Feature Extraction for Quality Modeling of Malang Oranges on an Automatic Fruit Sorting System

Karisma Trinanda Putra, Tony Khristanto Hariadi, Slamet Riyadi, Anna Nur Nazilah Chamim
Department of Electrical Engineering
Universitas Muhammadiyah Yogyakarta
Indonesia
e-mail: karisma@ft.umy.ac.id, tonykhariadi@umy.ac.id, riyadi@umy.ac.id, anna_nnc@umy.ac.id

***Abstract*—This paper evaluated some of the machine vision techniques to extract fruits characteristic, which in this case is Malang oranges. Appropriate algorithms are developed and implemented to extract features from local fruit based on Indonesian National Standard (SNI 3165: 2009) which is concerning to oranges. The research is done by analyzing fruit images, extracting HSV parameters and extracting feature using contour detection, hull convex and RGB histogram. A sensing machine which is consists of a photo box with a camera and a conveyor has been developed. The detection process can be done in real-time with the help of boxes equipped with adequate lighting. Convex hull analysis can be used to determine the diameter that has a great effect on the citrus fruit classification. While the red-green ratio can be used to label citrus fruits so that it can be used on a gradation-based fruit sorting machine. The performance was evaluated in terms of measurement accuracy which is above 88%. The research has the potential to be improved with the addition of an artificial intelligence-based decision system.**

Keywords-contour detection; hull convex; RGB histogram; red-green ratio

I. Introduction

Tangerine or mandarin orange (Citrus reticulate) is one type of orange that has long been known and cultivated by people in Asia, especially in Indonesia. Besides being a commodity trading in the country, the fruit is also traded in the international market. Indonesia is one of the producers of orange which has big enough potency to fulfill consumer demand inside and outside the country. The center of tangerine production is now mostly found in East Java especially in Malang area. Given its natural resources, this area has the advantage to produce export-quality oranges. To improve the quality in order to compete in the domestic and international market, it is necessary to have a quality standard that can be applied by Indonesian farmers and acceptable by the international market. Indonesian National Standard (SNI 3165: 2009), concerning Tangerines is formulated by the Technical Committee 65-03 of the Ministry of Agriculture Republic of Indonesia as an effort to produce citrus which has the qualities according to market demand [1].

Currently, Indonesia is the second largest citrus importer country in ASEAN after Malaysia especially sweet oranges, with an average import volume per year reaching 25,408 tons or equivalent to US $ 17,464,186 per year. As for the type of tangerine, it reached about 100,813 tons per year with a value reaching US $ 80,569,300 [2]. Not surprisingly, many imported oranges are found in almost all supermarkets including small traders. What is very sad is that imported oranges are laid out and better placed than national oranges. It seems that consumers are also less fond of their own oranges products because the quality is still inferior to imported oranges. The tendency to increase the import of citrus varieties indicates the existence of certain market segments (consumers) who want the type and quality of prime fruit that cannot be fulfilled by domestic producers. To supply these needs, import policy is opened, mostly from Australia, China and Pakistan. Actually, the condition of the oranges is not fresher than local oranges because it has been stored in cold storage for 6 months - 1 year.

Increased import of oranges can actually be a market opportunity as well as national oranges development opportunities along with increasing consumer preferences for good quality fruits. Opportunities to develop oranges cannot be separated from the national potential, such as the number of oranges production centers, the high diversity of citrus genetic resources, the availability of high quality national orange varieties including seeds, technology, and local market itself. It means that the national oranges have the opportunity to replace the existence of imported oranges as long as the government, citrus agribusiness and entrepreneurs have a high commitment to support the development of this national program. Therefore, the Indonesian government launched a strategic step "Sustainable National Oranges Development Program" which is systematically arranged, comprehensively and involves several institutions including policy makers.

To support the national oranges development movement, there needs to be technology development in line with the increase of production capability. Post-harvest processing and packaging technology is needed to accelerate and synchronize processes that are currently still done conventionally with human assistance. The quality of oranges can be seen from the diameter, shape and also the color of the fruit skin [3]. The camera is used as a visual sensing to detect the characteristics of fruits. By equating the patterns that have been compiled in SNI 3165: 2009, a sensing system was created to model the fruit quality sorter

978-1-5386-6981-5/18/$31.00 ©2018 IEEE

system and arrange them according to their color. This technology contributes to the development of fruit quality sorting machines as well as fruit preparations based on the color gradations of the fruit [4]-[8]. Fruit sorting system according to the color gradation can be applied to a fruit display on a supermarket. One of the characteristics of local fruit is its wide color variations. Preparation of oranges in columns according to the color gradation, will be a visual attraction for consumers to choose the fruit. In fact, in terms of taste, the quality of local oranges is not inferior to imported oranges.

