

Complex Background and Foreground Extraction in Color Document Images using Interval Type-2 Fuzzy

P.Murugeswari

Assistant Professor, Department of IT,
Sri Vidya College of Engineering and Technology,
Virudhunagar, Tamil Nadu, India

Dr.D.Manimegalai

Professor & Head, Department of IT,
National Engineering College, K.R. Nagar,
Kovilpatti, Tamil Nadu, India

ABSTRACT

This paper deals with the problem of extracting the text information from complex ground from color document images. Developing general framework for separating the foreground text and background information from complex document image is still a challenging problem because of its high unpredictability and complexity. In this paper a new interval type-2 fuzzy based thresholding method is proposed for processing color document images. The proposed method is experimented with varying background of multiple colors and texture and foreground text in any color, font, size and orientation. Experimental results show that the performance of the proposed approach is better than the existing ones.

Keywords- Document image preprocessing, Interval type-2 fuzzy, thresholding, binarization

1. INTRODUCTION

Analysis of document images for information extraction has become very important in recent years. Most of the traditional information is stored on the paper, but for the better storage and intelligent processing it is needed to convert it into electronic form. This aim is fulfilled by the image processing technique. Document image processing differs from the conventional image processing in its format and the information content. The objective of document image analysis is to recognize the text and graphics components in images and to extract the intended information as a human would. Document image analysis can be categorized by Textual processing and Graphics processing. Textual processing deals with the text components of a document image. The major tasks of textual processing are recognizing the text by optical character recognition (OCR), determining the skew, finding columns, paragraphs, text lines and words. Graphics processing deals with the non-textual line and symbol components that make up line diagrams, delimiting straight lines between text sections, company logos etc. Most of the graphics components consist of lines; such processing includes line thinning, line fitting, and corner and curve detection.

Document image processing need some preprocessing tasks. There are binarization, noise removal, skew detection & correction and marginal noise removal. The document image binarization means it automatically convert the image into bi-level form in such a way that the foreground information is represented by black pixels and the background by the white pixels. Several algorithms have been proposed for the document binarization task [1-4, 13, 14]. Fuzzy techniques are suitable for

the development of new algorithms because they are nonlinear knowledge based methods. It can be able to remove grayness ambiguities in a robust way.

Zadeh[2] first introduced the Type-1 Fuzzy Set(T1 FS) theory in 1965 and has been successfully applied in many areas including image processing, modeling and control, data mining, time-series prediction, etc. An example of a T1 FS is shown in Fig.1

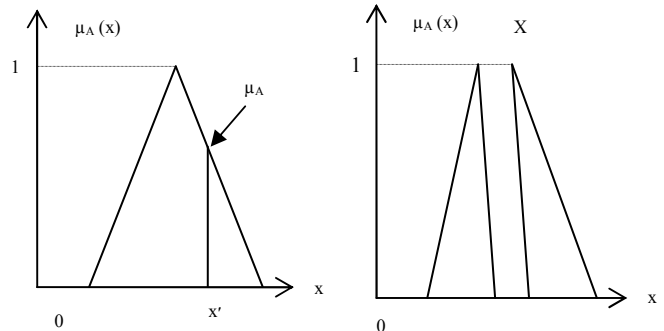


Fig .1 Type-1 Fuzzy Membership function

Many researches [3] have shown that there are limitations in the ability of T1 FSs to model and minimize the effect of uncertainties, because it's membership grades are crisp. Fig.2 shows type-2 FSs, characterized by Membership Functions (MFs) that are themselves fuzzy. Interval type-2 (IT2) FSs, a special case of type-2 FSs, are currently most widely used because of their reduced computational cost. An IT2 FS is bounded with above and below two T1 FSs, which are called Upper MF (UMF) and Lower MF (LMF) respectively. The area between X is Footprint Of Uncertainty (FOU).

