



# A Novel Algorithm for Image Thresholding Using non Parametric Fisher Information

Gamil Abdel-Azim, Member IEEE <sup>1</sup> and Z. A. Abo-Eleneen <sup>2,3,\*</sup>

<sup>1</sup> College of Computer & Informatics, Canal Suez University, Egypt; E-Mails: [gazim3@gmail.com](mailto:gazim3@gmail.com)

<sup>2</sup> Faculty of Computers and Informatics, Zagazig University, Zagazig 44519, Egypt

<sup>3</sup> Department of Mathematics, Faculty of Science, Qassim University, Al-Montazah, Buraydah 51431, Al Qassim, Saudi Arabia; \* E-Mail: [zaher\\_aboeleneen@yahoo.com](mailto:zaher_aboeleneen@yahoo.com)

Tel.: + 966 -502-275-461.

*Received: 28 October 2014 / Accepted: 17 November 2014 / Published: 27 November 2014*

---

**Abstract:** The Fisher information (FI) measure is an important concept in statistical estimation theory and information theory. However, it has received relatively little consideration in image processing. In this paper, a novel algorithm is developed based on the nonparametric FI measure. The proposed algorithm determines the optimal threshold based on the FI measure by maximizing the measure of the separability of the resultant classes over all of the gray levels. The algorithm is compared with several classic thresholding methods on a variety of images, including some nondestructive testing (NDT) images and text document images. The experimental results show the effectiveness of the new method.

**Keywords:** Image thresholding; Histogram; Fisher information; Information theory

---

## 1. Introduction

The segmentation of images into homogeneous regions is an important area of research in computer vision. Image thresholding, which is a popular technique for image segmentation, is also regarded as an analytic image representation method [1]. This technique plays an important role in many of the tasks that are required for pattern recognition, computer vision, and video retrieval [2]. Image thresholding is computationally simpler than other existing algorithms, such as boundary detection or region dependent techniques [3-6]. Its aim is to find an appropriate threshold for separating the object of interest from the background. The output of a thresholding process is a binary image in which all of the pixels with gray levels higher than the determined threshold are classified as object and the remaining of pixels are assigned to background, or vice versa. This technique can be

used in a variety of applications, including biomedical image analysis, handwritten character identification, automatic target recognition and change detection; see, for example [7-9].

Many approaches including comprehensive overviews and comparative studies of image thresholding can be found in [10-13]. Among these approaches, three of the most popular are Otsu's method, the minimum-error method and the entropy based threshold method. Otsu's method [14] can dynamically determine the threshold for image segmentation by maximizing the separability measures of the classes according to the image's gray-level histogram. Otsu's method has been adopted for automatic image thresholding in some free and commercial software, such as GIMP ([www.gimp.org](http://www.gimp.org)) and MATLAB (The Math Works, Inc.). The minimum-error method is ranked as the best in a comprehensive survey of image thresholding conducted by [13]. In image thresholding, the determination of an optimal threshold is often based on the estimation measures of the location and the spreading of the intensities of the object and background. As with many other approaches, both Otsu's method and the minimum-error method use the sample mean and the sample standard deviation when estimating a location and spreading, respectively.

Many studies on thresholding have utilized an information theoretical approach for threshold images. This approach constructs an optimization criterion that is based on the concept of entropy [15–19]. The principle of entropy is to use uncertainty as a measure for describing the information that is contained in a source. Maximum information is achieved when no a prior knowledge is available, in which case the result is maximum uncertainty. In other words, if there is no preference among the samples that result from an experiment, then the best decision is to treat all of the samples as equally important, without introducing any biased knowledge into the decision-making process. In this case, the probability distribution that describes the experiment is either uniformly distributed in continuous probability space or equally likely in discrete probability space, both of which yield the maximum entropy. This concept was further generalized to Renyi's entropy [19].

Entropy-based thresholding assumes that there are two probability distributions. One distribution is for the object class and another distribution is for the background class. If the sum of the entropies of the two classes is maximal then the image segmentation is performed.

Pun has proposed maximum entropy as an optimal criterion for image thresholding [20, 21]. Kapur et al. [22] found some flaws in Pun's derivations and presented an improved entropy-based method. Abutaleb [23] maximizes the 2D entropy also reported a method similar to Pun's method. Brink's method [24] maximizes the sum of the entropies computed from two autocorrelation functions of the thresholded histograms. Li and Lee's method [25] minimizes the relative cross entropy (Kullback-Leibler distance) between the original image and the thresholded images. Kitler and Illingworth [26] developed a thresholding method that minimizes the segmentation errors and that is derived using an information-theoretic approach. Cheng et al. [27] selected the threshold criterion based on maximizing the fuzzy entropy.

A quantity that is related to the Shannon entropy is the FI measure [28]. This quantity has two basic roles to play in theory. First, it is a fundamental principle of the statistical field of study called parameter estimation. Second, it is a measure of the state of disorder of a system or phenomenon. Like FI, Shannon's entropy is a function of an underlying probability density (mass) function  $p(x)$  and is an "information." Historically, FI predates the Shannon entropy by approximately 25 years (1922 versus

1948). There are some known relations that connect the two information concepts [29-31]. Shannon's entropy can be, but is not always, the thermodynamic, Boltzmann entropy [28].

While many criteria and algorithms for image segmentation are developed based on Shannon entropy and variance [13]. In this paper, we propose Fisher information-based image thresholding as a new algorithm for image segmentation. The Fisher based thresholding considers an image histogram to be a probability distribution, and then selects an optimal threshold value that yields the maximum FI.

