Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information). 7/2 (2014), 83-89
DOI: http://dx.doi.org/10.21609/jiki.v7i2.261

MULTITHRESHOLDING IN GRAYSCALE IMAGE USING PEAK FINDING APPROACH AND HIERARCHICAL CLUSTER ANALYSIS

Cahyono, Gigih P., Aprilio, Adrianus Y., and Ramadhan, Hani

Informatics Engineering Department, Faculty of Information Technology, Institut Teknologi Sepuluh Nopember, Jalan Raya ITS Kampus ITS Sukolilo, Surabaya, 60111, Indonesia

E-mail: gigih.prasetyo13@mhs.if.its.ac.id,yoza13@mhs.if.its.ac.id, hani13@mhs.if.its.ac.id

Abstract

Image segmentation is typically used to distinguish objects that exist in an image. However, it remains difficult to accommodate favourable thresholding in multimodal image histogram problem with specifically desired number of thresholds. This research proposes a novel approach to find thresholds in multimodal grayscale image histogram. This method consists of histogram smoothing, identification of peak(s) and valley(s), and merging process using hierarchical cluster analysis. Using five images that consisted of grayscale and converted-to-grayscale images. This method yields maximum value of accuracy, precision, and recall of 99.93%, 99.75%, and 99.75% respectively. These results are better than the similar peak finding method in multimodal grayscale image segmentation.

Keywords: *segmentation, multithresholding, peak finding, merging, cluster analysis*

Abstrak

Salah satu penggunaan segmentasi citra adalah membedakan objek-objek yang ada dalam suatu citra. Namun, untuk mengakomodasi suatu metode penentuan nilai ambang yang diinginkan dalam histogram multimodal citra masih sulit dilakukan. Maka dari itu, penelitian ini memberikan suatu pendekatan baru untuk menentukan nilai ambang dalam histogrammultimodal citra keabuan. Metode ini terdiri dari penghalusan histogram, identifikasi lembah dan puncak, dan proses penggabungan dengan analisis kluster hierarkis. Metode ini diuji dengan lima citra keabuan dan citra warna yang dikonversi ke citra keabuan dan menghasilkan nilai maksimum dari akurasi, presisi, dan *recall* masing-masing 99,93%, 99,75%, dan 99,75%. Hasil ini lebih baik daripada metode segmentasi citra keabuan dengan penentuan puncak yang mirip dengan metode dalam penelitian ini.

Kata Kunci: *segmentasi, multithresholding, pencarian puncak, penggabungan, analisis kluster*

1. Introduction

Image thresholding has been a challenging concept in image processing, especially in image segmentation. One of its usages is defining different objects in image and the background of image based on the color similarity. The main concept of image thresholding is composed by searching the best value of color intensity that able to distinguish a region of similar color and the different color. Although we have been discussing about colors before, image thres-holding is also important in grayscale images. The automatic object recognition in grayscale images is necessarily used because of hardly normal human observation. The grayscale image covers the cases of text recognition [1] or several other cases reviewed in literature by Chi, et al [2].

A research by Otsu[3], which has been popularly used until now, proposed a simple method based on the variances between and within so-called foreground class and background class to find the best value to differentiate an object and its background in a grayscale image. But, it lacked in cases of multimodal type of grayscale intensity distribution of the grayscale color intensity histogram. Otsu's approach of solving the problem of multimodal grayscale intensity histogram refers to longer execution time and more memory consumption.

A further approach by Papamarkos and Gatos [1] examined the hill-clustering process of grayscale intensity histogram, and then linearly approximated the histogram segments, and lastly used a one-dimensional search method to find the values of applicable thresholds. This method gave better differentiation by multilevel thresholds in multimodal histogram images, which compared by Otsu [3], Kapur [4], and Reddi [5]. Meanwhile, this method still utilized complex mathematical approach to find the thresholds and cannot adapt to flexible multithresholding on a multimodal image histogram in a desired number of thresholds.

