

SHAPE CLASSIFICATION OF HARUMANIS MANGO USING DISCRIMINANT ANALYSIS (DA) AND SUPPORT VECTOR MACHINE (SVM)

F.S.A.Sa'ad^{1*}, M.F.Ibrahim¹, A.Y.Md. Shakaff¹,
A. Zakaria¹, M.Z.Abdullah², A.H.Adom¹

¹Centre of Excellence for Advanced Sensor Technology,
Universiti Malaysia Perlis,
02600 Arau, Perlis, Malaysia

²School of Electrical and Electronic Engineering,
Universiti Sains Malaysia, 01000 Penang Malaysia.

ABSTRACT

The perceived quality of fruits, such as mangoes, is greatly dependent on many parameters such as ripeness, shape, size, and is influenced by other factors such as harvesting time. Unfortunately, a manual fruit grading has several drawbacks such as subjectivity, tediousness and inconsistency. By automating the procedure, as well as developing new classification technique, it may solve these problems. This paper presents the novel work on the using visible Imaging as a Tool in Quality Monitoring of Harumanis Mangoes. A Fourier-Descriptor method was developed from CCD camera images to grade mango by its shape. Discriminant analysis (DA) and Support vector machine (SVM) were applied for classification process and able to correctly classify 98.3% for DA and 100% for SVM.

KEYWORDS: *Infrared imaging; visible imaging; machine vision; fourier descriptor; harumanis mango; grading system; automated inspection*

1.0 INTRODUCTION

Mango (*Mangifera indica* L.) belongs to the family Anacardiaceae. It is the only species grown extensively and commercially in India, Philippines, tropical Australia, the lowlands of South-East Africa, in Hawaii and in the lowlands of Central and South America. The Malayan name of mango (*mangga*) attests its origin outside Malaya, being the same word as the Tamil *mangas*.

Mango is an important commercial fruit crop throughout the world particularly in South East Asian countries, e.g. Malaysia, India,

* Corresponding author email: fathinul@unimap.edu.my

Indonesia, Sri Lanka, and Thailand and also in African countries. The Mango is a popular evergreen fruit tree natural to South-Eastern Asia. It has been cultivated for over 4000 years during which time it has spread to other tropical and sub-tropical countries. The mango is probably a more important fruit in the tropics than the apple in the temperate zones. It is universally considered as one of the finest fruits in the world. The fruit is a good source of vitamins A and C.

Mangoes imported from other parts of the world, especially Thailand, Mexico and the Philippines, are usually available all year round but in Perlis, Malaysia there is one unique and famous mango label Harumanis and this fruit is seasonal. Harumanis is considered the "*King of Mangoes*" and is very popular in Malaysia because of its deliciousness, sweet and aromatic fragrance (Rosidah Musa et al., 2010). Compared with other mango varieties, Harumanis is a temperamental fruit. It can only be cultivated in Perlis because it needs a four-month-long dry season of at least 40 degrees Celsius to flower and fruit. Rain, even drizzles, during this crucial period will spoil the yield. After the fruit is set, it needs to be wrapped in waterproof paper. Precisely eight weeks after wrapping, the fruits must be manually harvested, washed to rinse off residues, treated in hot water for five minutes to eliminate fruit fly and seed weevil larvae, and then individually wrapped again before being placed in a ripening chamber for three days. It is a labour-intensive process and there are no shortcuts.

Japan showing strong interest in Harumanis, Perlis is hoping to increase its annual yield, which currently stands at 350 tonnes for local consumption. The bigger challenge, however, lies in persuading local farmers to apply good agricultural practices. Figure 1. show harumanis mango farm in Universiti Malaysia Perlis, Perlis Malaysia.



