A Threshlod Selection Method Based on Multiscale and Graylevel Co-occurrence Matrix Analysis

Yun Li[1], Mohamed Cheriet[2], Ching Y, Suen[1]

1. Center for Pattern Recognition and Machine Intelligence
Concordia University, GM606, 1455 de Maisonneuve Blvd. West
Montreal, Quebec, H3G 1M8, Canada

2. Laboratory for Imagery, Vision and Artificial Intelligence
Ecole de Technologie Supérieure, University of Quebec
1100, Notre-Dame West, Montreal, Quebec, H3C 1K3, Canada

yunli@cenparmi.concordia.ca

Abstract

Noises and complex backgrounds often make the thresholding of degraded document images a difficult job. In this paper, we propose a new threshold selection method to handle severely degraded document images. First, multiscale image description is adopted to analyze the image edge. From it, edge pixel pair information is derived and recorded by a graylevel co-occurrence matrix. An appropriate threshold value is chosen by measuring the edge pixel pair co-occurrence matrix. The new method is tested with degraded document images. The experimental results show it is resistant to noises and complex backgrounds.

1. Introduction

In image processing, binarization plays an important role and has attracted much interest for many years. This operation divides the image into two classes: the objects and the background. Compared with the grayscale image, a binary image decreases the storage and computational cost of subsequent processing. Binarization is usually the first step in most document image analysis systems. If poorly binarized images are used, document understanding would be difficult or even impossible. Many threshold selection methods [1-8] have been proposed which decide the threshold globally or adaptively. Different kinds of statistical characteristics, such as maximum likelihood [2], moment [7], entropy [6], variance between the foreground and background [4], are adopted for threshold selection. Some comprehensive and thorough overviews and evaluations about binarization method can be found in [9-11].

Although many kinds of threshold selection approaches exist, they seldomly yield satisfactory binarization result when applied on document images with noises and complex backgrounds. Even few adaptive methods may perform good binarization, they are too complicate and many parameters have to be fine-tuned. Here we propose a new simple threshold selection method. Instead of attempting to choose the threshold based on image histogram information like many other methods, we select the threshold in terms of multiscale edge analysis and graylevel co-occurrence matrix. The method can effectively select a proper threshold even the image is degraded by noises and disturbing backgrounds.

2. Threshold selection

2.1. Multiscale edge analysis

First, the Laplacian of the image is computed for locating the edges. For an image I, the Laplacian of I is

$$\nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} .$$

The discretion Laplacian at a point I(x, y) can be computed as:

$$\nabla^2 I(x,y) = I(x-1,y) + I(x+1,y) + I(x,y-1) + I(x,y+1) - 4 * I(x,y) .$$

After calculating the Laplacian value on all the points, we try to find the position where successive pixels have different signs of the Laplacian values, such as, - and +, +

and -, - and 0 and +, or + and 0 and -. The two connected pixels having signs of Laplacian + and – respectively are called an edge pixel pair and the edges will be located on these positions according to the zero-crossing edge detection method. An ideal threshold value should lie between the graylevel values of edge pixel pair so that objects and background can be segmented correctly. Figure 1 illustrates this idea in 1-D case.

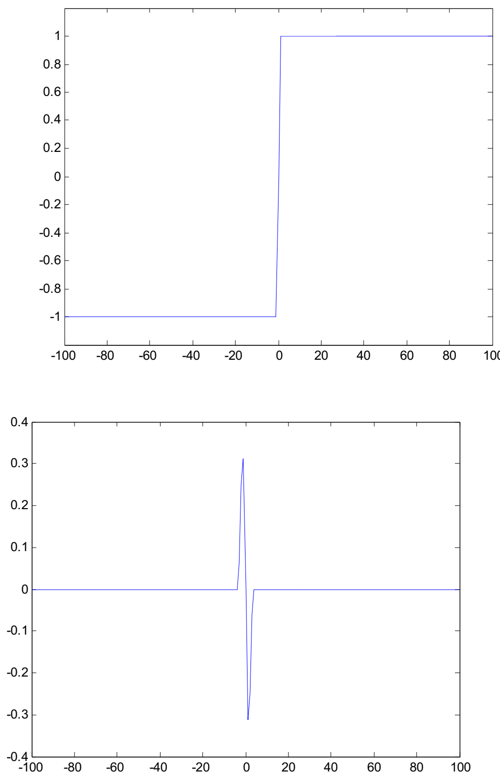

Figure 1. A 1-D step edge and convolution result with a LoG filter.