This paper proposes a novel approach to solve the multithresholding problem in more flexible way by taking histogram smoothing method, peak finding, inter-class variance, and hierarchical analysis, inspired by method of Arifin and Asano [6] into account. Peak finding is considered to be a better approach through the assumption that a threshold always resides between two peaks [7, 8] of mode in image histogram. Then, this method proposes a new approach in adaptability using merging technique through hierarchical cluster analysis, which happ-ens when the number of discovered threshold is more than the desired threshold.

This paper is composed as follows: second section reviews the algorithms and their basic concepts, third section presents the general implementation of the method and displays the comparison of result between this method and the two other methods is that Papamarkos and Gatos and Felzenszwalb and Huttenlocher's segmen-tation method, and last section concludes our method in addition of the discussion and further works.

2. Methods

Image Thresholding

Earlier part of this section describes about the image thresholding in multimodal image case. Then, the next part explains about histogram smoothing, peak finding, peaks and valleys identification, the proposed significant peaks identification, and the merging by hierarchical cluster analysis using the between class variance.

Image thresholds are generally found through the valley between two peaks of adjacent modes in a histogram [7, 8] of color intensity value of an image. In this paper, the approach to get the best value of valleys, which assumed as the thresholds of multimodal image, starts from the smoothed histogram. The smoothing is used because the original histogram of multimodal image contains noises which deliver to high and unnecessary computation [9, 10] to find a good threshold between them. It is possible that a valley between two peaks of mode is not a good threshold. Illustrated in Figure 1, the valleys T_1, T_3, and T_5 of an multimodal image histogram don't give a significant result if they are considered as threshold since they don't separate the modes well.

To find the good thresholds instead of unnecessary valleys, first, this paper finds all the possible valley(s) and peak(s), including unnecessary ones of smoothed histogram through the derivatives of the values, which can be calculated by the difference of adjacent value of grayscale intensity.

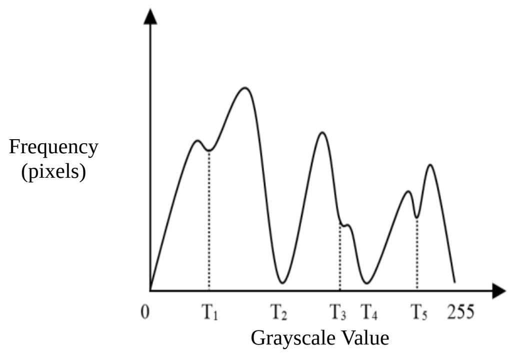

Figure 1. The illustration of the unnecessary valleys in multimodal image histogram.

After that, a mechanism of peak finding based on gradient change and identi-fication is used to get more robust peaks. Then, a value between two peaks, so-called valley, is decided as a threshold by iterating all the values between two peaks to differentiate two modes.

After a histogram of a grayscale image is extracted, the peak finding algorithm is executed, which starts from histogram smoothing, then peaks and valleys identification. The histogram smoothing is described as follows.

Histogram smoothing is used to remove rough curves, which considered as noises, which happen in the original histogram. The rough curves can yield into numerous valleys and peaks. Then, a simple averaging is used to acquire $\boldsymbol{h}'(\boldsymbol{i})$, a smooth histogram $\boldsymbol{h}'$ at level $\boldsymbol{i}$ from five original histogram value $\boldsymbol{h}(\boldsymbol{i}-\boldsymbol{2})$, $\boldsymbol{h}(\boldsymbol{i}-\boldsymbol{1})$, $\boldsymbol{h}(\boldsymbol{i})$, $\boldsymbol{h}(\boldsymbol{i}+\boldsymbol{1})$, and $\boldsymbol{h}(\boldsymbol{i}+\boldsymbol{2})$ in equation(1).

$$h'(i) = \frac{1}{5}\sum_{j=i-2}^{i+2} h(j) \qquad (1)$$

The denominator is determined using the method of Tan and Isa [11]. The new approach is if the histogram is not smooth enough, as the before and after smoothing process still has a big difference, the smoothing process is continuously executed until the before and after histogram smoothing yields a smooth histogram.

