Separation of the Mature Level of Papaya Callina Fruit Automatically Based on Color (RGB) uses Digital Image Processing

Z K Simbolon¹, S A Syakry¹, Mulyadi¹, M Syahroni²
¹Information Technology Department, Politeknik Negeri Lhokseumawe Aceh-Indonesia.
²Electrical Engineering Department, Politeknik Negeri Lhokseumawe Aceh-Indonesia.

E-mail: zulfan69@gmail.com

Abstract. The process of sorting papaya fruit based on quality is one of the processes that greatly determines the quality of papaya fruit that will be sold to consumers. The process of identifying the quality of fruits using the human eye has the disadvantages of requiring more energy to sort, the level of human perception in terms of different fruit sorting, the level of human consistency in assessing fruit quality is also unstable because humans can experience fatigue. Research on fruit using image processing is the current trend, especially for fruit conditions, both qualities weight and size because this system processes faster and avoids or reduces failures that occur as a result of human nature. The process of selecting the level of fruit maturity in the process of recognition and determination and classification of post-harvest agricultural products on papaya fruit, depends on how the system is built. This study aims to build a quality recognition system for papaya fruits using Digital Image Processing technology, to analyze the level of color values (RGB), to determine the maturity level of papaya-callina fruit, so that later can be used as a reference in determining the maturity level of papaya fruit. First, the image of papaya is taken, or the acquisition uses a camera to be used as a database based on the condition of its maturity level. Second, the separation of the fruit image with the background based on the pixels, calculating the pixel value looking for the mean value, min, the max that is used later in the reference in determining the fruit maturity condition: young ripe, the half mature, mature. The results of this study provide information about pixel data in which young ripe papaya, red value does not dominate that is 7.785495, the green value becomes the highest value of 10.23922, papaya the half mature, it can be seen that the red and green composition of the pixel value is almost the same, namely 12.56288 and 12.12431, while the fully mature condition of the papaya, average red pixel value becomes more dominant when compared to green, which is 24.111901 for red and 13.70812 for green.

1. Introduction
The process of identifying fruits and vegetables that are done manually is inefficient and inaccurate for a large enough amount. Today digital image processing techniques are growing rapidly with applications that are quite wide in various fields. One field of application is an image-based system used for identification. The development of information technology allows the identification of fruit with the help of computers. The problem is how to recognize the fruit. The condition of a papaya is determined by several parameters, including the level of maturity parameters, which is seen from the color of the papaya. In general, the classification of papaya is done manually, using human senses. On the other hand, humans have a weakness when humans perform sensory tasks in large numbers and long working hours such as when using the human sense of sight as a sensor to determine the level of maturity based on the colors contained in the fruit and also have some disadvantages, among others: human judgment subjective to the level of papaya maturity. Human judgment can differ from one assessor to another. For
this reason, the most important grouping of fruit quality levels is the basis for the value of the content of red, green, blue (RGB) as a reference for determining the maturity level of papaya callina fruit.

Research on fruits using image processing has become a current trend, especially for fruit conditions both quality and weight and size because this system processes faster and avoids or reduces failures that occur as a result of human nature. Some research has been done about papaya fruit using image processing, research on papaya was carried out by Eliyani with the title recognition of the maturity level of rabo papaya fruit using color-based (RGB) image processing with k-means clustering, this research was carried out by taking a sample of 30 pieces of papaya fruit [1]. Research conducted by Riyadi, about papaya under the title Papaya fruit grading based on size using image analyst using thresholding method and morphology as its main reference [2]. Ahmad conducted a study of oranges to bleach pontianak fruit based on size and color using image processing in 2009 [3]. Some important visual quality parameters of citrus fruit were studied, and five quality groups of citrus fruits obtained from the Kramatjati Market wholesaler were used as samples. Citrus fruit images were recorded, and the recorded images were analyzed by the image processing program using area projections to classify citrus fruits into quality groups A, B, C, D, and E, the conclusion of the study resulted in the whitening of 850 samples of citrus fruit by size. Using the built-in image processing program, color parameters can also be used as additional quality parameters in pontianak orange fruit bleaching to produce a quality class with a uniform color display, in addition to a uniform size, using a built-in image processing program showing compatibility with 95.9 %, while the results of manual bleaching show a compatibility of only 41.3% [6], where for other fruits such as strawberries Prasetyani and Ahmad have also been conducted with the title evaluation of strawberry fruit bleaching parameters (Fragaria chiloensis L.) [4]. Then for research that uses color as one of the objects of research in addition to the fruits that are most widely used to recognize an object both face and object and others.

