NEAR EAST UNIVERSITY

GRADUATE SCHOOL OF APPLIED SCIENCES

MD THRESHOLDING : A NOVEL IMAGE BINARIZATION METHOD

Boran Şekeroğlu

Ph.D. Thesis

Department of Computer Engineering

Nicosia - 2007

ACKNOWLEDGEMENT

I would like to thank everyone who provided help and advice during the preparation of this thesis.

First, I would like to thank my supervisor Assoc. Prof. Dr. Adnan Khashman for his invaluable advice and belief in my work and myself over the course of this Ph.D. Research.

Second, I would like to express my gratitude to Near East University and Thesis Supervision Committee Members; Prof. Dr. Fahreddin M. Sadıkoğlu, Assoc. Prof. Dr. Rahib Abiyev and Assist. Prof. Dr. Hüseyin Sevay for their advice.

Third, I would like to thank my family for their constant encouragement, support and patience during the preparation of this thesis.

Finally, I would also like to thank my wife Süsen D. Şekeroğlu and my daughter Dilara Naz Şekeroğlu, for their existence.

Boran Şekeroğlu

ABSTRACT

Thresholding is an efficient method for the binarization of grayscale scanned documents, where the relationship between pixel values in the document images can provide an effective single point for the separation of the background and foreground layers. Document analysis and effective separation of text may provide useful data for electronic storage systems, digital libraries and human readers. Many thresholding-based image binarization methods have been developed and used for document enhancement. However, the efficiency of these methods can be impaired by the variation of gray levels in different documents, thus causing over-thresholding, under-thresholding or noise addition.

This thesis presents a novel global single-stage thresholding method that enhances document images by clearly separating background and foreground layers within these images. The method, which is called Mass-Difference (MD) Thresholding, finds an optimum thresholding value or exact separation point for each image using the relationship between luminance value and mean intensity of the image without considering peak values in the gray level histogram. The proposed MD method is implemented using a database that was especially collected and constructed to have different types of challenging document images; comprising 174 historical documents, specially created words and handwritten text.

MD thresholding method will be compared to 13 benchmark and/or recently developed global and local thresholding methods. The evaluation of the thresholding methods aims at determining an optimum thresholding method that can be efficiently applied to a variety of images such as scanned documents. Evaluation is performed using visual inspection and computed noise analysis; which uses three new PSNR-derived metric parameters.

Experimental results suggest that the developed MD method is superior in providing a fast and efficient text separation in document images.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

IN : Image Negatives
LT : Log Transformations
PLT : Power-Law Transformations
PLTF : Piecewise-Linear Transformation Functions
HE : Histogram Equalization
FT : Fourier Transform
DFT : Discrete Fourier Transform
ILPF : Ideal Low Pass Filters
BLPF : Butterworth Low Pass Filter
GLPF : Gaussian Low Pass Filter
IHPF : Ideal High Pass Filter
BHPF : Butterworth High Pass Filter
GHPF : Gaussian High Pass Filter
CT : Computed Tomography
MRI : Magnetic Resonance Image
FFT : Fast Fourier Transform
PDF : Probability Density Function
PAT : Pattern Averaging Thresholding
ALT : Adaptive Logical Thresholding
WFM : Water Flow Model
MD : Mass-Difference
PSNR : Peak Signal-to-Noise Ratio
APAR : Average PSNR Accuracy Rate
APD : Average PSNR Deviation
CPR : Combined Performance Rate
MSE : Mean-Squared Error
RW : Recognized Word
WP : White Paper
WBM : White Board Marker
YP : Yellow Envelope Paper

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Digitized document analysis has recently become more significant with the advances in digital archiving and electronic libraries. Scanned document images, especially historical and handwritten documents, generally carry various levels of noise because of the age, paper, pen and pencil influences on the documents. Age factor adds irremovable noise and meaningless random shapes on the documents which prevent efficient separation and recognition of the layers. Paper properties such as patterned or colored papers; add different background layers to the scanned documents. In addition, the variety of pens and pencils produces different and various foreground layers for the documents. Therefore, efficient binarization of scanned paper-based documents is usually required prior to further processing. The efficiency of document image binarization depends on the efficient separation and classification of background and foreground layers. Thus, the initial purpose of document analysis techniques is the effective preparation and separation of various layers in documents in order to provide sufficient and clear data for recognition systems and human readers.

One of the simplest methods that can be used to separate foreground and background layers is thresholding. This is based on the assumption that objects and background layers in the image can be distinguished by their gray level values. Thresholding methods can be categorized into two groups as Global Thresholding and Local (Adaptive) Thresholding. Global thresholding is a simple and efficient method where a defined or computed threshold value is used to separate foreground objects from background and Local (Adaptive) Thresholding is the assigning of a value to each pixel to determine whether it is a foreground or background pixel using local information from the image. Several thresholding methods that belong to these two groups have been developed.

With the existence of many global and local thresholding methods, deciding upon an optimum method for document image binarization is a challenging task; because the efficiency of the existing thresholding methods is usually application-dependent where one method's performance appears superior when using a certain type of document, but fails on a different type of document. The solution to this problem would be in creating and using a comprehensive multi-applications document image database that accounts

for different types of documents, such as historical documents, degraded documents, artificially created words, and handwritten documents.

This thesis presents a new global thresholding method named as Mass-Difference (MD) Thresholding. Additionally, a comprehensive comparative evaluation of MD and 13 known or recent thresholding methods that can be used for document image binarization is provided. The objectives of the work presented in within this thesis can be summarized as follows:

- Design and development of an efficient thresholding method for image binarization.
- Creating and using a comprehensive multi-applications document image database that includes historical documents, degraded documents, handwritten and artificially created words within bright and low-contrast and dark images.
- Providing a larger document image database with sufficient number of images.
- Implementing document image binarization using 14 thresholding methods, including the proposed method, (seven global methods and seven local methods). The implementation and experiments are to be carried out using the C-programming language. The considered thresholding methods comprise known and recent methods.
- Defining and implementing two evaluation and comparison criteria: visual inspection and computed noise analysis of binarized images.
- Comparing the performance of the 14 methods and determining an optimum thresholding method.

The thesis is organized as follows: Chapter 1 briefly describes the fundamentals of image enhancement. Chapter 2 reviews the basics of image binarization, conventional methods and recently proposed methods. Chapter 3 introduces the proposed method and preliminary experiments and comparisons. Chapter 4 presents the multi-application document image database, the evaluation procedure (which includes three new evaluation parameters) and the performed comparative evaluation. Finally, the work that is presented within this thesis is concluded.

CHAPTER 1
FUNDAMENTALS OF IMAGE ENHANCEMENT

1.1 Overview

Image enhancement is the process that intended to increase the visual appearance of digital images, graphics or photographs. Consequently, the enhancement methods are application-specific and are often developed empirically [1]. Thus, method that is optimum for enhancing X-ray images may not necessarily the optimum for enhancing pictures of Mars transmitted by a space probe [2].

In this chapter, definitions of image enhancement, its techniques and application areas of these techniques will be explained. In addition, advantages and disadvantages of these techniques will be listed.

1.2 Image Enhancement Approaches

Image Enhancement approaches can be divided into two categories: spatial domain methods and frequency domain methods. Spatial domain is the normal image space and frequency domain is the continuous signal of an image. Basic difference between these two approaches is the processing way of enhancement techniques. In spatial domain approach, techniques are based on direct manipulation of pixels and in frequency domain; techniques are based on the modification of Fourier Transform [2].

1.2.1 Overview of Spatial Domain Image Enhancement Techniques

Spatial domain image enhancement techniques operate on pixels in image space and the processes are denoted as [2]:

$$g(x,y) = T[f(x,y)] \tag{1.1}$$

where *f(x,y)* is the input image, *g(x,y)* is the processed image, and *T* is an operator on *f*, defined over some neighborhood of *(x,y)*. So, *grayscale* (also called intensity and mapping [2]) *transformation function* can be obtained by determining neighborhood size *T* as *1x1*. Thus, in single pixel neighborhood, *T* becomes grayscale transformation function where *g* depends only on value of *f* at *(x,y)*. This form can be re-written as:

$$s = T(r) \tag{1.2}$$

where *s* and *r* are variables denoting, respectively the gray level of *f(x,y)* and *g(x,y)* at any point of *(x,y)* [2].

1.2.1.1 Basic Gray Level Transformations in Spatial Domain

Several transformation functions and techniques had been developed by modifying grayscale transformation function such as Image Negatives (IN), Log Transformations (LT), Power-Law Transformations (PLT) and Piecewise-Linear Transformation Functions (PLTF).

Image Negatives is used to obtain photographic negative of an image by applying the negative transformation which is given in Equation 1.3.

$$s = L - 1 - r \tag{1.3}$$

where L is the gray-level range of image which is defined as $[0,L-1]$.

Log Transformations is used to expand dark pixels while compressing higher value pixels in image. The general form can be seen in Equation 1.4.

$$s = c\log(1 + r) \tag{1.4}$$

where c is a constant. For specific applications, it is also possible to use inverse log transformation to expand higher value pixels while compressing dark pixels.

Power-Law transformations which can be seen in Equation 1.5, provide more flexible transformation curve than LT according to the value of c and γ. If $\gamma<1$, PLT produces expanded dark pixels while producing compressed higher value pixels, and in other case, if $\gamma>1$ it produces expanded higher value pixels while produces compressed dark pixels. Identity transformation is obtained if $\gamma=1$ (Note that $c=1$ for all cases).

$$s = cr^{\gamma} \tag{1.5}$$

where c and γ are positive constants.

Piecewise-Linear Transformation consists several functions such as Contrast stretching, Gray-Level slicing and bit-plane slicing which are used for image enhancement.

Contrast stretching is one of the simplest and most important piecewise linear transformation. During image acquisition, images may have low-contrast because of poor illumination. The idea of contrast stretching is to increase the dynamic range of the gray levels in the image being processed [2] and the typical formula can be seen in Equation 1.6 [3][4].

$$s = (r - c)\left(\frac{b - a}{d - c}\right) + a \tag{1.6}$$

where, s and r denotes output and input images respectively, a and b denotes lower and upper limits of image respectively (between 0 and 255 in 8 bit grayscale image) and c and d represents the lowest and highest pixel values in an image. Figure 1.1 shows the implementation of IN, LT, PLT and Figure 1.2 shows Contrast Stretching.

