

Papaya fruit grading based on size using image analysis

Slamet Riyadi[1,2*], Mohd. Marzuki Mustafa[1], Aini Hussain[1], Azman Hamzah[3]

[1] *Universiti Kebangsaan Malaysia, Bangi 43600 Selangor, Malaysia*
[2] *Universitas Muhammadiyah Yogyakarta, Jl. Lingkar Selatan, Tamantirto, Yogyakarta 55183, Indonesia*
[3] *Malaysian Agricultural Research and Development Institute, Serdang 43400 Selangor, Malaysia*

Papaya is the largest contributor to fruit export of Malaysia. Before exportation, these fruits for export are graded according to their size, maturity and defects. The major parameter which is used to classify fruit size is their weight. Currently, the papayas are being weight individually and such practice is time consuming and labor-intensive. The advent of computers and machine vision technology offers great potential to automate this process. Therefore, the objective of this research is to estimate the papaya weight using results of the image analysis and then classify them according to their grades. The methodology involves measuring the actual volume and weight of papaya samples, and capturing their images. The characterization results showed that the weight and volume parameters are highly correlated and therefore, the derived formulation based on the collected data could be used to estimate the size of papaya. In the image processing task, the morphological procedures and segmentation using excess green color filter allow papaya images to be precisely distinguished from the background and shadow. This in turn allows the computation of the estimated volume of papaya simply by measuring the radius of the object at specific area and integrating over the length. Finally, papaya weights are estimated using the volume information. The classification ability of the proposed system when tested yields above 90% accuracy.

1. Introduction

Papaya is the largest contributor of fruit export in Malaysia (10). Before being exported, all fruits are subjected to inspection for quality control purpose and graded according to their size, maturity and presence of defect (4). For grading according to size, weight is the common and popular parameter used because since it is easy to measure. Currently, the measurement is done manually by weighing individual papaya to get a uniform size prior to packaging. However, this manual weighing procedure is time consuming and labor-intensive. Furthermore, it can cause surface damage to the papaya yield and in addition, inefficiency in handling will result in low productivity due to huge production volume (9).

Computer vision systems have been widely used in the product quality inspection and grading. Heinemann et al has developed an automated machine vision for grading of potatoes that are capable of classifying moving potatoes with 88% accuracy and 97% accuracy for stationary potatoes (7). Another similar work was reported on apple classification with accuracy 89.2% (1). However, up to now no such work relating to machine vision papaya grading system has been reported Machine vision has been proven to offer great potential to overcome the disadvantages of the manual grading or inspection system. Based on the analysis and the characterization study performed onto the papaya images, we strongly believe that we can obtained meaningful descriptors to enable efficient automated papaya grading system be realized.

Therefore, this paper will discussed and provide details relating to the implementation of the proposed automated papaya grading system using the estimated papaya weight computed based on the papaya images. The grading of papaya is based on the grading system set by the Federal Agricultural Marketing Authority (FAMA) of Malaysia. Table 1 enlists the grading regulation for the exotica papaya for export (4).

Table 1. Size grading of exotica papaya

Size Grade	Weight (grams)
XL	> 850
L	650 – 850
M	450 – 640
S	250 – 450

2. Methodology

Methodology involving the papaya fruit grading system includes data acquisition task, determination of correlation formula, image processing task and computation of estimated weight. Fig. 1 summarizes the procedures involved.

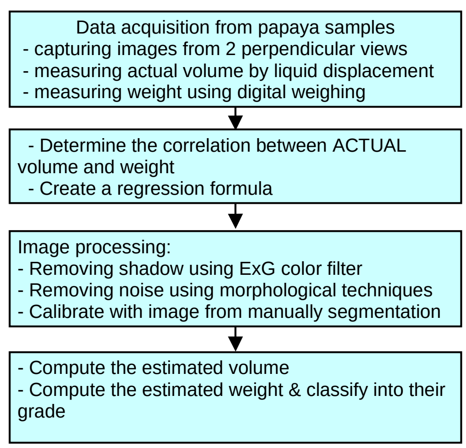


Fig. 1. Methodology of the research

2.1 Data acquisition

A total of 71 Exotica papayas of varying size were collected from one of the papaya orchard in Selangor, Malaysia. These fruits were then brought back to the laboratory for preliminary inspection to determine their weight digitally and measuring the actual volume using liquid displacement method. The characteristic of the exotica papaya in term of weight and volume are then studied to determine the correlation between the two parameters. A high correlation implies that if the volume of papayas were to be determined from image analysis then their weight could be estimated using the derived formula.

