

# Histogram Matching Schemes for Image Thresholding

Salah Ameer

Faculty of Applied Science and Technology, Humber College, Brampton, ON, Canada

## Article history

Received: 16-08-2019

Revised: 28-08-2019

Accepted: 07-09-2019

Email: salah.ameer@humber.ca

**Abstract:** This paper proposes several novel schemes for image thresholding. The idea is simply to compare the original image histogram to that of the thresholded image. Element by element comparison (sum of absolute difference between the two histograms) is found to be of better performance than a single feature (area or size) comparison. The optimum threshold is the one producing the best comparison. Cumulative histogram is introduced as a generalization to the area under the curve and found to be of better performance. In addition, a new performance measure is suggested based on percentage of correct assignments in both foreground and background. Comparative results with Otsu shows the effectiveness of the proposed schemes.

**Keywords:** Image Thresholding, Histogram Matching, Performance Measure

## Introduction

Image thresholding is vital in many applications and one of the effective methods for image segmentation. Various schemes have been proposed in the literature, a good review can be found in Sezgin and Sankur (2004).

The histogram plays a crucial role in many of these schemes. In general, the histogram is used as an approximation to the probability density function (Otsu, 1979; Kapur *et al.*, 1985). In these cases and their extensions, the threshold is selected as a solution to an optimization problem for some objective function dependent on features extracted from the histogram.

Due to the fact that a histogram does not carry spatial information (2 different images can have the same histogram), higher dimensional histograms have been proposed as in Zhang and Hu (2008; Abutaleb, 1989; Zheng *et al.*, 2017).

The aforementioned schemes can be generalized to multi-level thresholding as in Liu and Yu (2009). However, the computational price is too high. In addition, having many thresholds, the ensemble size for each region is reduced. This often results in inferior quality since statistics (or probability distribution, i.e., histogram) rely heavily on a large ensemble size.

There are many measures, as described in Sezgin and Sankur (2004), to evaluate the performance of a thresholding scheme. However, application dependant, a subjective decision may be preferred.

This research proposes few formulations that exploit simple features deduced from the 1D histogram or the cumulative histogram. The features investigated are simply the area under the (cumulative) histogram and the sum of the absolute difference between the histograms of the original and thresholded images. The cumulative histogram is of special interest due to the fact that it is a monotonic function. The threshold is the one producing an image having a (cumulative) histogram matching in some sense that of the original image.

In addition, a simple performance measure is suggested in this work to reduce the bias towards large background (foreground).

## Preliminaries

The following symbols are adopted in all subsequent sections:

- $g_i$  = The  $i^{\text{th}}$  gray level in the image.
- $h$  = histogram.
- $H$  = Cumulative histogram.
- $L$  = number of gray levels in the image.
- $N$  = Number of pixels in the image.
- $X_O$  = Feature from the original image.
- $X_T$  = Feature from the thresholded image.
- $T$  = Threshold.
- $n_{<T}$  = Size of gray levels below the threshold.
- $m_{<T}$  = Average of gray levels below the threshold.
- $V_{<T}$  = Variance of gray levels below the threshold.

$n_{>T}$  = Size of gray levels above the threshold.  
 $m_{>T}$  = Average of gray levels above the threshold.  
 $V_{>T}$  = Variance of gray levels above the threshold.

Without loss of generality, the gray levels are normalized to the interval [0, 1]. Also, the histogram is normalized such that its sum is equal to one.

For the purpose of comparing the performance of the thresholding schemes, many evaluation criteria have been suggested (Sezgin and Sankur, 2004). A common criterion is the Misclassification Error (ME) given by:

$$ME = 1 - \frac{|B_G \cap B_T| + |F_G \cap F_T|}{|B_G| + |F_G|} \quad (1)$$

where,  $F_G$  is the foreground of ground truth,  $B_G$  is background of ground truth,  $F_T$  is the foreground of thresholded image,  $B_T$  is the background of thresholded image and  $||$  is the cardinality of the set.

The ME measure given above is simply the mean of the absolute difference between the thresholded image and the ground truth provided that both images are binary.