Figure 1. Harumanis Mango farm

2.0 METHODOLOGY

2.1 Background

Mango quality indices include the uniformity of shape and size, freedom from decay and defects, skin colour that is characteristic of the cultivar, flesh colour, flesh firmness (juiciness, fibre content), and flavour (sweetness, acidity, aroma intensity). Sucrose and citric acid are the main sugar and organic acid, respectively, in mangos. Lactones contribute to preferred mango aroma and 2, 5-dimethyl - 4 - hydroxy - 3(2H) - furanone contributes to overripe aroma and flavour (Zainon, 1995).

Ripening process is the result of complex biochemical changes. During ripening, the green skin of mango will turn to yellow except for some cultivar such as Harumanis, due to breakdown of chlorophyll and unmasking of the carotenoid pigments. Harumanis is marketed as green-ripened fruit with green skin and soft-edible pulp. This green-ripened fruits fail to attract consumer attention as consumer perceived it as unripe (Zainon, 1995). It has been shown that the external color of the fruit is an important factor in consumer preference. Growers and farmers are inevitably having difficulties to sort out ripening Harumanis mango based on skin colour.

Maturity level determination based only on skin colour often leads to erroneous predictions. Hence, maturity level determination with information on fruits chemical and soluble solids changes, respiration, and ethylene production rate will be more scientific and appropriate when it goes for export markets (Miyachi & Perry, 1999). Currently most of the methods of estimating chemical component changes, respiration and ethylene changes are destructive, time consuming and tedious.

In recent years, the automatic visual inspection systems have become useful tools in industrial and agricultural process to grade the quality of fruit. Grading is an important operation to measure size, color, shape and defect of agricultural product. Fresh agricultural products are graded into quality categories according to their external features such as color, size, shape and texture (Leemans & Destain, 2004) reported that the Fresh market fruits like apples are graded into quality categories according to their size, colour and shape and to the presence of defects.

Presently, the quality inspection was done manually by the workers (Figure 2) and there are difficulties in enforcing these standards,

especially where it entails a large amount of mango to be evaluated, made no easier by increasing difficulty in hiring personnel who are adequately trained and willing to undertake the tedious task of inspection.



Figure 2. Current inspection for Mango

For this reason, this paper, we present the methods and techniques of an internal and external automatic grading system inspections using infrared and visible imaging to standardize the grading scheme of mango thus promoting the quality awareness amongst the planters and producers.

2.2 Elements of Machine Vision and conveyer System

The procedures and methods are implemented on machine vision workstation, which included a Personal Computer (PC), LabVIEW 2012 software, an illumination system, a charge coupled device (CCD) camera, VarioCam thermographic camera, NI CompactRIO and a conveyer system. Figure 3 shows NI compactRIO was used to control the conveyer system for inspection and sorting process, both CCD camera and VarioCam thermographic camera were used for capturing the image and LabVIEW 2012 software was used for image analysis and classification process.



Figure 3. Elements of machine vision and conveyer system.

The particular image analysis and processing developed in this study consisted of three levels: low-level processing, intermediate level processing and high-level processing (Figure 4). In summary, the first group includes image acquisition and pre-processing such as image enhancement, extraction and restoration. Meanwhile, the second group concerned with the image transformation such as RGB to HSI transformation, segmentation and filtering. Finally, the third group involved recognition and interpretation.

