BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Journal of Electronic & Information Systems
https://ojs.bilpublishing.com/index.php/jeis

ARTICLE

3D Reconstruction of Fruit Shape based on Vision and Edge Sections

Nasr Abdalmanan Nasr Ali[1] Kamarulzaman Kamarudin[1*] Chee Kiang Lam[1] Muhamad Safwan Muhamad Azmi[2] Abdul Halim Ismail[1] Norasmadi Abdul Rahim[1] Wan Mohd Nooriman Wan Yahya[1] Goh Kheng Sneah[3] Moey Lip Seng[3] Teoh Phaik Hai[3] Ong Thean Lye[3] Noor Zafira Noor Hasnan[4]

1. Faculty of Electrical Engineering Technology, Universiti Malaysia Perlis, Kampus Alam Pauh Putra, Arau, Perlis, 02600, Malaysia
2. Faculty of Mechanical Engineering Technology, Universiti Malaysia Perlis, Kampus Alam Pauh Putra, Arau, Perlis, 02600, Malaysia
3. Walta Engineering Sdn. Bhd., Prai, Pulau Pinang, 13600, Malaysia
4. Department of Process and Food Engineering, Faculty of Engineering, Universiti Putra Malaysia, UPM Serdang, Selangor, 43400, Malaysia

ARTICLE INFO

Article history
Received: 1 April 2022
Accepted: 28 April 2022
Published Online: 13 May 2022

Keywords:
3D Reconstruction
Machine vision
Fruit processing

ABSTRACT

The fruit industry has been known as one of the largest businesses in Malaysia, where most of the fruits pass through the peeling process well in advance before the final product as juice in a bottle or slices in a can. The current industrial fruit peeling techniques are passive and inefficient by cutting parts of the pulp of the fruit with peels leading to losses. To avoid this issue, a multi-axis CNC fruit peeler can be used to precisely peel the outer layer with the guidance of a 3D virtual model of fruit. In this work, a new cost-effective method of 3D image reconstruction was developed to convert 36 fruit images captured by a normal RGB camera to a 3D model by capturing a single image every 10 degrees of fruit rotation along a fixed axis. The point cloud data extracted with edge detection were passed to Blender 3D software for meshing in different approaches. The vertical link frame meshing method developed in this research proved a qualitative similarity between the output result and the scanned fruit in a processing time of less than 50 seconds.

1. Introduction

Malaysia is one of the world's leading exporters of tropical fruits. Malaysia, as the world's largest producer of durian, mangosteen, and starfruit, produced 1.49 million metric tonnes in 2017 and exported roughly 268,400 metric tonnes [1]. Fruits were the most often purchased organic

*Corresponding Author:
Kamarulzaman Kamarudin,
Faculty of Electrical Engineering Technology, Universiti Malaysia Perlis, Kampus Alam Pauh Putra, Arau, Perlis, 02600, Malaysia;
Email: kamarulzaman@unimap.edu.my

DOI: https://doi.org/10.30564/jeisr.v4i1.4585

food by Malaysian consumers, according to a survey done by Wong & Aini in Klang Valley [2]. Therefore, fruits are progressively becoming a significant part of Malaysia's agricultural production. Stripping is one of the post-harvest operations used to prepare horticulture crops and tubers for processing, and peeling is an important element of the food preparation process during the early phases of food processing. Reducing peeling loss is one of the most important objectives in the fruit industry. Therefore many studies have been conducted in this particular field to improve fruit peeling machines [3]. In this research an approach of utilizing machine vision and image processing was introduced in the pre-peeling process to optimize peeling quality by digitally reconstructing a 3D model of the real fruit to guide a CNC peeler. 3D image reconstruction is the process to convert a series or a single 2D image to a 3D model. Many different techniques were introduced to perform this task, nevertheless, they all share the same concept of staging from image acquisition to image reconstruction [4].

3D Image Reconstruction Methods

Laser triangulation scanning for 3D image reconstruction is a setup of a single fixed camera with low exposure settings and a linear (vertical line) laser is usually fixed in tangent to the camera circumference towards the object in the centre. The object is rotatable on a perpendicular axis with respect to the setup plane. The system's concept is to measure the object in 3D space by optically measuring the distortion of the laser line casted on the object. Due to its geometry, the object needs to be rotated in a number of degrees and capture multiple scans to accurately get the shape of the object. However, this process is relatively time consuming, but the time could be minimized by increasing the number of laser modules within the range of the camera view to capture more geometry data in one rotation [5].