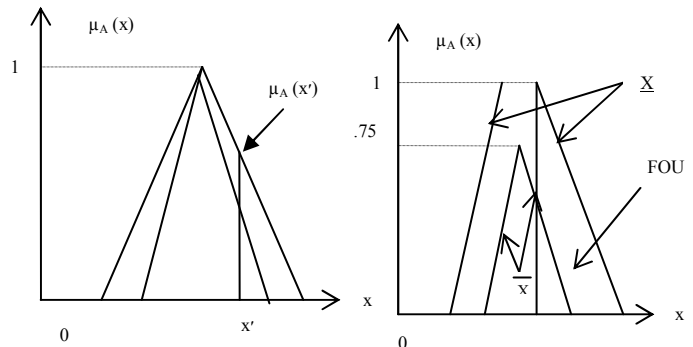


Fig .2 Type-2 Fuzzy Membership function

2. RELATED WORKS

In the literature, different existing conventional Type-1 and Type-2 fuzzy based document image binarization algorithms are studied. Document image binarization can be categorized as global or local binarization. The global binarization (thresholding) has a single threshold value to classify the image pixel into object or background classes. The local binarization (thresholding) means the local area information which is used to guide the threshold value for each pixel. Usually document images are degraded and have poor quality including shadows, non-uniform illumination, low contrast, large signal-dependent noise, smear and strains. Therefore global thresholding is not sufficient for binarization. Sezgin and Sankur[1] conducted an exhaustive survey of image thresholding methods. They categorized the thresholding methods according to the information they are exploiting, such as histogram shape, measurement space clustering, entropy, object attributes, spatial correlation, and local gray-level surface. Tizhoosh[6] carried out the survey of document image thresholding based on fuzzy. They categorized thresholding methods such as Fuzzy clustering, rule based approach, fuzzy geometrical approach and Information-theoretical approach. Fuzzy clustering means grouping the pixel value into classes. There are some algorithms such as fuzzy c-means (FCM), possibilistic c-means (PCM), etc., which are applied to image thresholding. Rule-based approach uses fuzzy if-then rules to find the suitable threshold and is suitable, if there exists an explicit expert knowledge about the image (e.g. in medical applications). Fuzzy-geometrical approach optimizes geometrical measure such as compactness, index of area coverage, etc. This approach uses spatial image information in contrast to other fuzzy techniques; Information-theoretical approach minimizes or maximizes measures of fuzziness and image information such as index of fuzziness or crispness, fuzzy entropy, fuzzy divergence, etc. Because of its simplicity and high speed, this approach is the most widely used fuzzy technique in the literature.

In recent years, more number of applications [5, 6] are developed using the type-2 fuzzy logic system. In the conventional (type-1) FLSs membership functions are scalar, but in the type-2 FLSs the membership function itself is fuzzy. This extra degree of fuzziness provides a more efficient way of handling uncertainty, which is certainly encountered in document image processing. Hence type-2 FLS may be utilized to design efficient thresholding selection which will give much better performance in document image binarization.

3. PROPOSED WORK

In this paper, a new approach is proposed to extract the foreground text from the complex background. The proposed approach has two steps; they are foreground and background separation and noise reduction.

3.1 Foreground and background separation

The Input color image is converted into the gray scale image. So the lowest intensity value $m = 0$ and the highest intensity value $M = 255$ (which means 256 gray scale). Two fuzzy sets need to be defined. The first is fuzzy set for foreground. The membership function of the set background is described in Fig.3.

The second fuzzy set is background. The membership function of the set foreground is described in Fig.4.

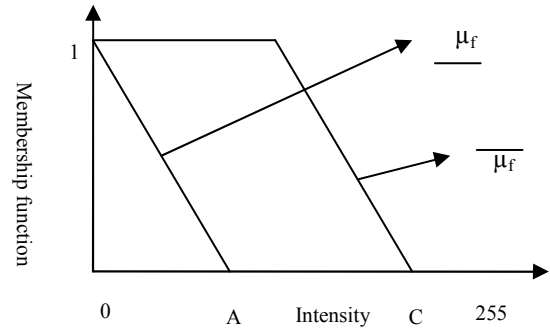


Fig.3 Membership function of fuzzy set background

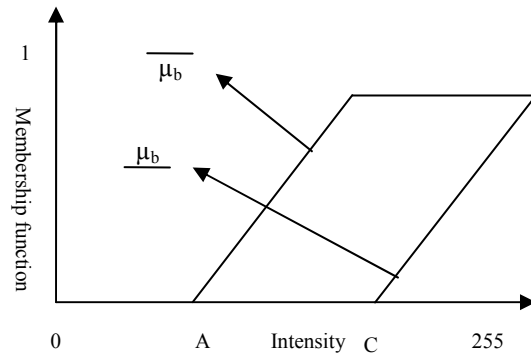


Fig.4 Membership function of fuzzy set foreground

3.2 Parameter estimation

Yan olinh [4] et. al is proposed to calculate the parameter estimation technique. Make use the same parameter estimation technique, FOU of Type-2 fuzzy set is calculated. The procedure of the technique is shown below:

Let x_i denotes pixel with intensity i and $f(x_i)$ denotes number of pixels with intensity i . The method to estimate parameters A and C is described below.

