Thresholding and Enhancement of Text Images for Character Recognition

W. W. Cindy Jiang*

Columbia University
Center for Telecommunications Research and
Department of Electrical Engineering
New York, NY 10027

ABSTRACT

A scheme which converts graytone text images of low spatial resolution to bi-level images of higher spatial resolution for character recognition are presented. Higher recognition rates are achieved when text images are processed using the proposed scheme, A good application of the proposed scheme is the recognition of characters in scene images.

1. INTRODUCTION

Most optical character recognition (OCR) [1] systems perform binarization of inputs before attempting recognition. However, a significant amount of information is lost during binarization. Text images, which are composed of fine strokes and are usually supposed to be binary, need be represented with high spatial resolution and low dynamic resolution for character recognition purposes. Our scheme make use of the information in graytone images and converts acquired graytone text images with low spatial resolution to bi-level images of higher spatial resolution. By trading dynamic resolution for spatial resolution, processed images are closer to ideal ones. Thus, better recognition rates can be achieved. In [2], topographical features, e.g. peak, pit, saddle, are extracted from graytone text images to make use of information in graytones. Our scheme is simpler and is effective for text image enhancement as well as for character recognition.

A survey on threshold selection techniques can be found in [3]. A global thresholding scheme using a correlation criterion was proposed in [4]. Adaptive thresholding techniques in [5, 6] are often used to segment text images with high resolution and non-plain background. Our proposed scheme is effective when spatial resolution of graytone images is not high enough for text to be recognizable in images. In scene images, spatial resolution is usually sacrificed in order to represent non-text parts of those images with high dynamic resolution. Thus, a good application of our scheme is the recognition of characters in scene images.

*Now with Kuang Wu Institute of Technology and Commerce, Taipei, Taiwan 112.

This work was supported in part by the National Science Foundation under grant #ECD-88-11111.

The rest of the paper is organized as follows. In the next section the proposed thresholding schemes are described. In section 3 morphological filtering is used to improve the image quality after thresholding. Section 4 presents the experimental results. A conclusion is given in section 5.

2. VARIABLE THRESHOLDING

Given the assumptions (1) text images have a purely white background and (2) the intensity of a graytone pixel is proportional to the size of the white area in that pixel, graytone text images are converted to bi-level text images with higher spatial resolution by the following proposed procedures:

(a) Reverse the intensity of the original graytone image by subtracting the intensity of each pixel from the full dynamic range, which will make its intensity proportional to the black area in the pixel.

(b) Construct a graytone image of nR dpi (dots per inch) from a graytone image of R dpi by linear interpolation, where nR is the desired spatial resolution of the converted bi-level image. Here we use bilinear interpolation for simplicity. Each pixel of the original graytone images is represented by $n \times n$ sub-pixels after the interpolation.

(c) Choose a threshold TH_i for each block i of size $n \times n$ in the interpolated images, such that the number of sub-pixels in that block with intensities larger than TH_i is proportional to the reversed intensity of the represented graytone pixel.

Because images of higher spatial resolution are obtained from interpolation and the neighboring pixels can have many possible intensities, no single fixed threshold can meet the proportion requirement of every block in an image at the same time. Therefore, variable thresholding is used to avoid this problem.

There are two ways of constructing threshold images for binarization of interpolated graytone images:

(d1) VV-thresholding: Apply variable thresholds on interpolated images. Construct threshold images with resolution nR dpi by linearly interpolating block thresholds (TH_i).

(d2) VF-thresholding: Apply different thresholds for different blocks while using a fixed threshold TH_i on all

0-7803-2431-5/95 $4.00 © 1995 IEEE

Table 1: Comparison of Similarity

	Number of processed sub-pixels which have intensities that differ from their originals		
	Resolution of test graytone image	100 dpi	133 dpi
	Total number of sub-pixels	652,960	1,163,722
I	Fixed thresholding without interpolation	42,324	52,650
II	Interpolation + Fixed thresholding	22,555	17,175
III	Interpolation + VV-thresholding	15,446	8,760
IV	Interpolation + VF-thresholding	14,338	5,838
V	Interpolation + VF-thresholding + Morphological filtering	12,654	5,548

sub-pixels of block i. Construct threshold images of resolution nR dpi by giving all sub-pixels of each block i the same TH_i.

Each sub-pixel in the interpolated graytone image has an associated threshold from step (d1) or (d2).

(e) If a sub-pixel intensity is larger than its threshold, binarize it into a black sub-pixel; otherwise make it a white sub-pixel.

3. MORPHOLOGICAL FILTERING

After VF-thresholding, many characters have chinked boundaries. A morphological filter is designed to smooth character boundaries. Each binarized sub-pixel is given a value of connectivity, which is the number of black sub-pixels next to it in the horizontal or the vertical direction. A black sub-pixel with low connectivity usually represents a non-smooth part of a character boundary. A white sub-pixel with high connectivity provides a good neighborhood to keep a black sub-pixel and make a character boundary smooth. In a 2x2 neighborhood of each sub-pixel, a black sub-pixel and a white sub-pixel are swapped if the black sub-pixel has lower connectivity than the white one. The same process is repeated afterwards in a neighborhood of size 4x4 to smooth character boundaries further.