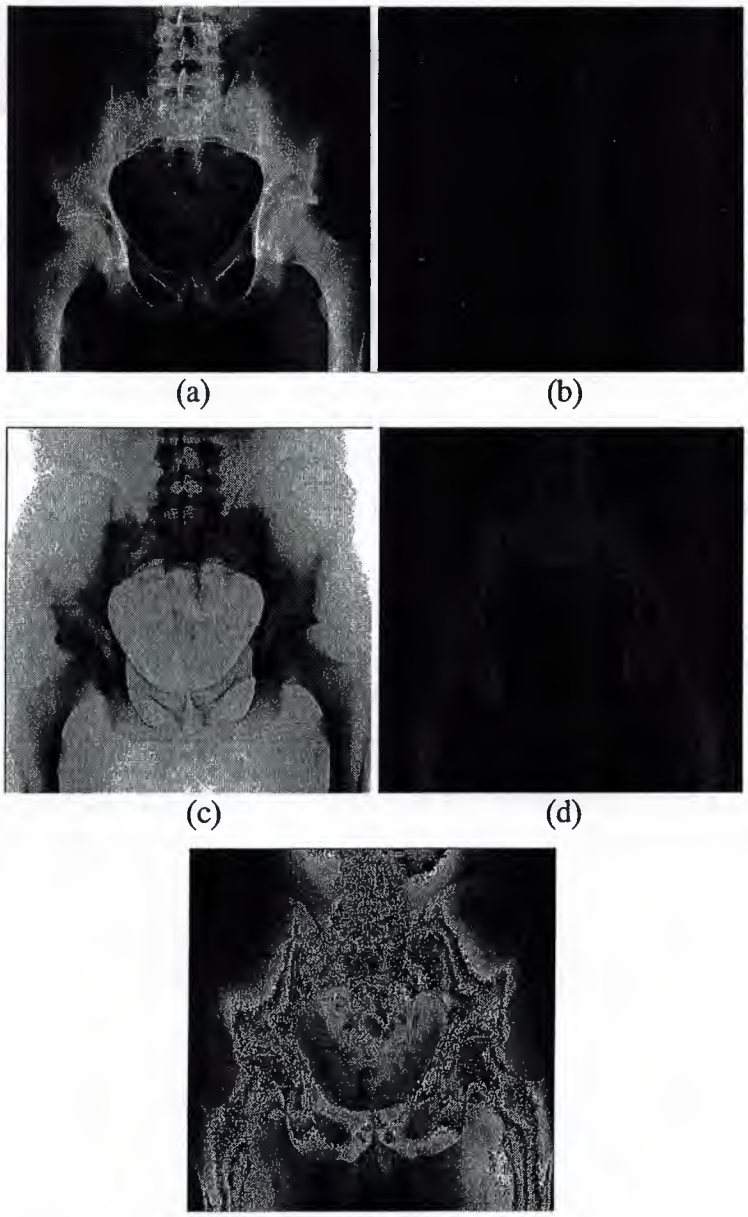


Figure 1.1 – Transformation Implementation of X-ray Image (a) Original Image, (b) Enhanced X-ray image using Log Transformation with $c=1$, (c) Enhanced Image using Image Negatives, (d) Enhanced Image using Power-Law Transformation with $\gamma=0.8$ $c=1$ and (e) Enhanced Image using Power-Law Transformation with $\gamma=1.2$ $c=1$.

(a) (b)

Figure 1.2 – Contrast Stretching on X-ray Image (a) Original Low Contrast Image and (b) Enhanced Image using Contrast Stretching.

1.2.1.2 Histogram Processing in Spatial Domain

In Spatial Domain, also Histogram Processing is an important approach for image enhancement and it is the basis for numerous processing techniques [2]. Histogram is the discrete function of digital image in the range k as $[0,L-1]$ and it is defined as :

$$h(r_k) = n_k \tag{1.7}$$

where r_k is the kth gray level and n_k is the number of pixels in the image having gray level r_k. Thus, it is easy to say that probability of occurrence of gray level r_k $(p(r_k))$estimated by dividing its values by total number of pixels in the image which is denoted as n in Equation 1.8. Also it is known as the *normalization* of a histogram.

$$p(r_k) = n_k / n \tag{1.8}$$

One of the basic usages of histograms is the determination of the contrast level (image types [2]) of the images such as dark image, bright image, low-contrast image and high-contrast image.

Dark image can be defined as the collection of image pixels in the range $[0,n]$, without having pixel values in the range $[n,L-1]$ where n is the gray-limit of image pixels and can be assumed as 128.

Bright image can be defined as the collection of image pixels in the range $[n, L-1]$, without having pixel values in the range $[0,n]$.

Low-contrast images have more complex relationship upper and lower limits of gray level values. An image can be classified as low contrast image if the image pixels are collected in the range $[n-z,n+z]$ where z is a variable to determine the upper and lower limits of image pixels.

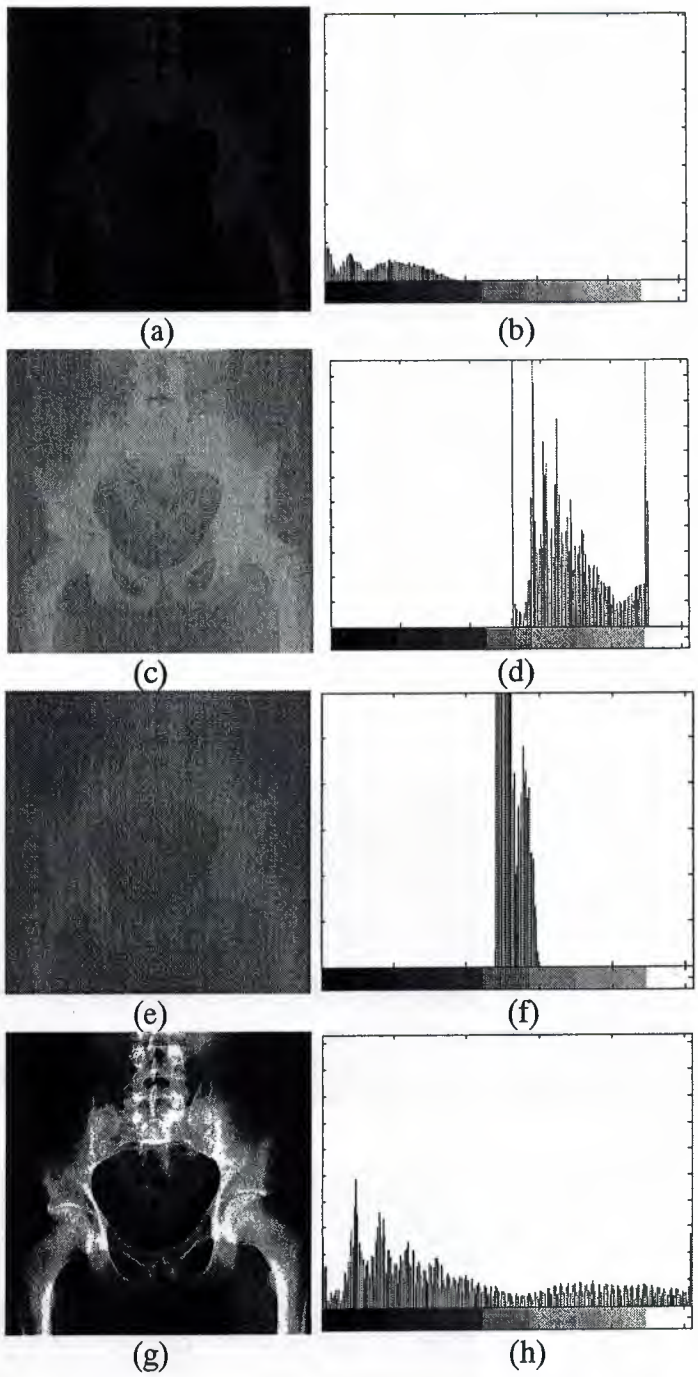


Figure 1.3 – Contrast Levels of X-ray Image and Corresponding Histograms (a)-(b) Dark Image and its Corresponding Histogram, (c)-(d) Bright Image and its Corresponding Histogram, (e)-(f) Low-Contrast Image and its Corresponding Histogram and (g)-(h) High-Contrast Image and its Corresponding Histogram.

In ideal case, high-contrast image can be defined as the equal distribution of image pixels in the range *[0, L-1]*. Examples of dark, bright, low-contrast and high contrast image with their corresponding histograms can be seen in Figure 1.3.

As it mentioned above, probability of occurrence of histogram can be computed by using equation 1.8 and histogram equalization can be defined as shown in equation 1.9:

$$s_k = T(r_k) = \sum_{j=0}^{k} p_r(r_j) \tag{1.9}$$

where T is the transformation function for histogram equalization, r_k is the kth gray level and n_k is the number of pixels in the image having gray level r_k, s_k is the histogram equalized image and $p(r_j)$ is the probability of the occurrence. By replacing equation 1.8 into the equation 1.9, we can simplify histogram equalization as shown in equation 1.10 and Histogram equalization applied bright and low contrast images of Figure 1.3 and their corresponding histograms can be seen in Figure 1.4.

