CHAPTER II

LITERATURE SURVEY

2.1 Introduction

An image is considered as a set of points or pixels distributed over a two-dimensional finite space. Segmentation, in precise form, can be considered as the process of assigning the label to each pixel in an image, such that all the pixels having same label share some common visual characteristics. Thresholding is the simplest method for picture segmentation. Segmentation is performed by assigning pixels having gray levels below the threshold to the background, and assigning those pixels having gray levels above the threshold to the objects or vice versa.

Image segmentation plays a crucial role in many computer vision applications. Computer pattern recognition, as its name suggests, is an artificial intelligence system that recognizes certain patterns such as characters, human faces, lines and their structures. But these simple tasks which are easy for our human recognition system are not so easy for a computer although the later has great capacity of computation. The features of these patterns have to be extracted from the distorted, blurry or faint images. It is an important task in pattern recognition because it is the basis of the recognition process and a good thresholded image is the essential step for efficient and accurate recognition. Though thresholding is a simple and old problem a robust and comprehensive solution is yet to be investigated. As part of investigating and developing a robust method, all the available
standard thresholding methods are considered for a review. And this chapter is devoted for the literature survey, since arrangement of literature is of first importance.

2.2 Global versus Local Threshold-based Methods

Many thresholding methods have been developed [1-10]. They can be classified into two groups: local and global [1]. A local thresholding method selects different threshold values for different regions, or even for each pixel. Tsai [19] suggested selecting a threshold at which the resulting binary images have the same first three moments. Local techniques are mostly time consuming because the threshold has to be calculated for each pixel according to the gray scale information around the pixel. Many local thresholding techniques have been developed [30-33] such as Nilblack’s method [31]. Nilblack’s method is considered to be very effective and simple method for local thresholding. The local threshold will be set at:

\[ T(i,j) = \mu(i,j) + w\sigma(i,j) \]  

(2.1)

Where \( \mu \) is local mean, \( \sigma \) is standard deviation, \( w \) is a weight and \( T \) is local threshold. Values of local mean and standard deviation are calculated on a local m×n window. The parameters to be considered are weight \( w \) and window size.

Yanowitz and Bruckstein method [32], Shen and Ip’s method [35] and Liu’s method [34]. A logical level technique [36] is proposed for document images which fully factors in the stroke width of an image and uses local gray scale averages in the calculation. The stroke width is usually uniform in most document images and the local averages have low sensitivity to noise. The technique achieves good results on document images with uneven background noise. In this method the gray scale of the processed
pixel is compared to the averages of four symmetrical surrounding windows. A threshold is chosen for the whole image. If the gray scale of the processed pixel is lower than the surrounding averages by the value of the threshold, it is set as the objects, otherwise, as background. Yang and Yan [30] extended this method to adaptively choose thresholds depending on the local average and other gray level information. This method achieves a better result on document images with poor illumination.

This thesis focuses on obtaining an optimal threshold using Global thresholding approach. In this approach one threshold value is selected for the entire image due to its simplicity, and less computational demand and used as a preprocessing step in various image processing applications. Points which have a gray level above the threshold are classified as background; alternatively, they are classified as objects.

Global thresholding methods select the threshold based on different criteria, such as Ridler and Calvard [14] uses an iterative technique for choosing a threshold value. An iteration $n$, a new threshold $T_n$ is established using the average of the foreground and background class means. The process is repeated until the change between two successive iterations is sufficiently small. The basic equation is:

$$T = \frac{\mu_b + \mu_f}{2}$$

(2.2)

This can be considered as the foundation of Ridler Calvard’s method. In practical life $\mu_b$ (mean of background pixels) and $\mu_f$ (mean of foreground pixels) are unknowns. We must estimate these based on suggested thresholds. This is what Ridler Calvard’s method tries to do. Assume that the histogram of the image is $p(z)$ where $z$ is the gray level. The simple idea:
• Start by choosing an initial threshold $T_0$ equal to the average gray level of the image.

• Then iterate and calculate new thresholds according to the following formula:

$$T_{k+1} = \frac{\mu_1(T_k) + \mu_2(T_k)}{2}$$  \hfill (2.3)

Here $\mu_1$ is the mean value of the gray levels below $T_k$ and $\mu_2$ the mean value of the gray levels above the threshold.

In references [15-17] and [23], Histogram modification and approximation techniques used for threshold selection and used gray level, local average gray level, amongst these methods image histogram is used, globally or locally, as a basis to perform the task of thresholding.