Then, we identify the peaks and valleys through simple comparison of the frequency of grayscale intensity, which ranged from 0 to 255, in the histogram by similar actions of Tan and Isa's. If the frequency of another intensity is higher than its adjacent intensities, it is considered as a peak $\boldsymbol{P}$. If the frequency of an intensity is lower than its adjacent intensities, it is considered as a valley $\boldsymbol{V}$. The requirement of peaks identification are shown in equation(2) and the requirements of valleys identification are shown in equation(3).

$$P = \{(i, h'(i)) | h'(i) > h'(i-1) \cap h'(i) > h'(i+1), 0 \le i \le 255\} \quad (2)$$

$$V = \{(i, h'(i)) | h'(i) < h'(i-1) \cap h'(i) < h'(i+1), 0 \le i \le 255\} \quad (2)$$

Here, $\boldsymbol{h'(i)}$ represents the frequency of smoothed histograms the grayscale intensity $\boldsymbol{i}$, valued 0-255, that appears in the image. In order of appearance, $\boldsymbol{h'(i-1)}$and $\boldsymbol{h'(i+1)}$represent the frequencies of the intensity before and after smoothed histogram at intensity$\boldsymbol{i}$.

Afterthat, a selection method which modifies the fuzzy rule of the research by Tan and Isa [11] is proposed to acquire more robust peaks and valleys identification. A $\boldsymbol{p}$-th peak $\boldsymbol{P_p}$that came from discontinued tail of histogram curve can be removed here because there's no valley after it, or if the intensity of $\boldsymbol{p}$-th peak $\boldsymbol{i_p}$ before intensity of a $\boldsymbol{v}$-th valley $\boldsymbol{i_v}$ shown in (4). And also, if there's $\boldsymbol{p}$-th peak $\boldsymbol{P(p}$that has frequency below a neighboring $\boldsymbol{v}$-th valley $\boldsymbol{V_v}$ and $\boldsymbol{v+1}$-th valley $\boldsymbol{V_{v+1}}$can't be considered as a peak because its frequency $\boldsymbol{h'(p)}$, unfortunately, is below the frequency of those valleys $\boldsymbol{h'(v)}$ and $\boldsymbol{h'(v+1)}$, shown in (5).

$$IF\left(i_p > i_v and P_p is last peak\right) THEN remove P_p \quad (3)$$

$$IF(h'(p) < h'(v) and\ h'(p) <) h'(v+1))\ THEN remove P_p \quad (4)$$

Hence, as discussed before, to find a good threshold, we use the between class variance (BCV), which can be found in Otsu's research [3]. The BCV of every pixel between two peaks of histogram is considered, using the probability histogram $\boldsymbol{Pr}$ that converts frequency into [0, 1] by dividing each frequency of intensity value $\boldsymbol{h'(i)}$ of a smoothed histogram $\boldsymbol{h'}$ in intensity level $\boldsymbol{i}$ by the number of pixel in the image $\boldsymbol{N}$, shown in equation (6) and equation(7).

$$N = \sum_{i=0}^{255} h'(i) \quad (5)$$

$$Pr(i) = \frac{h'(i)}{N}, i \le 0 \le 255 \quad (6)$$

The BCV is computed from sum of the probabilities $\boldsymbol{Pr(i)}$ at any intensity $\boldsymbol{i}$ of two regions $\boldsymbol{\omega_0}$ and $\boldsymbol{\omega_1}$ limited by intensity of two valleys, which are intensity of $\boldsymbol{v}$-th valley $\boldsymbol{i_v}$ and $\boldsymbol{v+2}$-th valley $\boldsymbol{i_{v+2}}$and separated by an intensity of$\boldsymbol{v+1}$-th valley $\boldsymbol{i_{v+1}}$, and then an optimal value of grayscale intensity $\boldsymbol{i_t}$ which has probability $\boldsymbol{Pr(i_t)}$ between two adjacent peaks, let's say $\boldsymbol{p}$-th and $\boldsymbol{p+1}$-th peaks which has probability $\boldsymbol{Pr(i_p)}$ and $\boldsymbol{Pr(i_{p+1})}$ which produces biggest BCV is chosen. A set of BCV is computed through every peaks in image histogram, which has $\boldsymbol{n_P}$ peaks. The process is shown from equation (8) through (10). Those processes iteratively computed as much as the number of peaks existed in the smoothed image histogram, which yields a set of BCV.

