K-mean Clustering for Segmentation of Irregular Shape Fruit Images under Various Illumination

Miss Monali R. Dahapute1
TGPCET, Nagpur
dahapute.monali@live.com

Mr. Amit welekar2
TGPCET, Nagpur
welekar.amit@gmail.com

***Abstract*—** Segmentation is the first step in analyzing or interpreting an image automatically. In particular applications, like image compression or image recognition, entire image can't be processed directly. Hence many segmentation techniques are proposed to segment an image before processing it. This made it possible to develop many techniques which are currently using in different industries and agriculture field. They are either applied for grading or inspecting quality of food products and Fruits. These developed techniques use thresholding and clustering approach to get proper segmented output. In this paper an image segmentation approach is developed based on k-means adaptive clustering. This approach segments the various shape fruit images particularly which are non-circular (like banana, mango, and pineapple) and captured in various illumination such as low, Medium and high intensity. Earlier segmentation methods were not apposite for fruit images captured in natural light; as they were responsive to various colour intensity predisposed by the sunlight illumination. Natural illumination tempts an uneven amount of light intensity on the surface of the object, resulting in poor quality image segmentation. This approach will deal with problem of light effect. K-means clustering is renowned method for image segmentation. This method is more efficient, robust than the others. It provides best result when dataset is well separated and distinct. Different shape fruit images are segmented properly along with grey scale. The analytical results are the evidence for the accurate segmentation of banana, mango pineapple using new approach developed here.

***Index Terms*—** *Color mapping, K-means Clustering, Fruit Images, Segmentation.*

*****

I. INTRODUCTION

Image segmentation is a process that groups together pixels that have similar attributes. Segmentation process splits an image into various regions, and provides all the relevant data required for analyzing and grading of a digital image. Finally it separate out the unwanted background from the digital image and shows the object of interest. Image segmentation process is a crucial stage in the image analysis where the results in this stage persuade the concert of the entire process. Segmenting the selected region guarantees that only the object of interest was processed during the analysis phase. However, the segmentation process has happen to a challenging issue because of the complex background and changeable illumination on the images. Therefore it is significant to have an efficient segmentation technique that will be able to divide an image into foreground and background, accurately and effectively in presence of natural illumination. Image segmentation has become an important step to analyze the various objects in automated manner. It provides vast application in the field of agriculture, industries, medical, and in recognition tasks. Image segmentation is most useful in agriculture specifically in detecting the infected fruit part [4], in classifying the fruits [5][7][11][14], in fruit quality management [8][9], for on plant fruit detection[10][17] and for estimating the crop [21] also segmentation is used in fruit maturity recognition [13][20]. Quality control is most important process in agriculture and food industries to grade them and to export the quality fruits. Manual examination of agricultural products is tiresome and indistinct, so it is necessary to replace this manual method with a computerized one.

There are many segmentation techniques available those are thresholding [5, 11, 14, and 22], clustering, FCM [3, 4, 8, 13, and 19] and [12]. An effective approach for image segmentation includes tools, and a comprehensive environment for data analysis, visualization, and an algorithm development. The researchers had given an extensive interest on threshold based segmentation. Thresholding partition the background and foreground for images having high contrast levels whereas clustering based approach separate out groups of objects. Among all clustering k-means is most renowned technique because it is very simple and can easily classify dataset or pixels throughout the certain number of clusters. Apart from these enhance segmentation techniques developed; they are not succeeding on images which are captured under natural illumination.

Segmentation methods can be classify as thresholding [1, 2, 3, 4], Color based segmentation [5] i.e. K-means clustering [6, 7, 8, 9, 10] and FCM Clustering [11], Region based segmentation [12], Edge based segmentation [13] i.e. watershed algorithm [14], Texture based segmentation i.e. texture filter [15]. Some are derived from above methods [12, 16 and 17] to attain more precise segmentation and to get rid of the drawbacks of existing methods. These methods are discussed here,

I. Thresholding Method – Otsu Method

Otsu's thresholding iterates through all possible threshold values and then it calculates a spread for the pixel levels on both sides of the threshold value, i.e. for the pixels that either falls in background or foreground. This is an efficient method because it operates directly on the gray level histogram.