$$\sum_{j=0}^{k} \frac{n_j}{n} \qquad k = 0,1,2,.....L-1 \tag{1.10}$$

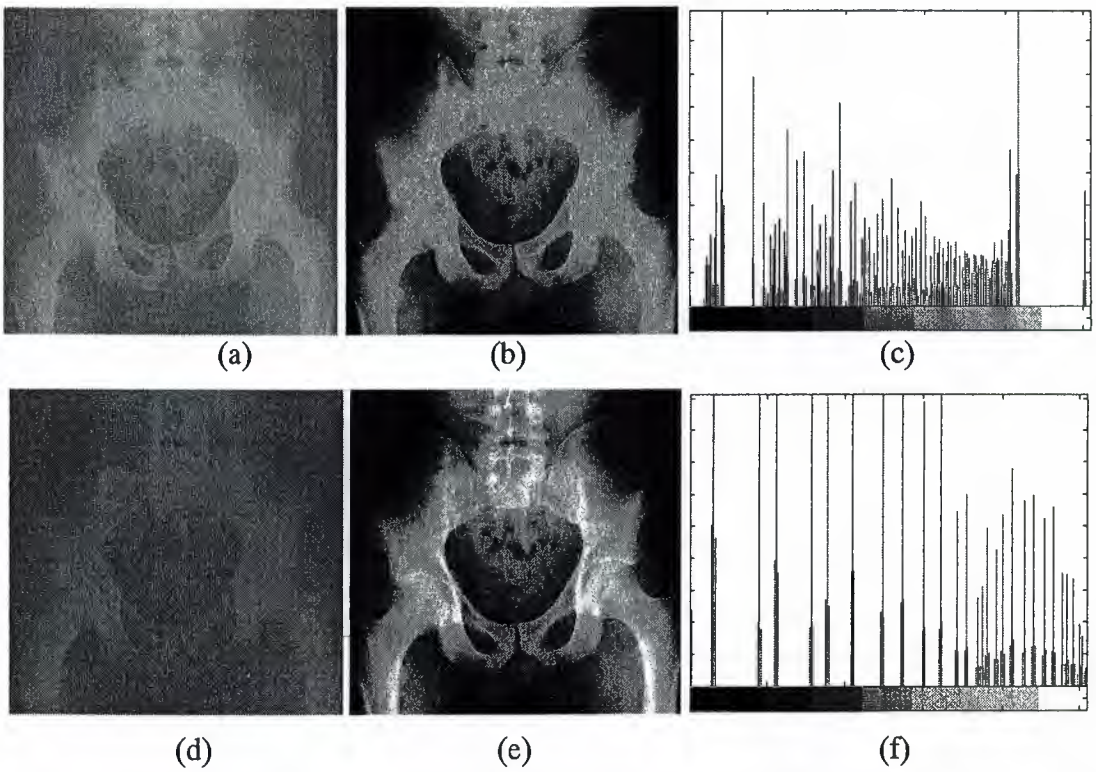


Figure 1.4 – Implementation of Histogram Equalization (HE) (a) Bright Image (b) Enhanced image of (a) using HE, (c) Corresponding Histogram of (b), (d) Low Contrast Image (e) Enhanced Image of (d) using HE, (f) Corresponding Histogram of (e).

1.2.1.3 Spatial Filtering : Smoothing and Sharpening Filters

The methods and approaches that were presented in previous sections are explained and simulated as global methods; however, it is easy to apply these methods in local kernels. For example, if transformation functions, such as Log and Power-Law transformations, or Histogram Equalization are applied in local kernels which are mostly defined as square or rectangular in a whole image, they become local enhancement methods that each of the defined kernels are independent from each other. Figure 1.5 shows the kernel operation on image with functions and coordinates.

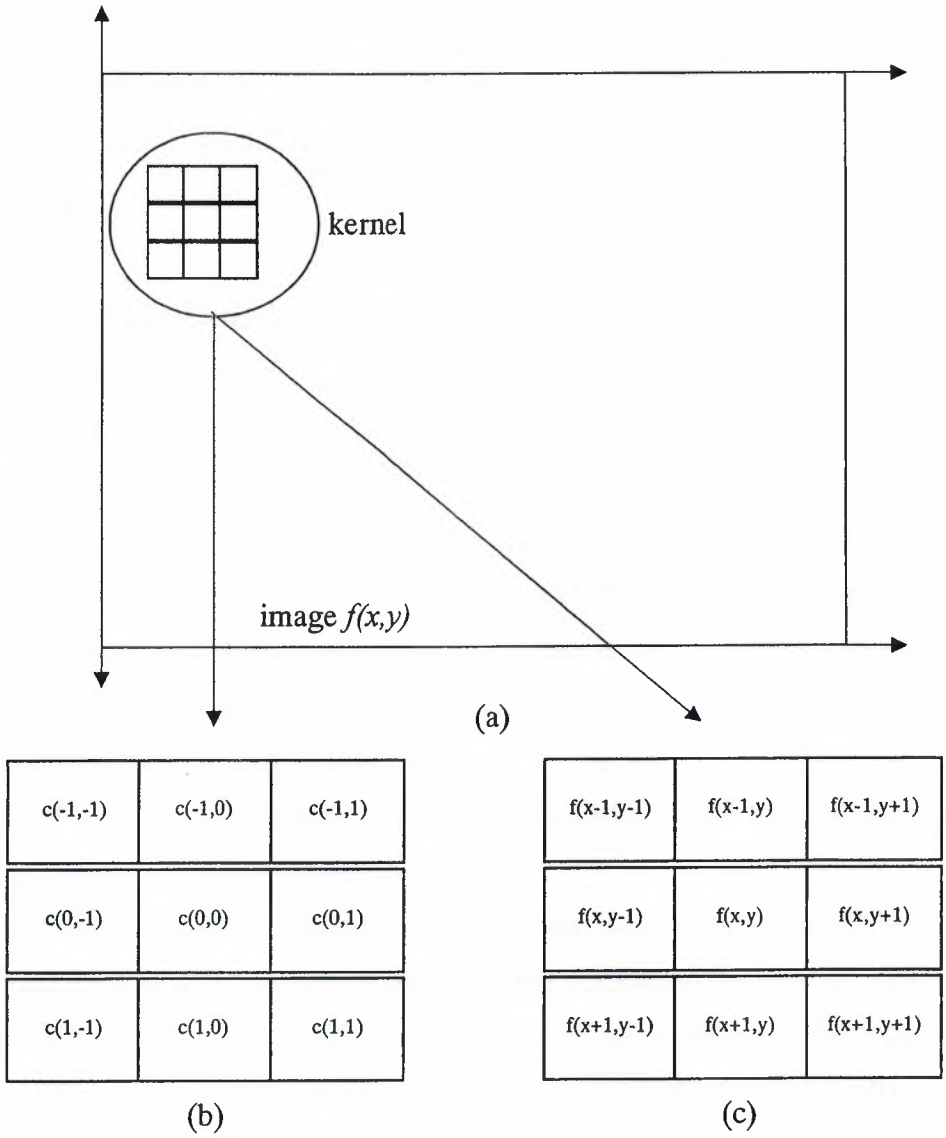


Figure 1.5 – Kernel Operation on Image (a) 3x3 kernel on image (b) represented coordinates of kernel and (c) operations in kernel. (Original drawing courtesy of R.C Gonzalez and R.E. Woods).

In spatial domain, main use of kernels are belong to the filtering approaches which can be classified into two groups as smoothing filters and sharpening filters.

Smoothing filters are used for blurring and for noise reduction [2]. Blurring is the removal of small details of image to make easy the extraction of objects or other interests and noise reduction is provided by applying some filters such as linear or non-linear.

Linear filters are straight forward methods which are directly applied to the defined kernels of image. They are generally the replacing the neighborhood pixels of kernel by the average of all pixels of kernel. Because of this reason, sometimes they are called *averaging filters*, however, mostly they are know as *lowpass filters*. Typical formulae of lowpass filters can be written as shown in equation 1.11.

$$R = \frac{1}{mxn} \sum_{i=1}^{mxn} z_i \tag{1.11}$$

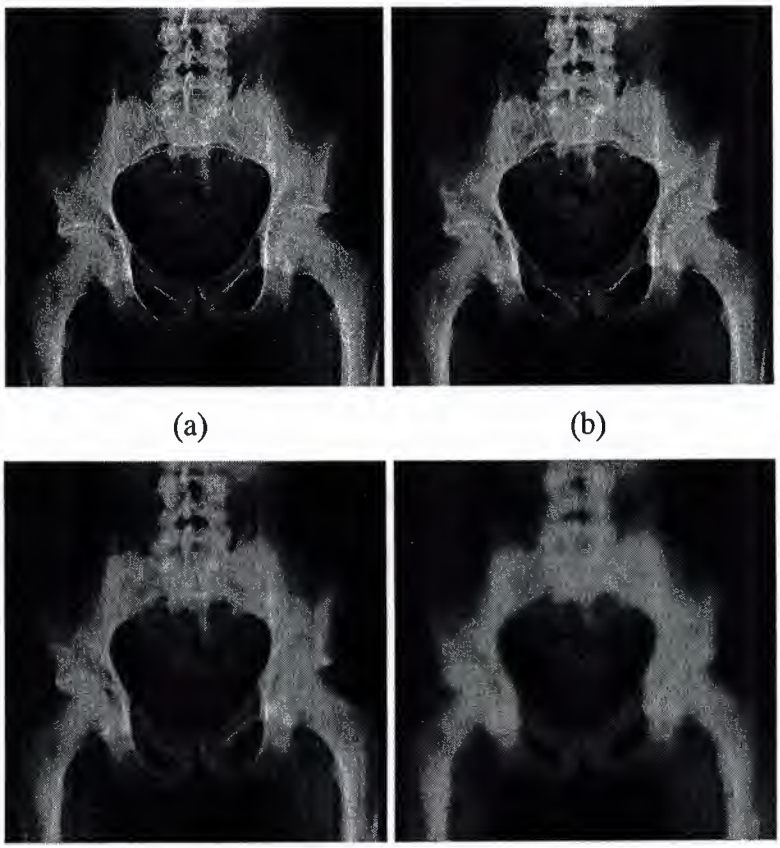
where R is the value to replace, m and n is kernel dimensions, and z is the pixel value within kernel neighborhood i.

Figure 1.6 shows the implementation of typical low-pass filter to the x-ray image by using different kernel sizes.

Non-linear filters which are generally called Order-Statistics Filters [2] in smoothing filters are based on the ranking of the pixels and replacing the center pixel with best ranking one. Most popular non-linear smoothing filter is median filter which is the best ranking was generally assumed the center pixel of sorted numbers which is 5th in *3x3* kernel and 13th in *5x5* kernel.

Figure 1.7 shows the implementation of median filter to the x-ray image by using *3x3* kernel size.

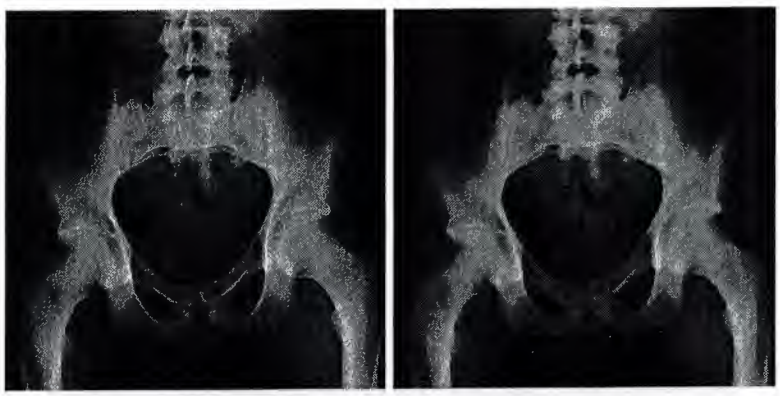
Another group of spatial domain filters is sharpening filters where the objective is the enhancement of noisy details of image. These noises can be blurring effect or the noises which is obtained during image acquisition. Sharpening filters are based on the first and second order derivatives of image which can be formulated basically as shown in Equation 1.12 and 1.13 respectively.



