ISSN 2347 - 3983
Volume 8. No. 5, May 2020
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter85852020.pdf
https://doi.org/10.30534/ijeter/2020/85852020

# A Study on Determination of Simple Objects Volume Using ZED Stereo Camera Based on 3D-Points and Segmentation Images

**Thien M. Tran[1], Khoa D. Ta[2], Manh Hoang[2], Tung V. Nguyen[2], Nhan D. Nguyen[3], Giao N. Pham[4]**
[1]Dept of Mechanical System Engineering, Pukyong National University,Busan, South Korea, thientm93@pukyong.ac.kr
[2]ICT Department, FPT University, Hanoi, Vietnam, khoatdhe130813@fpt.edu.vn, manhhhe130294@fpt.edu.vn, tungnvhe130151@fpt.edu.vn
[3]Dept. of Biomedical Engineering, Sungkyunkwan University, Suwon, South Korea, nhannd@skku.edu
[4]Dept. of Computing Fundamentals, FPT University, Hanoi, Vietnam, giaopn@fe.edu.vn

## ABSTRACT

Known visual information is needed to acquire exact pose of an object for controlling a picking system. There are several known visual information of objects such as depth, volume, mass, shape and boundary, etc. The volume and the depth of them are very important information for controlling the picking system. Until now, measuring 3D volume and depth rendering from one side of 2D projection using an image processing method are difficult. Previous researches using five cameras captured segmented image to estimate objects volume or concentrated only on 3D volume of objects around the principal axis. However, they did not estimate volume of several picking objects such as cubic and cylindrical shapes. To solve this problem, this paper present an image processing algorithm for the determining volume of simple objects using ZED stereo camera. In the proposed algorithm, 3D-points rendered from 2D projection combined with image segmentation are used to estimate the depth and surface areas of cubic and cylindrical objects, respectively. Experimental results show that object volume is compared with the conventional method measuring the object size by slide calipers.

**Key words:** ZED stereo camera, 3D-points, volume, depth value, segmentation image, image processing.

## 1. INTRODUCTION

Computer vision is defined as technology combined with a computational algorithm and image information. This technology is a crucial part of robotic, automatic system, and industrial robot to pick objects and products that satisfy certain requirements. However, there were several researches to measure depth value in recent year. Depth information was firstly researched by Nevatia et al. [1] to measure the value of 3D objects from an image. Moreover, Mrovlje et al. [2] measured distance based on stereoscopic pictures using special two lens stereo cameras or systems with two single lens cameras. The position of object was calculated by a geometrical derivation method. Typically, 3D point cloud generation from 2D depth camera images using successive triangulation is presented by Pal et al. [3]. Using image segmentation in evaluating volume in food industry and medical field has been a topic of great interest in recent year. Siswantoro et al. [4] presented an irregular shape food product volume measurement using the Monte Carlo method. The weakness of the Monte Carlo method is using five cameras to capture images. So it cannot synchronize image quality between cameras. In medical field, an image processing method for 3D volume rendering from one side of 2D projection applied to cancer spheroids was addressed by Piccinini et al. [5]. However, this method only concentrated on 3D volume of objects around the principal axis. Nevertheless, M. T. Tran et al. described an image processing method measuring injury rate on fish skin by comparing image segmentation between L*a*b* and HSV color space [6].

A picking tool system is integrated with robot kinematics and visual information to obtain orientation information more accurately and efficiently than robot kinematics or visual information alone. Therefore, known visual information obtained from camera is important to control a picking system. There are several known visual information of objects such as depth, volume, mass, shape and boundary, etc. However, 3D volume and depth are very important information for controlling a picking system. Until now, developing an image processing method for 3D volume and depth rendering from one side of 2D projection are difficult. To solve this problem, an image processing algorithm for determining volume of simple objects using ZED stereo camera based on 3D-points and segmented images is proposed. The image processing method consists of three parts, the configuration of an image processing system, 3D-points and a segmentation method. The object volume measured by the proposed method is accurate and close with those measured by the manual method measuring object size by slide calipers. The proposed method is described in detail in Sec. 2. Experimental results are shown in Sec.3, and conclusion will be explained in Sec. 4.

