Classification & Grading of Tomatoes using Image Processing Techniques

Santi Kumari Behera, Abhijeet Mahapatra, Amiya Kumar Rath and Prabira Kumar Sethy

Abstract: This examination manages the programmed arrangement and reviewing of tomatoes utilizing picture-preparing systems. Mostly tomatoes are of three assortments for example cherry, classic & cylindrical variety and of all assortment have little, medium & big size. Tomato order and evaluating is exceptionally troublesome precisely and in quick way because of their huge contrast in highlight, for example, size, shape and shading because of variable states of nature condition and manual components. Programmed characterization and evaluating of tomato dependent on picture preparing methods is the best arrangement as manual forecast is absence of objectivity, exactness and has lower proficiency. Here, we utilize the picture handling parameters, for example, major axis, minor axis, bounding box, perimeter & diameter for grouping and evaluating reason for tomato and furthermore confirmed with the ground truth measure by Vernier caliper. This examination accomplished coefficient of correlation (R²) 0.98 for length and 0.97 for width. Again, it effectively characterizes the variety of 96.67% and grade into three classes as indicated by size is 100%.

Index Terms: Classification, Grading, Tomato, and Image Processing.

I. INTRODUCTION

Natural item measure has been one of basic fixations for tomato cultivators, regular item retailers and cultivar genetic examiners. Directly off the bat, as a market surveying standard, regular item measure is considered as a quick pointer of tomato quality [1]. Regardless of the way, that natural item assessing for market is generally reliant on the conditions and the idea of the common items, retailers have particular expenses for size-investigated normal items as against size-ungrounded ones. For charge, tomatoes are commonly squeezed in cardboard boxes with a particular standard volume. Measure inspected characteristic items are configuration squeezed in layers to make best use of the case. Moreover, natural item measure imparted as new characteristic item mass or weight [3] [4], is a crucial parameter for yield desire and improvement mode depiction of nursery tomatoes [5]. Yield is vital to nursery cultivators for developing transient collect the official systems and completely deal exhibiting and works the administrators. Finally, natural item size of tomatoes has been one of fundamental spotlights for inherited research on tomato cultivars. Subdued tomatoes can be up to numerous occasions greater than their wild relatives because of chronic assurance of broad size cultivars and present-day transgenic collection improvement [6]. Despite the way that it has been understood that tomato measure is constrained by quality and can be gained by increase [7] [8], it has been difficult to clone these characteristics in tomatoes even with present day pushed genomics gadgets. The troubles are in light of the fact that natural item size of tomatoes is not influenced by one quality, anyway various characteristics acting together, which are called quantitative trademark loci (QTLs). Gauge estimation of natural item can help plant reproducers to separate phenotypic verbalization of target characteristics maybe associated with normal item measure.

Estimation of tomato characteristic items has been driven in history for a long time by mechanical way [9]. Human movement is as often as possible a work genuine, dull and passionate task. In perspective on the comfortable association between size and new natural item weight [3], electronic measure bowls for taking care of tomatoes subject to impact norms have been attempted in research office and in field for characteristic item looking into of tomatoes [8]. In any case, most of these characteristic item-assessing systems for tomatoes is routinely clumsy, available for off-vine natural item estimation, and cannot be used for in situ estimation of on-vine tomato regular items. With the headway of picture distinguishing and taking care of strategies, the estimation of regular item size could be coordinated subject to machine-vision strategy [2] [11]- [13]. One system for size estimation of natural items was to use the district having nourishments developed starting from the earliest stage float remove crosswise over after Binarization subject to the shading information [14]. On the other hand, automated perceptive close infrared imaging was examined for modified common item evaluating [15]. In spite of the way that these optical techniques for size estimation of regular items are much favorable for in-field conditions appeared differently in relation to the mechanical precursors, they oftentimes encountered the uncleanness of PC vision in view of the idea of got pictures. Consistently in light of the unstructured thought of regular green settings and normal assortment of plants inside them, question ID reliant on machine vision is astonishingly continuously troublesome [13].

Figure 1 shows the different varieties of tomatoes i.e. cherry, classic and cylindrical. The cherry variety of tomato is of very small size and mostly circular in dimension, the classic variety of tomatoes ranges from medium size to big size with circular dimension whereas the cylindrical variety of...
Classification & Grading of Tomatoes using Image Processing Techniques

Tomatoes are elongated length wise whose dimension is close to that of an oval shape.