Otsu’s method [16] suggests maximizing the weighted sum of between-class variances of foreground and background pixels to establish an optimal threshold according to the discrimination analysis and is further enhanced by W. Hongzhi and Dong Ying [29]. One of the common image processing tasks is to transfer a grayscale image to its monochrome version. Otsu's method, named after its inventor Nobuyuki Otsu [16], is one of the most popular automatic binarization algorithms, which is used in many computer vision and medical imaging applications. Otsu’s method is based on the shape of the histogram and is considered as a non-parameterized algorithm since it obtains the optimal threshold by maximizing the between-class variance with a comprehensive search. It also can be extended to multi level thresholding that will result in segmentation. Otsu’s thresholding method iterates through all the possible threshold values. It is an automated algorithm.
In each iteration, the weighted within-class variance of two classes (foreground and background on either sides of the threshold point) gets calculated. The optimal threshold is the one that minimizes the within-class variance, hence, maximizing the between class variance.

Consider an image with N pixels with gray levels in \([1, \ldots, L]\) interval. To divide this image into two classes, \(C_1\) with gray levels \([1, \ldots, t]\) and with \([t+1, \ldots, L]\), we need to calculate the gray level distribution for both classes as:

\[
C_1 : \frac{p_1}{w_1(t)}, \ldots, \frac{p_t}{w_1(t)} \quad (2.4)
\]

\[
C_2 : \frac{p_{t+1}}{w_1(t)}, \ldots, \frac{p_L}{w_2(t)} \quad (2.5)
\]

Where

\[
w_1(t) = \sum_{i=1}^{t} p_i
\]

\[
w_2(t) = \sum_{i=t+1}^{L} p_i
\]

Here, \(p_i\) represents the probability of gray level \(I\) with \(f_i\) pixels and is calculated as:

\[
P_i = \frac{f_i}{N} \quad (2.6)
\]

Now

\[
\mu_1 = \frac{\sum_{i=1}^{t} i p_i}{w_1(t)} \quad (2.7)
\]

\[
\mu_2 = \frac{\sum_{i=t+1}^{L} i p_i}{w_2(t)} \quad (2.8)
\]

Let \(\mu_T\) be the mean intensity of the whole image. It is easy to show that:

\[
\mu_T = w_1 \mu_1 + w_2 \mu_2 \quad (2.9)
\]

And \(w_1 + w_2 = 1 \quad (2.10)\)

Otsu defines the intra-class variance of the thresholded image as:

\[
\sigma_B^2 = w_1 (\mu_1 - \mu_T)^2 + w_2 (\mu_2 - \mu_T)^2 \quad (2.11)
\]

Otsu verified that for bi-level thresholding, the optimal threshold, \(t^*\), is chosen so that \(\sigma_B^2\) is maximized; that is,
\[ t^* = \text{Arg Max} \{ \sigma^2_{B}(t) \} \mid t \leq L \]  

(2.12)

It is easy to implement this method since it operates on histograms, which are arrays of length 256. This method is considered quiet fast, the thresholded images are shown in Figure 2.1. However, assumption of uniform illumination and bimodal histogram are the two drawbacks counted for this technique.

- Otsu’s method is one of standard methods for image segmentation.
- It works very well on bimodal or close to bimodal or/and multimodal images.
- However, it doesn’t work well on unimodal images and images with small objects on large background.

![Figure 2.1 Thresholded images of Otsu method.](image)

Rosenfeld’s convex hull method is based on analyzing the concavity structure of the histogram defined by its convex hull [18]. When the convex hull of the histogram is calculated, the deepest concavity points become candidates for the threshold value. J.Kittler and J.Illingworth’s [20] proposed minimum error thresholding method, that the image can be characterized by a mixture distribution of object and background pixels. J. Z. Liu and W. Q. Li [21] proposed an automatic thresholding of gray-level picture via
two-dimensional Otsu method. More recent thresholding algorithms have been developed using Fuzzy clustering ideas for thresholding. The works presented in [22] used fuzzy memberships based on pixels distance from each class’s mean to define which class a pixel belongs to and subsequently define the threshold as the cross over point of membership functions. The logical thresholding technique proposed by Yang and Yan [24] is adopted and further developed for the binarization of blueprint images because it outperforms other local techniques on suppressing the background noise in lowlight images. Julong Deng Gray system theory [25] has been successfully applied in the fields of system analysis, data processing, modeling. Gray relational analysis, a method of gray system theory is widely used in image processing [26]. Jawahar et al. [50] proposed a fuzzy thresholding scheme based on Fuzzy C-means clustering. The problem of fuzzy clustering is that of partitioning the set of n sample points into c classes. The algorithm is an iterative optimization that minimizes one cost function. Two extensions of this algorithm are found in [54] and [27]. Recently H. Bustince, E. Barrenechea, and M. Pagola proposed [28] Image thresholding using restricted equivalence functions and maximizing the measures of similarity.

