Selected RRIM 2000 Series Rubber Seed Clones Identification Through Statistical Analysis Using Image Manipulation

Mior Muhammad Ridzwan Bin Mior Khairuddin
Faculty of Electrical Engineering, Universiti Teknologi MARA,
40450 Shah Alam, Selangor, Malaysia
mior.m.ridzwan@gmail.com

***Abstract*— The rubber tree clones can be identified through the tree branch, the tree leaves, the rubber milk, and the rubber seeds. In this paper, the research in done to recognize the clone through the rubber seed by using image manipulation techniques based on the pixel from the image of the seed. The samples of seeds used are from the Hevea brasiliensis RRIM2000 series species. There are about 33 clones in this series. However, only five of the series is selected (RRIM2002, RRIM2015, RRIM2020, RRIM2023, and RRIM2024). Samples of this rubber seeds are captured by using a digital camera under a control light, stored and later to be processed by using the Matlab software. The data extracted from the histogram produced from the Matlab is then analyzed using the SPSS software. The statistical result from the error plot and the one-way ANOVA shows that it is hardly to differentiate the clones as the clone does not show a clear differences between each clones.**

Keywords: Keyword: Digital Image Processing, Rubber seed clones, MATLAB, SPSS.

I. INTRODUCTION

Rubber tree is an important plant in the world to supply the raw material for manufacturing. In order to increase the production from rubber tree, the seeds planted must be from the best quality of rubber tree. There are many differences of intensity of brown color and pattern on the seed's surface based on clone series. Previously the rubber seed's inspector can only look from the surface of seeds to decide the type of clones. However, these methods will consume time, percentage of accuracy and as well as cost in order to trained new worker or farmer with regards to the identification of rubber seed clones.

Previously the best rubber series clone was recognized based on shape and texture of the rubber series seeds through the image processing and wavelet technique that were introduced by the previous project student, Hajar Mohd Salleh and Erna Asmalina Bt Anua. This project is similar from the previous project, however for this research the analysis is based on the pixel projected from the image of the rubber seed surface. The factor that influenced of the pixel usually depends to the brightness of surface material which differs in grayscale image. Application from this factor was used in order to achieve the objective of this project.

II. CLONAL CHARACTERISTIC

Figure below shows the types of clones that been used as samples to make an analysis on this project.

Figure 1(a): RRIM2002

RRIM2002: The seed is medium in size, brownish with smooth and shining seed coat. It has square shape. Overall growth of this clone is considered good. This clone was recommended for both latex and timber production [1].

Figure 1(b): RRIM2015

RRIM2015: The seed size of this clone is medium. It was smooth, shining with light brownish seed coat. The shape is square to slightly ovoid. Overall growth and seed production of the clone is good. This clone was highly recommended for latex and timber production [1].

Figure 1(c): RRIM2020

RRIM2020: The seed features of this clone are most likely same with RRIM2015 but it has square to rectangular shaped of seed. Overall growth and seed production of the clone is vigorous. This clone was also highly recommended for latex and timber production [1].

Figure 1(d): RRIM2023

RRIM2023: The size of this clone is medium. It is smooth, shining with faint brownish seed coat. The shape is round to almost ovoid. Overall growth and seed production of the clone is good. This clone was also recommended for latex production [1].

Figure 1(e): RRIM2024

RRIM2024: The size of this clone is medium. It is smooth, shining with faint brownish seed coat. Overall growth of this clone is considered good. This clone is recommended for both latex and timber production [1].

III. METHODOLOGY

The process of recognition the rubber seed series clones can be simplified as below.

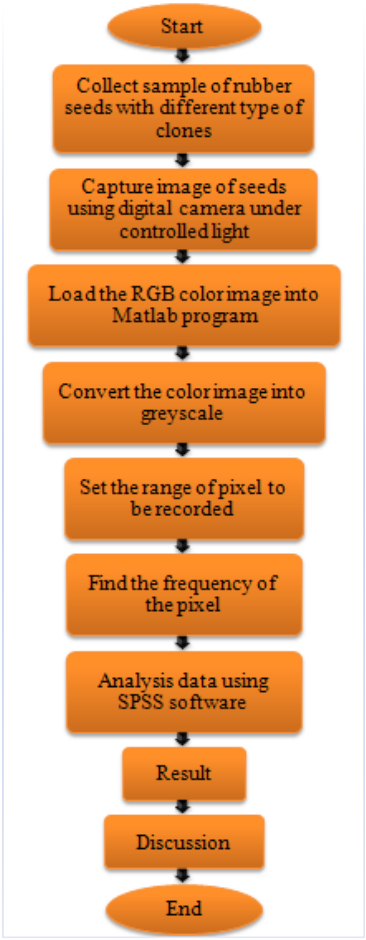

Figure 2: Methodology flowchart

A. Rubber seed samples

All the rubber seed samples that are used in this research were obtained from the Rubber Research Institute of Malaysia (RRIM), Sungai Buloh. Five clones of the RRIM2000 series is selected to for this research which are RRIM2002, RRIM2015, RRIM2020, RRIM2023, and RRIM2024. There are 45 samples of seeds for each clone used respectively for the data analysis.