(a) (b)

(c) (d)

Figure 1.6 – Lowpass Filter Implementation of X-ray Image (a) Original X-Ray Image, (b) Enhanced Image using *3x3* kernel, (c) Enhanced Image using *5x5* kernel and (d) Enhanced Image using *15x15* kernel.



(a) (b)

Figure 1.7 – Median Filter Implementation of X-ray Image (a) Original X-ray image and (b) Enhanced Image using *3x3* kernel.

$$\nabla f = \frac{\partial f}{\partial x} + \frac{\partial f}{dy} = f(x+1,y) - f(x,y) + f(x,y+1) - f(x,y) \tag{1.12}$$

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{dy^2} = f(x+1,y) + f(x-1,y) + 2f(x,y) + f(x,y+1) + f(x,y-1) + 2f(x,y) \tag{1.13}$$

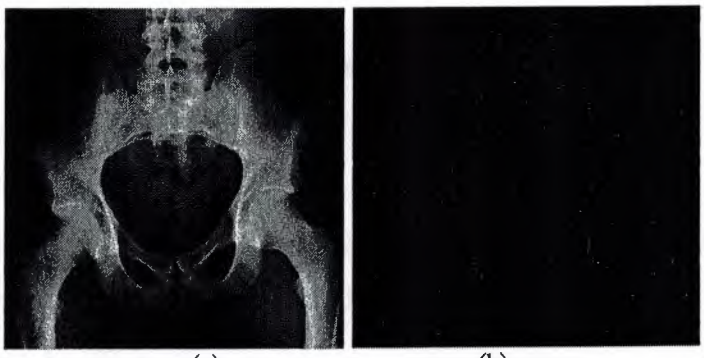
Implementation of second-order derivative of image which is called Laplacian Filtering can be obtained by using a mask which is shown in Figure 1.8.

However, in image enhancement, the use of Laplacian Filtering has some additional features to obtain enhanced image. These additional features can be seen in equation 1.14 and the result of Laplacian Filtering can be seen in Figure 1.8.

$$g(x,y) = \begin{cases} f(x,y) - \nabla^2 f(x,y) & \text{if the center coefficient of the Laplacian Mask is positive} \\ f(x,y) + \nabla^2 f(x,y) & \text{if the center coefficient of the Laplacian Mask is negative} \end{cases} \tag{1.14}$$

0	1	0
1	-4	1
0	1	0

Figure 1.8 – Laplacian Filtering Mask



(a) (b)

Figure 1.9 – Laplacian Filtering and Enhancement Implementation of X-ray Image, (a) Original X-ray Image, (b) the Result of Laplacian Filtering.

1.2.2 Overview of Frequency Domain Image Enhancement Techniques

In this section, basic definitions and the implementations of Discrete Fourier Transform (DFT) and the respected filters will be explained and presented.

In image processing, frequency domain always mentioned together with Discrete Fourier Transform (DFT) which is the discrete version of Fourier Transform (FT). The equations of single variable (one-dimensional) FT and DFT can be seen in Equation 1.15 and 1.16 respectively.

$$F(u) = \int_{-\infty}^{\infty} f(x) e^{-j2\pi ux} dx \tag{1.15}$$

where $j = \sqrt{-1}$.

$$F(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(x) e^{-j2\pi ux/M} \qquad \textit{for } u=0,1,2,3,\ldots\ldots M\text{-}1. \tag{1.16}$$

where $x=0,1,2,3,\ldots\ldots M\text{-}1$.

Also, it is possible to obtain *f(x)* by applying inverse Fourier Transformation which the continuous and discrete versions can be seen in Equation 1.17 and 1.18 respectively.

$$f(x) = \int_{-\infty}^{\infty} F(u) e^{-j2\pi ux} du \tag{1.17}$$

$$f(x) = \frac{1}{M} \sum_{u=0}^{M-1} F(u) e^{-j2\pi ux/M} \qquad \textit{for } x=0,1,2,3,\ldots\ldots M\text{-}1. \tag{1.18}$$

Thus, it is easy to express *F(u)* in polar coordinates as shown in Equation 1.19.

$$F(u) = |F(u)| e^{-j\phi(u)} \tag{1.19}$$

where

$$|F(u)| = \left[R^2(u) + I^2(u)\right]^{1/2} \tag{1.20}$$

is called the *magnitude* or *spectrum* of the Fourier Transform and,

$$\phi(u) = \tan^{-1}\left[\frac{I(u)}{R(u)}\right] \tag{1.21}$$

is called the *phase angle* or *phase spectrum* and the *power spectrum* defined as the square of the Fourier Spectrum as shown in Equation 1.22.

$$P(u) = |F(u)|^2 = R^2(u) + I^2(u) \tag{1.22}$$

where *R(u)* and *I(u)* are the real and imaginary part of *F(u)* respectively.

Also, it is easy to express two-dimensional continuous and discrete FT and their respecting inverse FT, phase angle and power spectrum as shown in Equations respectively.

$$F(u,v)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-j2\pi(ux+vy)}dxdy \tag{1.23}$$

$$f(x,y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)e^{-j2\pi(ux+vy)}dudv \tag{1.24}$$

$$F(u)=\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi(ux/M+vy/N)} \tag{1.25}$$

$$f(x,y)=\frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u)e^{-j2\pi(ux/M+vy/N)} \tag{1.26}$$

$$|F(u,v)|=\left[R^2(u,v)+I^2(u,v)\right]^{1/2} \tag{1.27}$$

$$\phi(u,v)=\tan^{-1}\left[\frac{I(u,v)}{R(u,v)}\right] \tag{1.28}$$

$$P(u,v)=|F(u,v)|^2=R^2(u,v)+I^2(u,v) \tag{1.29}$$

By using Euler's formula as shown in Equation 1.30, we can express the Equations 1.25 and 1.26 as shown in Equations 1.31 and 1.32.

$$e^{j\Theta}=\cos\Theta+j\sin\Theta \tag{1.30}$$

$$F(u)=\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\begin{bmatrix}\cos 2\pi(ux/M+vy/N)\\ -j\sin 2\pi(ux/M+vy/N)\end{bmatrix} \tag{1.31}$$

$$f(x,y)=\frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u)\begin{bmatrix}\cos 2\pi(ux/M+vy/N)\\ -j\sin 2\pi(ux/M+vy/N)\end{bmatrix} \tag{1.32}$$

Application of filtering in frequency domain generally has same procedure [2] which is started by the multiplication of input image by $(-1)^{x+y}$ (after preprocessing if necessary) to center the transform and continue by computing $F(u,v)$ (DFT) of the image by using Equations 1.25 or 1.31. Any filtering function which is denoted as $H(u,v)$ can be applied at this time by the multiplication with $F(u,v)$. Then it is easy to apply inverse DFT and to obtain the real part of the results by using Equations 1.26 or 1.32 which is followed by the multiplication of these results by $(-1)^{x+y}$ to normalize the centered transform. Thus, it is easy to show that the application of any filtering function can be written as shown in Equation 1.33.

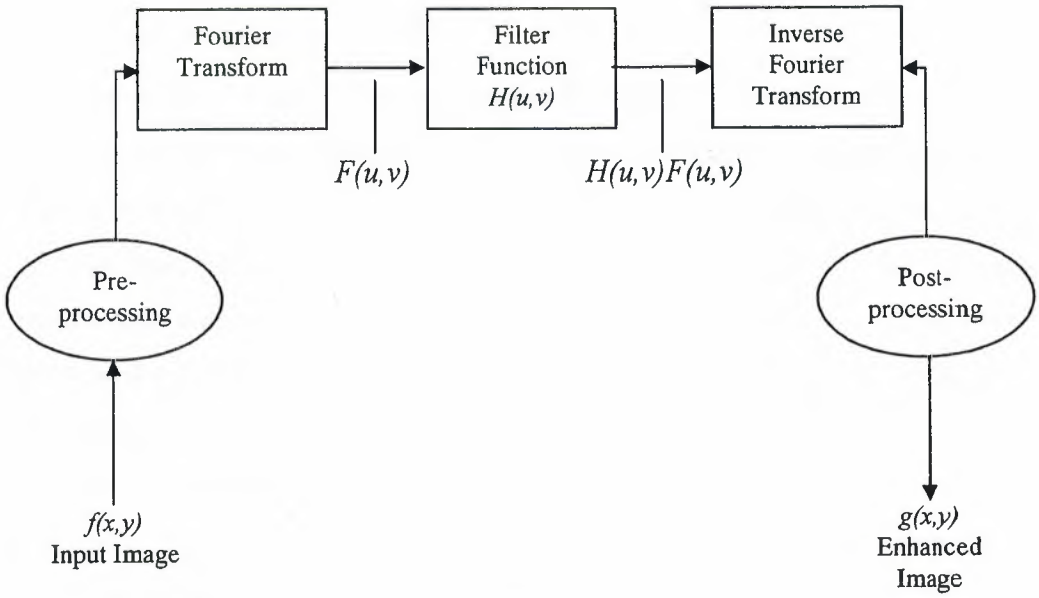


Figure 1.10 – Basic Filtering Operation Steps in Frequency Domain

$$G(u,v) = H(u,v)F(u,v) \tag{1.33}$$

General block diagram of filtering process in frequency domain can be seen in Figure 1.10.

Like spatial domain filters, we can divide frequency domain filtering approaches into two groups such as smoothing and sharpening filters.

1.2.2.1 Smoothing Filters in Frequency Domain

Smoothing can be obtained by the attenuation of high frequency signals by using a specified range in the DFT of image. As it was mentioned before, this attenuation can be achieved by applying filtering function which was defined in Equation 1.33.

Basic smoothing filters in frequency domain are Ideal Low Pass Filters (*ILPF*), Butterworth Low Pass Filter (*BLPF*) and Gaussian Low Pass Filter (*GLPF*).

One of the basic and simplest ILPFs is the 2D ILPF which is based on the defined distance D_0 from the centered DFT of an image. 2D ILPF cuts the higher frequency components of image which distance $D(u,v)$ is greater than D_0. Thus, transfer function can be written as shown in Equation 1.34 .

$$H(u,v) = \begin{cases} 1 & if \ D(u,v) \le 0 \\ 0 & if \ D(u,v) > 0 \end{cases} \tag{1.34}$$

Distance from any point (u,v) to the center of DFT can be expressed as :

$$D(u,v) = \left[(u - M/2)^2 + (v - N/2)^2\right]^{1/2} \tag{1.35}$$

Notice that, if the radius of a defined distance D_0 is relatively small, the image power will also be small and the result image will lose more information related to the loss of power. Thus, more blurred image will be obtained because of the more "cut-off" of high frequency components. However, if the radius of D_0 is relatively large value, power loss will be reduced and more detailed image which the visual appearance is increased will be obtained. Example of 2D Ideal Low Pass Filter implementation of X-ray image with Cut-off distance *10*, *50* and *150* can be seen in Figure 1.11.

