2009 10th International Conference on Document Analysis and Recognition

Threshold Correction of Document Image Binarization for Ruled-line Extraction

Hiroshi Tanaka
Fujitsu Laboratories Ltd.
4-1-1, Kamikodanaka, Nakahara-ku, Kawasaki, Kanagawa 211-8588, Japan
E-mail: htnk@jp.fujitsu.com

Abstract

In this paper, a new threshold correction method for document image binarization that is forcused on ruled-line extraction is presented. This method enhances the binary image of a ruled line, which is often adversely influenced by adjacent text pixels or background noise.

The threshold correction method consists of two submethods. One is a noise reduction method that is based on background determination, and the other is a threshold surface conversion method. Both these methods use the aspect of local straightness feature to distinguish ruled-line pixels from background pixels.

Keywords: adaptive binarization, ruled-line extraction, threshold correction, background noise

1. Introduction

One of the main objectives of document image binarization is to separate objects such as text or ruled lines from the document background. In a simple document model, each object is considered to be placed on the flat surface of the document background. According to this model, binarization can be reduced to a two-class discrimination problem for determining a global threshold [1]. However, complicated document images require adaptive binarization, in which the local threshold is calculated for each pixel. Such images have complex designs, which cannot be expressed using two classes; further, they could be severely degraded.

In the past, various adaptive binarization methods have been proposed. Trier [2] compared several binarization methods on the basis of their character recognition accuracies and concluded that Niblack's method [3] yields the optimal result when the noise reduction technique is applied. Sauvola [5] modified Niblack's method using region analysis, in which textual and nontextual regions were separated from each other. Sauvola's method has been the most popular binarization method for document images. These methods assume that pixels can be classified into two classes among local neighbors.

As adaptive binarization often generates small noise from flat background areas, noise reduction is essential to improve binarized images. Trier [2] used a method based on CCA (connected component analysis), and Eikvil [6] distinguished the background pixels from object pixels by using a flatness measure, which was calculated using the difference between the mean values of each class. Sauvola [5] used a measure similar to this flatness measure in order to separate textual and nontextual regions.

Our document recognition system recognizes ruled lines by extracting vertical or horizontal straight lines from the binary image, obtained by Niblack's binarization method and Eikvil's noise-reduction method. Eikvil's method has a problem that some pixels from faint ruled lines are often dropped off due to image deterioration (Fig.1). Therefore, we have developed a modified method that switches between two types of flatness parameters according to the morphological features of the image.

Adaptive binarization methods, including those developed by Niblack and Sauvola, also have a problem. Because these methods are based on the assumption that local neighbors can be classified into two classes, some pixels that have three or more pixel classes in each local area are often dropped off. This results in broken ruled lines near "strong" text pixels or an additional ruled line (Fig.2). We solve this problem by correcting the binarization threshold with respect to the neighboring threshold surface.

In Section 2, we describe the overall structure of our algorithm, along with the description of problems with the conventional method and our solutions to these problems. In Section 3, we present experimental results. Finally in Section 4, we conclude the paper.

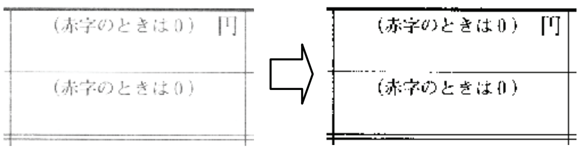


Figure 1. Broken faint line

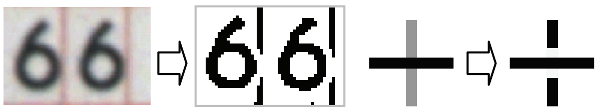


Figure 2. Effect of strong neighboring pixels

2. Binarization algorithm for ruled-line extraction

In our algorithm (Fig.3), initial thresholds are calculated for each pixel, and the pixels are initially classified as foreground or background pixels. The initial classification is used in the background determination process to determine whether each pixel is, in fact, a background pixel. For foreground pixels, binarization thresholds are obtained by correcting initial thresholds.

In the following subsections, the basic concept of the adaptive binarization method, which is used to calculate initial thresholds is described. Background determination, which is a modification of Eikvil's method, is then described in detail. Finally, the method for the correction of binarization thresholds is presented.