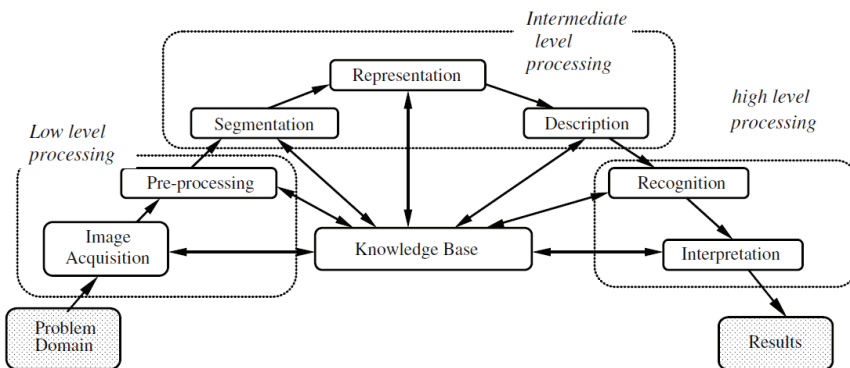


Figure 4. Different level in the image processing

2.3 Shape Analysis

Harumanis mango can be viewed adequately by 2-Dimensional perspectives; therefore they are most suitable for real-time machine processing. Figure 5 shows Grade of Harumanis mango.

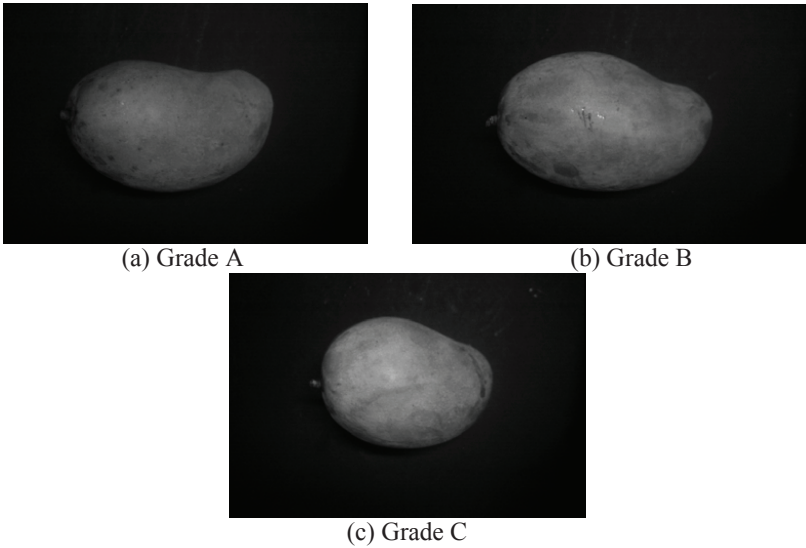


Figure 5. Grade of Harumanis mango by shape

Presently, there are many methods available for analyzing shape of an object, ranging from a simple multiple point features method to a complicated geometric features approach. The method used in this project was conceptualised by Zahn and Rookies (1972). It was based on Fourier Descriptors (FD). They provide details mathematical explanation of FD for object recognition, matching and registration. One unique feature of this method is that it uses global image descriptors instead on the local ones, making it more applicable to real-world images in which simple multiple point features may be difficult to extract, and eliminating the need for feature matching between the reference and observed images.

Before this method could be implemented, several image pre-processing operations were performed on Harumanis mango image. The image was firstly binarised with an adaptive threshold, and secondly, processed via a sequence of morphological image processing. Finally, the object centroid was extracted using first-order geometric moments and derived using Green's theorem shown in Figure 6.

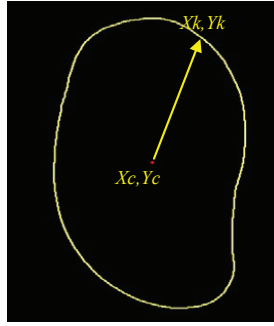


Figure 6. Object centroid and boundary

Mathematically, the two-dimensional centroid (x_c, y_c) is given;

$$x_c = \frac{\sum_{k=0}^N y_k (x_k^2 - x_{k-1}^2) - x_k^2 (y_k - y_{k-1})}{2 \sum_{k=0}^N y_k (x_k - x_{k-1}) - x_k (y_k - y_{k-1})} \quad (1)$$

and

$$y_c = \frac{\sum_{k=0}^N y_k^2 (x_k - x_{k-1}) - x_k (y_k^2 - y_{k-1}^2)}{2 \sum_{k=0}^N y_k (x_k - x_{k-1}) - x_k (y_k - y_{k-1})} \quad (2)$$