An advanced method of optical laser 3D reconstruction is based on Structured Light 3D scanning. Instead of using laser beams, the system utilises a light projector fixed next to the camera. The light projector will cast a series of high contrast light stripes patterns on the object and the processing part is to create a point cloud data based on the distortion of the light pattern casted on the object geometry captured by the camera [6]. Zhiping Xie performed a 3D image reconstruction of the Rosa Roxburghii fruit using structured light scanning method, where the point cloud processing algorithm developed achieved a volume error of 1.06% of the original fruit [7].

Another approach in 3D reconstruction field is using time-of-flight based RGB-D sensor to capture the depth information of a 2D scene. RGB-D camera consists of normal RGB camera and a depth sensor, where the depth sensor comprises of two major components: emitter and a receiver. The emitter blasts an infrared laser beam that hit the object and reflects back to the infrared receiver. Based on the delay time between emitting and receiving process, it is possible to calculate the distance between the sensor and the object acknowledging the fixed speed of light is 3×10^8 *m/s*. This process of emitting, receiving and calculating is to be done rapidly for each pixel of the RGB camera register. Therefore the overall system returns the data of Red, Green, Blue, and Depth for each pixel [8]. The combination of RGB camera and depth sensor allow the system to scan an object's geometrical data along with texture and colour data [9]. Satoshi Yamamoto et al., attempted to reconstruct 3D images of an apple fruit using Kinect RGB-D sensor [10]. However, such methods are relatively complex to compute or process, and the cost of infrared pattern sensors, light structure 3D scanners and laser scanners are relatively high compared to only optical solutions. Therefore Mitchell J. Feldmann approach a low-cost 3D reconstruction with stereo vision system, which is a setup of two RGB cameras with a fixed distance in between to extract depth information with aid of trigonometry based on the object optical features [11].

Feldmann utilised a multi-view stereo method to scan fruits by capturing multiple images of the fruit from different angles with overlapping and created a point cloud by extracting the features of the images. They successfully matched the overlapped features points to be measured and obtained in 3D space [12]. An advanced technique of stereo vision is Photogrammetry, which is the method of three-dimensional estimation and measurement of an object in space by capturing a series of RGB or Greyscale images of the object from various angles [13]. The camera or multiple cameras are spherically moved around the object and the processing part is responsible for collecting the features points of intersecting between the images sequence. Thus, able to measure the position of the camera at every shot relative to the object and estimate the point cloud in space. This method not only capture the geometrical information of the object but also carries the object texture information [14], however, the scanned object surface is required to have a relatively rough texture with various features patterns. Nevertheless if the object texture is transparent, highly reflective or shiny, it would lead to failing the process of images stitching due to lack of visual features [15]. Another factor of photogrammetry precision is environment illumination since the overall method is completely depends on optical sensory only [13]. An attempt of strawberry fruit 3D measuring system was conducted by Nobuo Kochi et al., utilizing photogrammetry method. A three-camera system was arranged vertically to obtain levelled view, above levelled view and below levelled view. The strawberry fruit is mounted on a rotational disc with a solid blue background colour. The overall process of image aquation, image processing and 3D modelling required 90 minutes to scan a single strawberry fruit with a matching error of 0.6 mm or less in 90% of the trials [16]. Table 1 compares the previous works of fruit 3D reconstruction. The key characteristics are the system cost and the complexity of the processing method.

This work proposes a method for 3D reconstruction of fruits based on edges sections with only optical sensing. A single RGB camera was used to extract the fruit geometry features via edge detection and convert it to point cloud data in three-dimensional space. The point cloud data were processed in Blender software and the results were evaluated in terms of the shape of the reconstructed 3D image.

2. Methods

Started by setting up the fruit on a rotating axis driven by a stepper motor as foreground, while the background is a solid colour plane, an RGB camera perpendicularly fixed to the axis is used to acquire the raw data set, which is a total of 36 images of the fruit taken from every 10 degrees of axis rotation. The reason of using a single RGB camera is to reduce the setup costs and complexity from the previous methods.

The current goal is to mask the fruit out from the background since the background is a solid chroma key colour. The alternative is to mask the background and then invert the mask values, however even though the background is in a single colour, the shadows and environment lightings will create a slight variation in colour hue and saturation which is very complex to process in RGB colour format. To begin the process, the images are converted from RGB colour format to HSV colour format. In RGB colour format every single pixel in the image is to be represented as a combination of (Red, Green and Blue) while in HSV colour format a single pixel is to be represented via three components Hue, Saturation and Value [17], this representation eases the work by clustering all Hues colours in one dimension instead of three dimensions, therefore selecting the range of the colour mask is less complex.