4. EXPERIMENTAL RESULTS

In this section we compare the experimental results of using various methods. In the first subsection, the test image is scanned as bi-level images and regarded as original printed images; while the test image is scanned as a graytone image to serve as an input image in the second subsection.

4.1. Test with Bi-level Inputs

A paragraph for text having 750 characters is first scanned at 300 dpi and 400 dpi as two bi-level images and then regarded as the two original printed bi-level images for testing. Original images obtained in this way make it easier to compute the difference after thresholding, and will also make valid the assumption of a purely white image background. Graytone text images are obtained by counting the number of black sub-pixels in each $n \times n$ block of the original bi-level images to make valid the¥proportion assumption. The intensities of the graytone images are then normalized to make use of the fully dynamic range. Here we consider a block sixe of 3x3. Therefore, two graytone test images are of 100 dpi and 133 dpi respectively.

Given graytone text images, our goal is to convert them to binary form and make them close to their originals. Table 1 illustrates the performance of various schemes by comparing the number of processed sub-pixels which have intensities (0 or1) that differ from their originals. The smaller the number, the closer the processed image is to its original. Row I of Table 1 gives the results of applying fixed thresholding on graytone images without interpolation. Row II shows the results of applying fixed thresholding on interpolated graytone images. The results of applying VV-thresholding on interpolated graytone images are given in row III. Row IV shows the results of applying VF-thresholding on interpolated images. Images thresholded with this scheme need to be processed to make character boundaries smooth. In row V, the results show further improvement in

Figure 1. The original bi-level test image.

(a) (b) (c) (d) (e) (f)

Figure 2. The test graytone images and their processed bi-level images.

restoration with morphological smoothing after VF-thresholding.

To demonstrate the improvement of the proposed scheme qualitatively, another test image and its processed images are given in Figure 1 and 2. Figure 1 is a test image representing the original bi-level image. Part (a) of Figure 2 illustrates the graytone images obtained under the proportion assumtion with nine possible offsets on 3×3 blocks. Parts (b), (c), (d), (e), (f) of Figure 2 show the processed bi-level images using the schemes in row I, II, III, IV, V of Table 1, respectively. In parts (b) and (c) of Figure 2, the dot over the 'i' is missing under some offset conditions, which shows the need to use variable thresholding.

4.2. Test with Graytone Inputs

In this subsection, the test image with 750 characters is scanned at 150 dpi as a graytone image and shown in Figure 3. Figure 4 is the resulting image after fixed thresholding without interpolation. The processed image using the adaptive thresholding scheme in [6] is shown in Figure 5. Figure 6 shows the image processed using our proposed VF-thresholding plus the morphological filtering. When the image in figure 4 is recognized using the OmniPage™ - a commercial OCR software by Caere Corp., the rate is 77.9%; while it is 93.3% when the image in Figure 6 is tested. This shows the effectiveness of our scheme. Although the recognition result is not perfect. the improvement using our scheme is significant.

5. CONCLUSION

A variable thresholding scheme which converts graytone text images to bi-level images of higher resolution is presented. A morphological filter is employed to smooth the boundaries of thresholded characters. Experimental results show that images processed using our VF-thresholding plus morphological filtering can be recognized at a higher rate. The proposed scheme is effective for image enhancement. Thus, it is better for feature extraction in character recognition.

6. REFERENCES

[1] S. Mori, C. Y. Suen and K. Yamamoto, "Historical review of OCR research and development," *IEEE Proc.* vol. 80, pp. 1029-1058, July 1992.

[2] L. Wang and T. Pavlidis, "Direct gray-scale extraction of features for character recognition," *IEEE Trans. Patt. Anal. Machine Intell.*, 1993, PAMI-15, pp. 1053-1067.

[3] J. S. Weszka, "A survey of threshold selection techniques," *Computer Graphics and Image Processing*, 7:259-265, 1990.

[4] A. D. Brink, " Gray-level thresholding of images using a correlation criterion," *Pattern Recognition Letters*, 9(5), June 1989.

[5] J. M. White and G. D. Rohrer, "Image thresholding for optical character recognition and other applications requiring character image extraction," *IBM J. Res. Develop.*, vol. 27, no. 4, pp. 400-411, July 1983.

[6] E. Giuliano, O. Paitra and L. Stringa, "Electronic character reading system," *U S Patent No 4,047,152* dated Sept. 6, 1977.