B. Capturing image of samples

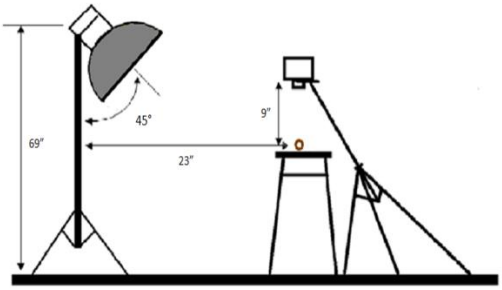

Figure 3: Arrangement of equipment during image capturing.

The inherent RGB color images were acquired by using Canon 1000D digital camera, with pixel resolution 1280x960 in JPEG format. The size is applicable for analysis. In Image Capturing Studio Room (ICSR) under environment at the Advanced Signal Processing (ASP) Lab, Faculty of Electrical Engineering UiTM Shah Alam, the image capturing process was done. To gain the sensitivity of image sensor, the International Standards Organization (ISO) was set at 800, the shutter was 40, the aperture and zoom was F8 and 55m respectively. While the photo was taken, the camera was placed at a distance of 9 inch directly above the rubber seed and light sources was provided by using (Digicolor K-2500) at 45 angle. A standard low flash having mean lux of 4.31$\pm$ 0.421 Klux by using Heavy Duty Light Marker controlling the intensity of light.

C. Data collection

The RGB color images were loaded into the Matlab program. The RGB color image were converted into a grayscale images. This grayscale images would produce a histogram with a range from 0 to 250. Which 0 is black and 250 is white. For the data collection, it is set into three ranges. First, is from 0 to 50. This represents the dark area (dark brown) which is classified as LOW. Secondly, is from 51 to 200. This represents the brighter area (light brown) which is classified as MED. Lastly; the range is from 201 to 250. This represents the white pixel which is classified as HIGH. However in this research this HIGH range of pixel is neglected as it is assumed the pixel of the background of the image. Therefore it is only been studied for the range of 0 to 200. The data obtained from the Matlab is recorded and arranged using Microsoft Office Excel for the statistical analysis later.

Figure 4(a): RGB image.

Figure 4(b): grayscale image.

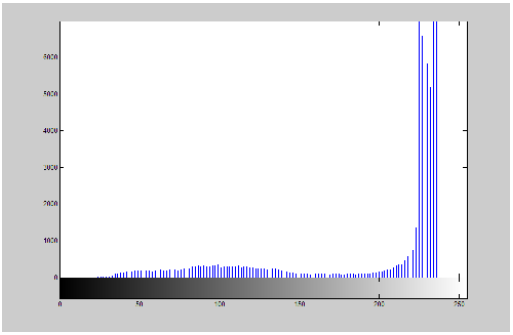

Figure 5: Pixel histogram

D. Statistical Measurement

The SPSS software was used as statistical measurement to represent the differences between the seed's image pixel. By using SPSS Software, the data was analyzed by using error plot and one-way ANOVA method.

IV. RESULT AND DISCUSSION

In order to recognize the different pixel produced between the five types of rubber tree seed clones, the data obtained earlier are analyzed using the SPSS software. Two methods which are error plot and the one-way ANOVA were utilized. There are five types of clones to be differentiating which are RRIM2002, RRIM2015, RRIM2020, RRIM2023 and RRIM2024. Both of the methods are able to prove that:

A. Error Plot

LOW pixel (0-50)

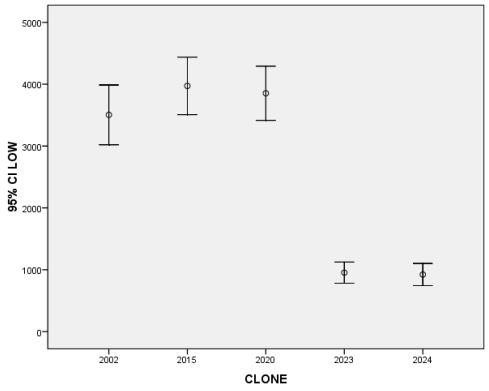

Figure 6(a): Error plot for LOW

Figure 6(a) shows the error plots based on the low range pixel of the rubber seeds. From observation, RRIM2002, RRIM2015 and RRIM2020 show almost similar to each other, so there is no significant different between these rubber seed clones. So does the RRIM2023 and RRIM2024, there is no significant different between these clone. However it can be observed that it is significantly different for RRIM2002, RRIM2015 and 2020 compared to RRIM2023 and RRIM2024. The author noted that only clones from these two separate groups (RRIM2002, RRIM2015 and 2020 compared to RRIM2023 and RRIM2024) can be recognize by using the LOW pixel of the seed image.

