An Analytical Review on Various Image Segmentation Methods Based on Transition Region-Based Thresholding

Sameer Kumar Behera¹, Priyadarsan Parida², Nilamani Bhoi³, Akash Kumar Athghara⁴

Graduate Student, Dept. of Electronics & Telecommunication, VSSUT, Burla, Sambalpur, Odisha, India¹
Research Scholar, Dept. of Electronics & Telecommunication, VSSUT, Burla, Sambalpur, Odisha, India²
Assistant Professor, Dept. of Electronics & Telecommunication, VSSUT, Burla, Sambalpur, Odisha, India³
Post Graduate Student, Dept. of Computer Engineering, San Jose State University, California, USA⁴

ABSTRACT: Object segmentation is still a very challenging problem in computer vision. Various gradient-based methods are used for their simplification which basically employ edge-structures that separate object region from background. But edge-detection process which uses conventional edge operators has a demerit as it is more sensitive to noise. Hence, transition region-based approach, being a recent boundary detection method is used as it is insusceptible to noise. In this paper, transition region-based thresholding methods such as local entropy (LE) [10], local gray level difference (LGLD) [11] & modified local entropy (MLE) [12] are studied and their various performance measures are analysed. Local entropy-based transition region extraction [10] catches the picture information very efficiently from an image by calculating the local entropy of the gray level of each and every pixel. Local gray level difference-based transition region extraction [11] considers the gray level changes in the transition region, thereby removing redundancy. Modified local entropy-based transition region extraction [12] employs local complexity and local variance for creation of a new and improved transition region descriptor for efficient segmentation. Simulation results of the various methods are demonstrated.

KEYWORDS: Thresholding, Transition Region, Local Entropy, Gray Level Difference, Modified Local Entropy.

INTRODUCTION

Image segmentation in the present era of digital computers plays a very significant role in various aspects of image processing. Image segmentation extracts an object from its background on the basis of various attributes such as gray level, color, texture and location [1-4]. Earlier object segmentation techniques basically used edge detection of the object using various operators such as Prewitt, Canny [5-6] etc. But these operators were sensitive to noise. Hence transition region-based approach, a recent technique for boundary detection was introduced by Gerbrands [7] as it was insensitive to noise. Some of the newly developed methods for image segmentation include gradient-based transition region extraction (GTREM) and transition region-based thresholding. Transition region was incorporated into image segmentation by Zhang and Gerbrands [8] which employed effective average gradient (EAG) through clip transformation function. But EAG couldn't describe the features of the transition region properly as it reflected only sudden gray level changes but not the frequent gray level changes [9]. Also it was highly unsuitable to handle complex images with multiple objects and was susceptible to noise [9]. On the contrary, transition region-based thresholding approach catches the natural properties of the transition region in a better manner. It has been a novel and effective process for image segmentation in recent years [8-12]. Thresholding can serve a variety of purposes such as biomedical image analysis, change detection, handwritten character identification and automatic target recognition [13-16].

The transition region-based thresholding methods include local entropy (LE) [10], local gray level difference (LGLD) [11] & modified local entropy (MLE) [12] for image segmentation. LE captures the nature of the transition region much better than the EAG method [10]. This robust and effective method represents frequent gray level changes in the transition region. This method significantly outperforms the conventional GTREM methods but it also faces a demerit
as it only considers the frequency of gray level changes while completely ignoring the degree of gray level changes. Also it has high computational complexity due to large amount of multiplications and logarithmic operations. LGLD overcomes the abovementioned limitations by deserting the gradient and analysing the characteristics of transition region [11]. In MLE, a new modified transition region descriptor is developed which describes both frequency and degree of gray level changes along an effective pre-processed image for better visual perception [12].

II. PROPERTIES OF TRANSITION REGION

The transition region present between foreground and background has the following characteristics:

(1) REGION CHARACTERISTICS: There always exists a transition region around an edge [7], whether for a step edge or a non-step edge. In step edge the intensity values changes abruptly, whereas in non-step edges the intensity change occurs over a finite distance. Near a non-step edge, the transition region has a width of certain pixels but around a step edge, the transition region must have at least one-pixel width.