Tests against a variety of images, including some nondestructive testing (NDT) and text document images, show that the objects are extracted successfully. This result is found irrespective of the complexity of the background and the difference in the class sizes. The remainder of this paper organized as follows: The proposed threshold algorithm is presented in section 2, the experimental results are presented in section 3 and the conclusions are presented in section 4.

## 2. The proposed threshold algorithm

In this section, the FI concept and global information measures of information are reviewed. A new thresholding objective function and the corresponding algorithm are then proposed.

### 2.1 Fisher information

FI measure is a key concept in information theory and the theory of statistical inference. FI essentially describes the amount of information data provide about an unknown parameter. It has applications both in finding the variance of an estimator through the Cramer–Rao inequality and in the asymptotic behavior of maximum likelihood estimates [32].

Let  $X$  be a random variable, and let  $p(x; \theta)$  be the probability density (mass) function for some model of the data that have parameter  $\theta$ . The FI  $I(\theta)$  measure is given by the following [32].

$$I(\theta) = -E\left(\frac{\partial^2 \log p(x; \theta)}{\partial \theta^2}\right) = E\left(\frac{\partial \log p(x; \theta)}{\partial \theta}\right)^2. \quad (1)$$

The FI measure arises as a measure of the expected error in a smart measurement. Consider the class of “unbiased” estimates, obeying  $E(\hat{\theta}) = \theta$ ; these estimates are correct “on average.” The mean-square error  $e^2$  in such an estimate  $\hat{\theta}$  obeys the following relation [33].

$$e^2 \geq \frac{1}{I(\theta)}. \quad (2)$$

Equation (2) is called the Cramer–Rao inequality. It expresses reciprocity between the mean-square error  $e^2$  and the FI in the intrinsic data. Hence, the Cramer–Rao inequality is an expression of intrinsic uncertainties. Eq. (2) shows that the estimation quality increases (e decreases) as  $I(\theta)$  increases. Therefore,  $I(\theta)$  is a quality metric of the estimation procedure. This consideration is the essential reason why FI is called an “information.”

### 2.2 Non parametric Fisher information measure versus global measures

If  $X$  is a discrete random variable that takes on a finite or countably infinite number  $N$  of values that are characterized by the probability density  $p_i, i \in \mathbf{N}$  where  $p_i$  is the probability of

$x_i$  and  $x_i \in (a, b) \subseteq \mathfrak{R}$ , is assumed to be normalized to unity so that  $\sum_{i=1}^N p_i = 1$ . In this case,  $X$  can be specified by a probability vector,  $P = \{p_1, p_2, \dots, p_N\}$ . its distribution over the interval (a; b) can be studied by using the following complementary spreading and information-theoretic measures: the variance, the FI measure [28] and the Shannon entropy [33]. The variance of the random variable  $X$  is given by the following.

$$V(X) = \sum_i (x_i - E(X))^2 p(x_i), \quad (3)$$

Where  $E(X)$  denotes the expected value of  $X$ . The nonparametric FI measure [28, 32, 34] and the Shannon entropy [33] of  $X$  are defined by the following respectively.

$$I(X) = \sum_i \frac{(p(x_{i+1}) - p(x_i))^2}{p(x_i)}, \quad (4)$$

and

$$H(X) = -\sum_i p(x_i) \log p(x_i). \quad (5)$$

These three quantities, which have a qualitatively different character, quantitatively measure the spreading of the random variable  $X$  in different and complementary ways. Shannon entropy  $H(X)$  uses uncertainty as a measure to describe the information that is contained in  $X$ . The sum in Eq. (5) can be taken in any order. Graphically, this scenario means that, if the curve  $p(x_n)$  undergoes a rearrangement of its points  $(x_n, p(x_n))$ , although the shape of the curve will drastically change the value of  $H$  remains constant.  $H$  is then said to be a global measure of the behavior of  $p(x_n)$ . The variance, which is commonly known to measure the distribution of the probability mass around the centroid, is a global measure that is, in a sense, stronger than the Shannon entropy. This relationship occurs because the variance gives a large weight to the tails of the density than the logarithmic Shannon functional. The strong dependence of the variance on the tails is not relevant, of course, when the tails fall off exponentially, which is the case for Gaussian or quasi-Gaussian distributions.

In contrast to the variance and the Shannon entropy, the FI is very sensitive to the difference in the density at adjacent points of the variable. Indeed, when the density  $p(x_n)$  undergoes a rearrangement of points  $x_n$ , although the shape of the density can change drastically the value of the entropy power remains constant according to Eq. (5), however the local slope values  $p(x_{n+1}) - p(x_n)$  change drastically and thus, the sum in Eq. (4), which defines the FI, will also change substantially [28, 34].

The preceding discussion implies that the analytical properties of the Shannon entropy and the FI measures are quite different. Thus, whereas Shannon entropy is a global measure of smoothness in  $p(x)$ , FI is a local measure. Hence, when extremized through the variation of  $p(x)$ , Fisher's form gives a differential equation whereas Shannon's form always gives directly the same form of solution, an exponential function [28]. Therefore, if one of the two measures Shannon entropy (global) or FI (local) is to be used in a variation principle in order to derive the physical law  $p(x)$  describing a general scenario, a preference is given to the local measure, FI [28, 31]. For different applications of FI measure and more comparisons between the FI measure and information-theoretic measures we refer the reader to the books by Frieden [28, 36].