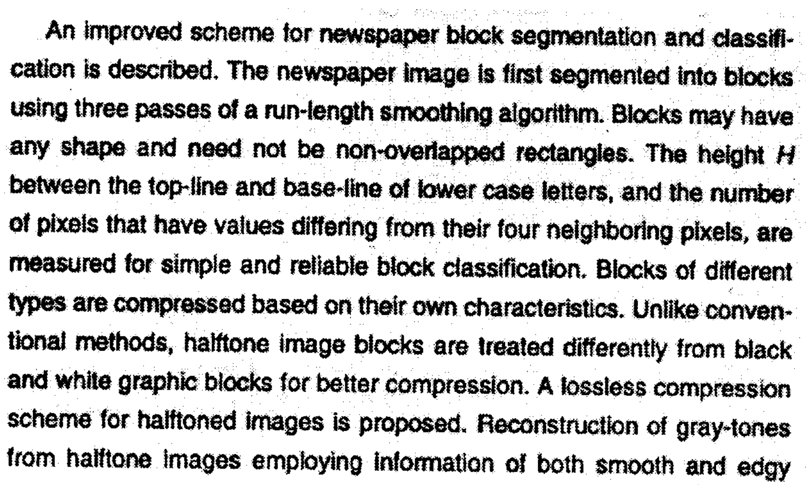

Figure 3. The original graytone test image.

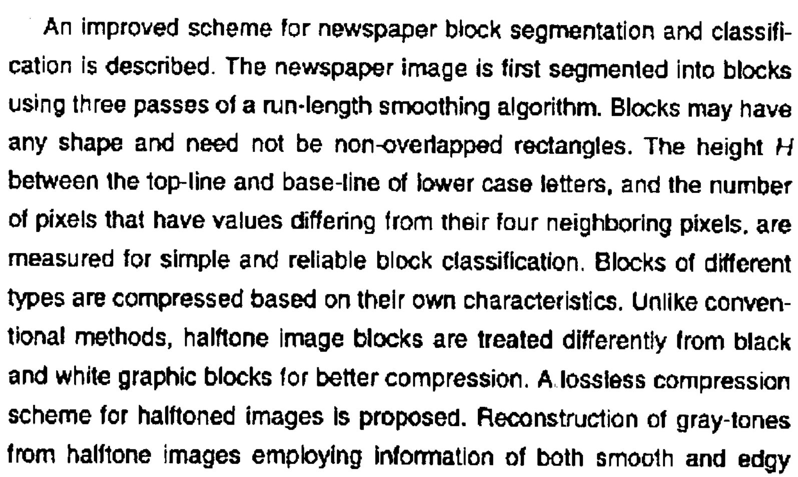

Figure 4 . The original image binarized using a fixed thresholding scheme.

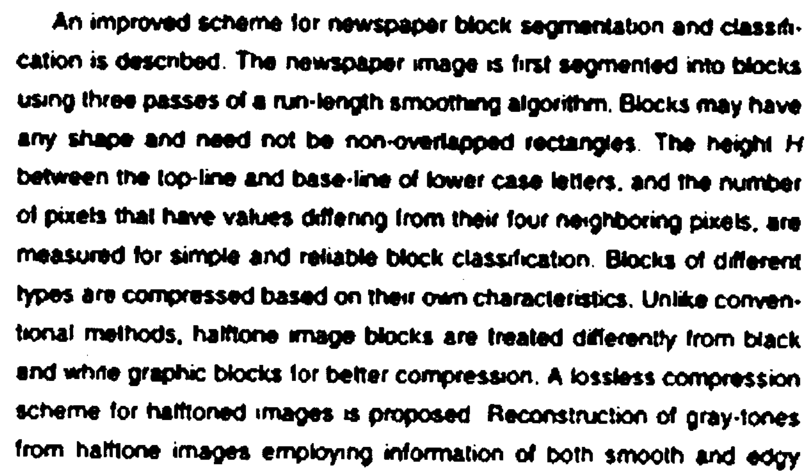

Figure 5. The test image binarized using an adaptive thresholding scheme.

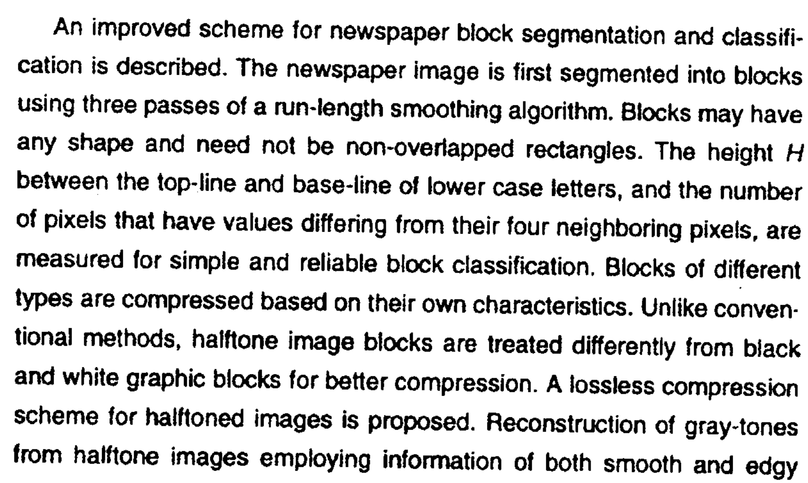

Figure 6. The image processed using our proposed scheme.