(2) BOUNDARY CHARACTERISTICS: The transition region is covered around the foreground and is located between the object and background, clearly defining a boundary between them.

(3) VARIANCE OF GRAY LEVELS: The gray levels in transition region undergo frequent changes bringing about abundant information for transition region description. Gradient has been a good measure for sudden gray level changes but cannot measure frequent gray level changes. As entropy can best represent the information content of an image, local entropy is better suited to depict the frequent gray level changes. However, it is unable to reflect the extent of gray level changes.

III. TRANSITION REGION BASED METHODS

III.1 LOCAL ENTROPY (LE) [10]:

Consider an image I with L gray levels \([0, 1, \ldots, L-1]\). The number of pixels at gray level \(i\) is denoted by \(n_i\) and the total number of pixels by \(N = n_0 + n_1 + \ldots + n_{L-1}\). Now let \(W = \{(i,j): i = 1,2,\ldots,n_0; j = 1,2,\ldots,n_a\}\), where \(n_0\) and \(n_a\) are the height and width of the image, respectively. Let the gray level at pixel \((i,j)\) be given by \(f(i,j)\). Following Shannon’s definition [17] of entropy, Pun [18] defined the entropy of an image as,

\[
E = -\sum_{i=0}^{L-1} P_i \log P_i (1)
\]

where \(P_i = \frac{n_i}{N}\) is the probability of gray level \(i\) appeared in the image.

If we take a \(m \times n\) neighbourhood window centred on the pixel \((i,j)\), \(\Omega\), its local entropy, can be calculated as follows:

\[
LE(i,j) = E(\Omega) = -\sum_{k=0}^{L-1} P_k \log P_k (3) \text{ where } P_k = \frac{n_k}{m \times n}
\]

is the probability of gray level \(k\) appeared in the neighbourhood \(\Omega\) and \(n_k\) is number of pixels with gray level \(k\) in its local entropy, \(\Omega\). Each pixel's local entropy value can be obtained by moving the neighbour window \(\Omega\) pixel by pixel within the image from left to right and top to bottom. The transition region can then be extracted by appropriate entropy threshold. The final segmentation threshold can then be determined by the peak or mean of the histogram of the transition region [8].

The algorithm is summarized in the following steps:

Step 1: Take a certain neighbour window size and appropriate entropy threshold.

Step 2: Calculate the local entropy.

Step 3: Extract the transition region.

Step 4: Obtain the segmentation threshold through the transition region's histogram.

Step 5: Segment image by the final threshold.

The threshold entropy can be computed as follows:

\[
E_T = \alpha E(\Omega)_{\text{max}} (5) \text{ where } E(\Omega)_{\text{max}} \text{ is the maximum entropy of the entropy image and } \alpha \text{ is a coefficient varying between 0 and 1.}
\]

III.2 LOCAL GRAY LEVEL DIFFERENCE (LGLD) [11]:

The mean of gray levels of the pixels in a neighborhood can be easily formulated as,
\( \overline{f}(i,j) = \frac{1}{m \times n} \sum_{x=1}^{m} \sum_{y=1}^{n} f(x,y) \) \hspace{1cm} (6) \) where \((x,y)\) is a pixel co-ordinate of the neighbourhood.

Now the absolute difference between \(f(x,y)\) and \(\overline{f}(i,j)\) is called the gray level difference of the pixel, denoted by \(\Delta f(i,j) = | f(i,j) - \overline{f}(i,j) | \) \hspace{1cm} (7) \). When the neighbourhood window is moved pixel by pixel within the image from left to right and top to bottom, we obtain an image matrix \(D\) composed of each pixel's gray level difference.

\[
D = \begin{bmatrix}
\Delta f(1,1) & \Delta f(1,2) & \ldots & \Delta f(1,n_w) \\
\Delta f(2,1) & \Delta f(2,2) & \ldots & \Delta f(2,n_w) \\
\vdots & \vdots & \ddots & \vdots \\
\Delta f(n_h,1) & \Delta f(n_h,2) & \ldots & \Delta f(n_h,n_w)
\end{bmatrix} \hspace{1cm} (8)
\]

