2009 Ninth International Conference on Hybrid Intelligent Systems

Otsu method and K-means

DongjuLiu, JianYu
Department of Computer Science
Beijing Jiaotong University
Beijing, China
liudongju0324@gmail.com jianyu@bjtu.edu.cn

***Abstract*—Otsu method is one of the most successful methods for image thresholding. This paper proves that the objective function of Otsu method is equivalent to that of K-means method in multilevel thresholding . They are both based on a same criterion that minimizes the within-class variance. However, Otsu method is an exhaustive algorithm of searching the global optimal threshold, while K-means is a local optimal method. Moreover, K-means does not require computing a gray-level histogram before running, but Otsu method needs to compute a gray-level histogram firstly. Therefore, K-means can be more efficiently extended to multilevel thresholding method, two-dimensional thresholding method and three-dimensional method than Otsu method. This paper proved that the clustering results of K-means keep the order of the initial centroids with respect to one-dimensional data set. The experiments show that the k-means thresholding method performs well with less computing time than Otsu method does on three dimensional image thresholding.**

Keywords-Otsu method; K-mean; K-means thresholding; two-dimensional thresholding; three-dimensional thresholding

I. INTRODUCTION

Thresholding is an important technique for image segmentation. So many researchers pay a lot of attention to the methods of how to select reasonable thresholds. Because gray levels characterize the objects in a gray image, many thresholding methods extract objects from their background based on the statistics of one-dimensional (1D) histogram of gray levels and two-dimensional (2D) histogram of gray levels. The frequently used method is Otsu method [12], which selects the global optimal threshold by maximizing the between-class variance. In bi-level thresholding, the pixel whose gray level is less than the threshold will be assigned to the background, else to the foreground. Unfortunately, it is not always good to find a single threshold that is "best" for an arbitrary gray-level. So there are many multilevel thresholding methods [2, 3, 6, 8, 13, 14, 15], which perform better than bi-level thresholding method when the picture has complex objects or background. One-dimensional (1D) histogram of gray levels only considers the gray levels. Therefore, different images with the same histogram will result in the same threshold value. Abutaleb extended a one-dimensional (1D) image histogram to two-dimensional (2D) image histogram [1], which consider about gray levels and local average. Liu *et al.* [9] proposed an automatic thresholding via two-dimensional Otsu method, which is good at reducing the noises but time consuming. Then Gong *et al.* [5] proposed a fast recursive algorithm for two-dimensional thresholding. In order to further improve the effect of segmentation, Jing *et al.* [7] extended the two-dimensional thresholding to three-dimensional thresholding. In [4], Fan *et al.* improved Jing's recursive method.

This paper proves that the objective function of Otsu mehod is equivalent to that of K-means method in multilevel thresholding, and shows that K-means can be extended to multilevel thresholding method, two-dimensional and three-dimensional thresholding method more efficiently than Otsu. As for one dimensional data, it is easy to prove that the clustering results of K-means keep the order of the initial centroids.

The paper is organized as follows: Section 2 reviews Otsu method. Section 3 discusses the relation between K-means thresholding and Otsu method, then extends K-means thresholding to two and three dimensional thresholding method. Section 4 presents the experiments. Section 5 draws the main conclusions.

II. OTSU METHOD

In this section, we rewrite Otsu method [12] as follows.

Assuming an image is represented in L gray levels [0,1,…L-1]. The number of pixels at level i is denoted by n_i , and the total number of pixels is denoted by $N = n_1 + n_2 + ... + n_L$. The probability of gray level i is denoted by

$$p_i = n_i / N,\ p_i \geq 0, \sum_{0}^{L-1} p_i = 1 \quad (1)$$

In the bi-level thresholding method, the pixels of image are divided into two classes C_1 with gray levels $[0,1,...t]$ and C_2 with gray levels $[t+1,...L-1]$ by the threshold t . The gray level probability distributions for the two classes are

