World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:2, No:6, 2008

Restoration of Noisy Document Images with an Efficient Bi-Level Adaptive Thresholding

Abhijit Mitra

Abstract—An effective approach for extracting document images from a noisy background is introduced. The entire scheme is divided into three substechniques – the initial preprocessing operations for noise cluster tightening, introduction of a new thresholding method by maximizing the ratio of standard deviations of the combined effect on the image to the sum of weighted classes and finally the image restoration phase by image binarization utilizing the proposed optimum threshold level. The proposed method is found to be efficient compared to the existing schemes in terms of computational complexity as well as speed with better noise rejection.

Keywords— Document image extraction, Preprocessing, Ratio of standard deviations, Bi-level adaptive thresholding.

I. Introduction

ELIMINATION of background noise in order to achieve a clean, perturbation-free image is very useful not only in the field of communication and image processing, but also in certain pattern recognition applications like restoration of valuable historical documents as well as different personal cognition schemes to meet the proliferating demands of crime science over the last few decades. In this paper, an effective restoration approach for different kinds of document images is proposed with an efficient thresholding scheme. While most of the works reported in the literature have preferred direct thresholding [1]-[3] on the test image, we have started with preprocessing [4] operations before applying the proposed thresholding method, which have exhibited better results than the existing schemes. Thresholding, in fact, is the most crucial technique for image segmentation which tries to identify and extract a target image from its background on the basis of the distribution of gray-level values in the same image object. Several thresholding methods have been reported [5]-[8] with different approaches for various kinds of images. Locating the suitable threshold among these methods can be broadly classified into two categories – parametric and non-parametric. In parametric approaches, the gray-level distribution of a particular object class leads to find the threshold level. For instance, in Wang and Haralick's study [7], the image pixels are classified into edge and non-edge pixel classes, followed by a bright and dark subclasses in the edge pixel class. The highest peaks of two histograms of these two subclasses are then chosen as the thresholds. In non-parametric techniques, the optimum thresholds are selected with the fulfillment of some criterion. As an example, Otsu's celebrated paper [8] chooses the optimum threshold by maximizing the between-class variance with an exhaustive search, while, Kittler and Illingworth's work [5] assumes two separate normal distributions and minimizes the error. In another study, Kapur *et al.* [6] have found the threshold by maximizing the entropy of the gray-level histograms of resulting classes. However, our proposed non-parametric scheme has assumed two different overlapping pixel clusters, one for background noise and the other for the image. Our goal is to restore the image cluster by, first, preprocessing the image to tighten the cluster distributions and then, searching further for the optimum bi-level threshold with a fast algorithm which basically maximizes the ratio of standard deviations of the combined weightage of these two clusters on the image to the sum of weighted classes. In other words, the ratio of combined weighted standard deviation of the image and weighted sum of standard deviations of these clusters shows a global maxima that we can choose as our required threshold point. Finally, a binarization method has been used to restore the document image. The organization of the paper is as follows: Section 3 deals with the primary approach for noise elimination with primary background noise reduction and then averaging while Section 4 discusses about the proposed low-complexity adaptive thresholding scheme. The final image extraction technique is dealt with in Section 5, which directly follows the effectiveness of proposed threshold. In Section 6, some experimental results are given considering three different kinds of noise and conclusions are drawn in Section 7 based on the above discussions and results.

II. Image Database

The document image is dealt with as a gray image. All the sample images are scanned within a limited space of 256×512 pixels and are digitized to matrix $\mathbf{A} = [a]_{ij}$, with each a_{ij} having one of the values from 2^8 gray levels. Usually in any standard MATLAB program, the lowest gray level value is assigned to black pixel and the highest value to the white pixel, which is just the opposite of our arithmetical convention that has been followed here. Hence, the image matrix must be complemented before starting and after finishing the proposed image extraction scheme. The algorithm gives satisfactory results with higher number of gray level bits as well. In the sequel, the matrices and vectors would be denoted with bold uppercase and bold lowercase letters respectively.

Manuscript received July 7, 2005.

A. Mitra is with the Department of Electronics and Communication Engineering, Indian Institute of Technology (IIT) Guwahati, North Guwahati - 781039, India. E-mail: a.mitra@iitg.ernet.in.

