Article 11

Automatic Preharvest Grading of Harumanis Fruits

Khairul Adilah bt Ahmad
Faculty of Computer Science & Mathematics
Universiti Teknologi MARA Kedah Branch

Sharifah Lailee Syed Abdullah
Faculty of Computer Science & Mathematics
UniversitiTeknologi MARA Perlis Branch

Abstract
Fruit size is one of the most important features for grading Harumanis fruits. However, harvesting the fruit at the correct size is problematic for fruits growers. The aim of this paper is to discuss the use of image processing technique to classify the grade of Harumanis fruits before harvesting. This research adopted a computer vision methodology which include image acquisition, image pre-processing, image segmentation, feature extraction and classification. The statistical analysis which used linear regression model showed that the size has high relationship with the Harumanis' weights and grades. The results showed that it is possible to estimate the weight of fruits before harvesting using image processing technique thus enabling fruits grower to grade their products efficiently.

Keywords: *Image processing, Feature Extraction, Size, Weight, Linear regression analysis, Harumanis Mango fruits*

Introduction

The quality of Harumanis mangoes is determine by the shape and size of the fruits. Harumanis shapes and sizes influence the market value and therefore, are important features in grading of the fruits (Ercisli, Sayinci, Kara, Yildiz, & Ozturk, 2012). However, as practiced by Federal Agriculture Marketing Authority (FAMA), Harumanis fruits' quality is determine by their weights which is a post-harvest processing. For fruits growers, it is important to be able to assess the quality of fruits before harvesting, as this situation has considerable influence on high quality fruits (Jha, Chopra, & Kingsly, 2007). Therefore, it is a need to change from traditional post harvest grading to automatic preharvest grading using computer vision technique. This technique imitate the abilities of humans by electronically and automatically grading fruits and offer non-destructive methods and produce more consistent results than humans.

In recent years, computer visions are most widely used for fruits sorting and size grading. Past researches have shown the success of applying image processing technique for non destructive grading of fruits such as papaya (Riyadi, Mustafa, Hussain, & Hamzah, 2007), mangoes (Ab Razak, Mahmod, Mohd Nazari, Khairul Adilah & Tajul Rosli, 2012; Ganiron, 2014; Naik, Patel, & Pandey, 2015; Teoh & Syaifudin, 2007), apple (Hazbavi, 2014) and banana (Mustafa et al., 2008).

The aim of this paper is to discuss the use of the image processing technique to classify the grade of Harumanis fruits before harvesting based on their sizes.

Methodology

This section discusses the material and methods use to extracts information from fruits' images to estimate the quality of Harumanis. This research adopted a computer vision methodology that includes image acquisition, image pre-processing, image segmentation, feature extraction and classification (Du & Sun, 2004; Gunasekaran, 1996). The Harumanis fruit samples used for this study were obtained from the Perlis Agriculture Department's farm in Bukit Bintang, Perlis, Malaysia.

i. Image acquisition and Pre-processing

Image acquisition is a process where images are captured using digital camera. For this study the distance between the camera and the fruit were fixed at 30 cm. The images captured were saved in RGB colour format before they are converted into grayscale image.

ii. Image Segmentation

Next, image segmentation was carried out using EmaBm algorithm (Mahmod, Sharifah Lailee, Khairul Adilah, Mohd Nazari & Ab Razak, 2016). This algorithm isolated the fruit images from the background. The result of this process is a single contour of the fruit image which provide a pixels value of "0" and "1" as shown in Figure 1. This values are used to extract the features of the fruit.

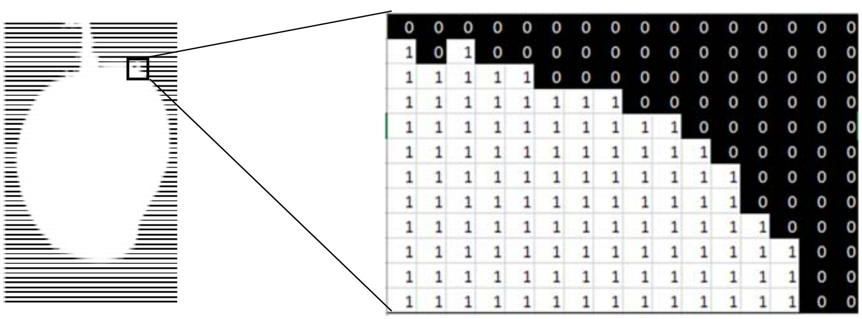

Figure 1: Pixel Value for Segmented Image

iii. Feature Extraction

For each fruit images, the features of the fruit were extracted by using image processing toolbox in the Matlab software. Area of the fruit was measured using total number of white pixels in the fruit contour region. The perimeter of fruit was measured as the total number of white pixels in the boundary contour region of binary image. The size estimation was used to determine the Harumanis grade. The size of Harumanis are graded as small, medium and large. The size of the fruits are measured using aspect ratio which are

$$Size = \frac{area}{perimeter}$$

iv. Size feature analysis

The size features was analyzed using SPSS. Variables which are area, perimeter and size were calculated and analyzed to determine the relationship between fruit weight and size. In regression analysis, a linear model was developed to estimate the weight of the Harumanis from size measurement. To evaluate the linear model, a 95% confidence and estimation intervals graph was used to verify that the linear model appropriately fit the data.

