Binary Adaptive Histogram Equalization Segmentation for Light Spot Detection in Tomato Maturity Classifications

Kartika Candra Kirana, Izzati Alvandiar, Muhamad Ainurrahman
Informatic Engineering
Universitas Negeri Malang
Malang, Indonesia
kartikacandra.k@gmail.com

Suastika Yulia Riska
Informatic Engineering
STMIK ASIA Malang
Malang, Indonesia
suastikayr@gmail.com

***Abstract*—Maturity classification of tomato is influenced by color. However, tomato texture which is smooth and glossy often leads production of light spots on tomato image acquisition whereas light spots reduce the accuracy of classification. In this research, we proposed Binary Adaptive Histogram Equalization (BAHE) for light spot detection in tomato segmentation. In this algorithm, some grey level is grouped into one grey level using histogram at relatively homogeneous regions of an image. Then the grey level is converted become binary image using Otsu Algorithm. Experiments were conducted using 24 tomato images. Accuracy of classification is used to evaluate performance of BAHE. Accuracy of classification which use BAHE algorithm is 100% and BHE is 83%. This experiment showed that BAHE outperform for the detection of light spot on tomato image.**

Keywords—binary adaptive histogram equalization, light spot, segmentation, tomato.

I. INTRODUCTION

Tomatoes are one of export commodities on international market [1]. Tomatoes which were chosen should have the appropriate maturity. Therefore, the classification of tomatoes requires decision making of maturity accurately whereas traditional classification which is performed by humans often leads misclassification. Hence, the development of a machine vision system (MVS) becomes essential technology to overcome aforesaid influences [2].

On machine vision systems, the first step is image acquisition. Tomato texture which is smooth and glossy often leads production of light spots on tomato image acquisition. Light spots are represented by white color close the actual color of tomatoes. Meanwhile tomato maturity classification is influenced by color. Therefore, the light spots can degrade the accuracy of classification [3].

Many algorithms were developed to detection spot. Tang et al (2009) proposed improved iteration centroid algorithm to increase light-spot location accuracy [4]. Some times light spots are irregular so the function is unknown and difficult computational processing. Agreal et al (2009) proposed color index analysis for detection of eggshell defect such as blood spots, dirt stains and cracks [5]. Xiaong et al (2011) proposed an adaptive threshold segmentation method based on BP neural network for paper defect detection such as dark spot, light spot and hole [6]. However, this algorithm is not suitable for few training data. Xiang proposed segmentation method based on normalized R-G for recognize light spot, although error segmentation would happen in some conditions [7].

The aim of this research was to develop Binary Adaptive Histogram Equalization (BAHE) for detection of light spot at tomato which is captured in outdoor. In this algorithm, some grey level is grouped into one grey level using histogram at relatively homogeneous regions of an image. Then the grey level is converted become binary image using Otsu Algorithm.

The rest of this paper is organized as follows. In section 2, we give the literature of standard BAHE. In section 3, we explain our research methodology about BAHE (data, proposed method, experiment and evaluation). We also give our result and discussion in Section 4. The section 5 concludes the result.

II. LITERATURE REVIEW

A. *Tomato Maturity Classification*

One of the aspects that affect the quality of tomatoes is by the level of maturity. Maturity level is determined by the color. There are six levels of maturity shown in Figure 1

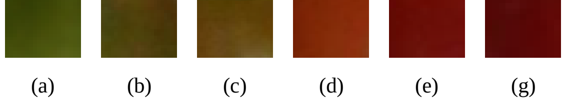

Fig. 1. Level of Tomato Maturity Color (a) Level 1 (b) Level 2 (c) Level 3 (d) Level 4 (e) Level 5 (f) Level 6 [3]

Tomato texture is smooth and glossy. This resulted in the light-spot when the tomatoes were captured outdoors. As with cherries, image capture resulted in the light-spot. Research conducted by Wang, et al (2012) is to determine the level of maturity of cherries based image enhancement of the light-spot on the cherries peel. Thus, produces accurate color rating. The enhancement affect the color rating results conducted outdoors. Before determining the level of maturity of the tomatoes by color, it is necessary to enhance the image.