Where, $H = \cos^{-1}\left[\frac{0.5\,[\,(R-G)+(R-B)]}{\sqrt{(R-G\,)^2+(R-B\,)(G-B\,)}}\right]$, $S = 1 - \frac{\min\,(R+G+B)}{V}$, and $V = \frac{R+G+B}{3}$.

The mask result is representing the separation of the fruit in the foreground and the image background. Since the fruit is rotating on axis, there is no need of using the full image and only one half is enough. There is no significant difference between the top or bottom half of the image since the axes are centred and the fruit is rotating. In order to define the outer contour of the mask, a method of contrast detection is used, where a weighted 3 by 3 kernel mask is used for the convolution of the binary image. This process is performed to detect the pixels that has abrupt changes in intensity with respect to its 8-neighbours, defined as an edge pixel. The x and y image coordinates of edge pixels are then stored in a variable.

To represent the edge pixels as a point cloud in three-dimensional space, a method of double conversion is used. The first step is to convert the edge pixels coordinate system from (x, y) to (r, θ, z) also known as cylindrical coordinates system, where a point in space is to be represented by three components (i.e., r radius from axis line to the point, θ is the angle of vector according to the axis and z is the elevation of the point).

Firstly, the range of z from z_1 to z_2 is determined, where

Table 1. Fruit 3D reconstruction prior approaches summary

Author	Sensory	Sensing method	Reconstruction technique	Findings
Zhiping Xie	3D laser scanner	Structured Light Scanning	Point cloud processing	The results have high accuracy (98.94%). The sensor is relatively costive.
Satoshi Yamamoto et al	Kinect RGB-D camera	Time of Flight	Point cloud processing	The setup cost is good in value. Relatively complex processing due to sensor accuracy.
Mitchell J. Feldmann	RGB cameras	Stereo Vision	Point cloud via image segmentation	Low-cost system. Relatively lower results accuracy. To enhance the accuracy additional cameras could be added to the system leading to a cost increase.
Nobuo Kochi et al	RGB cameras	Photogrammetry	Point cloud via features extractions	High step cost due to usage of multiple DSLR cameras. High accuracy results. Relatively complex/ slow processing.

z_1 could be defined as the x coordinate of the first point of the edge, and $z_2$is the x coordinate of the last point of the edge. Then to determine the r value which is the vertical distance between the edge point and the axis, the absolute value of edge pixel (y component) is subtracted by axis (y component).

$$r = |EP[1] - axis[1]|$$

The y-axis component reference is equal to the y coordinate of the edge pixel pair that has the maximum y value in the list variable. Finally, angle θ is determined, which is equal to the angle of the axis in the current image. After establishing all the previous rules to calculate r, θ, and z components of each edge pixel, the data is re-registered as a group in a variable list to be called a section.

The next step is to convert all sections points from cylindrical coordinates system (r, θ, z) to cartesian coordinates system (x, y, z)

$$x = r\ \cos\theta\,,\ \ y = r\ \sin\theta$$

The next procedure is to call the data, section by section in Blender software, where in each call, every point in the section is being plotted in space. Geometrical link is established between every subsequent point until all points in the section are linked (i.e., 1 to n points). The result of this process is a frame of connected point for each section as shown in Figure 1. The reason of connecting the relative points together vertically in each section is to aid the automated meshing process, where it generates a geometrical plane over the points and the linking lines guiding the algorithm to define the nearest shape. Then the observation method was used to verify the algorithm by observing the output image.

3. Results

The fruit in Figure 2 below is used to test the method. After the image processing operations were applied and all the sections points have been extracted, several meshing techniques were applied to obtain the 3D model. The first approach was by calling all the sections data at one time in Blender and directly generate the mesh using the python script with Blender to run the faces generator command. The faces generation algorithm in blender is a function that generates a geometrical plane that passes through three or more points in the point cloud object. Each point in the point cloud has a unique id in the memory, however when all the points are called at once, all the points will share the same id as one object. Therefore, the face generation function will attempt to randomly generate faces passed through the points leading to a random meshing as illustrated in Figure 3.