## 2. APPROACHES TECHNICAL PROCESS

### 2.1 Mechanical System of Image Processing

A computer vision system to calculate volume simple object using ZED stereo camera is illustrated as shown in Figure 1. ZED stereo camera has to put perpendicular with basement to capture the good quality image and avoids the scattering of light from the lamp. Moreover, the camera has to be calibrated because of the spherical effect and altitude.

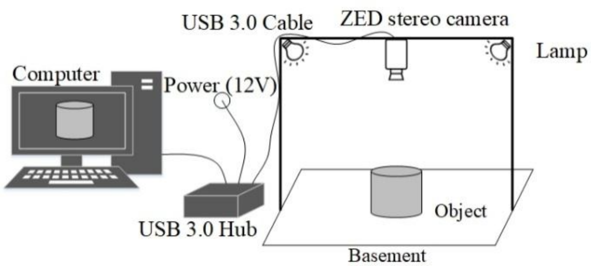
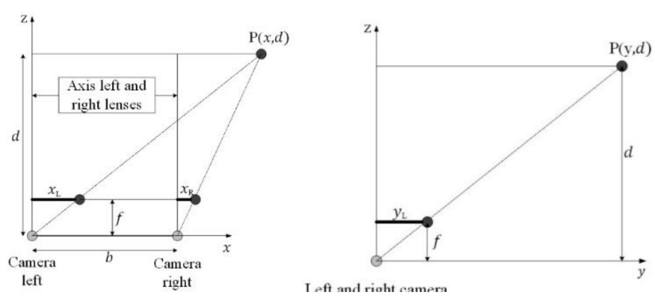



**Figure 1**: Components of the image processing system.

### 2.2 Stereoscopic Method for Measuring an Object Volume

#### *2.2.1 Depth measurement*

In figure 2a, the ZED stereo camera has two lenses of left lens and right lens. A ZED stereo camera with parallel optical axes has focal length baseline b, the depth value d, $x_L$ and $x_R$are distances from the object to the axis left and right camera lenses in image following $x$-axis and y-axis is perpendicular to surface zx.




(a) (b)

**Figure 2**: Triangulation model, (a) respect to depth and $x$ axis, and (b) respect to depth and y axis.

From similar triangles in figure 2a on the left camera and the right camera, the following is obtained:

$$\frac{d}{f}=\frac{x}{x_L}=\frac{x-b}{x_R} \quad (1)$$

In Fig. 2b, the relation between the depth and y axis is obtained as follows:

$$\frac{d}{f}=\frac{y}{y_L}=\frac{y}{y_R} \quad (2)$$

In figure 3, the disparity image is for matching the right image and the left image. The object must be in disparity image so that the depth can be measured. It cannot be measured if the object is in the left or right side of both of lenses camera. On the disparity image, the location of the 3D point P can be derived from Eq. (3) to Eq. (5):