The appropriate threshold \(T\), for extraction of transition region can be computed via the following steps:

1) Calculate the mean and standard deviation of the gray level difference image, \(\mu\) and \(\sigma\), by the following equations,

\[
\mu = \frac{1}{n_h \times n_w} \sum_{i=1}^{n_h} \sum_{j=1}^{n_w} \Delta f(i,j) \hspace{1cm} (9) \]

\[
\sigma = \left( \frac{1}{n_h \times n_w} \sum_{i=1}^{n_h} \sum_{j=1}^{n_w} (\Delta f(i,j) - \mu)^2 \right)^{\frac{1}{2}} \hspace{1cm} (10)
\]

2) Set the following abovementioned threshold, \(T = \mu + \alpha \sigma\) \hspace{1cm} (11) \) where \(\alpha\) is a parameter.

Gray level difference is related to the changes in the gray level in a neighbourhood. Thus, the pixels in transition region of an image will have bigger gray level differences than those in non-transition region of the image. The final segmentation threshold is finally determined by mean value of the histogram of the transition region and used to binarize the image.

**Algorithm for LGLD based image segmentation is given as follows:**

**Step 1:** Consider a certain window size and value of parameter, \(\alpha\).

**Step 2:** Calculate the gray level difference of each pixel.

**Step 3:** Create gray level difference image matrix, \(D\).

**Step 4:** Get the threshold \(T\) for transition region extraction.

**Step 5:** Extract the transition region.

**Step 6:** Compute the mean of gray levels of pixels in the transition region and use it as the final segmentation threshold.

**Step 7:** Segment the image through the final threshold.

### III.3 Modified Local Entropy (MLE) [12]:

#### III.3.1 Image Transformation

The human eye is insensitive to features present at both extremes of pixel intensity, but sensitive to the features in mid-range [19]. So, a pre-processing step called image transformation is suggested to simplify original images. Let the gray levels of transition region be within a range \((\theta_1, \theta_2)\) which can be determined in an unsupervised way, after which a transformation is later applied so that only pixels with gray levels inside the range will contribute to transition region.

The range can be found via image statistical characteristics.

For an image \(I\), \(\theta_1\) and \(\theta_2\) can be determined with the following ways:

1) Compute mean and standard deviation of the image by the following equations:

\[
\mu = \frac{1}{n-1} \sum_{i=0}^{n-1} i n_i \hspace{1cm} (12)
\]

\[
\sigma = \left( \frac{1}{n-1} \sum_{i=0}^{n-1} (i - \mu)^2 n_i \right)^{\frac{1}{2}} \hspace{1cm} (13)
\]

2) Now determine \(\theta_1\) and \(\theta_2\) by the following equations:

\[
\theta_1 = \mu - \alpha \times \sigma \hspace{1cm} (14)
\]

\[
\theta_2 = \mu + \alpha \times \sigma \hspace{1cm} (15)
\]

where \(\alpha\) is a parameter.

After \(\theta_1\) and \(\theta_2\) are determined, image transformation can be followed immediately,

\[
f_{t\varepsilon}(i,j) = \begin{cases} 
\theta_1 & \text{if } f(i,j) < \theta_1 \\
 f(i,j) & \text{if } \theta_1 \leq f(i,j) \leq \theta_2 \\
\theta_2 & \text{if } f(i,j) > \theta_2
\end{cases}
\]

The above transformation weakens gray level changes in both object and background simultaneously, thus simplifying the original image. The weakening effect is highly favorable for transition region extraction and image segmentation.

