Thresholding of tart-cherry image frame under uncontrolled daylight conditions

M.B. Lak

Young Researchers Club, Science and Research Branch of Islamic Azad University, Tehran, Iran


In order to detect a fruit, a well-defined segmentation algorithm is required. The segmentation process requires finding the adequate thresholds in order to extract the wanted features. In this project 15 methods were applied to determine the thresholds of thirty images in order to detect the tart-cherry fruits. Then, the resulted binary images were assessed. The images were acquired under uncontrolled daylight conditions and the main idea was to find at least one fruit feature in each image. All of the algorithms were based on image gray-levels, i.e. R (red), G (green), and B (blue). Maximum accuracy (90%) was obtained by means of two algorithms which both were based on division of standard deviation of R and G values on the mean of i, meanwhile the minimum accuracy (0%) was based on division of mean of i on 255.

Key words: image segmentation, natural luminance, tart-cherry

Introduction

Image segmentation describes the process of dividing an image into non-overlapping, connected image areas (Koschan and Abidi, 2008). The goal of image segmentation is to extract meaningful objects from an input image (Tao et al., 2003) and thresholding is a technique for segmentation of colored or grey scaled images based on the color or grayscale value, which transforms an image into a binary image by transforming each pixel according to whether it is inside or outside a specified range (Huang and Chau, 2008).

Thresholding is a fundamental technique applied in many image processing applications and many relatively simple and computationally effective algorithms have been developed and used for change detection in video sequences (Rosin and Ioannidis, 2003). Tao et al. (2003) presented a three-level thresholding method for image segmentation, based on probability partition, fuzzy partition and entropy theory. Bazi et al. (2007) presented a novel parametric and global image histogram thresholding method and Huang

*Corresponding author: M.B. Lak; e-mail: mbagher_lak@yahoo.com

Application of image processing based methods in agricultural activities has been developed for years. The applications involve activities such as auto-guidance (Benson et al., 2003; Han et al., 2004), weed control (Nieuwenhuizen et al., 2007; Ghazali et al., 2009), and post harvest (Naderiboldaji et al., 2008; Jahromi et al., 2008). One of the most important applications of image processing is to detect fruits for operations such as, robotic harvesting (Lak, 2010; Bulanon and Kataoka, 2010), ripeness assessment (Junkwon et al., 2009), and yield mapping (Chinchuluun, 2007) which are done under natural daylight conditions and there is no control on the luminance when the images are acquired. Therefore, different thresholds would be required to segment image frames.

The paper suggested coefficient of variation of each gray-scale image as the threshold of that gray-scale image. Then, different methods based on image grayscales means and standard deviations were assessed in order to find thresholds of tart-cherry image frames. The thresholds are required to segment images into binary form. The main idea of this paper was to use variables originated from RGB images to estimate the thresholds.

**Materials and methods**

Thirty digital images were obtained under natural luminance. Image frames were 3072 × 2304 pixels in the JPEG format. A digital camera (Sony, DSC-H5, Color CCD Camera) was used to acquire the RGB images. One of the images is shown in Fig. 1.

![Fig. 1. Original image in RGB color space (rgb-image).](image-url)
The images were divided into their gray-level images \((R, G, \text{ and } B)\) and \(i\) which calculated by Eq.1 (Fig. 2). Descriptive statistics of means and standard deviations of gray-level images are listed in Table 1.

\[
i = 3 \times I = 3 \times \frac{r + g + b}{3} = r + g + b (1)
\]

Where:

\(I = \) image intensity; \(r = \) first gray-level of original image in \(RGB\) color space (red); \(g = \) second gray-level of original image in \(RGB\) color space (green); \(b = \) third gray-level of original image in \(RGB\) color space (blue).

![Fig. 2. Gray-level images](image-url)
Table 1. Descriptive statistics of mean and standard deviation of grayscales