1. Pick a local maximum that is nearest to the lowest intensity (black) and call it p .
2. Pick the second local maximum that is nearest to the lowest intensity and call it q .
3. Maximize the following function by varying u from 1 to $(q-p)/2$. Then choose $u_{max} = u$ such that it maximizes the function $g(u)$.

$$g(u_{max}) = \max_{u=1..(p-q)/2} \{g(u)\} \quad (1)$$

Where

$$g(u) = \frac{\sum_{i=p}^{i=p+u-1} f(x_i)}{\sum_{i=p+u}^{i=p+2u-1} f(x_i)} \quad (2)$$

4. Choose $A = p + u_{max} - 1$
5. Maximize the following function by varying v from 1 to $(p-q)/2$. Then choose $v_{max} = v$ such that it maximizes the function $h(v)$.

$$h(v_{max}) = \max_{v=1..(p-q)/2} \{h(v)\} \quad (3)$$

Where

$$h(v) = \frac{\sum_{i=q-v+1}^{i=q} f(x_i)}{\sum_{i=q-2v+1}^{i=q-v} f(x_i)} \quad (4)$$

6. Choose $C = q - v \max + 1$

3.3 Binarization with Interval Type-2 fuzzy

The general algorithm for image thresholding based on interval type-2 sets and measures are formulated as follows:

1. Define the membership function and parameter for the foreground and background type-2 fuzzy set.
2. Following rule bases are defined for membership function
 - 2.1 If the pixel belongs strictly to the background then make it brighter (white)
 - 2.2 If the pixel belongs to the background then make it brighter
 - 2.3 If the pixel strictly belongs to the foreground then make it dark (black)
 - 2.4 If the pixel belongs to the foreground then make it darker
3. For computing the weight for the given inputs, compute the firing interval for the n^{th} rule.
4. Calculate in each position the upper and lower membership values
5. The threshold factor T_i of the rule is calculated by evaluating the membership expressions.

$$\overline{T}_i = \overline{\mu(x_i)} \cdot \mu(x_i) \quad (5)$$

6. The output of the T_i is also the interval type-2 set. The centroid defuzzifier is used to convert the interval type-2 fuzzy set into type-1 fuzzy, which is the final threshold value.

$$Td_i = \frac{T_i + \overline{T}_i}{2} \quad (6)$$

3.4 Noise reduction

Gatos et. al proposed to eliminate noise, improve the quality of text regions and preserve stroke connectivity by isolated pixel removal and filling of possible breaks, gaps or holes. This post processing has the two filters which are shrink filter and swell filter. Shrink filter is used to remove noise from the background and swell filter is used to fill breaks, gaps or holes in the foreground.

4. EXPERIMENTAL RESULTS

The proposed algorithm was tested using complex background which belongs to three categories: Microsoft background design in power point, old newspapers and poor quality modern documents. All images have varying resolutions, stroke sizes, illumination contrast, background complexity and noise levels. Table.1 shows the details of documents taken for experiment. The performance of the proposed method is compared with four well-known binarization techniques using MATLAB tool. The proposed method is evaluated in the following techniques: Otsu's global thresholding method, Niblack's adaptive thresholding method, and FCM thresholding.

Table 1. Experimental document data

Document Type	Background complexity	Foreground Details
Pages from old news paper	Non uniform patterned	Text titled in any orientation
Powerpoint slides	Background designs from Microsoft power point	Text of varying sizes
Invitation cards		
Airticles/Journals cover pages	Single or multicolored background	Text of varying size and dense & sparse text

Table 2 shows that the application of the proposed binarization technique has demonstrated the best performance regarding the final OCR results. Overall results show that the proposed method achieves at least 20% improvement compared to the other approaches.

$$\text{character(or word)recognition} = \frac{\text{Number of character(or words) correctly recognized}}{\text{Total number of characters (or words) in source document image}}$$

Table 2. OCR result comparison

	Recognition rate %				
	Original document	Otsu's method	Niblack's method	FCM method	Proposed method
No. of Characters	42%	55%	54%	60%	84%
No. of Words	36%	42%	41%	52%	70%

Table.3 shows the result of the various methods applied on a complex document.


