Turkish Journal of Agriculture and Forestry

http://journals.tubitak.gov.tr/agriculture/

Research Article

Turk J Agric For
(2013) 37: 62-68
© TÜBİTAK
doi:10.3906/tar-1201-10

Online tomato sorting based on shape, maturity, size, and surface defects using machine vision

Omid Omidi ARJENAKI, Parviz Ahmadi MOGHADDAM*, Asad Moddares MOTLAGH
Department of Agricultural Machinery Engineering, Faculty of Agriculture, Urmia University, Urmia, Iran

Received: 09.01.2012 • **Accepted:** 04.07.2012 • **Published Online:** 15.01.2013 • **Printed:** 15.02.2013

Abstract: Online sorting of tomatoes according to their features is an important postharvest procedure. The purpose of this research was to develop an efficient machine vision-based experimental sorting system for tomatoes. Relevant sorting parameters included shape (oblong and circular), size (small and large), maturity (color), and defects. The variables defining shape, maturity, and size of the tomatoes were eccentricity, average of color components, and 2-D pixel area, respectively. Tomato defects include color disorders, growth cracks, sunscald, and early blight. The sorting system involved the use of a CCD camera, a microcontroller, sensors, and a computer. Images were analyzed with an algorithm that was developed using Visual Basic 2008. In order to evaluate the accuracy of the algorithms and system performance, 210 tomato samples were used. Each detection algorithm was applied to all images. Data about the type of each sample image, including healthy or defective, elongated or rounded, small or large, and color, were extracted. Results show that defect detection, shape and size algorithm, and overall system accuracies were 84.4%, 90.9%, 94.5%, and 90%, respectively. System sorting performance was estimated at 2517 tomatoes h^{-1} with 1 line.

Key words: Grading, machine vision, sorting, tomato

1. Introduction

Tomato is one of the important products in human nutrition that is consumed by millions of people daily. In addition, this product has a special place among Iranian families. According to FAO statistics, world tomato production was 314 million t in 2010, and Iran ranks sixth in the world with 5 million t of tomato harvested. About 75% of tomatoes produced are consumed fresh, and so appropriate appearance is very important. Immaturity and ripening disorders in tomatoes are common defects seen in markets. According to Velioglu et al. (1998), vulnerabilities and defects in tomatoes increase with overuse of pesticides and toxins and incorrect storage. One of the most important processes in packaging and product supply to the market is sorting. This operation requires different parameters for quick identification and management. Parameters include maturity, color, shape, size, and defects. According to Jarimopas and Jaisin (2008), the efficiency and effectiveness of sorting governs the quality standard of the packing lines and the product, which, in turn, determines the marketability of the product. Accordingly, it is necessary to have a rapid, consistent, effective, and robust method for sorting. Manual sorting is the most common method for sorting the fruits. The following problems emerge in quality control carried out by humans: high labor costs, labor fatigue, inconsistency, and low precision due to various factors such as variations in ambient light intensity, differences in personal perception of quality, and scarcity of trained labor. Using machine vision will contribute to the automation of sorting and reduce the labor costs and number of employees required. The best technique for quality evaluation and fruit sorting is machine vision. Among the advantages of machine vision are nondestructiveness, accuracy, and consistency. According to Yud et al. (2002), a machine vision system can accurately identify the internal and external characteristics of agricultural products, including the degree of maturity, defects, moisture, and nutrients. The charge-coupled device (CCD) digital cameras used in previous studies assessed the characteristics of color or monochrome grades to determine the quality of products illuminated by a light source. This technique was used by Lino et al. (2008) in a study that classified lemons and tomatoes according to color, defects, and volume. Equatorial diameter was measured in millimeters, and the surface area was expressed in pixels as a mean of diameter optical evaluation. The correlation coefficient between these 2 parameters (equatorial diameter and