Results and Discussion

Sample size measurements of three grade levels are shown in Table1. Mean and standard deviation of the 105 training data set were 255.33g and 17.32 respectively.

Table 1: Contour Features for the Three Grades of Harumanis

Grade / Features	Large (Grade A)	Medium (Grade B)	Small (Grade C)
Area	1159157	1080690	929711
Perimeter	4188.7274	4063.0878	3799.8401
Size	276.7325	265.9775	244.6711
Actual Weight (g)	373	342	299

For each relationship, a scatter plot was produced as shown in Figure 2, Figure 3 and Figure 4, then a regression analysis was performed. Each scatter plot was fitted with a least square regression line and the equation of the coefficient of determination (r^2). This equation is the estimation equation for weight value based on size.

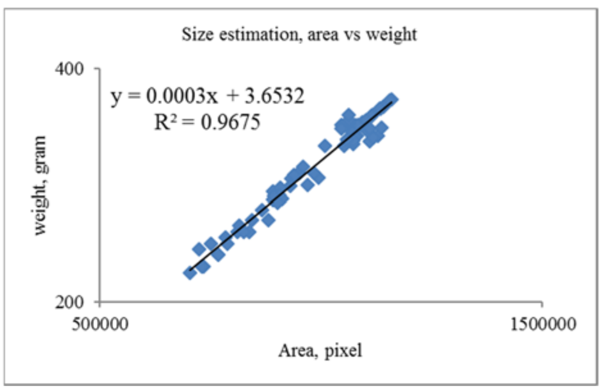

Figure 2: Scatter Plot for Area, A and Weight, W

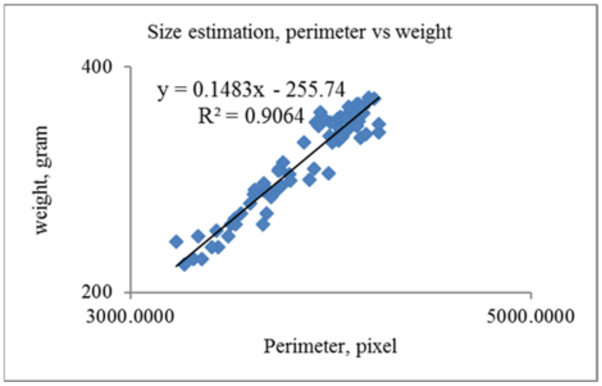

Figure 3: Scatter plot for Perimeter, P and Weight, W

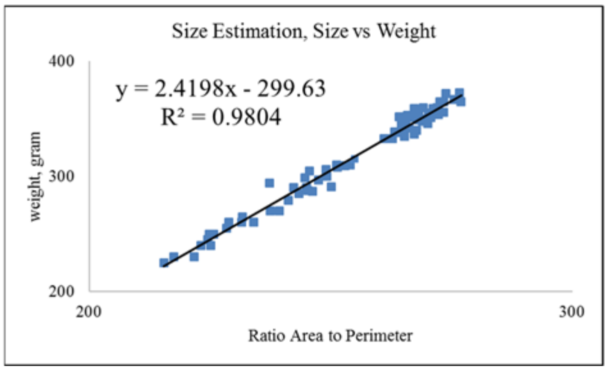

Figure 4: Scatter Plot forSize and Weight, W

A good estimation equation was selected based on R-squared (r^2) obtained for the correlation equation. Table 2 shows the estimation equations obtained for the contour area, perimeter, and size (in pixels) with their respective r^2.

Table 2: Weight Estimation Equations and their Respective r^2Values

Estimation Equations	r^2 values
Weight (gram) = 0.0003 x area + 3.6532	0.9675
Weight (gram) = 0.1483 x perimeter – 255.74	0.9064
Weight (gram) = 2.4198 x size + 184.33	0.9804

The results showed the varying correlation relationships for the features. This was seen in the high correlation coefficient values r^2 = 0.9804 obtained for the equations with weight used as the independent variable, in relation to the estimation equations with size measurement using aspect ratio. This indicates that the calculations of size measurement using aspect ratio area to perimeter performed a good mathematical descriptors to estimate weight from size measurements for determining the fruits grades.

