Applied Soft Computing 11 (2011) 5630–5638

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Modified local entropy-based transition region extraction and thresholding

Zuoyong Li[a,b,*], David Zhang[a,c], Yong Xu[c], Chuancai Liu[d]

[a] Biometrics Research Centre, Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[b] Department of Computer Science, Minjiang University, Fuzhou 350108, China
[c] Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen 518055, China
[d] School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China

ARTICLE INFO

Article history:
Received 25 October 2009
Received in revised form 26 February 2011
Accepted 7 April 2011
Available online 16 April 2011

Keywords:
Local entropy
Human visual perception
Transition region
Thresholding
Image segmentation

ABSTRACT

Transition region-based thresholding is a newly developed image binarization technique. Transition region descriptor plays a key role in the process, which greatly affects accuracy of transition region extraction and subsequent thresholding. Local entropy (LE), a classic descriptor, considers only frequency of gray level changes, easily causing those non-transition regions with frequent yet slight gray level changes to be misclassified into transition regions. To eliminate the above limitation, a modified descriptor taking both frequency and degree of gray level changes into account is developed. In addition, in the light of human visual perception, a preprocessing step named image transformation is proposed to simplify original images and further enhance segmentation performance. The proposed algorithm was compared with LE, local fuzzy entropy-based method (LFE) and four other thresholding ones on a variety of images including some NDT images, and the experimental results show its superiority.

1. Introduction

Image segmentation attempts to extract an object (foreground) from a background on the basis of some characteristics such as gray level, color, texture and location [1–5]. It is a critical preprocessing step in image analysis and pattern recognition. Thresholding is one of the most important and effective image segmentation techniques, which is suitable for those images with distinctive gray levels in object and background. Its aim is to find an appropriate threshold to separate both parts. Thresholding result is a binary image where all pixels with gray levels higher than the determined threshold are classified as object and the rest of pixels as background, or vice versa. Thresholding can serve a variety of applications, such as biomedical image analysis [6], character identification [7], automatic target recognition [8] and quality inspection of materials [9].

Transition region-based thresholding is a kind of approaches for image segmentation in recent years [10–12]. Gerbrands [10] first demonstrated the existence of transition region in an image. Zhang and Gerbrands [11] introduced transition region into image segmentation, and presented a transition region descriptor, i.e., effective average gradient (EAG). It first calculates two image matrices via clip transformation function, then computes effective average gradients of the two matrices and gets two EAG(L) ~ L curves, finally determines two gray levels according to the peaks of the two curves and obtains a gray level interval. The pixels with gray levels in the interval are classified into transition region. EAG has some limitations: (1) gradient reflects only sudden gray level changes, which can not depict transition region accurately, (2) gradient-based method is much sensitive to noise, (3) in some cases, EAG can not extract transition region for an incorrect gray level interval as proved by Groenewald et al. [13]. In order to eliminate the above limitations, a descriptor, local entropy (LE) [12], is introduced. LE represents frequent gray level changes, and better captures nature of transition region than EAG.

LE outperforms EAG in terms of transition region description. Nevertheless it suffers a deficiency, i.e., LE takes only frequency of gray level changes into consideration, but neglects the degree of these changes. Real images may have frequent gray level changes in some small neighborhoods of object or background, but the degrees of these changes are so slight that they would not cause wrong judgment by visual perception. In this case, the slight but frequent gray level changes will increase local entropies of these neighborhoods and make the pixels corresponding to the neighborhoods be erroneously divided into transition region. To eliminate the limitation of LE, Zhang et al. [14] developed a modified version, namely LFE, by treating an image as a fuzzy set and defining its local fuzzy entropy. LFE indirectly reflects degree of gray level changes via membership function and fuzzy entropy. Unfortunately, the indirect reflec-

* Corresponding author at: Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong; and Department of Computer Science, Minjiang University, Fuzhou 350108, China. Tel.: +86 13906926400; fax: +86 059183761607.
E-mail addresses: fzulzytdq@126.com, fzulzytdq@yahoo.com.cn (Z. Li).

1568-4946/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2011.04.001

4	3	5		25	46	98
6	2	4		78	9	25
3	5	6		46	98	78

Fig. 1. Gray level changes in different local neighborhoods.

tion for the degree of changes is inadequate and instable, causing unsatisfactory effect. In this paper, a modified local entropy-based method (MLE) is presented for improving performance on transition region extraction and thresholding. A new descriptor depicting both frequency and degree of gray level changes is developed in MLE. Furthermore, in the light of human visual perception, an effective preprocessing step, namely image transformation, is also introduced to simplify original images and assists transition region extraction and thresholding. The performance of MLE was compared with LE, LFE and four other classic thresholding methods by testing a variety of real world images. Experimental results show that MLE greatly outperforms LE and LFE in terms of transition region extraction and image thresholding. In addition, comparisons against four classic methods [15–18] also demonstrate its superiority.

The remainder of this paper is organized as follows: Sections 2 and 3 review properties of transition region and human visual perception, respectively. The proposed method is presented in Section 4. The performance of the new method is tested on a variety of real world images and compared with LE, LFE and other thresholding methods in Section 5. Conclusions appear in Section 6.

2. Transition region

Transition region located between object and background has the following characteristics:

(1) Region characteristic: whatever for step edge or non-step edge, there always exists transition region near an edge [10]. Transition region near a non-step edge has certain pixel width, while transition region around a step edge should have at least one pixel width. In real world images, transition region near a step edge usually has several pixel widths, owing to sampling error.
(2) Boundary characteristic: transition region is located between object and background, and covers around the object.
(3) Variation of gray levels: pixels' gray levels in transition region usually change frequently and intensively, bringing about abundant information for transition region description. Gradient is a good measure for sudden gray level changes, but inapplicable in measuring frequent gray level changes. Local entropy is suitable to represent frequent gray level changes, but unable to reflect degree of gray level changes. For example, there are two local neighborhoods in Fig. 1. The number in both neighborhoods denotes gray level of pixel. From Fig. 1, one can see that the two neighborhoods have the same changing frequency in gray level, but the degree of changes in the right one is obviously sharper than that in the left one. Since a slight difference in gray level is not easy to be observed by visual perception, the right neighborhood is more likely divided into transition region, with the left one liable to being classified into non-transition region. Because LE only considers frequency of gray level changes, the two neighborhoods have the same probability of being divided into transition region. This is obviously unreasonable. Therefore, a new descriptor considering both frequency and degree of gray level changes is developed to depict transition region more adequately in this paper.