$$w_1 = \Pr(C_1) = \sum_{i=0}^{t} p_i \quad (2)$$

978-0-7695-3745-0/09 $25.00 © 2009 IEEE
DOI 10.1109/HIS.2009.74

$$w_2 = \Pr(C_2) = \sum_{i=t+1}^{L-1} p_i \quad (3)$$

The means of class C_1 and C_2 are

$$u_1 = \sum_{i=0}^{t} ip_i / w_1 \quad (4)$$

$$u_2 = \sum_{i=t+1}^{L-1} ip_i / w_2 \quad (5)$$

The total mean of gray levels is denoted by u_T

$$u_T = w_1 u_1 + w_2 u_2 \quad (6)$$

The class variances are

$$\sigma_1^2 = \sum_{i=0}^{t} (i - u_1)^2 p_i / w_1 \quad (7)$$

$$\sigma_2^2 = \sum_{i=t+1}^{L-1} (i - u_2)^2 p_i / w_2 \quad (8)$$

The within -class variance is

$$\sigma_W^2 = \sum_{k=1}^{M} w_k \sigma_k^2 \quad (9)$$

The between-class variance is

$$\sigma_B^2 = w_1 (u_1 - u_T)^2 + w_2 (u_2 - u_T)^2 \quad (10)$$

The total variance of gray levels is

$$\sigma_T^2 = \sigma_W^2 + \sigma_B^2 \quad (11)$$

Otsu method chooses the optimal threshold t by maximizing the between-class variance, which is equivalent to minimizing the within-class variance, since the total variance (the sum of the within-class variance and the between-class variance) is constant for different partitions.

$$t = \arg\left\{ \max_{0 \le t \le L-1} \left\{ \sigma_B^2(t) \right\} \right\} = \arg\left\{ \min_{0 \le t \le L-1} \left\{ \sigma_W^2(t) \right\} \right\} \quad (12)$$

Otsu method can be extended to multilevel thresholding method. Assuming that there are M-1 thresholds [$t_1, t_2 ,\ldots t_{M-1}$] that divide the pixels in the image to M classes{ $C_1, C_2, \ldots C_M$ }.

$$\{t_1, t_2 \ldots t_{M-1}\} = \arg\left\{ \max_{0 \le t \le L-1} \left\{ \sigma_B^2 (t_1, t_2 \ldots t_{M-1}) \right\} \right\}$$
$$= \arg\left\{ \min_{0 \le t \le L-1} \left\{ \sigma_W^2 (t_1, t_2 \ldots t_{M-1}) \right\} \right\} \quad (13)$$

Where

$$w_j = \sum_{i=t_{j-1}+1}^{t_j} p_i \quad (14)$$

$$u_j = \sum_{i=t_{j-1}+1}^{t_j} ip_i / w_j \quad (15)$$

$$\sigma_j^2 = \sum_{i=t_{j-1}+1}^{t_j} (i - u_j)^2 p_i / w_j \quad (16)$$

$$\sigma_B^2 = \sum_{j=1}^{M} w_j (u_j - u_T)^2 \quad (17)$$

$$\sigma_W^2 = \sum_{j=1}^{M} w_j \sigma_j^2 \quad (18)$$

III. K-MEANS THRESHOLDING METHOD

K-means method proposed by MacQueen [11] partitions the data set by minimizing the within-class variance. In this work, we will prove that the objective function of Otsu method is equivalent to that of K-means method in multilevel thresholding .

First, we retell K-means again. K-means method partitions N data points into k disjoint subsets $\{C_1, C_2, \ldots C_k\}$ by minimizing the sum-of-squares criterion (within-class variance)

$$J = \sum_{j=1}^{k} \sum_{n \in C_j} (x_n - U_j)^2 \quad (19)$$

where x_n is a vector representing the n^{th} data point, j is the index of classes and $1 < j < k$, U_j is the centriod of the data points in C_j , C_j contains N_j data points.

A. K-means thresholding method

K-means can be used for thresholding as follows. Assuming that there are $m \times n$ pixels in the image with gray levels [0,1,…L-1]. The gray levels of pixel at the position (x, y) is denoted by $f(x, y)$ .Then the image can be presented by a data matrix:

$$\left[F(x, y)\right]_N = \left[f(x, y)\right] \quad (20)$$

where N= $m \times n$. K-means method partitions the points in the data matrix $\left[F(x, y)\right]_N$ into k classes with initial centroids $\left[ic_1, ic_2, \ldots ic_k\right]$ and final centroids $\left[cc_1, cc_2, \ldots cc_k\right]$.