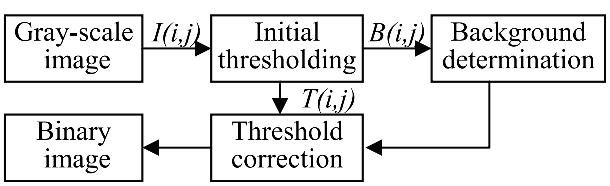


Figure 3. Binarization algorithm.

2.1. Initial thresholding

By initial thresholding, threshold values for each pixel are calculated and a threshold surface is formed using Niblack's method. The threshold for a pixel is obtained using the neighboring pixel values *I(i, j)*, reverse values of brightness, in the $w \times w$ local areas. The initial threshold *T* is calculated using the following formulas.

$$m = \sum_{i}^{w} \sum_{j}^{w} I(i,j) / w^2 \qquad (1)$$

$$\delta^2 = \sum_{i}^{w} \sum_{j}^{w} \{I(i,j) - m\}^2 / w^2 \qquad (2)$$

$$T = m + k\delta \qquad (3)$$

A binarized image is obtained by comparing each pixel value with its associated threshold (Fig.4). The value $k\delta$ in formula 3 is used to shift the threshold value so as to reduce background noise (Fig.5).

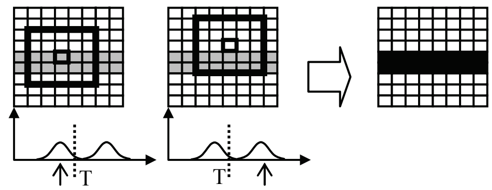


Figure 4. Thresholding by Niblack's method

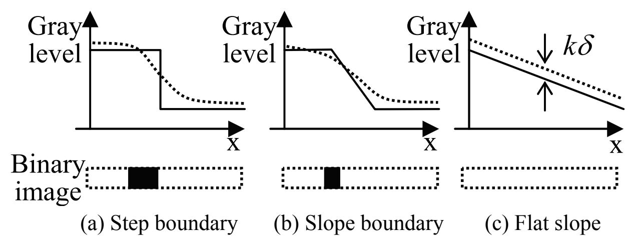


Figure 5. Binarization example

2.2. Background determination

When a background surface has small fluctuations, many tiny noises will be extracted from the background because the value $k\delta$ is so small that even minute noises cannot be suppressed (Fig.6). To reduce these noises, the value of background flatness is calculated, and noises in the flat region are eliminated.

We calculate the flatness value on the basis of Eikvil's method [6], which is described as follows. In the neighboring $w \times w$ local areas, each pixel can be classified according to the initial threshold (formula 3), and the mean values of each class, i.e., μ_1 and μ_2, can be obtained. The flatness value *F* of the central pixel of the local area is defined as $F = |\mu_1 - \mu_2| - d$, where *d* is a predefined background threshold, and the pixel is considered to be flat when *F* is negative.

For reference, Bernsen [4] has defined *F* as follows: $F = \max_{i,j}^{w}\{I(i,j)\} - \min_{i,j}^{w}\{I(i,j)\} - d$.

Sauvola has proposed a flatness measure called "transient difference" to distinguish between flat and nonflat regions.

Although Eikvil's method is effective for most images, it encounters problems while binarizing weak (faint or low-contrast) ruled-line images. Fig.7 shows an example of a weak ruled-line image. When d is set to a normal value (e.g., 10), some of the ruled-line pixels are dropped off because the difference between the two classes i.e. $|\mu_1 - \mu_2|$ is smaller than d (Fig.7(b)). However, when d is set to a lower value (e.g., 2), the background region produces noisy pixels (Fig.7(c)). To solve this problem, we use a morphological mask (Fig.8) to detect a straight line. The mask is set in the same area as the local area using which d is calculated. Let the number of pixels belonging to class 1 in the areas A, B, and C be denoted by *cntA*, *cntB*, and *cntC*, respectively. The value of d is reduced if the following condition is true (written using the C-Language notation), where L_1 and L_2 are predefined parameters.

$$(cntB > L_1)\ \&\&(cntA > L_2 \ || \ cntC > L_2)$$

In our experiments, $L_1 = w \times 0.75$ and $L_2 = w \times 0.25$, where w is the local area size. When the above condition is true, the value of d changes to $d = d \times 0.3$. Using this method for the conversion of d, a corrected binary image (such as that shown in Fig.7(d)) can be obtained.