III. Primary Approach for Noise Elimination

A. Background Noise Reduction

The initial background noise reduction phase is a combinational process of equalization and then subtraction from the original values. Equalization can be done either by row or by column uniforming process [9] with no loss of generality. Here, column equalization has been used and noise is reduced by the equation

$$\overline{a}_{ij} = a_{ij} - \langle \mathbf{e_q}, \mathbf{a_j} \rangle \tag{1}$$

where a_{ij} is the (i, j)th pixel of the original image martix with the range $1 \leq i \leq M$ and $1 \leq j \leq N$, $\overline{a}_{ij}$ denotes the same of the reduced matrix $\overline{\mathbf{A}}$, '$\langle . \rangle$' is the inner-product notation with the M element qualization vector

$$\mathbf{e_q} = [\frac{1}{M}\frac{1}{M}\cdots\frac{1}{M}]^T \tag{2}$$

and $\mathbf{a_j}$ being the jth column vector of the image matrix $\mathbf{A}$ with size $M \times N$. The reduction operation can alternately be expressed in matrix notation as

$$\overline{\mathbf{A}} = \mathbf{A} - \mathbf{\Psi} \tag{3}$$

where

$$\mathbf{\Psi} = [\psi_\mathbf{1}\ \psi_\mathbf{2}\ \cdots\ \psi_\mathbf{N}] \tag{4}$$

with $\psi_\mathbf{j} = \langle \mathbf{e_q}, \mathbf{a_j} \rangle [1\ 1\ \cdots\ 1]^T$, $j = 1, 2, ..., N$.

B. Background Noise Clipping

The above noise reduction technique leads to convert most of the low density noise pixels to non-positive values which should be eliminated as a first procedure for achieving a clean image. This can be achieved by clipping the reduced image according to the following equation

$$\tilde{a}_{ij} = \begin{cases} \overline{a}_{ij} & \forall\, \overline{a}_{ij} \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $\tilde{a}_{ij}$ is the reduced image pixel after clipping.

C. Noise Averaging

In this phase, we first take a 1-definable square lattice [11] with $n = 2$, i.e., 2D Cartesian space, with direction vectors $\{(1, 0), (-1, 0), (0, 1), (0, -1)\}$ and another 1-definable square lattice with a new set of direction vectors $\{(1, 1), (-1, 1), (1, -1), (-1, -1)\}$. A test area is set by the superposition of these two square lattices to get a 2-definable square lattice. Then the clipped pixel cluster distributions are further tightened by the relation

$$\hat{a}_{ij} = \mathbf{w^T A' w} \tag{6}$$

where $\hat{a}_{ij}$ is the averaged image pixel, the averaging vector $\mathbf{w} = [\frac{1}{3}\frac{1}{3}\frac{1}{3}]^T$ and

$$\mathbf{A'} = [\tilde{a}]_{kl} = \begin{pmatrix} \tilde{a}_{k-1,l-1} & \tilde{a}_{k-1,l} & \tilde{a}_{k+1,l} \\ \tilde{a}_{k,l-1} & \tilde{a}_{k,l} & \tilde{a}_{k,l+1} \\ \tilde{a}_{k+1,l-1} & \tilde{a}_{k+1,l} & \tilde{a}_{k+1,l+1} \end{pmatrix} \tag{7}$$

is the *lattice-superposition matrix* of the test pixel $\tilde{a}_{kl}$. In general, a test area for averaging purpose is set by taking into acccount the x and y adjacent pixels to the test pixel $\tilde{a}_{kl}$ in both directions, i.e., the size of the test area is $(2x + 1) \times (2y + 1)$. In our case, the lattice-superposition method can be viewed as the *nearest-neighbourhood* approach around the test pixels to select only the eight juxtaposed pixels to the respective $\tilde{a}_{kl}$s and the averaging area thus becomes 3×3. This of course leads towards a better result with satisfactory corner noise reduction. The lattice-superposition process along with two specific cases have been shown in Fig. 1 and 2. In Fig. 1, the lattice-superposition process is shown to select the eight direct neighbour pixels of a test pixel. The process sets the size of test area as 3×3, while Fig. 2 shows two specific cases where the test pixels are situated in two corner positions. This figure helps to understand the logic of isolated corner noise reduction graphically for the presence of less number of neighbour pixels. After this phase, an almost separated image pixel cluster can be expected from the background.