II. Related Work

Research that has been done previously discussed a lot about the classification of mandarin oranges. With modern computer vision techniques, manual sorting of fruits is being replaced with automated low cost and consistent approach with accuracy around 90% [9]. Other studies indicate that RGB parameters can be used to perform characterization of maturity of oranges [10]. RGB parameter has a weakness in the process of recording an image which is depend on the lighting and the quality of the camera. For this research, we used a box equipped with artificial lighting to help the camera capture a more constant image. Other research that has been done to classify the Thompson citrus fruit, demonstrates the power of computer vision techniques, which can extract about fourteen parameters from a single fruit [11].

For local fruit varieties, the research conducted is still minimal. Local oranges especially Malang oranges have color characteristics that cannot be used as a reference to determine fruit maturity, in contrast to imported oranges [12], [13]. Local orange color that has been ripe varies ranging from greenish to yellowish. The diameter and defects on the fruit are the only reference to determine the quality of local oranges. Color variations in local oranges can be an advantage for the commodity, so it is necessary to examine the process of orange selection based on size and color automatically. It is intended that the fruit can be arranged in a rack in accordance with the color gradation which is more representable when marketed to the costumer.

III. Method

A fruit sorting device has been developed at laboratory of Electrical Engineering, Universitas Muhammadiyah Yogyakarta. The device consists of a mechanical conveyor and a sensing device. This paper focuses on sensing technology applied to Malang oranges on a fruit sorting system. The result is data characteristic of fruit, which consist of diameter, area of object and red/green ratio based on its color. A photo box with certain lighting is used as a fruit detection area. The conveyor will pull the fruit into the photo box. A computer is used to process images from a camera inside photo box and extract data related to fruit characteristics.

A. *Hardware and Software Setup*

To produce an optimal image, we have designed a photo box to isolate the fruits from the environmental condition. The box is fitted with lighting from the top using LED array lights. The box also has an entrance and an exit door so it can be placed on the conveyor. The fruit will be pulled by the conveyor into the photo box. Furthermore, the system will analyze the condition of the fruit without stopping the conveyor. The box is designed to have a size as shown in Figure 1. By using the box, the distance of the camera with the object is about 15 cm, varies depending on the size of the fruit. The background color is designed using white to make it easier for the program to distinguish between objects against the background. A camera is placed on the upper side of the box facing the fruit to be analyzed. The camera is connected to the computer using a USB interface as shown in Figure 2. On the computer, embedded program, written in Python language, record images with a resolution of 960 x 720 pixels. The research uses Python programming language because of its comprehensive support, huge community and simplicity of its codes. Python is an open source language so researchers do not need to buy a license.

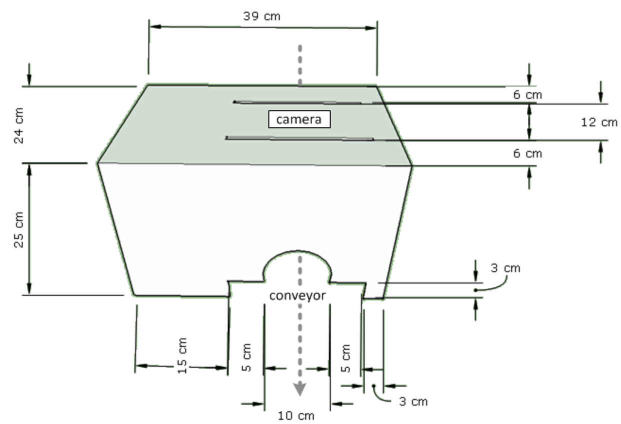


Figure 1. Photo box design to accommodate the conveyor rails and produce shooting uniformity.