The Yule coefficient has been suggested by Sneath and Sokal, (1973) to avoid the bias against small foregrounds, given by:

$$\frac{|B_G \cap B_T|}{|B_G \cup B_T|} + \frac{|F_G \cap F_T|}{|F_G \cup F_T|} - 1 \quad (2)$$

Many algorithms, however, tend to be biased toward one side. Unfortunately, Equation (2) will tend to produce higher values for the performance. In addition, the measure can be negative for highly misaligned foreground and/or background. As a remedy, the dual similarity measure DSM is thereby proposed as a modification to the Yule coefficient:

$$DSM = 1 - \min \left( \frac{|B_G \cap B_T|}{|B_G \cup B_T|}, \frac{|F_G \cap F_T|}{|F_G \cup F_T|} \right) \quad (3)$$

Obviously, a value of 0 indicates best match and a value of 1 indicates completely misaligned foreground or background. In addition, algorithms having similar losses for foreground and background are preferred over those having good foreground detection with bad background detection or vice-versa.

### Algorithms Based on the Histogram

Intuitively, the best thresholded image would be the one having the highest similarity (in some sense) with the original image. Similarity is performed in this work

through a simple comparison (absolute difference) between a feature belonging to the original image histogram and a corresponding one belonging to the thresholded image histogram.

All of the proposed schemes implement exhaustive search to find the optimum threshold.

#### Linearized Histogram Area LHA

The first scheme is the comparison of the area under the histogram curve. The area under the original histogram can be approximated using trapezoidal rule as:

$$LHA_o = \left\{ \sum_1^L h(i) - (h(1) + h(L)) / 2 \right\} / L \quad (4)$$

The area under the thresholded histogram can be represented by that of two triangles (assuming 0 for the histogram at gray levels 0, T and 1), resulting in:

$$LHA_T = 0.5 * \{ T * n_{<T} + (1-T) * n_{>T} \} \quad (5)$$

The optimum threshold is given by:

$$T_{LHA} = \arg \left\{ \min_T |LHA_T - LHA_o| \right\} \quad (6)$$

#### Linearized Histogram Difference LHD

The piece-wise linear approximation to the histogram of the thresholded image in the previous section can be normalized to have a sum of 1. The result is then matched with the original histogram in a similar manner to the L-norm formulation. The optimum threshold can then be found as:

$$T_{LHD} = \arg \left\{ \min_T \left( \sum_1^L |h_o(i) - h_T(i)|^2 \right) \right\} \quad (7)$$

A value of 2 was chosen for the exponent. However, the scheme has tendency to produce better performance with values more than 2. Other functions can be explored as well, not necessarily of power or polynomial type.

#### Area under 2 Gaussians A2G

Similar to scheme LHA above, with the exception that 2 Gaussians are used instead of 2 triangles. The thresholded histogram is now given by:

$$h_T(i) = a_{<T} e^{-(g_i - m_{<T})^2 / (2V_{<T})} + a_{>T} e^{-(g_i - m_{>T})^2 / (2V_{>T})} \quad (8)$$

The values of  $a_{<T}$  at  $a_{>T}$  are the solutions to the linear system resulting from fitting  $h_T$  to  $h_o$ . The optimum threshold is then given by:

$$T_{A2G} = \arg \left\{ \min_T \left( \left| \sum_1^L h_o(i) - \sum_1^L h_T(i) \right| \right) \right\} \quad (9)$$

### Difference between 2 Gaussians D2G

Similar to scheme A2G above, however,  $h_T$  should now be normalized to have a sum of 1. The optimum threshold is formulated as:

$$T_{D2G} = \arg \left\{ \min_T \left( \sum_1^L |h_o(i) - h_T(i)|^2 \right) \right\} \quad (10)$$

Similar extensions suggested for LHD scheme above are of interest in this scheme as well.

### Truncated 2 Gaussians T2G

Following the same procedures as in A2G and D2G above, we formulate the histogram as a sum of two truncated Gaussians; one belongs to the foreground and one to the background. Hence, we have:

$$h_T(i) = \begin{cases} a_{<T} e^{-(g_i - m_{<T})^2 / (2V_{<T})} & g_i < T \\ a_{>T} e^{-(g_i - m_{>T})^2 / (2V_{>T})} & g_i > T \end{cases} \quad (11)$$

Similarly,  $a_{<T}$  at  $a_{>T}$  are the solutions to the linear system resulting from fitting  $h_T$  to  $h_o$ . Resulting in:

$$T_{ATG} = \arg \left\{ \min_T \left( \left| \sum_1^L h_o(i) - \sum_1^L h_T(i) \right| \right) \right\} \quad (12)$$

$$T_{DTG} = \arg \left\{ \min_T \left( \sum_1^L |h_o(i) - h_T(i)|^2 \right) \right\} \quad (13)$$

In a similar to scheme D2G above,  $h_T$  should now be normalized to have a sum of 1 in scheme DTG.