$$d=\frac{f.b}{x_L-x_R} \quad (3)$$

$$x=\frac{x_L.d}{f} \quad or \quad x=b+\frac{x_R.d}{f} \quad (4)$$

$$y=\frac{y_L.d}{f} \quad or \quad y=\frac{y_R.d}{f} \quad (5)$$

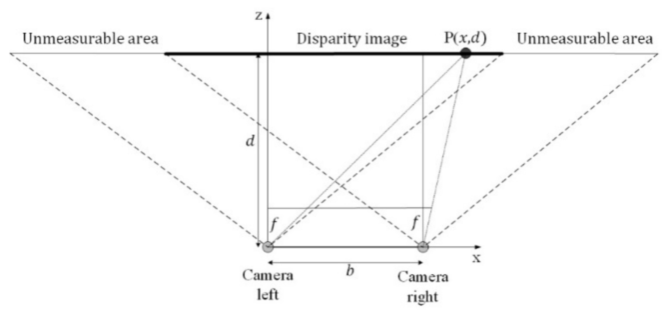



**Figure 3**: Disparity image model of object's depth value.

#### *2.2.2 Image segmentation process*

A ZED stereo camera captures an original color full HD image which is cropped to fix an objects body. After that, Median and Bilateral filters are applied to filter white, salt-pepper and random noises due to environmental light on the cropping color image.

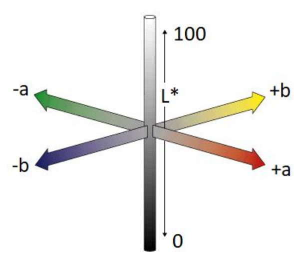



**Figure 4**: L*a*b* color space [7].

L*a*b* color space is defined by CIE and specified by the International Commission on Illumination organization [7]. In figure 4, the central vertical axis is Luminance (L*) which has range of value from 0 to 100 (black to white). The coordinate axes follows channel "a". The value of channel" a" is from positive value as red to negative value as green. Channel "a" and "b" is called chromaticity layers.Median filter can run through each element of the signal and replace each pixel's value with the median of its neighboring. This method is good in dealing with "salt and pepper noise" [8].

$$y[m,n]=median\{x[i,j],(i,j)\in w\} \quad (6)$$

where *w*is a neighborhood defined by user, $x[i, j]$ is a neighborhood element and $y[m,n]$ is a center location in an image.

Bilateral filter not only dissolves noise but also smooth edges. The Bilateral make edges less sharper be removed. This is an advanced Gaussian filter. It represents how two pixels can be close to one another in value. Bilateral filter can keep edges sharp while blurring image [8].

$$g\left(i,\ j\right)=\sum_{k,l} f\left(i+k,\ j+l\right)h\left(k,l\right) \tag{7}$$

where $g\left(i,\ j\right)$ is an output pixel's value, $f\left(i+k, j+l\right)$ is an input pixel value and $h\left(k,l\right)$ is a coefficient of filter defined by a user.

The cropping frame image is transformed into L*a*b* color space. From an original channel "a" separated from L*a*b* color space, a method to adjust value of channel "a" is used to realize injury on fish clearly - an image applying the adjusted value is called as a new channel "a" image. The formula for adjusting value of channel "a" is as follows [8]:

$$g\left(i,j\right)=\alpha\cdot f\left(i,j\right)+\beta \tag{8}$$

where $g(i,j)$ is the output image pixels, $f\left(i,j\right)$ is the source image pixels. The parameters $\alpha>0$ and $\beta$ are usually called the gain and bias parameters.Because the original channel "a" image is separated and adjusted value to become the new channel "a" image, this generates many random noises in the image. Before the new channel "a" image is transferred into binary image by Otsu's method, it is filtered by Gaussian filter to reduce the random noises. This step follows from Eq. (7) to Eq. (9).

For Gaussian filter, pixels located in the middle injury have bigger weight, and the weights decrease with distance from the neighborhood center, so pixels located on sides have smaller weight [8].

$$G_0(x,y)=Ae^{\frac{-(x-\mu_x)^2}{2\sigma_x^2}+\frac{-(y-\mu_y)^2}{2\sigma_y^2}} \tag{9}$$

where $\mu_x$ and $\mu_y$ is the mean, and $\sigma^2_x$ and $\sigma^2_y$ describe variances with respecting to the variables *x* and *y*, respectively.

