3D Color co-occurrence texture features as tool to evaluate quality of fruits

B Sathya Bama*, T Harinie, I Janani Chellam, S Raju and V Abhaikumar
Department of Electronics and Communication Engineering, Thiagarajar College of Engineering, Madurai, India

Received 18 February 2011; revised 26 August 2011; accepted 29 August 2011

This paper proposes a novel quality evaluation method to categorize fruits using color and texture features. In proposed method, input image is segmented using histogram based thresholding technique to identify whether the portion of image is defective or non defective. For defected samples, texture features are extracted from 3D color co-occurrence distribution. Sum of squared distance (SSD) is calculated between texture features of training and test fruits. Non defective fruits are considered as extra class, whereas classes I and II are moderately and highly defective fruits respectively. A database of 150 fruit samples was collected from mozambi and orange fruits with all quality classes for marketable – grade fruit inspection and distribution system. Images (80) were manually classified for training and the rest were used for testing. Proposed method (classification accuracy, > 93%) outperforms existing methods.

Keywords: 3D Color co-occurrence distribution, Histogram based thresholding, Sum of squared distance (SSD)

Introduction
Classification of agricultural products is necessity for agricultural marketing to increase the speed and to minimize misclassification. Fresh fruits must have required external and internal quality when it is harvested and until it reaches the consumer\footnote{1}. Most part of the external quality attributes are currently inspected visually; machine vision provides a means to perform this task automatically\footnote{2}. Current commercial systems for citrus fruits are capable of calculating size, color, shape or ripeness\footnote{3}. Ashutosh et al\footnote{4} studied the effect of soil stress on quality of citrus. Barwal & Shre\footnote{5} revealed a link between juice content and maturity of fruit. Several pixel and region oriented image segmentation and classification algorithms, commonly used in on-line fruit grading\footnote{6}, use different features (color changes, boundaries, texture etc.) and are robust methods against color variations\footnote{7}. Martínez-Uso et al\footnote{8} developed techniques to evaluate citrus fruit's quality, but these techniques were unfit for fruits with uneven reflectance. Yang\footnote{9} used neural networks to classify segmented areas as patch-like defects, elongated defects or non-defective areas. Later, other methods\footnote{10,11} were developed, including automatic skin defect detection methods using multivariate analysis.

This study proposes a region based approach for quality inspection of citrus fruits using color and texture features.

Experimental Section

Material
All types (defective & non defective) of oranges and mozambis were collected. Three different qualities (extra, high quality; class I, medium quality; and class II, poor quality) of fruits were considered. Extra and class I were edible but class II fruits were not.

Proposed Method
In proposed method (Fig. 1), a two step algorithm for quality inspection of fruits, input image is segmented using histogram based thresholding technique to identify whether portion of image is defective or non defective. If fruit is good, histogram formed is unimodal. If fruit is defective, then histogram may have multiple modes. Then defected fruit is clustered and segmented in HSV (hue saturation value) space. Texture features are extracted from segmented defected part using 3D color co-occurrence matrix (CM). A database of 150 citrus fruit images [oranges (80) and mozambi (70)] was created and its textural parameters were calculated.

Image Acquisition
First step consists of acquiring images of fruit surface, while it goes through grading machine. In order
to grade fruits, images should cover the whole surface of fruit and a high contrast has to be created between defects and healthy tissue, while maintaining a low variability for healthy tissue. On common systems, fruits placed on rollers are rotating while moving, and viewed by a camera set above. In this case, fruit parts near the points where rotation axis crosses its surface are not observed\(^1\). This can be overcome by placing mirrors on each side of fruit lines oriented to reflect pole images to the camera\(^13\).

**Histogram-based Thresholding**

Classification problem is basically partitioning the feature space into regions, one region for each category. So, high discriminating features will lead to high and accurate classification rates. Histogram thresholding is applied to separate fruits into defected and non-defected lots. Reduced brightness helps in finding defective fruit. It is based on the shape of histogram where defects can be found using the modes. If bimodal curve is obtained, the dip in curve forms separation between defective and non-defective parts.

**Role of Color and Texture**

Computer vision strategies, used to recognize a fruit, rely on four basic features\(^1\) (intensity, color, shape and texture) to characterize the object. Many kind of citrus fruits [oranges and mozambi (sweet orange)] are subjected to significant variation in color and texture, depending on fruit ripeness. For example, color of mozambi ranges from being uniformly green, yellow, to patchy and brown, whereas its texture varies from medium to smooth. Coloring is an external reference to internal maturity of citrus fruit under the production conditions.

**Color Space Selection**

3D space could be constructed using many color models [RGB (red green blue), LUV (L-Luminescence, U,V-Chrominance) and HSV]. Usually, color shade and brightness matter for a fruit. Since RGB space is not accord with human eye, so HSV and LUV are good selections. Although color is an intrinsic attribute of an image and provides more information than a single intensity value, there have been attempts to incorporate chrominance information into textural features. A color texture can be regarded as a pattern for fruits described by the relationship between its chromatic and structural distribution. Two fruit images, consisting of same color but different textures or same texture but different colors, form four different color texture patterns (smooth & fine yellow, rough & fine yellow, rough & yellow, rough & green yellow).