<table>
<thead>
<tr>
<th></th>
<th>(m)</th>
<th>(m_1)</th>
<th>(m_2)</th>
<th>(m_3)</th>
<th>(s)</th>
<th>(s_1)</th>
<th>(s_2)</th>
<th>(s_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td>52.38</td>
<td>36.21</td>
<td>31.19</td>
<td>55.10</td>
<td>33.95</td>
<td>20.21</td>
<td>20.47</td>
<td>50.87</td>
</tr>
<tr>
<td>Min</td>
<td>177.28</td>
<td>79.38</td>
<td>105.29</td>
<td>42.87</td>
<td>44.29</td>
<td>52.21</td>
<td>55.14</td>
<td>50.67</td>
</tr>
<tr>
<td>Max*</td>
<td>229.66</td>
<td>115.59</td>
<td>136.48</td>
<td>97.97</td>
<td>78.25</td>
<td>72.42</td>
<td>75.60</td>
<td>101.54</td>
</tr>
<tr>
<td>Mean</td>
<td>195.96</td>
<td>97.03</td>
<td>115.62</td>
<td>68.58</td>
<td>66.33</td>
<td>66.99</td>
<td>63.90</td>
<td>64.20</td>
</tr>
<tr>
<td>Std.</td>
<td>14.61</td>
<td>9.47</td>
<td>7.95</td>
<td>14.79</td>
<td>9.41</td>
<td>4.72</td>
<td>5.44</td>
<td>13.73</td>
</tr>
<tr>
<td>C.V.</td>
<td>7.46</td>
<td>9.76</td>
<td>6.87</td>
<td>21.57</td>
<td>14.18</td>
<td>7.75</td>
<td>8.52</td>
<td>21.38</td>
</tr>
</tbody>
</table>

- *Minimum
- **Maximum
- + Standard Deviation
- ++ Coefficient of Variation

\(m, m_1, m_2, m_3, s, s_1, s_2, s_3\) were defined as follow:

- \(m\) = mean of the \(i\)
- \(m_1\) = mean of the \(r\)
- \(m_2\) = mean of the \(g\)
- \(m_3\) = mean of the \(b\)
- \(s\) = standard deviation of the \(i\)
- \(s_1\) = standard deviation of the \(r\)
- \(s_2\) = standard deviation of the \(g\)
- \(s_3\) = standard deviation of the \(b\)

RGB images converted into YCbCr color space (Fig. 3) and the third gray-level of yeber-image \(l\) (Fig. 4) was extracted. Then, image \(t\) was obtained using Eq. 2 (Fig. 5).

\[ t = 3 \times g - 0.5 \times l \quad (2) \]

Where:

\(l\) = third gray-level of the image in YCbCr color space
A “disk” filter was applied on the \( t \) to obtain a noise reduced image (\( t_2 \)) (Fig.6). The filter diameter was defined as the measure of mean of \( i \) in each image frame.

Means of 1\textsuperscript{st}, 2\textsuperscript{nd}, and 3\textsuperscript{rd} gray-levels of rgb-image (\( m_1 \), \( m_2 \), and \( m_3 \)) were extracted, and the standard deviations (\( s_1 \), \( s_2 \), and \( s_3 \)) were calculated. \( m \) and \( s \) demonstrated the mean and standard deviation of \( i \), respectively.

Filtered images (\( t_2 \)) then converted into the binary form. 15 thresholding methods were developed in order to obtain binary form of \( t_2 \). The methods were all based on \( m, m_1, m_2, m_3, s, s_1, s_2, \) and \( s_3 \) and their equations are listed in Table 2.

**Results and discussions**

Fifteen methods were suggested in order to make binary form of the processed images (\( t_2 \)) (Table 2). Best image segmentation resulted from
thresholds of \( bw_{10} \) and \( bw_{11} \) (Fig. 7-8), while \( bw_1 \) resulted in the worst result (Fig. 9). The accuracy of \( bw_{10} \) and \( bw_{11} \) where both 90% while \( bw_1 \) resulted in the accuracy of 0%.