* Responsible author. Email: riyadi@umy.ac.id

ISBN 978-979-16338-0-2

The images of papayas were captured at random orientation from two perpendicular views using the Olympus Camedia C-5050 digital camera. The papaya is first laid flat on the surface and the first view was taken at a random horizontal position followed by the second view in which the papaya is rotated 90^0 from the first captured view. Image features such as diameter, area and volume were calculated by averaging the features of both views. Images were taken with camera flash in standard room lighting. The camera was setup in a fix position to get an appropriate silhouette of object. A bright yellow paper was used as background surface to facilitate and simplify the segmentation task.

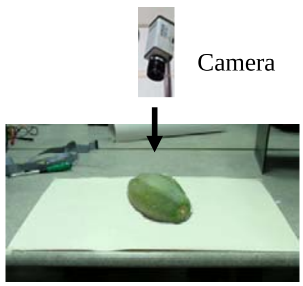


Fig. 2. Configuration of capturing images of papaya

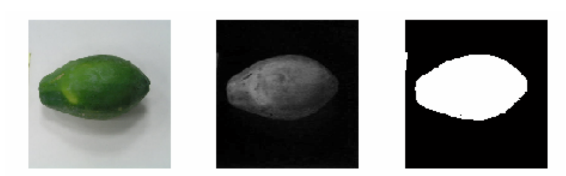
2.2 Image processing

In this work, volume estimation are measured based on silhouette image of objects. The presence of shadow will affect the silhouette. Therefore, the object shadow should be removed before performing the volume estimation. For this task, excess green (ExG) color filter was implemented to separate the object from both its shadow the background.

ExG filter has successfully been applied to separate plant and soil region for weed species identification research (5). The ExG is defined as follows:

$$\text{ExG} = 2\text{G} - \text{R} - \text{B} \quad (1)$$

with constraints: if (G<R or G<B or G<120) then ExG=0. Fig. 3(a) depicts the original papaya image that includes the background and shadow. The result of ExG color filter implementation is a gray scale image as shown in Fig. 3(b) and 3(c) displays the binarized image of the papaya obtained via thresholding.



(a) (b) (c)

Fig. 3. ExG color filter implementation results (a) original image (b) ExG filtered image and (c) Binarized image

Fig. 4 shows the matrix of the gray scale values of the fruit object. By setting the threshold value to 20, a high gray scale value (>20) can be separated from the low gray scale value(<20) in an attempt to separate values greater than 20 represent the object where as values less than 20 symbolize the shadow and background.

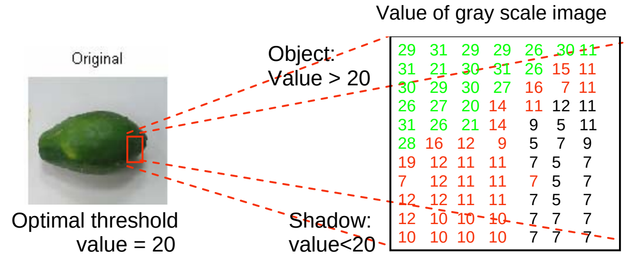


Fig. 4. Determining the optimal thresholding level

The room lighting or camera flash acts as noise and can affect the image quality (Fig. 5.a). Morphological algorithm based on dilation, complementation and intersections was implemented to remove the noise using region filling technique. As shown in Fig. 5 (b) the black dot ('0') inside the white object represents noise. Beginning with a point *p* of the noise, the objective is to fill the entire region of noise with '1' using the following procedure (8)

$$X_k = (X_{k-1} \oplus B) \cap A^C \qquad k = 1, 2, 3, \ldots \quad (2)$$

where $X_0 = p$ and

B : represents the symmetric structuring element

A : represents the original binary image.

The algorithm is terminated if $X_k = X_{k-1}$.

a. Original b. Binary image c. Free noise

Fig. 5. Step-by-step results of the region filling technique to remove noise inside the object

In order to calibrate the result, segmentation and noise removal tasks were performed manually using photo editor software as a reference image. Then, the area of the object obtained from the proposed technique and the reference image were compared.

2.3 Computation of estimated volume and weight

Grading based on size will provide accurate results if the volume can be accurately estimated from the image (2). The fruit produce should also have a constant density, or at least consistent within a batch. Without these two conditions, fruits can be of the same size but their weights may vary. In (11), to estimate the volume of an orange, a spherical approximation was used and it is defined as

$$Vorange \approx \frac{\pi d^3}{6} \quad (3)$$

Volume estimation using the above mentioned approximation formula varies greatly when compared to the liquid displacement technique. Another formulation was developed to estimate potato volume as a function of the area profile (6). The potato volume estimation formula is as shown below:

$$Vpotato \approx \frac{(projected\_area)^2}{lenght} \quad (4)$$

ISBN 978-979-16338-0-2

This formula has been applied to 12 different images of potatoes that were rolled along the conveyor. With the assumption that the density of potato is relatively constant, the estimated potato weight differs slightly from the actual with an average error of ±7%.