IV. Proposed Thresholding Scheme

After completing the preprocessing phase, the problem of having two clusters of gray pixels still sustains, one with a specific range of values while the other with another and these two spans have the possibility of being overlapped. Our goal is to find an optimum threshold level so that the ratio of combined weighted standard deviation of the image and weighted sum of standard deviations of these clusters is maximum. In our proposed thresholding scheme, we follow the initial steps of Otsu [8] to find out the mean and variance of the whole image. Then a weighted sum of standard deviations is defined, with the help of which the maxima is calculated by a fast formula.

The noisy image is initially considered as a 2D gray scale intensity function which contains N number of pixels with L gray levels starting from 0 upto $L - 1$ (in our case, $L = 256$). If the total number of pixels with any gray level i is n_i, then the probability of gray level i would be

$$p_i = \frac{n_i}{N} \tag{8}$$

where $\forall i \in \mathbb{Z}$,

$$0 \leq p_i \leq 1 \qquad \text{and} \qquad \sum_{i=0}^{L-1} p_i = 1 \tag{9}$$

with $\mathbb{Z}$ being the integer set $\{x | x = 0, 1, \cdots, L - 1\}$. The mean and the standard deviation of this image can be denoted as μ_r and σ_r respectively.

For bi-level thresholding of a documentation image, the intensity pixels are divided into two classes C_1 and C_2 with gray level sets $\mathbb{Z}_1$ and $\mathbb{Z}_2$ respectively so that $\mathbb{Z}$=$\mathbb{Z}_1 \cup \mathbb{Z}_2$ and $\mathbb{Z}_1 \cap \mathbb{Z}_2$= ϕ, i.e., we set $\mathbb{Z}_1 = \{x_1 | x_1 = 0, 1, \cdots, t - 1\}$ and $\mathbb{Z}_2 = \{x_2 | x_2 = t, t + 1, \cdots, L - 1\}$ respectively. With such

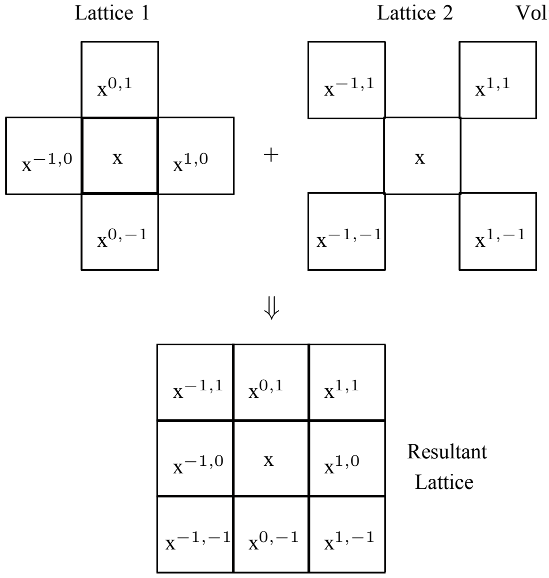


Fig. 1 The lattice-superposition process to form the test area for any test pixel x. The superscripts on different x denote the direction vectors associated with those particular pixels.

definitions, the gray level probability distributions for these two classes can be written as

$$\phi_1(i) = \frac{p_i}{\lambda_1} \qquad \forall i \in \mathbb{Z}_1 \tag{10}$$

and

$$\phi_2(i) = \frac{p_i}{\lambda_2} \qquad \forall i \in \mathbb{Z}_2 \tag{11}$$

where

$$\lambda_1 = \sum_{i=0}^{t-1} p_i \quad \text{and} \quad \lambda_2 = \sum_{i=t}^{L-1} p_i \tag{12}$$

with the mutual relation

$$\lambda_1 + \lambda_2 = \sum_{i=0}^{L-1} p_i = 1. \tag{13}$$

From the above equations, we can write

$$\lambda_1\mu_1 + \lambda_2\mu_2 = \mu_r \tag{14}$$

where

$$\mu_1 = \sum_{i=0}^{t-1} i\phi_1(i) \tag{15}$$

and

$$\mu_2 = \sum_{i=t}^{L-1} i\phi_2(i) \tag{16}$$

are the means of the classes C_1 and C_2 respectively. At this point, let us define a *weighted-class sum of standard deviations* (WCSSD) as