#### III.3.2 Modified Transition Region Descriptor

In order to reduce the computational complexity of local entropy, a new function called local complexity [20], is used to describe frequency of gray level changes, i.e.
Detailed algorithm of MLE based transition region extraction and thresholding is as follows:

**Step 1:** Find the gray level range (\(b_0, b_2\)) for image transformation.

**Step 2:** Now calculate each pixel’s S value in the transformed image to construct an image matrix S.

**Step 3:** Now obtain the following threshold \(S_T\) for transition region extraction,
\[
S_T = \gamma \times S_{\text{max}}(25) \quad \text{where} \quad S_{\text{max}} = \max_{0 \le i < h, 0 \le j < w} S(i,j)(26) \quad \text{and} \quad \gamma \text{ is a coefficient between 0 and 1.}
\]

**Step 4:** Now extract transition region as follows,
\[
TR(i, j) = \begin{cases} 
1 & \text{if } S(i, j) \ge S_T \\
0 & \text{otherwise}
\end{cases}
\]

**Step 5:** Final segmentation threshold \(T^*\) is computed as gray level mean of transition region \([9, 10]\),
\[
T^* = \frac{\sum_{(i,j) \in TR} f(i, j)}{|TR|}
\]

**Step 6:** Now, binarize the image by \(T^*\)

### IV. PERFORMANCE MEASURES

Performance measures give regard to image segmentation as a pixel classification process and help in qualitative evaluation of the quality of image segmentation. Some of them are:

#### IV.1 MISCLASSIFICATION ERROR (ME) [21]:
ME [21] reflects the percentage of background pixels incorrectly classified into the foreground and conversely, object pixels erroneously assigned to the background. For a two-class segmentation, ME [21] can be simply calculated as,
\[
ME = 1 - \frac{B_T |F_0| + F_T |B_0|}{|B_0| + |F_0|}(29)
\]

where \(B_0\) and \(F_0\) are the background and foreground of the ground truth image, \(B_T\) and \(F_T\) are the background and foreground in an image segmentation result and \(|\cdot|\) is the cardinality of a set.


#### IV.2 FALSE POSITIVE & NEGATIVE RATE (FPR & FNR) [22]:
FPR [22] is the rate of the number of background pixels misclassified into foreground pixels to the total number of background pixels in the ground truth image. Also, FNR [22] is the rate of the number of foreground pixels
misclassified into background pixels to the total number of foreground pixels in the ground truth image. For a two-class segmentation, FPR and FNR are given as,

\[ \text{FPR} = \frac{|B_G \cap F|}{|B_G|} \quad \text{(30)} \text{and} \quad \text{FNR} = \frac{|F \cap B|}{|F|} \quad \text{(31)} \]

**IV.3 GRAY LEVEL CONTRAST (GC) [23]:**

An adequate segmentation should also produce images having higher contrast across adjacent regions [23]. In a simple case that a gray-level image \( f(x,y) \) consists of the object with average gray-level \( f_o \) and the background with average gray-level \( f_b \), a gray-level contrast measure \( GC \) [23] is computed by:

\[ GC = \frac{|f_o - f_b|}{f_o + f_b} \quad \text{(32)} \]

**V. RESULTS & DISCUSSION**

For LE [10], a window size of 7×7 or 15×15 is suggested to be appropriate [24]. Sufficient pixels can be extracted for transition region by taking value of \( \alpha \) between 0.6 and 1. Better or more precise ranges for \( \alpha \) can be empirically chosen between 0.8 and 0.9. For LGLD [11], the neighborhood window size is chosen as 3 × 3 and the parameter, \( \alpha \) is set to 1.5 respectively. MLE [12] uses a 3×3 window size and the parameters, \( \beta \) and \( \gamma \) are set as 0.3 and 0.1 respectively. The bigger \( \beta \) is, larger is the weight of local complexity & bigger \( \gamma \) is, lesser is the number of pixels in transition region [12]. The ME [21] gives the discrepancy measure, i.e. the percentage of misclassified pixels which should be as small as possible [25]. The large values of FPR [22] and FNR [22] indicate over-segmentation and under-segmentation of the segmented image respectively [25].

![Image](image.jpg)

**Fig.1.** Image segmentation results on: Airplane & Tree:
(a) Original grayscale, (b) Ground truth, (c) LE’s transition region, (d) LE’s result, (e) LGLD’s transition region, (f) LGLD’s result, (g) MLE’s transition region, (h) MLE’s result.
In Fig. 1, the images of the original grayscale, the ground truth, the transition region & the final segmentation results of LE [10], LGLD [11] and MLE [12] of two scenarios, i.e. (i) Airplane and (ii) Tree are shown in order to achieve a comparative study between the different segmentation results of LE [10], LGLD [11] and MLE [12] with the ground truth images of the two scenarios.