The procedure of K-means as a threshold selection method can be expressed as follows:

1. Select k points as the initial class centroids.
2. Assign each object to the class whose center it is closest to.
3. When all objects have been assigned, recalculate the positions of the k centroids.
4. Repeat step 2 and 3 until the positions of centroids no longer change.
5. Find thresholds from the final partition.

In the bi-level thresholding method, k=2, the smallest and largest gray levels in the image are selected as the initial centroids. The threshold t can be gotten by calculating the average of the final centroids.

The k-means thresholding method can be extended to multilevel thresholding method by setting k=M, where M-1 is the number of thresholds. In the method, we don't select the initial centriods randomly as usual. The value of initial centroid ic_1 and ic_k are the smallest and largest gray level in the image; the i^{th} initial centroid is

$$ic_i = ic_1 + \frac{ic_k - ic_1}{k-1} \times (i-1) \quad (21)$$

where $1 < i < k$.

The j^{th} thresholds t_j is selected as follows:

$$t_j = 0.5\left(cc_j + cc_{j+1}\right) \quad (22)$$

where $1 \le j \le k-1$.

It should point out that the clustering results of K-means keep the order of the initial centroids with respect to one-dimensional data set. Therefore, threshold $t_j < t_{j+1}$ for $1 \le j \le k-2$.

Proof: assuming that the centroid of i^{th} class produced in the r^{th} iteration is denoted by $cc_i^{(r)}$, the new i^{th} class is denoted by $C_i^{(r)}$. Set $\theta_i^{(r)} = 0.5(cc_i^{(r)} + cc_{i+1}^{(r)})$.

Step 1:

In the first iteration, the initial centroids have a sequence from low to high gray levels. The pixels with gray levels belonging to $[\theta_{i-1}^{(1)}, \theta_i^{(1)}]$ are partitioned to $C_i^{(1)}$, where $2 \le i \le k-1$.The pixels with gray levels smaller than $\theta_1^{(1)}$ are partitioned to $C_1^{(1)}$,and the pixels with gray levels larger than $\theta_{k-1}^{(1)}$ are partitioned to $C_k^{(1)}$ .Obviously, the new centroid $cc_i^{(1)}$ which is the mean of class $C_i^{(1)}$ satisfies this: $\theta_{i-1}^{(1)} \le cc_i^{(1)} \le \theta_i^{(1)}$, where $2 \le i \le k-1$. The new centroid $cc_1^{(1)}$ satisfies this: $cc_1^{(1)} \le \theta_1^{(1)}$, and the new centroid $cc_k^{(1)}$ satisfies this: $\theta_{k-1}^{(1)} \le cc_k^{(1)}$ .Therefore $cc_1^{(1)} \le \theta_1^{(1)} ... cc_{i-1}^{(1)} \le \theta_{i-1}^{(1)} \le cc_i^{(1)} \le \theta_i^{(1)} \le cc_{i+1}^{(1)} ... \theta_{k-1}^{(1)} \le cc_k^{(1)}$.

Step 2:

Assuming the centroids produced in the r^{th} iteration have a sequence: $cc_1^{(r)} \le ... cc_{i-1}^{(r)} \le cc_i^{(r)} \le cc_{i+1}^{(r)} ... \le cc_k^{(r)}$.