B. Binary Adaptive Histogram Equalization (BAHE)

Binary Adaptive Histogram Equalization is development of Binary Histogram Equalization (BHE). Histogram equalization is one of the means used to perform image enhancement [10]. The purpose of image enhancement is to generate visual information from images that have been processed [10]. The main idea of this algorithm is grouping of some grey level into one grey level using histogram at relatively homogeneous regions of an image. BHE is used to improving the appearance of a poor image. That is similar of a histogram stretch, but often provides more visually pleasing results across a wider range of image[14]. The main steps are performed on the BHE, among others, (1) find the value histogram, (2) normalize the histogram value by dividing the total pixel of the image, (3) multiplying the result of normalizing the maximum value of gray level and round, (4) using a one-to-one correspondence to map the gray level [14].

In the example, the image BHE has L gray levels, ie 0 to L-1, and the gray level i appearing as ni times. if n represents a total pixel of the image, then the equations used to transform the gray level in order to contrast possessed a better image shown in equation 1.

$$Transform = \left(\frac{n_0 + n_1 + \cdots + n_i}{n}\right)(L-1) \quad (1)$$

The results obtained from the equation 1 is rounded to the nearest integer [16].

In the digital image processing, adaptive histogram equalization is used to enhance the contrast of an image. Methods used to calculate the histogram with adaptive histogram equalization is by calculate the contextual region of pixel image [11]. The main steps are performed on the BAHE, among others, (1) measuring the grid of image based on maximum dimensions, (2) defaulf the size of of the window by the size of grid, if the window size is not predetermined, (3) determine the grid point on the image that starts from the left corner of the image, so that each grid point separated the size of a predetermined pixel grid, (4) calculate the area around it for each grid point that has been determined, (5) find the four nearest grid points surrounding the pixel, (6) using a pixel intensity value as the index, so it can find the four neighboring grid points based on cumulative distribution function (cdf), (7) interpolate these values to get the value of the new location, (8) looping stage 5 to 7 at each image pixel.

By using BAHE an image is divided into several sections, which some gray color grouped into one value using a histogram. BAHE illustration shown in figure 1.

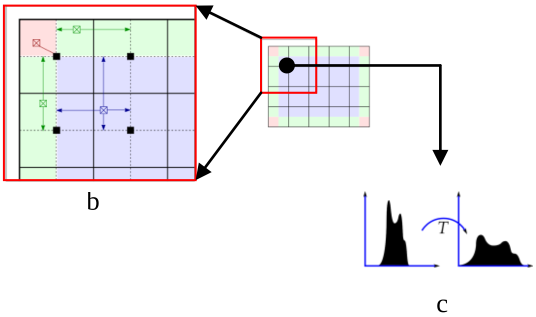

Fig. 2. Illustration of BAHE (a) Image that devide into some grid (b) Zoom in of grid (c) histogram is equalized

Based on figure 1, the conversion value of several gray-level into one value of gray-level using the equation shown in equation 1. In Figure 1 (c) there is a clip limit, the state of the histogram is clipped, which can limit the slope of the cdf by using the transformation function. The value of the clip limit depends on the normalization of the histogram and the size of the neighborhood region.

C. Otsu Algorithm

The basic concept using Otsu method is to find a threshold that minimizes weight class variance. In which Otsu method divides the gray level histogram of the image into two different areas automatically to insert threshold value. Thus, it can maximize the between class variance [8]. Otsu method approach is to conduct a discriminant analysis is to determine the variables that can distinguish between two or more groups emerging. Discriminant analysis will maximize these variables in order to divide the object, ie the foreground and background [9].