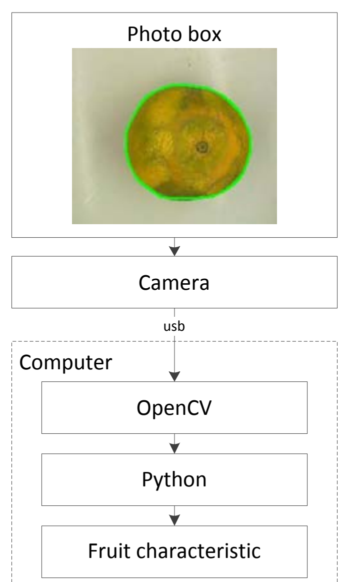


Figure 2. Block diagram of sensing device which is consist of an USB camera and a computer with OpenCV-Python installed.

Python also has support with other libraries, such as OpenCV. OpenCV is used as a library that will enclose image analysis program with camera hardware used. OpenCV (Open Source Computer Vision Library) is a software library that aims to process images in real-time or not real-time. In addition, pyQt is used to generate a GUI that can be operated in a windows environment. PyQt is an API that acts as a binding between Python and Qt cross-platform framework. For future projects, research development can be done by integrating artificial intelligent libraries such as Tensorflow for more accurate fruit classification system.

B. *Object Detection*

Object detection is the process of distinguishing objects to be observed against a background. Object detection is considered successful if the object can be separated from the background color so that the characteristics of the object can be observed more accurately. With the condition of the photo box which is white, then the oranges can be easily detected. The method used is with HSV thresholding. The method read the image color gradations in HSV format based on a predefined threshold value. HSV format has advantages in its sensitivity to distance changing, so it is suitable to be applied in object detection system. After the object can be separated from the image, then the shape of the object can be estimated.

In the first step, the process of color image analysis will be done. The camera captures the image of the object along with the background of the conveyor used. Next, the user will manually mark the area of the object. The marked area will be converted to HSV image. By using HSV histogram, it can be determined the upper and lower limits of the image. The process is only done once to get the HSV threshold value. Furthermore, the process of detecting the object can be done repeatedly using the threshold value that have been obtained, as shown in Figure 3.

Object detection is process of separating objects which is in this case, oranges with its background. The purpose of this process is to produce an accurate recognition to the object, not affected by the background color. In the process of detecting the object, the image captured, then converted into HSV format. After that, HSV thresholding is used to convert the image into binary images using threshold values which is obtained from the previous process. Furthermore, noise removing is done on the binary image, using morphological transformation, which is erosion and dilation. The basic idea of erosion is like soil erosion. It erodes away the boundaries of foreground object. Meanwhile, dilatation is an opposite of erosion. It increases the size of foreground object. Generally, in noise removal, erosion is followed by dilation, because erosion removes white noises, but it also shrinks the object. The object appear as a white image surrounded by black pixels as the background. Finally, a contour detection function is applied. Contours can be explained as a curve joining all the boundary, having the same color and intensity. The contours are useful tool for object detection and recognition. Contours that have been formed can be used as a reference to know the area of the object and its location on the image, captured by the camera.

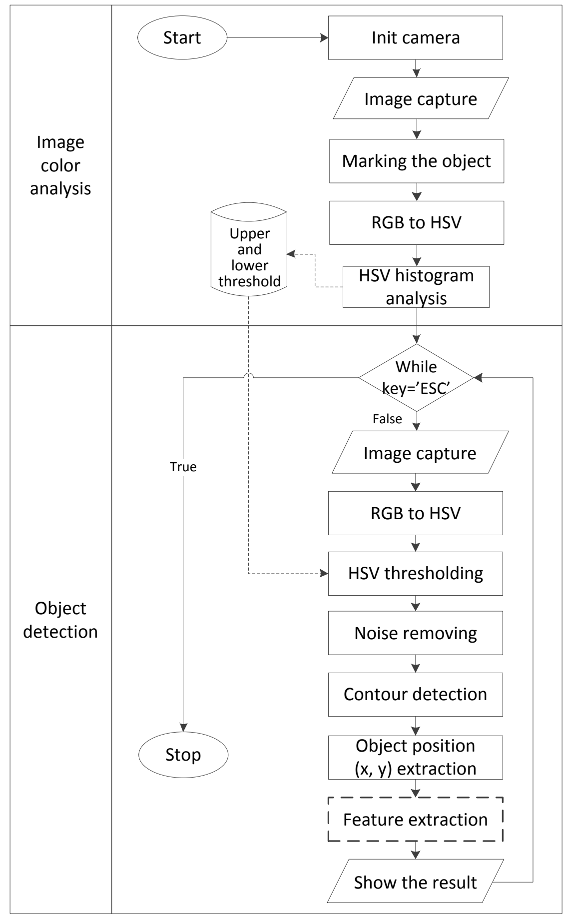


Figure 3. Flowchart of object detection using HSV thresholding and contour detection.