## 2.2 Fisher information thresholding

Let  $I$  denote a gray-scale image with  $L$  gray levels  $[0, 1, \dots, L-1]$ . The number of pixels with gray level  $i$  is denoted by  $n_i$  and the total number of pixels is denoted by  $N = n_0 + n_1 + \dots + n_{L-1}$ . The probability of gray level  $i$  appearing in the image is defined as the following:

$$p_i = \frac{n_i}{N}, \quad p_i \geq 0, \quad \sum_{i=0}^{L-1} p_i = 1.$$

Suppose that the pixels in the image are divided into two classes  $A$  and  $B$  by a gray level cutoff  $t$ .  $A$  is the set of pixels with levels  $[0, 1, \dots, t]$ , and the remaining pixels belong to  $B$ .  $A$  and  $B$  usually correspond to the object class and the background class, or vice versa. Then the probabilities of the two classes are given by the following.

$$P_A = \frac{p_1}{w_1}, \frac{p_2}{w_1}, \dots, \frac{p_t}{w_1},$$

$$P_B = \frac{p_{t+1}}{w_2}, \frac{p_{t+2}}{w_2}, \dots, \frac{p_{L-1}}{w_2},$$

where

$$w_1(t) = \sum_{i=0}^t p_i, \quad w_2(t) = 1 - w_1(t).$$

The mean gray levels of the two classes can be defined as follows

$$m_1(t) = \sum_{i=0}^t \frac{ip_i}{w_1},$$

$$m_2(t) = \sum_{i=t+1}^{L-1} \frac{ip_i}{w_2},$$

and the corresponding class variances are given by the following:

$$\sigma_1 = \sum_{i=0}^t \frac{(i - m_1)^2 p_i}{w_1},$$

$$\sigma_2 = \sum_{i=t+1}^{L-1} \frac{(i - m_2)^2 p_i}{w_2}.$$

The within-class variance can be defined by Otsu's method [14]

$$\sigma_w^2 = w_1 \sigma_1^2 + w_2 \sigma_2^2,$$

Otsu selects a threshold  $t$  that minimizes the within-class variance  $\sigma_w^2$ , which is the following:

$$t = \arg \min_{1 \leq t \leq L} \{ \sigma_w^2 \}. \quad (7)$$

Kapur [22] has employed the entropy criterion method for determining whether optimal thresholding can provide histogram-based image segmentation with satisfactory desired characteristics. It selects a threshold  $t$  that maximizes the function

$$f(t) = H_A + H_B, \quad (10)$$

where

$$H_A = \sum_{i=0}^t \frac{p_i}{w_1} \ln \frac{p_i}{w_1} \quad \text{and} \quad H_B = \sum_{i=t+1}^{L-1} \frac{p_i}{w_2} \ln \frac{p_i}{w_2}.$$

The new method proposes the FI measure as an optimality criterion. The priori nonparametric FI for each distribution is defined as

$$I_A(t) = \frac{1}{w_1} \sum_{i=1}^t \frac{(P(x_{i+1}) - P(x_i))^2}{P(x_i)},$$

$$I_B(t) = \frac{1}{w_2} \sum_{i=t+1}^L \frac{(P(x_{i+1}) - P(x_i))^2}{P(x_i)}.$$

The FI  $I(t)$  is parametrically dependent on the threshold value  $t$  for the foreground and the background. We define the FI measure within the two classes as the following.

$$I(t) = w_1 I_A(t) + (1 - w_1) I_B(t).$$

We maximize the information measure within the two classes (the object and the background). When  $I(t)$  is maximized, the luminance level  $t$  is considered to be the optimum threshold value. This criterion can be achieved with a small computational effort.

$$t_{opt} = \arg \max [w_1 I_A(t) + (1 - w_1) I_B(t)]. \quad (12)$$

Recently Abo-Eleneen and Abdel-Azim [37, 38], proposed FI-based approach to image thresholding, to extend Kittler and Illingworth's MET method and to developed a criterion that employs the combination between the FI measure and the intensity contrast to segment a normal MRI brain images and on a glioma MRI brain images respectively.

### 2.3 Algorithm

The proposed algorithm is a simple and effective thresholding method. This technique defines a new criterion that is based on the FI corresponding to two thresholded classes and determines the optimal threshold by maximizing the criterion. The following steps describe the proposed algorithm for image segmentation:

1. Let  $max=0$  be the optimal threshold, and let  $max I$  be the maximum value of the objective function.
  2. For  $t = 1$  to Maximum of gray intensities
  3. Compute the function *objective* value that corresponds to the gray level  $t$   
If  $I(t) > max$ ,  
Then  $max = I(t)$ ,  $Topt = t$ .
  - end
- Take  $Topt$  as the optimal threshold for segmenting the image.

### 3. Experimental results

In this experiment, we implement threshold optimization based on the FI measure. To evaluate the performance of the proposed method, we apply it to a variety of images including, NDT images, text images and general real-world images. All of the images used are  $256 \times 256$  and have 8-bit (i.e. 256 gray levels) types. The results from the proposed method compare with the most commonly used methods in the literature, namely Otsu's method [14], Kapur's method [22], MEC [25], MET [26] and

Tsai's method [39]. The quality of the segmentation result is quantitatively evaluated by a misclassification error (ME) measure [13] and a relative foreground area error (RAE) measure [5, 13]. ME reflects the percentage of the background pixels incorrectly classified as, the foreground, and conversely, the percentage of foreground pixels that are erroneously assigned to background. For two-class segmentation, ME can be simply formulated as

$$ME = 1 - \frac{|B_O \cap B_T| + |F_O \cap F_T|}{|B_O| + |F_O|}, \quad (13)$$

where  $B_O$  and  $F_O$  are the background and foreground of the ground truth image,  $B_T$  and  $F_T$  are the background and foreground pixels in the thresholded image, and  $|\cdot|$  is the cardinality of a set. The value of ME varies between 0 for a perfectly classified image and 1 for a totally erroneously classified image. A lower value of ME means better quality of corresponding thresholded image. RAE reflects a feature measurement such as the area or shape, as obtained from the segmented image with respect to the reference image accuracy, which is defined as

$$RAE = \begin{cases} \frac{A_O - A_T}{A_O}, & A_T < A_O \\ \frac{A_O - A_T}{A_O}, & A_T \geq A_O \end{cases} \quad (15)$$