A binary image is a digital image that has only two possible values for each pixel shown in Eq. (10). Typically, the two colors used for a binary image are black and white. The color used for the objects in the image is the foreground color, while the rest of the image is the background color.

$$Binary\ value(x,y)=\begin{cases}1 & if\ src(x,y)\geq thresh\\ 0 & otherwise\end{cases} \tag{10}$$

After that, the injury binary image is filtered by Median filter to filter salt-pepper noises which is generated in transferring binary image step. This step follows Eq. (7). Counting pixels method is used on injury binary image to count and sum pixels of top objects' area. The total pixels of top objects' area is calculated as follows:

$$Top\ area=\sum_{i=0}^{max\,i}\sum_{j=0}^{max\,j} x\left(i,j\right) \qquad \left(pixels\right) \tag{11}$$

where *max i* is the height of image and *max j* is the width of image. $x\left(i,j\right)$ is the pixels of top areas.

The unit of the object's top area is pixels after the calculation of Eq. (11). Fig. 5 shows a relationship between millimeters and pixels by calibrating "chess king" method.From section 2.2.1, the objects' depth value is calculated by Eq. (3), whose unit is millimeters. Area of the objects are calibrated by segmentation image process whose unit is millimeters as shown in figure 5.Volume of simple objects is calculated as following:

$$Volume=d\cdot(\,top\ area\,) \tag{12}$$

The tolerance of information objects is calculated as following:

$$Tolerance=\frac{original\ data-proposed\ data}{original\ data}\times 100(\%) \tag{13}$$

Figure 6 shows the flowchart of the proposed method.

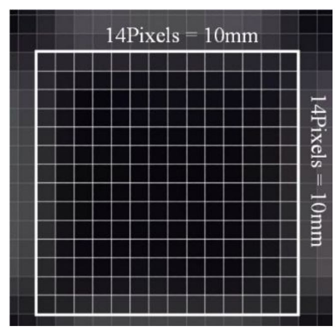



**Figure 5:**Calibration between millimeters and pixels.

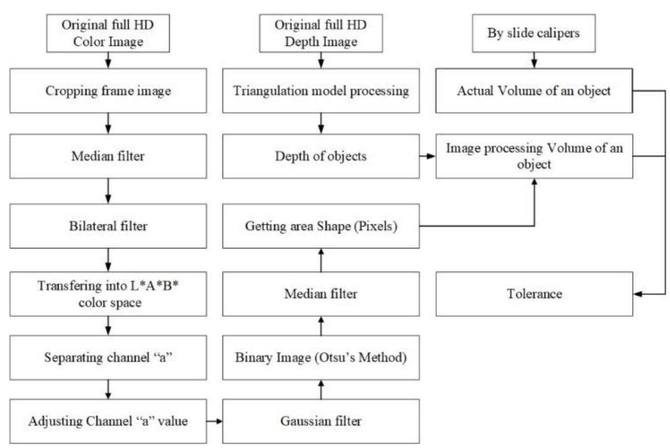



**Figure 6**: Flowchart of the proposed method.

## 3. EXPERIMENTAL RESULTS

### 3.1 Depth Measurement

Original full HD color images are captured by ZED stereo camera as a cubic object and a cylindrical object shown in figure 7 and figure 9, respectively. And then use a 3-D point cloud that is a collection of 3D points (*x*, *y*, *d*) mentioned in section 2.2.1. Moreover, each object is measured more 5 times to reduce error of depth information during experimental time. Table 1 shows the simple specification of components of the proposed image processing system.Figure 8 and figure 10 show the depth measurement process of a cubic object and cylindrical object, respectively. Figure 8a and figure 10a show the depth map of 3D point cloud image. Markers A-B-C-D-E-F contain pixels of a depth value of the top surface of an object and makers X-Y-W-Z present pixels of a depth value of basement. However, the depth value is easily effected by external noises such as light, shape, electric power source of camera, etc. Therefore, several points are marked and taken on surface of objects and basement to reduce the error of the depth value. To match the depth image and the original HD color image, the marker points on object are observed clearly as shown in figure 8b and figure 10b. In figure 8c and figure 10c, the 3D point cloud can be seen clearly as a depth map in three dimensions (*x*, *y*, *d*).