C. *Feature Extraction*

After the object is successfully detected, the feature extraction process can be done, shown in Figure 4. At first, convex hull is applied to the extracted contours. A polygon is considered a convex, if it is drawn a line connecting between points, then no line cuts the line that becomes the outer boundary of the polygon. From the convex polygon, then, min-max diameter and area of the polygon can be obtained. After that, the area outside the polygon will be removed, while the area inside the polygon is replaced with the original image, which is in RGB format. Furthermore, RGB histogram analysis can be done on the object without including background colors. The extractable feature is the amplitude of each color component that is red, green and blue based on its intensity range. Based on Figure 7, the most dominant color on the Malang orange fruit is red and green. The research uses red-green ratio to generate label on each fruit so that the fruit can be qualified based on its size and arranged based on its color. From Equation (1), r/g ratio can be calculated, using red and green color intensity.

$$r/g = \frac{\sum_{i=0}^{255} i \times r_i}{\sum_{i=0}^{255} i \times g_i} \tag{1}$$

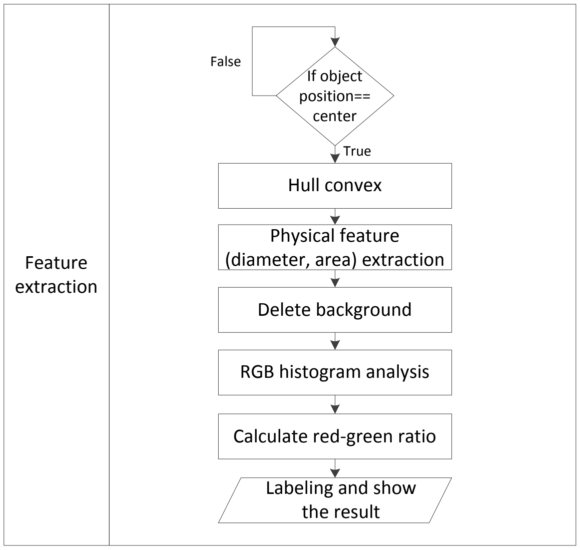


Figure 4. Flowchart of extraction process to extract physical feature and red-green ratio.

IV. RESULT

Hull convex can be applied effectively on object selection process especially Malang oranges. Polygons can follow the border of the fruit with the background. Thus, the area outside the polygon can be removed so that the histogram analysis process can be done accurately, focus on the color of the object only. From the polygon that was successfully detected, it can extract the min-max diameter and the area inside the object.

Testing is done by using 7 spherical objects with different sizes. Each object is tested 5 times. Then the average data is calculated based on the measurement of the object diameter. The results are presented in graphical form in Figure 5. In the graph, the highest accuracy value is obtained using an object with a diameter of 52 mm x 52 mm. It because calibration is only done using objects with those sizes. The smallest accuracy value obtained is 88%. Accuracy is still good with an average value in the range of 90%. The measurement accuracy can be improved by calibrating of all samples. By using regression analysis, the measurement error due to non-linearity problems resulting from the camera, can be minimized.