3. Human visual perception

Human visual perception [19] has the following characteristics.

(1) Human eye is insensitive to features present at the both extremes of pixel intensity, whereas sensitive to distinguishing features at the mid-range intensities. This suggests a focus on mid-region of a gray scale image, i.e., around the mean, when segmenting images.
(2) A lot of images may have either histograms with high intensity values or more structures near a certain value (usually the mean) than that farther from the mean. A rough estimation of such a histogram exhibits a Gaussian distribution.

Transition region is geometrically located between object and background, and composed of pixels having intermediate gray levels between that of object and of background [20]. In the light of human visual perception, a preprocessing step called image transformation is suggested to simplify original images. Its basic idea is as follows: suppose that gray levels of transition region are within a range (θ_1, θ_2), which can be determined in an unsupervised way, a transformation is then applied so that only pixels with gray levels inside this range will contribute to transition region. The transformation preserves gray level changes of mixture between object and background, meanwhile weakens the changes of non-transition region. Thus simplifies originals, which should be helpful for transition region extraction and subsequent thresholding.

4. Transition region-based thresholding

4.1. Local entropy

Local entropy was first used to describe transition region in [12]. Without losing generality, let I be an image with L gray levels [0,1, ..., $L-1$]. The number of pixels at gray level i is denoted by n_i and the total number of pixels by $N = n_0 + n_1 + \ldots + n_{L-1}$. Suppose $V = \{(i,j): i = 1,2,\ldots,n_h; j = 1,2,\ldots,n_w\}$, where n_h and n_w are the height and width of the image, respectively. Let $f(i, j)$ be the gray level at pixel (i, j). Following Shannon's definition [21] of entropy, Pun [22] defined the entropy of an image as

$$E = -\sum_{i=0}^{L-1} P_i \log P_i \tag{1}$$

where

$$P_i = \frac{n_i}{N} \tag{2}$$

is the probability of gray level i appeared in the image.

Given a $m \times n$ neighborhood window centered on pixel (i, j), Ω, its entropy, also named as local entropy of the pixel, can be formulated as

$$Le(i, j) = E(\Omega) = -\sum_{k=0}^{L-1} P_k \log P_k \tag{3}$$

where

$$P_k = \frac{n_k}{m \times n} \tag{4}$$

is the probability of gray level k appeared in the neighborhood, and n_k the number of pixels with level k in Ω.

4.2. Modified transition region descriptor

Local entropy can depict frequency of gray level changes. However, its computational complexity is high, as the calculation of local entropy involves statistical analysis for the pixels' gray levels and computes each gray level's probability appeared in a neighborhood. Furthermore, the process involves logarithm and multiplication operations. To reduce the computational complexity, a simple form with similar effect, local complexity [23], is used to describe frequency of gray level changes,

$$Lc(i,j) = C(\Omega) = \sum_{k=0}^{L-1} \mathrm{sgn}(k) \tag{5}$$

where

$$\mathrm{sgn}(k) = \begin{cases} 1 & \text{if } \exists f(x,y) = k, \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

and (x, y) is a pixel coordinate in the neighborhood Ω.

Both local entropy and local complexity consider only frequency of gray level changes, while neglect degree of the changes. Therefore, variance, a common statistical measure reflecting degree of deviations between mean and individuals, is used to describe the degree of gray level changes. For the neighborhood Ω, its local variance can be formulated as

$$Lv(i,j) = \sigma^2(\Omega) = \frac{1}{m \times n - 1} \sum_{x=1}^{m} \sum_{y=1}^{n} (f(x,y) - \bar{f})^2 \tag{7}$$

where $\bar{f}$ is the gray level mean of Ω.

According to this way, when we move the neighborhood window pixel by pixel within the image from left to right and top to bottom, each pixel's local complexity and variance can be obtained and constitutes two following image matrices

$$Lc = \begin{bmatrix} Lc(1,1) & Lc(1,2) & \ldots\ldots & Lc(1,n_w) \\ Lc(2,1) & Lc(2,2) & \ldots\ldots & Lc(2,n_w) \\ \cdots & \cdots & \ldots\ldots & \cdots \\ Lc(n_h,1) & Lc(n_h,2) & \ldots\ldots & Lc(n_h,n_w) \end{bmatrix} \tag{8}$$

$$Lv = \begin{bmatrix} Lv(1,1) & Lv(1,2) & \ldots\ldots & Lv(1,n_w) \\ Lv(2,1) & Lv(2,2) & \ldots\ldots & Lv(2,n_w) \\ \cdots & \cdots & \ldots\ldots & \cdots \\ Lv(n_h,1) & Lv(n_h,2) & \ldots\ldots & Lv(n_h,n_w) \end{bmatrix} \tag{9}$$

To depict gray level changes of transition region adequately, local complexity and local variance are synthesized as a new descriptor. In the process, the two factors are first normalized via the following way for avoiding one factor being neglected due to large differences between their values.

$$NLc(i,j) = \frac{Lc(i,j) - \min\limits_{\forall(x,y)} Lc(x,y)}{\max\limits_{\forall(x,y)} Lc(x,y) - \min\limits_{\forall(x,y)} Lc(x,y)} \tag{10}$$

$$NLv(i,j) = \frac{Lv(i,j) - \min\limits_{\forall(x,y)} Lv(x,y)}{\max\limits_{\forall(x,y)} Lv(x,y) - \min\limits_{\forall(x,y)} Lv(x,y)} \tag{11}$$

Both normalized factors are then synthesized as a new transition region descriptor

$$S(i,j) = \beta \times NLc(i,j) + (1-\beta) \times NLv(i,j) \tag{12}$$

where β is a weight balancing contributions of normalized local frequency and local variance. When $\beta = 1$, the new descriptor degenerates to local complexity. And conversely, when $\beta = 0$, it is equivalent to local variance. Therefore, β should be between 0 and 1. Each pixel's S value constructs an image matrix S.

Both local complexity and local variance are related to gray level changes. The more frequent and intensive the changes are, the larger their values. Accordingly, the pixels in transition region have larger S values than those in non-transition region. This should be used to extract transition region of an image.

4.3. Image transformation

Separation of object and background in a gray level image could be attributed to their difference on gray level. That is, transitional pixels' gray levels are usually located between those of object and of background. This coincides with the property of transition region, i.e., sensitiveness on the features present at the mid-range of a gray scale image, as the range usually faces more frequent and intensive gray level changes. However, if the rough mid-range (i.e., mid-gray level interval) is directly used to extract transition region just like EAG [11], it is unreasonable. Here, this rough range is only utilized to preprocess an image for simplifying it by image transformation.