Where N is the total number of boundary pixel defined in a clockwise direction from any starting point; are the coordinates of the boundary pixel, k . The distance of each boundary point to the centroid was calculated as follows:

$$R(k) = \sqrt{(x_k - x_c)^2 + (y_k - y_c)^2} \quad (3)$$

The $R(k)$ was then subjected to Discrete Fourier Transform (DFT), yielding a one-dimensional feature vector of Harumanis mango. In Fourier space, such transformation was mathematically implemented as follows:

$$|F(m)| = \frac{1}{N} \sqrt{\left[\sum_{k=0}^N R(k) \cos\left(\frac{2\pi mk}{N}\right) \right]^2 + \left[\sum_{k=0}^N R(k) \sin\left(\frac{2\pi mk}{N}\right) \right]^2} \quad (4)$$

Since the descriptors are influenced by the curve shape and by the initial point of the curve, therefore, calculating and examining each harmonic component in $|F(m)|$ provide some clues of the shape. For a given shape, the plot of Fourier descriptors produces a pattern or

fingerprint which uniquely describe this shape. In theory, the order of Fourier descriptors ranges from zero to infinity. However, one favourable property common to Fourier descriptors is that the high-quality boundary shape representation can be obtained using only a few lower-order coefficients. Therefore, only the first few components of $|F(m)|$ are distinct and generally required to distinguish the difference between mango shapes.

3.0 RESULTS AND DISCUSSION

3.1 Shape analysis

The experiment conducted graded the fruits to three grades, from grade A, representing the best quality grade, to grade C. The quality of each fruit was inspected by inspectors, who studied the mango one at a time and made judgment as to fruit quality. Damaged or injured fruits and fruits infected by insects, diseases, blemishes such as scars, scabs and staining were discarded.

Approximately 300 fruits were sampled; the first 300 samples were allocated to test set, comprising of 100 samples for each grade. These samples were then imaged and later used to train machine vision system for shape inspection and recognition. The test sets were used to evaluate the machine vision accuracy for shape, weight and maturity inspection. In the same way, the inspectors looked at the mangoes one at the time and performed classification based on the quality feature of each fruit.

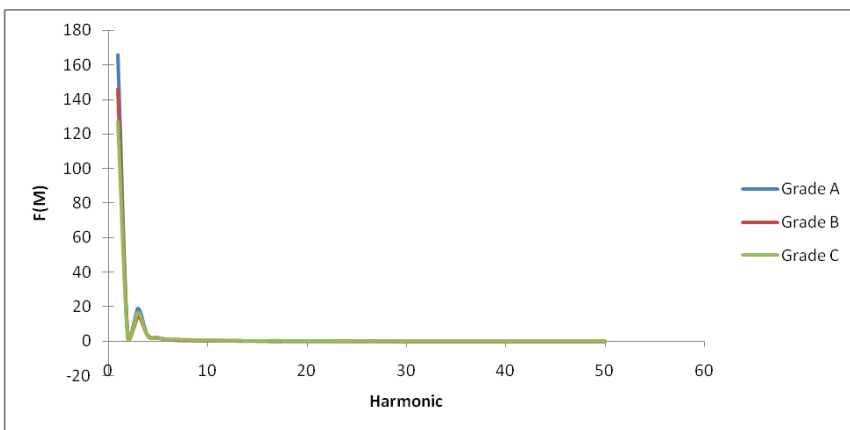


Figure 7. shape analysis based Fourier Descriptor

The resulting image was used to calculate the moments and finally the Fourier of the object. Figure 7 shows the plot of normalised Fourier descriptors for each shape category. Clearly from Figure 7, the Grade A, B and C can be characterised by $|F(3)|$. The method based on direct thresholding cannot be used because of the difficulty in establishing a single and effective shape threshold. A different approach is needed to solve this type of pattern recognition problem by applying discriminant analysis (DA) and Support Vector Machine (SVM) to establish classification

3.2 Artificial inspector

3.2.1 Discriminant Analysis (DA) Learning for Shape Analysis

The objective of DA learning is principally to identify a subset of dominant features that are most responsible for splitting a set of observations into two or more groups. Frequently, the successful application of shape recognition using machine vision system relies strongly upon the choice of proper spectral range and the number of variables employed in the calibration model.