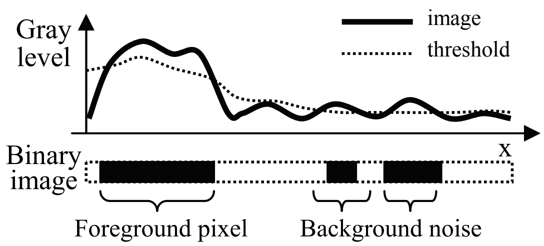


Figure 6. Background noise


(a) Gray image (b) d = 10 (c) d = 2 (d) Adaptive

Figure 7. Background decision adaptation

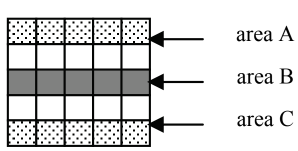


Figure 8. Morphological mask for horizontal line

2.3. Threshold correction

As adaptive binarization methods, such as those developed by Niblack and Sauvola, assume that a local area contains pixels of two classes, local areas containing more than two classes of pixels often encounter classification failures. Fig.9 shows the effect of neighboring pixels on the binarization threshold. When a local area contains a third class of pixels (black pixels in the figure), the threshold shifts because of the change in the mean value m (see formula (3)). This causes a discontinuity in the ruled line shown in Fig.2.

The effect of the third class of pixels can be prevented if the local area is shifted away from this class as shown in Fig.10. This shift is equivalent to using the binarization threshold of the shifted position at a shift width S from the current pixel. We define the threshold conversion using a shift operation as follows.

$$T'(i,j) = \max_{k,l=-S}^{S}\{T(i+k, j+l)\} \qquad (4)$$

By carrying out conversion using formula (4), the result of binarization of the image shown in Fig.11(a) changes from that shown in Fig.11(b) to that shown in Fig.11(c). Please note that S must be smaller than $w/2$, or the current pixel should be outside the shifted local window. Although the gap created by the dropped ruled-line pixels is bridged, many of the other white pixels are incorrectly converted into black pixels. In our system, all the converted black pixels are considered as "reserved" and original black pixels are considered as "fixed." The reserved pixels are converted into fixed pixels by hysteresis only when they are connected to a sequence of sufficient number of fixed pixels. This conversion is shown in Fig.11(d)–(f), and an example of threshold correction is shown in Fig.12.

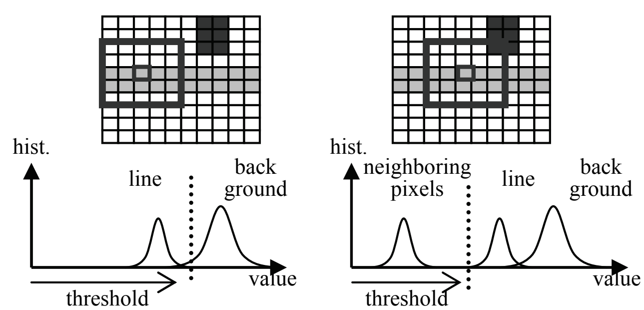


Figure 9. Affect of neighboring pixels

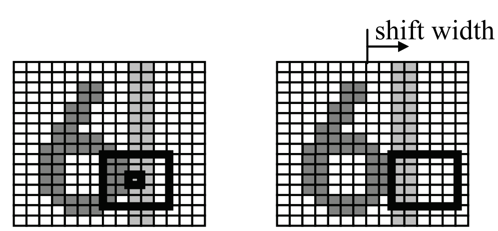
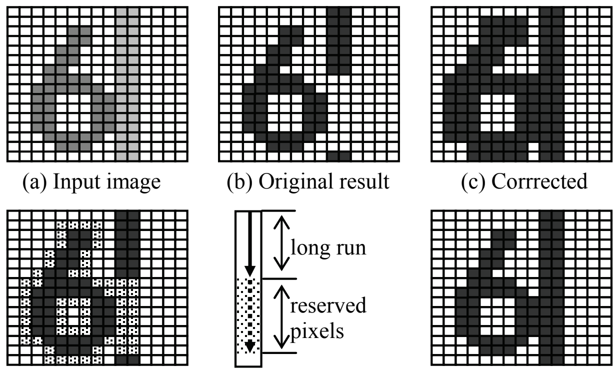