In order to implement image transformation, a range (θ_1, θ_2) should first be obtained. In the light of human visual perception, the range may be found via image statistical characteristics. For an image I, θ_1 and θ_2 can be determined via the following steps:

(1) Compute mean and standard deviation of the image according to the following equations

$$\mu = \frac{1}{N} \sum_{i=0}^{L-1} i n_i \tag{13}$$

$$\sigma = \left(\frac{1}{N-1} \sum_{i=0}^{L-1} (i-\mu)^2 n_i \right)^{1/2} \tag{14}$$

(2) Determine θ_1 and θ_2 as

$$\theta_1 = \mu - \alpha \times \sigma \tag{15}$$

$$\theta_2 = \mu + \alpha \times \sigma \tag{16}$$

where α is a parameter and its value can be automatically determined by optimizing the statistical criterion σ_S in the literature [24]. Take material image in Fig. 2(a) as an example. The region confined by (θ_1, θ_2) is shown in Fig. 2(b), where bright pixels are our focuses. The figure shows that the region confined by the range is a rough mixture between object and background.

Once θ_1 and θ_2 are determined, image transformation can be followed immediately

$$f_{tr}(i,j) = \begin{cases} \theta_1 & \text{if } f(i,j) < \theta_1, \\ f(i,j) & \text{if } \theta_1 \le f(i,j) \le \theta_2, \\ \theta_2 & \text{if } f(i,j) > \theta_2. \end{cases} \tag{17}$$

The transformation weakens gray level changes in both object and background simultaneously, thus simplifying the original image. The weakening effect is favorable to transition region extraction and image segmentation. Take transformed form of the material image in Fig. 2(c) as an example. Corresponding gray level range is (157, 165). From the figure, one can conclude that gray level changes of object and background have been weakened obviously, and the transformed image becomes much simpler than the original. The simplifying effect would be helpful to transition region extraction and thresholding.

4.4. Transition region extraction and thresholding

The proposed method aims to improve performance of transition region extraction and thresholding from two aspects. One is to simplify an image by image transformation in a preprocessing step. The other is to seek better transition region descriptor. Detailed process of transition region extraction and thresholding is as follows:

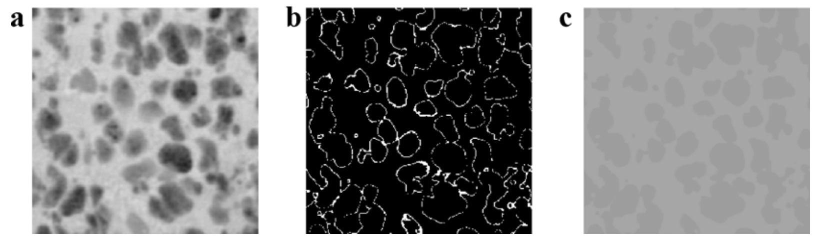

Fig. 2. Image transformation: (a) original material image, (b) the region confined by (θ_1, θ_2), (c) transformed form.

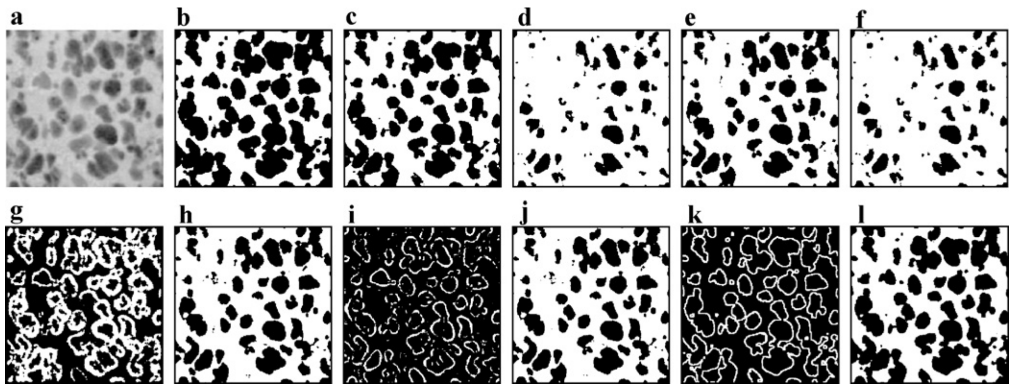

Fig. 3. Segmentation of material image: (a) original, (b) ground truth image, (c) HOU, (d) KAPUR, (e) TSALLIS, (f) PARZEN, (g) transition region extracted by LE, (h) LE's segmentation result, (i) transition region extracted by LFE, (j) LFE's segmentation result, (k) transition region extracted by MLE and (l) MLE's segmentation result.

(1) Find gray level range (θ_1, θ_2) by Eqs. (15) and (16), and implement image transformation by Eq. (17).
(2) Calculate each pixel's S value in the transformed image by Eq. (12), and construct an image matrix S.
(3) Obtain the following threshold S_T for transition region extraction.

$$S_T = \gamma \times S_{\max} \tag{18}$$

where

$$S_{\max} = \max_{\forall(i,j)} S(i,j) \tag{19}$$

and γ is a coefficient between 0 and 1.
(4) Extract transition region via the following way

$$TR(i,j) = \begin{cases} 1 & \text{if } S(i,j) \geq S_T, \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

(5) Final segmentation threshold T^* is taken as gray level mean of transition region [11,12], i.e.,

$$T* = \frac{\sum_i \sum_j TR(i,j) \times f(i,j)}{\sum_i \sum_j TR(i,j)} \tag{21}$$

(6) Binarize the image by T^*.

5. Experimental results

To evaluate the performance of the new approach (MLE), a variety of real world images including some nondestructive testing (NDT) images were chosen as testing samples. The results were compared with those obtained by HOU [16], KAPUR [17], TSALLIS [15], PARZEN [18], LE [12] and LFE [14]. Quality of thresholding is quantitatively evaluated via misclassification error (ME) [25], which regards image segmentation as a pixel classification process. ME reflects the percentage of background pixels incorrectly classified into foreground, and conversely, foreground pixels erro-

Table 1
Thresholds, numbers of misclassified pixels, ME values and running times obtained by applying various methods to the NDT images.