In "Shape Characteristics Analysis for Papaya Size Classification" [3], thresholding is classifying grayscale pixels into two categories, for foreground and for background resulting in a binary image; since intensity values are different in each image, but global threshold value of images could not perform precise segmentation. In "Segmentation Of Natural Images Using An Improved Thresholding-Based Technique" [4], Otsu method is used to calculate threshold value automatically, and thus researchers are able to extract objects of interest from its background. "Quality Analysis and Classification of Bananas" [2], used thresholding for segmentation of bananas pictures as this is a simple method and also thresholding is more suitable for such type of segmentation.

II. Color based Segmentation – Data Clustering

The color based segmentation uses a centroid to represent each cluster and it classifies based on the similarity with the centroid of the cluster. Papers [6,7,8] used color based K-means clustering to identify defect in fruits using segmentation, as this is more acceptable method for images containing blur boundary. In "Segmentation of apple color images utilizing fuzzy clustering algorithms" [11], the author is using FCM clustering for segmenting an image of apple. K-means and FCM are very sensitive to find out initial cluster values and this may produce different segmentation results for those images captured in natural illumination. In "Adaptive K-Means Method for Segmenting Images under Natural Environment" [10], an improved clustering based segmentation is used i.e. adaptive K-means, which may overcome limitation of K-means and FCM. An approach used in adaptive K-mean clustering is to divide an image into sub-images and further these sub-images are segmented individually on the basis of local intensity values.

III. Region based Segmentation

Region-based segmentation mainly assumes that, neighboring pixels within one region have similar value; this is basically pixel based segmentation. The procedure regarding region based segmentation is to compare one pixel with its neighbors and if found identical, pixel can be set belong to the cluster as one or more of its neighbors.

In "On Plant Detection of Intact Tomato Fruits Using Image Analysis and Machine Learning Methods" [12], a method is developed which exactly detects individual intact tomato in mature, immature and at young stage on plant. This method comprised pixel based image segmentation to make a GUI application with which a color pixel and its 8-neighbour pixels are automatically extracted with a label such as "fruit", "leaf", "stem" and "background".

IV. Edge based Segmentation

This method is generally based on edge in an image, which distinguishes object of interest from its background. Some techniques include edge detection method like gradient operators and Hilbert transform, and some are using watershed algorithm using concept of edge. In watershed algorithm, waterlines are found to separate out the distinct regions.

In "Image Segmentation and Maturity Recognition Algorithm Based on Color Features of Lingwu Long Jujube" [14], images are segmented based on features of Lingwu long jujubes; like shapes and colors. Based on color fusion technology, they have abstracted the red and non-red areas by summering the areas. As there would be the problem of adhesion and occlusion phenomena, watershed transformation was used to combine with distance and gradient algorithm; which deals with over-segmentation and under-segmentation. Marker approach which connects component belonging to image, is used for resolving the problem of over-segmentation occurred in the watershed algorithm. It consists of two markers internal and external, for area belongs to the object of interest and background respectively. The marker selection is based on pre-processing, in which image is filtered using smoothing filter; this filtering is useful in reducing the over segmentation, occur in watershed algorithm as shown in figure 2[18]. In "A Novel Approach to Image Segmentation" [13], image segmentation is used which is based on detecting the edges and performed some steps over the edge detected.

Figure 2: The filtering process in Marker method

V. Texture Method

For, segmenting highly textures images, Grey level or color pixel values are not sufficient; a spatial property of texture is used to characterize group of pixels. A local measure of texture is computed over a neighborhood. In "Estimation Of Mango Crop Yield Using Image Analysis - Segmentation Method" [15], A.B. Payne and K.B. Walsh brought an approach to count mangos of a tree. For this they had captured images for period of three days. These images were then segmented into fruit pixels and background pixels, using color segmentation in color ranges of RGB and YCbCr; further they

have performed texture segmentation, based on adjacent pixel variability.

VI. Improved Algorithm

In last many years, several segmentation approaches are developed to get accurate segmentation of fruit images. These methods are not widely applied for fruit images captured under natural illumination; as some of the colours of fruit are very sensitive to the variation in colour intensity influenced by sunlight, and hence some improved approaches are brought to overcome this limitation.