As shown in Figure 1, a step edge is formed where image changes from foreground to background, for example 1 represents foreground, -1 represents background. After blurring by the Gaussian kernel, the Laplacian of the edge will have a positive peak and a negative peak. There is a zero-crossing on the position 0. The two points at postions 1 and -1 form an edge pixel pair, which have different signs of Laplacian value. The ideal threshold will lie between the graylevels of these two points. If the threshold exceeds the graylevel of point 1, part of the foreground will be classified as background. On the other hand, if the threshold is lower than the graylevel of point -1, part of the background will be considered as foreground. Hence, the edge pixel pair information can help us to select a proper threshold.

In general, a document image is often degraded by noises which may produce the oscillation of the second derivative of the image. Many pixels related to noises will be falsely detected as an edge pixel pair, which influences the estimation of the threshold. It is desirable to suppress the noises and focus on true character stroke edges. For this purpose, we lower the image to a coarse scale and compute the Laplacian on a coarse resolution image. It is well known that multiscale descriptions of image are very useful in revealing the different image details [15]. On the fine scale, we can see more details, also noises. On the coarse scale, we omit some details and only see some significant boundaries. For noise suppression, we lower the image resolution by convolving the original image with a Gaussian kernel.

Another problem occurs when we analyze the image on a coarse scale. The location of the boundaries may shift a little compared with true boundaries on the original image. Therefore, the edge pixel pairs only produced from the coarse scale image can not precisely estimate the threshold of the original image. For accurately locating the edge pixel pair, we also find those edge positions from the original image. Comparing the edge positions detected from the original fine scale image and those from the coarse scale image, we remove those edge positions on the fine scale which do not have corresponding edge points at the same positions or 4-connected positions on the coarse scale. By this means, we attempt to reduce the false detection occurred by noises and acquire the accurate location of edge position.

2.2. Graylevel co-occurrence matrix

We will select a threshold based on the edge pixel pair information acquired by multscale analysis. In a document image, the foreground usually contains some characters which consist of many strokes. Because the graylevels of these character strokes are similar, after blurring by point spread function, the pixel values of the blurred edges for those stroke patterns look very similar too. Instead, the graylevels related to the noises and complex background patterns vary considerably. Therefore, the blurring at the edge of noises and background will show different pixel values. This makes the edge pixel pairs related to the character strokes gather in a narrow interval, and the edge pixel pairs related to background distribute sparsely.

The graylevel co-occurrence matrix is adopted to measure the characteristics of the edge pixel pairs. We use co-occurrence matrix because sometimes information provided by histograms is not good enough for selecting a proper threshold, and “In such case, the graylevel co-occurrence matrix can be an aid to threshold selection” [14]. The graylevel co-occurrence matrix has been proposed for threshold selection by many researchers [12, 13], but all the pixels in the image, including noises and undesired background pixels, are counted in the previous works. When there are lots of noises and variation in the background, the old methods can not give an appropriate threshold. In our method, only edge pixel pairs are considered. This information represents the characteristics of the stroke edges. Hence, our method is accurate and robust.

First, we give a brief description about the graylevel co-occurrence matrix. Let m represent the graylevel of pixel (x, y) and n represent the graylevel of pixel (x- sinθ, y+ cosθ) of an L-level gray scale image I, 0 <= x <= M-1, 0 <= y <= N-1, and 0 <= m, n <= L-1. Now the elements of the graylevel co-occurrrence matrix $[C_\theta]$ for the direction θ can be defined as

$$c_{\theta,m,n} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} N(I(x,y)=m \;\&\; I(x-\sin\theta, y+\cos\theta)=n),$$