Where  $A_O$  is the area of the reference image, and  $A_T$  is the area of the thresholded image. Obviously, for a perfect match of the segmented regions, RAE is zero, whereas if there is zero-overlap of the object areas, the penalty is maximum one. To provide more details about the quantitative evaluation, we introduce other two measures [40] false positive rate (FPR) and false negative rate (FNR). FPR is the rate of the number of background pixels misclassified into foreground to the total number of background pixels in the ground truth image. Similarly, FNR is the rate of the number of foreground pixels misclassified into background to the total number of foreground pixels in the ground truth. For two-class segmentation, FPR and FNR can be respectively formulated as

$$FPR = \frac{|B_O \cap F_T|}{|B_O|}, \quad (16)$$

$$FNR = \frac{|F_O \cap B_T|}{|F_O|}. \quad (17)$$

### 3.1. Experiments on NDT and Text images

The first set of experiments is related to the problem of the analysis of NDT image and text image. NDT means to detect an object and quantify its possible defects without harmful effects on it by special equipments and methods. NDT is used in an abroad variety of applications, such as aeronautics and astronautics, the nuclear industry, chemistry and civil constructions [35].

In this study, two real NDT and text images are used in the experiments to assess the performance of the proposed method. The first image represents a light microscopy image of a material structure. Light microscopy is frequently used for inspecting the microstructures of materials to derive information about their properties such as the porosity, the particle sizes, the distribution uniformity, etc. The second image is a lenience plate image. Accurate segmentation of text in a plate lenience image plays a crucial role in a lenience plate recognition system. Obviously, the histogram of this gray

level data is non-Gaussian in nature. The results in terms of the thresholds, the numbers of misclassified pixels, and the values of ME, RAE, FPR and FNR are obtained by applying various methods to the images as listed in Table 1. This table shows that the segmentation results yielded by the proposed method have fewer misclassified pixels, and lower values for ME, RAE, FPR and FNR for the material image in Fig 1 and text image in Fig 2. Kapur and MET have serious under-segmentation, and conversely. Visually thresholded results are displayed in Figs. 1 and 2. From these figures, one can observe that our results are closest to the ground truth images. Hence, the proposed method in general performed reasonably well in the NDT and text experiments because of its classification-based thresholding criterion, which results in the less quantitative measure values.

**Table 1**

Image	Methods	Thrshold	FTP	FNP	MIS	ME	RAE
Material	Kapur	128	5.84E-01	0	18793	2.87E-01	3.58E-01
	Otsu	150	2.85E-01	0	9169	1.40E-01	2.15E-01
	MCE	148	3.16E-01	0	10183	1.55E-01	2.33E-01
	MET	187	0	2.76E-01	9208	1.41E-01	2.75E-01
	Tsai	146	3.46E-01	0	11143	1.70E-01	2.49E-01
	Propose	172	0	2.41E-02	804	1.23E-02	2.40E-02
Plate Lenience	Kapur	217	0	9.50E-01	33703	5.14E-01	9.49E-01
	Otsu	115	1.97E-01	0	5915	9.03E-02	1.42E-01
	MCE	106	2.23E-01	0	6687	1.02E-01	1.57E-01
	MET	83	2.90E-01	0	8705	1.33E-01	1.95E-01
	Tsai	132	1.01E-01	0	3049	4.65E-02	7.93E-02
	Propose	153	0	1.07E-02	380	5.80E-03	1.07E-02