In "Identification of red apples in field environment with over the row machine vision system" [17], a machine vision system which identifies red apples in RGB at outdoor environment in clusters is developed. In addition to the clustering algorithm, this approach is a fusion of blob analysis and Circular Hough Transform (CHT) to identify apples in over the row sensor platform. "A Rule-Based Segmentation Method For Fruit Images Under Natural Illumination" [16], brought an improved clustering based approach TsNKm, which is combination of an improved Thresholding and Adaptive K-means; this algorithm works only on circular shaped fruit images captured under natural illumination.

This paper is systematized as follows. Part II describes the proposed approach for image segmentation and part III explains methods and material utilized for this research. Part IV discusses about recital of the segmentation carried out on different fruits with various illumination. Towards the end part V puts about the conclusion and prospects.

II. Modified K-means Clustering Based Segmentation Method

In this paper we have developed the segmentation approach namely GsKm which uses the K-means adaptive clustering for different cluster generation as per cluster index. In clustering technique the objects of different set classify a given a data set which divides into clusters in order that the data in each cluster shares some common defined distance measurement. K-means clustering is an unsubstantiated clustering method in which the input data objects are classified into multiple classes on the basis of their intrinsic distance from each other. The GsKm segmentation approach involves Grayscale Segmentation and K-means adaptive clustering. The flowchart of GsKm is shown in Fig

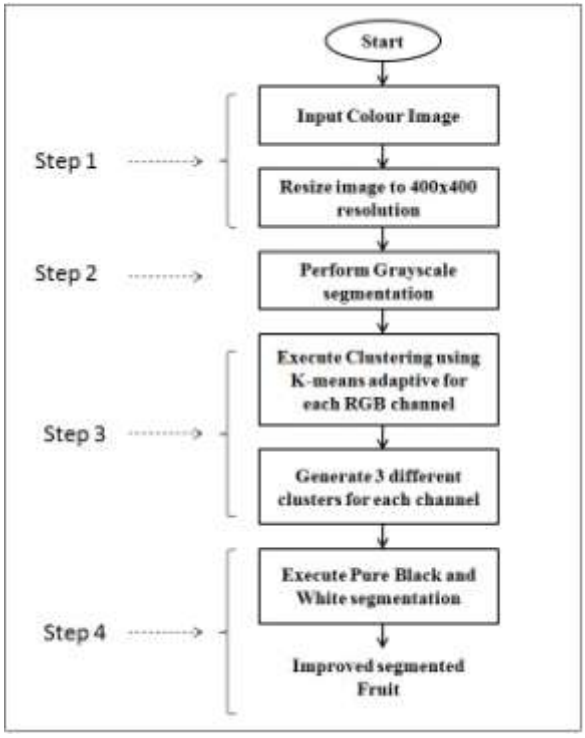


Fig Flowchart of GsKm

Step 1 refers to resizing the input colour image to 400 x 400 resolutions to fit image evenly. This is required to keep the even intensity of all pixel values which will prevent the consumption of time.

Step 2 refers to perform the grayscale segmentation, which convert the RGB image to grayscale.

Step 3 refers to convert the grayscale image into binary image. Execute clustering using K-means adaptive for each RGB channel. Generate 3 different clusters for each channel.

Step 4 refers to execute pure black and white segmentation.

The GsKm produced a good segmented image.

III. Methods and Material

In this section methods and material are discussed which are utilized for developing an approach to segment the fruit images with an intension to provide better recital. The segmentation methods make the use of three important steps to segment an image. Figure 1 shows system flow of image segmentation.

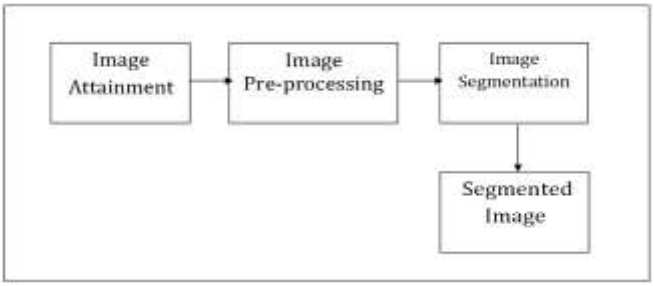
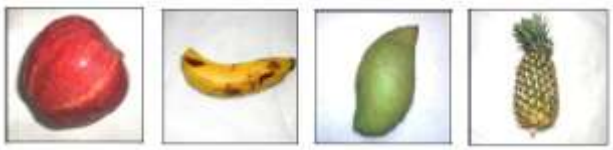


