Abstract

Potato grading is one of the evolving areas of research due to infeasibility of various algorithms and designs developed to work in real time industrial environments. Various problems which arrive in real time environment include shadow in images, high complexity and inefficient speed of algorithms. The fuzzy c-mean clustering based algorithms can be efficiently exploited for potato grading due to their low complexity. So this paper introduces a potato defect detection method using fuzzy c-mean clustering based segmentation. The Euclidean Distance based on RGB color space is used to segment the potato image into defected and healthy areas. An anti-shadow wings model is introduced for the process of image acquisition. The potato is graded on the basis of defects like greening, cracks and rotten. The potatoes used to prepare the database were taken from local vegetable markets of India. The algorithm segments the potato images on the basis of RGB color values of the image pixels. The experimental results show that Fuzzy c-mean clustering is very effective to segment the potato images into defected and healthy areas. The novelty of the work is represented by novel wings design to remove the shadow in the images without using any pre-processing algorithm. Along with this modified Fuzzy c-mean algorithm is efficiently used first time to grade the potatoes successfully. Combination of both image acquisition design and algorithm proves to be a novel feature to grade potatoes successfully.

Keywords: Defect Detection, Fuzzy C-Mean Clustering, Potato, RGB Color Space, Segmentation

1. Introduction

Potato being one the popular crop in the world is widely consumed in the form of raw or processed. Almost 80% of the countries are involved in potato cultivation and worldwide production exceeds 300 million ton per year. There is large demand of potato for making processed products rather that raw usage. Thus high quality potato is raising demand of the market. Potato grading is mostly done on the basis of its external appearance, which involves its shape, color, texture and various external defects. With the rising quality demands of potato market, grading and sorting of potatoes become an essential asset to increase the revenue generated by in. External appearance of potato has large dependency on environmental conditions and crop handling. Traditionally the grading is done by the trained human inspectors who grade the potatoes by seeing and feeling the product. But using human inspectors for grading is an inconsistent, expensive and inefficient method. Due to downfall in the availability of labor manual grading process is infeasible sometimes. So industry demands an accurate, fast and efficient method for potato grading. As grading process is largely dependent on the color, shape and texture of the potato skin, machine vision systems can be efficiently used for potato grading. Machine vision systems could match the high accuracy and speed demands of the current industry. As the demand for potato is rising day by day there is need for good quality batches of potatoes. This is not only necessary for industrial requirements, but also for raw potatoes sale in the local market. All the gradation of potato is done on the basis of its external quality. There is strong need to develop standardized parameters for grading potatoes. External quality parameters of potato...
Potato Defect Detection using Fuzzy C-mean Clustering based Segmentation

Indian Journal of Science and Technology

Toes can be easily exploited by the machine vision system to automate the grading process. The various surface defects which are recognized to grade the potatoes are size distribution, surface roughness, tuber germination, green spots, bruises and tuber disease due to viral, fungal or bacterial infection. Size parameter plays an important role in gradation of potatoes. Potatoes having abnormal shape and size lead to huge loss during the process of peeling. Common Scab and Silver Scurf are common examples of surface roughness. Bad storage environments lead to tuber germination, potatoes which are exposed to air during the growth process lead to potato greening. Bruises are color marks which even remain after the peeling and are most responsible factor affecting potato loses. So which such variance found in the potato defects there is need to develop an effective machine vision system which could effectively grade the potato crop. In developed an image processing algorithm to grade the potatoes without simulation. Basically the potatoes were separated from clods without simulation of potatoes. The grading of potatoes was based on the clustering. The clusters of potatoes and clods were formed which were then exploited to grade the potatoes. The proposed scheme proves to be very successful. In another potato sorting algorithm based on shape sorting. This algorithm sorts irregular potatoes using machine vision system very fast and accurately. In this author used eight shape parameters extracted from size features and Fourier transform. Linear discriminant analysis was used to select most important parameters which were most effective to sort regular and irregular potatoes. In this 228 potatoes were used to test the algorithm which provides 96.5% accuracy. In the hierarchical potato grading method. In this author developed machine vision system for potato defect detection and size sorting. In this classification algorithms were used to solve the mathematical equations on the basis of which potato grading was done. In this color based classification was used and proposed system proves to be accurate. In suggested a new mechanism based on helix with variable distance for grading potatoes. Effective parameters on the gradation helix were found after various attempts. Experimental result shows this scheme to be successful. In a genetic algorithm for feature selection to classify potatoes. Developed system classifies potatoes on the basis of external defects. Potato image is segmented and given to classifier which tells which segment is defected or healthy. An ad-hoc genetic algorithm is developed which chooses those features which enhance the performance of classifier. In developed novel potato blemish detection algorithm using machine vision. In this pixel wise classifier is used to detect the blemish in potatoes. The features extracted from the segmented image are provided to the classifier on the basis of which blemish is detected. An adapted boosting algorithm is used to select the best suited features from to total feature set. Experimental results show that the proposed system achieved 89.6% accuracy. In developed a high-speed machine vision system grading and quality inspection of potatoes. The machine vision system was able to grade potatoes on the basis of size, shape and external defects like rizoctonia, silver scab, greening, common scab, cracks, growth cracks and mechanical damages. Linear Discriminant Analysis along with Mahalanobis distance classifier was used to classify the pixels. Fourier based shape classification technique was used for detecting misshapen potatoes also area, central moments and eccentricity features were used to differentiate between similar colored defects. In the field of potato grading most of the work done till now is based upon classification. Various classifiers are used to develop a machine vision system but there is yet need of an accurate machine vision system which could be feasible to work in real time environments. So we used fuzzy c-mean segmentation algorithm to segment the potatoes into defected and healthy areas, on the basis of this segmentation potatoes can be classified as defected or healthy. In provided modified Euclidian Distance based on the RGB color images and incorporated it into existing fuzzy c-mean segmentation algorithm. Euclidian Distance based on RGB color images was successfully used to segment the defected areas of potato image in our work.