$$\boldsymbol{\omega_0} = \sum_{i=i_v}^{i_{v+1}} \boldsymbol{Pr(i)}, \boldsymbol{\omega_1} = \sum_{i=i_{v+1}}^{i_{v+2}} \boldsymbol{Pr(i)} \quad (8)$$

$$\boldsymbol{i_t} = \mathbf{arg}\,\boldsymbol{max}\left(\boldsymbol{\omega_0}\left(\mathrm{Pr}(\boldsymbol{i_p}) - \mathrm{Pr}(\boldsymbol{i_t})\right)^2 + \boldsymbol{\omega_1}\left(\boldsymbol{Pr}(\boldsymbol{i_{p+1}}) - \boldsymbol{Pr}(\boldsymbol{i_t})\right)^2\right) \quad (9)$$

$$\mathrm{BCV} = \{\, \mathrm{p}\,|\omega_0\left(\mathrm{Pr}(\mathrm{i_p}) - \mathrm{Pr}(\mathrm{i_t})\right)^2 + \omega_1\left(\mathrm{Pr}(\mathrm{i_{p+1}}) - \mathrm{Pr}(\mathrm{i_t})\right)^2, \mathbf{p} < \mathbf{n_P} - \mathbf{1}\} \quad (10)$$

The hierarchical cluster analysis happens while the number thresholds found in this approach is more than desired threshold actually. The hierarchical merging starts from the removal of the peak $\boldsymbol{P_x}$ which has the smallest BCV, or we can say $\boldsymbol{x}$-th BCV $\boldsymbol{BCV_x}$, which considers peak $\boldsymbol{P_x}$ of set $\boldsymbol{P}$, and merges it to adjacent classes, left or right, depends on which side that produces the bigger BCV as described in equation(11).

$$P_x = \min(BCV_x), x \in P \quad (7)$$

Finally, the desired threshold is achieved by executing a loop of iterative BCV computation which initially grows from peaks in grayscale image histogram.

3. Results and Analysis

The methods explained before are implemented using MATLAB and several original grayscale and color-converted-to-grayscale images are used for this research. The method is compared by similar peak finding method [1], which uses linear programming. The flow of the algorithm is shown by Figure 2.

The images used in this research are Lena, Cameramen, Rice, Peppers, and Mountain. They are displayed in Figure 3. To evaluate the result of

