Grading of fruits on the basis of quality using Image Processing

Vardhini Sharma
Computer Science and Engineering
PIET, Jaipur
India

Smarika Dutta
Computer Science and Engineering
PIET, Jaipur
India

***Abstract*—In the era of computer age, with the invention of new technologies the need tocompute with accuracy is increasing. The natural approach of detection of fruit quality is done by the experts on the basis of human eye. Automation of quality assessment of fruits is important in order to reduce human efforts and save time. Image processing can be used to detect the quality of fruit which includes extraction of morphological features. After feature extraction, feature vectors are formed on which K-Means clustering is applied to form clusters. The formation of different clusters helps in grading of fruit.**

Keywords: Image Processing, morphological features, K-means clustering

I. Introduction

An image being worth a thousand words can be used to extract quality information, which can be of utmost importance in various fields like machine learning, robotics, computer vision and medical field.Worth of a product depends upon its quality so quality assessment came into existence. Traditionally, the quality was assessed under human supervision which is quite cumbersome so automating it will contribute in development of smart cities. Computerization in work has made the lives of human easier. Nowadays there is automation in various sectors like traffic control, medical field, education, telecom companies and food industries. Fruit being a major cultivable product of India is exported overseas so its quality is a major concern. Automating this area will enhance its performance and efficiency which would play a major role in development of smart cities.

For grading and quality assessment of fruit, a fully automated system can be developed by using image processing. This will not only reduce human efforts but will also save time and increase efficiency of fruit based industries where fruits are used in large quantities for making variousfruit products [1]. This approach can be used in import and export industry.

For detecting the quality of fruit, different morphological features are determined from the image using MATLAB and this information is used to evaluate its quality.MATLAB is a high-performance language used for methodological computing. It assimilates computation, visualization, and programming in a user friendly environment where results are expressed in mathematical notation.

Morphological features like eccentricity, orientation, major axis length, minor axis length, area, area of spots, red color intensity, blue color intensity, green color intensity, entropy, solidity, equidiameter, hue and saturation are extracted using functions of MATLAB [2] [3] [4].All the values of these features are combined together to form a 1-dimensional array which is termed as feature vector for a particular image.

K-Means Clustering is a partitioning method which partitions feature vectors into k mutually exclusive clusters [1]. This clustering methodology operates on observations and creates a single level of clusters, which makesit suitable for large amounts of data. K-Means Clustering is applied on the set of feature vectors of dataset so that fruits with similar features fall under one cluster and hence they get segregated from others.

Clusters are formed on the basis of morphological features. The quality of the fruit can be judged according to the cluster to which it belongs. To determine the cluster of the test image, the distance between the feature value of test image and the centroid of the cluster is calculated. By comparing these distances, the minimum distance among them is obtained and the nearest cluster to which the fruit belongs is determined to assess the quality of the fruit.

II. PROPOSED WORK

The methodology of fruit grading involves image acquisition, extraction of morphological features, determining the spot percentage and k-means clustering.

Figure 1 summarizes the steps involved in categorization of fruits on the basis of spot whose steps are as follows:

A. Collection of Dataset

The images of dataset should have white background with good contrast. They have fixed formats of JPG, TIF and BMP. Images which are used in the dataset are taken in good light conditions so that their quality is not degraded. There should be no noise in the image.