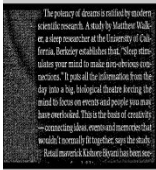
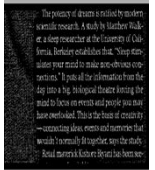

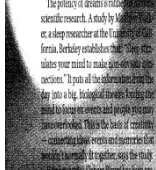
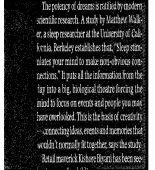
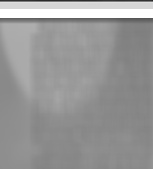
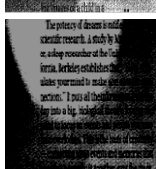
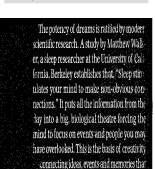

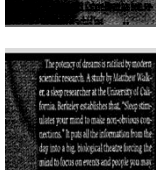

5.conclusion

In this paper, a new binarization method is proposed using interval type-2 fuzzy for document images. The proposed method has automatic parameter estimation so there is no need to tune manually and it can deal with any background complex documents. The proposed method has two steps: background and foreground separation and noise reduction. In order to improve the quality of text regions and preserve stroke connectivity the noise reduction is performed. It is observed that the performance of the proposed approach is better than the existing ones. Further research will focus on degraded document binarization using the Type-2 fuzzy.

6. ACKNOWLEDGEMENT

The authors would like to acknowledge the Tijuana Institute of Technology and Baja California Autonomous University, Tijuana Campus, Mexico for providing the Interval Type-2 Fuzzy toolbox.

Table 3. Comparison of binarization Result

Method	background	foreground	After Noise reduction
Niblack's adaptive thresholding method			
Otsu's global thresholding method			
Fuzzy C-Mean Method (Type-1 Fuzzy)			
Proposed method			

7. REFERENCES

[1] Mehmet Sezgin, Bu' lent Sankur, "Survey over image thresholding techniques and quantitative performance evaluation" Journal of Electronic Imaging 13(1), 146–165 (January 2004).

[2] B. Gatos, I. Pratikakis, S.J. Perantonis,, Adaptive degraded document image binarization Pattern Recognition 39 (2006) 317 – 327

[3] J. Sauvola, M. PietikaKinen ,”Adaptive document image binarization” Pattern Recognition 39 (2006) 317 – 327

[4] Reza Farrahi Moghaddam Mohamed Cheriet ,” A multiple framework for adaptive binarization of degraded document” images Pattern Recognition 43 (2010) 2186–2198

[5] Yan olinhin,C.G.Leedham,V.K.Sagar “A Fuzzy Based Handwriting Extraction Technique for Handwritten Document preprocessing”, 1996 IEEE TENCON-[6] Digital Signal Processing Applications

[6] HamidR. Tizhoosh,”Image thresholding using type II fuzzy sets”Pattern Recognition 38 (2005) 2363 – 2372

[7] Olivia Mendoza1 Patricia Melin2 Juan R. Castro1 “The Use of Interval Type-2 Fuzzy Logic as a General Method for Edge Detection”, IFSA-EUSFLAT 2009

[8] Dongrui Wu,” A Brief Tutorial on Interval Type-2 Fuzzy Sets and Systems “July 22, 2010

[9] G. Leedham, Y. Chen, K. Takru, J. H. N. Tan, and L. Mian, “Comparison of some thresholding algorithms for text/background segmentation in difficult document images,” Proceedings of 7th International Conference on Document Analysis and Recognition, pp. 859–864, 2003.

[10] S. Nirmala, P. Nagabhushan, “Isolation of foreground- text in document images having known complex back-ground,” Proceedings of 2nd International Conference on Cognition and Recognition, pp. 99–106, 2008.

[11] P. Nagabhushan, S. Nirmala ,”Text Extraction in Complex Color Document Images for Enhanced Readability”, Intelligent Information Management, 2010, 2, 120-133

[12] M. Seeger, C. Dance, Binarising Camera Images for OCR, Sixth International Conference on Document Analysis and Recognition (ICDAR’01), Seattle, Washington, 2001, pp. 54–58.

[13] N.Ostu, “A threshold selection method from gray-level histograms”,IEEE Trans. System Man Cybernet,9(1) 1979, 62-66

[14] M.Chang, S.Kang,W.Rho,H.Kim ,D.Kim, “Improved binarization algorithm for document image by histogram and edge detection”,ICDAR’95,1993,636-643