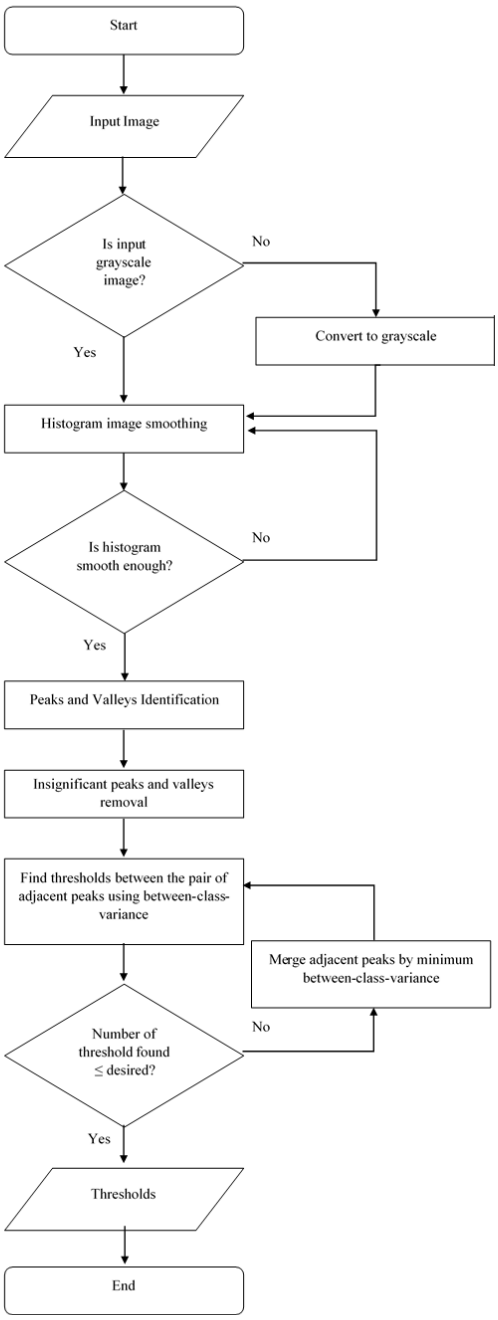

Figure 4. The illustration of the unnecessary valleys in multimodal image histogram.

image segmentation, we use the accuracy, precision and recall of image segmentation, which derived from confusion matrix as illustrated in Figure 4.

It can be seen that there are j true segments of ground truth G and k found segments of classified segments S by the method. If j is more than k, it means that there is a segment from the evaluated method which is not considered as a segment. To choose the unconsidered classified segment, we use the minimum value of two or more

Figure 2. Images used in this research. First row, left-to-right: Rice, Cameraman, Mountain. Second row, left-t

Ground truth segments	Segments found				
	S_1	S_2	S_3	…	S_k
G_1	$C(1,1)$	$C(1,2)$	$C(1,3)$	…	$C(1,k)$
G_2	$C(2,1)$	$C(2,2)$	$C(2,3)$	…	$C(2,k)$
…	…	…	…	…	…
G_j	$C(j,1)$	$C(j,2)$	$C(j,3)$	…	$C(j,k)$

Figure 3. The confusion matrix of image segmentation.

segment which then removed from the evaluation. Otherwise, if j is less than k, the ground truth segment, which has two or more segments classified into it, is forced to choose the segment that has more pixel that correctly classified into it. Each cells $C(x,y)$ in confusion matrix represents the number of pixels that the method classifies into segment x and actually it should be classified into segment y.

The accuracy, precision, and recall are presented in the percentage numbers and respectively can be seen on equation(12), equation(13), and equation(14). If the metric values are higher, the performance of the method is better.

$$Accuracy = \frac{\sum_{x=1}^{j} c(x,x)}{\sum_{x=1}^{j}\sum_{y=1}^{k} c(x,y)} \tag{8}$$

$$Precision = \frac{\sum_{x=1}^{j} \frac{c(x,x)}{\sum_{y=1}^{k} c(x,y)}}{j} \tag{9}$$

$$Recall = \frac{\sum_{y=1}^{k} \frac{c(x,x)}{\sum_{x=1}^{j} c(x,y)}}{k} \tag{10}$$

Figure 5 shows the multimodal grayscale image histogram of each sample images and also the result of the thresholding process using the de-

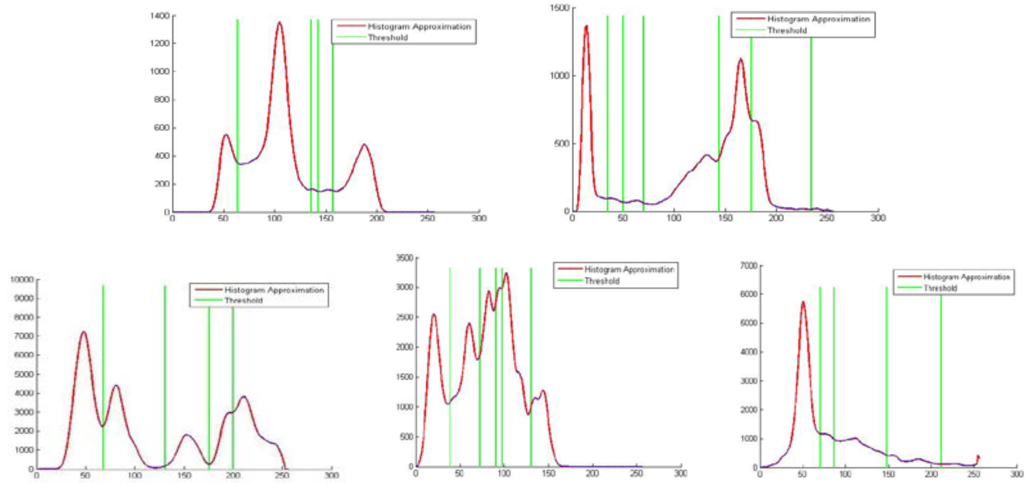

Figure 5. Sample images histogram. First row left-to-right**:** Rice, Cameraman, Mountain. Second row left-to-right: Lena, Pepper.