One of the most important and widely used low-pass filtering is Butterworth Low Pass Filtering (BLPF) which can be applied in *n*th order of image. Transfer function of BLPF is defined as shown in Equation 1.36.

$$H(u,v) = \frac{1}{1 + \left[D(u,v)/D_0\right]^{2n}} \tag{1.36}$$

Like ILPF, the effect of radius value D_0 is almost same in BLPF. Example of Butterworth Low Pass Filter implementation of X-ray image in 2nd order with Cut-off distance *10*, *50* and *150* can be seen in Figure 1.12.

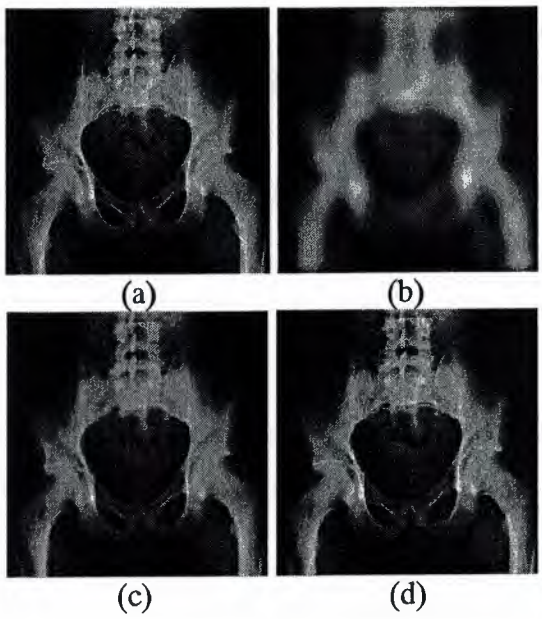
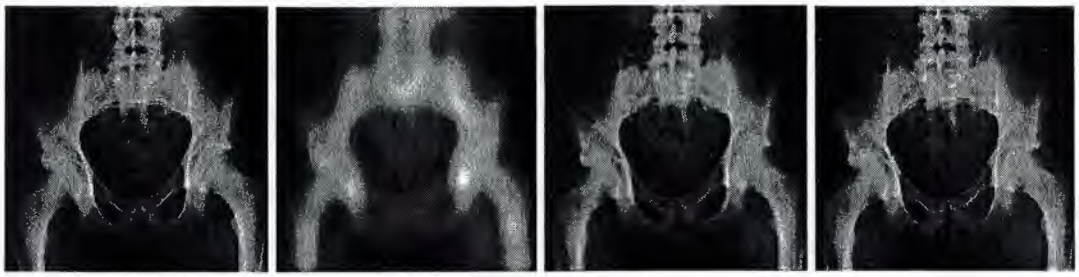
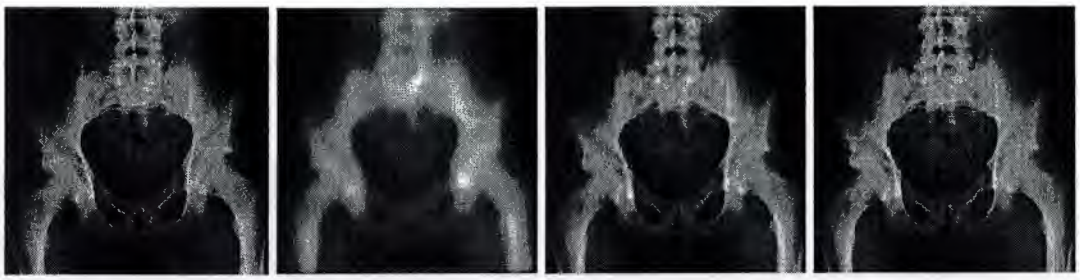


Figure 1.11 – 2D ILPF Implementation of X-ray Image with Various Cut-off Points D_0 (a) Original X-ray Image, (b) Filtering Result with Cut-off Point *10*, (c) Filtering Result with Cut-off Point *50* and (d) Filtering Result with Cut-off Point *150*. Note that the blurring effect in (b) with small size of cut-off point D_0.



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

Figure 1.12 – BLPF Implementation of X-ray Image in 2nd Order with Various Cut-off Points D_0 (a) Original X-ray Image, (b) Filtering Result with Cut-off Point *10*, (c) Filtering Result with Cut-off Point *50* and (d) Filtering Result with Cut-off Point *150*.



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

Figure 1.13 – GLPF Implementation of X-ray Image with Various Cut-off Points D_0 (a) Original X-ray Image, (b) Filtering Result with Cut-off Point *10*, (c) Filtering Result with Cut-off Point *50* and (d) Filtering Result with Cut-off Point *150*.

Another important Lowpass Filter in Frequency Domain is Gaussian Low Pass Filter (GLPF) which uses D_0 and $D(u,v)$ like other low-pass filters. The general formulae of Gaussian Low Pass Filter can be seen in Equation 1.37.

$$H(u,v) = e^{-D^2(u,v)/2\sigma^2} \tag{1.37}$$

where σ is the standard deviation of Gaussian Curve. However, it is possible to let $D_0 = \sigma$ and to express Equation 1.37 as shown in 1.38.

$$H(u,v) = e^{-D^2(u,v)/2D_0^2} \tag{1.38}$$

Example of Gaussian Low Pass Filter implementation of X-ray image with Cut-off distance *10, 50* and *150* can be seen in Figure 1.13.

1.2.2.2 Sharpening Filters in Frequency Domain

Sharpening can be achieved in Frequency Domain by high-pass filtering process, with attenuating low frequency components without disturbing high frequency components

[2]. Generally, high pass filtering is the reverse operation of low pass filtering and basically they can be described as shown in Equation 1.39.

$$H_{HP}(u,v) = 1 - H_{LP}(u,v) \tag{1.38}$$

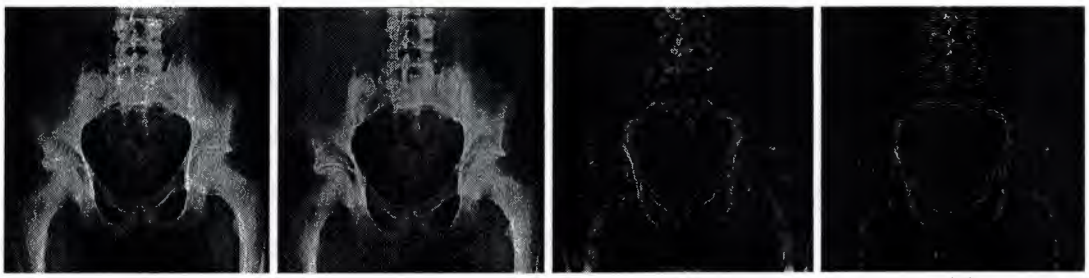
where H_{LP} is the low pass filtering transfer function.

Thus Ideal High Pass Filter, Butterworth High Pass Filter and Gaussian High Pass Filter can be expressed by using Equation 1.38 as shown in Equations 1.39, 1.40 and 1.41 respectively.

$$H(u,v) = \begin{cases} 0 & if\ D(u,v) \le 0 \\ 1 & if\ D(u,v) > 0 \end{cases} \tag{1.39}$$

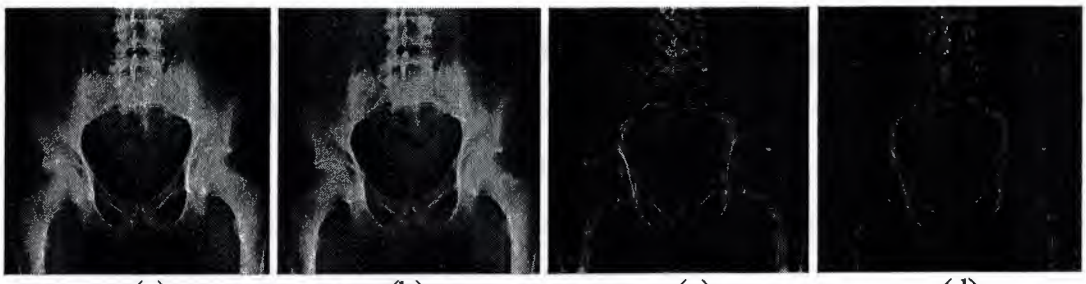
$$H(u,v) = \frac{1}{1 + [D_0 / D(u,v)]^{2n}} \tag{1.40}$$

$$H(u,v) = 1 - e^{-D^2(u,v)/2D_0^2} \tag{1.41}$$



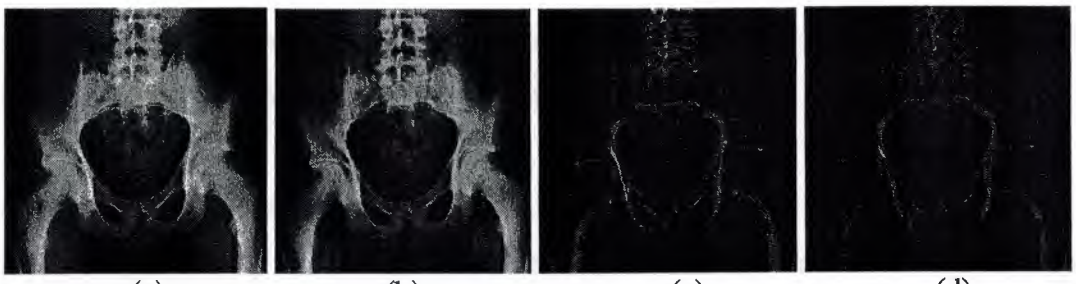
(a) (b) (c) (d)

Figure 1.14 – IHPF Implementation of X-ray Image with Various Cut-off Points D_0 (a) Original X-ray Image, (b) Filtering Result with Cut-off Point *1*, (c) Filtering Result with Cut-off Point *10* and (d) Filtering Result with Cut-off Point *20*.



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

Figure 1.15 – BHPF Implementation of X-ray Image with Various Cut-off Points D_0 (a) Original X-ray Image, (b) Filtering Result with Cut-off Point *1*, (c) Filtering Result with Cut-off Point *10* and (d) Filtering Result with Cut-off Point *20*.