Then in the $(r+1)^{th}$ iteration, the pixels with gray levels belonging to $[\theta_{i-1}^{(r+1)}, \theta_i^{(r+1)}]$ are partitioned to $C_i^{(r+1)}$,where $2 \le i \le k-1$.The pixels with gray levels smaller than $\theta_1^{(r+1)}$ are partitioned to $C_1^{(r+1)}$,and the pixels with gray levels larger than $\theta_{k-1}^{(r+1)}$ are partitioned to $C_k^{(r+1)}$ Obviously, the new centroid $cc_i^{(r+1)}$ which is the mean of class $C_i^{(r+1)}$ satisfies this: $\theta_{i-1}^{(r+1)} \le cc_i^{(r+1)} \le \theta_i^{(r+1)}$, where $2 \le i \le k-1$. The new centroid $cc_1^{(r+1)}$ satisfies this: $cc_1^{(r+1)} \le \theta_1^{(r+1)}$, and the new centroid $cc_k^{(r+1)}$ satisfies this: $\theta_{k-1}^{(r+1)} \le cc_k^{(r+1)}$. Therefore

$cc_1^{(r+1)} \le ... cc_{i-1}^{(r+1)} \le \theta_{i-1}^{(r+1)} \le cc_i^{(r+1)} \le \theta_i^{(r+1)} \le cc_{i+1}^{(r+1)} ... \le cc_k^{(r+1)}$

As K-means converges in a finite iterations, the above proof shows that $cc_1 \le ... cc_{i-1} \le cc_i \le cc_{i+1} ... \le cc_k$

By the above analysis, the thresholds selected as the averages of nearby centroids satisfies this:

$$cc_1 \le t_1 ... cc_{i-1} \le t_{i-1} \le cc_i \le t_i \le cc_{i+1} \le t_{i+1} ... t_{k-1} \le cc_k$$

That is to say： threshold $t_j < t_{j+1}$ for $1 \le j \le k-2$.

Liu *et al.* [10] proved that the objective function of Otsu method is equivalent to that of K-means method in bilevel thresholding, we prove the equivalence in multilevel thresholding.

Assuming an image is represented in L gray levels [0, 1,…L-1]. The number of pixels at level i is denoted by n_i, and the total number of pixels is denoted by N. The gray levels are partitions to k classes { $C_1, C_2, ... C_k$ }, then the objective function (18) can be substituted by follows:

$$J = \sum_{j=1}^{k} \sum_{i \in C_j} n_i (i - U_j)^2 \quad (23)$$

where $$U_j = \left(\sum_{i \in C_j} i n_i \right) \Big/ \left(\sum_{i \in C_j} n_i \right) \quad (24)$$

(24) (25) Can be gotten by combining (16) and (18) :

$$\sigma_W^2 = \sum_{j=1}^{M} \sum_{i=t_{j-1}+1}^{t_j} (i - u_j)^2 p_i \quad (25)$$

where p_i can be substituted by (1), then

$$\sigma_W^2 = \frac{1}{N} \sum_{j=1}^{M} \sum_{i=t_{j-1}+1}^{t_j} n_i (i - u_j)^2 = \frac{1}{N} \sum_{j=1}^{M} \sum_{i \in C_j} n_i (i - u_j)^2 \quad (26)$$

where $$u_j = \sum_{i=t_{j-1}+1}^{t_j} i p_i / w_j = \left(\sum_{i=t_{j-1}+1}^{t_j} i p_i \right) \Big/ \left(\sum_{i=t_{j-1}+1}^{t_j} p_i \right) = \left(\frac{1}{N} \sum_{i \in C_j} i n_i \right) \Big/ \left(\frac{1}{N} \sum_{i \in C_j} n_i \right) = \left(\sum_{i \in C_j} i n_i \right) \Big/ \left(\sum_{i \in C_j} n_i \right) \quad (27)$$

When the number of classes k=M, minimizing (23) is equivalent to minimizing (26). That is to say, the objective function of the K-means method is equivalent to that of Otsu method.

B. *Two-dimensional thresholding*

Two-dimensional Otsu method [9] performs better than one-dimensional Otsu methos when image has noises. Two dimensional Otsu method considers not only original gray levels of pixels, but also the local average gray levels. K-means thresholding method can be extended to two-dimensional thresholding method (2D K-means) too.

The 2D method is based on the assumption that "the probability sum of the object and background regions is close to 1". More details can be seen in [5, 9].