A 3-D model used to estimate the volume of irregular shape objects is as shown in Fig. 6 (3)

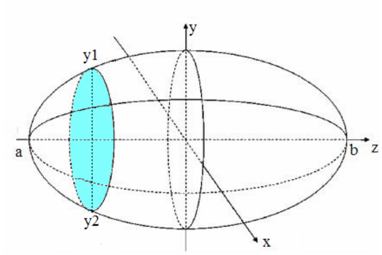


Fig. 6. A 3-D model of irregular shape object

While the major axis of papaya lies on the horizontal position, the shaded area is

$$A(z) = \pi(d/2)^2 \quad (5)$$

where $d = y_2 - y_1$ is the diameter of shaded area. This approximation formula assumes that the papaya has a circular cross sectional area. For selected *dz*, the volume can be defined as the integral of the cross-sectional area *A(z)* with respect to *dz*. Therefore, the object volume can be computed using Eq. 6 below

$$V = \int_a^b A(z)dz = \sum_{z=a}^{b} A(z).\Delta z \quad (6)$$

If $\Delta z = 1$ pixel is chosen, then the volume in cubic pixels can be calculated using Eq. 7.

$$V = \sum_{z=a}^{b} A(z) \quad (7)$$

According to (3), this volume estimation technique has the ability to estimate correctly volume of an apple with an average error of ±2.5%. In this work, volume of papaya was computed by averaging two 2-D images using Eq. (7).

The estimated volume of papaya is calculated in milliliters and compared to its actual volume. A reference value was taken from a 3-D object with known actual volume.

Before computing the estimated volume, major diameter of papaya is normalized to horizontal axis using Hotteling transform. Hotteling transform rotates the data so its primary axis lies along the axes of the coordinate space and moves it so that its center of mass lies on the origin. Graphical illustration of the Hotelling transform is as shown in Fig. 7.

The algorithm involves determining the coordinate vectors of the object $x=[x_1, x_2, \ldots, x_N]^T$ where N is the number of data. The mean and covariance of the vector can be defined as

$$m_x = E\{x\} = \frac{1}{N}\sum_{k=1}^{N} x_k \quad (8)$$

$$C_x = \frac{1}{N}\sum_{k=1}^{N} x_k x_k^T - m_x m_x^T \quad (9)$$

respectively (8). As such, the Hotteling transform can be defined as

$$y = V(x - m_x)$$

where *V* is the eigenvector of the covariance matrix.

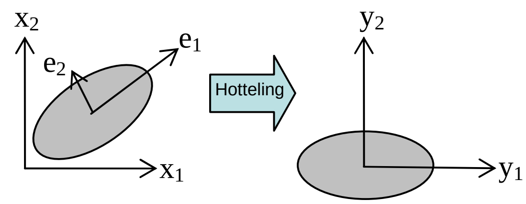


Fig. 7. Principle of Hotteling tansform

(a). Original image (b). After transformation

Fig. 8. Hotteling transform

Result obtained after the Hotelling transformation is as shown in Fig. 8. Finally, papaya weights are estimated using the correlation formula between volume and weight obtained earlier and the exotica papayas are graded accordingly using Table 1 as guidelines.

3. Result and discussion

The characterisation results of papayas showed that the volume and weight parameters are highly correlated (Fig. 9). The correlation formula derived based on the collected data is determined as

$$w = 0.8484V + 34.353 \quad (10)$$

where *w* is estimated weight in grams and *V* is estimated volume in milliliters. This weight formula is then used to grade the size of papaya using the estimated volume information of the papaya that was calculated via image analysis approach.

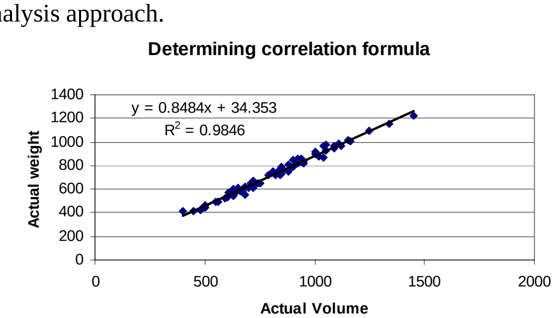


Fig. 9. Correlation between actual volume and weight of papaya samples

Accuracy of area estimation is a significant factor that allows for accurate volume estimation be calculated based on silhouette image.

ISBN 978-979-16338-0-2

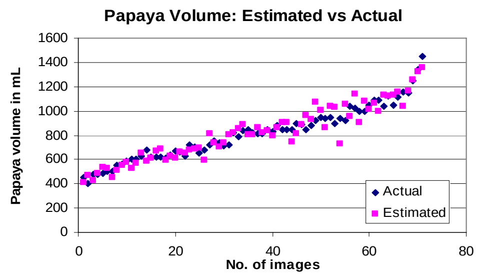


Fig. 10. Volume estimation results

Fig. 10 shows both the estimated volume results and the actual volume of the papayas. It can be seen that the estimated volume of the papayas do not differ too much from the actual. On the average, the estimated papaya volume has less than 7% error. The large error mainly originated from the irregular shape papaya since our model is based on a 3D model with circular cross sectional area. Sample image of an irregular shape papaya is as shown in Fig. 11.