Assumptions of Otsu method among others bimodal histogram of the image, do not use the spatial coherence as well as other ideas of object structure, in the form of stationary statistics but can be modified into a local adaptive, and there is an equal illumination. So that the bimodal brightness can be derived from the difference in the appearance of the object [8]. The weighted within-class variance can be calculated by equation 2.

$$\sigma_w{}^2(t) = q_1(t)\sigma_1{}^2(t) + q_2(t)\sigma_2{}^2(t) \quad (2)$$

q_1 and q_2 indicate the probability of occurrence class and class mean levels of pixels t, and σ shows the class variances. Where,

$$q_1(t) = \sum_{i=1}^{t} P(i) \quad (3)$$

and

$$q_1(t) = \sum_{i=1}^{t} P(i) \tag{4}$$

Individual class variances are given by:

$$\sigma_1{}^2(t) = \sum_{i=1}^{t} [i - \mu_1(t)]^2 \frac{P(i)}{q_1(t)} \tag{5}$$

and

$$\sigma_2{}^2(t) = \sum_{i=t+1}^{t} [i - \mu_2(t)]^2 \frac{P(i)}{q_2(t)} \tag{6}$$

,where class means μ_1 and μ_2 indicated by:

$$\mu_1(t) = \sum_{i=1}^{t} \frac{iP(i)}{q_1(t)} \tag{7}$$

and

$$\mu_2(t) = \sum_{i=t+1}^{t} \frac{iP(i)}{q_2(t)} \tag{4}$$

III. Methodology

A. Data

Dataset of this research are 36 tomatoes which were capture using Samsung Tab 2.7. 24 tomatoes are testing data and 12 tomatoes are training data. Example of data is showed at Figure 2.

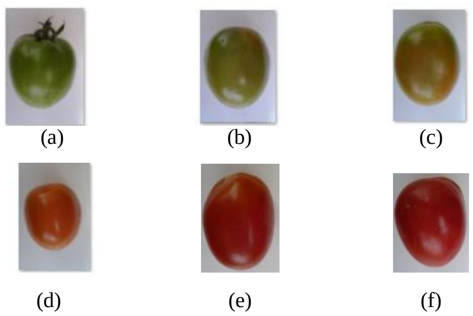

Fig. 3. Example of Data (a) Level 1 (b) Level 2 (c) Level 3 (d) Level 4 (e) Level 5 (f) Level 6

B. Methodology

In this paper, there are 7 main step are conducted by image acquisition, preprocessing, segmentation, feature extraction, feature reduction, and classification. The main processing is showed at Figure 3.

1) Image Acquisition

There are 36 tomatoes was used in this study were capture using Samsung Tab 2.7. Camera is perpendicular to tomato. Tomatoes image was captured in outdoor. Tomato image acquisition is done by the camera perpendicular to the tomatoes. this is done to reduce the shadow effect of the tomato. Captured image is one side of tomatoes which represents the maturity of the tomatoes.

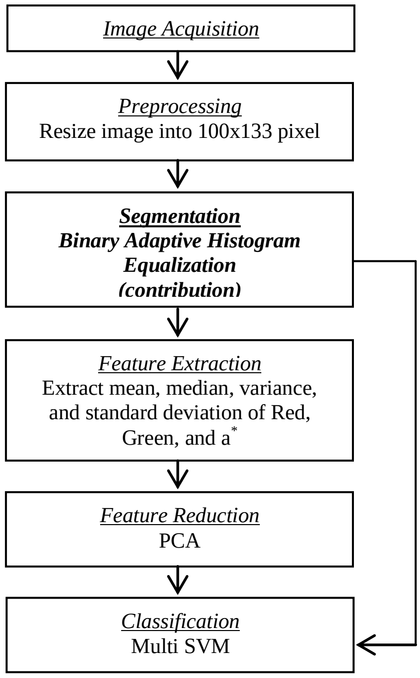

Fig. 4. Main Steps of Tomato Maturity Classification in this Research

2) Pre Processing

Images were resized into 100X133 pixels. It used to speed up the processing of computation.