where N(·) = 1 if the argument is true, otherwise, N(·) = 0. $[C_\theta]$ is an L*L matrix, and the element $c_{\theta,m,n}$ gives the frequency that a pixel having graylevel m is adjacent to a pixel having graylevel n in direction θ. Usually θ is chosen as 0, π/2, π, 3 π/2, which are 4-connected directions. Considering all these directions, the co-occurrence matrix is

$$[C] = \frac{1}{4}([C_0]+[C_{\frac{\pi}{2}}]+[C_\pi]+[C_{\frac{3}{2}\pi}]).$$

We adopt co-occurrence matrix to analyze the edge pixel pair. Therefore, only two adjacent pixels which form an edge pixel pair will be counted in our new method. The definition of the elements of the matrix will be modified as

$$c_{\theta,m,n} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} N\begin{pmatrix} I(x,y)=m \,\&\, I(x-\sin\theta, y+\cos\theta)=n \,\&\ldots \\ I(x,y), I(x-\sin\theta, y+\cos\theta) \text{ are edge pixel pair}\end{pmatrix}.$$

Then we can calculate the edge pixel pair graylevel co-occurrence matrix [C] as mentioned before.

The edge pixel pair graylevel co-occurrence matrix is used to select a threshold. Because there are many stroke patterns which have similar graylevels in the document image, many edge pixel pairs related to character strokes will occur in some approximate intervals. However, the edge pixel pairs related to noises or background will occur in many different intervals. We want to select a threshold which will be in the graylevel values of most edge pixel pairs. Note [C] is a symmetric matrix, we only consider the upper triangle of the matrix. The threshold T is chosen as:

$$T = \arg\max_{0\le t\le L-1}\left(\sum_{m=0}^{t}\sum_{n=t}^{L-1} c_{m,n}\right)$$

By this means, the threshold T can separate most edge pixel pairs. Figure 2 illustrates the threshold selection by the edge pixel pair graylevel co-occurrence matrix.

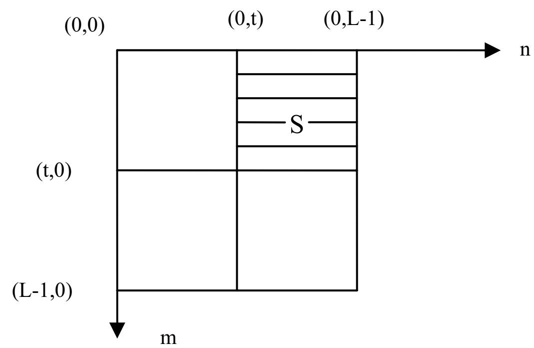

Figure 2. Graylevel co-occurrence matrix, where t is value of threshold. Select T to maximize the sum of entries in area S. The T will be between the graylevels of most edge pixel pairs.

2.3. New algorithm

Now we summarize our new threshold selection algorithm as follows:

1. Lower the image scale and compute the Laplacian of the coarse scale image. Find edge points according to the change of the sign of Laplacian.
2. Detect edge points in the original fine scale image. Remove those points which have no corresponding edge points at the same positions or 4-connected positions on the coarse scale, because those points are probably produced by noises.
3. Get the edge pixel pair related to each edge point. An edge pixel pair is formed by two pixels which are near edge points and whose signs of Laplacian change.
4. Record these edge pixel pairs to a graylevel co-occurrence matrix C. Select the threshold T as

$$T = \arg\max_{0\le t\le L-1}\left(\sum_{m=0}^{t}\sum_{n=t}^{L-1} c_{m,n}\right)$$

3. Experimental results

Our method has been tested with 25 document images which are degraded by noise, low contrast, or complex background. In Figure 3-4, we show some binarization

Proceedings of the 2005 Eight International Conference on Document Analysis and Recognition (ICDAR’05)
1520-5263/05 $20.00 © 2005 **IEEE**

results by our method. For comparison, the results by Otsu's method [4], Tsai's moment preservation method [7] and Kapur's maximization of entropy of the image [6] are also illustrated. Otsu's binarization algorithm is one of the most popular globally thresholding methods. It is simple and easy to implenment. In most simple cases, Otsu's method is optimal. Tsai's moment preservation and Kapur's entropy maximization are also frequently cited. They use different statistical characteristics of the image.