Images	Thresholding methods						
	HOU	KAPUR	TSALLIS	PARZEN	LE	LFE	MLE
Material							
Threshold	157	119	143	120	147	146	161
Misclassified pixels	5343	22021	12328	21640	10377	10905	3607
ME	0.081528	0.33601	0.18811	0.3302	0.15834	0.1664	0.055038
Running time (s)	2.86	0.062	1.125	16.578	14.703	21.516	5.391
Cell							
Threshold	149	172	171	171	183	229	224
Misclassified pixels	8752	7276	7375	7375	5513	5864	3357
ME	0.13354	0.11102	0.11253	0.11253	0.084122	0.089478	0.051224
Running time (s)	4.047	0.031	2	86.156	15.172	20.484	5.188
PCB							
Threshold	112	158	161	161	137	112	89
Misclassified pixels	2029	13852	14426	14426	6828	2029	1224
ME	0.03096	0.21136	0.22012	0.22012	0.10419	0.03096	0.018677
Running time (s)	2.782	0.031	0.047	30.469	13.797	20.297	5.329

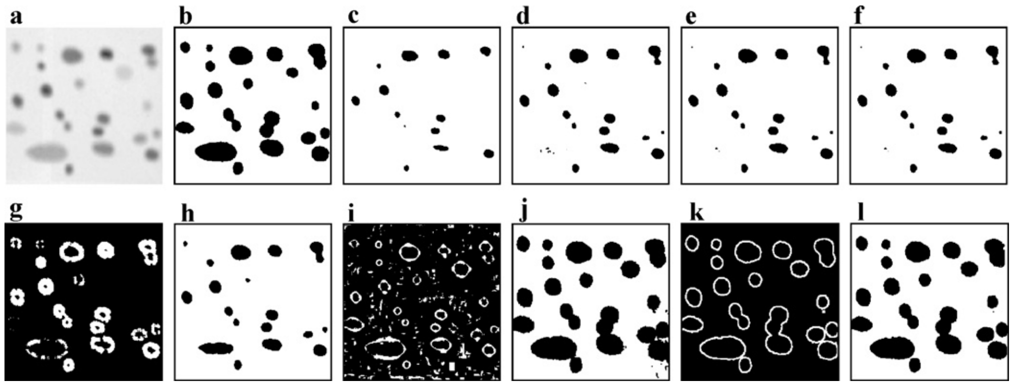

Fig. 4. Segmentation of cell image: (a) original, (b) ground truth image, (c) HOU, (d) KAPUR, (e) TSALLIS, (f) PARZEN, (g) transition region extracted by LE, (h) LE's segmentation result, (i) transition region extracted by LFE, (j) LFE's segmentation result, (k) transition region extracted by MLE and (l) MLE's segmentation result.

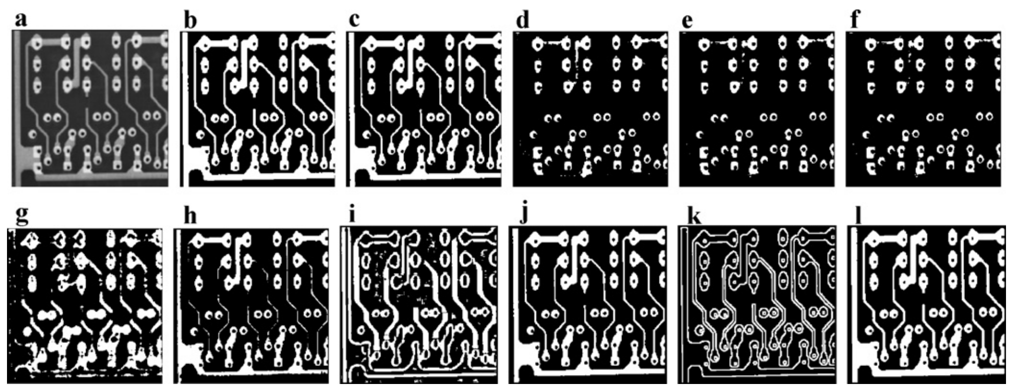

Fig. 5. Segmentation of PCB image: (a) original, (b) ground truth image, (c) HOU, (d) KAPUR, (e) TSALLIS, (f) PARZEN, (g) transition region extracted by LE, (h) LE's segmentation result, (i) transition region extracted by LFE, (j) LFE's segmentation result, (k) transition region extracted by MLE and (l) MLE's segmentation result.

neously assigned to background. For a two-class segmentation, ME can be simply formulated as

$$\mathrm{ME} = 1 - \frac{|B_{\mathrm{O}} \cap B_T| + |F_{\mathrm{O}} \cap F_{\mathrm{T}}|}{|B_{\mathrm{O}}| + |F_{\mathrm{O}}|} \tag{22}$$

where B_{O} and F_{O} are the background and foreground of the ground truth image, B_{T} and F_{T} the background and foreground pixels in the thresholded image, and $|\cdot|$ cardinality of a set. The value of ME varies between 0 for a perfectly classified image to 1 for a totally erroneously classified one. A lower value of ME means better quality. In the proposed method, the neighborhood size is 3×3, β and γ are 0.3 and 0.1, respectively. For LE and LFE, the neighborhood sizes are 7×7 and 11×11, the parameter about entropy threshold is 0.85 and 0.6, respectively. All experiments are performed on a notebook PC with 2.13G Intel Core 2 Duo CPU and 3G RAM. All the images used in the experiments are of 256×256 pixels and 8-bit (i.e., 256 gray levels).

5.1. Experiments on NDT images

Three NDT images were first chosen and compared. They are a light microscopy form of a material structure, a cell image and a PCB. NDT means to detect an object and quantify its possible defects without harmful effects on it by special equipments and methods. It is used in a broad variety of applications, such as aeronautics and astronautics, nuclear industry, chemistry and civil constructions.

The results in terms of thresholds, numbers of misclassified pixels, ME values and running times obtained by applying various methods to the images are listed in Table 1. The table shows that our segmentation results have less misclassified pixels and lower ME values, implying better performance. This observation can further be judged by comparing visual segmentation results in Figs. 3–5. From the figures, one can conclude that MLE extracts transition regions more accurately and generates results closest to the ideal ones. By comparison, LE and LFE fail to accurately extract transition regions of the images and get bad results. Other methods are also

Table 2
Thresholds, numbers of misclassified pixels and ME values obtained by applying various methods to the simple images.