2.3 Fuzzy sets and Fuzzy measures

In this thesis fuzzy sets are computed with gray-level spatial correlation to estimate the initial fuzzy seed subsets for the fuzzy s-function as inception. Fuzzy set theory, fuzzy measures are used for the purpose finding optimal threshold. It is evident from the literature, that fuzzy set theory has been widely used in many fields where the ambiguous phenomena exist from the point it was proposed by L. A. Zadeh [38] in 1965. Many of the thresholding methods have been developed by using this theory. The concept
being used here is called fuzzy measures to find the limits of fuzzy region, also called initial seed subsets. Fuzzy measure is a reasonable approach to estimate the average ambiguity in fuzzy sets and hence measuring its fuzziness [48]. Kaufmann in [39] & [40] introduced an index of fuzziness (IF) comparing a fuzzy set with its nearest crisp set, and proposes a fuzziness index as a normal distance. Another author Yager [41] understands the entropy measures as the difference between a fuzzy set and its complementary fuzzy set. From [42] the index of fuzziness and entropy of an image reflect a kind of quantitative measure of its enhancement quality. Their values are found to decrease with enhancement of an image when different sets of S-type membership functions with appropriate crossover points were considered for extracting the fuzzy property plane from the spatial domain of the image. The threshold is obtained by comparing the association errors and assigning each gray level bin of the undefined region to one of the defined regions that corresponds to the lower association error [43]-[55]. H. J. Zimmermann [44] mentioned good number of applications of fuzzy set theory. A. De Luca, S. Termini [46] introduced definition of non-probabilistic entropy in the setting of fuzzy set theory. Kuriyama used fuzzy c-partition entropy as the criterion to measure the fitness of a fuzzy partition in [45]. J. C. Bezdek, S. K. Pal[47] used fuzzy set theory for pattern recognition. In reference [48] a reasonable approach is employed to estimate the average ambiguity in fuzzy sets is measuring its fuzziness. Huang and Wang [49] assign the memberships to the pixel with the help of the relationship between its gray value and mean gray value of the region to which it belongs. A more robust method is suggested by O. J. Tobias and R. Seara [51] which deals with object edges and ambiguity and avoids the problems involved in finding the minimum of a function. Q. Wang et al. [52]
proposed a method using index of non-fuzziness of the 2-D gray scale histograms. H. R. Tizhoosh [53], introduced a new Image thresholding using type II fuzzy sets with the help of his new measure of ultrafuzziness.

A generalized fuzzy entropy definition is given by Jiulun Fan, Feng Zhao [56] and they have applied it to image segmentation, this is further enhanced using the principle that the membership degree of the threshold point is to \( m \) where \( 0 < m < 1 \) and to determine the parameter \( m \) Lei Bo, Fan Jiu-Lun [57] used particle swarm optimization. In the research performed by Tizhoosh et al. [58], the authors introduced opposition-based fuzzy thresholding, called OFT henceforward, and combine the concepts of fuzzy memberships and opposition-based computing to extract some local information of the image that leads to selecting a threshold value.

2.4 Fuzzy Measures and Initial fuzzy seed subset

Nuno vieira Lopes [59] et al. presented a method to segment the image without finding the minimum function. Similarity between gray levels is the key to find an optimal threshold. Two initial regions of gray levels, located at the boundaries of the histogram are defined. Then using an index of fuzziness, a similarity process is started to find the threshold point. A significant contrast is assumed between object and background. In this approach, the initial subsets are defined and they are large enough for providing a minimum number of pixels at the beginning of the process. This minimum number depends on the shape of the histogram and it is the function of the number of the pixels in the gray level interval [0,127] and [128,255] in the total interval of [0, 255]. The calculation based on the equation as follows:

\[
\text{MinPix}_{Bseed} = P_1 \sum_{i=0}^{127} h(z_i)
\]  

(2.13)
MinPix_{\text{Wseed}} = P_2 \sum_{i=1}^{255} h(z_i) \tag{2.14}

Where P_1, P_2 \in [0,1] since P_1 + P_2 = 0 and h(z_i) denotes the number of occurrences of gray level z_i. However, images with low contrast are equalized for significant contrast. The probability of occurrence of gray level z_i in an image is approximated by the following equation.

\[ p(z_i) = \frac{h(z_i)}{MN} \]

For discrete values

\[ T(z_i) = \sum_{k=0}^{h(z_i)} p(z_k) = \frac{\sum_{k=0}^{h(z_k)} h(z_k)}{MN} \]

Where M, N are image dimensions.