Table 1(a): Descriptive Statistics of LOW

	N	Mean	Minimum	Maximum
2002	45	3504.11	1407	8803
2015	45	3972.62	1427	7605
2020	45	3853.60	1371	6718
2023	45	953.04	137	2504
2024	45	922.07	27	2813
Total	225	2641.09	27	8803

Table 1(a) shows the Descriptive Statistics low pixel due to minimum, maximum and mean value for all type of clone.

MED pixel (51-200)

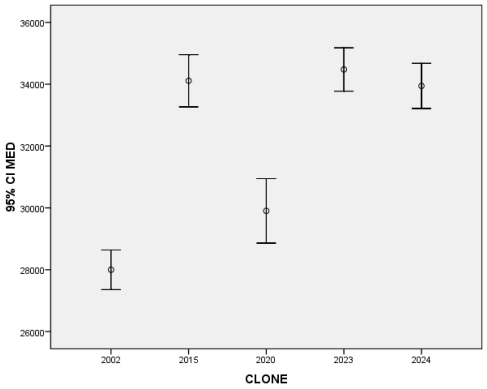

Figure 6(b): Error plot for MED

Figure 6(b) shows the error plots based on the medium range pixel of the rubber seeds. From observation, RRIM2015, RRIM2023 and RRIM2024 show almost similar to each other, so there is no significant different between these rubber seed clones. However, it is significantly different with RRIM2002 and RRIM2020 clone series. The author noted that only rubber seed clones for RRIM2002 and RRIM2020 can be recognize by using the MED pixel of the seed image.

Table 1(b): Descriptive Statistics of MED

	N	Mean	Minimum	Maximum
2002	45	2.80E4	21374	31513
2015	45	3.41E4	28567	41490
2020	45	2.99E4	22836	36577
2023	45	3.45E4	30709	40224
2024	45	3.39E4	29304	40232
Total	225	3.21E4	21374	41490

Table 1(b) shows the Descriptive Statistics medium pixel due to minimum, maximum and mean value for all type of clone.

B. Rubber seed samples

LOW pixel (0-50)

H_0 = There are no significant differences in low pixel between the rubber seeds clones' mean

H_1 = There is a significant difference in low pixel between the rubber seeds clones' mean

Table 2(a): ANOVA Table For LOW

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.407E8	4	1.102E8	70.543	.000
Within Groups	3.436E8	220	1561678.148		
Total	7.842E8	224			

Table 2(a) shows there is a significant difference in low pixel between the types of clones for all rubber seed series, the null (H_0) hypothesis can be rejected and the alternative (H_1) hypothesis is accepted since the F_{LOW} (4, 220) = 70.543, p-value < 0.05.

MED pixel (51-200)

H_0 = There are no significant differences in medium pixel between the rubber seeds clones' mean

H_1 = There is a significant difference in medium pixel between the rubber seeds clones' mean

Table 2(b): ANOVA table for MED

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.563E9	4	3.908E8	54.448	.000
Within Groups	1.579E9	220	7177291.031		
Total	3.142E9	224			

Table 2(b) shows there is a significant difference in medium pixel between the types of clones for all rubber seed series, the null (H_0) hypothesis can be rejected and the alternative (H_1) hypothesis is accepted since the F_{MED} (4, 220) = 54.448, p-value < 0.05.

C. Multiple Comparisons

LOW pixel (0-50)

Table 3(a): Multiple Comparisons of LOW

(I) CLONE	(J) CLONE	Mean Difference (I-J)	Std. Error	Sig.
2002	2015	-468.511	263.454	.389
	2020	-349.489	263.454	.675
	2023	2551.067*	263.454	.000
	2024	2582.044*	263.454	.000
2015	2002	468.511	263.454	.389
	2020	119.022	263.454	.991
	2023	3019.578*	263.454	.000
	2024	3050.556*	263.454	.000
2020	2002	349.489	263.454	.675
	2015	-119.022	263.454	.991
	2023	2900.556*	263.454	.000
	2024	2931.533*	263.454	.000
2023	2002	-2551.067*	263.454	.000
	2015	-3019.578*	263.454	.000
	2020	-2900.556*	263.454	.000
	2024	30.978	263.454	1.000
2024	2002	-2582.044*	263.454	.000
	2015	-3050.556*	263.454	.000
	2020	-2931.533*	263.454	.000
	2023	-30.978	263.454	1.000

*. The mean difference is significant at the 0.05 level.