3) Segmentation using Binary Adaptive Histogram Equalization

The contribution of this paper is propose Binary Adaptive Histogram Equalization (BAHE) for detect the light spot at tomato segmentation. The main idea of BAHE was concluded by: grouping grey level into some level based on histogram at relatively homogeneous regions of an image and grouping some grey level into background and foreground. Foreground is tomato without light spot and background are white yard and light spot. The step of BAHE is showed at Figure 5.

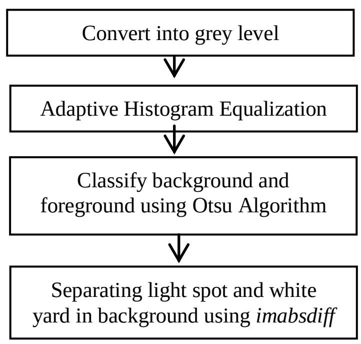

Fig. 5. Steps of Binary Adaptive Histogram Equalization

First, RBG image is converted into grey scale image. Then, grey scale image is contrasted to separate the foreground and background clearly. After that, background and foreground is classified using Otsu. In this paper, foreground is tomato without light spot and background are concluded by white yard and light spot. For next research, we need separate light spot and white yard for repairing the color of tomato which containing light spot so in this research we separate light spot and white yard using *imabsdif. Imabsdif* is used to subtract the reverse of tomato without light spot from the corresponding element in all of tomato region for getting light spot.

4) Feature Extraction

After detect light spot from tomato, 12 feature is extracted which is included by mean, median, variance, and standard deviation of red, green, and a* from foreground (tomato without light spot). These features are used because they are color features whereas tomato classification is influenced by color features [3] [17].

5) Feature Reduction

12 features are reduced into 6 features using PCA. Principal component analysis (PCA) is an efficient algorithm to reduce the dimensionality of a data set consisting of a large number of interrelated variables while retaining the most significant variations [18].

6) Classification

Tomato maturity is classified using Support Vector Machines (SVM). SVM are classification methods based on machine learning theory which have significant advantages because of their high accuracy mathematical tractability. It also does not need a large number of training samples to avoid over fitting [12]. Accuracy of classification is calculated by Equation 1.

$$Accuracy = \frac{TP+FP}{TP+TN+FP+FN} \tag{1}$$

Accuracy is counted by True Positive (TP), TN True Negative (TN), FP (False Positive), and FN (False Negative). TP and FP is the tomato which has the same class between estimate class and actual class. TN and FN is the tomato which has the different class between estimate class and actual class. In this paper, we calculate 6 classes, so we can to modify Equation 1 become Equation 2.

$$Accuracy = \frac{True\ classification}{All\ of\ data} \tag{2}$$

IV. RESULT AND DISCUSSION

. The main steps performed for the classification of tomato maturity shown in Figure 6.

At this research, classification is used to perform testing on BAHE algorithm. Testing is done by comparing the classification results of segmentation using BAHE and BHE. Classification results are shown in Table 1.