Fig.1 System flow of Image Segmentation

A. Image Attainment

For this approach good quality images of four fruits i.e. apple banana, mango, and pineapple are captured. Each fruit represents different shape along with different colour. Apple is of red colour, banana is of yellow, mango as green and pineapple is of greenish yellow colour. A digital camera with high resolution has been used. These fruit images are captured under the Natural illumination condition to get the realistic data. The RGB values of these images are stored.



a) Red Apple b) yellow Banana c) Green Mango d) Greenish Yellow Pineapple

Fig: Original Fruit images

B. Pre-processing of Image

All the RGB values of the fruit images captured are resized to 400x400 pixels. This is required to keep the even intensity of all pixel values which will prevent the consumption of time. Then all these images are converted to grayscale.

C. Segmentation of Image

In this step, Segmentation of various shape Fruit images has divided the image into two part one containing unwanted background and one which is required segmented image of exact fruit. This is an important step is performed by GsKm technique. This technique has segmented the fruit images with great accuracy

IV. Results and Discussion

This section discusses the results produced by GsKm. In this paper fruit image segmentation is carried out using Grayscale segmentation and K-means adaptive and black and white segmentation. The four fruit images are captured under natural illumination and fruits considered in this paper are apple, banana, mango and pineapple. These fruits show different shape with different color. Various segmentation approaches are developed which were able to segment the circular fruit images like orange lemon etc, but these developed methods were not that accurate to segment irregular shape fruit. In case of segmentation of irregular shape images the main problem took place while edge detection. When those images are captured under natural illumination Natural illumination tempt unequal amount of light intensity on the surface of the fruits, which results in low quality image segmentation. Hence GsKm method is developed which is able to segment the different shape fruit images captured under natural illumination more accurately.

Fig. 4 Segmented images using various techniques

Fig. 5 Final Output

V. Conclusion

An improved method which can segment various shape fruit images captured under the natural illumination. Image grayscale image segmentation and colour classification are very important steps to be carried out. This paper has been developed an approach using K-means adaptive clustering which form different cluster and colour classification for segmentation which can execute both the steps accurately and can provide the required segmentation of fruits.

The various segmentation techniques such as Otsu, K-

mean, and FCM are used in the different paper studied. They were able to segment an image in various conditions with some pros and cons, but they made it possible to reach more close to the accuracy. Each segmentation method and its approach for segmentation are restricted for a particular fruit only along with the provision for illumination. Some of the improved and hybrid approaches had overcome the illumination effect but effort must be taken to develop an approach which would be able to segment the images of various irregular shape fruit in natural illumination accurately. To develop an improved method, the hybrid approach using K-means clustering, improved thresholding and adaptive k-means can be used for getting the required segmentation.

References

[1] Miss Monali R. Dahapute and Mr. Amit Welekar, "A Review: Segmentation Approaches for fruit Images " 2nd International Conference Technosoft 16, TGPCET, Nagpur, Special Issue, ISSN : 2454-1958, Jan 2016

[2] Miss Monali R. Dahapute and Mr. Amit Welekar, "Improved K-means Clustering with colour Classification for Segmentation of Fruit images" International Journal of Research, Volume 3, Issue 5, ISSN : 2348-6848, March 2016

[3] Manjinder Kaur1, Navjot Kaur2, Harkamaldeep Singh, "Adaptive K-Means Clustering Techniques For Data Clustering" International Journal of Innovative Research in Science, Engineering and Technology, ISSN: 2319-8753, Vol. 3, Issue 9, September 2014

[4] Shiv Ram Dubey, Pushkar Dixit, Nishant Singh, Jay Prakash Gupta4 1glau, Mathura, India "Infected Fruit Part Detection Using K-Means Clustering Segmentation Technique", International Journal Of Artificial Intelligence And Interactive Multimedia, Vol. 2, Nº 2. Doi: 0.9781 / Ijimai. 2013

[5] D Surya Prabha and J Satheesh Kumar, "A Study on Image Processing Methods for Fruit Classification" roc. Int. Conf. on Computational Intelligence and Information Technology, CIIT, 2012