* Correspondence: p.ahmadi@urmia.ac.ir

surface area) was 0.89. Blasco et al. (2009) developed a machine for singulating, inspecting, and sorting satsuma mandarin (*Citrus unshiu*) segments using morphological features. Their system automatically identified pieces of skin and other raw material, separated whole segments from broken ones, and was able to correctly classify 93.2% of sound segments on conveyor belts at 600 mm s^{-1}. A practical application was demonstrated by Zhang et al. (2009) in a study that developed a machine vision system to automatically sort cherry tomato according to maturity. Nine features were extracted from each image. Tomatoes were classified into 3 categories (unripe, half-ripe, and ripe). Images were captured in the RGB color space. The principle component analysis (PCA) results showed that ripe tomatoes were distinguished from immature and half-ripe tomatoes. The machine was able to correctly classify 93.2% of tomato samples.

In industry today tomatoes are sorted manually, as are satsuma, limes, pomegranate, and other fruits. The objective of this research was to develop an efficient automated sorting system for tomatoes based on the image processing techniques that were effectively used with limes, pomegranate, and other products.

2. Materials and methods

2.1. Hardware and software design

The hardware included a conveyor, power drive with inverter, light source, CCD camera, mechanical segregator, control unit, and computer. The software consisted of separate algorithms for shape, size, maturity, and defect detection.

2.2. Hardware and operation

Figure 1 shows the experimental tomato sorting system. It featured a black conveyor belt 25 cm wide and 500 cm long with 2 receivers for sorted tomatoes. The conveyor speed was 105 mm s^{-1}, and conveyor speed could be increased using industrial cameras. The conveyor was driven by a 1.49-kW, 3-phase electric motor that was adjusted by an inverter (LG Inverter SV-iG5). On the left side of the belt (Figure 1) was a box with a CCD camera (Sony, Japan). It was equipped with a circular polarizing filter and mounted on top of the conveyor. There were 4 LED lamps of 220 V (500–700 nm wavelength) with 4 polarizing films on the right and left sides (45° from horizontal) and above (perpendicular to the surface) the box to provide uniform light intensity with minimum shadow and light reflection. The camera, with a focal length of 40.6–406 mm, was mounted 53 cm above the belt and provided a resolution of 2M pixels (spatial resolution: 640 × 480). In this study, tomatoes were classified into 2 categories (desirable and undesirable) based on maturity, defects, shape, and size (2-D area) (Table 1).

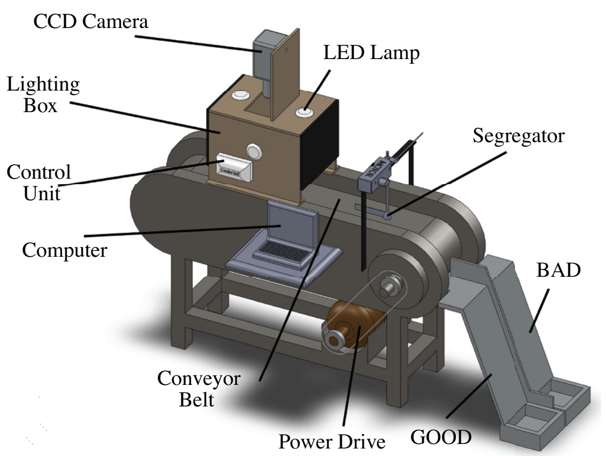

Figure 1. An experimental machine vision system for sorting tomatoes.

Table 1. Minimum thresholds of expectation.