**Table 1:** Specification of components

| Components | Specifications |
|---|---|
| ZED stereo camera | Visual distance: 0.3-20 meters |
| Lamp | White LED 8W |
| USB 3.0 Hub | Source: 12V |
| Computer | Intel Core i7-6700 – RAM: 16GB |

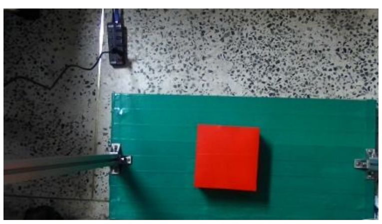


**Figure 7:** Original HD color image of a cubic object.

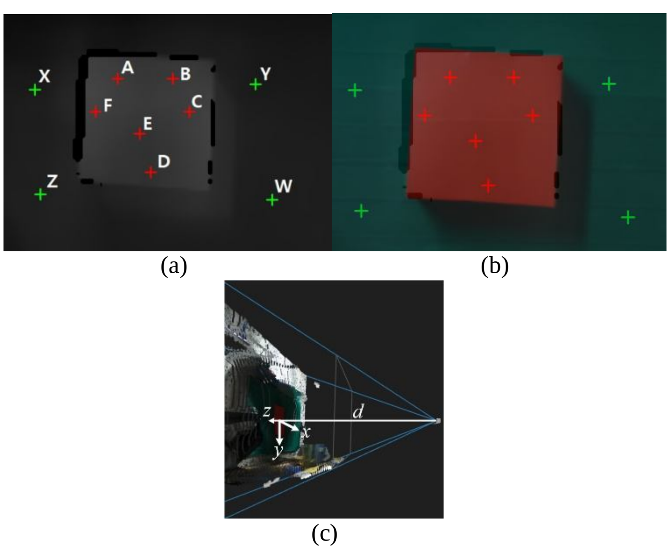



**Figure 8:** Depth measurement process of a cubic object, (a) depth map of 3-D point cloud image, (b) matching with HD color image and depth image, and(c) 3-D point cloud image.

The depth data cubic objects is shown in table 2. The average distance from camera to 6 markers on the top surface of the object are shown measurement time. After 5 measurement times, the average distance from camera to the top surface of cubic object is 1010.3 mm. Similarly to the top surface, the average distance from camera to basement is 1108.8 mm. Average depth value of a cube object is 98.5 mm as the difference between the surface and the basement.

**Table 2:** Data of depth information of a cubic object (mm)

| Measurement times | Experimental results | |
|---|---|---|
| | Cubic object surface (A-B-C-D-E-F) | Basement surface (X-Y-W-Z) |
| 1 | 992-1020-1008-1034-1017-995 | 1092-1123-1127-1090 |
| | Average: 1011 | Average: 1108 |
| 2 | 995-1015-1002-1031-1012-992 | 1093-1121-1125-1092 |
| | Average: 1007.8 | Average: 1107.8 |
| 3 | 997-1013-1005-1034-1024-990 | 1094-1126-1123-1091 |
| | Average: 1010.5 | Average: 1108.5 |
| 4 | 994-1018-1007-1030-1018-994 | 1096-1123-1126-1095 |
| | Average: 1010.1 | Average: 1111 |
| 5 | 995-1016-1015-1033-1022-994 | 1094-1125-1126-1091 |
| | Average: 1012.5 | Average: 1109 |
| Average | 1010.3 | 1108.8 |