### Algorithms Based on Cumulative Histogram

The encouraging results of the previous section motivates the author to investigate the cumulative histogram as it can be considered as a generalization to the area under the histogram. In addition, the cumulative histogram has better grounds in terms of comparing the unavailable gray levels in the thresholded image. In other words, the histogram of the thresholded image has only two nonzero values, while the cumulative histogram has nonzero values for all gray levels greater than or equal to  $m_{<T}$ .

The histogram of the thresholded image has only two nonzero entries:  $n_{<T}$  at  $m_{<T}$  and  $n_{>T}$  at  $m_{>T}$ . Therefore, the resultant cumulative histogram is a two-step function.  $H_T$  will be zero until  $m_{<T}$ , then  $n_{<T}$  until  $m_{>T}$ , after that it is 1. In other words:

$$H_T(i) = \begin{cases} 0 & g_i \leq m_{<T} \\ n_{<T} & m_{>T} > g_i > m_{<T} \\ 1 & g_i \geq m_{>T} \end{cases} \quad (14)$$

Exhaustive searches were used in the following schemes in a similar fashion to that of the previous section.

### Cumulative Histogram Size CHS

This scheme is simply the difference between the total sums of the cumulative histograms. Hence, the threshold is given by:

$$T_{CHS} = \arg \left\{ \min_T \left( \left| \sum_1^L H_o(i) - \sum_1^L H_T(i) \right| \right) \right\} \quad (15)$$

A similar outcome can be obtained by comparing the areas (trapezoidal approximations) under the cumulative histograms of the original and the thresholded images.

### Cumulative Histogram Power CHP

The objective function here is through power comparison given by:

$$T_{CHP} = \arg \left\{ \min_T \left( \left| \sum_1^L H_o^2(i) - \sum_1^L H_T^2(i) \right| \right) \right\} \quad (16)$$

### Cumulative Histogram Difference CHD

In this scheme, the optimum threshold is obtained when the resultant image has a matching cumulative histogram to that of the original in the following sense:

$$T_{CHD} = \arg \left\{ \min_T \left( \sum_1^L |H_o(i) - H_T(i)|^{0.1} \right) \right\} \quad (17)$$

The value of 0.1 for the exponent was chosen through experimentation. Other functions of the absolute difference can be used. This can open the path to a family of algorithms.

### Experimental Results

The suggested comparative schemes based on Equation (4-17) are compared with Otsu (1979) thresholding due to its popularity. The images in Fig. 1 were used.

Table 1 lists ME values, see Equation (1), for the proposed schemes and that of Otsu (1979) using the images in Fig. 1.

Results are encouraging as seen from Table 1 and 2. However, area (or size) schemes are inferior in performance to that of absolute difference schemes. Power scheme is somewhat in between. This clearly highlights the superiority of element by element

match over a single feature match. Generalization of the last statement to other features in the literature

requires an extensive testing. Nevertheless, it seems attainable given the strong results from Table 1 and 2.