Type of sorting	Threshold
Defect	72
Shape	0.722
Size	35,696 (11 ± 0.2 cm^2)
Maturity	R = 45–104, G = 23–50, B = 26–46

The electric control unit (Figure 2) comprised a microcontroller (ATmega8) and an IR sensor (made in the Urmia University Agricultural Machinery Engineering Department workhouse) with wavelength of 840 nm. A computer (CPU speed: 2.8 MHz, dual core) was used for signal processing and capturing/processing images. While

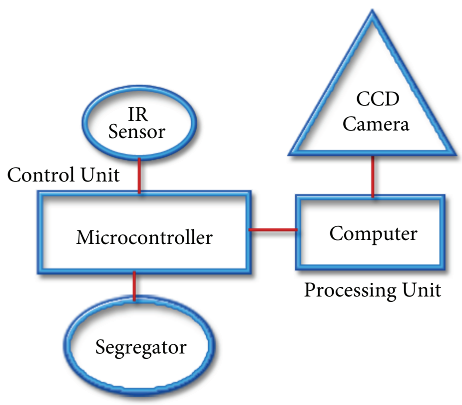

Figure 2. Block diagram depicting functional units.

traveling on the conveyor, the tomatoes passed an IR sensor comprising a 3-IR diode and a receiver. The signal produced by the sensor was sent to the microcontroller, processed, and sent onwards through a USB port to the computer. The software commanded the CCD camera driver to take an image using the IR diode and receiver.

2.3. Image preprocessing

Algorithms for processing the images were developed in Visual Basic 2008. In order to separate the tomato from the background, images were converted to HSI space (hue, saturation, intensity). Image pixels with color outside of the specified HSI range were filtered and filled with black. The processing duration for each image was 0.7 s. The suitable ranges for HSI components were: 71–31 for H, 0–1 for S, and 0–1 for I. The new image was converted to RGB space, and its pixel colors were filtered in (0, 25) for red and (0, 64) for the green and blue range. The image was then transformed to grayscale and thresholded with Otsu's algorithm. The final image was of tomato (Figure 3).

2.4. Determination of shape

The shape of the tomato can be identified from its curvature. Shapes were categorized into "rounded" or "oblong". In order to find the shape index of a tomato, its eccentricity was calculated. In mathematics, the eccentricity, denoted E or ε, can be thought of as a measure of how much the 2-D object (section) deviates from being circular. As the eccentricity of a circle is 0, the eccentricity of an ellipse is greater than 0 but less than 1 (Weisstein 2011). The visualization of eccentricity requires the center of gravity coordinates and the width, height, and area of the image. It was calculated using Eq. (1).

$$E = 2\frac{\sqrt{\left(\frac{L_{Maj}}{2}\right)^2 - \left(\frac{L_{Min}}{2}\right)^2}}{L_{Maj}} \tag{1}$$

Here, L_{Maj} is the length of the major axis, and L_{Min} is the length of the minor axis (Gonzalez and Wood 2002).

In addition, the center of the tomato could be calculated as the mean value of the X and Y coordinates of the tomato's points (van Assen et al. 2002). As E approaches zero, the tomato becomes more circular.

2.5. Determination of size, maturity, and defects

The segmentation algorithm that was explained in the section on shape determination was used. However, images were not thresholded since the black area (tomato only) was not considered. The area value was the size index. Tomato maturity was inspected by 3 experts. After determination of 50 mature and 50 immature tomatoes, the tomato images were captured, histograms of those images were made, and the mean of color component was calculated. This value was used as a base for identifying mature tomato status in the algorithm.

Defects of tomatoes include color disorders, growth cracks, sunscald, and early blight. With the exception of being crushed, defects in tomato cause a color change in the tomato skin. Images captured from defects were thresholded with Otsu's algorithm, and images captured from immature tomatoes were thresholded with the simple image statistic (SIS) algorithm. In the final image, the background and defects were black, and the intact section of the tomato was white. The identification index of defects was "fullness". This index was calculated as the ratio of object area to the multiplied value of the width and