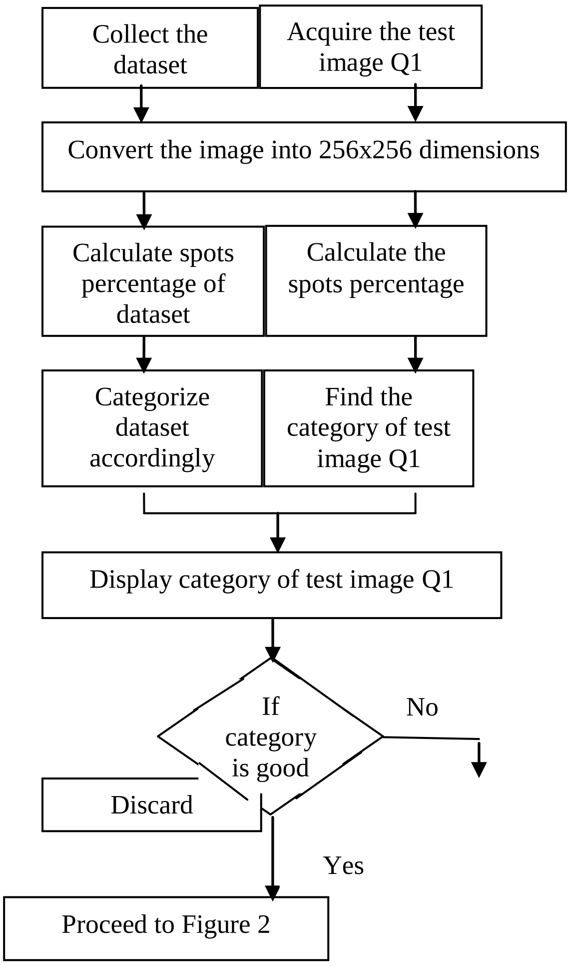

Figure 1: Categorization on the basis of spots

B. Calculation of spots percentage

Extract the red colour intensity from a RGB image and convert it into a grey scale image. The resultant grayscale image is segmented into foreground and background regions based on segmentation. The output image will be a binary image which displays the spots in the fruit. Finally, the percentage of spots is calculated on the basis of total area of fruit [8].

C. Categorization of dataset

On the basis of spots percentage, the dataset is categorized into excellent, good, average, bad and worst categories.

D. Determination of category of test image

According to the percentage of spots in the test image, it is categorized in the above mentioned categories.Now this categorization of test image can be classified further on the basis of clusters.

Figure 2 summarizes the steps involved in formation of clusters using K-means clustering for good category appleswhich are as follows:

A. Extraction of morphological features

Various extracted morphological features are eccentricity, orientation, major axis Length, Minor Axis Length, Area of fruit, Red Colour Intensity, Aspect Ratio, Solidity, Equi Diameter, HSI (Hue Saturation Intensity), HSV (Hue Saturation Value), Green Colour Intensity, Blue Colour Intensity, Entropy[7] [9]. All these features are extracted by performing various operations on the test image and dataset [2] [5] [6].

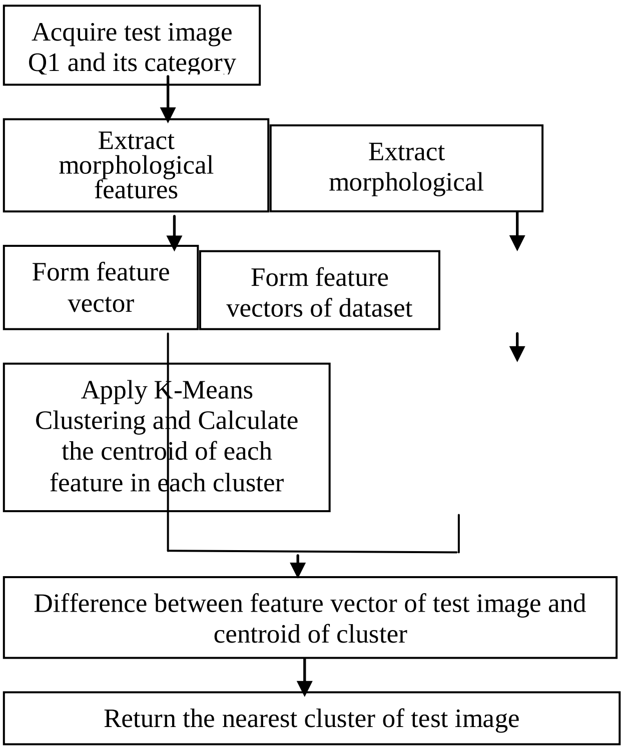

Figure 2: Determining cluster on the basis of K-Means

B. Formation of Feature Vector

After extraction of all the morphological features, their values are stored in the form of a one dimensional array which is termed as feature vector. These value of feature vector are stored in excel sheet.

C. Applying K-Means Clustering to obtain centroids

K-Means clustering is applied on all the feature vectors of the dataset to form different clusters of images which are identical to each other and different from other clusters[1]. Centroids for each feature are obtained with respect to their clusters.

D. Determination of cluster of test image

Distance between the centroid of a feature in a cluster and the value of that particular feature for test image is calculated. On the basis of minimum distance, cluster value is assigned to the test image.

III. EXPERIMENTAL WORK

A query image Q1 is taken for analysis in which first its spots percentage is calculated and according to the result its features are extracted to form a feature vector. Then this feature vector is compared with the cluster centres and the closest cluster is assigned to the query image Q1. The procedure followed for the analysis is as follows:

A. Data set

The dataset comprises of 200 images of apples, a t of which is shown in figure 3.