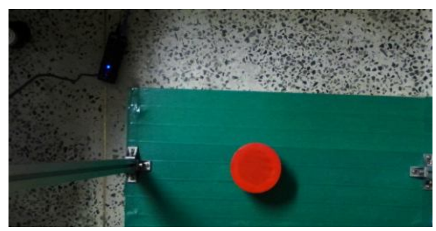


**Figure 9:** Original HD color image of a cylindrical object.

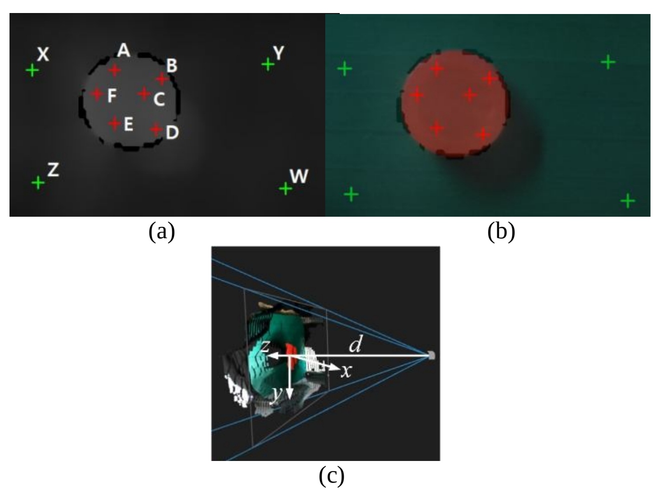



**Figure 10:** Depth measurement process of a cylindrical object,(a) depth map of 3-D point cloud image,(b) matching with HD color image and depth image, and(c) 3-D point cloud image.

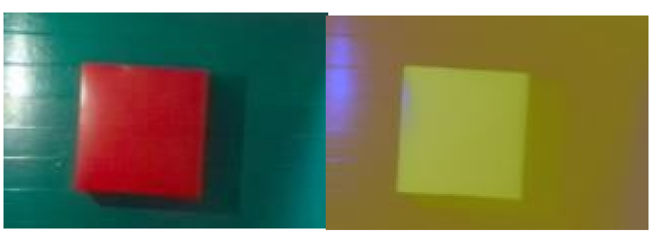
With similar steps to cubic object, the depth data of cylindrical objects is shown in Table 3. The average distances from camera to basement and to cylindrical object surface are is 1103.8 mm and 997.7 mm, respectively. Average depth value of cylindrical object is 106.1 mm as the difference between the surface and the basement.

**Table 3:** Data of depth information of a cylindrical object (mm)

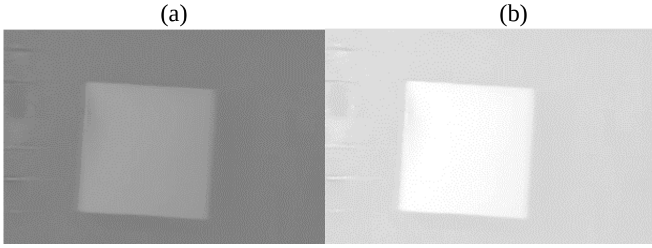
| Measurement times | Experimental results | |
|---|---|---|
| | Cylindrical object surface (A-B-C-D-E-F) | Basement surface (X-Y-W-Z) |
| 1 | 997-1004-998-994-1002-994 | 1081-1125-1136-1088 |
| | Average: 998.1 | Average: 1107.5 |
| 2 | 1000-1008-996-993-995-990 | 1093-1101-1121-1080 |
| | Average: 997 | Average: 1098.8 |
| 3 | 998-1006-992-991-999-996 | 1080-1124-1127-1080 |
| | Average: 997 | Average: 1102.8 |
| 4 | 999-1002-997-991-995-987 | 1089-1114-1124-1099 |
| | Average: 995.1 | Average: 1106.5 |
| 5 | 1003-1013-1003-995-1000-995 | 1082-1108-1130-1093 |
| | Average: 1001.5 | Average: 1103.3 |
| Average | 997.7 | 1103.8 |