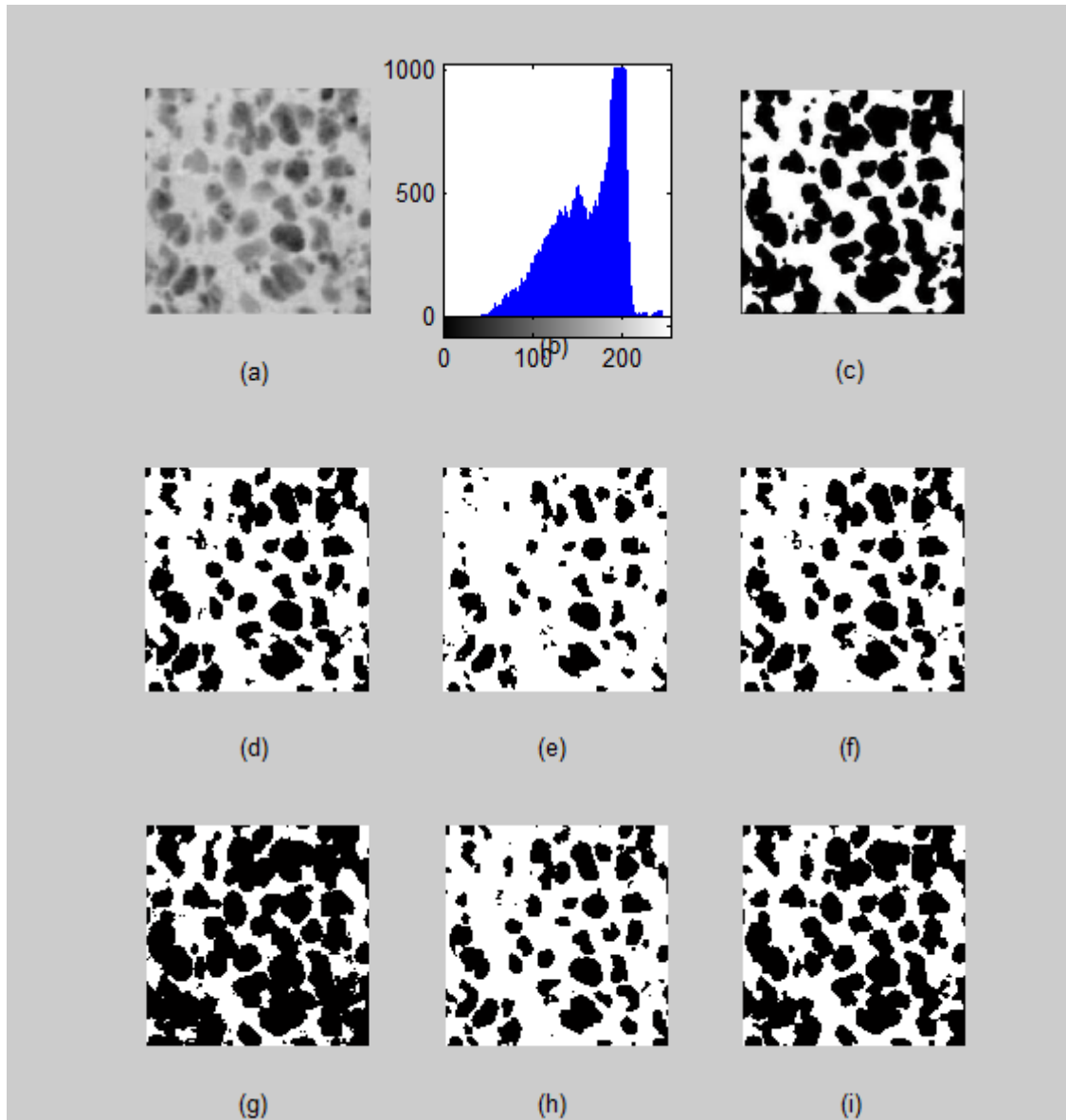


Fig.1. Thresholding results of material image: (a) original, (b) histogram, (c) ground truth image, (d) Otsu's method ( $t = 150$ ), (e) Kapur ( $t = 128$ ), (f) MCE ( $t = 148$ ), (g) MET ( $t = 187$ ), (h) Tsai ( $t = 146$ ) and (i) the proposed method ( $t = 172$ ).