**Texture Feature Extraction**

Pattern analysis discriminates different patterns of images by extracting the dependency of intensity between pixels and their neighboring pixels\(^1\). Feature extraction techniques for texture description can be classified into four major categories (statistical, model based, signal processing and structural). Recently, different features of color and texture are combined together for their applications in fruit industry\(^16\). Textural analysis can improve detection of defects, since roughness of sound and defective areas of fruit surface usually have different appearances. Kavdir & Guyer\(^18\) used textural features based on spatial grey-level dependence matrices (SGDM) to process monochromatic images of apples. Pydipati et al\(^19\) used a similar technique based on SGDM to inspect defective and sound leaves from citrus trees using computer vision. Mäenpää et al\(^20\) presented a fast method that combines color and texture properties, and chose color percentiles as color features and simple statistics computed in the spatial domain as texture descriptors. Two alternatives to extract features for color texture analysis that appear to be most often used consist of: i) Processing each color band separately by applying grey level texture analysis techniques; and ii) Deriving textural information from luminance plane along with pure
chrominance features. Proposed scheme uses the latter method.

3D Co-occurrence Matrix (CM) Distribution

CMs characterize relationship between the values of neighboring pixels. Therefore, CMs represent a second order statistical measurement of textured surfaces. 2D CM is always used in CBIR systems, and co-occurrence textural features describe some periodicity of an image. Traditional 2D CM is always for gray level images, but color information will be lost. Therefore, a 3D matrix to describe the texture of color image is redefined. In 3D space, angle interval and one pixel distance are not good methods to describe the feature. Therefore, 9 directions are used to describe the feature. Every pixel on H plane is selected; each pixel on H plane corresponds to 9 pixels on S plane and 9 pixels on V plane. A pixel in H plane has 9 neighbor pixels in both S and V planes. There are 9 directions in the matrix space. In this case, color images are coded on three channels, leading to 6 different matrices [(H,H), (S,S) & (V,V) that are the same as grayscale CMs computed on one channel; and (H,S), (H,V) & (S,V)] that account to the correlations between channels]. There are 9 directions, which form 3D matrix space.

Classifier

Classifier works based on the goodness-of-fit measurement by considering a feature set from testing fruits and training fruits. Consequently, sum of squared distance (SSD) metric between test and training fruit texture features are computed. Correct classification is said to occur if SSD value is minimum between actual training and test texture. It is also used to calculate retrieval efficiency by ranking SSDs in ascending order.

Results and Discussion

Data Sets

Images of fruits were taken at various times. Digital camera (resolution, 8.2 mega pixels) was used to capture fruit images. A total of 150 images [oranges (80) and mozambi (70)] were taken at a distance of 20 cm and with proper lighting. In created database, of the first 10 sample oranges, samples 1, 2, 3, 4, 5, 8 and 10 belong to class I, whereas samples 6, 7 and 9 belong to class II. Similarly, of the first 10 sample mozambis, samples 1, 2, 3, 4, 7, 8 and 10 belong to class I, whereas samples 5, 6 and 9 belong to class II. Training set was formed from the database with classified fruits in the collection of fruits. An image for each class was standardized. Then query image was compared with these standardized images and sorted.

Segmentation

For image segmentation, histogram based thresholding was done for query image in training set. In histogram plot (Fig. 2), if the plot is unimodal, then fruit has no defect and it is of extra class; for a fresh and defectless fruit, the intensity will be uniform and graph will be unimodal. Since defected fruit may have non uniform intensities, histogram plot will be bimodal or multimodal. In this way, extra class fruits were sorted out in the first stage itself. Class I and class II images (Fig. 3) were classified based on color and texture properties of the image.

Physical Characteristics related to Color and Texture

Juice content, treatments and juice acidity are the three most important parameters used to determine orange quality. Juice content reveals healthiness of fruit; correspondingly the texture of skin is medium rough. If juice content is less, the fruit will be shrunk and texture becomes very rough. For rotten fruit, juice content will
be excess but is more acidic. For this class of fruit, texture will be very smooth. If quality is related with color, skin color of good orange is smooth and uniformly distributed. But the intensity of top most layer of blemished orange is not uniformly distributed. Skin is normally darker than the good skin. For quality of extra class, it should not be drier or more rotten and also it should be a ripe fruit. This could be effectively shown from color and texture characteristics of dry and rotten fruits. Average intensity of orange color skin was brighter than green color skin. Intensity of blemished skin is usually darker than normal skin. Orange skin has different layers of colors ranging from orange to green. Color transaction area should not be identified as blemishes, although intensity is usually darker on green skin color area. Thus combined color texture features will clearly discriminate the fruits.