B. Feature set

Feature set include the features that are described in proposed work. Some of them are discussed below:

Figure 3: Dataset of apple

- *Entropy*– Entropy is anarithmeticalmeasure of arbitrariness that can be used to illustrate the texture of the input image. It is defined as:
  sum(p.*log2(p)) (1)
  wherep contains the histogram counts.
- *Area*– To calculate the area of apple, RGB image is converted into greyscale image and then object is detected in the image which is then converted to binary image, hence area is calculated [4].
- *Area of spots*– Red colour intensity is extracted from the image andthe image thus obtained is segmented into foreground (spots) and background. The output image is a binary image where the spots are white (logical true) and the background is black (logical false). The boundaries of the spot region in mask define the initial contour position used for contour evolution to segment the image.
- *HSI*– Hue, Saturation and Intensity components of a fruit is calculated by the given formula [2]:
  1/2*((R-G)+(R-B))/((R-G).^2+((R-B).*(G-B))).^0.5 (2)
- *Equidiameter*– Equidiameter of a spherical particle is similar to its diameter. To calculate the equidiameter of the fruitformula is:

$\sqrt{4 * area/\pi}$(3)

Other features like eccentricity, orientation, major axis length, minor axis length, solidity are calculated in a similar manner.

Calculation of the above mentioned features is done for the query image shown in figure 4 and the corresponding values are shown in table 1.

Figure 4: Query Image

Table 1: Feature Vector

Eccentricity	0.4890	Equidiameter	172.0732
Orientation	1.0265	Hue(HSV)	0.0459
Major Axis Length	44.4570	Saturation (HSV)	0.2743
Minor Axis Length	27.4032	Value	0.9516
Hue(HSI)	0.1519	Saturation(HIS)	0.1748
Area	23255	Intensity	0.8035
Red Color Intensity	242.5683	Green Color Intensity	191.2201
Aspect Ratio	1.6045	Blue Color Intensity	180.8911
Solidity	0.9893	Entropy	1.7310
Area of spots	3.7135e+03	Spots percentage	15.9686

C. K-Means Clustering Algorithm

The main idea of K-Means clustering is to define k centers, one for each cluster[1].

The steps to find the cluster are as follows:

Let A = {a_1, a_2, a_3,……..,a_n} be the set of feature values and B = {b_1,b_2,…….,b_c} be the set of centers.

- Arbitrarily select *'c'* cluster centers.
- Calculate the distance between each value of feature vector and cluster centers.
- Allocate the value to the cluster center whose distance from the cluster center is minimum among the cluster centers.
- Evaluate the new cluster center using:

$$v_i = (1/c_i)\sum_{j=1}^{c_i} x_i \quad (4)$$

where, *'c_i'* represents the number of features in i^{th} cluster.

- Recalculate the distance between each feature vector value and new obtained cluster centers.
- If no value was reassigned then stop, otherwise repeat from step 3).

On the basis of K-Means clustering, center of each feature in each cluster is determined and shown in figure 5. After calculating the center value of each feature, we calculate theEuclidean distance between query image and center value of corresponding feature as shown in figure 6.

	Eccentricity	Orientation	MajorAxisLength	MinorAxisLength	Area	RedColorIntensity	AspectRatio	Solidity	EquiDiameter	Hue	Saturation	Value
	Number	Number	Number	Number	Number	Number	Number	Num...	Number	Num...	Number	Nu...
1	0.4445	10.2377	27.9300	23.9109	5.9377e+05	183.7748	0.9982	0.9583	869.3817	0.3954	0.3312	0.7243
2	0.4809	14.6272	17.4493	13.2939	2.0918e+04	212.8087	1.0949	0.9497	159.5002	0.2182	0.3073	0.8372
3	0.4634	9.0781	16.6939	13.6514	6.8657e+05	187.7651	1.0568	0.9463	934.6758	0.3152	0.3673	0.7399
4	0.4762	8.6833	39.0561	31.5535	4.2619e+05	181.0497	1.0506	0.9451	736.1881	0.2550	0.3897	0.7138
5	0.4429	8.2058	25.9882	21.7446	5.2694e+05	183.3361	1.0107	0.9534	818.9152	0.3273	0.3615	0.7231