Table 2. Thresholding methods equations and their descriptive statistics

<table>
<thead>
<tr>
<th>Method</th>
<th>Equation</th>
<th>Max.</th>
<th>Mean</th>
<th>Min.</th>
<th>Std.</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( bw_{1} )</td>
<td>( m/255 )</td>
<td>0.9006</td>
<td>0.77</td>
<td>0.6952</td>
<td>0.0573</td>
<td>0</td>
</tr>
<tr>
<td>( bw_{2} )</td>
<td>( m_1/255 )</td>
<td>0.4533</td>
<td>0.3803</td>
<td>0.3113</td>
<td>0.0371</td>
<td>83.33</td>
</tr>
<tr>
<td>( bw_{3} )</td>
<td>( m_2/255 )</td>
<td>0.5352</td>
<td>0.4496</td>
<td>0.4129</td>
<td>0.0312</td>
<td>73.33</td>
</tr>
<tr>
<td>( bw_{4} )</td>
<td>( m_3/255 )</td>
<td>0.3842</td>
<td>0.2686</td>
<td>0.1681</td>
<td>0.058</td>
<td>80</td>
</tr>
<tr>
<td>( bw_{5} )</td>
<td>( s/255 )</td>
<td>0.3069</td>
<td>0.2625</td>
<td>0.1737</td>
<td>0.0369</td>
<td>83.33</td>
</tr>
<tr>
<td>( bw_{6} )</td>
<td>( s_1/255 )</td>
<td>0.284</td>
<td>0.2383</td>
<td>0.2048</td>
<td>0.0185</td>
<td>80</td>
</tr>
<tr>
<td>( bw_{7} )</td>
<td>( s_2/255 )</td>
<td>0.2965</td>
<td>0.2494</td>
<td>0.2162</td>
<td>0.0213</td>
<td>80</td>
</tr>
<tr>
<td>( bw_{8} )</td>
<td>( s_3/255 )</td>
<td>0.3982</td>
<td>0.2308</td>
<td>0.1987</td>
<td>0.0538</td>
<td>80</td>
</tr>
<tr>
<td>( bw_{9} )</td>
<td>( s/m )</td>
<td>0.4254</td>
<td>0.3691</td>
<td>0.1929</td>
<td>0.0671</td>
<td>76.67</td>
</tr>
<tr>
<td>( bw_{10} )</td>
<td>( s_1/m )</td>
<td>0.3917</td>
<td>0.3228</td>
<td>0.2358</td>
<td>0.0394</td>
<td>90</td>
</tr>
<tr>
<td>( bw_{11} )</td>
<td>( s_2/m )</td>
<td>0.4119</td>
<td>0.335</td>
<td>0.2435</td>
<td>0.0419</td>
<td>90</td>
</tr>
<tr>
<td>( bw_{12} )</td>
<td>( s_3/m )</td>
<td>0.5477</td>
<td>0.3162</td>
<td>0.2296</td>
<td>0.0823</td>
<td>83.33</td>
</tr>
<tr>
<td>( bw_{13} )</td>
<td>( s_1/m_3 )</td>
<td>0.7494</td>
<td>0.6491</td>
<td>0.4947</td>
<td>0.0737</td>
<td>16.67</td>
</tr>
<tr>
<td>( bw_{14} )</td>
<td>( s_2/m_3 )</td>
<td>0.6664</td>
<td>0.5597</td>
<td>0.406</td>
<td>0.0575</td>
<td>30</td>
</tr>
<tr>
<td>( bw_{15} )</td>
<td>( s_3/m_3 )</td>
<td>1.3498</td>
<td>0.9762</td>
<td>0.5645</td>
<td>0.1619</td>
<td>5.56</td>
</tr>
</tbody>
</table>

Fig. 7. Image \( bw_{10} \)  
Fig. 8. Image \( bw_{11} \)
Between the methods, \(bw_{10}\) and \(bw_{11}\) (Table 2) were the best. The methods were based on \(R\) and \(G\) standard deviations and the mean of the \(i\). Therefore, \(R\) and \(G\) gray-levels are more effective than \(B\) when an image frame of tart-cherry is analyzed. According to Table 1, measures with fewer coefficient of variation resulted in more reliable thresholds. While, fewer coefficients of variations of means and standard deviations of gray-levels may help to find more reliable thresholds.

The worst results (accuracy of 0% and 5.56%) obtained from \(bw_{1}\) and \(bw_{15}\), respectively. They were based on division of the mean of the \(i\) on 255 (and the mean of \(b\) on standard deviation of \(b\), respectively.

The error of 10% in \(bw_{10}\) and \(bw_{11}\) may be due to the uncontrolled luminance conditions. Daylight conditions make shiny and dark regions in the image which lower the algorithm accuracy. Changing color space and filtering reduced the images features and resulted in noise reduced images to obtain better segmentation.

Tao et al. (2003) mentioned that their proposed method gave good performance. Xue and Titterington (2010) demonstrated that their two methods (One approach was an extension of Otsu’s method, and the other was an extension of Kittler and Illingworth’s MET method) could accomplish more robust performance than that of the originals. The main disadvantage of the proposed technique by Bazi et al. (2007) is the higher computational time required compared to standard algorithms (which depends on the genetic algorithm used for the identification of the best initial conditions). While the method developed here is low-time consuming, simple, and relatively with high-accuracy.

**Fig. 9. Image bw1**
Conclusion

Thirty images were acquired. The images color space was changed in to YCbCr. Then a mathematical equation \( t = 3 \times g - 0.5 \times i \) was developed. \( t \) was filtered by a disk filter which diameter was \( m \). Fifteen methods were developed and used to thresholding the images when they were converted into binary form. Two methods were better than the others. They were based on \( R \) and \( G \) standard deviations and mean of \( i \). Their accuracy was both 90%. It is suggested to more focus placed on developing methods based on \( R \) and \( G \) gray-level images. Fewer coefficients of variations of means and standard deviations of gray-levels also help find more reliable thresholds. For further development, it is suggested to develop three-level thresholding based on maximum fuzzy entropy and genetic algorithm but the required time for the purposed methods must be reduced.

References


(Published in July 2012)