$$\sigma_{wc} = \langle \mathbf{w}, \mathbf{s} \rangle \tag{17}$$

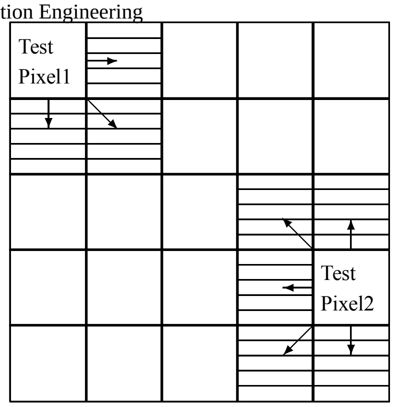


Fig. 2 The neighbour pixels (shown by the shaded area) of two kinds of terminal test pixels. Test Pixel1 has three neighbour pixels while Test Pixel2 has five.

with $\mathbf{w} = [\lambda_1\ \lambda_2]^T$ and $\mathbf{s} = [\sigma_1\ \sigma_2]^T$ where

$$\sigma_1^2 = \sum_{i=0}^{t-1} (i - \mu_1)^2 \phi_1(i) \tag{18}$$

and

$$\sigma_2^2 = \sum_{i=t}^{L-1} (i - \mu_2)^2 \phi_2(i) \tag{19}$$

denote the variances of C_1 and C_2 respectively. Then, if $\lambda_1\lambda_2\sigma_r$ represents the combined weightage of these two classes on the standard deviation of the image where $\sigma_r^2 = \sum_{i=0}^{L-1}(i - \mu_r)^2 p_i$, the ratio of standard deviation (RSD) of the combined effect and the weighted class comes as

$$RSD = \frac{\sigma_{comb}}{\sigma_{wc}} = \frac{\lambda_1\lambda_2\sigma_r}{\lambda_1\sigma_1 + \lambda_2\sigma_2}. \tag{20}$$

Our goal is then to select the optimum threshold (t_{opt}) so that the above equation gives the optimum result, i.e., that maximizes the RSD. In other words, we have to find out an adaptive t_{opt} such that

$$\begin{aligned} t_{opt} &= Arg_{0 \le t \le L-1}\ Max\{\frac{\lambda_1\lambda_2\sigma_r}{\lambda_1\sigma_1 + \lambda_2\sigma_2}\} \\ &= Arg_{0 \le t \le L-1}\ Min\{\frac{\sigma_1}{\lambda_2} + \frac{\sigma_2}{\lambda_1}\} \end{aligned} \tag{21}$$

as σ_r is a fixed quantity. This searching algorithm of eq. (21) has exhibited better thresholding results with less number of operations compared to most of the existing methods. The above algorithm takes a considerable speed up over Otsu's method as the proposed scheme has the complexity of $O(n)$ in comparison with $O(n^2)$ of the latter one. For such a choice of threshold, the overall probability of error, i.e., the probability of erroniously classifying the background as an object point as well as classifying an object point as background, comes as

$$E(t_{opt}) = \lambda_1 \sum_{i=t_{opt}}^{L-1} \phi_1(i) + \lambda_2 \sum_{i=0}^{t_{opt}-1} \phi_2(i) \qquad (22)$$

which will vary depending upon the nature of background noise in any test image.

V. Image Extraction Approach

A. Binarization

After finding the optimum threshold level, the averaged image is binarized according to the equation

$$a_{ij}^{bin} = \begin{cases} 1 & \text{if } \hat{a}_{ij} \geq t_{opt} \\ 0 & \text{otherwise} \end{cases} \qquad (23)$$

where a_{ij}^{bin} represents the binarized image pixel. A good quality texture body of the image is extracted after this phase.

B. Original Image Restoration

This phase extracts the original gray density information of the document image using the following equation

$$a_{ij}^{*} = \begin{cases} a_{ij} & \text{if } a_{ij}^{bin} = 1 \\ 0 & \text{otherwise} \end{cases} \qquad (24)$$

where a_{ij}^{*} is the restored image pixel. This information is needed in some writer identification applications.