The second attempt was to call all the sections at once,

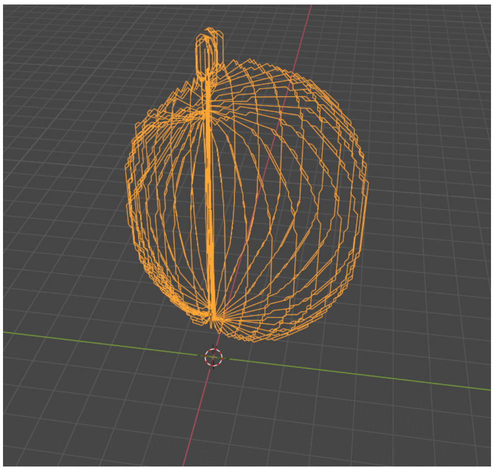
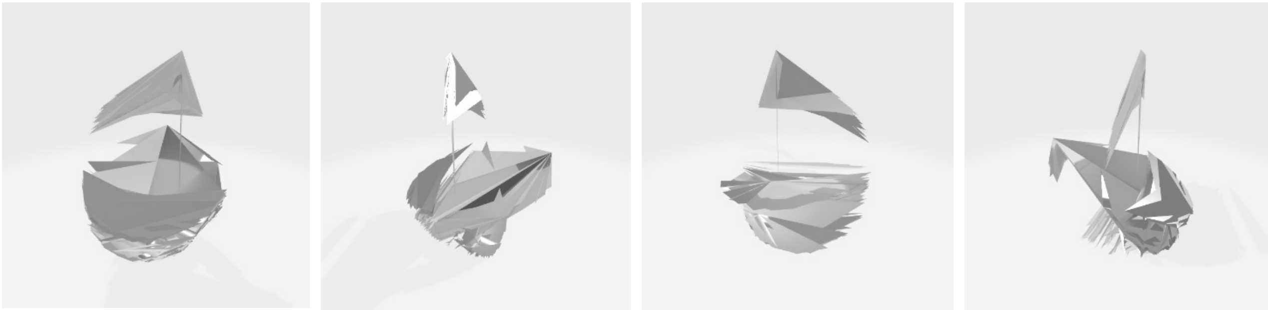


Figure 1. Framing process.

Figure 2. Apple fruit rotate around axis with solid colour background setup.



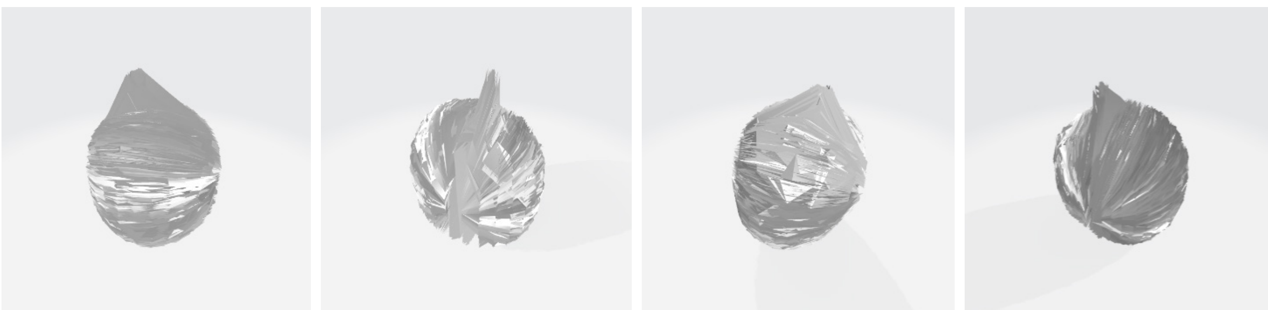
(a) **(b)** **(c)** **(d)**

Figure 3. Result of attempt 1 (a) front side view, (b) right side view, (c) back side view, (d) left side view.

followed by link framing and face generation operations. The framing operation was performed before filling faces in order to help to hold a structure that the faces will be constructed on. This way, the blender software would be guided with paths to perform face generation function (i.e., placement of the face shape, size, direction, and orientation). In this attempt, the points linking was complete but in a random way, means that every point is linked by a wire to all the others point which created a structured like a wool ball. Then the face generation algorithm worked on covering the structure with geometrical planes. However, the images rendered in Figure 4 are not similar in shape to the test sample.