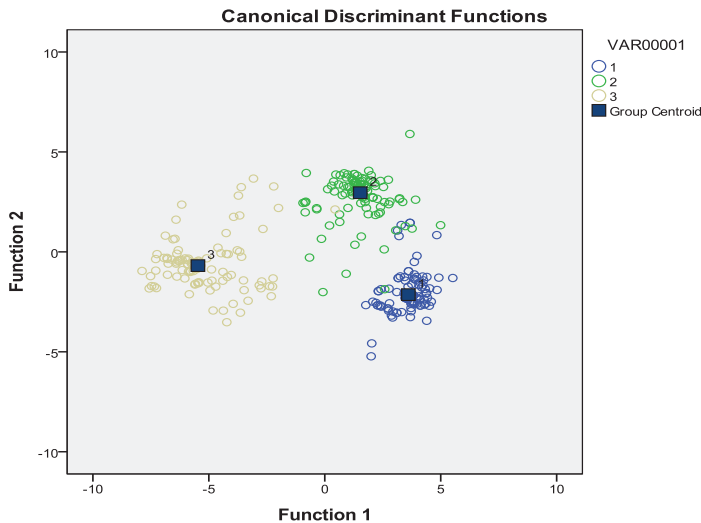


Figure 8. Canonical Plot for Harumanis mango grade

The discrimination power of these principal FD can be examined by studying the Mahalanobis' distances which are shown canonically in Figure 8. This plot clearly demonstrates that the mango separate into three groups, corresponding to three different grade. A close inspection of Figure 8 reveals that Grade A, B and C groups separate distinctively among themselves and Table 2 show the classification result for grade A, B and C are 98.3% classify.

Table 2. Classification Result for Harumanis Grade by shape analysis

				Predicted Group			Total
				Membership			
VAR00001				1.00	2.00	3.00	
Cases Selected	Original	Count	1.00	62	1	0	63
			2.00	2	57	0	59
			3.00	0	0	58	58
		%	1.00	98.4	1.6	.0	100.0
			2.00	3.4	96.6	.0	100.0
			3.00	.0	.0	100.0	100.0
	Cross-validated ^a	Count	1.00	61	2	0	63
			2.00	5	50	4	59
			3.00	1	3	54	58
		%	1.00	96.8	3.2	.0	100.0
			2.00	8.5	84.7	6.8	100.0
			3.00	1.7	5.2	93.1	100.0
Cases Not Selected	Original	Count	1.00	33	4	0	37
			2.00	3	38	0	41
			3.00	0	4	38	42
	%	1.00	89.2	10.8	.0	100.0	
		2.00	7.3	92.7	.0	100.0	
		3.00	.0	9.5	90.5	100.0	

100 % of selected original grouped cases correctly classified.

97.62% of unselected original grouped cases correctly classified.

100 % of selected cross-validated grouped cases correctly classified.

3.2.2 Support Vector Machine(SVM) Learning for Shape Analysis

SVM are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. The basic SVM takes a set of input data and predicts, for each given input, which of two possible classes forms the output.