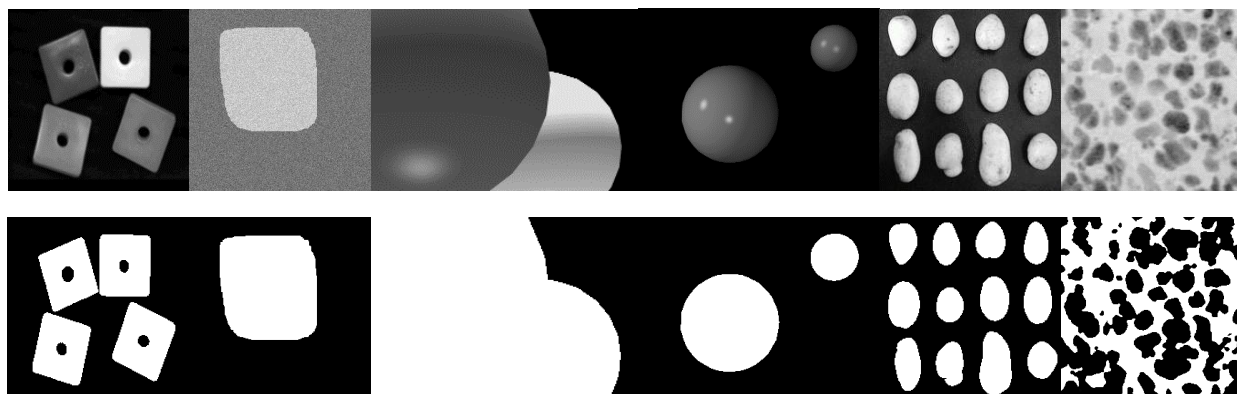


Fig. 1: Test images used and their ground truth

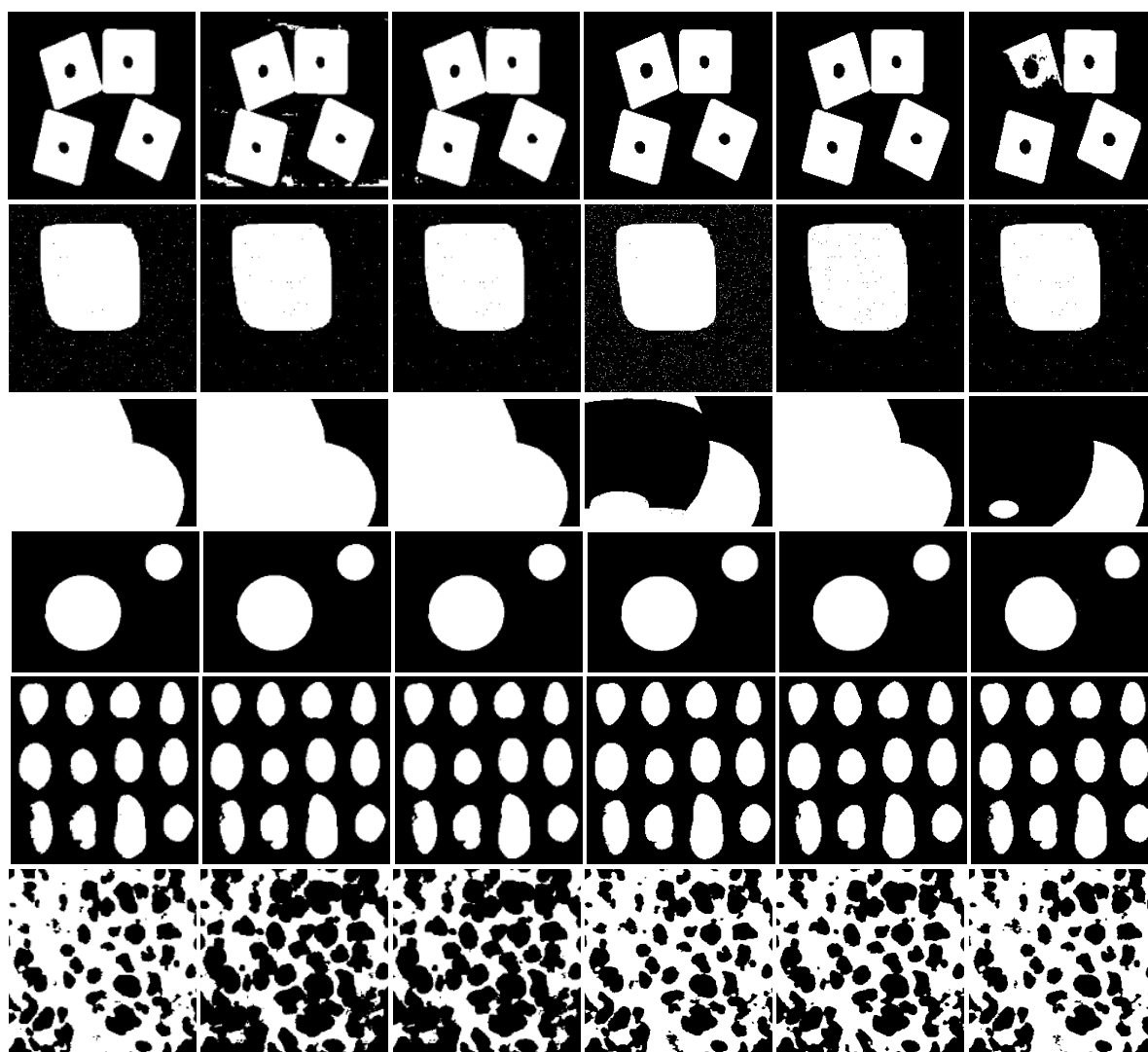
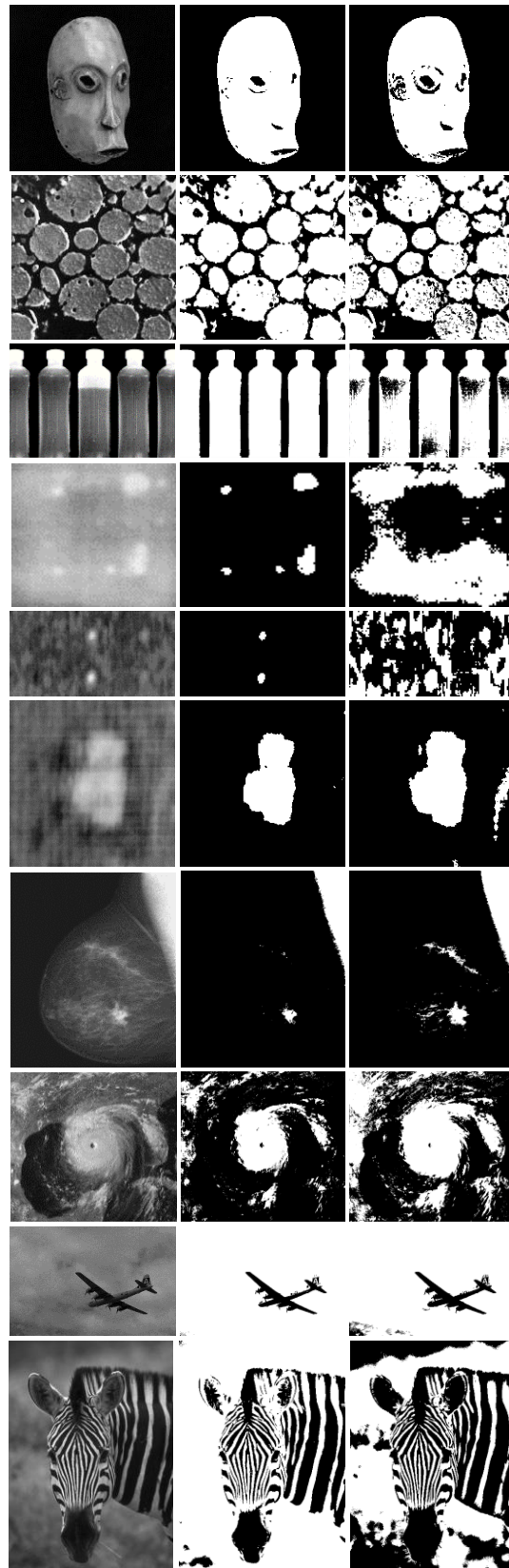