The next attempt was to call each section alone and operate a vertical frame linking algorithm, then eventually generate the faces. This means that every section will have a unique geometry id ranging from 0 to 35, where the total of all sections is 36. Instead of connecting all the points to each other, a method of connecting each section points only as group is used. This will function as a guide to the face generation algorithm to generate faces along the sections vertically instead of random. However due to the relatively large number of sections, the face generation function crashes at the middle of the process. The program was not able to continue generating the entire sections leaving a half empty frame without meshing as shown in Figure 5.

The final attempt was to call each section alone, followed by operating a vertical frame linking algorithm. The face generator was applied before calling the next section



(a) **(b)** **(c)** **(d)**

Figure 4. Result of attempt 2 (a) front side view, (b) right side view, (c) back side view, (d) left side view.

(a) **(b)** **(c)** **(d)**

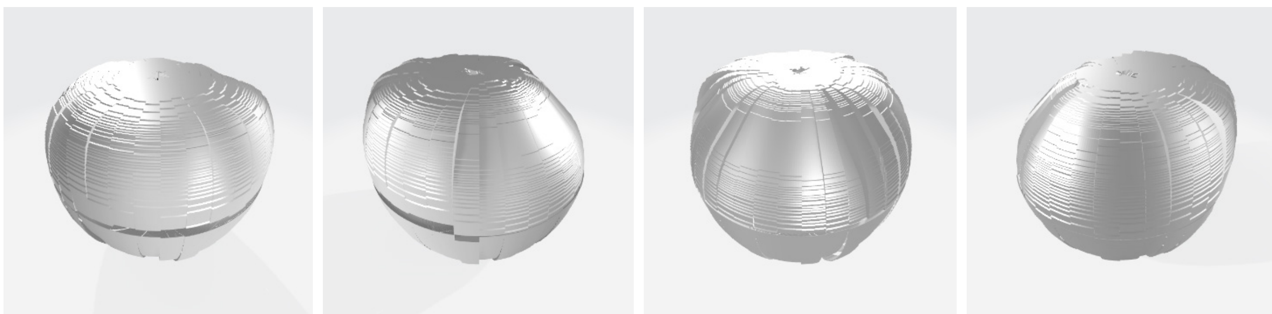
Figure 5. Result of attempt 3 (a) front side view, (b) right side view, (c) back side view, (d) left side view.

so that the face generation command shall deal with relatively small data at a time. This process took longer than usual due to operation of face generation command executed 36 times between calling the sections, however the result turned better than previous approaches where the generated mesh is relatively similar to the test object as shown in Figure 6.

Table 2 shows the summarized results of four attempts of meshing sequences in Blender software. The first three attempts showed distorted images of the input as the Blender software failed to build the correct surfaces without proper guides. However, the fourth attempt was successful as the output 3D image was less distorted and relatively similar to the shape of real fruit. However, the most computation time took among all four attempts is the fourth one.

4. Conclusions

To conclude, this work discussed the possibility of 3D fruit image reconstruction with one camera by the developed new method of sections slicing and meshing. Only qualitative analysis is included in the scope of this study. This algorithm is promising and have the possibility to reconstruct a 3D fruit model in Blender software with the tuneable parameters of frame generation and face generation, however the method proposed is limited to only axial symmetrical shaped fruits. In addition, it was found that, the face generator needs to be applied step by step for each section in order for the Blender software to generate more accurate 3D fruit reconstruction.



(a) **(b)** **(c)** **(d)**

Figure 6. Result of attempt 4 (a) front side view, (b) right side view, (c) back side view, (d) left side view.

Table 2. Overall results summary

Attempt	Sections calling	Framing	Face generation technique	Process time	Visual representation
First	All at once	No	Random / All at once	3 Seconds	Distorted
Second	All at once	Yes / Random	Along frame / All at once	7 Seconds	Distorted
Third	One at once	Yes / Vertical	Along frame / All at once	9 Seconds	Distorted
Fourth	One at once	Yes / Vertical	Along frame / One at once	44 Seconds	Less distorted

Conflict of Interest

There is no conflict of interest.

Acknowledgement

The authors would like to acknowledge the support from the University-Private Matching Fund (UniPRIMA) from the Research Management Centre, UniMAP and Walta Engineering Sdn. Bhd.

References

[1] Kamarubahrin, A.F., Haris, A., Mohd Daud, S.N., et al., 2020. Indigenous alternate fruit for dates: Critical review on the selection criteria for local prophetic fruits in Malaysia. Journal of Tropical Agriculture and Food Science. 48(2), 43-58.

