The implementation of k-means algorithm to determine the quality of teak wood in image based on the texture

Erlan Darmawan¹, Panji Novantara², Gentur Priguna Suwarto³, Rio Andriyat⁴, Yati Nurhayati⁵

Universitas Kuningan, Kuningan, Indonesia

*erlan.darmawan@uniku.ac.id

Abstract. Teak is a type of tree that produces high quality wood. Large tree, straight trunk, can grow to a height of 30-40 m. In the developing teak parquet industry, the process of selecting raw materials is very important. This is due to the many types, types and classes of wood texture. The choice of wood texture is still widely used by human workers, most of which require a long time and the results obtained are still inaccurate. This study aims to solve the problem of quality teak grouping based on image texture by creating a computerized system or application using the K-means algorithm. With this method, it is hoped that the classification of the quality of teak can be described more easily based on the texture of the image. Based on the results of system design and testing, the application can group the quality of teak wood based on the texture and color of the image

1. Introduction

Teak is a kind of high-quality wood production. Large tree, straight trunk, can grow to a height of 30-40 m. Teak grows in deciduous forests, which shed their leaves in the dry season. Teak or Tectona Grandis is first class wood because of its strength, durability and beauty. Although hard and strong, teak wood is easy to cut and work, so it is preferred for making furniture and carvings. Several products that have the characteristics of wood texture that are commonly used in indoor applications, such as furniture and floors [1] [2] [3] [4], Indonesian teak wood has several advantages that other types of wood may not have either. In terms of strength or texture. [5] researching image processing to find out the strength of the wood under study is easier, cheaper and certainly does not damage the wood being tested.

In the teak parquet industry that is currently developing, the process of selecting raw materials is very important. This is because there are many types, types and classes in the wood texture. The choice of wood texture is still widely used by human labor, most of which take a long time and the results obtained are still not accurate enough. In addition, the process of selecting raw teak wood is carried out by an expert (grader) conventionally, this causes dependence on the expertise and experience of a grader in selecting teak raw materials. This dependence causes the problem to become more complex when the amount of teak that will be tested visually is increasing. To overcome this problem, a wood texture imaging system is used to classify the texture of the parquet.

Texture is a function of the spatial variation of pixel intensity in an image. Based on the structure texture can be divided into macrostructure and microstructure textures. Macro structural textures have
repeating patterns periodically in an area, usually in man-made patterns and tend to be easy to represent mathematically. Meanwhile, the microstructure texture has unclear looping, so it is not easy to provide a comprehensive texture definition.

The difficulty in the process of grouping image textures that usually occurs is determining the texture boundaries, because the boundaries between textures are often unclear. So the segmentation process is needed to properly separate or group objects so that the next results get good results. In classifying the quality of teak based on texture in this image using the K-means method.

2. Methodology
In this research, the problem solving method applied is the k-means algorithm. With the hope that the system can classify teak color clusters as shown in Figure 1.

![Figure 1. Flowchart of teak classification system](image)

3. Result and Discussion
The system that is designed aims to provide convenience in grouping the quality of teak wood based on the texture which in the program that is made will further accelerate the process of determining the quality of wood quickly, accurately and objectively. Figure 2 shows the system to be built.

![Figure 2. Rich picture of the system under study](image)

The process of calculating the K-means algorithm begins by determining the Teak Color Dataset which will be grouped into 3 parts, namely Grade A, Grade B and Grade C.

3.1 RGB Process
a. The user captures (photographing teak wood) then the captured results are stored in the cellphone storage in the form of a JPEG image.
b. The captured image is a 2D image, which means that this image has dimensions of length and width. The size of this dimensional image is expressed in pixels. for example, the captured image has a length of 640 pixels and a width of 480 pixels, so this image can be said to have a resolution of 640x480 pixels. In other words, the image is composed of 640 columns x 480 rows.
c. Each pixel consists of 3 main color components. The main colors are R (red = red), G (green = green), and B (blue = blue).

d. RGB color notation can be written in the form of 6 hex digits. digits 1 and 2 represent red color (R), digits 3 and 4 represent green (G), then digits 5 and 6 represent blue (B). example: 225588

e. In addition to the 6 hex digit form, color pixels can be written in RGB format. RGB Missal (34, 85, 136). 34 represents R, 85 represents G, and 136 represents B. Each segment R, G, or B has a range from 0 – 255

f. For example, if we want to take RGB pixels from the image, we can do this by:
   ```java
   pixel[image.width x image.height] // eg image size 640x480
   n=0 // initialize iterator
   For (x = 0 to image.width)
     For (y = 0 to image.height)
       pixel[n] = image.get(x,y) // takes the pixel from the image in index
       n++ // increment iterator
     next y
   next x
   ```
   From the results of the RGB pixel capture process, a total of 307200 RGB pixels are obtained.