2. Materials and Methods

2.1 Image Acquisition

Image acquisition phase is one of the most important steps involved in designing machine vision system for potato grading. The design of this step tells weather the system is feasible in the industrial environments or not. For enhancing the performance of image segmentation algorithm we have proposed a novel wings design for image acquisition. The image acquisition system consist of color camera (Nikon Coolpix S3100), four 3 watt natural light LED bulbs, white base, four wings to remove the uneven lightness and shadow of potato. Color camera is
fixed above the sample and distance between sample and camera is 45 cm. Four LED bulbs are placed at the center of four walls of white square box. Four wings are placed in front of the bulbs at 5 cm distance. Potato sample is placed between the wings. The proposed image acquisition system is shown in Figure 1. The software system used for this process was Visual Studio 2012 and language used to develop the algorithm was C#.

2.2 Ground Truth
The potatoes used for algorithm development and testing were purchased from various agriculture fields of Punjab where potatoes were grown. Total 100 potato samples were used for testing the algorithm, which contain healthy, cracked and rotten potatoes. After initially marked as healthy or defected the potato images were used to get the results and at last results were compared with ground truth data.

2.3 Basic Algorithm
The Fuzzy C-Means clustering (FCM) algorithm \(12\) is one of the popular image clustering algorithms. It is very sensitive to noise as it is based on iterations which compare the properties of each individual pixel to each cluster defined in the feature domain. It is based on objective function and can be obtained by minimizing it. The objective function is given in Equation (1):

\[
A = \sum_{i=1}^{n} \sum_{k=1}^{c} \mu_{ik} |p_i - v_k|^2
\]  

(1)

Where:
- \(A\) is the objective function.
- \(c\) is the number of clusters.
- \(n\) is the number of pixels in the image \(E\).
- \(\mu\) is a fuzziness factor (a value \(> 1\)).
- \(\mu_{ik}\) is the fuzzy membership value from table.
- \(v_k\) is the centroid of the \(k\)th cluster.
- \(p_i\) is the \(i\)th pixel in \(E\).

Figure 1.

\(|p_i - v_k|\) is the Euclidean distance between \(p_i\) and \(v_k\) given by Equation (2):

\[
|p_i - v_k| = \sqrt{\sum_{i=1}^{n} (p_i - v_k)^2}
\]  

(2)

The fuzzy C-mean segmentation has been successfully used to segment the images in the fields of medical, geological and even for satellite images. We proposed a system which uses modified fuzzy C-mean segmentation for segmenting the potato images into healthy and defected areas.

2.4 Modified Fuzzy C-Mean Algorithm
There has been many modifications done by various researchers in basic Fuzzy C-mean algorithm since it has been developed. But all the changes are not effective to segment the images for food quality evaluation. Since the major factor effecting the segmentation of food images is color, so we have used the modification done by Equation (1) to segment the potato images. The algorithm has been slightly modified to decrease the complexity of algorithm and make it feasible to grade the potato in the real time environments. The fuzzy C-mean algorithm is almost similar in working with the K-mean algorithm. In both the pixel value is compared with the cluster center to determine whether pixel belongs to particular cluster or not. Only difference is that in the fuzzy C-mean algorithm instead of making hard decisions about pixel belongs to which cluster, it assigns a value between 0 and 1. This value describes the strength with which pixel belongs to that cluster. According to fuzzy rule the sum of membership value of pixel to all clusters must be 1. The higher is the membership value for particular cluster, the pixel more likely falls into that cluster.