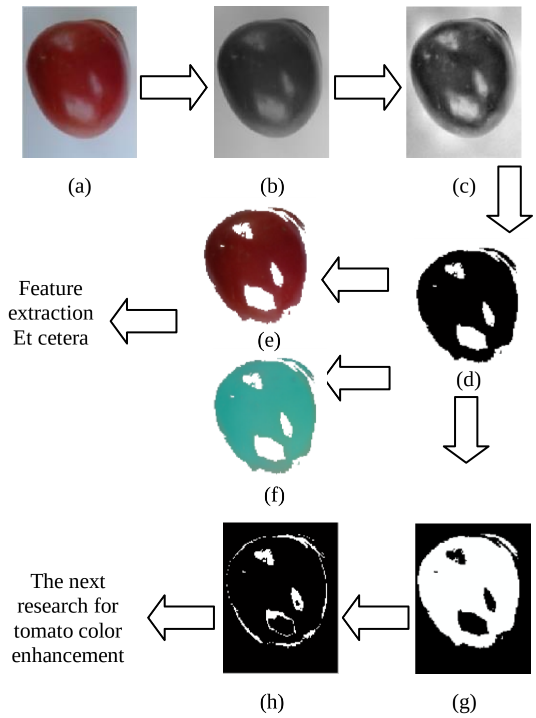

Fig. 6. Implementation of BAHE algorithm (a) RGB image (b)grayscale image (c) enhance using Adaptive Histogram Equalization (d) Threshold using Otsu (e) extract Red and Green feature (f) extract a* feature (g) reversce of tomato without light spot (h) light spot

TABLE I. TOMATO MATURITY CLASSIFICATION RESULT

Segmentation Methods	True Classification	False Classification	Percentage of Accuracy
Binary Histogram Equalization	20	4	83%
Binary Adaptive Histogram Equalization	24	0	100 %

Based on Table 1, the results of tomato maturity classification using segmentation with BAHE showed higher accuracy than the segmentation by BHE. The difference is due to the segmentation of BAHE can divide the gray-level between light-spot and tomatoes area better than BHE. Comparison of the results of segmentation between BAHE and BHE shown in figure 7.

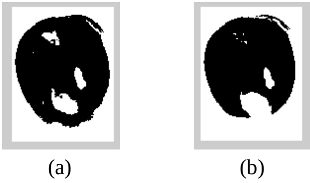

Fig. 7. Region Segmentation (a) Binary Adaptive Histogram Equalization (b) Thresholding of Histogram Equalization

Based on figure 7, the segmentation results of the BAHE better than BHE. Because the grouping of BAHE gray level of histogram calculations are based on relatively homogeneous regions of an image, that can generate new gray level sharper than the BHE. Before thresholding, that can change black-white images, the gray levels can be shown by comparing the histogram BAHE and BHE. Thus, the results obtained are histogram on BAHE more pointed than the BHE. Comparison histogram is shown in figure 8.

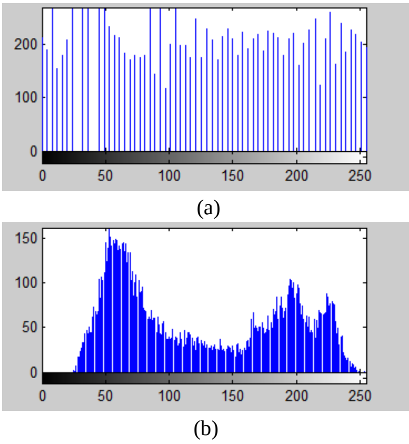

Fig. 8. Histogram before Thresholding (a) Binary Adaptive Histogram Equalization (BAHE) (b) Binary Histogram Equalization (BHE)

V. CONCLUSSION

Tomato is one of the fruits that have a rapid maturation process. Tomatoes have a smooth and glossy texture. When arrested outside the emerging light-spot on the image. Tomato maturity classification is influenced by color. The existence of light-spot would reduce the accuracy of the classification. Thus, this research conducted by applying image enhancement Bahe to detect light-spot on the image of tomatoes. Steps taken in this study include image acquisition, preprocessing, segmentation, feature extraction, feature reduction, and classification. In this research, segmentation by using Bahe better than the BHE, because the histogram is generated by using Bahe more pointed than the BHE. Testing is done by comparing the accuracy of the classification of image segmentation using BAHE and BHE. The accuracy of tomato maturity classification by using BHE is 83.3% and BAHE is 100%.

For the next research related to improving the image that is to repair the color of tomatoes image on light-spot.

REFERENCES

[1] Hanadita, Nisa. “Analisis Eksport Tomat Segar di Indonesia”. Bandung: IPB, 2008.