Table 3(a) shows the comparison of ventral between all the series clones with each other in terms of mean of low pixel and significant values. The result shows that none of he clones have the entire mean significant below than 0.05. So, all the evidence is proved to assume none of the clones are significantly different between each other.

MED pixel (51-200)

Table 3(b): Multiple Comparisons of MED

(I) CLONE	(J) CLONE	Mean Difference (I-J)	Std. Error	Sig.
2002	2015	-6112.022*	564.793	.000
	2020	-1904.711*	564.793	.008
	2023	-6477.244*	564.793	.000
	2024	-5946.267*	564.793	.000
2015	2002	6112.022*	564.793	.000
	2020	4207.311*	564.793	.000
	2023	-365.222	564.793	.967
	2024	165.756	564.793	.998
2020	2002	1904.711*	564.793	.008
	2015	-4207.311*	564.793	.000
	2023	-4572.533*	564.793	.000
	2024	-4041.556*	564.793	.000
2023	2002	6477.244*	564.793	.000
	2015	365.222	564.793	.967
	2020	4572.533*	564.793	.000
	2024	530.978	564.793	.881
2024	2002	5946.267*	564.793	.000
	2015	-165.756	564.793	.998
	2020	4041.556*	564.793	.000
	2023	-530.978	564.793	.881

*. The mean difference is significant at the 0.05 level.

Table 3(b) shows the comparison of ventral between all the series clones with each other in terms of mean of medium pixel and significant values. The result shows that only RRIM2002 and RRIM2020 have the entire mean significant below than 0.05. So, all the evidence is proved to assume only RRIM2002 and RRIM2020 are significantly different between each other. Whereas other series clones are not significantly different between each other.

V. CONCLUSION AND FUTURE RECOMMENDATION

Conclusion

In this research, it is focused on recognizing the different types of rubber tree seed clones using the factor of their texture. The five types of rubber seed clones which are RRIM2002, RRIM2015, RRIM2020, RRIM2023, and 2024 have being tested in order to recognized the differences between the clones in terms of the pixel projected by the image of the rubber seed. Data extracted from the Matlab software were analyzed using statistical method. Observation from error plots and one-way ANOVA shows that each clone only differs with certain clone. Yet still the accuracy is too low to differentiate the clones. Therefore, it can be concluded that by using this method of pixel recognition is not recommended as from the results it is shown that this method may not show a clear differences between the clones.

Future Recommendation

From the result obtained, the accuracy is not high enough in order to recognize different type of rubber tree seed clones. Therefore, as a recommendation for future analysis, it is recommended that more types of rubber seed clones should be use as sample and other parameters should be used in order to increase the accuracy because as we can see the result in this project only a certain clones can be recognized by using this method it's self. Hence, from the results obtained it is known that this method cannot be used for this parameter and in the future this method may be used for other parameters in obtaining a better accuracy.

ACKNOWLEDGEMENT

The author would like to acknowledge Mr. Fairuz Nazmie Osman, the project supervisor for the important and useful guidance in completing this project. Also, special thanks to Miss Nor Ezan from the Digital Image Processing group, Faculty of Electrical Engineering for her help and valuable information. Finally, author would like to show his gratitude to his beloved family, friends and who ever that has contributed in completing this project.

REFERENCES

[1] M. Z. A. Aziz, D. R. Othman, D. M. Benong, and D. O. S. Huat *RRIM2000 SERIESCLONES:Characteristic and Description*: RUBBER RESEARCH INSTITUTE OF MALAYSIA, 1997.

[2] H. M. Salleh, "An Automated Shape Recognition for Rubber Seed Clones Through Imaging Techniques," in *Faculty of Electrical Engineering*,vol. Bsc.Eng (Hons.) Electrical. Shah Alam: Universiti Teknologi Mara, 2007.

[3] A.Bittorf, M.Fartasch, G.Schuler, and T.L.Diepgen, *Resolution requirements for digital images in Dermatology*, 1997.

[4] E. A. Anua, "An Automated Texture Recognition for Rubber Seed Clones through Wavelet Techniques," in *Faculty of Electrical Engineering*, vol. Bsc.Eng (Hons.) Electrical. Shah Alam: Universiti Teknologi Mara.

[5] Li Jinwei, Liao Guiping, Xiao Fen, "Rapeseed Seeds Colour Recognition by Machine Vision," in Institute of Agricultural Information, Human Agricultural University, Changsa 410128, P. R. China.

[6] MATLAB, "Lavenrberg Marquardt, Pixel," in *MATLAB 7.0*, 2007.

[7] Mendenhall Beaver Beaver, *Introduction to Probability and Statistic,Twelfth Edition*.