Images	Thresholding methods						
	HOU	KAPUR	TSALLIS	PARZEN	LE	LFE	MLE
Potatoes							
Threshold	143	71	64	65	135	122	106
Misclassified pixels	1805	936	1823	1593	1475	951	341
ME	0.027542	0.014282	0.027817	0.024307	0.022507	0.014511	0.0052032
Running time (s)	2.985	0.046	0.39	17.938	13.875	19.953	5.266
Block							
Threshold	192	138	17	128	85	47	60
Misclassified pixels	19431	16794	2589	13910	3814	660	46
ME	0.29649	0.25626	0.039505	0.21225	0.058197	0.010071	0.0007019
Running time (s)	2.734	0.032	0.422	58.719	12.578	19.25	5.203

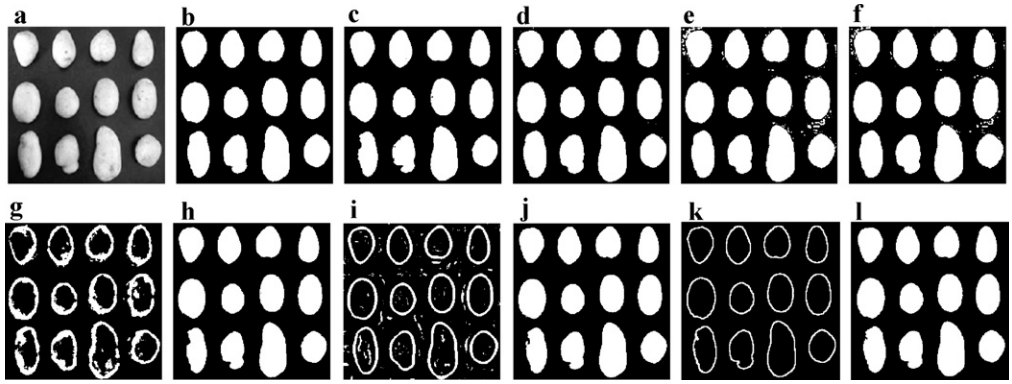

Fig. 6. Segmentation of potatoes image: (a) original, (b) ground truth image, (c) HOU, (d) KAPUR, (e) TSALLIS, (f) PARZEN, (g) transition region extracted by LE, (h) LE's segmentation result, (i) transition region extracted by LFE, (j) LFE's segmentation result, (k) transition region extracted by MLE and (l) MLE's segmentation result.

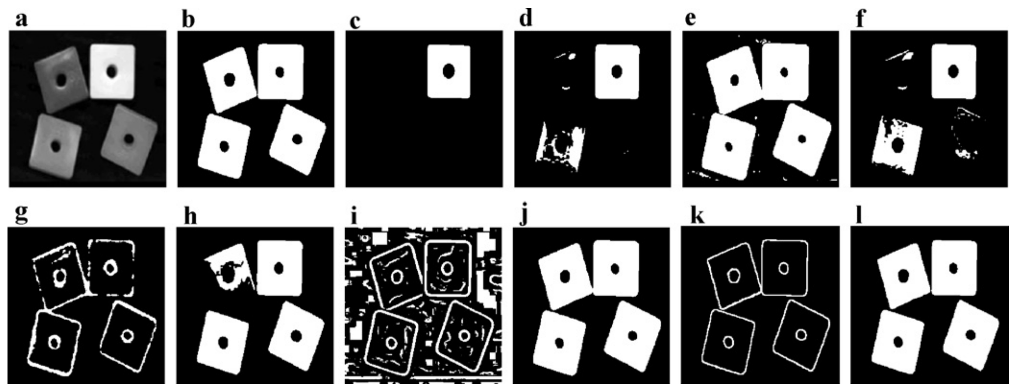

Fig. 7. Segmentation of block image: (a) original, (b) ground truth image, (c) HOU, (d) KAPUR, (e) TSALLIS, (f) PARZEN, (g) transition region extracted by LE, (h) LE's segmentation result, (i) transition region extracted by LFE, (j) LFE's segmentation result, (k) transition region extracted by MLE and (l) MLE's segmentation result.

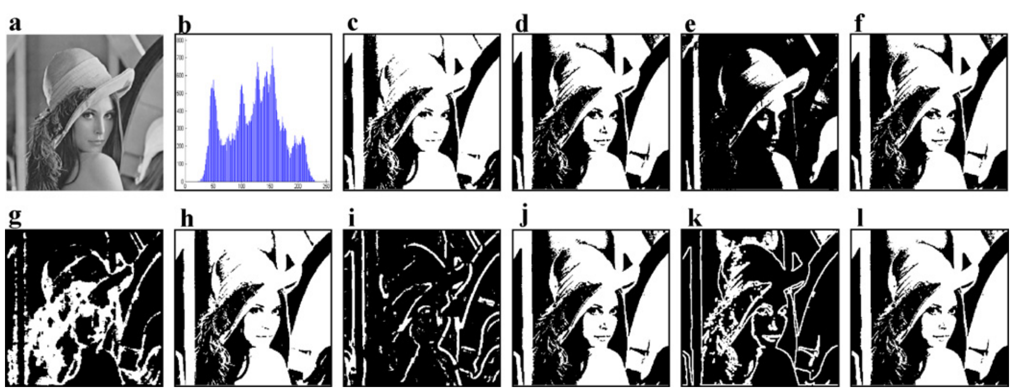

Fig. 8. Segmentation of lena image: (a) original, (b) histogram, (c) HOU, (d) KAPUR, (e) TSALLIS, (f) PARZEN, (g) transition region extracted by LE, (h) LE's segmentation result, (i) transition region extracted by LFE, (j) LFE's segmentation result, (k) transition region extracted by MLE and (l) MLE's segmentation result.

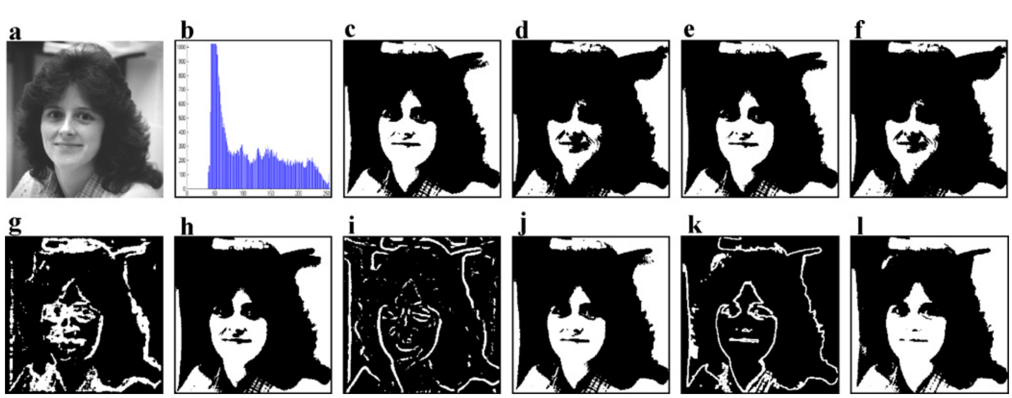

Fig. 9. Segmentation of woman image: (a) original, (b) histogram, (c) HOU, (d) KAPUR, (e) TSALLIS, (f) PARZEN, (g) transition region extracted by LE, (h) LE's segmentation result, (i) transition region extracted by LFE, (j) LFE's segmentation result, (k) transition region extracted by MLE and (l) MLE's segmentation result.

