A WAVELET APPROACH TO DOUBLE-SIDED DOCUMENT IMAGE PAIR PROCESSING

Ruini Cao, Chew Lim Tan, Peiyi Shen

School of Computing, National University of Singapore
3 Science Drive 2, Singapore 117543
Email: {caorn, tancl, shenpy}@comp.nus.edu.sg

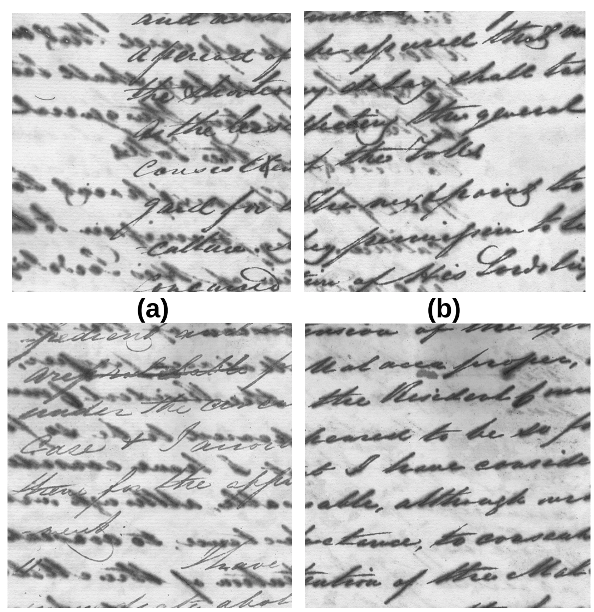
ABSTRACT

In this paper, we present a novel method for processing double-sided historic handwritten documents using wavelets. The method is specially designed to remove the interfering strokes from the reverse side due to ink sipping through pages after long periods of storage. The proposed method works by first matching both sides of a document page such that the interfering strokes are mapped with the originating strokes from the reverse side. This is to facilitate the identification of the foreground and interfering strokes. A wavelet reconstruction process then iteratively enhances the foreground strokes and smears the interfering strokes so as to strengthen the discriminating capability of an edge detector to extract clear foreground image. This is an improvement of our earlier work.

1. INTRODUCTION

Many archives keep a large number of historic handwritten documents. After long period of storage, the characters may sip through pages due to moisture etc. As a result the characters on one page or on one side the page may be smeared and interfere with the characters on the other page or on the other side of the same page. Four such sample images are shown in Figure 1. Given this problem, we have three classes of objects to separate here: foreground text, interfering strokes from the reverse side, and the background. Usually, the foreground writing appears darker than the interfering stroke. However, there are cases where the foreground and interfering strokes have similar intensities, or worst still, the interfering strokes are more prominent than the foreground.

Many segmentation and binary approaches have been reported in the literature [1][2]. Negishi et al. [3] presented several automatic thresholding algorithms to extract the character bodies from the noisy background in old manuscript. Liu and Srihari [5] presented a thresholding algorithm based on texture features to extract characters from the run-length featured texture background. Similar works can be seen in Liang and Ahmadi's algorithm [6] that adopts a mathematical morphological approach to extract text strings from regular periodic overlapping text/background images. White and Rohrer's [7] method is a thresholding technique based on boundary characteristics to suppress unwanted background patterns. Very similar work can be seen in Don's work [8] that segments the double-sided images based on the isolated gray-scale range of interfering images and noise characteristics. Lu et al.'s [9] method not only enhances the contrast of the edges in the low contrast area but also changes the intensity of the gray level of the edges. Lu also presented another wavelet method in decreasing the edge contrast and smearing the direct component of the edges with its neighboring pixels [10].



(a) (b)

(c) (d)

Figure 1. Sample images: sample 1, (a) the front, (b) the reverse side; sample 2, (c) the front, (d) the reverse side

The nature of our problem is the similar structures of the foreground and the interfering characters, and their overlapping ranges of gray-scale in histogram. It is hard to apply the above methods directly to solve our problem. In our earlier work [14], we adopted an edge detection algorithm based on the observation that the interfering strokes are not as sharp as the normal strokes. The orientation information of the edges is also employed to favor foreground strokes that are predominantly slanting at an angle [15]. However, there are still rooms for improvements as some strong interfering strokes remain on the resultant images. The intensity values of these interfering strokes are not reduced. This affects the final binarization results. A new technique using wavelet is now proposed to use the correspondence between the front image and the reverse side image to process both sides cooperatively. This is the major focus of this paper.

The remainder of the paper is organized as follows. In section 2, we explain the proposed method in detail. In section 3, we present some experiment results and discussions. Section 4 gives the conclusion and future work.