VI. Experimental Results and Discussions

Several old and withered documents were scanned within a limited space to be used as iamges and few fresh signatures and documents were also taken and added up with different kinds of noise. These types include:
(i) **Gaussian Noise**: with zero mean and different variances σ^2, i.e., $N(0, \sigma^2)$,
(ii) **Salt and Pepper Noise**: which means *on* and *off* pixels added to the image with noise density ρ that affects approximately $\rho.M.N$ pixels of the image, and,
(iii) **Speckle Noise**: which is a multiplicative noise, used in the floating-point arithmetic as the relative roundoff error. We have used the variances 0.04 (default), 0.06 and 0.08.

Our proposed noise elimination scheme has exhibited satisfactory results for almost all the cases. Note that the proposed thresholding, when applied without preprocessing the test image, yielded a different threshold value (t'_{opt}) which couldn't remove almost all the noise pixels in the final results in most of the cases.

Fig. 3 to Fig. 7 are given in the following to exhibit the effectiveness of our proposed technique considering all the three types of above mentioned noise as well as the importance of preprocessing the document before applying the proposed threshold method to the image. Among these figures, Fig. 3(a) and Fig. 4(a) are scanned old images contaminated with unknown noise and Figs. 5(a), 6(a) and 7(a) have been made by applying those three kinds of noise sequences to clean test images. For these figures, i.e., Figs. 5, 6 and 7, the overall probability of errors after applying the proposed scheme have been computed with the knowledge of background noise and are given in Table 1. Table 2 enlists the values of t_{opt} and t'_{opt} for the same figures. These two tables show the effectiveness of the method along with preprocessing technique. The proposed bi-level thresholding scheme has also given satisfactory results compared to the other existing methods in terms of computational counts.

TABLE I
PROBABILITY OF ERROR VALUES FOR DIFFERENT DOCUMENT IMAGES UNDER TEST.

Figure number	Probability of error ($E(t_{opt})$)
Fig. 5	0.0001
Fig. 6	0.0012
Fig. 7	0.0047

TABLE II
THRESHOLD VALUE WITH PREPROCESSING (t_{opt}) AND WITHOUT PREPROCESSING (t'_{opt}) FOR DIFFERENT TEST IMAGES.

Figure number	t_{opt}	t'_{opt}
Fig. 5	107	208
Fig. 6	123	205
Fig. 7	114	193

VII. Conclusions

An efficient adaptive thresholding method with reduced operational complexity has been proposed in this paper which can be treated as a better alternative compared to the existing techniques. The results have shown the satisfactory performance of bi-level thresholding for any kind of document image contaminated in different kinds of noise environments. It has also been shown that preprocessing the image led to better results instead of applying the direct threshold scheme on the image. Extending such a simplified computational approach to multidimensional thresholding can serve as a good topic for future research.

References

[1] Ping-Sung Liao *et al.*, "A Fast Algorithm for Multilevel Thresholding," *Journal Info. Sc., Engg. 17*, pp. 713-727, 2001.

[2] Y. Yang and H. Yan, "An Adaptive Logical Method for Binarization of Degraded Document Images," *Pattern Recognition*, vol. 33, no. 5, pp. 787-807, May 2000.

[3] S. Dizenzo *et al.*, "Image Thresholding using Fuzzy Entropies," *IEEE Trans. Systems, Man, Cybern.*, vol. 28, no. 1, pp. 15-23, Jan. 1998.

[4] A. Mitra, Signature Extraction from a Noisy Environment and Signature Verification using Pressure Features, M. E. Tel. E. Dissertation, Jadavpur University, India, Feb. 1999.

[5] J. Kittler and J. Illingworth, "Minimum Error Thresholding," *Pattern Recognition*, vol. 19, no. 1, pp. 41-47, Jan. 1986.

[6] J. N. Kapur *et al.*, "A New Method for Gray-Level Picture Thresholding using the Entropy of the Histogram," *Computer Vision Graph. Image Proc.*, vol. 29, pp. 273-285, 1985.

[7] S. Wang and R. Haralick, "Automatic Threshold Selection," *Computer Vision Graph. Image Proc.*, vol. 25, pp. 46-67, 1984.