So, a processed image is attained by mapping each pixel with h level z_i in the input image into a corresponding pixel with level x_i = T(z_i) in the output image using previous image.

In reference [60] automatic fuzzy seed subsets based on image characteristics is proposed to identify threshold using ultrafuzzyness. However, the limited works already reported in literature leave many questions open. There is still room for investigations and introduction of new technologies as no single technique can threshold all kinds of images. The fusion of fuzziness, maximum entropy and spatial correlation on gray levels has scope to explore to obtain the optimal threshold.

**2.5 Maximum Entropy Techniques**

In the last 50 years there has been much debate between proponents of Bayesian and maximum entropy approaches [61]. There is a good reason for this debate. Both approaches produce different results that cannot be duplicated using the other. The
Bayesian and maximum entropy approaches make different assumptions and answer different questions. The Bayesian approach finds expectations while making some assumption about the prior probabilities of the distributions. The maximum entropy approach attempts to find the distribution in which the variables are least dependent on each other, i.e., the distribution that is closest to the uniform distribution. Due to this factor many researchers chose maximum entropy approach for the research, and the following works from the literature were studied.

T. Pun [62] first applied the concept of entropy to image thresholding, and assumed that an image is the outcome of a symbol source. Thus, in order to select the threshold he maximized an upper bound of total a posteriori entropy of the binary image.

Kapur et al. [63] propose a method based on the previous work of Pun. This method interprets the image object and background as two different information sources. When the sum of the object and background entropies reaches its maximum, the image is said to be thresholded optimally.

Kapur et al. also uses Shannon’s concept of entropy but from a different point of view. Two probability distributions of the entire image instead of one probability distribution are considered in their method. One probability distribution is for the object and the other is for the background. The sum of the individual entropy of the object and the background is then maximized. In other words, this will result in equal probability gray levels in each region, thus maximizing the sum of homogeneities in gray levels within object and background by making the gray levels equally probable in either region. The probability distribution of the gray levels over the background of the image

\[
\frac{p_o}{p_b}, \frac{p_1}{p_b'}, \ldots, \frac{p_r}{p_b'}
\]
and that of the object is

\[ P_{O_1}, P_{O_2}, ..., P_{O_{n-1}}. \]

In which, \( t \) is the threshold; \( p_i \) where \( (i=0,1,2, ..., n-1) \) is the statistical probability of pixels with gray level \( i \) in the whole image; \( P_B \) is the probability of pixels with gray level less than or equal to threshold \( t \).

\[ P_B = \sum_{i=0}^{t} p_i \]

The entropy of the background of the image is

\[ H_B(t) = - \sum_{i=0}^{t} \frac{p_i}{P_B} \ln \left( \frac{p_i}{P_B} \right) \]  
\[(2.15)\]

and that of the object of the image is

\[ H_O(t) = - \sum_{i=t+1}^{n-1} \frac{p_i}{1-P_B} \ln \left( \frac{p_i}{1-P_B} \right) \]  
\[(2.16)\]

The total entropy of the image is then defined as

\[ \phi(t) = H_O(t) + H_B(t) \]  
\[(2.17)\]

The optimal threshold, \( t^* \) is selected as the one which maximizes \( \phi(t) \).

The experiment results of Kapur’s method are shown in Figure 2.2.

![Figure 2.2 The thresholded images of coins and rice with Kapur’s method.](image)