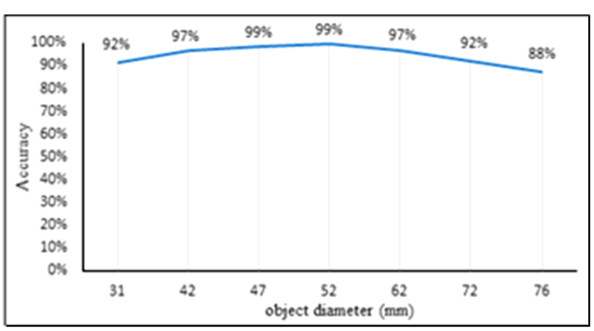
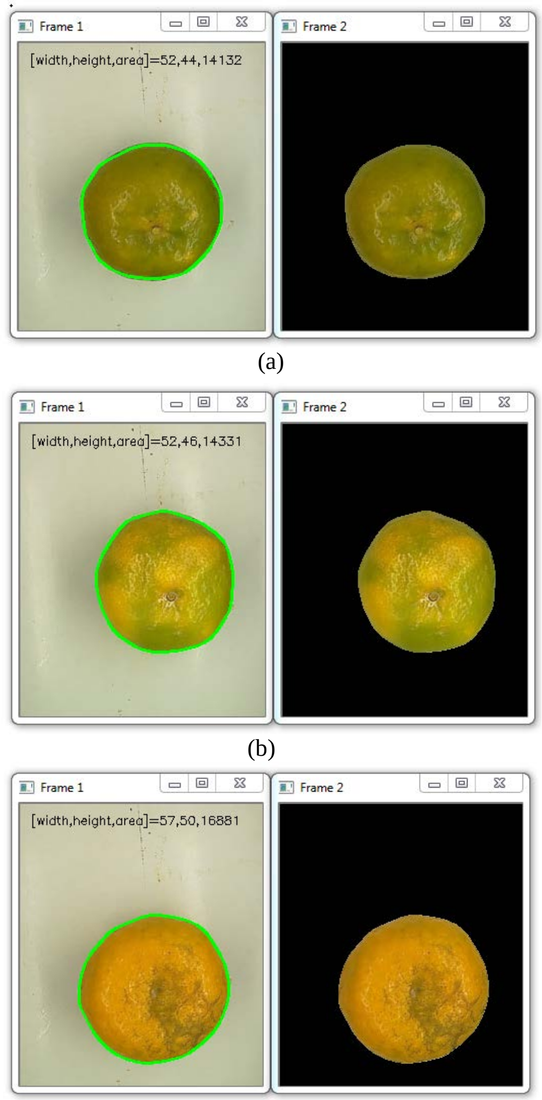


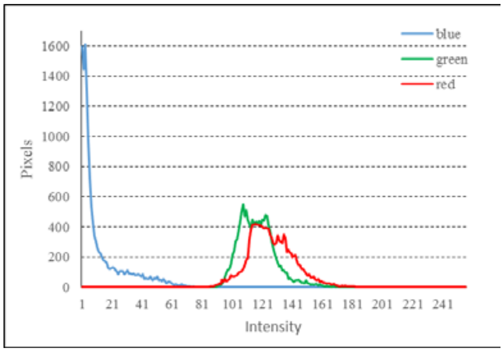
Figure 5. Accuracy of fruit measurement process using 7 spherical objects with different sizes.



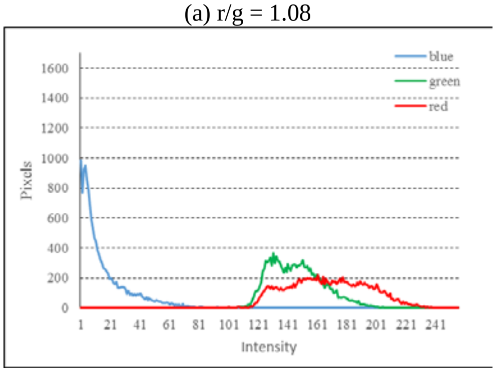
(c)

Figure 6. The result of the object selection generates min-max diameter and the area inside the hull polygon of an orange with (a) low ripeness, (b) medium ripeness and (c) high ripeness.

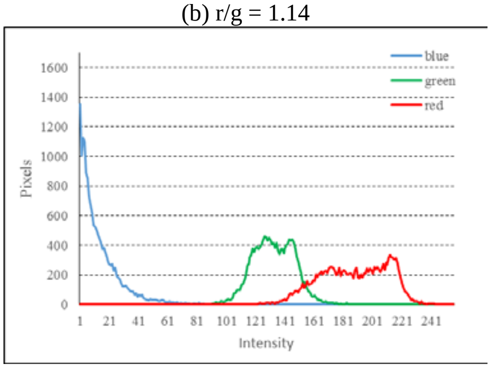
Malang oranges has very large color variations, as shown in Figure 6. Fruit gradation cannot be used as a reference to recognize the fruit maturity level. According to existing standards, the size and defects in the fruit become parameters to determine the quality of the fruit. Based on the fruit diameter, the classification can be done. The harvest will be categorized as class A for fruit with a diameter more than 4 cm and smooth. While the 3-4 cm diameter fruit goes to class B and less than 3 cm for class C. Whereas, from the histogram data, it can be used as a reference for fruits labeling, so that, the fruit can be arranged according to the color gradation. This is done by calculating the red-green ratio from histogram data. From Figure 7, the red-green ratio of the three oranges, generated different values. Visually, the low red-green fruit ratio has a greener color, while the high red-green ratio has a more yellow color.