For an image with $m \times n$ pixels, $f(x, y)$ presents the gray levels at position (x, y), and $g(x, y)$ presents the local average gray levels of 3×3 neighboring pixels.

$$g(x, y) = \frac{1}{3 \times 3} \sum_{\Delta x=-1}^{1} \sum_{\Delta y=-1}^{1} f(x + \Delta x, y + \Delta y) \quad (28)$$

The image can be expressed by a data matrix as follows:

$$[F(x, y)]_{N \times 2} = [f(x, y), g(x, y)] \quad (29)$$

where N= $m \times n$ .In the 2D K-means method, the points in the data matrix are partitioned into two classes, k=2.The initial centroids can be selected as follows:

$$ic_1 = (f_{\min}, g_{\min}) \quad (30)$$

$$ic_2 = (f_{\max}, g_{\max}) \quad (31)$$

where $f_{\min}$ and $f_{\max}$ are the smallest and largest gray levels of pixels in the image , $g_{\min}$ and $g_{\max}$ are the smallest and largest local average gray levels. The optimal threshold is gotten by calculating the average of final two centroids.

C. Three-dimensional thresholding

K-means thresholding method can be extended to three-dimensional thresholding method (3D K-means).

For an image with $m \times n$ pixels, the representations of $f(x,y)$ and $g(x,y)$ are showed in 2D K-means. $h(x,y)$ represents the median gray levels of 3×3 neighboring pixels.

$$h(x,y) = med\{f(x+\Delta x, y+\Delta y), \quad \Delta x = -1,0,1; \Delta y = -1,0,1\} \tag{32}$$

The image can be denoted by an data matrix

$$\left[F(x,y)\right]_{N\times 3} = \left[f(x,y), g(x,y), h(x,y)\right] \tag{33}$$

where N= $m\times n$ .The selection of initial centroids and the calculation of threshold is similar to the 2D K-means.

IV. THEORETIC ANALYSIS AND EXPERIMENTS REPARE

The K-means thresholding method is a local optimal method , while Otsu method uses an exhaustive search. Otsu method is time consuming, especially the multilevel thresholding method, two-dimensional thresholding method and three-dimensional thresholding method.

The experiments in the paper are conducted on a computer with 2.2GHz Inter(R) core (TM) 2 Duo CPU and 2 GB physical memory. The programs are coded in matlab R2008a.

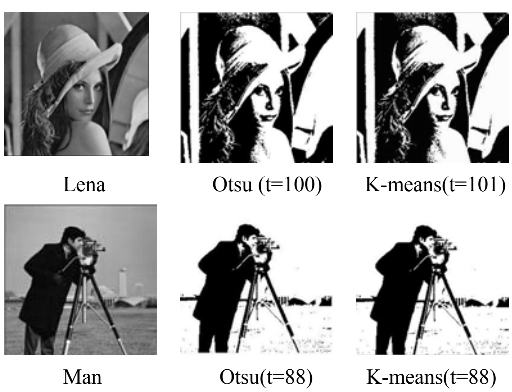

Figure 1. Bi-level thresholding.

Figure 2. K-means for multilevel thresholding.

TABLE I. COMPUTER TIME OF MULTILEVEL THRESHOLDING

time (s) \ k / image	k=2	k=3	k=4	k=5	k=6
lena	0.235	0.260	0.287	0.355	0.669
man	0.222	0.326	0.432	0.351	0.421

TABLE II. COMPUTATION TIME OF TWO METHODS USING FOR 3D THRESHOLDING

time (s) \ image / method	Lena+ G noise	Lena+ S & P noise	Man+ G noise	Man+ S & P noise
3D Otsu	19.817	19.803	19.854	19.941
3D K-means	1.118	1.198	0.983	0.991

("G" is short for "Gassian", "S & P" is short for "Salt and Pepper")

For segmenting an image with L gray levels using M-1 thresholds, Otsu's exhaustive method searches $\binom{L}{M-1}$ combinations of thresholds, which can be approximated to $O\left(L^{M-1}\right)$ for (M-1) ≪ L. The time complexity is exponential in the number of thresholds. While the complexity of K-means thresholding method which segments an image with N pixels using M-1 thresholds is $O\left(MNDI\right)$. D is the number of dimensions, I is the number of iterations. Then the complexity of the K-means thresholding method is linear. *Table Ⅰ* shows the running time of K-means threshlolding method for multilevel thresholding. The computation time is less than 1 second. But Huang et al.[6] showed that for Otsu method, the variations in runtimes are small when k = 4; but they become large when k = 5 with the runtime being from 651.1 s to 689.3 s, and 47,399 s to 54,168 s for k = 6； for the recursive Otsu method(liao et al.[8]), the runtimes are close when k⩽ 5; for k= 6, the runtimes varied from 1024 s to 1041 s running on the a Pentium PC with a 3.4 GHz Core Duo processor and 4GB DDR II memory.