<table>
<thead>
<tr>
<th>Images</th>
<th>Transition Region-based Thresholding Methods</th>
<th>PERFORMANCE MEASURES</th>
<th>Algorithm Runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Airplane</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LE</td>
<td>0.0684</td>
<td>0.0048</td>
</tr>
<tr>
<td></td>
<td>LGLD</td>
<td>0.0569</td>
<td>0.0482</td>
</tr>
<tr>
<td></td>
<td>MLE</td>
<td>0.0142</td>
<td>0.0023</td>
</tr>
<tr>
<td>b) Tree</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LE</td>
<td>0.0714</td>
<td>0.0619</td>
</tr>
<tr>
<td></td>
<td>LGLD</td>
<td>0.4769</td>
<td>0.1113</td>
</tr>
<tr>
<td></td>
<td>MLE</td>
<td>0.1529</td>
<td>0.3388</td>
</tr>
</tbody>
</table>

Table 1. Quantitative comparison of the performance measures of various transition region-based thresholding methods for the segmentation of the following images: a) Airplane, b) Tree.

In Table 1, a quantitative comparison of the different segmentation results of LE [10], LGLD [11] and MLE [12] with the ground truth images of the two scenarios, i.e. (i) Airplane and (ii) Tree in accordance with various performance measures i.e. ME [21], FPR & FNR [22] and GC [23] along with the Algorithm Runtime (in seconds) is shown.

In LE [10], a too large window loses localization while a too small window results in imprecise estimation of local entropy because of lack of sampling. Due to bigger window size and large multiplications and logarithmic operations for the calculation of entropy, the computational time of LE [10] is very high. LGLD [11], on the other hand operates very fast having the least computational time as the difference matrix, D gets created by simple subtractions. MLE [12], extracts transition region more accurately and achieves better results with less misclassified pixels. Its computational time is nearly half of that of LE due to simple additions and multiplications, but is more than that of LGLD [11].

![More segmentation results on some complex images](image1)

Fig. 2. More segmentation results on some complex images: Chopper, Island, Caterpillar, Flower & Leaf; (a) Original grayscale, (b) LE’s result, (c) LGLD’s result, (d) MLE’s result.
In Fig.2, the images of the original grayscale & the final segmentation results of LE [10], LGLD [11] and MLE [12] of various complex scenarios, i.e. (i) Chopper, (ii) Island, (iii) Caterpillar, (iv) Flower and (v) Leaf are shown only for visual analysis and not at all for quantitative analysis based on performance measures.

The images used for segmentation using transition region based thresholding are contrast-dependent. These methods may not be applicable for all types of images. Only those images which have considerable amount of contrast difference between foreground and background are liable for good segmentation.

All the experiments are performed on a notebook PC with 2.53 GHz Intel(R) Core(TM) i5 CPU and 4 GB RAM. All the experiments used 8-bit images of 225x300 pixels from the Weizmann image dataset [26].

**VI. CONCLUSION**

In this paper we surveyed various transition region based thresholding methods for image segmentation by transition region extraction of object from the background. All of these methods used multiple thresholds for image segmentation.

One threshold is used for extraction of transition region and is calculated by considering local parameters from the transition region and the other threshold is used for final segmentation which is calculated from the gray level mean of the histogram of the transition region.

The LE-based method [10] takes into consideration only the frequency of gray level changes while completely ignoring the degree or extent of these changes. The LGLD-based method [11] considers both frequency and degree of gray level changes which helps in accurate extraction of transition region and produces a good segmentation result. Furthermore, it takes very little time for image segmentation. The MLE-based method [12] performs image transformation for simplification of the original images by weakening the gray level changes of non-transition region while also taking into account the frequency and degree of gray level changes. This method also extracts the transition region more accurately which leads to better image segmentation.

These transition region-based thresholding methods discussed so far are very well-suited for single object segmentation, but can also be further extended to be used for multi-object segmentation.

**REFERENCES**