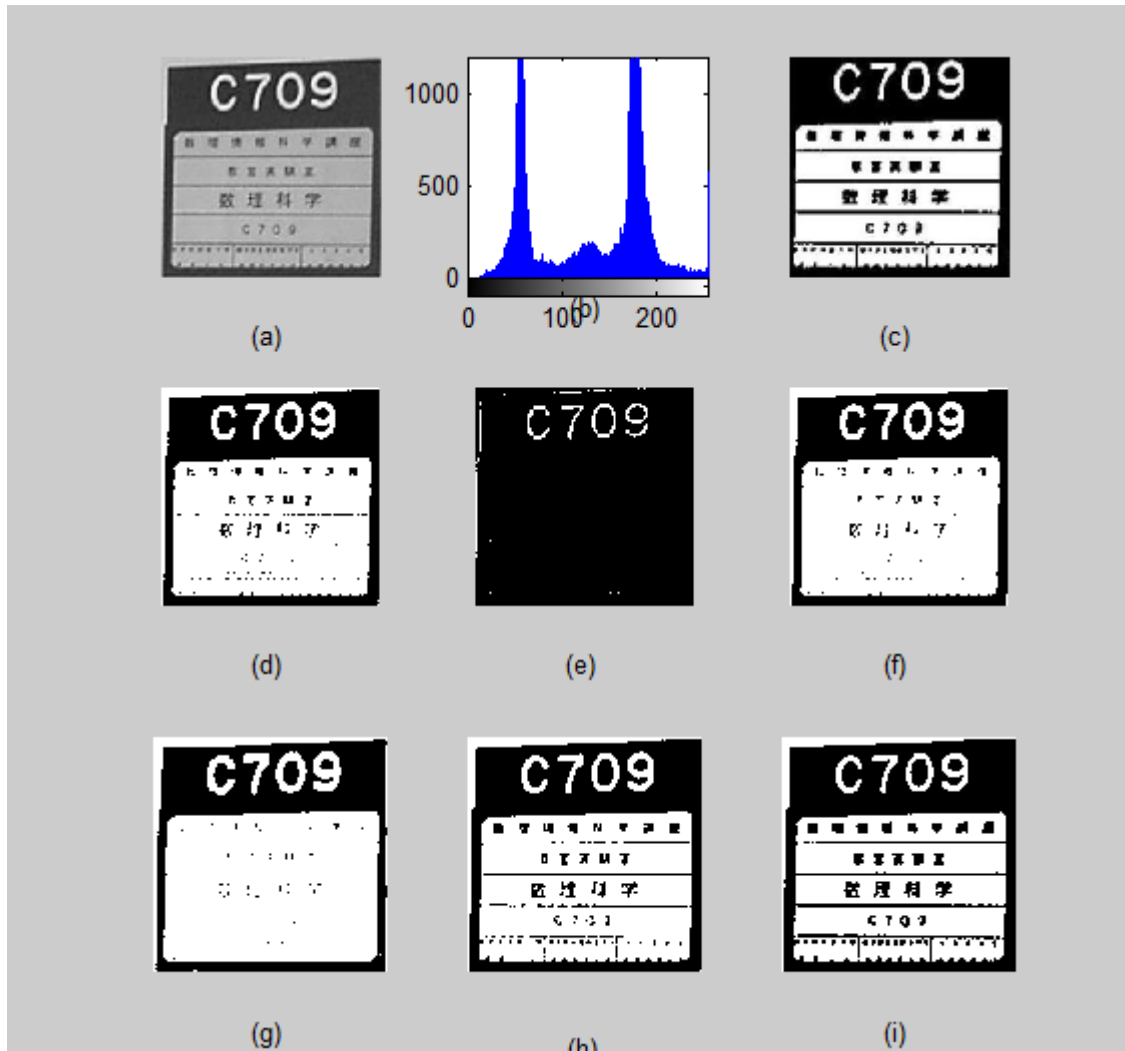


Fig.2. Thresholding results of licence plate image: (a) original, (b) histogram, (c) ground truth image, (d) Otsu's method ( $t = 115$ ), (e) Kapur ( $t = 217$ ), (f) MCE ( $t = 106$ ), (g) MET ( $t = 83$ ), (h) Tsai ( $t = 132$ ) and (i) the proposed method ( $t = 153$ ).

### 3.2. Experiments on real-world images

In this section, a variety of real-world images is used for assessing the relative performance. Four sample images are used, namely Lena, Peppers, Grapes and Girl. These images are of more complex structures than the images in NDT and Text, which are not suitable for performing a quantitative measurement of the segmentation quality. The quality of the results is compared only by visual perception. Figures 3, 4, 5 and 6 displayed the segmentation results of the three methods. From these figures, it can be easily observed that the segmentation image based on the proposed method not only segments the target from the background, but also provides a thin and similar gray level segment very nicely. Moreover, it provides prominence to the interested regions. We summarize the thresholds determined by various methods in Table 3.

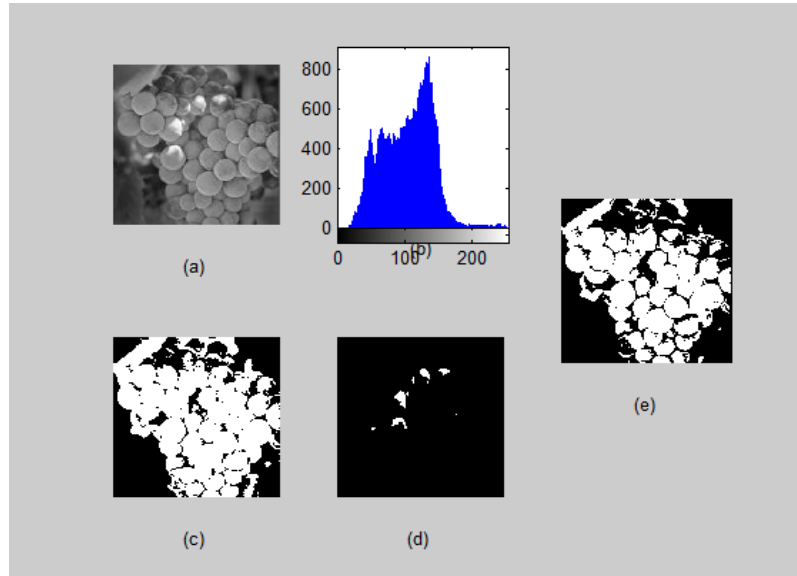


Fig.5. Thresholding results on the Grapes image: (a) original, (b) histogram, (c) Otsu's method ( $t = 98$ ), (d) Kapur ( $t = 194$ ), (e) the proposed method ( $t = 111$ ).

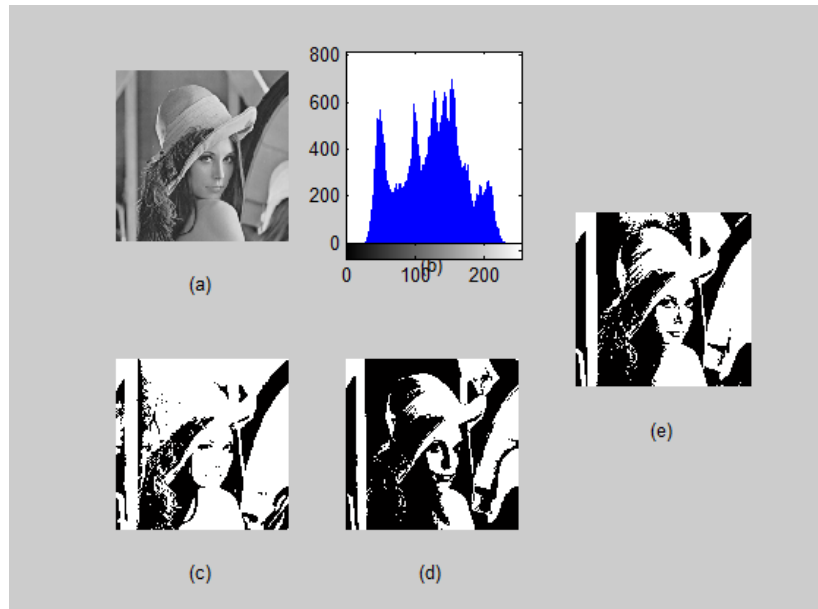


Fig.6. Thresholding results of the Lena image: (a) original, (b) histogram, (c) Otsu's method ( $t = 99$ ), (d) Kapur ( $t = 142$ ), (e) the proposed method ( $t = 129$ ).