[2] Wong, S.S., Aini, M.S., 2017. Factors influencing purchase intention of organic meat among consumers in Klang Valley, Malaysia. International Food Research Journal. 24(2), 767-778.

[3] Kohli, D., Champawat, P.S., Mudgal, V.D., et al., 2021. Advances in peeling techniques for fresh produce. Journal of Food Process Engineering. 44(10), e13826.

DOI: https://doi.org/10.1111/JFPE.13826
[4] Ham, H., Wesley, J., Hendra, H., 2019. Computer vision based 3D reconstruction: A review. International Journal of Electrical and Computer Engineering. 9(4), 2394-2402.
DOI: https://doi.org/10.11591/ijece.v9i4
[5] Wijayanto, I., Hadiyoso, S., Gunawan, T.M., 2018. Video Processing for Dual Laser 3D Scanner Prototype based on Cloud to Cloud Method. In 2018 International Conference on Control, Electronics, Renewable Energy and Communications (ICCEREC). pp. 149-153. IEEE.
DOI: https://doi.org/10.1109/ICCEREC.2018.8712107
[6] Zhang, S., 2018. High-speed 3D shape measurement with structured light methods: A review. Optics and Lasers in Engineering. 106, 119-131.
DOI: https://doi.org/10.1016/j.optlaseng.2018.02.017
[7] Xie, Z., Lang, Y., Chen, L., 2021. Geometric Modeling of Rosa roxburghii Fruit Based on Three-Dimensional Point Cloud Reconstruction. Journal of Food Quality.
DOI: https://doi.org/10.1155/2021/9990499
[8] Zabatani, A., Surazhsky, V., Sperling, E., et al., 2019. Intel® realsense™ sr300 coded light depth camera. IEEE transactions on pattern analysis and machine intelligence. 42(10), 2333-2345.
DOI: https://doi.org/10.1109/TPAMI.2019.2915841
[9] Fu, Y., Yan, Q., Yang, L., et al., 2018. Texture mapping for 3d reconstruction with rgb-d sensor. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4645-4653.
DOI: https://doi.org/10.1109/CVPR.2018.00488
[10] Yamamoto, S., Karkee, M., Kobayashi, Y., et al., 2018. 3D reconstruction of apple fruits using consumer-grade RGB-depth sensor. Engineering in Agriculture, Environment and Food. 11(4), 159-168.
DOI: https://doi.org/10.1016/J.EAEF.2018.02.005
[11] O'Riordan, A., Newe, T., Dooly, G., et al., 2018. Stereo vision sensing: Review of existing systems. In 2018 12th International Conference on Sensing Technology (ICST). pp. 178-184. IEEE.
DOI: https://doi.org/10.1109/ICSensT.2018.8603605
[12] Feldmann, M.J., Tabb, A., 2021. Cost-effective, high-throughput phenotyping system for 3D reconstruction of fruit form. The Plant Phenome Journal.
DOI: https://doi.org/10.1002/ppj2.20029
[13] Reljić, I., Dunđer, I., Seljan, S., 2019. Photogrammetric 3D scanning of physical objects: tools and workflow. TEM Journal. 8(2), 383.
DOI: https://doi.org/10.18421/TEM82-09
[14] Dostal, C., Yamafune, K., 2018. Photogrammetric texture mapping: A method for increasing the Fidelity of 3D models of cultural heritage materials. Journal of Archaeological Science: Reports. 18, 430-436.
DOI: https://doi.org/10.1016/j.jasrep.2018.01.024
[15] Li, J., Berglund, J., Auris, F., et al., 2018. Evaluation of Photogrammetry for Use in Industrial Production Systems. In 2018 IEEE 14th International Conference on Automation Science and Engineering (CASE). pp. 414-420. IEEE.
DOI: https://doi.org/10.1109/COASE.2018.8560496
[16] Kochi, N., Tanabata, T., Hayashi, A., et al., 2018. A 3D shape-measuring system for assessing strawberry fruits. International Journal of Automation Technology. 12(3), 395-404.
DOI: https://doi.org/10.20965/ijat.2018.p0395
[17] Oinosho, T., Kameyama, M., Taguchi, A., 2021. Color Conversion Formulae between RGB Color Space and HSI Color Space for Color Image Processing. In 2021 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS). pp. 1-2. IEEE.
DOI: https://doi.org/10.1109/ISPACS51563.2021.9651118