The results indicate that our method performs well even when the document images have been degraded by lots of noises and undesirable background. The foreground characters are accurately segmented. From the images, we can observe the edge pixel pairs related to the noises and background are very different, but the blurred edges of strokes have similar pixel values. Therefore, by considering only those edge pixel pairs which occur frequently, we can successfully remove the influence of noises and background.

4. Conclusions

We have proposed a new threshold selection method for document images based on multiscale and graylevel co-occurrence matrix. The multiscale analysis is used to find edge pixel pairs related to foreground strokes. Then graylevel co-occurrence matrix of the edge pixel pairs is adopted to determine a proper threshold. Because edge pixel pairs represent the characteristics of the strokes and we acquire edge pixel information with noise suppressed, our method is accurate and robust. Experimental results show this method performs well even on severely degraded documents.

References

[1] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recogntion,* Vol. 19, No. 1, pp 41-47, 1986.

[2] T. Kurita, N. Otsu, and N. Abdelmalek, "Maximum likelihood thresholding based on population mixture medels," *Pattern Recogntion,* Vol. 25, No. 10, pp. 1231-1240, 1992.

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

[4] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Systems, Man, and Cybernetics,* Vol. 9, No. 1, pp. 62-66, 1979.

[5] S.D. Yanowitz and A.M. Bruckstein, "A new method for image segmentation," *Computer Vision, Graphics and Image Processing,* Vol. 46, No. 1, pp. 82-95, 1989.

[6] J.N. Kapur, P.K. Sahoo, and A.K.C. Wong, "A new method for gray-level picture thersholding using the entropy of the histogram," *Computer Vision, Graphics and Image Processing,* Vol. 29, pp. 273-285, 1985.

[7] W. Tsai, "Moment-preserving thresholding: a new approach," *Computer Vision, Graphics and Image Processing,* Vol. 29, pp. 377-393, 1985.

[8] J.R. Parker, "Gray level thresholding in badly illuminated images," *IEEE Trans. Pattern Analysis and Machine Intelligence,* Vol. 13, No. 8, pp. 813-819, 1991.

[9] Ø.D. Trier and T. Taxt, "Evaluation of binarization methods for document images," *IEEE Trans. Pattern Analysis and Machine Intelligence,* Vol. 17, No. 3, pp. 312-315, 1995.

[10] Ø.D. Trier and A.K. Jain, "Goal-directed evaluation of binarization methods," *IEEE Trans. Pattern Analysis and Machine Intelligence,* Vol. 17, No. 12, pp. 1191-1201, 1995.

[11] B. Sankur, M. Sezgin, "A survey over image thresholding techniques and quantitative performance evaluation," *Journal of Electronic Imaging,* Vol. 13 No. 1, pp. 146-165, 2004.

[12] J.S. Weszka and A. Rosenfeld, "Threshold evaluation techniques," *IEEE Trans. Systems, Man, and Cybernetics,* Vol. 8, pp. 622-629, 1978.

[13] F. Derevi and S.K. Pal, "Graylevel thresholding using second order statistics", *Pattern Recogntion Letter,* Vol. 1, pp. 417-422, 1983.

[14] B. Chanda and D.D. Majumder, "A note on the use of the graylevel co-occurrence matrix in threshold selection," *Signal Processing,* Vol. 15, pp. 149-167, 1988.

[15] J. Babaud, A. Witkin, M. Baudin, and R. Duda, "Uniqueness of the Gaussian kernel for scale-space filtering," *IEEE Trans. Pattern Analysis and Machine Intelligence,* Vol. 8, No. 1, pp.26-33, 1986.