The objective function used for proposed technique is given in Equation (3):

\[
A = \sum_{i=1}^{n} \sum_{k=1}^{c} \mu_{ik} |p_i - v_k|^2
\]  

(3)

Where:
- \(A\) is the objective function.
- \(c\) is the number of clusters.
- \(n\) is the number of pixels in the image \(E\).
- \(\mu\) is a fuzziness factor (a value \(> 1\)).
- \(\mu_{ik}\) is the fuzzy membership value from table.
- \(v_k\) is the centroid of the \(k\)th cluster.
- \(p_i\) is the \(i\)th pixel in \(E\).
|pi – vk| is the Euclidean distance between pi and vk.

Equation (4) is used to calculate the centroid of kth cluster

$$v_k = \frac{\sum_{i=1}^{n} \mu_{ik}^m p_i}{\sum_{i=1}^{n} \mu_{ik}^m}$$  \hspace{1cm} (4)

The fuzzy membership is obtained by using the Equation (5):

$$\mu_{ik} = \frac{1}{\sum_{i=1}^{n} (\frac{|p_i - v_k|}{|p_i - v_l|})^{\frac{m}{2}}}$$  \hspace{1cm} (5)

To make the algorithm feasible for segmenting the potato images we have used extended algorithm which performs clustering of color images in the RGB color space Equation (1). So the Equation (2), which is given to compute the Euclidean distance between pi and vk is modified to include RGB colors. The modified Euclidean distance is given by Equation (6):

$$|p_i - v_k| = \sqrt{\sum_{i=1}^{n} (p_{ir} - v_{kr})^2 + (p_{ig} - v_{kg})^2 + (p_{ib} - v_{kb})^2}$$  \hspace{1cm} (6)

Pseudo code is given as follows:

1. Initially the number c of clusters, the stopping condition ε and the fuzzy parameter m are set.
2. Then initialize the fuzzy partition matrix.
3. Loop counter is set to b = 0
4. Cluster centroids are calculated using Equation (4)
5. Membership values of each cluster are calculated for each pixel using Equation (5).
6. Objective function A is calculated. If between consecutive iterations value of A is less than ε then stop, else set b = b+1 and go to Step 4.
7. Defuzzification and segmentation.

The algorithm forms the clusters of defect and healthy parts of potato, which can be exploited to grade the potato successfully.

### 3. Results and Discussion

The proposed algorithm successfully identified the defected and healthy areas of potato. In the proposed technique the main focus was given on the image acquisition phase. The acquisition system was novel wing design which removes uneven lightness and shadow, due to this there is no need of image pre-processing step. The complexity of the grading system was decrease by removing the image pre-processing step. The proposed algorithm was successful in grading potatoes without any image pre-processing. Another factor was using color based Euclidian distance in the Fuzzy C-mean segmentation algorithm. As all the external defects in potato largely vary the color values of pixels, so this feature was exploited to segment the healthy areas from defected ones. The proposed system was tested and validated using ground truth data. Total 100 potato images were used to test the algorithm. Samples used for experimental work consist of cracked, rotten and healthy potatoes. The sample was collected from various agriculture fields of Punjab, India. Experimental results are displayed in Table 1. The proposed system provides 95% accuracy. Also proposed technique was able to detect the defected potatoes with 100% accuracy although few healthy potatoes were misclassified as defected. The misclassification was due to variability in crop as well as human errors. The potatoes were washed

**Table 1.** Experimental results and analysis of proposed technique

| Defect Type | Total number of potatoes | Number of potatoes misclassified | Number of potatoes correctly classified | Percentage (%) |
|-------------|--------------------------|---------------------------------|---------------------------------------|----------------|
| Rotten      | 30                       | 0                               | 30                                    | 100            |
| Cracks      | 40                       | 0                               | 40                                    | 100            |
| Healthy     | 30                       | 5                               | 25                                    | 83.33          |
| Total       | 100                      | 5                               | 95                                    | 95             |

![Figure 2](image1.png) **Figure 2.** Left image displays original healthy potato image. On the right is segmented healthy potato image.