Papaya fruit is picked. must at the time the fruit gives signs of maturity, namely: the color of the fruit skin begins to turn yellow. But there are still many farmers who pick it when the fruit is not too ripe [5]. Papaya must be harvested at the right time according to the level of maturity so that the harvested fruit will ripen normally and produce fruit with good aroma and taste. Skin discoloration is usually used by farmers and traders. The level of maturity is determined by the degree of yellow seen and harvesting done depending on the market objectives. The fruits that will be sent to a distant market are usually harvested when the color of the new fruit skin is slightly yellow. At this color level, fruit can last longer (not quickly decay). For local market purposes, the fruit is harvested at a higher level of maturity, when three-fourths of the fruit peel is yellow, and the color of the fruit stalk begins to turn yellow. Fruits like this must be marketed quickly because the fruit will not last long with a very short life span of only 3-4 days. A systematic method in describing / describing the color signs for papaya is explained by Lam (1987), where the fruit color index is given a value according to the fruit maturity level as shown in Figure 1. Fruit for a distant market must be harvested at 2 & 3 color index the fruit that is in the color index 1 cannot be consumed directly because the fruit is still very green and if picked the fruit will not ripen normally. The fruit harvested in the 4 & 5 color index is only suitable for the local market [6].

1.1 Image Processing

Image is an intensity function in a two-dimensional plane. Because the intensity in question originates from a light source, and light is a form of energy, the state in which the intensity function is located between0<f(x,y)<∞. Basically, the image that is seen consists of files of light reflected by objects around it. So scientifically, the function of light intensity is a function of light sources that illuminate objects, as well as the amount of light reflected by objects [7], or written:
\[ f(x,y) = i(x,y) \cdot r(x,y) \]

that is:

\[ 0 < i(x,y) < \infty \quad \text{(Illumination of light sources)} \]

\[ 0 < r(x,y) < 1 \quad \text{(The object's reflection coefficient)} \]

The image can be represented numerically, so the image must be digitized, both to the coordinate space \((x,y)\) and to the gray scale \(f(x,y)\). The digitization process of coordinates \((x, y)\) is known as "image sampling", while the digitization process of gray scale \(f(x,y)\) is referred to as "gray degree quantization". A continuous image \(f(x, y)\) will be approximated by samples that are uniform in distance and form the \(NxN\) matrix. The value of the matrix elements states the degree of gray image, while the position of the element (in rows and columns) represents the coordinates of the points \((x, y)\) of the image [9].

\[
f(x, y) = \begin{bmatrix}
    f(0,0) & f(0,1) & \cdots & f(0,N-1) \\
    \vdots & \ddots & \vdots & \vdots \\
    f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1)
\end{bmatrix}
\]  

(2)

1.2 RGB image

An image usually refers to an RGB image. Actually, how images are stored and manipulated on a computer derived from television technology, which first applies it to computer graphics display. When viewed with a magnifying glass, the computer monitor display will consist of a number of triplet dots in red (R), green (G), and blue (B). The image in the computer is no more than a set of triplets where each triplet consists of brightness variations of the red, green and blue elements. Its representation in the image, the triplet will consist of 3 numbers that regulate the intensity of the red (R), green (G), and blue (B) of a triplet. Each triplet will represent 1 pixel. These RGB numbers are often referred to as color values. In the .bmp format, each pixel image is represented by 24 bits, 8 bits for R, 8 bits for G, and 8 bits for B, with the settings shown in Figure 2.

![Figure 2. Color triplet](image)

1.3 Normalization

Normalization process is needed to uniformize a size, in image processing one of them is used to uniform image size (height and width of the image). This is to meet the standards or uniformity of the data to be analyzed and to be examined.

1.4 Image segmentation

Image segmentation here is to separate the image between the R value, G value, and B value so that the color values can be obtained on an object that is observed later.