2. THE PROPOSED METHOD

2.1. Enhancement and Smearing Features

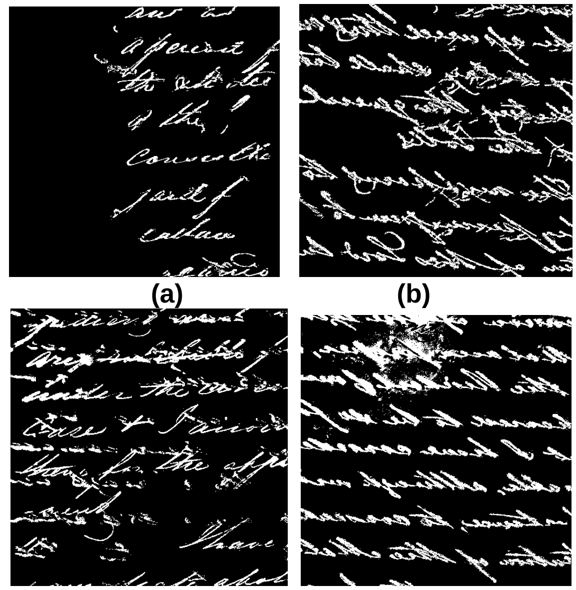
To improve the readability of the documents, we need to enhance the foreground strokes and meanwhile smear the interfering strokes from the reverse side. We apply the method described in [14, 15] to roughly detect the foreground strokes on both sides of the page. The foreground strokes on the front will be the enhancement feature for the image on the front. And the foreground strokes on the reverse side will be the smearing feature for the image on the front. To find the correspondence between the images on the front and the reverse side of the same page, we first flip the reverse side image horizontally to obtain a mirror image and then superimpose the front and mirrored reverse image by hand such as to map the corresponding strokes.

Let *f(x,y)* be the original image, *E(i,j)* be the enhancement feature image and *S(m,n)* be the smearing feature image. The three square images have the same dimension *N*N*. The enhancement and smearing features may be described as follows:

$$E(i,j) = \begin{cases} 0, background \\ 255, detected \quad stroke \end{cases}$$

$$S(i,j) = \begin{cases} 0, background \\ 255, detected \quad stroke \end{cases} \quad (1)$$

Figure 2 shows the detected enhancement and smearing features for the front images of sample1 and sample2.



(a) (b) (c) (d)

Figure 2. Enhancement and smearing features: (a) enhanced feature for the front of sample 1 (Figure 1a); (b) smeared feature for the front of sample 1 (Figure 1a); (c) enhanced feature for the front of sample 2 (Figure 1c); (d) smeared feature for the front of sample 2 (Figure 1c)

2.2. Iterative Wavelet Reconstruction

Let *f(x,y)* be the original image, the wavelet decomposition can be written as in [16],

$$\begin{cases} C_j f(m,n) = (< f(x,y), \Phi_{j,m,m}(x,y) >)_{(m,n)\in Z^2} \\ D_j^1 f(m,n) = (< f(x,y), \Psi^1_{j,m,m}(x,y) >)_{(m,n)\in Z^2} \\ D_j^2 f(m,n) = (< f(x,y), \Psi^2_{j,m,n}(x,y) >)_{(m,n)\in Z^2} \\ D_j^3 f(m,n) = (< f(x,y), \Psi^3_{j,m,n}(x,y) >)_{(m,n)\in Z^2} \end{cases} \quad (2)$$

A 10-scale wavelet decomposition of the original image *f(x,y)* may be described as follows:

$$Wf(x,y) = \{C_9 f(x,y), D_0^1 f(x,y), D_0^2 f(x,y), D_0^3 f(x,y), \ldots, D_9^1 f(x,y), D_9^2 f(x,y), D_9^3 f(x,y)\} \quad (3)$$

Given the image wavelet representation *Wf(x,y)*, the enhancement feature *E(i,j)* and the smearing feature *S(m,n)*, the iterative wavelet reconstruction may be described as follows:

Step 1. Multi-scale decomposition of the image on the front.

$$Wf(x,y) = \{C_9(x,y), D_j^k(x,y), j = 0,\ldots,9, k = 1,2,3\} \quad (4)$$

Step 2. The magnitudes of the wavelet coefficients in all the scales are revised by the enhancement and smearing coefficients. The following pseudo code shows the algorithm (cf. equation 5). The coefficients in equation (5), $\{e_j^k > 1,\ 1 > s_j^k > 0, j = 0,...,9, k = 1,2,3,\}$, are the enhancement and smearing coefficients that are set empirically. The enhanced/smeared image *f'(x,y)* is reconstructed from the modified coefficients (cf. equation 6).