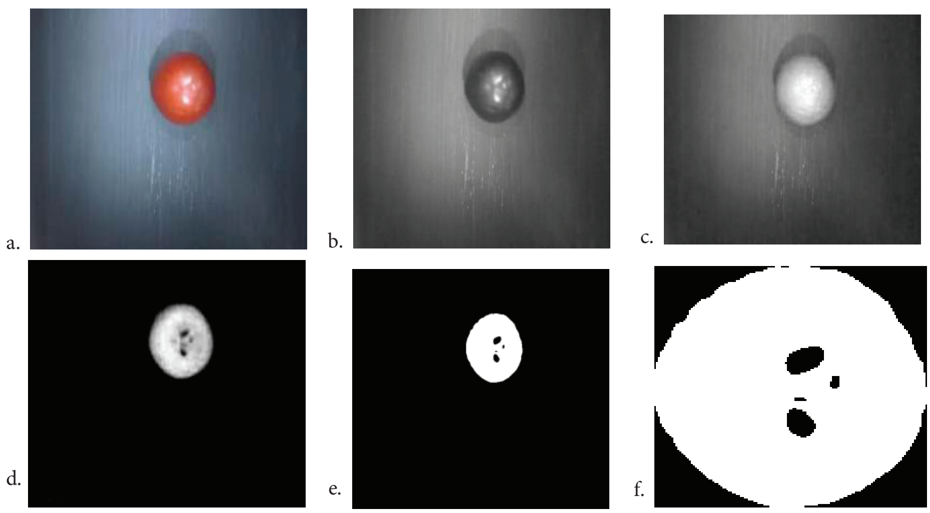

Figure 3. Image processing steps: a) original image, b) green component, c) red component, d) G-R component, e) thresholded image, f) final image.

height. If it equaled 1, the entire rectangle of the object was filled by the pixels of the object (no black areas), which is true only for white rectangles. If it equaled 0.5, only half of the bounding rectangle was filled by white pixels. As the fullness value decreased, the defective area increased.

By comparing the experimental results with the threshold values for size, color, defective area, and eccentricities, which were typed into the software, the signal for a relevant category (desired or undesired) of sorting was produced.

2.6. Performance test

The influences of conveyor speed, tomato spacing on the belt, and light intensity on the total performance of the machine vision system were evaluated. The optimum speed of the belt was determined by analyzing the images that were captured at different belt speeds. The spacing distance of tomatoes in feeding was related to IR-sensor sensitivity and the visual field of the camera. After tomato sample images from the light and dark regions of the conveyor were captured, the tomato colors in both regions were analyzed in terms of gray level profiles of RGB components.

In order to remove the reflected light from tomatoes, 5 polarizing filters, 4 for lamps and 1 for the camera, were used. Samples were imaged, and the gray-level profiles of bright spots and other tomato sections were obtained at the same pixel distance and were analyzed.

2.7. Performance of the machine vision system

The sorting system was operated under conditions considered to be optimum with regard to belt speed and tomato spacing. A total of 210 tomatoes (Red Cloud) of various degrees of quality (good, defective, and immature; rounded and oblong; different colors and sizes) were randomly selected from markets by an expert. Initially, sorting type (size, shape, maturity, and defects) was selected in the software. The tomato was then sorted out into 2 categories (desirable and undesirable). The number of correctly and incorrectly sorted tomatoes in each receiver was recorded. The mean value of missorting (error) and the throughput capacity of the sorting system were evaluated. The mean value of missorting was evaluated using the following equation:

$$\overline{C}_E = \frac{\sum N_{ij}}{\sum N_i} \qquad (2)$$

where N_{ij} is the number of class j tomatoes in the class i receiver, and Ni is the number of class i tomatoes in the class i receiver.

3. Results

The most appropriate conveyor belt speed for image capturing was 10.5 cm s^{-1}. Considering the ability of the IR sensor to accurately sense all tomatoes and the evaluation of the camera's visual field for the imaging of 1 tomato, the perfect spacing distance for feeding was 15 cm. Spacing distance consisted of the distance between the sensor and the end of the camera's visual field (10 cm) and the minimum confidence distance (5 cm).

3.1. Minimum threshold values of expectation

All data gained from sorting experiments for each sorting criterion were distributed normally (in accordance with Kolmogorov–Smirnov analysis). Minimum of fullness, top quartile value of eccentricity, down quartile value of 2-D area, and range between minimum and maximum of mean color were assumed as thresholds of expectation in defect, shape, size, and maturity sorting type, respectively (Table 1).