1.5 Mean, min, and max

Mean is the mean of calculating the size of the central tendency which provides a general description of the data. The average count is obtained by summing all data values with the amount of data. The average population count is the average value of the population. Formula: the sum of all population values / amount of data (observation).

\[
\mu = \frac{\sum x}{N} \quad (3)
\]

\(\mu = \text{average population count}
\]

\(\Sigma = \text{Sum}\)
x = value of data in the population
N = number of data
Σx = the total number of data values in the population

Min shows the smallest data of the variable while the max shows the greatest value of the variable

2. Research methods
The research method for the process of obtaining a reference database is carried out as shown by the fishbone diagram in Figure 3. The number of images of papaya fruit as many as 300 pieces, where each papaya fruit in repeated photos to replace if there is an image that is not in accordance with the image standards for research. The process of taking pictures uses the same light, position and distance. From the fishbone chart in Figure 3, the research stages are as follows: 1) Image data retrieval of papaya fruit, 2) Image normalization, 3) Segmentation (Separation of foreground with background), 4) RGB image segmentation, 5) pixel value retrieval, 6) Searching the mean value of each RGB channel located in the fruit image.

3. Results and discussion
3.1 Photo capture of papaya fruit image data for the database
The first process is capturing is done with the help of equipment with the camera position for distance, the shooting light and zoom used are 2x. With the distance from the location of the digital photo to the floor 74 cm, then the image retrieval obtained here is 300 images where 100 raw images are ripe, 100 images mature, 100 images are fully ripe. After the capture process, the creation of the data name is adjusted to the condition of the fruit image and the data type of JPG. This is done to facilitate the recognition of the maturity level so that the RGB value obtained can be mapped. To produce the average value of each fruit image on each component value of Red Green and Blue.

3.2 Image Normalization.
The initial image used has a size of 4608x3456x3 uint8 then normalized to make the computation process size to 704x528x3 uint8 shown in Figure 4 and Figure 5, this avoids the size of the image that is not the same, the normalization process does not reduce the color value information contained in the image.

3.3 Image segmentation
After the normalization process then color segmentation. This is done by separating RGB values directly, this can be done if the black background image (0) is shown in Figure 6, so that the RGB value of the object can be obtained, then the value contained for the background must be zero this is intended to not affect the search the average value of each RGB component. Then segment to separate values R, G, B into their respective channels as shown in Figure 7, 8, 9.

Figure 4. Initial image with actual size

Figure 5. Image After Changing Size

Figure 6. Image After Background Pixel Value Is Changed To 0

Figure 7. Segmentation of Red Pixel Value

Figure 8. Segmentation of Green Pixel Value

Figure 9. Segmentation of Blue Pixel Value
As shown in the figure values at the same coordinates, namely X = 1438 and Y = 791, the values of R = 0.196, G = 0.302 and B = 0.0941 are obtained.

3.4 Taking pixel values and finding the mean value, min, the max of each RGB.

The mean value, min, max is sought for later reference to the process of introducing maturity levels. The results of fruit images with the condition of raw ripe fruit were obtained and shown in Table 1.

| NamaFile  | MeanR | MeanG | MeanB |
|-----------|-------|-------|-------|
| 01-kates-01 | 4,7658 | 6,4608 | 3,2126 |
| 01-kates-02 | 4,2452 | 6,0076 | 2,8937 |
| 01-kates-03 | 4,188  | 5,8476 | 2,9287 |
| 01-kates-100 | 7,3682 | 9,5283 | 3,8239 |
| Average     | 7,7855 | 10,2392| 3,8628 |
| Min         | 3,7809 | 5,3947 | 1,9336 |
| Max         | 13,4588| 16,0290| 5,6631 |

From the data in Table 1 and Figure 10, the average values for each component of color data are obtained from the image of raw ripe fruit (100 images) with a red average value of 7.7855, a green average of 10.2392 and a blue average of 3.8628, for the minimum / maximum average value obtained for the average red is 3.7809 / 13.4588, the minimum / maximum average of green is5,3947 / 16,0290, the minimum / maximum blue average is 1,9336 / 5,6631. Under mature conditions, the value of fruit with the condition of raw ripe fruit is obtained and shown in Table 2 and in Figure 11. From the data in Table 2 above, the average value for each component of color data from the image of ripe fruit (100 images) with a red average value of 12.5629, a green average of 12.1243 and a blue average of 3.1737. For the minimum / maximum average value obtained for the average red is 7.7971 / 23.4206, the minimum / maximum average of green is 9.7780 / 17.8270, the minimum / maximum average for blue is 2.1239 / 4.4239, while the median for the red average is 11.5788, for the green average is 11.8724, for the blue average is 3.1446. In full ripe conditions the value of fruit images with the condition of raw ripe fruit is obtained and shown in Table 3 and in Figure 12.