Fig. 1 shows the resulting images by K-means thresholding method and Otsu method. It is found that the bi-level threshold gotten by K-means thresholding method is close to the bi-level threshold gotten by Otsu method. By testing 100 images in Berkeley's database, we find that the bi-level thresholds of an image gotten by the two methods are same for 42 images; for 57 images, the difference between the thresholds gotten by the two methods only less than three gray levels; but for one image, the thresholds are so different. *Fig. 2* shows the resulting image by multi-level thresholding method.

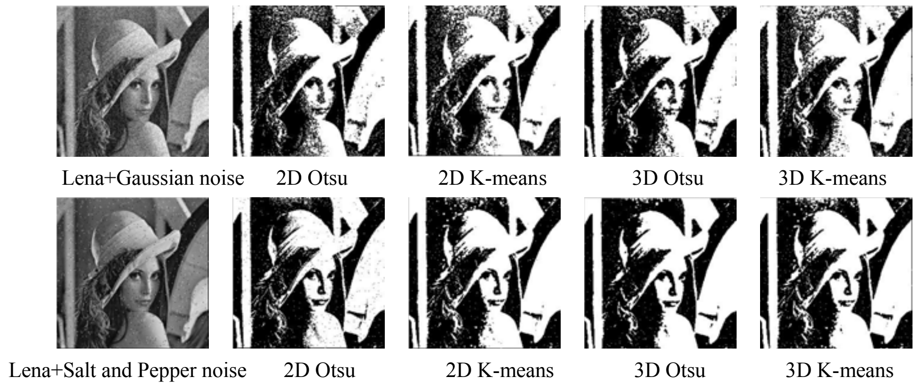

Figure 3. thresholded images for image-Lena.

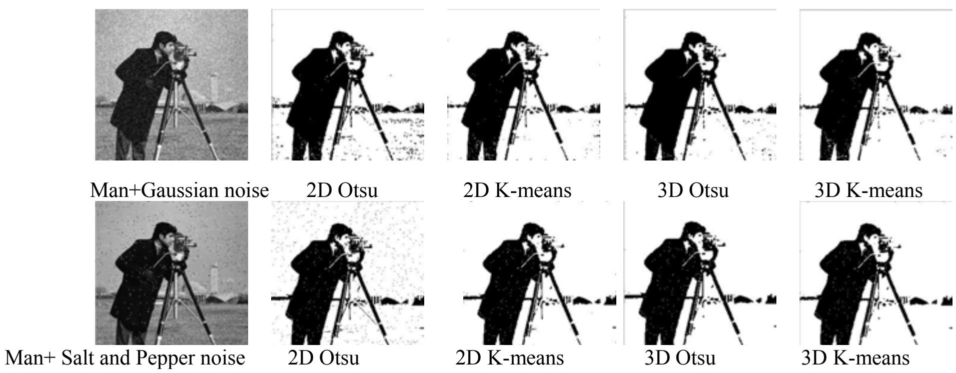

Figure 4. thresholded images for image-Man.

The complexity of 2D K-means is $O(MNDI)$, where the number of thresholds M=2, the number of dimensions D=2.In the 3D K-means M=2, D=3. The number of iteration I is usually smaller than 20 for our initial centroids. When the number of pixels in an image is less than 256×256 , K-means thresholding method without a histogram runs fast, but if the number of pixels is very large, we can use a histogram in order to speed up the algorithm.

K-means thresholding method runs fast, especially the three dimensional thresholding . *Table II* shows the experimental results that the K-means thresholding method is 20 times faster than Otsu method for three dimensional thresholding .