Fig. 11. Example of an irregular shape papaya

Next, the estimated weight of the papaya is computed using the estimated volume information. Weight estimation results of each papaya via the image analysis approach are shown in Fig. 12. Using the estimated weight information, the papayas are graded according to size. The classification of papayas by size grading using our proposed method yields above 90% accuracy (Fig. 13). Misclassification occurs mainly due to irregular shaped papaya that resulted in inaccurate volume estimation which in turn affects the weight estimation computation. Additionally, some of the estimated weights fall in between the grade categories and as such causing the misclassification.

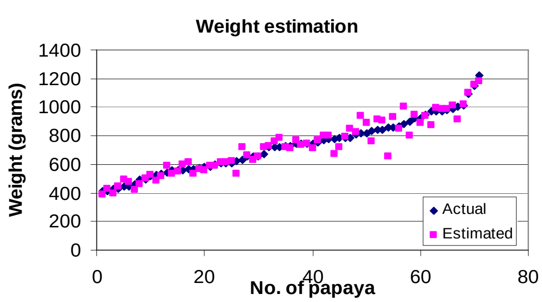


Fig. 12. Weight estimation results

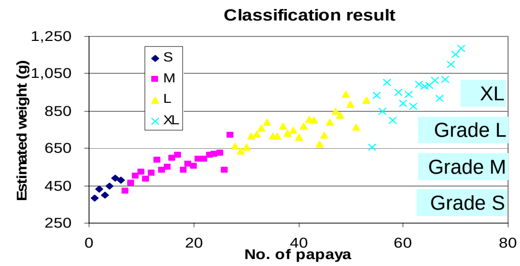


Fig. 13. Result of papaya classification

4. Conclusion

This paper affords an automated papaya fruit grading system which uses an algorithm to estimate papaya weight using image analysis approach. The method involves ExG color filter and morphological operators implementation to segment the object from the shadow, noise and background. The area information of the segmented object is then used to estimate the object volume and thereafter the weight is estimated. Finally, papaya samples are graded according to size using the estimated weight information. The capability of the developed system looks promising with more than 90% correct classification accuracy based upon our preliminary grading results. More testing will be done which require more papaya samples in the near future to validate our proposed system.

5. Acknowledgment

The authors would like to acknowledge the Ministry of Science, Technology and Innovation Malaysia (Project No. 03-01-02-SF0178) for the support of this research.

References

(1) D. Unay and B. Gosselin: Thresholding-based segmentation and apple grading by machine vision, *Proc. of EUSIPCO* (2005)
(2) D.G. Baely, K.A. Mercer, C. Plaw, R. Ball, H. Barraclough: High speed weight estimation by image analysis, *Proceedings of the 2004 New Zealand Conference on Non Destructive Testing* (2004)
(3) D.J. Lee, X. Xu, J. Eifert, P. Zhan: Area and volume measurements of objects with irregular shapes using multiple silhouettes, *Journal of Optical Engineering*, 45(2), 027202-1 - 027202-11 (2006)
(4) Federal Agricultural Marketing Authority: Quality guidelines series of Exotica Papaya, *www.famaexchange.org* , in Malay (2004)
(5) G.E. Meyer, T. Mehta, D.A. Mortensen, A. Samal: Textural imaging and discriminant analysis for distinguishing weeds for spot praying, *Transaction of ASAE*, 41(4), 1189 – 1197 (1998)
(6) J.A. Marchant: A Mechatronic Approach to produce grading, *In Mechatronics: Designing Intelligent Machin*e, 159 – 348 (1990)
(7) P.H. Heinemann, N.P. Pathare and Ch.T. Morrow: An automated inspection station for machine-vision grading of potatoes, *Machine Vision and Applications Journal*, 9, 14-19 (1996)
(8) R. Gonzalez, R.E. Woods: *Digital Image Processing*. 2^{nd} Ed. New Jersey: Prentice Hall (2001)
(9) T. Brosnan and D.W. Sun: Inspection and grading of agricultural and food products by computer vision systems – a review, *Computer and Electronic in Agriculture Journal*, 36, 193 – 213 (2002)
(10) Treasury Malaysia: Economic Report 2002/2003, *www.treasury.gov.my*, (2003)
(11) W.M. Miller, K. Peleg, and P. Briggs: Automatic density separation for freeze damaged fruit, Applied Engineering in Agriculture Journal, 4(4), 344 – 348 (1998)