| NamaFile  | MeanR | MeanG | MeanB |
|-----------|-------|-------|-------|
| 02-kates-01 | 10,601 | 12,515 | 3,2876 |
| 02-kates-02 | 11,407 | 12,543 | 2,9616 |
| 02-kates-03 | 14,999 | 14,961 | 2,7982 |
| 02-kates-100 | 11,409 | 11,882 | 4,2193 |
| Rata-rata   | 12,5629| 12,1243| 3,1737 |
| Min         | 7,7971 | 9,7780 | 2,1239 |
It can be seen that the composition of red and green means that the average score is evenly equal, namely 12.56288 and 12.12431, then in full ripe conditions the average red value becomes more dominant when compared to green, ie 24.111901 for red and 13.70812 for green. In the graph Figure 13 shows a meeting point on the average value of the condition of cooked papaya when using the eye as a sensor to determine the maturity level of the fruit. There is a change in the average pixel value at full maturity in blue. The value obtained is feasible to be used as a reference in determining the maturity level of Papaya callina fruit.

### 4. Conclusion

The results of the study can be concluded as follows: The color value of the data obtained is in accordance with the actual conditions when using the eye as a sensor to determine the level of fruit maturity. There is a change in the average pixel value at full maturity in blue. The value obtained is

Max 23,4206 17,8270 4,4239

| NamaFile         | MeanR | MeanG | MeanB |
|------------------|-------|-------|-------|
| 03-kates-01      | 31,244| 15,63 | 3,0121|
| 03-kates-02      | 31,083| 15,65 | 2,9564|
| 03-kates-03      | 32,622| 14,733| 3,7738|
| 03-kates-100     | 20,486| 11,915| 3,5601|
| Rata-rata        | 24,1190| 13,7081| 2,8550|
| Max              | 37,2177| 18,4233| 4,9775|

Table 4. Average value of RGB components on each maturity level condition

| Matang | Rata-rata | Min   | Max       |
|--------|-----------|-------|-----------|
| Mentah | 7,785495  | 10,23922 | 3,862826  |
|        | 3,7809    | 5,3947  | 1,9336    |
|        | 13,4588   | 16,029  | 5,6631    |
| Mentah | 12,56288  | 12,12431| 3,17368   |
|        | 7,7971    | 9,778   | 2,1239    |
|        | 23,4206   | 17,827  | 4,4239    |
| Mentah | 24,11901  | 13,70812| 2,854977  |
| penuh   | 15,6722   | 10,3999 | 1,3772    |
|        | 37,2177   | 18,4233 | 4,9775    |
feasible to be used as a reference for determining the maturity level of Papaya callina fruit (California kates / honey kates).

Reference

[1] Eliyani, Tulus and Fahmi 2013 The recognition of the maturity level of the Paya Rabo papaya fruit uses image processing based on RGB color with K-Means Clustering (Singuda ensikom journal. North Sumatera University. Medan. Indonesia.

[2] S. Riyadi, M.M. Mustafa, A. Hussain, A. Hamzah, Proceeding of WSEAS International Conference on Circuits, Systems, Electronics, Control & Signal Processing, Cairo, Egypt, 2007, p.348

[3] U. Ahmad, S. Mardison, A. Nurhasanah, S.B. Sulistyo, Proceedings of the National Seminar Association Indonesian Agricultural Informatics, 2009.

[4] E. Prasetyani, Bachelor's Thesis, Department of Agricultural Engineering, Bogor Agricultural University, Indonesia, 2008.

[5] M. Khojastehnazhand, M. Omid, A. Tabatabaeefar, J. Plant Sci. 4/4 (2010) 122.

[6] S. Ahmad, Harvesting and Post Harvesting Papaya, http://kliksatori.blogspot.com/2011/09/penanganan-panen-dan-pasca-panen-pepaya.html, 2013.

[7] M.C.H. Wijaya, A. Prijono, Digital Image Processing Using Matlab. Informatika Publishing. Bandung. Indonesia. 2017. P.78.