### 3.2 Area surface measurement

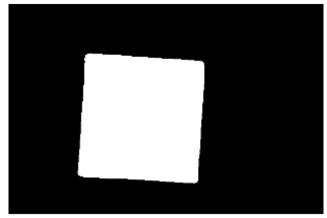
There are two objects of a cubic object and a cylindrical object which are tested by the proposed process. An original full HD color image is captured by ZED are processed by the flowchart of Fig. 6 and all steps are presented in the section 2.2.2 are shown in figure 11a ~ figure 11e, and figure 12a ~ figure 12e, respectively. The experimental results of the cubic object are shown in figure 11. Figure 11 and figure 12 show the area measurement processes of a cubic object and cylindrical object, respectively. Table 4 and table 5 show the areas of top surfaces of the cubic object and cylindrical object. They are 53734 pixels and 20961.6 pixels, respectively.



(a) (b)



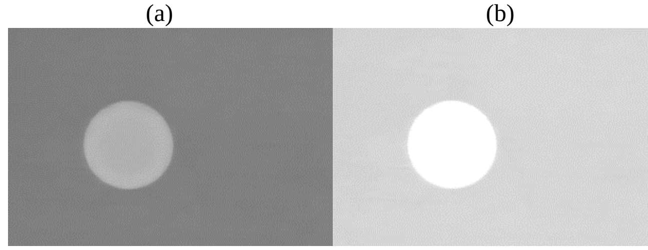
(c) (d)



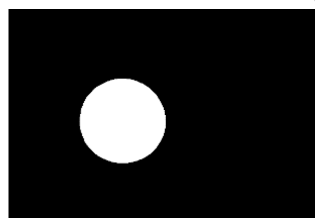
(e)

**Figure 11:** Are measurement result of a cubic object, (a) cropped original HD color image,(b) L*A*B* color space image,(c) original channel “a” image,(d) new channel “a” image, and(e) binary injury image after being filtered by being Median filter.

(a) (b)



(c) (d)



(e)

**Figure 12**: Are measurement result of a cylindrical object, (a) cylindrical object cropped original HD color image,(b) L*A*B* color space image,(c) original channel “a” image,(d) new channel “a” image,and (e) binary injury image after being filtered being by Median filter.

**Table 4:** Area of top surface of a cube object (pixels)

| Measured times | Experimental results |
|---|---|
| 1 | 53730 |
| 2 | 53723 |
| 3 | 53759 |
| 4 | 53745 |
| 5 | 53711 |
| Average | 53734 |

**Table 5:** Area of top surface of a cylindrical object (pixels)

| Measured times | Experimental results |
|---|---|
| 1 | 20961 |
| 2 | 20953 |
| 3 | 20956 |
| 4 | 20975 |
| 5 | 20963 |
| Average | 20961.6 |

### 3.3 Volume Measurement

The volumes of the cubic and cylindrical objects are calculated with the data measured actually by slide calipers are compare with their volumes measured by our proposed image processing method as shown in figure 13.The specification of the slide calipers is 300 ± 0.05mm.Table 6 shows the results of the volumes of the cubic object and cylindrical object measured by the proposed method are 270407.65 $mm^3$ and 1134707.67 $mm^3$, respectively. Their volume measured by the slide calipers are 2755500 $mm^3$ and 1112059.6 $mm^3$, respectively. The tolerance between of them are approximately ±2% in both objects. The result of the cubic depth is 98.5 mm in the proposed method and 100 mm in slide calipers method. The cylindrical depth are 106.1 mm and 110 mm, respectively. The tolerance of the cubic area and the cylindrical area are 0.51% and -5.78%, respectively.