Figure 5: Centers of clusters

	Eccentricity	Orientation	MajorAxisLength	MinorAxisLength	Area	RedColorIntensity	AspectRatio	Solidity	EquiDiameter	Hue	Saturation	Value
	Number	Number	Number	Number	Number	Number	Number	Number	Number	Nu...	Number	Nu...
1	0.0708	5.1388	20.8678	18.9066	5.5485e+05	13.1822	0.0851	0.0114	646.7866	0.0273	0.2804	0.0486
2	0.0344	0.7493	10.3871	8.2896	1.7997e+04	15.8517	0.0116	0.0027	63.0949	0.2044	0.3043	0.0644
3	0.0519	6.2984	9.6318	8.6471	6.4765e+05	9.1919	0.0265	6.4372e-04	712.0807	0.1075	0.2443	0.0329
4	0.0391	6.6932	31.9939	26.5492	3.8727e+05	15.9074	0.0328	0.0019	513.5930	0.1677	0.2218	0.0591
5	0.0724	7.1707	18.9260	16.7402	4.8802e+05	13.6209	0.0727	0.0065	596.3201	0.0954	0.2501	0.0497

Figure 6: Difference from cluster centers

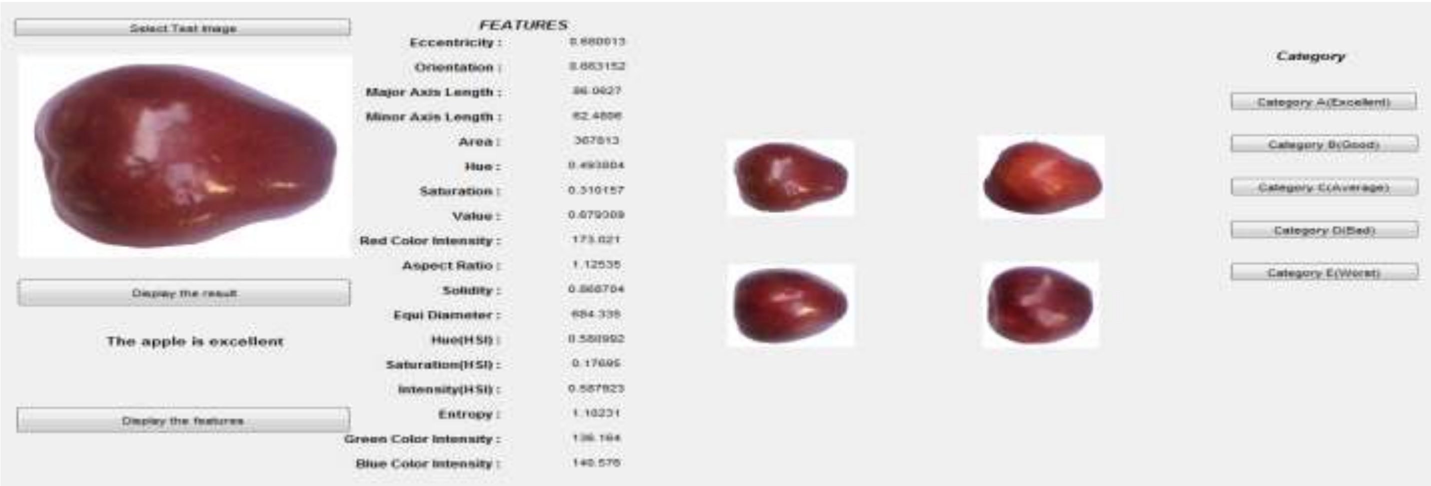

Figure 7: Analysis of apple on the basis of spots

Figure 8: Analysis of apple on the basis of spots

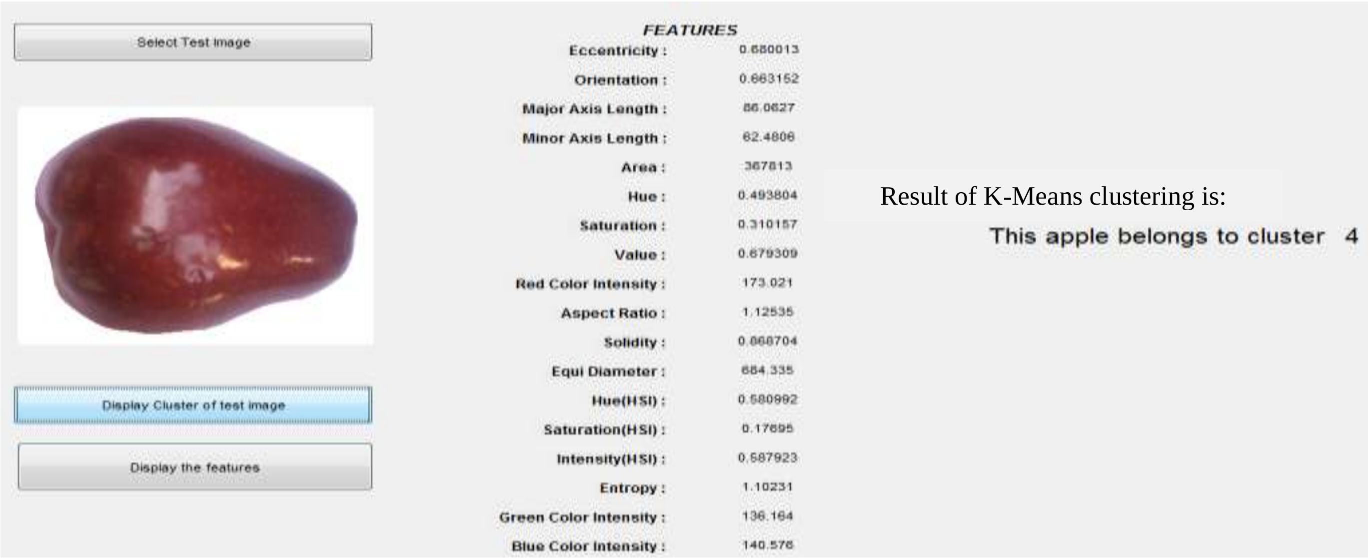

Figure 9: Analysis of apple on the basis of clustering

IV. RESULT

The result is the final step in which the outcome of the research is being displayed. The GUI is created for the analysis of spots percentage of a fruit as shown in figure 7 and figure 8. In figure 7 and figure 8, first the query image is selected by the user and then the image is categorized according to its spots percentage in 5 categories that are excellent, good, average, bad and worst. Feature vector of query image is displayed. Simultaneously, the images from dataset of similar category are displayed.

If the category of image is excellent, K-Means clustering is applied for that query image and the cluster of query image of excellent category is determined using K-Means clustering as shown in figure 9.

V. CONCLUSION AND FUTURE WORK

This paper presents integrated techniques for grading of fruits on the basis of spots and various morphological features. The quality of different fruits can be assessed using this approach. Future work includes implementation of a fully automated system which consists of conveyor belt, camera, IR sensors and LCD.

REFERENCES