Fig. 2: Resultant images (left to right): LHD, D2G, DTG, CHP, CHD and Otsu. Figure 1 for original images



**Fig. 3:** Comparisons with no gold standard, (left to right): Original, proposed CHD and Otsu images

**Table 1:** ME values for OTSU and proposed schemes for the images in Fig. 1

Image/Method	1	2	3	4	5	6
<b>Otsu</b>	0.047	0.002	0.599	0.020	0.019	0.149
<b>LHA</b>	0.027	0.063	0.555	0.000	0.011	0.194
<b>LHD</b>	0.015	0.006	0.000	0.000	0.034	0.165
<b>A2G</b>	0.028	0.186	0.001	0.205	0.011	0.035
<b>D2G</b>	0.048	0.002	0.002	0.000	0.017	0.092
<b>ATG</b>	0.067	0.017	0.001	0.191	0.011	0.014
<b>DTG</b>	0.032	0.002	0.001	0.000	0.017	0.082
<b>CHS</b>	0.255	0.719	0.384	0.209	0.354	0.477
<b>CHP</b>	0.003	0.017	0.488	0.001	0.005	0.116
<b>CHD</b>	0.008	0.002	0.000	0.001	0.012	0.049

**Table 2:** DSM values for OTSU and proposed schemes for the images in Fig. 1

Image/Method	1	2	3	4	5	6
<b>Otsu</b>	0.133	0.008	0.751	0.097	0.052	0.290
<b>LHA</b>	0.070	0.183	0.736	0.000	0.030	0.377
<b>LHD</b>	0.041	0.022	0.002	0.000	0.093	0.320
<b>A2G</b>	0.073	0.664	0.006	0.981	0.031	0.070
<b>D2G</b>	0.121	0.008	0.008	0.001	0.047	0.188
<b>ATG</b>	0.190	0.058	0.007	0.914	0.031	0.029
<b>DTG</b>	0.083	0.008	0.006	0.000	0.047	0.169
<b>CHS</b>	0.747	0.999	0.420	0.984	0.981	0.982
<b>CHP</b>	0.008	0.058	0.710	0.003	0.013	0.226
<b>CHD</b>	0.023	0.006	0.001	0.006	0.034	0.096

In reference to the above two tables, the proposed DSM has a wider range of values compared to that of ME. In essence, ME is related to the percentage of error to the whole image, while DSM is related to the worst of the percentage errors obtained from either foreground or background.