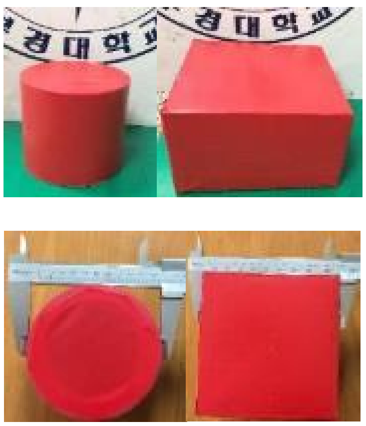


**Figure 13**: Cubic object and cylindrical object measured by slide calipers.

**Table 6:** Depth, area and volume measured by the proposed method and the manual method.

| Objects | | By slide calipers | By image processing | Percentage of tolerance (%) |
|---|---|---|---|---|
| Cube | Vol. | 2755500 $mm^3$ | 2700407.65 $mm^3$ | 1.999 |
| | Depth | 100 mm | 98.5 mm | 1.5 |
| | Area | 27555 $mm^2$ | 27415.31 $mm^2$ | 0.51 |
| Cylinder | Vol. | 1112059.6 $mm^3$ | 1134707.67 $mm^3$ | -1.996 |
| | Depth | 110 mm | 106.1 mm | 3.54 |
| | Area | 10109.63 $mm^2$ | 10694.7 $mm^2$ | -5.78 |

## 4. CONCLUSION

In this paper, a study on determination of objects volume using ZED stereo camera based on 3D-points and image segmentation was presented. 3D volume and depth rendering from one side of 2D projection for a cubic object and a cylindrical objects were calculated. Moreover, the average method was applied in the proposed image processing method. The depth was 98.5mm in a cubic object and 106.1mm in a cylindrical object and close to the real depths of 100mm and 110mm measured by the slide calipers, respectively. The experimental results verified that the volume measured by the proposed image processing method was close to the real volume whose tolerance was 1.999% in a cubic object and -1.996% in a cylindrical object. In further study, a 3D image processing method for measuring the depth and volume of large and complicated shape using a stereo camera should be considered. Moreover, combing 2D and 3D image processing methods integrated with interface monitoring on the picking system will be done.

## ACKNOWLEDGEMENT

This work is supported FPT University, Hanoi, Vietnam;Pukyong National University, Busan, South Korea; and Sungkyunkwan University, Suwon, South Korea.

## REFERENCES

1. R. Nevatia.**Depth measurement by motion stereo**, *Computer Graphics and Image Processing*, Vol.1976, No. 5, pp. 203-214, 1976. https://doi.org/10.1016/0146-664X(76)90028-9
2. J. Mrovlje and D. Vrancic.**Distance measuring based on stereoscopic pictures**, i*n Proc. 9th International PhD Workshop on Systems and Control*, Izola, 2008.
3. B.Pal, S. Khaiyum and Y. S. Kumaraswamy.**3D point cloud generation from 2D depth camera images using successive triangulation**, *in Proc. IEEE International Conference on Innovative Mechanisms for Industry Applications*, 2017, pp.129-133. https://doi.org/10.1109/ICIMIA.2017.7975586
4. S. Joko, S. P. Anton, A. Azizi andI. Bahari.**Monte Carlo methos with heuristic adjustment for irregularly shaped food product volume measurement**, *The Scientific World Journal*, Vol. 2014, pp. 1-10, 2014. https://doi.org/10.1155/2014/683048
5. P. Filippo, T. Anna, Z. Wainer andB. Alessandro.**Image processing method for 3D volume rendering from one 2D projection: Application to cancer spheroids,***in Proc. 4th International Conference on Image Processing Theory*, Tools and Applications, Paris, 2014.
6. M. T. Tran, J. Rantung, T. H. Nguyen, H. K. Kim, andS. B. Kim.**Measurement of injury rate on fish skin and performance comparison based on L*A*B* and HSV color spaces,** *in Proc.IJCAET & ISAMPE*, Bali, 2017, pp.1-6.
7. R. S. Hunter.**Photoelectric color difference meter***. *Journal of the Optical Society of America,*Vol. 48, No. 12, pp. 985-995, 1958. https://doi.org/10.1364/JOSA.48.000985
8. **The OpenCV Tutorials**, Release 2.4.13.2.