The weight estimation from size measurements of Harumanis was computed using the ratio area to perimeter values via the feature extraction in the images analysis process. Statistical analysis of standard deviation was used for variability or diversity measurement. It shows how much variation there is from the mean. The lowest value of standard deviation of data indicates that the data points of size estimation are tend to be very close to the mean.

Root Mean Square Error (RMSE) represents the deviation between the actual weight and estimated weight from size measurements using image analysis. This parameter is used to compare two measurements from different level of errors. RMSE is frequently used to measure the differences between values predicted by a model and the values actually observed from the thing being modeled.

Table 3: Actual weight, Estimated Weight and Weight Different of Harumanis data set

Sample No	Actual Weight, gram	Estimation Weight, pixel	Weight Different
H1	373	370.01	2.9745
H2	351	346.44	5.5676
H3	356	361.99	0.0325
H4	354	349.48	1.5186
H5	355	355.87	(0.8696)
H6	353	343.66	(4.6738)
H7	367	367.15	(0.1217)
H8	365	360.01	4.9925
H9	372	362.91	9.0888
Mean			(0.0045)
Standard Deviation			5.9227
RMSE			5.8825
Min			(14.7988)
Max			19.2384

Table 3 shows nine sample Harumanis fruits with actual weights in gram, estimation weight in pixels and the weight difference between both values. The mean values of weight different between actual weights and estimation weight was 0.0045 grams. The standard deviation value of weight difference was 5.9227 grams. A low RMSE indicates high accuracy of the model. Results show that the size estimation gave the lowest values of RMSE at 5.8825. The weight estimated by the mathematical approximation was about 14.8 grams lower or 19.23 grams higher than the weight measured with digital scale. This means that the image analysis method based on the fruits' size measurements used in this experiment was sufficiently reliable to estimate the weights of the Harumanis mangoes before harvesting,

Conclusion

Measurements of the weight for Harumanis mangoes can be accomplished using EmaBm technique and size measurement. The present work has achieved the high-level in image processing and analysis technique to determine weight of Harumanis mangoes before harvesting.

References

Ab Razak, M., Mahmod, O., Mohd Nazari, A. B., Khairul Adilah, A., & Tajul Rosli, R. (2012). Mango size classification using RGB color sensor and fuzzy logic technique. *In Regional Conference on Science, Technology and Social Sciences*, pp. 287–296.

Du, C.-J., & Sun, D.-W. (2004). Shape extraction and classification of pizza base using computer vision. *Journal of Food Engineering*, 64(4), pp. 489–496.

Ercisli, S., Sayinci, B., Kara, M., Yildiz, C., & Ozturk, I. (2012). Determination of size and shape features of walnut (Juglans regia L.) cultivars using image processing. *Scientia Horticulturae*, 133, pp. 47–55.

Ganiron, T. J. (2014). Size Properties of Mangoes using Image Analysis. *International Journal of Bio-Science and Bio-Technology*, 6(2), pp. 31–42. http://doi.org/10.14257/ijbsbt.2014.6.2.03

Gunasekaran, S. (1996). Computer vision technology for food quality assurance, *71*(August).

Hazbavi, I. (2014). Shape and size grading of apple fruit (cv . Fuji) based on geometrical properties. *International Journal of Biosciences*, 4(12), pp.269–273.

Jha, S. N., Chopra, S., & Kingsly, a. R. P. (2007). Modeling of color values for nondestructive evaluation of maturity of mango. *Journal of Food Engineering*, 78(1), pp. 22–26.

Mahmod, O., Sharifah Lailee, S. A., Khairul Adilah, A., Mohd Nazari, A. B., & Ab Razak, M. (2016). The fusion of edge detection and mathematical morphology algorithm for shape boundary recognition. *Journal of Information and Communication Technology (JICT)*, 15(1), pp. 133–144.

Mustafa, N. B. A., Fuad, N. A., Ahmed, S. K., Abidin, A. A. Z., Ali, Z., Yit, W. B., & Sharrif, Z. A. M. (2008). Image Processing of an Agriculture Produce : Determination of Size and Ripeness of a Banana. *In ITSim 2008. International Symposium in Information Technology* (Vol. 1, pp. 1–7). IEEE. Retrieved from http://ieeexplore.ieee.org/

Naik, S., Patel, B., & Pandey, R. (2015). Shape, size and maturity features extraction with fuzzy classifier for non-destructive mango (Mangifera Indica L., cv . Kesar) grading. *In IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development*, pp. 1–7.

Riyadi, S., Mustafa, M. M., Hussain, A., & Hamzah, A. (2007). Papaya fruit grading based on size using image analysis, pp. 645–648.

Teoh, C. C., & Syaifudin, A. R. M. (2007). Image processing and analysis techniques for estimating weight of Chokanan mangoes. *Journal of Tropical Agriculture and Food Science*, 35(1), pp.1 83–190.