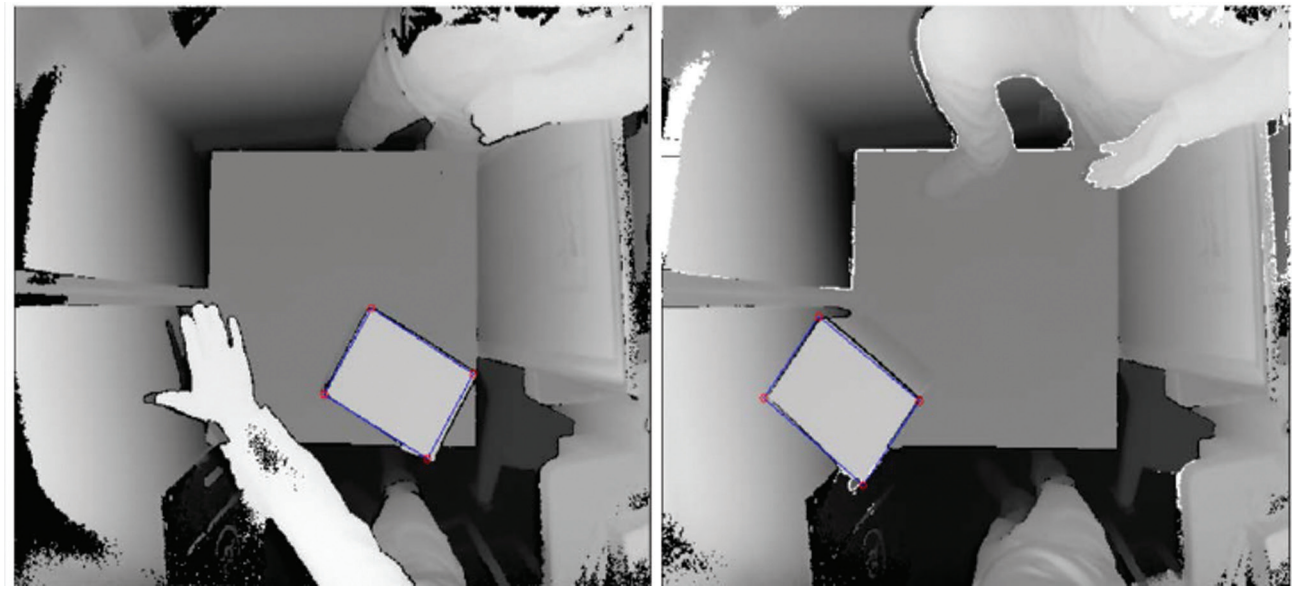
Supplemantary Figure 7. Sample Measurements 7.



Supplemantary Figure 8. Sample Measurements 8.



Supplemantary Figure 9. Sample Measurements 9.



Supplemantary Figure 10. Sample Measurements 10.