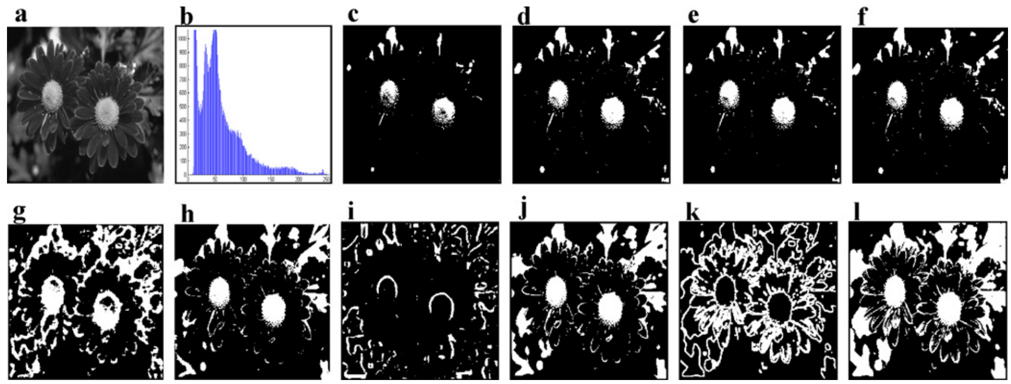

Fig. 10. Segmentation of flower image: (a) original, (b) histogram, (c) HOU, (d) KAPUR, (e) TSALLIS, (f) PARZEN, (g) transition region extracted by LE, (h) LE's segmentation result, (i) transition region extracted by LFE, (j) LFE's segmentation result, (k) transition region extracted by MLE and (l) MLE's segmentation result.

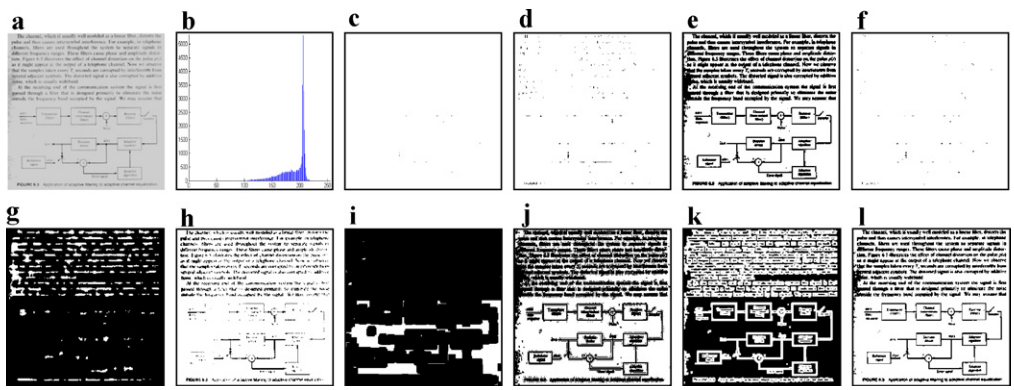

Fig. 11. Segmentation of text image: (a) original, (b) histogram, (c) HOU, (d) KAPUR, (e) TSALLIS, (f) PARZEN, (g) transition region extracted by LE, (h) LE's segmentation result, (i) transition region extracted by LFE, (j) LFE's segmentation result, (k) transition region extracted by MLE and (l) MLE's segmentation result.

unable to obtain satisfactory results. In addition, we can observe that MLE is obviously faster than LE and LFE, and its running time are about half of LE's and one third of LFE's, respectively. The main reason is that its neighborhood size is 3×3, while neighborhood sizes of LE and LFE are 7×7 and 11×11. When calculating each pixel's local characteristic under three transition region descriptors, MLE, LE and LFE involves 9, 49 and 121 pixels, respectively. This saves a lot of time for MLE. Among three transition region-based approaches, LFE is slowest, owing to its largest neighborhood size and extra calculation for each pixel's membership.

5.2. Experiments on real world images

In this section, 6 images were selected for test purpose. They fall into two groups. One group is for simple images, and the other for complex ones.

Quantitative comparisons of segmentation results for the first group are listed in Table 2. It shows that MLE achieves better results with less misclassified pixels and lower ME values. Visual thresholding results are displayed in Figs. 6 and 7. From the figures, one can observe that MLE successfully extracts transition regions and gets satisfying results. LE misclassifies partial object regions into transition regions. The reason is that there exist some object regions with frequent but slight gray level changes in the images, which are easy to be erroneously assigned to transition regions by LE as explained in Section 2. In addition, LFE also wrongly divides some non-transition regions in foreground or background into transition regions, resulting in unsatisfactory segmentation results.

Experimental results on the complex images are shown in Figs. 8–11. The first three images have complex structures and contain multiple transition regions of different gray levels, which make transition region extraction more difficult. The last image is a document with low contrast and uneven illumination. Since they are complex and have not specific objects, quantitative measurement is inapplicable to them. Quality of segmentation results is only judged by visual perception. Figs. 8–11 indicate that MLE extracts transition regions more accurately and obtains better results than LE and LFE. Our method also outperforms other four methods.

Table 3
Noise density of salt and pepper noise (d) and mean misclassification error (MME) of various methods.

PCB image	Thresholding methods						
	HOU	KAPUR	TSALLIS	PARZEN	LE	LFE	MLE
d	MME						
0.1	0.085527	0.053619	0.17035	0.10278	0.13668	0.20253	0.054379
0.2	0.30163	0.11896	0.22994	0.15903	0.17569	0.18331	0.10108
0.3	0.32418	0.18343	0.26424	0.21077	0.21706	0.19967	0.15592
0.4	0.34647	0.23878	0.29865	0.26119	0.25383	0.23326	0.20515
0.5	0.36904	0.29069	0.3339	0.30924	0.29095	0.27111	0.25956

Table 4
Noise density of Gaussian noise (σ^2) and mean misclassification error (MME) of various methods.