In terms of computation (hence fast implementation), more consideration should be given to LHA followed by CHP, CHD and LHD.

Figure 2 shows the thresholded images using LHD, D2G, D2G, CHP, CHD and Otsu schemes. Clearly, the results of these scheme are subjectively comparable to that of the ground truth in Fig. 1.

Interestingly, schemes dependant on the histogram have tendency to use larger exponents, giving more influence to outliers. On the contrary, smaller exponents are better with cumulative histogram schemes. The reason is simply the smoothing of the outliers thanks to the cumulative operation.

The scheme CHD has remarkable performance and hence more images were compared in Fig. 3. Unfortunately, no ground truth is available. In all these cases, the performance is superior to that of Otsu. However, more test data is needed to draw a stronger conclusion.

## Conclusion and Future Work

New algorithms for image thresholding has been proposed in this work using simple (cumulative) histogram comparison.

Results are promising, however, more test images are needed to explore the limits of the proposed schemes. The domain of application for each of the proposed schemes as well as their extension to higher dimensional histograms are currently under investigation.

The results clearly indicate the superiority of schemes based on absolute difference over those dependant on a single value (or objective function).

The proposed schemes have many local minima (the LHA scheme is an exception) with objective function values comparable to that of the global minimum. This observation is also noticed (in some cases) for other schemes including Otsu. However, the number of minima is far less in Otsu case. More investigation is needed to decide whether these minima are due to histogram noise or can be used to refine the threshold. Suggestions can be: Weighted average, iterative thresholding, or the possibility of multi-level thresholding.

The remarkable performance of schemes D2G, DTG, CHP and CHD, see Equation 10, 13, 16 and 17, should encourage further investigations to find the optimum exponent to these schemes. Other non-linear functions may be worthy of some insight.

As clearly demonstrated in Fig. 3, CHD performance is outstanding. Referring to Equation 14, more thorough investigations are need to compare the CHD scheme with a piece-wise linear approximation or even a higher order one. This can also include exponential functions.

## Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

## References

- Abutaleb, A., 1989. Automatic thresholding of gray-level pictures using two-dimensional entropy. *Comput. Vis. Graph. Image Process.*, 47: 22-32. DOI: 10.1016/0734-189X(89)90051-0
- Kapur, J., P. Sahoo and A. Wong, 1985. A new method for gray-level picture thresholding using the entropy of the histogram. *Comput. Vis. Graph. Image Process.*, 29: 273-285. DOI: 10.1016/0734-189X(85)90125-2
- Liu, D. and J. Yu, 2009. Otsu method and k-means. *Proceedings of the 9th International Conference on Hybrid Intelligent Systems*, Aug. 12-14, IEEE Xplore Press, Shenyang, China, pp: 344-349. DOI: 10.1109/HIS.2009.74
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybernet.*, 9: 62-66. DOI: 10.1109/TSMC.1979.4310076
- Sezgin, M. and B. Sankur, 2004. Survey over image thresholding techniques and quantitative performance evaluation. *J. Electronic Imag.*, 13: 146-165. DOI: 10.1117/1.1631315
- Sneath, P. and R. Sokal, 1973. *Numerical Taxonomy: The Principle and Practice of Numerical Classification*. 1st Edn., W.H. Freeman, ISBN-10: 9780716706977, pp: 573.
- Zhang, J. and J. Hu, 2008. Image segmentation based on 2D Otsu method with histogram analysis. *Proceedings of the International Conference on Computer Science and Software Engineering*, Dec. 12-14, IEEE Xplore Press, Hubei, China, pp: 105-108. DOI: 10.1109/CSSE.2008.206
- Zheng, X., H. Ye and Y. Tang, 2017. Image bi-level thresholding based on gray level-local variance histogram. *Entropy*, 19: 1-8. DOI: 10.3390/e19050191