K-means thresholding method using for two dimensional and three dimensional thresholding can reduce noise in the image. *Fig. 3*，*Fig. 4* show the resulting images by K-means thresholding method and Otsu method. They are both good at reducing Gaussian noise, especial for Salt and Pepper noise, K-means thresholding method is better than Otsu method for two and three dimensional thresholding.

The K-means thresholding method doesn't need a histogram before calculation while the Otsu method needs a histogram, so the data matrix partitioned by the K-means thresholding can be decimal although the gray levels is integer, the K-means thresholding is more general in this sense.

V. CONCLUSION

This paper proves that the objective function of the Otsu method is equivalent to that of the K-means method in multilevel thresholding. They both evaluate the same criterion for minimizing the within–class variance. K-means method is a general thresholding method as it does not need a histogram before calculation. The K-means thresholding method performs efficiently, and it runs fast especially in multilevel thresholding method and three-dimensional thresholding method.

ACKNOWLEDGMENT

This work is partially supported by NSFC Grant 6087503, 90820013; 973 Program Grant 2007CB311002; the Program for New Century Excellent Talents in University of China Grant NECT-06-0078.

REFERENCES

[1] A.S. Abutaleb, "Automatic thresholding of gray level pictures using two-dimensional entropy," Computer Vision, Graphics, and Image Processing, 47: 22–32, 1989.

[2] S.Arora, J. Acharya, A.Verma, Prasanta K. Panigrahi , "Multilevel thresholding for image segmentation through a fast statistical

recursive algorithm," Pattern Recognition Letters ,Vol.29: 119–125, 2008.

[3] L.J. Dong, G.Yu, P. Ogunbona , W.Q. Li, "An efficient iterative algorithm for image thresholding," Pattern Recognition Letters. 29:1311–1316, 2008.

[4] J.L.Fan, F.Zhao, X.F.Zhang, "Recursive algorithm for three-dimensional Otsu's thresholding segmentation method," Acta Electronical Sinica,Vol. 35, No.7:1398-1402, 2007.

[5] J. Gong, L. Li, and W. Chen, "Fast recursive algorithms for two-dimensional thresholding,"Pattern Recognition, Vol. 31, No. 3: 295-300, 1998.

[6] D.Y.Huang, C.H.Wang, "Optimal multi-level thresholding using a two-stage Otsu optimization approach," Pattern Recognition Letters, 30 : 275−284, 2009.

[7] X.J. Jing, J.F.Li, Y.L.Liu, "Image segmentation based on 3-D maximum between-cluster variance,"Acta Electronic Sinica, 31 (9) : 1281-1285, 2003.

[8] P.S.Liao, T.S.Chen, P.C. Chung, "A fast algorithm for multilevel thresholding," Journal of Information Science and Engineering, 17: 713-727, 2001.

[9] J. Z. Liu and W. Q. Li, "The automatic thresholding of gray-level pictures via two-dimensional Otsu method,"Acta Automatica Sinica, 19(1): 101-105, 1993.

[10] J. Z. Liu and and Y. Q. Tu, "Thresholding of images using an efficient c-means clustering algorithm," Journal of electronics,Vol. 14, No.4: 424-427,1992.

[11] J. MacQueen. "Some Methods for Classification and Analysis of Multivariate Observations," Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, volume 1:281–296, 1967.

[12] N. Otsu, "A threshold selection method from gray-level histogram,"IEEE Transactions on System Man Cybernetics, Vol. SMC-9, No. 1: 62-66, 1979.

[13] N.Papamarkos, B.Gatos, "A new approach for multilevel threshold selection,"Graphics Models Image Process, 56:357–370, 1994.

[14] S.S.Reddi, S.F. Rudin, H.R. Keshavan, "An optimal multiple threshold scheme for image segmentation," IEEE Trans. System Man Cybernet. 14(4): 661–665, 1984.

[15] S. Wang, R. Haralick,"Automatic multithreshold selection,"Computer Vision,Graphics, and Image Processing, Vol. 25:46-67, 1984.