Table 3. Classification Result for Harumanis Grade by shape analysis

				Predicted Group Membership			Total		
				GredA	GredB	GredC			
Case Selected	Original	Count	Group						
			GredA	70	0	0	70		
			GredB	0	70	0	70		
				GredC	0	0	70	70	
				%	GredA	100.00	0.00	0.00	100
				GredB	0.00	100.00	0.00	100	
	Cross-validated	Count	Group	GredA	69	1	0	70	
				GredB	0	66	4	70	
				GredC	0	1	69	70	
%		GredA	98.57	1.43	0.00	100			
		GredB	0.00	94.29	5.71	100			
		GredC	0.00	1.43	98.57	100			
Case Not Selected	Original	Count	GredA	30	0	0	30		
			GredB	0	30	0	30		
			GredC	0	0	30	30		
	%	GredA	100.00	0.00	0.00	100			
		GredB	0.00	100.00	0.00	100			
		GredC	0.00	0.00	100.00	100			

From the Table 3, Clearly demonstrates that the mango separate into three groups, corresponding to three different grade. A close inspection of table 3 reveals that Grade A, B and C groups separate distinctively among themselves show the classification result for grade A, B and C are 100% classify.

4.0 CONCLUSIONS

A Fourier description method was developed from CCD (visible) data to grade mango by its shape and the results were able to correctly classify 98.3% using DA and 100% using SVM by its' shape. The data can provide a real alternative to human expert panel in non-destructive fruit quality assessment for mango.

ACKNOWLEDGMENT

This research has been supported by a Fundamental Research Grant Scheme (FRGS) Multi Modalities Sensor Fusion for Quality Assessment of Agro based product (9003-00250).

REFERENCES

- Alfatni, M. Shariff, et al. (2008). "Oil Palm Fruit Bunch Grading System Using Red, Green and Blue Digital Number." *Journal of Applied Science* 8(8):1444-1452.
- Gonzalez, R.C. & Woods, R.E. (2002): *Digital Image processing*. 2nd edition, Prentice Hall, New Jersey, USA.
- Heinemann, P.H., Varghese, Z.A., Morrow, C.T., Sommer III, H.J. & Crassweller, R.M. (1995). Machine vision inspection of golden delicious apples. *Transactions ASAE* 11(6):901-906.
- Heinemann, P.H., Pathare, N.P. & Morrow, C.T. (1996). An automated inspection station for machine-vision grading of potatoes. *Machine Vision and Applications*. 9:14-19.
- Leemans, V. & M.-F. Destain (2004). "A real-time grading method of apples based on features extracted from defects". *Journal of Food Engineering*. 61:83-89.
- Malamasa, E. N., E. G. M. Petrakisa, et al. (2003). "A survey on industrial vision systems, applications and tools" *Image and Vision Computing*. 21:171-188.
- M.Z. Abdullah, J. Mohamad-Saleh, A.S. Fathinul-Syahir, B.M.N. Mohd-Azemi (2006). Discrimination and classification of fresh-cut starfruits (*Averrhoa carambola* L.) using automated machine vision system, *Journal of Food Engineering*, 76(4):506-523.
- Miyauchi, Y. & Perry, C. (1999). Marketing fresh fruit to Japanese consumers: exploring issues for Australian exporters. *European Journal of Marketing*. 33(1/2):196-214.
- Rosidah Musa, Faridah Hassan, Jamaliah Mohd Yusof & Norzaidi Mohd Daud (2010). "Examining market accessibility of Malaysia's Harumanis mango in Japan:challenges and potentials" *Business Strategy Series*. 11(1):3-12.
- Tao, Y., Morrow, C.T., Heinemann, P.H. & Sommer III, H.J. (1995). Fourier-based separation technique for shape grading of potatoes using machine vision. *Transactions ASAE*. 38(3):949-957.
- Wang, Z., Chu, Z., & Feng, D. (2003). Shape based leaf image retrieval. *IEEE Proceedings Vision, Image and Signal Processing*. 150(1):34-43.
- Zahn, C.T. & Rookies, R.Z. (1972). Fourier descriptors for plane closed curves. *IEEE Transactions of Computers*. C(21):269-281.
- Zainon Mohd Ali, Santhi Armugam Fl-Galactosidase And Its Significance In Ripening Mango Fruit, (5):1109-1995.