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

Figure 1.16 – GHPF Implementation of X-ray Image with Various Cut-off Points D_0 (a) Original X-ray Image, (b) Filtering Result with Cut-off Point *1*, (c) Filtering Result with Cut-off Point *10* and (d) Filtering Result with Cut-off Point *20*.

Example of Ideal High Pass Filtering, Butterworth High Pass Filtering and Gaussian High Pass Filter implementation of X-ray image with Cut-off distance *1,10* and *20* can be seen in Figure 1.14, 1.15 and 1.16 respectively.

1.3 Main Application Areas of Image Enhancement

The usage of image enhancement has increasing popularity in almost every field of life. It can be used in everywhere that requires optimum visual appearance of images or objects. Most important application areas of image enhancement are Medical Imaging, Military-Security-Forensic Sciences, Document Analysis, and Image Preprocessing.

1.3.1 Enhancement in Medical Imaging

Medical Imaging consists several areas that enhancement of images are necessary. Widely used medical imaging techniques are Digital X-Ray, Digital Mammography [5,6,7], CT Scans [8,9,10] and MRI [10]. The aim of image enhancement in medical imaging is to improve visual appearance of image to provide optimum diagnosis of diseases. For example, in X-ray image, it is important to enhance image to see if there is any broken bones in patient and in mammography, it is important to show all cells clearly to see if there are any cancer cells or tumor.

In enhancement of medical imaging, either existing spatial domain approaches or frequency domain approaches can be used or new techniques can be developed based on these domains. For, example J.K.Kim et al. [5] developed a technique by using first derivatives and local statistics of images which are belong to spatial domain approaches to improve the appearance of mammographic images and a techniques that was based

Fast Fourier Transform (FFT) was presented by E.W. Abel et al. [11] to increase the visual appearance of cancellous bones of x-ray image.

1.3.2 Enhancement in Military, Security and Forensic Sciences

In military, security and forensic sciences, main application areas of image enhancement are the improvement of night-vision images [12], fingerprint images [13][14], face components [15], and satellite images[16].

In night vision and satellite images, generally it is important to increase the visuality of each component of dark or noisy image, however in fingerprint and face images, it is more important to clear unnecessary data to extract features from the images.

Like all enhancement applications, any spatial or frequency domain approaches can be efficient to increase the visual appearance of images, however, it is not guaranteed that a method should produce optimum results for all night-vision, fingerprint, face or satellite images.

1.3.3 Enhancement in Document Analysis

In document analysis, the aim can be the extraction of the characters after providing effective reducing of the noises and the additional layers within documents and to provide readability of documents or to prepare the document for optical character recognition modules.

Thus, both aims of document analysis require different enhancement methods to achieve readable and separable documents. For example, the improvement of readability of the documents can be useful for fax documents to eliminate added noises while transmission [17], however separation can be useful for digitizing documents [18].

1.4 Summary

The visual appearance of images can be increased by using several enhancement methods which are belong to one of the two domains, spatial or frequency domain. In spatial domain, methods are applied directly to the image, however in frequency domain, images are considered as signals and methods or filters can be applied after obtaining Discrete Fourier Transform of image.

For both domains, output images can be different or same according to the applied techniques, applications and the characteristics of an image. Thus, it is almost impossible to determine which domain produces optimum results. In next chapter, binarization process will be explained in details which are spatial-domain image processing technique.

CHAPTER 2
REVIEW OF BINARIZATION METHODS

2.1 Overview

Image binarization (thresholding) is the low-level image processing method to separate and to enhance the region of interest to provide increased visual appearance of image. This enhancement and separation is provided by dividing image into two regions as background (logical 1) and foreground (logical 0). Ideally, separated image of foreground is expected to have a region of interest or object in image with a minimum loss of information and fuzziness. Thus, it should not consist any pixels belong to the background and several techniques are developed to achieve this aim.

In this chapter, basic definitions of image binarization, chronological development, detailed explanation about selected thirteen methods and application areas will be presented.

2.2 Fundamentals of Image Binarization

Image Binarization is one of the basic spatial domain image processing technique that is used to segment or enhance the region of interest with in image. It is based on the assumption that object and background can be distinguished by their gray level values [19] and the result of this assumption is caused for the development of several thresholding methods that use various properties of images. General image binarization can be expressed as shown in Equation 2.1.

$$g(x,y) = T[f(x,y)] \tag{2.1}$$

where *f(x,y)* is the input image, *g(x,y)* is the processed image, and *T* is an operator on *f*, defined over some neighborhood of *(x,y)*.

However, the difference between the spatial domain techniques that were explained in Chapter 1 and image binarization, is the output image, where in binarization, it consists only *0* (binary 0) and *255* (binary 1). Thus characteristic formulae of image binarization with threshold point *Θ* can be defined as shown in Equation 2.2.

$$g(x,y) = \begin{cases} 0, if & g(x,y) \le T(f[x,y]) = \Theta \\ 255, else & \end{cases} \tag{2.2}$$

General properties of binarization methods are mostly common for all methods, especially for global ones. Gray-level image histogram, probability density function and its corresponding standard deviation, mean, priori probability and image entropy should be understood before implementing and analyzing any method.

Gray level image histogram which was defined in Equation 1.7 is plotted distribution of the number of pixels that have same gray value and for binarization methods generally defined as follows:

$$h(g) = n_g \tag{2.3}$$

where g is the gray level and n_g is the number of pixels in the image having gray level g.

In image processing and binarization, probability density function is used to normalize gray level histogram of images and it was defined as below:

$$PDF = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{2.4}$$

where σ and μ are the variance and the mean of image and was defined in Equation 2.5 and 2.6 respectively:

$$\sigma^2(T) = \left[\sum_{g=a}^{b} \{g - \mu\ (T)\}^2 p_a(g)\right] \tag{2.5}$$

where g is the gray level, μ is the mean, $h(g)$ is the gray level histogram, $p_a(g)$ is the gray level distribution and a and b are the lowest and highest gray level value of distribution.

$$\mu(T) = \left[\sum_{g=a}^{b} h(g)g\right] \Big/ P(T) \tag{2.6}$$

Gray level distribution is defined as follows:

$$p(T) = \sum_{g=a}^{b} h(g) / NxM \tag{2.7}$$

where $h(g)$ is the gray level histogram, a and b are the lowest and highest gray level values of distribution and N and M are the x and y dimension of image or kernel. Priori probability $P(T)$ was defined as follows:

$$P(T) = \sum_{g=a}^{b} p(g) \tag{2.8}$$

Image entropy is the other way to perform binarization methods. Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image and was defined as shown in Equation 2.9:

$$H(T) = -\sum_{g=0}^{T} p(g) \log p(g) \tag{2.9}$$

In order to provide efficient separation and enhancement of region of interest within image, several thresholding methods which can be classified into two groups such as Global Binarization Methods and Local Binarization Methods, were proposed.

Global Thresholding methods consider whole image and its global characteristics to determine a single threshold value and Local Thresholding divides image into kernels to determine individual threshold values for each kernel. However, both groups carry out some disadvantages beside their advantages. Global methods have generally faster execution time and less noise in resultant image than local methods, however, according to the characteristics of document images, they can be over or under thresholded that cause some loss of relevant information. Local methods generally produce images with less loss of relevant information than global methods; however, the kernel size which is the main disadvantages of local methods brings some additional noises to these images in small sizes and they behave like global methods and can be over-thresholded in large sizes.

In literature, one of the first proposed thresholding methods is Riddler and Calvard [20] method (1978) which is based on the change of the foreground and background class means at iteration *n*. This method was followed by Otsu [21] method (1979) which became one of the most popular global methods and uses variances within image to determine final threshold point (see Section 2.3.1) and Nakagawa and Rosenfeld [22] proposed one of the first local thresholding methods which is known as Nakagawa and Rosenfeld implementation of Chow and Kaneko [23]. Then Pun [24], proposed the use of image entropy in threshold selection in 1980. In same year, Yasuda et al. [25] proposed another local thresholding method.

In 1983, White and Rohrer [26] proposed local thresholding which compares the gray value of the pixel with the average of the gray values in some neighborhood and if the pixel is significantly darker than the average, it is denoted as character; otherwise, it is classified as background. Also in same year, Rosenfeld et al. [27] proposed

histogram-based global thresholding method that based on analyzing the concavities of the histogram *h(g)* vis-and its convex hull.

In 1985, Kapur et al. [28] proposed entropy based thresholding method that later become one of the most famous entropy-based method (See Section 2.3.5). At that time, Lloyd [29] proposed another global method that divided image histogram into two clusters and minimizes misclassification error between these clusters.

Then in 1986, Kittler and Illingworth [19] proposed their Minimum Error Thresholding technique (See Section 3.2.2) which is based on clustering of image histogram like Lloyd method. Also, Niblack [30] and Bernsen [31] independently proposed their local thresholding methods, which are still most popular and mostly compared methods (see Section 2.4.1 and 2.4.6) in 1986. Palumbo et al. [32] proposed another local threshold method in same year which consists in measuring the local contrast of five neighborhoods.

In 1989, Abutaleb [33] proposed global thresholding method which was based on two-dimensional entropy of image and Yanowitz and Bruckstein [34] proposed local thresholding method that uses the discrete Laplacian of the surface, produced by using the combination of edge and gray level information. Again in 1989, Taxt et al. [35] proposed local thresholding method for document image segmentation.

In 1991, Eikvil et al. [36] proposed local thresholding which is based on image clustering of small window in larger concentric window. In that year, Parker [37] proposed another local thresholding that first detects edges and the interior of objects is filled.

In 1993, Li and Lee [38] proposed another entropy based method that minimizes the theoretic distance of information. In that year, Kamel and Zhao [39] proposed another local thresholding method that measures the difference of local mean and local pixel and compare it by predetermined value to determine threshold point for each segment.

Yanni and Horne [40] proposed global thresholding method in 1994 which uses the midpoint of two assumed peaks to determine final threshold (see Section 2.3.3).

In 1995, Ramesh et al. [41] proposed global thresholding that uses a simple functional approximation to minimize histogram (see Section 2.3.4).

In 1995 Yen et al. [42], in 1996 Pal [43] and in 1997 Sahoo et al. [44] proposed another entropy based thresholding methods and recently Albuquerque et al. [45] proposed another entropy based method that uses Tsallis Entropy (see Section 2.3.6).

Oh and Lindquist [46] proposed local method in 1999 and this method was followed by Sauvola et al. [47] method in 2000 which recently become popular while improving Niblack method (see Section 2.4.1).

In 1999, Solihin and Leedham [48] proposed global thresholding technique which is based on integral ratio.