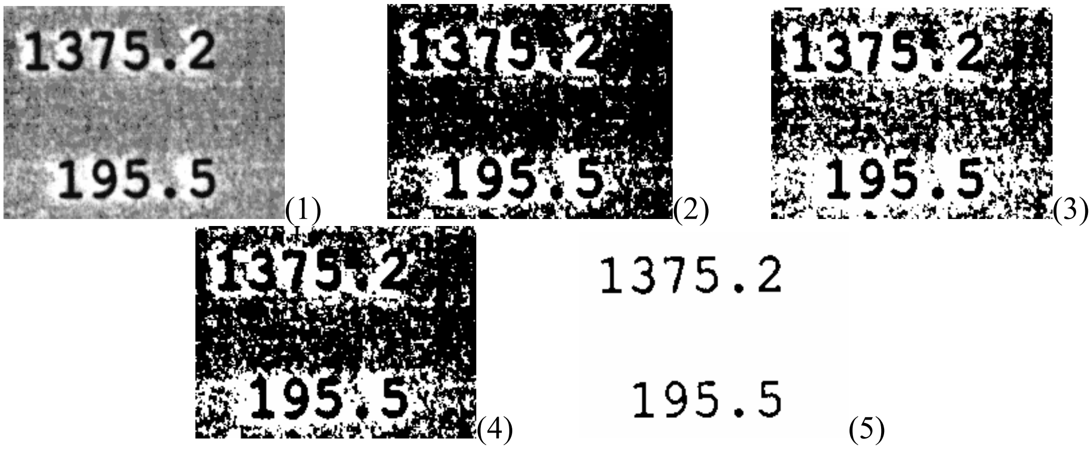

Figure 3. (1) original image, (2)-(5) binarization by Kapur's entropy method, Tsai's moment preservation, Otsu's method and our new method respectively.

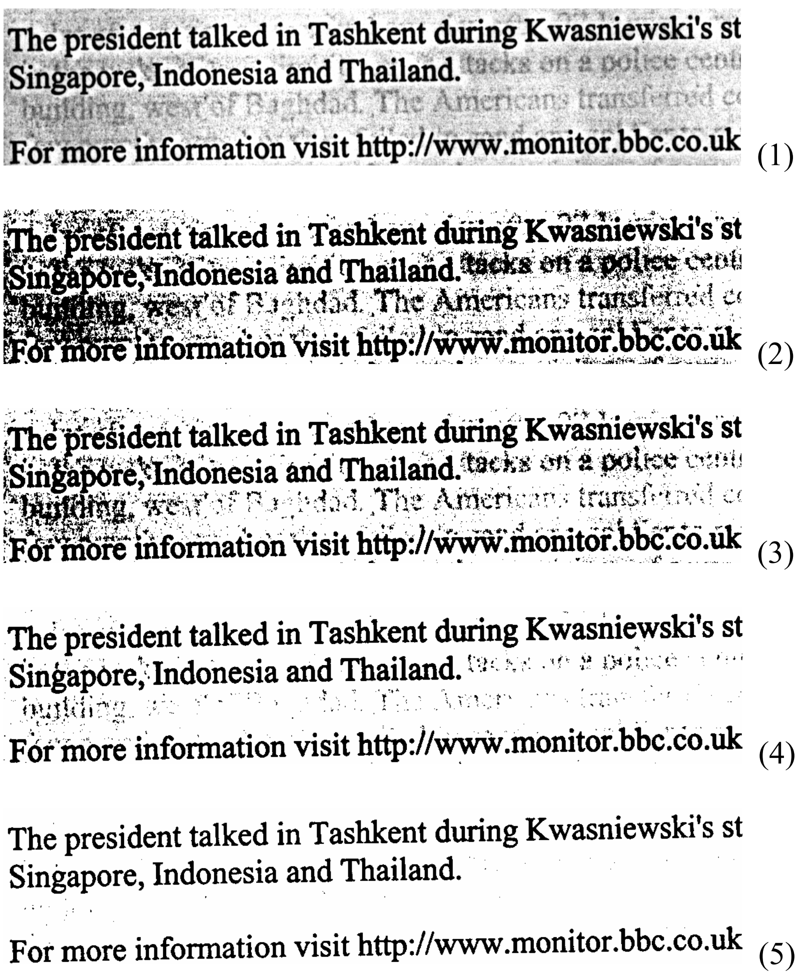

Figure 4. (1) original image, (2)-(5) binarization by Kapur's entropy method, Tsai's moment preservation, Otsu's method and our new method respectively.