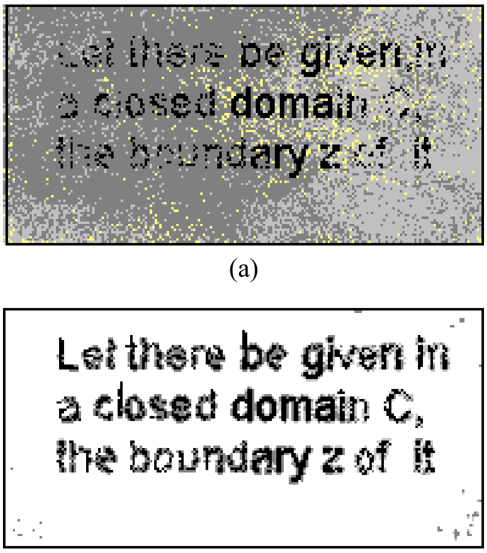
[8] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. Systems, Man, Cybern.*, vol. 9, no. 1, pp. 62-66, Jan. 1979.

[9] M. Ammar *et al.*, "A New Effective Approach for Automatic Off-line Verification of Signatures by using Pressure Features," in *Proc. 8th Int. Conf. Pattern Recognition (ICPR)*, Paris 1986, pp. 566-569.

[10] A. D. Brink, "Thresholding of digital images using two dimensional entropies," *Pattern Recognition*, vol. 25, no. 8, pp. 803-808, 1992.

[11] P. Meyer (2001, February). Lattice Geometries [Online]. Available: http://www.hermetic.ch/compsci/lattgeom.htm.

For a brief biography of the author, please see IJSP, vol. 2, no. 2, 2005, page 125.



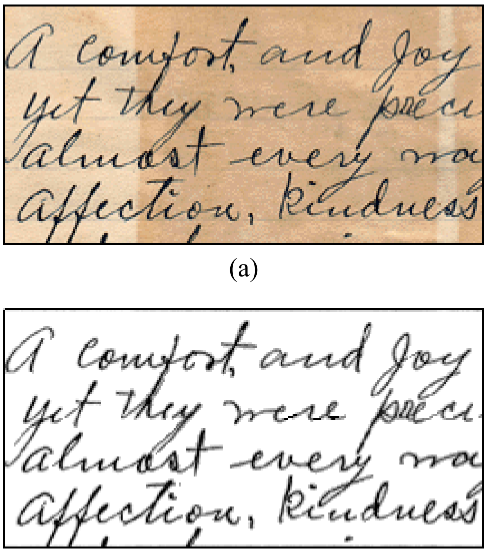
Let there be given in a closed domain C, the boundary z of it

(a)

Let there be given in a closed domain C, the boundary z of it

(b)

Fig. 3 (a) An old faded document image scanned within the specified limited space. The document can hardly be read. (b) The same image after applying the proposed thresholding technique along with preprocessing.



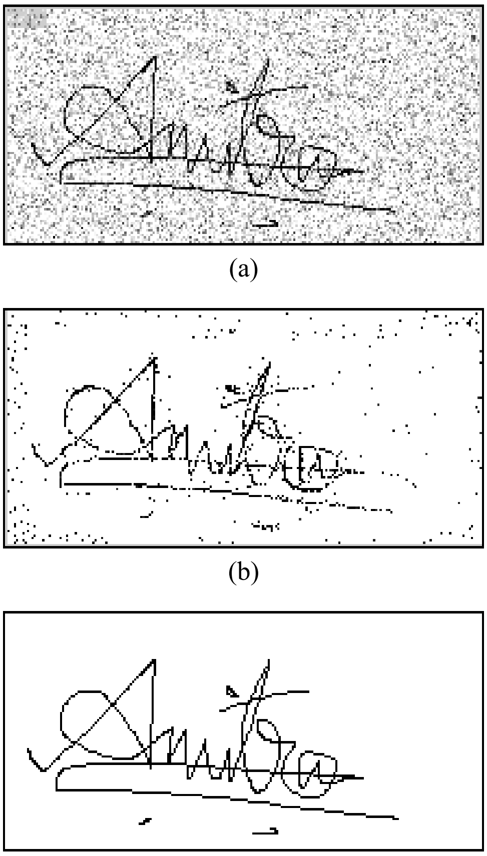
A comfort and Joy
yet they were preci
almost every ma
affection, kindness

(a)

A comfort and Joy
yet they were preci
almost every ma
affection, kindness

(b)

Fig. 4 (a) Another old handwritten document. (b) The recovered document after applying the same proposed scheme.