In 2000, Yibing and Yang [49] improved the Kamel and Zhao logical thresholding technique (see Section 2.4.5) to determine required parameters automatically. In 2002, Wold and Jolion [50] improved Sauvola method to normalize contrast and the local mean of image to decrease the amount of noise.

In 2003, Leedham et al. [51] proposed Mean-Gradient technique which is based on local mean and local mean gradient of image (see Section 2.4.3) and in same year, Badekas and Papamarkos [52] improved Adaptive Logical Thresholding of Yibing and Yang and Sezgin and Sankur [53] proposed global thresholding method which is based on sample moment function..

Park et al. [54] proposed a new method that uses 3D terrain of grayscale image and simulates waterfall to binarize images in 2004 (see Section 2.4.7).

In 2005, Kavallieratou [55-56] proposed iterative global thresholding, which was designed especially for document images and calculates the difference of mean value and current pixels and uses histogram equalization in each iteration to clean and binarize images. Also in that year, Leedham and Chen [57] proposed decompose algorithm which requires several processing steps that includes Mean Gradient method of Leedham et al.

In 2006, authors proposed local thresholding method [58], that use local mean value as threshold level for each segment (see Section 2.4.4). Table 2.1 and 2.2 show chronological order of basic and recently proposed global and local thresholding methods respectively.

2.3 Global Binarization Methods

Global thresholding methods use a defined or computed threshold value for the entire image and several techniques that intend to achieve optimum thresholding point.

In next subsections, most popular conventional and recently proposed six global methods will be explained and in Section 2.3.7 advantages and disadvantages of global binarization methods will be presented.

Table 2.1 – Chronological Order of Basic and Recently Proposed Global Thresholding Methods

	Author	Year	Features
1	Riddler and Calvard	1978	Iterative clustering
2	Otsu	1979	Class separability
3	Pun	1980	Maximum Shannon's entropy
4	Rosenfeld et al.	1983	Histogram concavities and convex hull
5	Kapur et al.	1985	Entropy
6	Lloyd	1985	Clustering and minimizing error
7	Kittler and Illingworth	1986	Minimum error between clusters
8	Abutaleb	1989	High order entropy
9	Li and Lee	1993	Entropy and theoretic distance
10	Yanni and Horne	1994	Clustering and peak values
11	Ramesh et al.	1995	Functional approximation
12	Yen et al.	1995	Entropic correlation
13	Don	1995	Noise Attribute
14	Pal	1996	Maximum entropy
15	Sahoo et al.	1997	Renyi entropy
16	Solihin and Leedham	1999	Integral ratio
17	Sezgin and Sankur	2003	Sample Moment Function
18	M. Portes de Albuquerque et al	2004	Tsallis entropy
19	Kavallieratou	2005	Iterative histogram equalization

Table 2.2 – Chronological Order of Basic and Recently Proposed Local Thresholding Methods

	Author(s)	Year	Features
1	Nakagawa and Rosenfeld	1979	
2	Yasuda et al.	1980	Local intensity change
3	White and Rohrer	1983	Based on local mean and neighbors
4	Niblack	1986	Local mean and deviation
5	Bernsen	1986	Local based on neighbors
6	Palumbo et al.	1986	Local contrast
7	Yanowitz and Bruckstein	1989	Threshold surface
8	Taxt et al.	1989	Mixture of two Gaussian distribution
9	Eikvil et al.	1991	The pixels inside a small window are thresholded on the basis of clustering in larger window
10	Kamel and Zhao	1993	Local contrast and logical level
11	Oh and Lindquist	1999	Two pass algorithm
12	Sauvola and Pietikainen	2000	Improvement of Niblack
13	Yibing and Yang	2000	Adaptive logical level
14	Wold and Jolion	2002	Improvement of Sauvola et. al.
15	Badekas and Papamarkos	2003	Improvement of adaptive logical level
16	Leedham et al.	2003	Local mean and gradient
17	Park et al.	2004	Rainfall simulation
18	Chen and Leedham	2005	Decompose algorithm
19	Khashman and Sekeroglu	2006	Local mean

These methods are: Otsu Method [21], Kittler and Illingworth Minimum Error Technique [19], Yanni and Horne method [40], Ramesh et al. method [41], Kapur et al. Entropy Method [28] and Albuquerque et al. Entropy Method [45].

2.3.1 Otsu Method

Otsu method [21] was proposed in 1979 as a selection method which was based on image histogram. It uses discriminant analysis to divide foreground and background by maximising the discriminant measure. According to Ng and Lee [59], the threshold operation is regarded as the partitioning of the pixels of an image into two classes C_0 and C_1 (e.g., objects and background) at grey-level t, i.e., $C_0 = \{0, 1... t\}$ and $C_1 = \{t + 1, t+2 ...l-1\}$. An optimal threshold point can be determined by minimizing one of the following equations using within-class variance, between-class variance, and the total variance, σ_w^2, σ_b^2, σ_t^2 respectively.

$$\lambda = \left(\sigma_b^2/\sigma_w^2\right),\ \eta = \left(\sigma_b^2/\sigma_T^2\right) \text{ and } k = \left(\sigma_T^2/\sigma_w^2\right) \tag{2.10}$$

Thus, the optimal threshold value can be found using only the term:

$$\sigma_B^2(k).\sigma_B^2(k) = [\mu_T\omega(k) - \mu(k)]^2/\omega(k)[1-\omega(k)] \tag{2.11}$$

$$k^* = ArgMin\eta \tag{2.12}$$

Operations of Otsu method can be seen in Figure 2.1.

2.3.2. Kittler and Illingworth method

Kittler and Illingworth method [19], which is based on clustering the image, starts by choosing an arbitrary initial threshold T and compares both sides of T to determine error. Then, T is shifted and determined errors are compared to find a minimum error point which is assigned as a threshold point. The simplest formulae can be written as:

$$J(\tau) = \min_T J(T) \tag{2.13}$$

where $J(\tau)$ is minimum error threshold and $J(T)$ is the criterion function. Also $J(T)$ can be written directly as:

$$J(T) = 1 + 2[P_1(T)\log\sigma_1(T) + P_2(T)\log\sigma_2(T)] - 2[P_1(T)\log P_1(T) + P_2(T)\log P_2(T)] \tag{2.14}$$

where P_1 and P_2 denote Priori Probability and σ_1 and σ_2 denote standard deviations of left and right sides of T respectively. Operations of Kittler and Illingworth method can be seen in Figure 2.2.

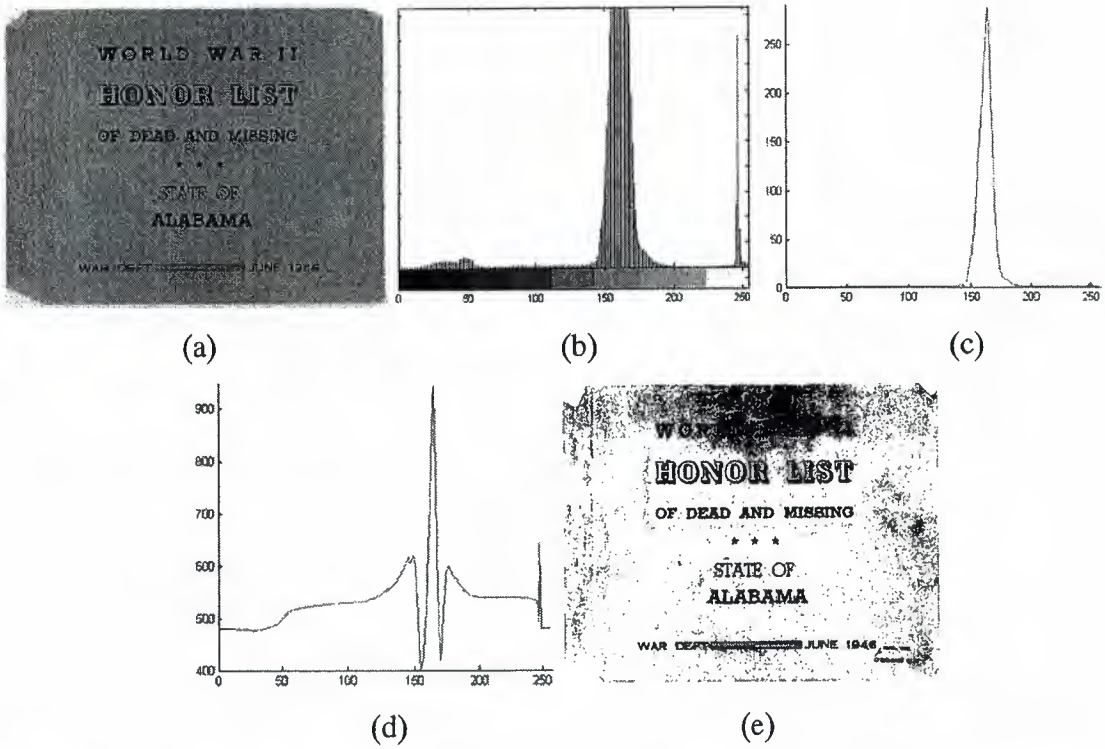


Figure 2.1 – Otsu Thresholding Operations (a) Original Image, (b) Corresponding Gray-level Histogram, (c) Gaussian Distribution of Histogram, (d) Minimum Arguments of 154 and (e) Binarized Image.