C. H. Li and others [70-71] exploit cross-entropic thresholding techniques as the minimization of an information theoretic distance. Fuzzy entropic thresholding is another
branch of this group. H. D. Cheng and J. Chen [72] in their paper, the fuzzy set “brightness of gray levels” of an image is used to illustrate how the membership function can be determined automatically. Started with the concept of fuzzy event and use the maximum entropy principle as the criterion to find a membership function which will best represent the membership of brightness for each gray level in an image. This is further enhanced to determine Automatic bandwidth selection of fuzzy membership function by H. D. Cheng and Y. M. Lui [73]. The fuzzy entropy thresholding proposed by Cheng and et al. [74] employs the concept of fuzzy C-partition and the maximum fuzzy entropy principle. However, the search procedure for the parameters of the fuzzy sets in this method is not efficient and is time consuming. H.D. Cheng, Y.H.Chen and X.H. Jiang [75] presented a threshold based segmentation method using two-dimensional histogram and Fuzzy entropy, another variant is also found in [76]. Reference [77] provides another segmentation method based on Renyi’s entropy by P. K. Sahoo, C.Willkins, and J.Yeager. In reference [78] The two-dimensional Renyi’s entropy was obtained from the two-dimensional histogram which was determined by using the gray value of the pixels and the local average gray value of the pixels and includes a previously proposed global thresholding method [66]. M. Portes de Albuquerque et al. [79] presented a novel method using Tsallis entropy gives extensive and non extensive entropies with the introduction of a new parameter in the entropic criterion function. P.K. Sahoo, and G. Arora [80] presents a thresholding method for image segmentation by using an improved thresholding output function on a two-dimensional (2-D) histogram based on Tsallis-Havrda-Charvat entropy principle. The Tsallis-Havrda-Charvat entropy is obtained from two-dimensional histogram which has determined by using the gray
value of the pixels and the local average gray value of the pixels. Yonghao Xiao, Weiyu Yu [81] presented a method in which fuzzy entropy function is simplified with single parameter and Bee colony algorithm is applied to search the minimum value of fuzzy entropy function.

2.6 Maximum Entropy and Spatial Correlation Techniques

Originating from the work of [63] and that of Kirby and Rosenfeld [65], Abutaleb [66] extended the entropic thresholding technique by using two-dimensional histogram (2-D histogram). The two-dimensional histogram is obtained by the gray value and local average gray value of the immediate neighborhood pixels. Hence, the threshold becomes a vector having gray level value of the pixel and average gray level value of its neighborhood. Finally, the vector that maximizes the two-dimensional entropy is used as the two-dimensional threshold.

As following works, Brink [67], N. R. Pal et al. [68] refined abutaleb’s method. A. D. Brink also made use of two-dimensional histogram or scatterplot to make use of more information available in the image in pursuance of accurate threshold value. His approach is popularly known as maximin approach. According to Brink the normalized probabilities of background and objects, respectively, are:

$$P_0(t, s) = \sum_{j=0}^{s} \sum_{i=0}^{t} p_{ij}$$

$$P_1(t, s) = \sum_{j=s+1}^{n} \sum_{i=t+1}^{n-1} p_{ij}$$

But Abutaleb considers the second one that is object probability $P_1(t, s)$ as

$$P_1(t, s) = 1 - P_0(t, s)$$
In this approach, the process estimates a \((T,S)\) as threshold vector that maximizes the smaller value of \(H_0(t,s)\) and \(H_1(t,s)\) as

\[
H(T,S) = \max (\min \{ H_0(t,s), H_1(t,s) \}) \tag{2.18}
\]

Finally, this threshold vector is used for segmentation of a given image.

Liu et al. [21] suggested 2-D Otsu technique. Moreover, Wang et al. was the first to employ 2-D histogram in fuzzy thresholding [52]. According to A.K.C. Wong and P.K. Sahoo [69], the optimal threshold value is determined by maximizing the a posteriori entropy subject to certain inequality constraints which are derived by means of spectral measures characterizing uniformity and the shape of the regions in the image. For this purpose, the authors used both the gray-level distribution and the spatial information of an image. Yang Xiao et al., [82] & [83], defined gray level spatial correlation histogram. The local property of the image is considered differently when compared with traditional two dimensional histograms and with a weight function used in entropic criterion function calculation yielded better results. We, in reference [84] came out with fuzzy probability partition using GLSC histogram. The authors of [81] came forward with a novel idea to compute the varying similarity measure with human visual nonlinearity characteristics [85] and also tried with type-II fuzzy logic to compute maximum entropy. Due to the computational requirements of human visual nonlinearity characteristics and the time complexity involved in the process hereafter attention is paid on the gray images derived from the spatial parameters of the gray level spatial correlation histogram is proposed, but it is just an idea presented and needs to throw some additional light on this theory.
2.7 Chapter summary

This chapter presents a wide range of literature survey about the segmentation of the gray level images using threshold. There are pure entropy based methods, fuzzy based methods, methods dependent on the pixel arrangement in the spatial domain and methods with entropy and spatial features. Our main focus is on recent research that combines the spatial features with entropy based approaches. So, an in-depth survey is conducted on available resources and all of them are referred and presented and briefly explained in this chapter. Hence, the literature review induces innovative idea to study on combination of spatial concepts and ultrafuzzy approach for image segmentation. This is the topic of the 5th chapter where the proposed method is illustrated and the obtained results are compared with results of standard methods to prove its efficiency in the 6th chapter.