[1] Shiv Ram Dubey, Pushkar Dixit, Nishant Singh, Jay Prakash Gupta, "Infected Fruit Part Detection using K-Means Clustering Segmentation Technique", International Journal of Artificial Intelligence and Interactive Multimedia, Vol. 2, Nº 2.

[2] Seema, A Kumar, G.S. Gill, "Automatic Fruit Grading and Classification System Using Computer Vision: A Review", 2015 Second International Conference on Advances in Computing and Communication Engineering, 978-1-4799-1734-1/15$31.00©2015 IEEE DOI 10.1109/ICACCE.2015.15.

[3] Mohana S.H., Prabhakar C.J., and Praveen Kumar P.U., "Surface Defect Detection and Grading of Apples", Proc. of Int. Conf. on Multimedia Processing, Communication& Info. Tech., MPCIT, DOI: 03.AETS.2013.4.33 © Association of Computer Electronics and Electrical Engineers, 2013.

[4] Manoj B. Avhad, Satish M. Turkane, "ARM Based Fruit Grading and Management System Using Image Processing ",ISSN: 2278 – 1323 International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 1, January 2013

[5] Alok Mishra, Pallavi Asthana, Pooja Khanna,"The Quality Identification of Fruits in Image Processing using MATLAB", IJRET: International Journal of Research in Engineering and Technology eISSN: 2319-1163 | pISSN: 2321-7308, Volume: 03 Special Issue: 10 | NCCOTII 2014 | Jun-2014.

[6] Slamet Riyadi, Ashrani A. Abd. Rahni, Mohd. Marzuki Mustafa, and Aini Hussain, "Shape Characteristics Analysis for Papaya Size Classification", The 5th Student Conference on Research and Development –SCOReD 2007 11-12 December 2007, Malaysia, 1-4244-1470-9/07/$25.00 ©2007 IEEE.

[7] P.Deepa, Dr.S.N.Geethalakshmi,"A ComparativeAnalysis of Feature Extraction Methods for Fruit Grading Classifications", International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS) IJETCAS 13-138; © 2013.

[8] Dr.Vilas D. Sadegaonkar, Mr.Kiran H.Wagh," Quality Inspection and Grading Of Mangoes by Computer Vision & Image Analysis", Int. Journal of Engineering Research and Applications, ISSN : 2248-9622, Vol. 3, Issue 5, Sep-Oct 2013, pp.1208-1212

[9] Dr. Vilas D. Sadegaonkar, Kiran H. Wagh,"Improving Quality of Apple Using Computer Vision & Image Processing Based Grading System", International Journal of Science and Research (IJSR),Paper ID: SUB152911 Volume 4 Issue 4, April 2015