Fig. 1: Different varieties of tomatoes : (a) Cherry (b) Classic (c) Cylindrical

This paper consists of four sections: first section reports the literature survey done in order to gather knowledge about the classification of various fruits & vegetables from earlier studies, second section reports the materials and methodology used to design the system, third section reports the results obtained from the system and at last fourth section concludes the paper.

II. RELATED WORKS

Zhi-yuan Wen et. al. [13] a framework to enhance the evaluating precision of citrus natural product by portraying the shading and shape highlights. The citrus organic product tests included Citrus unshiu Marc. cv.unbergii Nakai. Pictures from peduncle, calyx and two inverse sides were caught for the exploratory reason. A sum of 120 example natural products were tried and the precision for shading and shape evaluating was observed to be 95.83%.

Sudhir Rao Rupanagudi et. al. [14] presented a perceptive development-reviewing framework for tomato utilizing modest materials and picture handling methods for recognizing the 6 phases of tomato aging. The framework was produced utilizing Simulink, which is a piece of MATLAB 2011b on a 2.5GHz CPU. A general precision of 98% alongside an execution speed more noteworthy than 7.6 occasions when contrasted with past two famous strategies was acquired.

V.Pavithra, R.Pounroja et. al. [15] proposed a framework for programmed and non-damaging reviewing of cherry tomatoes, which depends on quality and development. The evaluating procedure comprised of two stages, first being the reviewing of tomatoes as for development level and second being the reviewing of developed tomatoes concerning quality. For the extraction procedure of tomatoes from foundation and leaves, another shading-based division technique was proposed which depends on Euclidean separation. Regarding exactness and calculation time, it outperforms both Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) classifier.

Ujjwal Verma et. al. [16] presented a total framework to dissect the development of tomatoes in open field, which was a testing undertaking as the impediment was serious and the differentiation was poor in the pictures obtained in the open field. Division technique was utilized at first to recognize the tomatoes after which the size was estimated. The mistake was around 4% among manual and programmed division process and the vigor of the framework was great (up to half) as for impediment. The precision of the proposed framework was observed to be at 96.5%.

Megha. P. Arakeri et. al. [17] displayed a programmed evaluating framework for tomato dependent on PC vision procedures, which included two stages, i.e., equipment improvement and programming advancement. The equipment created was useful in catching the pictures and moving the tomatoes to address receptacles with no manual obstruction, while the product created utilizing picture handling systems helped in dissecting the tomatoes for imperfections and readiness. The proposed framework accomplished a precision of 96.47% in reviewing the nature of tomatoes.

Katrin Utai et. al. [18] presented a created calculation to compute the mass utilizing an Artificial Neural Network (ANN) model of the mango cultivar ‘Nam Dokmai by changing the pictures of the mango to disentangle the item acknowledgment undertaking and separated the highlights like width, zone, thickness and length which are utilized as contributions to the ANN demonstrate. Here they have talked about seven unique ways to deal with get the measurement and compute the mass of the fruit. The precision rate was 97% with most noteworthy coefficient of efficiencies of 0.99 utilizing two (region and thickness) or three (thickness, length and width) input parameters.

Oktaviana Rena Indriani et. al. [19] presented a tomato order technique utilizing KNN dependent on Gray Level Concurrence Matrix (GLCM) and Hue Saturation Value (HSV). The reason GLCM and HSV were picked is on the grounds that dependent on the estimation of differentiation, connection, vitality and homogeneity GLCM has the most elevated level of acknowledgment and HSV can be utilized for shading investigation. Presently the result from GLCM and HSV was given as contribution to KNN to characterization since it uses the separation (k) as an examination of the closeness dimension of the pictures. The examination was finished utilizing 100 informational collections out of which 75 were utilized for preparing and 25 for testing reason. 100% exactness was acquired with p esteem in GLCM being nine and the participation esteem (k) in KNN being three.

M. Karkee et. al. [20] carried out an examination for building up a machine vision framework which contained shading CCD camera and a period of-flight (TOF) light-based 3D camera, which helped in assessing the extent of apples in tree shelters. To ascertain the natural product measure, significant hub (longest pivot) was resolved utilizing 3D organizes, whose exactness was observed to be 69.1%.