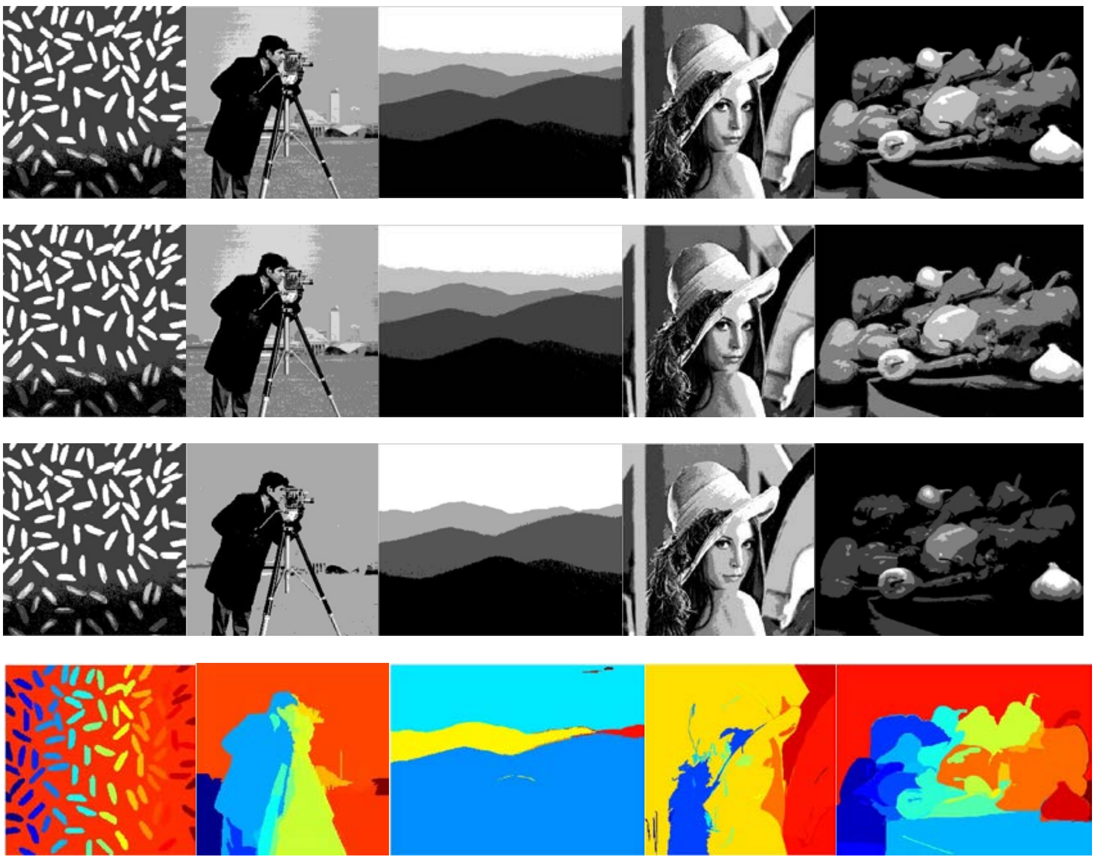

Figure 6. Segmented images used in this research. Left-to-Right: Rice, Cameraman, Mountain,Lena, Pepper. First row: Ground Truth. Second row: Result of proposed method. Third row: Other method[1]. Fourth row: Other method[12].

sired input of thresholds. Even though the thresholds does not seem correctly drives to the valley, which considered the best threshold, it approximates the value of the valley. The result of other method [12] is colored because its output is really in colors. This research compares the result of proposed segmentation method with the ground truth segmentation, Papamarkos and Gatos' segmentation method [1].