PCB image	Thresholding methods						
	HOU	KAPUR	TSALLIS	PARZEN	LE	LFE	MLE
σ^2	MME						
0.1	0.077153	0.059793	0.10407	0.060289	0.098727	0.10928	0.14405
0.2	0.078049	0.088467	0.33295	0.11793	0.13671	0.064482	0.14864
0.3	0.069206	0.1951	0.41073	0.36024	0.18693	0.069948	0.15976
0.4	0.058754	0.46165	0.41521	0.54954	0.26626	0.13371	0.18291
0.5	0.07206	0.58844	0.4154	0.59643	0.35462	0.23571	0.22391

5.3. Experiments on noisy images

The performance of various methods in the presence of noise is studied in this section. Salt & pepper noise and Gaussian noise with certain noise density are considered. The PCB image is used to assess the performance of various approaches in the presence of degeneration. Since salt & pepper noise or Gaussian noise added to an image is random, we run the simulation 10 times to get mean misclassification error (MME) for each noise density. Quantitative results in Table 3 indicate that MLE obtains best results under salt and pepper noise. Table 4 shows that MLE has intermediate effect under Gaussian noise.

5.4. Parameter selection

In the proposed method, there are four parameters: α, β, γ and neighborhood size. However, α can be automatically determined by the statistical criterion in the literature [24]. Hence, only β, γ and neighborhood size are left uncertain. The parameter β is used to balance contributions of local complexity and local variance in transition region description. Higher β corresponds to larger weight of local complexity. Another parameter γ aims to control the number of pixels in transition region. The bigger the γ is, the less the number. Image transformation is an independent preprocessing step, which is irrespective of the latter three parameters. To discuss the three parameters, a series of experiments on the first five images without image transformation under their different combinations have been carried out. Table 5 lists experimental results. From the table, one can observe that: neighborhood size of 3×3 results in lowest average MME (mean misclassification error) value. However, maximum difference of average MME values under different neighborhood sizes is only 0.0089, which shows that our method is little affected by selection of parameters. In order to reduce computational complexity, a 3×3 is chosen as neighborhood size in our method. Accordingly, β and γ are 0.3 and 0.1, as there is minimum MME in this case.

5.5. Computational complexity

Theoretical analysis of MLE's computational complexity is presented in this section. For segmenting an image of N pixels and L gray levels with a neighborhood size $n \times n$, the first step is to compute the mean and standard deviation of the image for implementing image transformation, which has a time complexity of $O(L)$. The next step is to calculate each pixel's local complexity and local variance in the transformed image and construct their respective image matrix, which has complexity of $O(n^2N)$, where $n << N$. The third step is to normalize the above two matrices and form a new matrix under the synthesized descriptor, whose complexity is $O(N)$. The next step is to find a threshold for the synthesized image matrix and extract transition region, which has complexity $O(N)$. The last step is to determine the segmentation threshold as gray level mean of transition region, and corresponding complexity is $O(N_t)$, where N_t is the total number of the pixels in transition region ($N_t < N$). Thus the total computational complexity of MLE is $O(N)$. MLE, LE and LFE have the same computational complexity $O(N)$. But practical experimental results show that MLE is faster than other two methods. As compared with LE and LFE, MLE needs extra time $O(L)$, where $L << N$, to perform image transformation. However it is worth mentioning that neighborhood size of our method is only

Table 5
Mean misclassification error (MME) and average MME obtained by applying the proposed method to the first five images (i.e., material, cell, PCB, potatoes and block) under different combinations of β, γ and neighborhood size.

	γ=0.1	0.2	0.3	0.4	0.5
Size = 3					
β = 0.1	0.0596	0.0815	0.1105	0.1297	0.1356
0.2	0.0367	0.0649	0.0866	0.1158	0.1339
0.3	0.0253	0.0554	0.0662	0.0929	0.1225
0.4	0.0262	0.0367	0.0569	0.0696	0.0984
0.5	0.0262	0.0367	0.0471	0.0635	0.0703
Average MME		0.0740			
Size = 5					
β = 0.1	0.0607	0.0805	0.1037	0.1264	0.1347
0.2	0.0549	0.0666	0.0874	0.1136	0.1306
0.3	0.0428	0.0621	0.0719	0.0950	0.1204
0.4	0.0311	0.0567	0.0667	0.0772	0.1032
0.5	0.0264	0.0504	0.0625	0.0703	0.0838
Average MME		0.0792			
Size = 7					
β = 0.1	0.0583	0.0796	0.1027	0.1198	0.1355
0.2	0.0548	0.0687	0.0901	0.1142	0.1316
0.3	0.0495	0.0628	0.0771	0.1006	0.1233
0.4	0.0424	0.0595	0.0691	0.0874	0.1122
0.5	0.0358	0.0546	0.0641	0.0770	0.0976
Average MME		0.0827			
Size = 9					
β = 0.1	0.0533	0.0766	0.0992	0.1189	0.1345
0.2	0.0518	0.0675	0.0897	0.1114	0.1321
0.3	0.0485	0.0598	0.0794	0.1027	0.1237
0.4	0.0433	0.0569	0.0708	0.0908	0.1166
0.5	0.0399	0.0544	0.0644	0.0821	0.1050
Average MME		0.0829			
Size = 11					
β = 0.1	0.0515	0.0745	0.0980	0.1188	0.1360
0.2	0.0476	0.0653	0.0893	0.1116	0.1308
0.3	0.0448	0.0562	0.0776	0.1019	0.1246
0.4	0.0437	0.0532	0.0694	0.0922	0.1154
0.5	0.0408	0.0498	0.0622	0.0812	0.1041
Average MME		0.0816			
Size = 13					
β = 0.1	0.0494	0.0707	0.0949	0.1158	0.1350
0.2	0.0462	0.0612	0.0854	0.1088	0.1317
0.3	0.0429	0.0546	0.0766	0.0995	0.1235
0.4	0.0415	0.0512	0.0676	0.0896	0.1140
0.5	0.0394	0.0477	0.0609	0.0804	0.1032
Average MME		0.0797			
Size = 13					
β = 0.1	0.0479	0.0672	0.0918	0.1137	0.1348
0.2	0.0449	0.0608	0.0822	0.1064	0.1307
0.3	0.0414	0.0544	0.0731	0.0957	0.1203
0.4	0.0391	0.0498	0.0647	0.0864	0.1117
0.5	0.0375	0.0463	0.0581	0.0780	0.0989
Average MME		0.0774			

3×3, while those of LE and LFE are 7×7 and 11×11, respectively. Therefore, their time complexities in calculating local characteristic value are $O(9N)$, $O(49N)$ and $O(121N)$, respectively. The differences are $O(40N)$ and $O(112N)$, which are larger than $O(L)$. Thus MLE can run faster than LE and LFE, and its running time is almost half of LE's and one third of LFE's.