3.2. Color analysis in maturity sorting

Images obtained from 50 samples of tomatoes, with the component value of mean color related to each sample, were collected for investigation. The averages of each component value were calculated (Table 2).

3.3. Sorting system performance

By applying the minimum thresholds of expectation, samples were sorted, and the sorting accuracy of the system for each type of sorting was calculated (Table 3). In this study, it was assumed that defects appeared only on the side of the tomato seen by the camera.

4. Discussion

4.1. Effects of polarizing filters

The results obtained from gray-level profiles indicated that the color components of tomato in all analyzed distances were uniformed by the use of polarizing filters, and the

Table 2. Result of study and analysis of tomatoes in sorting based on maturity.

Type of samples	Average of R	Average of G	Average of B
All samples	85.12	43.9	38.3
Healthy samples	82.68	40.22	37.76
Defective samples	92.08	42.08	39.23

Table 3. Calculated accuracy of each sorting type.

Sorting type	Expert identification		System identification		Accuracy
	Good	Bad	True	False	
Defect	68	32	85	15	85.00%
Shape	40	15	50	5	90.90%
Size	28	27	52	3	94.54%
Maturity	37	13	46	4	92.00%

red component was higher. However, in the absence of a polarizing filter, at certain points in the distance the gray level of color components was close together or equal (Figure 4). As a result, these points were blacked and were identified as defects. The use of polarizing filters to remove the effects of light reflection from the tomatoes was effective and necessary. It was very important, especially when the defects were on bright spots. Polarizing filters were used in 3 different modes: 1) filter used only on camera, 2) filter used only on light sources, and 3) filter used on both camera and light sources. Use of the filter in the third mode provided the best results.

The results obtained from gray-level profiles indicated that along with sampling distance, gray levels of green (G) and black (B) were affected by light intensity (Figure 5). As the intensity of light became greater, the gray levels of G and B became greater and moved closer to red (R). Under these conditions, B variations were greater than those of G on the conveyor. The difference between R and G for tomato separation from the background was almost constant under both light conditions on the conveyor. As a result, use of the difference between R and G for segmentation was more appropriate.

4.2. Color analysis in maturity sorting

Results (Table 2) showed that the average of R was the maximum among all averages. The average of B was less than the others. The average values of all components in defective samples were greater than in healthy samples (57.8 for defective samples; 53.55 for healthy samples). Two profiles of gray level in defective and healthy tomatoes are shown in Figure 6.

In order to separate the tomato from the background, the G component of the image was subtracted from the R component. Because the difference between R and G in the tomato was 30 times greater than in the background and defects, as a result of subtracted images, the tomato appeared more significant and specific than the background. Therefore, the background was removed and the defects were effectively identified.

4.3. Sorting system performance

The average of sorting accuracy was 90.61% (Table 3). Zhang et al. (2009) reported 94.9% accuracy of identification of ripe tomatoes from immature and half-ripe tomatoes. Those authors sorted tomatoes based only on maturity. The current study simultaneously considered maturity, defects, shape, and size in one algorithm as