Do {

if E(x,y)==255 $D_j^k(x,y) = e_j^k D_j^k(x,y)$;

if S(x,y)==255 $D_j^k(x,y) = s_j^k D_j^k(x,y)$;

} while(j=0, j<=9, j++; k=1, k<=3, k++; x=0, x<N, x++; y=0, y<N, y++)

(5)

$f'(x,y)$ = *inverse wavelet transform(*

$\{\{C_9(x,y), D_j^k(x,y), j = 0,...,9, k = 1,2,3\}\})$ (6)

Step 3. Perform the wavelet transform again on the reconstructed image *f'(x,y)* (cf. equation 7). Note that in the inverse wavelet transform in equation (8), the revised D_j^k obtained in equation (5) is used.

$$Wf'(x,y) = \{C_9'(x,y), D_j'^k(x,y), j = 0,...,9, k = 1,2,3\} \quad (7)$$

$f'(x,y)$ = *inverse wavelet transform(*

$\{C_9'(x,y), D_j^k(x,y), j = 0,...,9, k = 1,2,3\})$ (8)

Step 4. Iteratively process the wavelet decomposition and reconstruction using equation (7) and (8). we would get the final enhanced/smeared gray-scale image.

Step 5. Rescale the enhanced/smeared image using the following function:

$$f'(x,y) = \begin{cases} 0, & f'(x,y) <= 0 \\ 255, & f'(x,y) >= 255 \\ & f'(x,y) \end{cases} \quad (9)$$

In our implementation, we used wavelet transform up to 10 scales for images of size 512x512. In the reconstruction process, we set the maximum number of iterations to 15.

Figure 3(a) and (b) show the results after 15 iterations. After enhancement and smearing, the resultant images are further processed by our text extraction method [14][15] to pick up the foreground strokes as shown in Figure 3(c) and (d). The Niblack's method [17] is finally applied to binarize the images to produce clear readable copies as shown in Figure 3(e) and (f).

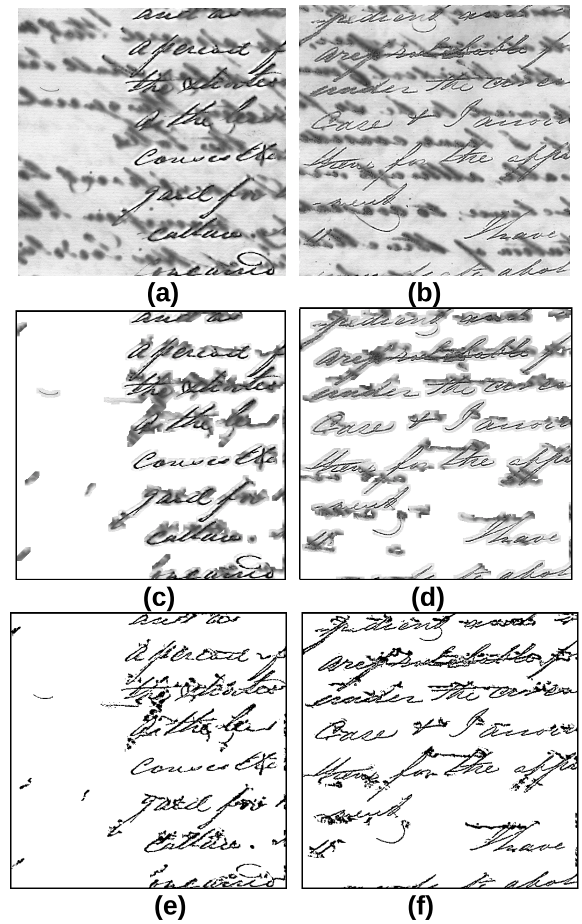


Figure 3. Iterative wavelet reconstruction: (a) enhanced/smeared image of the front of sample 1 (Figure 1a); (b) enhanced/smeared image of the front of sample 2 (Figure 1c); (c) the segmentation result of (a); (d) the segmentation result of (b); (e) the binarized image of (c); (f) the binarized image of (d)

3. EXPERIMENTAL RESULTS

The performance of our approach has been evaluated on scanned images of historical handwritten documents provided by the National Archives Board. The images were scanned at 150 dpi and saved as TIF format without compression. Since the final purpose of our process is for human viewers, the cleaned-up images were visually inspected to assess the readability of the extracted words.

We selected 12 severely interfering images for testing. The well-known IR standard measures, *precision* and *recall*, are used to measure the performance of the proposed method. Precision and Recall are defined as follows:

Precision = # of correctly detected words / # of all words detected by the system,

Recall = # of correctly detected words / total # of words present in the document.

Table 1 shows the evaluation results of the 12 images. Precision reflects the performance of removing the interfering strokes of the system and recall reflects the performance of restoring the foreground words of the system. The average precision and recall for the 12 testing images were 84% and 96% respectively. With this improvement, we found that the readability improves significantly due to the effect of enhancement of the foreground strokes and smearing of the interfering strokes.