6. Conclusions

In this paper, a modified local entropy-based transition region extraction and thresholding method has been presented. By analyzing the properties of transition region, we find that local entropy could not completely depict transition region. This is because local entropy only considers frequency of gray level changes, but neglects the degree of these changes. This causes those non-transition regions with frequent but slight gray level changes to be misclassified into transition regions. In order to eliminate this limitation, a new transition region descriptor integrating local complexity and local variance is proposed to depict frequency and degree of gray level changes adequately. In addition, in the light of human visual perception, a preprocessing step named image transformation is introduced to simplify original images. This transformation retains image details within a gray level range determined in an unsupervised way while excluding the contribution of transitional pixels outside this range. In other words, this transformation preserves gray level changes in the mixture between object and background, and weakens gray level changes of non-transition region, simplifying original images. Thus, it is helpful for transition region extraction and subsequent image segmentation. The proposed method can extract transition region more accurately, which naturally leads to a good segmentation result. Furthermore, the proposed method runs faster than classic local entropy-based method and local fuzzy entropy-based approach for needing smaller neighborhood window in local characteristic calculation. Experimental results on a variety of images demonstrate the effectiveness and efficiency of the new method.

Acknowledgements

This work is supported by National Natural Science Foundation of China (Grant Nos. 90820004, 60875010), Technology Project of Education Department of Fujian Province (JA10226), Technology Project of Fujian Province (JK2010046), Nanjing Institute of Technology Internal Fund (KXJ06037) and Guangxi Natural Science Fund of China (No. 2011GXNSFB018070)

References

[1] Y. Zhang, F.L. Chung, S. Wang, Robust fuzzy clustering-based image segmentation, Appl. Soft Comput. 9 (1) (2009) 80–84.
[2] S. Chen, M. Wang, Seeking multi-thresholds directly from support vectors for image segmentation, Neurocomputing 67 (2005) 335–344.
[3] Y. Qiao, Q.M. Hu, G.Y. Qian, S.H. Luo, W.L. Nowinski, Thresholding based on variance and intensity contrast, Pattern Recognit. 40 (2) (2007) 596–608.
[4] M. Sezgin, B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, J. Electron. Imaging 13 (1) (2004) 146–165.
[5] Y. Hata, S. Kobashi, Fuzzy segmentation of endorrhachis in magnetic resonance images and its fuzzy maximum intensity projection, Appl. Soft Comput. 9 (3) (2009) 1156–1169.
[6] T. Sund, K. Eilertsen, An algorithm for fast adaptive binarization with applications in radiotherapy imaging, IEEE Trans. Med. Imaging 22 (1) (2003) 22–28.
[7] Y. Solihin, C.G. Leedham, Integral ratio: a new class of global thresholding techniques for handwriting images, IEEE Trans. Pattern Anal. Mach. Intell. 21 (8) (1999) 761–768.
[8] B. Bhanu, Automatic target recognition: state of the art survey, IEEE Trans. Aerosp. Electron. Syst. 22 (4) (1986) 364–379.
[9] M. Sezgin, R. Tasaltin, A new dichotomization technique to multilevel thresholding devoted to inspection applications, Pattern Recognit. Lett. 21 (2) (2000) 151–161.
[10] J.J. Gerbrands, Segmentation of noisy images. Ph.D. Dissertation, Delft University, The Netherlands, 1988.
[11] Y.J. Zhang, J.J. Gerbrands, Transition region determination based thresholding, Pattern Recognit. Lett. 12 (1) (1991) 13–23.
[12] C.X. Yan, N. Sang, T.X. Zhang, Local entropy-based transition region extraction and thresholding, Pattern Recognit. Lett. 24 (16) (2003) 2935–2941.
[13] A.M. Groenewald, E. Barnard, E.C. Botha, Related approaches to gradient-based thresholding, Pattern Recognit. Lett. 14 (7) (1993) 567–572.
[14] C. Zhang, J. Zhang, H. Chen, Local fuzzy entropy-based transition region extraction and thresholding, Int. J. Inf. Technol. 12 (6) (2006) 19–25.
[15] M.P. Albuquerque, I.A. Esquef, A.R.G. Mello, Image thresholding using Tsallis entropy, Pattern Recognit. Lett. 25 (9) (2004) 1059–1065.
[16] Z. Hou, Q. Hu, W.L. Nowinski, On minimum variance thresholding, Pattern Recognit. Lett. 27 (2006) 1732–1743.
[17] J.N. Kapur, P.K. Sahoo, A.K.C. Wong, A new method for grey-level picture thresholding using the entropy of the histogram, Comput. Vis. Graph. Image Process. 29 (1985) 273–285.
[18] S. Wang, F. Chung, F. Xiong, A novel image thresholding method based on parzen window estimate, Pattern Recognit. 41 (1) (2008) 117–129.
[19] S. Arora, J. Acharya, A. Verma, P.K. Panigrahi, Multilevel thresholding for image segmentation through a fast statistical recursive algorithm, Pattern Recognit. Lett. 29 (2) (2008) 119–125.
[20] Q. Hu, S. Luo, Y. Qiao, G. Qian, Supervised grayscale thresholding based on transition regions, Image Vision Comput. 26 (12) (2008) 1677–1684.
[21] C.E. Shannon, A mathematical theory of communication, Bell Syst. Technol. J. 27 (1948) 379–423.
[22] T. Pun, A new method for grey-level picture thresholding using the entropy of histogram, Signal Process. 2 (1980) 223–227.
[23] C.X. Yan, N. Sang, T.X. Zhang, K. Zeng, Image transition region extraction and segmentation based on local complexity, J. Infrared Millim. Waves 24 (4) (2005) 312–316.
[24] Z. Li, J. Yang, G. Liu, Y. Cheng, C. Liu, Unsupervised range-constrained thresholding, Pattern Recognit. Lett. 32 (2) (2011) 392–402.
[25] W.A. Yasnoff, J.K. Mui, J.W. Bacus, Error measures for scene segmentation, Pattern Recognit. 9 (4) (1977) 217–231.