The Performance Evaluation of Thresholding Algorithms for Optical Character Recognition

A.T.Abak[1], U.Barış[1], B.Sankur[2]
[1] TÜBİTAK MAM Software Processing Group, 41470 Gebze-Kocaeli
{toygar, ufuk}@mozart.mam.gov.tr
[2] Boğaziçi University Electrical-Electronic Eng. Department, 80815 Bebek-Istanbul
sankur@boun.edu.tr

Abstract

This paper presents performance evaluation of thresholding algorithms in the context of document analysis and character recognition systems. Several thresholding algorithms are comparatively evaluated on the basis of the original bitmaps of characters. Different distance measures such as, Hausdorff, Jaccard, and Yule are used to measure the similarity between thresholded bitmaps and original bitmaps of characters.

I. Introduction

A document image analysis and recognition system includes several image processing techniques, beginning with digitizing the document and ending with character recognition and postprocessing.

Thresholding is a low-level image processing technique used, before document analysis step, for obtaining the binary image from its gray scale one. The result of thresholding affects the performance of successive image operations on the document character recognition. Images binarized via thresholding may contain noise pixels both in the background and foreground spoiling the original bitmaps of characters. Furthermore thresholding may cause various character deformations. Both spurious pixels as well as shape deformation are known to affect the recognition rate. Therefore criteria to assess thresholding algorithms must take into consideration both the noisiness of the image as well as the shape similarity of the characters. To this effect Hausdorff [1] metric is proposed for shape similarity measurement. On the other hand Jaccard and Yule [2] metrics focus on the differences of bitmaps. Both metrics will be eventually correlated with the correct recognition probability.

In Section II candidate thresholding algorithms are briefly described. Section III presents evaluation criteria and tests conducted. In section IV the results and conclusions are presented.

II. Thresholding Algorithms

Thresholding aims to separate a gray scale document image into two groups, that is foreground and background. There are two broad classes of thresholding algorithms, namely global methods and locally adaptive methods. These methods are listed in the sequel with brief reminders about each of them. We believe these algorithms are representative enough of the type of algorithms extant in the literature.

II.1 Global Thresholding

Lloyd: This algorithm places the optimum threshold at the middle point of the valley of the image bimodal histogram, by trying to minimize iteratively the misclassification error.

Otsu: Otsu minimizes the within-group variance or alternatively maximizes the between-group variance of the image.

Local Average Thresholding (LAT): The gray level image is thresholded by using the Otsu algorithm which calculates the optimum thresholding value over the local average histogram.

0-8186-7898-4/97 $10.00 © 1997 IEEE

Moment-preserving: This algorithm selects the threshold value in such a way as to maintain the equality between the first three moments of the gray scale image and those of the binarized image.

Pun1: Pun assumes that the histogram of the gray scale image is an L-symbol source. The optimum threshold is the threshold value where the total entropy of foreground and background pixels after thresholding is maximum.

Pun2: This algorithm uses an anisotropy coefficient to maximize the entropy of the thresholded images.

Kapur: Kapur also assumes that the foreground and background pixels constitute an L-symbol source. The optimum threshold is the one which maximizes the sum of the entropies of these two groups.

PP1: These two algorithms (PP1 and PP2) are based on the co-occurence matrix of the gray scale image. The sum of the entropies of the co-occurence matrix corresponding to within-class transitions is maximized to determine the optimum threshold.

PP2: This algorithm deals also with the co-occurence matrix as in PP1, but partitions the co-occurrence matrix to maximize the entropy of the between-class transitions.

YHM: The two maximum points of the histogram of the original image are calculated, then the middle point of these two peaks is found. The optimum threshold is found by using this middle point and the probability distribution function of the histogram of the equalized image.

Dynamic Thresholding (DTM): The range of the gray values of the image is determined and the optimum threshold is found by using the percentage of the dynamic of the histogram and the middle point found in YHM method.

II.2 Locally Adaptive Thresholding

Yasuda-Dubois-Huang (YDH): This algorithm includes the steps of dynamic range expansion (normalization), smoothing, adaptive thresholding and dynamic range expansion, and segmentation.

Local Contrast Technique (LCT): In this method the image is divided into 9x9 overlapping windows and the center pixels in these windows are binarized according to the contrast difference of pixels in the 3x3 corner windows and the center window.