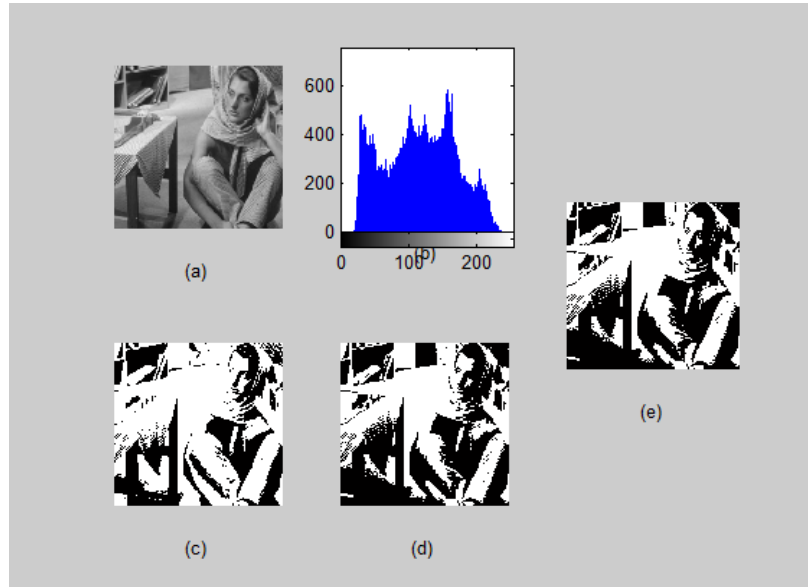


Fig.7. Thresholding results of the Girl image: (a) original, (b) histogram, (c) Otsu's method ( $t = 97$ ), (d) Kapur ( $t = 120$ ), (e) the proposed method ( $t = 121$ ).

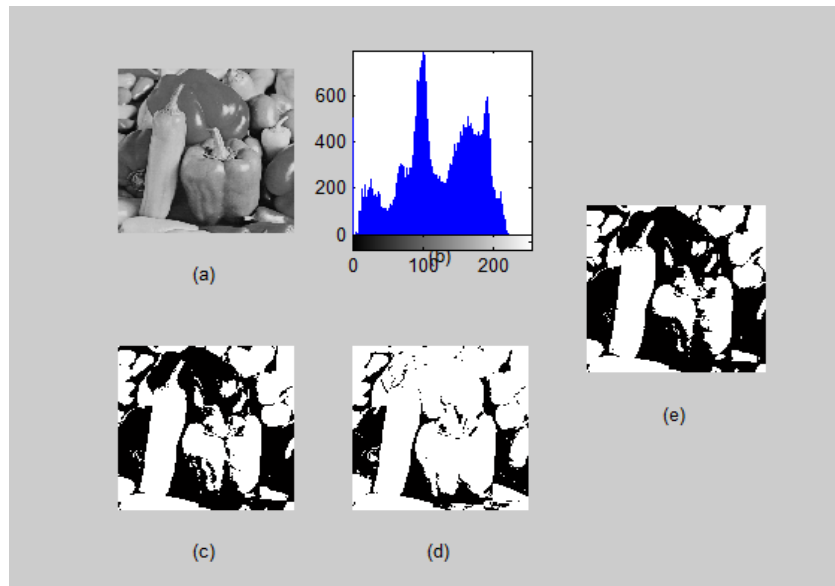


Fig 8. Thresholding results of the Peppers image: (a) original, (b) histogram, (c) Otsu's method ( $t = 114$ ), (d) Kapur ( $t = 89$ ), (e) the proposed method ( $t = 124$ ).

**Table 2**

	Otsu	Kapur	Proposed method
<b>Grapes</b>	98	194	111
<b>Lena</b>	99	142	129
<b>Girl</b>	97	120	121
<b>Peppers</b>	114	89	109

#### 4. Conclusions

FI is a measure of the state of disorder of a system or phenomenon thus, it plays an important role in terms of physical theory. In this paper, we have developed a simple but effective method of image segmentation that employs the FI measure. This method assumes that there are two probability distributions; one distribution is for the object (class), and the other distribution is for the background class. The underlying idea of the proposed method is to maximize the FI within the object and the background classes. The proposed method has the following advantages:

- 1- It is characterized by its nonparametric and unsupervised nature of threshold selection.
- 2- The implementation of the method is very simple.
- 3- The extension of the proposed method into multi-level thresholding and color images is an open problem for future exploration.

#### Conflicts of Interest

The authors declare no conflicts of interest.

#### References

- [1] R. C. Gonzalez, R. E. Woods, Digital Image Processing, Publishing House of Electronics Industry, Beijing, 2002.
- [2] D. S. Huang, Systematic Theory of Neural Networks for Pattern Recognition, Publishing House of Electronic Industry of China, Beijing, 1996.
- [3] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annu. Rev. Biomed. Eng.* 2000, Vol. 2, pp. 315-337.
- [4] S. Mallat and S. Zhong, "Characterization of Signals from Multiscale Edges," *IEEE Trans Pattern Anal Mach Intell.*, 1992, vol. 14, pp. 710-732.
- [5] Y. Zhang, "A survey of evaluation methods for image segmentation." *Pattern Recogn.* 1996, 29:1335-1346.
- [6] P. K. Saha and J. K. Udupa, Optimum image thresholding via class uncertainty and region homogeneity. *IEEE Trans Pattern Anal Mach Intell.*, 2001, 23:689–706.
- [7] T. Sund and K. Eilertsen, An algorithm for fast adaptive binarization with applications in radiotherapy imaging. *IEEE Trans Med Imaging*, 2003, 22:22–28.
- [8] Y. Solihin and C. G. Leedham, Integral ratio: a new class of global thresholding techniques for handwriting images. *IEEE Trans Pattern Anal Mach Intell.* 1999, 21:761-768.
- [9] B. Bhanu, Automatic target recognition: state of the art survey. *IEEE Trans Aerosp Electron Syst.* 1986, 22:364–379.
- [10] C. A. Glasbey, An analysis of histogram-based thresholding algorithms, *CVGIP: Graphical Models and Image Processing* 1993, 55 (6) 532–537.
- [11] P. K. Sahoo, S. Soltani, A.K.C. Wong, Y.C. Chen, A survey of thresholding techniques, *Comput. Vis. Graph. Image Process*, 1988, 41 (2) 233–260.