Yet, in the pixel-measure based technique, the...
exactness of deciding the real pivot was expanded to 84.8%.

Reza Ehsani et al. [21] proposed a strategy to distinguish the dimension of development of crisp tomatoes (Roma and Pear assortments) from market where the component shading esteem was joined with Back Propagation Neural Network (BPNN) classifier. The development identification gadget was created utilizing PC vision innovation where the information tomato pictures were prepared utilizing picture handling systems. The precision rate for this technique was observed to be 99.31% with a standard deviation of 1.2%.

Nay Zar Aung et al. [22] reported a picture preparing calculation for estimation of shape and size of strawberry natural product with and without calyx. Here, right kites and straight-forward kite geometry was utilized to find the head point, peak point, and two side focuses after which pinnacle edge was drawn. At that point, the calculation looks through the tringle whose apex angles were more noteworthy than 75 degrees. After that, by considering width and proportion of length and distance across, strawberries were arranged into four classes. The outcomes demonstrate that the exactness for distance across and length estimation are 94% and 93% individually for strawberries without calyx impediment and 94% and 89% for that with calyx impediment. The grouping exactness was somewhere in the range of 94% and 97% and the normal preparing time for one strawberry (one piece) was underneath 0.45secs to 0.50secs.

Eduardo Jr Piedad et al. [23] developed a handy framework for order of banana (Musa acuminate AA Group ‘Lakatan’) with the assistance of machine realizing which depended on level-based arrangement as opposed to characterizing separately (“finger”) for reasonable reason. Here the natural products were characterized into four classes, in particular, additional class, class I class II and reject class. The after effect of this framework was then contrasted and ANN, SVM and Random Forest (RF) classifiers and was discovered that the precision of this framework was 97% if the reject class was overlooked.

Table 1: Summary of related work.

<table>
<thead>
<tr>
<th>Sl. No</th>
<th>Source</th>
<th>Fruit(s)</th>
<th>Methodology</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Zhi-yuan Wen et al. [13]</td>
<td>Citrus</td>
<td>Machine vision with fractal dimension</td>
<td>95.83% accuracy</td>
</tr>
<tr>
<td>02</td>
<td>Sudhir Rao Rupanagudi et al. [14]</td>
<td>Tomato</td>
<td>Image processing techniques</td>
<td>98% accuracy</td>
</tr>
<tr>
<td>03</td>
<td>V.Pavithra, R.Pounroja et al. [15]</td>
<td>Cherry Tomato</td>
<td>Colour based segmentation based on Euclidean distance, KNN based on SVM classifier</td>
<td>SVM and KNN classifier perform well.</td>
</tr>
<tr>
<td>04</td>
<td>Ujjwal Verma et al. [16]</td>
<td>Tomato</td>
<td>Image segmentation and size estimation</td>
<td>96.5% accuracy</td>
</tr>
<tr>
<td>05</td>
<td>Megha P. Arakeri et al. [17]</td>
<td>Tomato</td>
<td>Computer vision techniques</td>
<td>96.47% accuracy</td>
</tr>
<tr>
<td>06</td>
<td>Katrin Utai et al. [18]</td>
<td>Mango</td>
<td>Image processing and Artificial neural network (ANN)</td>
<td>97% accuracy</td>
</tr>
<tr>
<td>07</td>
<td>Oktaviana Rena Indriani et al. [19]</td>
<td>Tomato</td>
<td>KNN based on GLCM and HSV color space</td>
<td>100% accuracy</td>
</tr>
<tr>
<td>08</td>
<td>M. Karkee et al. [20]</td>
<td>Apple</td>
<td>3D machine vision system</td>
<td>84.8% accuracy</td>
</tr>
<tr>
<td>09</td>
<td>Reza Ehsani et al. [21]</td>
<td>Tomato</td>
<td>Computer vision with BPNN classifier</td>
<td>99.31% accuracy</td>
</tr>
<tr>
<td>10</td>
<td>Nay Zar Aung et al. [22]</td>
<td>Strawberries</td>
<td>Image processing techniques</td>
<td>Accuracy between 94% to 97%</td>
</tr>
<tr>
<td>11</td>
<td>Eduardo Jr Piedad et al. [23]</td>
<td>Banana</td>
<td>Tier-based machine learning</td>
<td>97% accuracy</td>
</tr>
</tbody>
</table>