Bernsen: The image is divided into overlapping windows having a predetermined size. The center pixels are thresholded according to the average of the maximum and the minimum values within these windows.

Nonlinear Dynamic Method (NDA): In this method the mean of the overlapping windows and a bias value are used to threshold the central pixels.

Logical Level Technique (LLT): Average stroke width of the characters (w) is predetermined and the central pixels are thresholded according to logical relations between the local averages of the pixels in (2w+1)x(2w+1) overlapping windows.

III. Evaluation

Several document images were produced using a variety of fonts, character sizes, typefaces, and pitch, using proportional spacing. These documents were than thresholded after printing and scanning. Both the original characters and the thresholded characters were placed in the bounding boxes. As the bitmap of original documents are known exactly the thresholded characters are checked vis-a-vis their originals using Hausdorff, Jaccard, and Yule similarity measures.

Hausdorff distance: Give two finite point sets, $A=\{a_1, a_2,...a_p\}$, $B=\{b_1, b_2,...b_q\}$, the directed Hausdorff distance from A to B. $h(A,B)$ is defined as

$$h(A,B)= \max_{a\in A} \min_{b\in B} \|a-b\|.$$

The Hausdorff distance itself is defined as

$$H(A,B)=\max\{h(A,B), h(B.A)\}.$$

Jaccard and Yule distances: Let Y and X be n-dimensional binary vectors; on the elements of these vectors following operations can be defined:

$$\delta_m(i,j) = \begin{cases} 1 & \text{if } x_m = i \,\&\, y_m = j \\ 0 & \text{otherwise} \end{cases}$$

$$n_{ij} = \sum_{m=1}^{n} \delta_m(i,j)$$

for $i, j = 0, 1$ and y_m, x_m are the mth elements of Y and X.

That is, n_{ij} is the number of occurrences where $X = i$ and $Y = j$. Then Jaccard and Yule similarity measures are defined as:

Jaccard: $n_{11} / (n_{11} + n_{10} + n_{01})$,

Yule: $(n_{11}n_{00} - n_{10}n_{01}) / (n_{11}n_{00} + n_{10}n_{01})$.

Finally the background contamination (pepper noise) and the foreground distortion can be assessed separately using the proportional match in the background and foreground pixels over all the document.

B_O : Background in the original image.
B_T : Background in the test image.
F_O : Foreground in the original image.
F_T : Foreground in the test image.

$$\beta = \frac{|B_O \cap B_T|}{|B_O|} \qquad \theta = \frac{|F_O \cap F_T|}{|F_O|}$$

Thus the total rate of misclassified pixels in the image is given by

$$\alpha = 1 - \frac{|B_O|\beta + |F_O|\theta}{|B_O| + |F_O|}$$

IV. Results and Conclusions

The result of extensive tests on scanned documents using measures given in the previous section are shown in Table 1.

The main conclusions that can be extracted from the experiments are as follows:

- The algorithms least contaminating the background are (highest β): Lloyd, Otsu, LAT, YHM, DTM, LCT, LLT.
- The algorithms least chipping away from the foreground are (highest θ): All of them.
- The highest ranking algorithms according to Hausdorff metric: NDA, Kapur, Lloyd, LAT, PP1, LCT, Otsu, Moment-preserving, PP2.
- The highest ranking algorithms according to Jaccard distance: LCT, Lloyd, Otsu, NDA, Moment-preserving, PP2, DTM, LLT.
- The highest ranking algorithms according to Yule distance: LAT, Kapur, PP1, YHM, NDA, Lloyd, Otsu, Moment-preserving, PP2, DTM, LCT, Bernsen, LLT.
- The fastest algorithms according to processing time are: Lloyd, Otsu, LAT, Moment-preserving, Pun1, Pun2, Kapur, YHM, DTM.
- The slowest algorithms according to processing time are: YDH, Bernsen, and NDA.

To reach to a final selection, we have summed the ranks for each method from the four distance criteria, namely, α-measure, Hausdorff, Jaccard, and Yule. When the algorithms were ranked according to these sums, we have observed that:

- The highest scoring five algorithms are: Lloyd, LCT, NDA, Otsu, LAT.
- The lowest scoring five algorithms are: Pun2, Pun1, YDH, Bernsen, PP1.

Acknowledgments

This work is supported by the GARİLDİ project of TÜBİTAK MAM, the Scientific and Technical Research Council of Turkey Marmara Research Center.