(a) r/g = 1.08



(b) r/g = 1.14



(c) r/g = 1.41

Figure 7. Histogram analysis of an orange with (a) low, (b) medium and (c) high red-green ratio.

V. SUMMARY

Feature extraction is a process that utilizes image processing to detect the specific characteristics contained in an object inside an image. A lot of variable which can affect the image quality such as light and distance. Therefore, all measurements are made with a camera that has the same distances and fixed illumination. To get a fixed lighting, then the measurement is done utilizing LED lighting with color temperature of 6500 Kelvin. The measurement result using convex hull method has a high accuracy of more than 88%. From the experiment, the highest measurement accuracy was obtained on the object with a diameter of 52 mm because the calibration was performed at that point. The area of the object can be detected effectively because polygons can follow the shapes of the edge of the fruit well. Malang oranges classification system can be done only based on its diameter because the color does not affect the maturity and fruit grade. In addition, it was found that the color of Malang orange affects the dominant red and green components of the RGB histogram. Red-green ratio can be used as a reference to label the fruit, so it can be used to arrange the fruit based on the color gradation.

ACKNOWLEDGMENT

Support was given by Universitas Muhammadiyah Yogyakarta, who funded the work in all its stages. We would like to thank Microcontroller and Robotic Lab at Department of Electrical Engineering for their support, without whose help this work would never have been possible.

REFERENCES

[1] Badan Standarisasi Nasional, ‘Jeruk Keprok’, SNI 3165:2009, 2009.

[2] Top Production of Oranges 2014-choose 'Production Crops World’ in the left margin and picklist.

[3] F. López-García, G. Andreu-García, J. Blasco, N. Aleixos, J. M. Valiente, "Automatic detection of skin defects in citrus fruits using a multivariate image analysis approach", Comput. Electron. Agric., vol. 71, no. 2, pp. 189-197, 2010.

[4] U. Ahmad, S. Mardison, R Tjahjohutomo, A. Nurhasanah, "Development of automatic grading machine prototype for citrus using image processing", Australian Journal of Agriculture Engineeringl, no. 5, pp. 165-169, 2010.

[5] U. Ahmad, "The use of color distribution analysis for ripeness prediction of Golden Apollo melon", J. Appl. Hortic., vol. 19, Nov. 2017.

[6] Sirisathitkul Yaowarat, Naphasorn Thumpen, Weerayut Puangtong, "Automated Chokun orange maturity soritng by colour grading", alailak Journal of Science and Technology (WJST), vol. 3, no. 2, pp. 195-205, 2011.

[7] C. Damarjati, S. Riyadi, W. I. Triyani, L. M. Azizah and T. K. Hariadi, "Statistical features extraction of discrete curvelet transform for surface quality evaluation of mangosteen," *2017 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, Penang, 2017, pp. 236-241.

[8] L. M. Azizah, S. F. Umayah, S. Riyadi, C. Damarjati and N. A. Utama, "Deep learning implementation using convolutional neural network in mangosteen surface defect detection," 2017 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), Penang, 2017, pp. 242-246.

[9] C. P. Diana Carolina and N. T. Deivis David, "Classification of oranges by maturity, using image processing techniques," 2014 III International Congress of Engineering Mechatronics and Automation (CIIMA), Cartagena, 2014, pp. 1-5.

[10] A. Wajid, N. K. Singh, P. Junjun and M. A. Mughal, "Recognition of ripe, unripe and scaled condition of orange citrus based on decision tree classification," 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, 2018.

[11] S. Sabzi, Y. Abbaspour-Gilandeh and J. I. Arribas, "Non-intrusive image processing Thompson orange grading methods," 2017 56th FITCE Congress, Madrid, 2017, pp. 35-39.

[12] B. Kanimozhi, R. Malliga, Classification of Ripe or Unripe Orange Fruits Using the Color Coding Technique, vol. 1, no. 3, pp. 43-47, 2017.

[13] E Kane Kevin, Suk Lee Won, "Multispectral Imaging for In-field Green Citrus Identification", American Society of Agricultural and Biological Engineers, no. 073025, 2007.