Table 1. Evaluation of the system in 12 testing images

Image number	1	2	3	4	5	6	7
Total no. of words	132	124	103	125	125	123	121
Precision	91%	86%	76%	94%	92%	80%	79%
Recall	98%	100%	90%	94%	98%	94%	89%
Image number	8	9	10	11	12		Average
Total no. of words	128	112	113	114	114		
Precision	75%	84%	82%	91%	78%		84%
Recall	97%	97%	98%	96%	96%		96%

4. THE CONCLUSION AND FUTURE WORK

This paper describes a method for removal of interfering strokes from historic handwritten documents. The method superimposes a matching between both sides of the same page and employs wavelets to enhance the foreground strokes and meanwhile smear the reverse-side strokes. The present manual matching may not result in perfect matching due to document skews, different scales during image capture, and warped surfaces at books' spine areas. Our future work is to find an automated way for matching the front and reverse side image.

5. REFERENCES

[1] G. Nagy, "Twenty Years of Document Image Analysis in PAMI," *IEEE Trans. on PAMI*, vol. 22, no. 1, pp38-62, Jan. 2000

[2] R.G. Casey and E. Lecolinet, "A Survey of Methods and Strategies in Character Segmentation," *IEEE Trans. on PAMI*, vol.20, no.7, pp.690-706, July 1996

[3] H. Negishi, J. Kato, H. Hase, and T. Watanabe, "Character Extraction from Noisy Background for an Automatic Reference System," *Proc. of the 5th Int. Conf. on Document Analysis and Recognition*, Bangalore, India, pp.143-146, 1999

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

[5] Y. Liu, and S.N. Srihari, "Document Image Binarization Based on Texture Features," *IEEE Trans. on PAMI*, vol. 19, no. 5, pp.540-544, May 1997

[6] S. Liang, and M. Ahmadi, "A Morphological Approach to Text String Extraction from Regular Periodic Overlapping Text/Background Images," *Graphical Models and Image Processing, CVGIP*, vol. 56, no. 5, pp 402-413, Sep. 1994

[7] J.M. White, and G.D. Rohrer, "Image Thresholding for Optical Character Recognition and Other Applications Requiring Character Image Extraction," *IBM J. Res. Dev.* 27(4), pp.400-410, 1983

[8] H. Don, "A Noise Attribute Thresholding Method for Document Image Binarization," Proc. of the 3rd Int. Conf. on Document Analysis and Recognition, Canada, pp231-234, 1995

[9] J. Lu, D.M. Healy, and J.B. Weaver, "Contrast Enhancement of Medical Images Using Multiscale Edge Representation," *Optical Engineering*, 33(7), pp.2151-2161, 1994

[10] J. Lu, "Image Deblocking via Mulitscale Edge Processing," *Proc. of SPIE: Wavelet Applications in Signal and Image Processing IV*, vol. 2828, Part two, M.A. Unser, A. Aldroubi, A.F. Laine, (Eds), Denver, Colorado, pp.742-751, 6-9 Aug., 1996

[11] S. Mallat and S. Zhong, "Characterization of Signals from Mulitscale Edges," *IEEE Trans. on PAMI*, vol.14, no.7, pp.710-732, July 1992

[12] W.L. Hwang, and F. Chang, "Character Extraction from Documents Using Wavelet Maxima," *Proc. of SPIE: Wavelet Applications in Signal and Image Processing IV*, vol. 2828, Part two, M.A. Unser, A. Aldroubi, A.F. Laine, (Eds), Denver, Colorado, pp.1003-1015, 6-9 Aug. 1996

[13] K. Etemad, D. Doerman, and R. Chellappa, "Multiscale Segmentation of Unstructured Document Pages Using Soft Decision Integration," *IEEE Trans. on PAMI*, vol. 19, no. 1, pp92-96, Jan. 1997

[14] R. Cao, C. L. Tan, Q. Wang, and P. Shen, "Segmentation and Analysis of Double-Sided Handwritten Archival Documents," *Prof. Of the 4th IAPR Int. Workshop on Document Analysis Systems*, Rio de Janeiro, Brazil, Dec. 10-13, 2000

[15] C. L. Tan, R. Cao, P. Shen, Q. Wang, J. Chee, and J. Chang, "Removal of Interfering Strokes in Double-Sided Document Images," *Proc. of the 5th IEEE Workshop on Applications of Computer Vision*, Palm Springs, California, Dec. 4-6, 2000

[16] L. Feng, Y.Y. Tang, L.H. Yang, "A Wavelet Approach to Extracting Contours of Document Images," *Proc. of the 5th Int. Conf. on Document Analysis and Recognition*, Bangalore, India, pp71-74, 1999

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