(a)

(b)

(c)

Fig. 5 (a) A sample signature image embedded in Gaussian noise $N(0, .16)$. (b) Recovered signature by applying the proposed thresholding without preprocessing operation. (c) Recovered clean signature by applying the proposed thresholding after preprocessing the same.

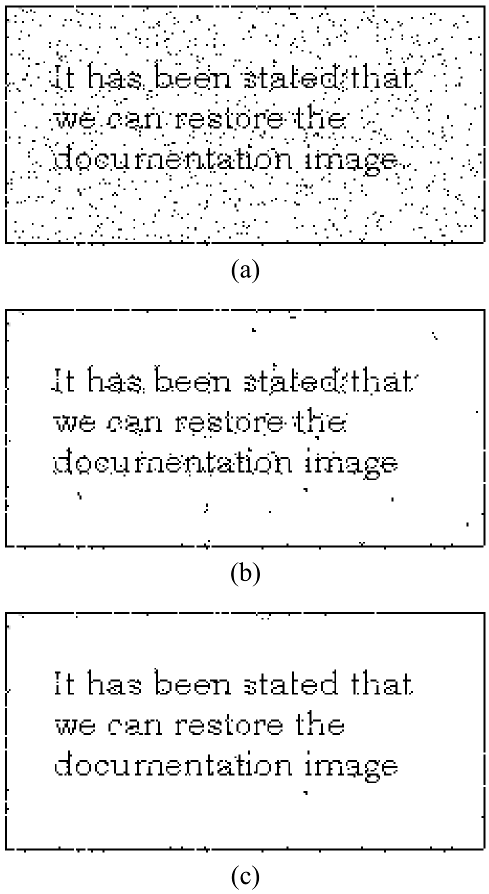

It has been stated that
we can restore the
documentation image
(a)
It has been stated that
we can restore the
documentation image
(b)
It has been stated that
we can restore the
documentation image
(c)

Fig. 6 (a) Another document image contaminated by ‘salt and pepper’ noise with density $\rho = 0.07$. (b) The same image with thresholding scheme only. (c) The recovered image after applying the proposed scheme on it along with preprocessing.

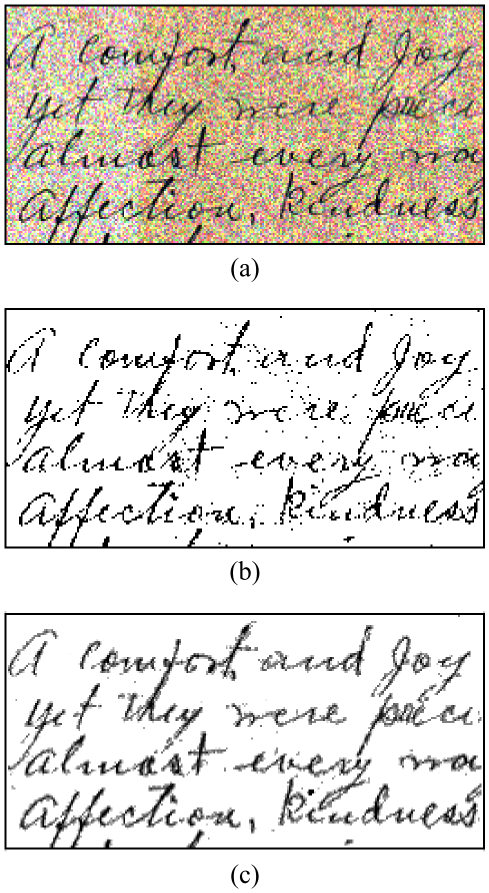

A comfort and Joy
yet they were prec
almost every ma
affection, kindness
(a)
(b)
(c)

Fig. 7 (a) The handwritten document of Fig. 4 contaminated in ‘Speckle Noise’ with variance 0.08. (b) Recovered binarized document by applying the proposed thresholding without preprocessing operation. (c) Recovered handwriting after applying the proposed thresholding with proper preprocessing operations.