For ripe oranges, average intensity gradually increases from edges to centre of the image due to its rounded convex contour. Most significant variation occurred on red-green channel (Fig. 4a). Average intensity on green channel was lower than red channel. Pixel value (Fig. 4b) varied as follows: red channel, 12-255; and green channel, 10-104. Average value of blue channel was very low. For unripe oranges, average intensity on red channel was lower compared with ripe orange, and pixel value varied as follows: red channel, 68-114; and green channel, 70-118. Due to the nature of unripe orange texture, intensity variation on green channel became more important in this case. Texture was medium rough for ripe oranges and medium smooth for unripe oranges.

In proposed approach, statistical parameters (Table 1) were computed (Table 2) for various degrees (0, π/4, π/2, and 3π/4 radians). Energy for class I was very high due to the occurrence of repeated pixel pairs. Contrast for class I was lesser than class II due to greater variation in pixel values of class II images. Correlation values were higher between two different classes and were less among the same class. Since color variation is directly proportional to entropy, here class I had larger entropy than class II. From these values, SSD was calculated (Tables 3). Finally, sorting was done based on this value. If obtained value of class I is less, it will be

![Fig. 4— a) Variation of red-green channel; and b) Variation of blue channel](image)

<table>
<thead>
<tr>
<th>Table 1—Texture features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature</td>
</tr>
<tr>
<td>ASM</td>
</tr>
<tr>
<td>Contrast</td>
</tr>
<tr>
<td>Correlation</td>
</tr>
<tr>
<td>Entropy</td>
</tr>
</tbody>
</table>

Where, $ P(i,j) $ is co-occurrence probability, $ \mu $ and $ \sigma $ are mean and standard deviations respectively.
grouped under class I itself, else it will be grouped under class II. In this way, classification in HSV was done.

Fruit samples 1, 8 and 10 were analyzed by manual classification and found to belong to class I. All the 4 methods classify them under Class I. When compared with other methods of classification, 3D color GLCM gave a very high discrimination between classes for both samples. By manual classification, samples 4 and 5 belonged to class I and sample 9 belonged to class II. Moments and color coocurrence algorithms classified them under class I, whereas GLCM and 3D color GLCM classified them under class II. All three samples had rotten parts inside. Only 3D color GLCM classified them correctly under class II with high discrimination.

By manual classification, samples 2 and 3 belonged to class I. By using moments, these samples had been classified under class II. 3D color GLCM and color coocurrence classified them as class I. Samples 6 and 7 had been classified under class II by both moments and color coocurrence. But 3D color GLCM classified sample 6 under class II and sample 7 under class II. GLCM gave exact opposite result to that of 3D color GLCM. Manual classification and high level of discrimination of 3D color GLCM confirmed that sample 6 belongs to class I and sample 7 under class II. GLCM gave exact opposite result to that of 3D color GLCM. Manual classification and high level of discrimination of 3D color GLCM confirmed that sample 6 belongs to class I and sample 7 under class II. GLCM gave exact opposite result to that of 3D color GLCM. Manual classification and high level of discrimination of 3D color GLCM confirmed that sample 6 belongs to class I and sample 7 under class II. GLCM gave exact opposite result to that of 3D color GLCM.
features proved better than other features by providing classification more accurately than others (Table 5).

Conclusions
A new approach for quality inspection of orange fruits has been proposed using combined color texture features. Inner quality, juice content and acidity of oranges can be qualitatively measured by external appearance because color and texture of orange peel varies for ripe, unripe, rotten and dry fruit. The approach segregates oranges into three quality classes; ripe and healthy oranges as extra class, and unripe and defective oranges as class I and II according to severity. This has been clearly discriminated by texture features extracted in HSV color domain. Performance of proposed method as compared with other existing approaches has proved to be efficient with an classification accuracy of more than 93% for both oranges and mozambis.

Acknowledgement
Authors thank Management and Department of ECE of Thiagarajar College of Engineering for their support and assistance to carry out this work.

References
8 Martý’nez-Uso’a, Pla F & Garé’y a P, A quadtree-based unsupervised segmentation algorithm for fruit visual inspection, in First Iberian Conf on Pattern Recognition and Image Analysis (Mallorca, Spain) 2003, 510-517.

Table 5—Comparison of classification accuracy

<table>
<thead>
<tr>
<th>Fruits</th>
<th>No of Samples</th>
<th>GLCM</th>
<th>Moment</th>
<th>Color co occurrence method</th>
<th>Proposed 3D GLCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orange</td>
<td>10</td>
<td>50</td>
<td>60</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td>Orange</td>
<td>80</td>
<td>50</td>
<td>60</td>
<td>80</td>
<td>95</td>
</tr>
<tr>
<td>Mozambi</td>
<td>10</td>
<td>60</td>
<td>70</td>
<td>90</td>
<td>100</td>
</tr>
<tr>
<td>Mozambi</td>
<td>70</td>
<td>60</td>
<td>70</td>
<td>72</td>
<td>93</td>
</tr>
</tbody>
</table>