![Figure 3](image2.png) **Figure 3.** Left image displays original potato image with green areas. On the right are segmented green areas of potato image.
before grading and were labeled before preparing ground truth data. Figure 2 shows the original image of a healthy potato on the left and on the right result after segmentation of healthy potato image is shown. Figure 3 and Figure 4 shows original potato image with green areas on the left side and on the right side green area is segmented successfully by the algorithm. Figure 5 on the left shows original potato image with green and rotten areas and right image displays defected potato image after segmentation. Figure 6 on the left original potato image with rotten areas is displayed and on the right shows segmented image. Figure 7 on the left shows original potato image with cracks and cracks are segmented successfully in the right image.

4. Conclusion

Provided results show that modified fuzzy C-mean segmentation algorithm (1) is successful in segmenting potato image into the healthy and defected areas. Also the proposed scheme was very successful to detect the cracks, rotten and green areas in the potato image. The modified Euclidean Distance was very successful in segmenting the defected areas from healthy areas in the potato images. Although some of the healthy areas were wrongly detected as defected due to which some healthy potatoes were classified as defected. Also novel wings design used for image acquisition proves to be very successful. There was no disturbance found in the process of segmentation due to uneven lightness. Due to this design there was no need to incorporate the image pre-processing step. So the complexity of the algorithm was decreased to great extent. Also the LED bulbs used prove the feasibility of the proposed scheme to work successfully in the industrial environments. The proposed scheme provides 95% accuracy in grading the potato on the basis of three basic external defects i.e., greening, rotten and cracks. All three defects were detected with 100% accuracy but some of the healthy potatoes were misclassified as defected. For the future work there is need to develop an algorithm which could detect more number of defects found on potato skin. Instead of using only color feature there is need to exploit more features to detect large set of defects found in potato. There is need to develop a system which could work on internal defects such as measuring sugar content of tubers. Also there is need to improve an algorithm so that healthy potatoes are not misclassified as defected. There is need to work on large variety of potatoes found in the world. As with the changing varieties the external properties of potato also changes abruptly, So more there is need to focus on these changing varieties.
5. Acknowledgement

This work has been technically supported by Digital Image processing Laboratory, Department of computer engineering, Punjabi University, Patiala. Also there has been great support of all the members of Computer Center, Punjabi University, Patiala.

6. References

1. Gauge C, Sasi S. Modified fuzzy C-means clustering algorithm with spatial distance to cluster center of gravity. IEEE International Symposium on Multimedia (ISM); Taichung, 2010 Dec. p. 310–3.

2. Pedreschi F, Mery D. Computer vision classification of potato chips by color. Journal of Food Process Engineering. 2011 Oct; 34(5):1714–28.

3. Noordam JC, Otten GW, Timmermans TJM, Zwol BHV. High-speed potato grading and quality inspection based on a color vision system. Machine Vision Applications in Industrial Inspection VIII; San Jose, CA. 2000 Mar. p. 1–12.

4. Marique T, Pennincx S, Kharoubi A. Image segmentation and bruise identification on potatoes using a Kohonen's self-organizing map. Journal of Food Science. 2005 Sep; 70(7):415–7.

5. Rousselle P, Robert Y, Crosnier JC. La pomme de terre. Production, amélioration, ennemis et maladies, utilisations. INRA éd. Paris; 1996.

6. Al-Mallahi A, Kataoka T, Okamoto H, Shibata Y. An image processing algorithm for detecting in-line potato tubers without singulation. Computers and Electronics in Agriculture. 2010 Jan; 70(1):239–44.

7. ElMasry G, Cubero S, Molto E, Blasco J. In-line sorting of irregular potatoes by using automated computer-based machine vision system. Journal of Food Engineering. 2012 Sep; 112(1-2):60–8.

8. Razmjooy N, Mousavi BS, Soleymani F. A real-time mathematical computer method for potato inspection using machine vision. Computers and Mathematics with Applications. 2012 Jan; 63(1):268–79.

9. Farhadi R, Sakenian N, Azizi P. Design and construction of rotary potato grader, (Part I). Bulgarian Journal of Agricultural Science. 2012 Apr; 18(2):304–14.

10. Dacal-Nieto A, Vazquez-Fernandez E, Formella A, Martin F, Torres-Guijarro S, Gonzalez-Jorge H. A genetic algorithm approach for feature selection in potatoes classification by computer vision. 35th Annual Conference of IEEE Industrial Electronics, IECON ’09; 2009 Nov. p. 1955–60.

11. Barnes M, Duckett T, Cielniak G. Boosting minimalist classifiers for blemish detection in potatoes. 24th International Conference Image and Vision Computing; New Zealand. 2009 Nov. p. 397–402.

12. Bezdek JC. Pattern recognition with fuzzy objective function algorithms. MA, USA: Kluwer Academic Publishers Norwell; 1981.