The ground truth segmentations results of segmented image and the other method [1] and [12], are presented in Figure 6. The threshold re-

TABLE 1
COMPARISON OF MULTITHRESHOLD VALUES ACQUIRED FROM GROUND TRUTH IMAGES, OUR PROPOSED METHOD OUTPUT, AND OTHER METHOD OUTPUT

Sample Images	Multithresholds Acquired		
	Ground Truth	Proposed Method	Other Method [1]
Rice	67, 134, 144, 160	64, 136, 143, 157	66, 145, 161
Cameraman	33, 53, 78, 141, 177, 233	35, 50, 70, 144, 176, 235	50, 75, 213
Mountain	68, 128, 176, 200	68, 131, 176, 200	68, 115, 176
Lena	36, 69, 88, 97, 127	39, 73, 91, 98, 131	35, 69, 88, 127
Pepper	72, 94, 151, 217	71, 87, 148, 212	98, 171, 219, 239

TABLE 2
COMPARISON OF NUMBER OF SEGMENTS VALUES ACQUIRED FROM GROUND TRUTH IMAGES, OUR PROPOSED METHOD OUTPUT, OTHER METHOD OUTPUT [1] AND [12]

Sample Images	Multithresholds Acquired			
	Ground Truth	Proposed Method	Other Method [1]	Other Method [12]
Rice	5	5	4	80
Cameraman	7	7	4	23
Mountain	5	5	4	5
Lena	6	6	5	38
Pepper	5	5	5	49

TABLE 3
EVALUATION COMPARISON BETWEEN PROPOSED METHOD AND OTHER METHOD [1]

Sample images	Evaluation					
	Accuracy (%)		Precision (%)		Recall (%)	
	Proposed Method	Other method [1]	Proposed Method	Other method [1]	Proposed Method	Other method [1]
Rice	**97,47**	94,53	**87,68**	78,99	**99,31**	98,89
Cameraman	**98,89**	82,08	**96,09**	60,27	**96,09**	59,38
Mountain	**99,93**	95,24	**99,75**	82,79	**99,75**	99,23
Lena	**96,89**	92,95	**90,55**	88,56	90,55	**91,68**
Pepper	**97,68**	94,52	**94,18**	81,89	**94,18**	85,26

sults of our proposed method, other method [1] and the ground truth are presented in Table 1. Table 2 presents the number of segments in the output produced in each segmentation method. The three metrics are used to evaluate the quality of our method compared with the other method in case of the sample images. The performance evaluation is presented in Table 3.

Table 1 shows that the proposed method is more adaptable to do image multithresholding into favorable number of segments of ground truth than other method [1]. The other method can't approach the ground truth's number of threshold of most of sample images, except the pepper image. This shows that our method is more adaptable into determining the number of desired threshold even though by input. Initially it finds more or equal number of thresholds than the desired thresholds and merges them into equal number of thresholds. Meanwhile, the other method automatically acquires the thresholds but it drives to different number of thresholds comparing to the ground truth. On the other hand, the method of graph-base segmentation [12] are not presented because it does not produce the threshold for peak finding methods.

Table 2 shows the number of segments produced by every segmentation method tested in this research with the ground truth. Our proposed method results a completely exact number of segment within the ground truth. Meanwhile, the other method by Papamarkos and Gatos [1] produced a nearly same number of segments within ground truth. But, the method by Felzenszwalb [12], which focused in segmentation based on the graph efficiency concept, produced too-many number of segments (except for Mountain image). Consequently, it does not apply to a desired number of segment, which leads to the inappropriate approach to the case. Hence, in order of success, the proposed method and Papamarkos and Gatos' method are nearly adaptable towards the sample images.

Table 3 shows the comparison of evaluation between the proposed method and the other method [1], in which better value is bold formatted. Graph based segmentation [12] evaluation is not calculated due to its numerous segments that it had produced, except for the mountain case. It can be inferred that the proposed method, in majority, has better accuracy, precision and recall than the other. So that, in the term of classification by the ground truth, our method is better than the method of Papamarkos and Gatos, except the recall in case of Lena image. In the case of Lena image, our proposed method did not approximate the threshold too far from the exact ground truth, while the method of Papamarkos and Gatos has better approximation to the ground truth.