References

[1] D.P.Huttenlocher, G.A.Klanderman, W.J. Rucklidge. Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):850-863, September 1993.

[2] J.D.Tubbs. A note on binary template matching. *Pattern Recognition*, 22(4):359-365, 1989.

[3] P.K.Sahoo, S.Soltani, A.K.C.Wong, Y.C.Chen. A survey of thresholding techniques. *Computer Vision, Graphics, and Image Processing*, 41:233-260, 1988.

[4] M.Kamel, A.Zhao. Extraction of binary character/graphics images from grayscale document images. *CVGIP: Graphic Models and Image Processing*, 55(3):203-217, 1993.

[5] N.B.Venkateswarlu, R.D.Boyle. New segmentation techniques for document image analysis. *Image and Vision Computing*, 13(7):573-583, September 1995.

[6] Ø.D.Trier, A.K.Jain. Goal-directed evaluation of binarization methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(12):1191-1201, December 1995.

[7] N.R.Pal, S.K.Pal. Entropic thresholding. *Signal Processing*, 16(2):97-108, 1989.

[8] M.K.Yanni, E.Horne. A new approach to dynamic thresholding. *EUSIPCO-94*, 1:33-44, Edinburg 1994.

[9] L.Li, J.Gong, W.Chen. Gray-level image thresholding based on Fisher linear projection of two-dimensional histogram. *Pattern Recognition*, 30(5):743-749, 1997.

Methods	Subjective Comments	β, θ, α	Hausdorff	Jaccard	Yule	Time
LLOYD	satisfactory for images having bimodal histogram, but not satisfactory for very noisy images.	β=0.98 θ=0.99 **α=0.01**	**2.05**	**0.72**	**0.87**	**1**
OTSU	satisfactory for images having bimodal histogram, but not good enough for very noisy images or with creased papers.	β=0.98 θ=0.99 **α=0.01**	**2.08**	**0.72**	**0.87**	**1**
LAT	satisfactory for images having bimodal histogram, but not good enough for very noisy images or with creased papers.	β=0.98 θ=0.99 **α=0.01**	**2.06**	0.66	**0.89**	**1.1**
MOMENT	not always satisfactory for images having bimodal histogram and fails for images having active background.	β=0.93 θ=0.99 α=0.05	**2.10**	**0.71**	**0.87**	**1**
PUN1	poor, causes thickening of characters and cannot remove active background.	β=0.70 θ=1.00 α=0.24	6.61	0.45	0.80	**1.3**
PUN2	poor, causes characters become merged due to background noise.	β=0.60 θ=1.00 α=0.33	11.25	0.36	0.32	**1**
KAPUR	poor, causes thickening of characters and cannot remove active background.	β=0.75 θ=0.99 α=0.20	**2.04**	0.67	**0.89**	**1.3**
PP1	cause character merges with background noise.	β=0.64 θ=1.00 α=0.29	**2.06**	0.65	**0.89**	4.3
PP2	satisfactory for images having bimodal histogram, also good enough for images having active background.	β=0.95 θ=0.99 α=0.04	**2.10**	**0.71**	**0.87**	3.3
YHM	satisfactory for images having bimodal histogram, not good enough for images having active background.	β=0.98 θ=0.99 **α=0.01**	2.47	0.56	**0.88**	**1.5**
DTM	satisfactory for images having bimodal histogram, also good enough for images having active background.	β=0.99 θ=0.97 **α=0.01**	2.23	**0.68**	**0.87**	**1.5**
YDH	satisfactory, but thickens the characters. Processing time is large.	β=0.87 θ=0.99 α=0.1	2.40	0.56	0.75	78
LCT	satisfactory, parameter setting is easy, processing time is acceptable.	β=0.96 θ=0.99 **α=0.02**	**2.06**	**0.73**	**0.87**	3.8
BERNSEN	success of the method increases by the increasing window size, but processing time increases proportionally.	β=0.92 θ=0.99 α=0.06	2.80	0.67	**0.87**	275
NDA	satisfactory only when the parameters of the method are correctly adjusted.	β=0.92 θ=0.99 α=0.06	**2.00**	**0.72**	**0.88**	28
LLT	satisfactory, parameter setting is easy, processing time is acceptable.	β=0.97 θ=0.96 **α=0.02**	2.38	**0.68**	**0.87**	4.3

Table 1: Score table of thresholding algorithms.