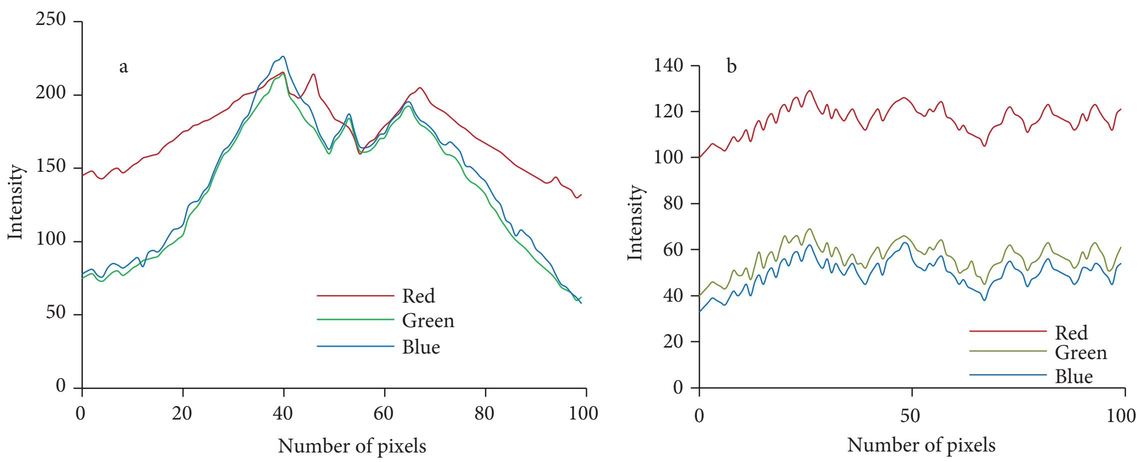

Figure 4. Image profile of captured tomato: a) without polarized filter, b) with polarized filter.

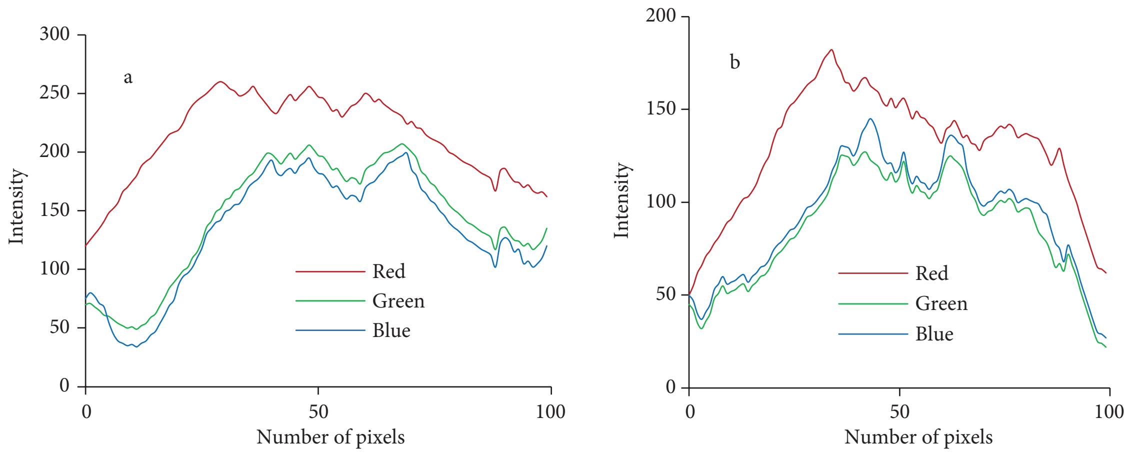

Figure 5. Image profile of captured tomato: a) in bright region, b) in dark region.