In summary, it has been proved that this method is adaptable to find multithresholds than the other method that separates modals in the multimodal grayscale image histogram. The approach of peak finding is used because a good threshold resides between two peaks of image histogram [7, 8]. The iterative search of the value depends on the biggest distance measured by between class variance of that value, which later called a threshold of the image. After numerous thresholds are found, it is possible that there are more thresholds than the desired, so a strategy of merging and peak removal considering the minimum between class variance value is used to get better multithreshold.

Unfortunately, during the process of the execution of the segmentation, which been stated before, an iterative computation of all pixel between pair of adjacent peaks causes several inefficient computation. So that, a better approach to approximate the value of BCV in merging process could be considered as a good extension. On the other hand, our proposed method, which proved an excellent result in grayscale image, can be developed into color image segmentation in future.

4. Conclusion

In this research, we have presented a new and simple multithresholding by peak finding method and the hierarchical analysis using between class variance in grayscale images. It has been found that the method proposed in this paper used simple statistical operation such as averaging to yield the smooth histogram, and finding minimum value of between class variance which used to merging process later.

Evaluation of five sample images, which were originally grayscale images and converted-to-grayscale images, had showed that the proposed method, which initially has smaller number of peaks to be processed later, was more adaptable by the desired number of threshold than the automatic multithresholding method. And also, in the term of the image segmentation, it yielded better thresholds than the other methods by Papamarkos and Gatos.

This method had a disadvantage in recomputing whole between class variances of all pixels between two peaks in merging process. This leads to an improvement to minimize the computation using more efficient approach. Meanwhile, the proposed method lead chances to extend this method to color image segmentation, instead of grayscale image only.

References

[1] N. Papamarkos dan B. Gatos, “A New Approach for Multilevel Threshold Selection,” CVGIP: Graphical Models and Image Processing, pp. 357-370, 1994.

[2] Z. Chi, H. Yan dan T. Pham, Fuzzy Algorithms: With applications to images processing and pattern recognition, Singapore: World Scientific, 1996.

[3] N. Otsu, “A threshold selection method from gray-level histogram,” IEEE Transaction of Systems, Man, and Cybernetics 9, pp. 62-66, 1979.

[4] J. Kapur, P. Sahoo dan A. Wong, “A new method for gray-level picture thresholding using the entropy of the histogram,” Computer Vision Graphics Image Processing, vol. 29, pp. 273-285, 1985.

[5] S. Reddi, S. Rudin dan H. Keshavan, “An optimal multiple threshold scheme for image segmentation,” IEEE Transaction Systems Man Cybernet, vol. 14, no. 4, pp. 661-665, 1984.

[6] A. Arifin dan A. Asano, “Image segmentation by histogram thresholding using hierarchical cluster analysis,” Pattern Recognition Letters, 2006.

[7] Jain, Fundamentals of Digital Image Processing, Upper Saddle River, New Jersey: Prentice Hall, 1989.

[8] P. Sahoo dan S. Soltani, “A survey of thresholding techniques,” Computer Vision Graphics Image Processing, vol. 41, no. 2, pp. 233-260, 1988.

[9] H. Cheng dan Sun.Y, “A hierarchical approach to color image segmentation using homogeneity,” IEEE Transaction Image Processing, vol. 9, no. 12, pp. 2071-2082, 2000.

[10] D. Tseng dan C. Chang, “Color segmentation using perceptual attributes,” IEEE International Conference Image Processing, pp. 228-231, 1988.

[11] K. S. Tan dan N. A. M. Isa, “Color image segmentation using histogram thresholding-Fuzzy C-means hybrid approach,” Pattern Recognition, vol. 44, pp. 1-15, 2011.

[12] P. Felzenszwalb dan D. Huttenlocher, “Efficient graph-base image segmentation,” International Journal of Computer Vision, vol. 59, no. 2, pp. 167-181, 2004