g. The bigger the image size, the more pixels will be.
h. To get the dominant color from the image using Kmeans, of course, the pixels will be read.
i. These pixels are then grouped into clusters. The number of clusters can be predefined examples in the program:
   ```java
   kmeans = new Kmeans(sizeImage 10, fileImage);
   kmeans.initClusters();
   kmeans.startKmeans();
   mClusters = kmeans.getClusters();
   ```
j. Below is a wooden image, 10 clusters are defined, then the kmeans process will produce the 10 most dominant colors from the image.

| NO | Color   | R  | G  | B  |
|----|---------|----|----|----|
| D1 | 5a3c1e  | 90 | 60 | 30 |
| D2 | 5A3C3C  | 90 | 60 | 60 |
| D3 | 785a3c  | 120| 90 | 60 |
| D4 | 96785a  | 150| 120| 90 |
| D5 | 965A3C  | 150| 90 | 60 |
| D6 | B4785A  | 180| 120| 90 |
| D7 | b4b496  | 180| 180| 150|
| D8 | D2B4B4  | 210| 180| 180|
| D9 | b49678  | 180| 150| 120|

After determining the dataset, determine the color clusters that will be formed as follows:

| Cluster   | Color   | Grade |
|-----------|---------|-------|
| Cluster (C1) | 785a3c | A     |
| Cluster (C2) | 96785a | B     |
| Cluster (C3) | b49678 | G     |
After determining the cluster then determine the Centroid as follows:

Table 3. Centroid

| Cluster | R   | G   | B   |
|---------|-----|-----|-----|
| C1      | 120 | 90  | 60  |
| C2      | 150 | 120 | 90  |
| C3      | 180 | 150 | 120 |

3.2 Cluster Center Distance Calculation

Euclidian Distance is used to measure the distance between the center, then the distance matrix is obtained, namely C1, C2 and C3 as follows:

Euclidian Distance formula:

(1)

Perform the above calculations until D9. After doing the calculations above, the following results will be obtained:

Table 4. Results of Cluster Distance

| No  | Color   | Distance to Cluster | Grade |
|-----|---------|---------------------|-------|
| D1  | 5a3c1e  | 51.96152            | 103.923 | 155.8846 | A     |
| D2  | 5A3C3C  | 42.42641            | 90     | 140.7125 | A     |
| D3  | 785a3c  | 0                   | 51.96152 | 103.923 | A     |
| D4  | 96785a  | 51.96152            | 0      | 51.96152 | A     |
| D5  | 965A3C  | 30                  | 42.42641 | 90     | A     |
| D6  | B4785A  | 73.48469            | 30     | 42.42641 | B     |
| D7  | b4b496  | 140.7125            | 90     | 42.42641 | C     |
| D8  | D2B4B4  | 174.9286            | 123.6932 | 73.48469 | C     |
| D9  | b49678  | 103.923             | 51.96152 | 0      | C     |

Then re-determine the center point of the new cluster based on the average of the new cluster obtained from the formula = value of results / number of results (Grade)

Cluster A (R) = (90+90+120+150) / 4 = 112.5
Cluster A (G) = (60+60+90+90) / 4 = 75
Cluster A (B) = (30+60+60+60) / 4 = 52.5
Cluster B (R) = (150+180) / 2 = 165
Cluster B (G) = (120+120) / 2 = 120
Cluster B (B) = (90+90) / 2 = 90
Cluster C (R) = (180+210+180) / 3 = 190
Cluster C (G) = (180+180+150) / 3 = 170
Cluster C (B) = (150+180+120) / 3 = 150

Then we will get the results of the new cluster as follows:
Table 5. New Cluster

| Cluster | R    | G    | B    |
|---------|------|------|------|
| C1      | 112.5| 75   | 52.5 |
| C2      | 165  | 120  | 90   |
| C3      | 190  | 170  | 150  |

Then do the calculations again with the formula that has been done above and will produce the following results:

Table 6. Distance to New Cluster

| No | Color  | Distance to Cluster | Grade |
|----|--------|---------------------|-------|
|    |        | C1      | C2       | C3       |       |
| D1 | 5a3c1e | 35.17812 | 113.2475 | 191.0497 | A     |
| D2 | 5A3C3C | 36.57185 | 100.6231 | 142.5096 | A     |
| D3 | 785a3c | 18.37117 | 61.84658 | 139.2839 | A     |
| D4 | 96785a | 69.55214 | 15      | 87.74964 | B     |
| D5 | 965A3C | 41.07919 | 45      | 126.8858 | A     |
| D6 | B4785A | 89.37281 | 15      | 78.74008 | B     |
| D7 | b4b496 | 158.3903 | 86.16844 | 14.14214 | C     |
| D8 | D2B4B4 | 191.8007 | 117.1537 | 37.41657 | C     |
| D9 | b49678 | 121.3981 | 45      | 26.92582 | C     |

The results of the first and second stages do not change, so the results are in accordance with the cluster grouping. Following are the results of the grouping (Grade):

Table 7. The Final Result of K-Means Calculations

| No | Color  | R    | G    | B    | Grade |
|----|--------|------|------|------|-------|
| D1 | 5a3c1e | 90   | 60   | 30   | A     |
| D2 | 5A3C3C | 90   | 60   | 60   | A     |
| D3 | 785a3c | 120  | 90   | 60   | A     |
| D4 | 96785a | 150  | 120  | 90   | B     |
| D5 | 965A3C | 150  | 90   | 60   | A     |
| D6 | B4785A | 180  | 120  | 90   | B     |
| D7 | b4b496 | 180  | 180  | 150  | C     |
| D8 | D2B4B4 | 210  | 180  | 180  | C     |
| D9 | b49678 | 180  | 150  | 120  | C     |

The following is a table describing the taking of wood parts for the specified grade:

Table 8. Teak wood Grade

| Grade | Description |
|-------|-------------|
| A     | Grade A wood is located in the core of the teak tree trunk. has the densest fiber characteristics, darker color and has a lubricating oil that can protect teak wood from termites and fungi attack. Grade A teak is only 25% of all teak trees and is the most expensive type of teak among other grades. |
This type of teak wood is located in the middle of the teak tree trunk. This teak wood has medium quality. It has physical characteristics of fibers that are not too dense and not too regular, the color is almost the same as grade A teak wood and is still able to produce lubricating oil to protect it from various termite and fungal attacks. Grade B teak is only 35% of the total teak tree trunk and is a type of teak wood that is quite expensive.

The type of teak that is found on the outer part of the teak tree trunk. Grade C teak wood has a less good quality than grade A and grade B teak wood. Its characteristics, this wood is the brightest color, namely yellowish white and irregular and less dense fibers also does not have lubricating oil so it is prone to termites and mold. This grade C teak accounts for 40% of the entire teak tree trunk and is the cheapest teak wood.

Figure 3 is the result of detection of teak wood grade by the system:

4. Conclusion
Based on the results of implementation and testing, the application system for grouping the quality of teak wood based on the texture of the image using the Android-based K-means Logic Algorithm can function properly on Android devices, the K-means algorithm can be applied to classify the quality of teak wood to be faster, more accurate and objective. Applications that run well can respond according to program needs, and the application cannot distinguish the type of wood used.

References
[1] Nakamura, M., Matsuo, M., and Nakano, T. (2010). "Determination of the change in appearance of lumber surfaces illuminated from various directions," Holzforschung 64(2), 251-257. DOI: 10.1515/HF.2010.031
[2] Cristea, M. V., Riedl, B., and Blanchet, P. (2011). "Effect of addition of nanosized UV absorbers on the physico-mechanical and thermal properties of an exterior waterborne stain for wood,” Progress in Organic Coatings 72(4), 755-762. DOI: 10.1016/j.porgcoat.2011.08.007
[3] Burnard, M. D., and Kutnar, A. (2015). "Wood and human stress in the built indoor environment: a review,” Wood Science and Technology 49(5), 969-986. DOI: 10.1007/s00226-015-0747-3
[4] Lozhechnikova, A., Bellanger, H., Michen, B., Burgert, I., and Österberg, M. (2017). "Surfactant-free carnauba wax dispersion and its use for layer-by-layer assembled protective surface coatings on wood,” Applied Surface Science 396(7), 1273-1281. DOI: 10.1016/j.apsusc.2016.11.132
[5] O. D. Sukrisdyanto, I. K. E. Purmana and S. M. S. Nugroho, "Wood Strength Classification Based on RGB Color and Image Texture Using KNN Method," 2019 International Seminar on
Intelligent Technology and Its Applications (ISITIA), Surabaya, Indonesia, 2019, pp. 360-365, doi: 10.1109/ISITIA.2019.8937239.