[2] Savakar, Dyanand., “Identification and Classification of Bulk Fruits Images using Artificial Neural Networks”, March 2012, vol.1, no.3, 36-41.

[3] Riska, S.Y., Ginardi, R.V.H., Suciati, Nanik. “Segmentasi Variasi Pencahayaan Citra Tomat Menggunakan Marker Controlled Watershed dan Arimoto Entropy untuk Perbaikan Citra” Jurnal Buana Informatika, vol. 6, no. 3, 2015, pp. 213-224.

[4] Tang,Shengjin., Zhou, Zhaofa., Guo, Xiaosong., Xiao, Yongchao., “Improved iteration centroid algorithm based on linear CCD light-spot location”, IEEE, 2009, pp. 4.38-440.

[5] Alegre, M.C.G., Riberio, Angela., Guinea, Domingo., Cristobal, Gabriel., “Color Index Analysis for Automatic Detection of Eggshell Defect”, The International Society For Optics And Photonics, 2009.

[6] Xiaong, Wang., Qinghua, Li., Jun, Li., “An Adaptive Threshold Segmentation Method based on BP Neural Network for Paper Defect Detection” IEEE, 2011, pp. 405 – 408.

[7] Xiang,Rong., Ying, Yibin., Jiang, Huanyu., “Research on Image Segmentation Methods Of Tomato In Natural Conditions”, IEEE, 2011, pp. 1268 – 1272.

[8] Otsu, Nobuyuki. “A Threshold Selection Method from Gray-Level Histogram” IEEE, 1979, pp. 62 – 66.

[9] Mizushima, Akira. Lu, Renfu. “An Image Segmentation Method for Apple Sorting and Grading Using Support Vector Machine and Otsu’s Method”. Elsevier, 2013, pp.29 – 37.

[10] Sasi, Neethu M. Jayasree,V.K. 2013. Contrast Limited Adaptive Histogram Equalization for Qualitative Enhancement of Myocardial Perfusion Images. SciRes, Engineering 326-331.

[11] Pisano, Etta D. Zong, Shuquan. Hemminger, Bradley M. Deluca, Maria. Johnston, R. Eugene. Muller, Keith. Braeuning, M. Patricia. Pizer, Stephen M. 1998. Contrast Limited Adaptive Histogram Equalization Image Processing to Improve the Detection of Simulated Spiculations in Dence Mammograms. Journal of Digital Imaging, Vol 11, No 4, pp 193-200.

[12] Wang, Xinzhong., Mao, Hanping., Han, Xu., Yin, Jianjun. “Vision-based Judgment of Tomato Maturity Under Growth Conditions”. African Journal of Biotechnology, 2011, pp. 3616 – 3623.

[13] Wang, Qi. Wang, Hui. Xie, Lijuan. Zhang, Qin. “Outdoor Color Rating of Sweet Cherries Using Computer Vision”. Elsevier, 2012, pp. 113 – 12.

[14] Umbaugh, Scott E. “Digital Image Processing and Analysis”. CRC Press, Taylor & Francis Group, 2011.

[15] Pizer, Stephen M. Amburn, E Philip. Austin, John D. Cromartie, Robert. Geselowits, Ari. Greer, Trey. Romeny, Bart Ter Haar. Zimmerman, John B. Zuiderveld, Karel. “Adaptive Histogram Equalization and Its Variations”. Computer Vision, Graphics, and Image Processing. 1987, pp.355 – 368.

[16] Sianipar, R.H. 2013. “Pemrograman MATLAB Dalam Contoh dan Terapan”. Bandung: Informatika.

[17] Deepa, P., Geethalakshmi, S.N., “A Comparative Analysis of Feature Extraction Methods for Fruit Grading Classfications”

[18] Zang, Yudong., Wu, Lenan., “Classification of Fruits Using Computer Vision and a Multiclass Support Vector Machine” Sensor, 2012. vol. 12, pp. 2489-12505