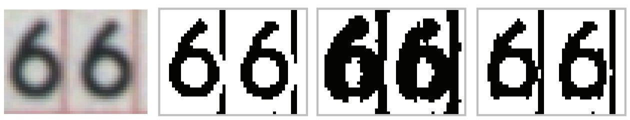
Figure 10. Shifted local window



(a) Input image (b) Original result (c) Corrrected

(d) Reserved pixels (e) Hysteresis (f) Final result

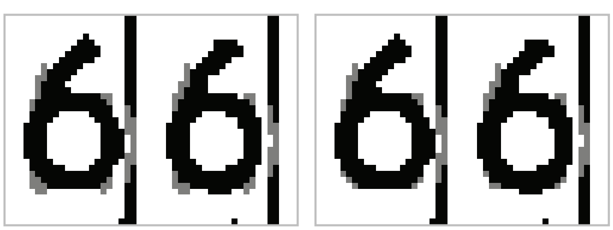
Figure 11. Threshold correction



(a) Input (b) Original (c) Corrected (d) Result

Figure 12. Example of threshold correction for vertical line detection

The hysteresis may produce insufficient deformation in the image of “6,” as shown in Fig.12(d). This deformed image can be improved by introducing a limitation on fluctuations within the reserved pixels converted to fixed pixels. As the converted pixels are assumed to belong to the same class, each pixel value should be within a certain range (such as $\pm d$). Fig.13 shows an example of the improvement of an image.



(a) Hysteresis (b) Hysteresis+limited fluctuation

Figure 13. Example of conversion of reserved pixels into fixed pixels (gray)

3. Evaluation results

We tested our method using three types of document image sets. Set 1 consists of scanned images of shipping slips (20 images), set 2 consists of images of shipping slips captured by using a digital camera (28 images), and set 3 consists of scanned images of invoices (21 images).

Table.1 shows the results of ruled-line extraction along with the extraction accuracy and relevance rate. Although the absolute values do not imply any thing because they depend on the test images, the results of the developed method are relatively more accurate than those of the conventional method. It is observed that two-thirds of the ruled lines are broken because of the weak image and four-fifth of the broken lines caused by threshold change are corrected. This result indicates the effectiveness of our method.

Table 1. Ruled-line extraction rate

	slip (scan)		slip (camera)		invoice (scan)	
	Acc.	Rel.	Acc.	Rel.	Acc.	Rel.
Conventional method	82.0%	89.6%	81.5%	84.8%	93.5%	90.2%
Developed method	92.7%	94.7%	89.9%	91.3%	97.4%	90.5%

Table 2. Causes of misrecognition

	conv.		developed		difference	
	drop	add	drop	add	drop	add
Broken faint line	62	32	24	13	-38	-19
Neighboring strong pixels	25	0	5	0	-20	0

4. Conclusion

In this paper, we have described a threshold correction method that can be used to binarize ruled-line images. This method consists of two submethods: one is an improvement of background determination based on Eikvil’s noise reduction method and the other is a threshold conversion method. Both these submethods use morphological information regarding a straight-line image in order to suppress the increase in noise.

The basic concept of the developed method is to combine simple parameter conversion and morphological information. Since the shape of a ruled line is very simple, it is easy to utilize morphological information for ruled-line extraction.

As the next step, it is necessary to think about applying our method to images that do not consist of ruled lines, such as text images. As text does not necessarily have long straight lines, our method cannot be applied in its present form. For example, low-contrast texts are often dropped off, and they cannot be recovered as in the case shown in Fig.7. We are currently studying some other morphological features that can be applied to images other than ruled-line images.

References

[1] N. Otsu, "A Threshold Selection Method from Gray-level Histograms", *IEEE Trans. Systems, Man, and Cybernetics*, vol.9, no.1, 1979, pp. 62-66.

[2] O.D. Trier and A.K. Jain, "Goal-directed Evaluation of Bainarization Methods", *IEEE Trans. PAMI*. vol.17, no.12, Dec. 1995, pp. 1191-1201.

[3] W. Niblack, *An Introduction to Digital Image Processing*, Prentice Hall, Englewood Cliffs, N.J., 1986, pp. 115-116.

[4] J. Bernsen, "Dynamic Thresholding of Gray-level Images", Proc. 8th ICPR, Paris, France, Oct. 1986, pp. 1251-1255.

[5] J. Sauvola, T. Seppanen, S. Haapakoski, and M. Pietikainen, "Adaptive Document Binarization", Proc. 4th ICDAR, Ulm, Germany, Aug. 1997, pp. 147-152.

[6] L. Eikvil, T. Taxt, and K. Moen, "A Fast Adaptive Method for Binarization of Document Images", Proc. 1st ICDAR, Saint-Malo, France, Sep. 1991, pp. 435-443.