III. MATERIAL AND METHODOLOGY

A. Machine Vision System

The machine vision system consists of a box with white background, camera and a computer with MATLAB shown in Figure 2. Images of different varieties of tomato were captured using a digital camera of 16 megapixels. The distance of the camera from the fruit is maintained at 10cm for each picture taken. At first, the whole setup was done, which included LED lights from every side (side & top) to reduce the shadow of the fruit, and then a digital camera of 16 megapixels was setup, which was then used to capture the images that assured the images were of high quality. After that, these pictures were then taken as contribution for estimation of different parameters, for example, zone, centroid, bounding box, significant hub, minor pivot, centroid, length and width. For the estimation of these parameters, MATLAB R2017b was utilized on a framework having Windows 10 working framework with RAM size of 4GB. The figuring procedure incorporated the transformation of info RGB picture to Gray Scale picture; at that point, thresholding was finished. After which the pictures were changed over to Binary pictures and little point disposal was improved the situation the simplicity of the figuring of the different parameters. From the after effect of little point disposal and the edge of the organic product determined, the framework could decide the measure of the tomato natural product, i.e., Small, Medium or Big size. Of course, a bouncing box was fitted, which at that point helped in deciding the assortment of the tomato natural product, i.e., Cherry, Classic or Cylindrical.
B. Field Data Collection

Tomatoes of different sizes and shapes were considered intentionally to cover different types of tomatoes available. The tomatoes have been handpicked to assure the quality of the tomato fruit being the best. Three different varieties of tomato were considered, i.e., Cherry, Classic & Cylindrical variety.

C. Algorithm of the proposed system:

**Step 1:** Images are taken as input.

**Step 2:** RGB images are then converted into GRAYSCALE.

**Step 3:** Thresholding is then used where gray level is set below 150 as the background is white.

**Step 4:** Then the images are converted into binary images.

**Step 5:** “bwareaopen (BW, 800)” function is used to take objects with pixel value more than 800.

**Step 6:** “region props ()” function is used to calculate area, centroid, bounding box, major axis, minor axis, perimeter, length & width of the input image.

**Step 7:** if Perimeter<6500 pixels

- Tomato is small.
- Else if 6500 pixels <Perimeter<900 pixels
- Tomato is medium.
- Else
- Tomato is big size.

**Step 8:** After the calculation of all the parameters in Step 7, a bounding box was fitted around the fruit image.

**Step 9:** if major axis/minor axis<1.0

- Tomato is of Cherry variety.
- Else if 1.0<major axis/minor axis<1.2
- Tomato is of Classic variety.
- Else
- Tomato is of cylindrical variety.

**Step 10:** At last, the result is displayed on the command window.

IV. RESULT AND DISCUSSION

The proposed approach was approved by 300 number of tests i.e., 100 number from every assortment of tomatoes. The procedure has two stages, first it orders the assortment and after that, it classifies as indicated by its size. The characterization depended on the proportion of length and width (L/W) and reviewing as indicated by size by evaluating its edge. The order and evaluating ranges are represented in Table 3 and Table 4 separately. Before expectation of grouping and reviewing we’ve checked the estimation honesty of our strategy by cross approving the anticipated length and width with the ground truth esteem i.e., estimated utilizing Vernier caliper likewise with the blunder %, for which we took 12 number of tomatoes and accomplished the coefficient of connection (R²) esteem 0.98 for length and 0.97 for width, which are outlined in Table 2 and Figure 4 & 5.

**Table 2:** Actual and Predicted measurement of Length and Width of Tomatoes

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Actual length (mm.)</th>
<th>Predicted length (mm.)</th>
<th>Error % for length</th>
<th>Actual width (mm.)</th>
<th>Predicted width (mm.)</th>
<th>Error % for width</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>45.8</td>
<td>42.4865</td>
<td>7.234%</td>
<td>45.4</td>
<td>41.4528</td>
<td>8.694%</td>
</tr>
<tr>
<td>2</td>
<td>34</td>
<td>31.8287</td>
<td>6.863%</td>
<td>33.9</td>
<td>30.9727</td>
<td>8.634%</td>
</tr>
<tr>
<td>3</td>
<td>36.9</td>
<td>32.5932</td>
<td>11.67%</td>
<td>35.7</td>
<td>31.8566</td>
<td>10.76%</td>
</tr>
<tr>
<td>4</td>
<td>62.6</td>
<td>68.1024</td>
<td>-7.192%</td>
<td>56.3</td>
<td>60.5053</td>
<td>-7.465%</td>
</tr>
<tr>
<td>5</td>
<td>62</td>
<td>69.3191</td>
<td>22.87%</td>
<td>64.6</td>
<td>65.8139</td>
<td>-1.879%</td>
</tr>
</tbody>
</table>
Table 3: Range of L/W for classification