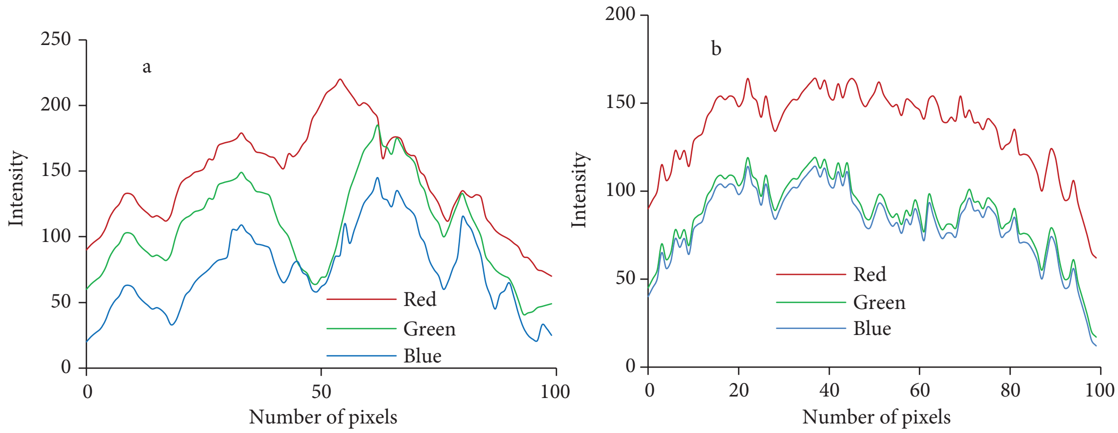

Figure 6. Gray-level profiles for: a) defective tomato, b) healthy tomato.

sorting factors. For this reason, the accuracy obtained in this study is slightly lower than that reported by Zhang et al. (2009).

Accuracy for size sorting was higher than overall accuracy and the accuracies for maturity, shape, and defect sorting. The reasons for this difference are manifold. One reason could be that immature tomatoes were assumed to have defects. In this case, 4 thresholding methods were used. The Otsu method gave optimum values for healthy tomatoes; however, the SIS method and a value of 16 for threshold resulted in immature and defective tomatoes in thresholding, respectively. Only the Otsu method was chosen for sorting defects. The second reason could be incorrect identification of tomatoes by shape and color by experts. The third reason could be that there were only 2 classes to be sorted in all types of sorting, while more classes were sorted in combined sorting, which suggests that less contamination was likely to occur during size sorting. Based on optimum belt speed and tomato spacing values, the carrying time on the conveyor belt for 1 tomato to be sorted was 1.43 s. As a result, the throughput capacity of the system was 2517 tomatoes h^{-1}.

In this study, an image processing technique was developed to sort tomatoes according to 4 quality criteria: maturity, defects, shape, and size. The software developed in this study evaluated tomato shape by its eccentricity, tomato size by its 2-D image area, tomato maturity by its mean color, and tomato defect by its fullness parameter. An experimental sorting system equipped with machine vision was constructed to test the ability of the software to sort tomatoes under 3 operational conditions: belt conveyor speed, tomato spacing, and light intensity. After optimum operating conditions were defined, the sorting machine was used to separate tomato samples according to their shape, color, size, and defects. The evaluation of experimental data indicated that sorting accuracy changed with the quality criteria considered, but overall accuracy was remarkably high (90.61%).

References

Blasco J, Aleixos N, Cubero S, Gómez SJ, Moltó E (2009) Automatic sorting of Satsuma (*Citrus unshiu*) segments using computer vision and morphological features. Comput Electron Agr 66: 1–8.

Gonzalez RC, Woods RE (2002) Digital Image Processing, 2nd ed. Prentice Hall, Inc., New Jersey.

Jarimopas B, Jaisin N (2008) An experimental machine vision system for sorting sweet tamarind. J Food Eng 89: 291–297.

Lino LAC, Sanches J, Dal FMI (2008) Image processing techniques for lemons and tomatoes classification. Bragantia Campinas 67: 785–789.

van Assen HC, Egmont PM, Reiber JC (2002) Accurate object localization in gray level images using the center of gravity measure: accuracy versus precision. IEEE T Image Process 11: 1379–1384.

Velioglu SY, Mazza G, Gao L, Omah BD (1998) Antioxidant activity and total phenolics in selected fruits, vegetables, and grain products. J Agr Food Chem 46: 4113–4117.

Weisstein E (2011) Eccentricity. A Wolfram Web Resource. MathWorld, Wolfram Research Inc., available online at http://mathworld.wolfram.com/Eccentricity.html.

Yud RC, Kuanglin C, Moon KS (2002) Machine vision technology for agricultural applications. Comput Electron Agr 28: 173–191.

Zhang Y, Yin X, Zou X, Zhao J (2009) On-line sorting maturity of cherry tomato by machine vision. IFIP AICT 295: 2223–2229.