[6] Prabhishek Singh, Ramneet Singh Chadha Computer Sc. & Engg, , CDAC Noida India. "A Novel Approach to Image Segmentation", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 4, April 2013

[7] Guo Feng; Cao Qixin and Nagata Masateru, "Fruit Detachment and Classification Method for Strawberry Harvesting Robot", International Journal of Advanced Robotic Systems, Vol. 5, No. 1,ISSN 1729-8806, pp. 41-48, 2008

[8] J.Ramprabhu1, S.Nandhini2,,"Embedded System Based Fruit Quality Management Using PIC Micro Controller" International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 3 Issue I, ISSN: 2321-9653, January 2015

[9] Singh Dhillon, Er. Ashok Kumar Bathla, "Detecting Guava Quality Using Gradient Function Histogram Plotting " International Journal of Engineering and Technical Research (IJETR) ISSN: 2321-0869, Volume-2, Issue-9, September 2014

[10] Kyosuke Yamamoto 1, Wei Guo 1, Yosuke Yoshioka 2 and Seishi Ninomiya, "On Plant Detection of Intact Tomato Fruits Using Image Analysis and Machine Learning Methods" Sensors 2014, 14, 12191-12206; doi:10.3390/s140712191, July 2014

[11] Sanjay Chaudhary*, Bhavesh Prajapati, "Quality Analysis and Classification of Bananas" International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 1, ISSN: 2277 128X, January 2014

[12] Ghabousian, A., & Shamsi, M., "Segmentation of apple color images utilizing fuzzy clustering algorithms," Advances in Digital Multimedia, vol. 1(1), pp. 59-63, 2012

[13] D. I. Amarasinghe and D. U. J. Sonnadara, "Surface color variation of Papaya fruits with maturity", Proceedings of the Technical Sessions, 25 , 21-28,Institute of Physics – Sri Lanka, 2009

[14] 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, Malaysia, 1-4244-1470-9/07 IEEE. 11-12 December 2007

[15] Leemans, V., magein, H. & Destain, M.F. On-line fruit grading according to their external quality using machine vision. Biosystem Engineering 83(4): 397 – 404, 2002.

[16] Hamirul'Aini Hambali, Hazaruddinharun, University Utara Malaysia ,Sharifah Lailee Syed Abdullah, Nursuriati Jamil, University technology Mara, "A Rule-Based Segmentation Method For Fruit Images Under Natural Illumination", Ieee ,978-1-4799-4575-7/14/ 2014

[17] Abhisesh Silwal, Aleana Gongal, Manoj Karkee., "Identification of red apples in field environment with over the row machine vision system", Agric Eng Int: CIGR Journal Vol. 16, No. 4, December, 2014

[18] Tomas U. Ganiron Jr. International Association of Engineers (IAENG)," Size Properties of Mangoes using Image Analysis" International Journal of Bio-Science and Bio-Technology Vol.6, No.2 (2014), pp.31-42 , 2014.

[19] Sharifah Lailee Syed Abdullah1, Hamirul'aini Hambali2, And Nursuriati Jamil3 Sarawak, Malaysia. Universiti Utara Malaysia "Adaptive K-Means Method For Segmenting Images Under Natural Environment" Proceedings Of The 4th International Conference On Computing And Informatics, Icoci 2013, 28-30 August, 2013.

[20] Yutan Wang, Jiangming Kan, Wenbin Li And Chuandong Zhan, "Image Segmentation And Maturity Recognition Algorithm Based On Color Features Of Lingwu Long Jujube", Advance Journal Of Food Science And Technology 5(12): 1625-1631, Issn: 2042-4868; E-Issn: 2042-4876, 2013.

[21] Payne, A. B., Walsh, K. B., Subedi, P. P., & Jarvis, D., "Estimation Of Mango Crop Yield Using Image Analysis - Segmentation Method," Computers And Electronics In Agriculture, Vol, 91, Pp. 57-64, 2013.

[22] Sharifah Lailee Syed Abdullaha, Hamirul'aini Hambalia,B*, Nursuriati "Segmentation Of Natural Images Using An Improved Thresholding-Based Technique" Jamilc International Symposium On Robotics And Intelligent Sensors 2012 (Iris 2012)