<table>
<thead>
<tr>
<th>Range of L/W in mm</th>
<th>Classification Variety</th>
</tr>
</thead>
<tbody>
<tr>
<td>L/W&lt;1.0</td>
<td>Cherry</td>
</tr>
<tr>
<td>1.0&lt;L/W&lt;1.2</td>
<td>Classic</td>
</tr>
<tr>
<td>L/W&gt;1.2</td>
<td>Cylindrical</td>
</tr>
</tbody>
</table>

Table 4: Range of perimeter for grading

<table>
<thead>
<tr>
<th>Range of Perimeter in mm</th>
<th>Size Variety</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perimeter&lt;6500pixels</td>
<td>Small</td>
</tr>
<tr>
<td>6500pixels&lt; Perimeter&lt;9000pixels</td>
<td>Medium</td>
</tr>
<tr>
<td>Perimeter&gt;9000pixels</td>
<td>Big</td>
</tr>
</tbody>
</table>

Table 5: Performance analysis of classification of tomatoes

<table>
<thead>
<tr>
<th></th>
<th>Cherry</th>
<th>Classic</th>
<th>Cylindrical</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cherry</td>
<td>100 %</td>
<td>×</td>
<td>×</td>
<td>100 %</td>
</tr>
<tr>
<td>Classic</td>
<td>10</td>
<td>90</td>
<td>×</td>
<td>90 %</td>
</tr>
<tr>
<td>Cylindrical</td>
<td>×</td>
<td>×</td>
<td>100</td>
<td>100 %</td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>96.67 %</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For validation purpose, we took 100 numbers of each variety of tomato, which in total is 300 numbers with all ranges of size. The methodology successfully classified cherry and cylindrical with 100% of accuracy. However, for classic variety it predicted erroneously i.e., 10 numbers of classic variety were predicted as cherry, which implies 90% of accuracy. In addition, this error only occurs when the classic variety was of very small size. Again, the methodology successfully graded all tomatoes in proper class according to their size. Therefore, the methodology achieved accuracy of 96.67% and 100% for classification and grading respectively.

V. CONCLUSION

This paper reports a strategy of characterization and evaluating of tomato dependent on picture preparing procedures with the assistance of image processing, for example, major axis, minor axis, bounding box, perimeter & diameter. The technique effectively characterized the tomato in three noteworthy assortments i.e., cherry, classic and cylindrical with 96.67% of exactness and reviewed into three class as indicated by their size i.e., little, medium and big size with 100% of precision. This technique may plan a programmed arranging machine for order and evaluating motivation behind tomatoes which will be valuable for vegetable sellers to get the real cost for their items.
REFERENCES


AUTHORS PROFILE

Santi Kumari Behera has received her Master of Technology degree in Computer Science & Engineering from NIT Rourkela. She is working as Assistant Professor in Veer Surendra Sai University of Technology, Odisha, India. Her research of interest in Computer Graphics and Image Processing.

Ahibjeet Mahapatra is pursuing his Master of Technology degree in Computer Science & Engineering from VSSUT Burla. His research of interest is Agricultural Image processing.

Amiya Kumar Rath has already received Ph.D. in Computer Science from the Utkal University, Odisha, India. He is working as Professor at Veer Surendra Sai University of Technology, Odisha. He has published over 70 number of research articles in various areas of Computer Science, concentrating on Artificial Intelligence, Image Processing, and Embedded System.

Prabir Kumar Sethy has received his Master of Technology degree in Communication Engineering from IIT (ISM) Dhanbad. He is working as Assistant Professor in Electronics, Sambalpur University, Odisha. His interest of research area is Image Processing. He has published more than 20 number of research article in International Journal and Conferences.