- [12] Ø. D. Trier, A.K. Jain, Goal-directed evaluation of binarization methods, *IEEE Trans Pattern Anal Mach Intell.*, 1995 17 (12) 1191–1201.
- [13] M. Sezgin, B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, *Journal of Electronic Imaging* 2004, 13 (1) 146–165.
- [14] N. Otsu, A threshold selection method from gray-level histogram, *IEEE Trans. Syst. Man Cybern.*, 1978, 8 62–66.
- [15] S. U. Lee, S. Y. Chung, R.H. Park, A comparative performance study of several global thresholding techniques for segmentation', *Comput. Vis. Graph. Image Process.*, 1990, 52,171–190.
- [16] R. M. Haralick, , L.G. Shapiro, *Computer and robot vision*, vol. 1' Addison Wesley, 1992.
- [17] N. R. Pal, S.K. Pal, A new definition and its application', *IEEE Trans. Syst. Man Cybern.*, 1991, 21, (5), 1260–1270.
- [18] P. K. Sahoo, D.W. Slaaf, T.A. Albert, Threshold selection using a minimal histogram entropy difference', *Opt. Eng.*, 1997, 36,1976–1981.
- [19] A. Rényi. On measures of entropy and information. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob.*, 1961, Vol. I, pp. 547–561. Univ. California Press, Berkeley, Calif.
- [20] T. Pun, A new method for grey-level picture thresholding using the entropy of the histogram, *Signal Process.*, 1980, 2, 223–237.
- [21] T. Pun, Entropic thresholding: a new approach', *Comput. Graph. Image Process.*, 1981, 16, 210–239.
- [22] J. N. Kapur, P. K. Sahoo, A.K.C. Wong, A new method for gray level picture thresholding using the entropy of the histogram, *Comput. Vision Graphics Image Process.* 1985, 29, 273–285.
- [23] A. S. Abutaleb, Automatic thresholding of gray-level pictures using two-dimensional entropies, *Pattern Recogn.*, 1989, 47 pp 22–32.
- [24] A. D. Brink, Thresholding of digital images using two-dimensional entropies, *Pattern Recogn.*, 1992, 25, 803–808.
- [25] C.H. Li, C.K. Lee, Minimum cross entropy thresholding, *Pattern Recogn.*, 1993, 26, 617–625.
- [26] J. Kittler and J. Illingworth, Minimum Cross Error Thresholding, *Pattern Recogn.*, 1986, vol. 19, pp. 41-47.
- [27] H. d. Cheng, J. R. Chen and J. G. Li, Threshold selection based on fuzzy  $c$ -partition entropy approach, *Pattern Recogn.*, 1998, 31,857–870.
- [28] B. R. Frieden, *Science from Fisher information: a unification*; 2004, Cambridge University Press: Cambridge, UK.
- [29] A. J. Stam, *Some Mathematical Properties of Quantities of Information*, 1959, Ph.D. dissertation, Technical University of Delft.
- [30] N.M. Blachman. The convolution inequality for entropy powers. *IEEE Trans. Information Theory*, 1965, 11, 267–271.
- [31] B. R. Frieden, Fisher information, disorder, and the equilibrium distribution of physics. *Phys. Rev. A* 1990, 41, 4265–4276.
- [32] R. A. Fisher. Theory of statistical estimation. *Proc. Cambridge Phil. Soc.*, 1925, 22:700-725.
- [33] T. M. Cover and J .A. Thomas, *Elements of Information Theory*. 1991, Wiley, N.Y.

- [34] J. S. Dehesa, R.J. Yanezz, R. Alvarez -Nodarsex and P. S\_anchez-Morenoy. Information-theoretic measures of discrete orthogonal polynomials. *Journal of Difference Equations and Applications*, 2004, 1-17.
- [35] Z. Hou, Q. Hu, W.L. Nowinski, On minimum variance thresholding, *Pattern Recogn. Lett.*, 2006, 27, 1732–1743.
- [36] B. R. Frieden, *Exploratory Data Analysis Using Fisher Information: 2007* Springer-Verlag London Limited.
- [37] Z. A. Abo-Eleneen and Gamil Abdel-Azim. An Improved Image Segmentation Algorithm Based on MET Method, *International Journal of Computer Science* 2012, Vol. 9, Issue 5, No 3.
- [38] Z. A. Abo-Eleneen and Gamil Abdel-Azim. A Novel Approach for MRI Brain Images Segmentation, *I.J. Image, Graphics and Signal Processing*, 2013, 3, 10-18
- [39] W. Tsai. Moment-preserving thresholding: a new approach, *Comput. Vision Graphics Image Process.*, 1985, vol. 29, 377-393.
- [40] W. A. Yasnoff, J. K. Mui, J. W. Bacus, Error measures for scene segmentation. *Pattern Recogn.* 1977, 9, 217–31.
- © 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).