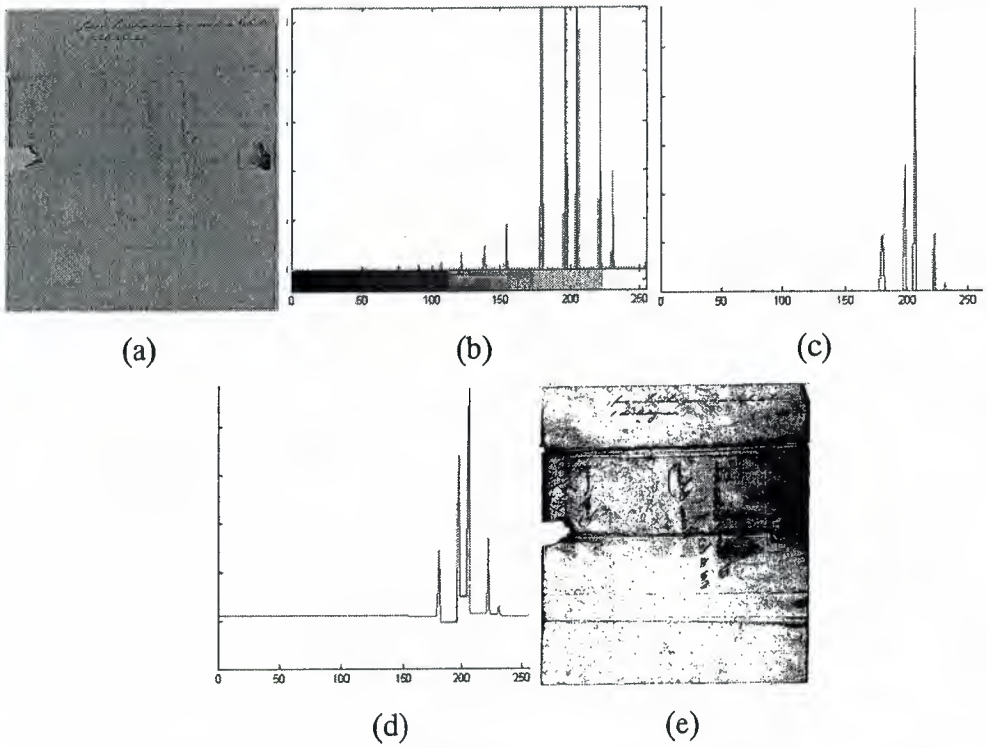


Figure 2.2 – Kittler and Illingworth Thresholding Operations (a) Original Image, (b) Corresponding Gray-level Histogram, (c) Gaussian Distribution of Histogram, (d) Error Graph *J(T)* with minimum error point *T=195* and (e) Binarized Image.

2.3.3. Yanni and Horne method

Yanni and Horne method [40] initializes the midpoint of two peaks of image histogram which is defined as:

$$g_{mid} = (g_{\max} + g_{\min})/2 \tag{2.15}$$

where g_{mid} is the midpoint of assumed peaks of image histogram and g_{max} and g_{min} are highest and lowest gray level respectively. The midpoint is updated using the mean of the two peaks on the right and left which can be written as:

$$g^*_{mid} = (g_{peak1} + g_{peak2})/2 \tag{2.16}$$

where g^*_{mid} is updated midpoint and g_{peak1} and g_{peak2} are the mean values of left and right sides of initial midpoint respectively. Finally, optimum threshold is calculated as shown in Equation 2.17:

$$T_{opt} = (g_{\max} - g_{\min}) \sum_{g=g_{\min}}^{g^*_{mid}} p(g) \tag{2.17}$$

2.3.4. Ramesh et al. Method

Ramesh et al. method [41] is based on the approximation of distributed grey-level histogram and it divided this distributed histogram into two parts T_0 and T_1, and finds the minimum argument of the summation of these parts, which is defined as:

$$T_{opt} = \arg\min[T_0 + T_1] \tag{2.18}$$

where T_0 and T_1 is the left and right sides of histogram and can be defined as :

$$T_0 = \sum_{g=0}^{T} (\mu_0(T)/P(T) - g)^2 \tag{2.19}$$

$$T_1 = \sum_{g=T+1}^{L-1} ((\mu_1(T)/1 - P(T)) - g)^2 \tag{2.20}$$

Operations of Ramesh method can be seen in Figure 2.3.

2.3.5. Kapur et al. Entropy Method

Kapur et al. method [28] divides an image into two classes such as background and foreground, and assumes these classes have different signal source. Maximum summation of these two classes entropies is considered as the optimum threshold, which is defined as:

$$T_{opt} = \arg\max\left[H_f(T) + H_b(T)\right] \tag{2.21}$$

where $H_f(T)$ and $H_b(T)$ is the foreground and background entropies of image and defined as:

$$H_f(T) = -\sum_{g=0}^{T} p(g)/P(T) \log p(g)/P(T) \tag{2.22}$$

$$H_b(T) = -\sum_{g=T+1}^{G} p(g)/P(T) \log p(g)/P(T) \tag{2.23}$$

where $p(g)$ and $P(T)$ are probability mass function and area probability, respectively. Operations of Kapur et al. method can be seen in Figure 2.4.

2.3.6. Albuquerque et al. Entropy Method

Albuquerque et al. Tsallis entropy thresholding [45] is based on Kapur et al. entropy method however it uses Tsallis entropy form due to the presence of non-additive information in some classes of images.

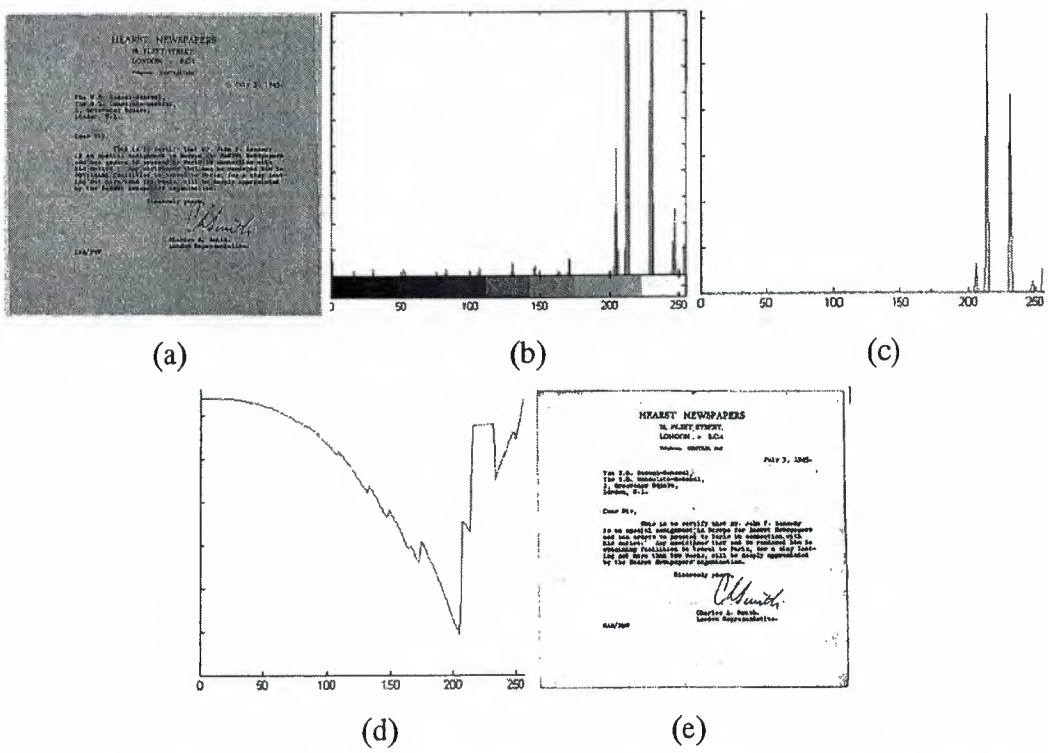


Figure 2.3 – Ramesh et al. Thresholding Operations (a) Original Image, (b) Corresponding Gray-level Histogram, (c) Gaussian Distribution of Histogram, (d) Error Graph with minimum argument point $T=204$ and (e) Binarized Image.

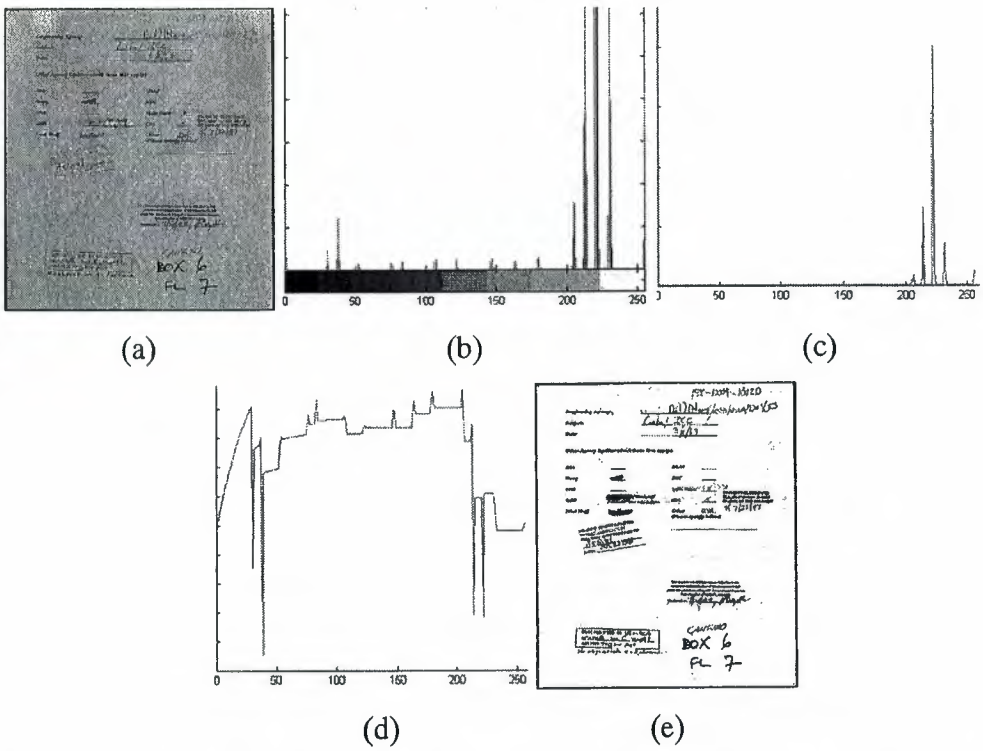


Figure 2.4 – Kapur et al. Thresholding Operations (a) Original Image, (b) Corresponding Gray-level Histogram, (c) Gaussian Distribution of Histogram, (d) Summation Graph of Two Classes with Maximum Argument point *T=204* and (e) Binarized Image.

Similar to Kapur et al. method, image divided into two classes such as background and foreground, and maximum argument of calculated T is selected as optimum threshold value. General formulae can be seen in Equation (2.24).

$$t_{opt} = \arg\max\left(S_q^A(t) + S_q^B(t) + (1-q).S_q^A(t).S_q^B(t)\right) \tag{2.24}$$

where q is an entropic index that characterizes the degree of non-extensivity, S_q^A and S_q^B are Tsallis entropy of image foreground and background which were defined as shown in Equation (2.25) and (2.26).

$$S_q^A(t) = 1 - \sum_{i=1}^{t}\left(\frac{p_i}{p^A}\right)^q \Big/ q-1 \tag{2.25}$$

$$S_q^B(t) = 1 - \sum_{i=1}^{t}\left(\frac{p_i}{p^B}\right)^q \Big/ q-1 \tag{2.26}$$

where p_i, p^A and p^B are probability distribution level, and probability distribution of foreground and background respectively.

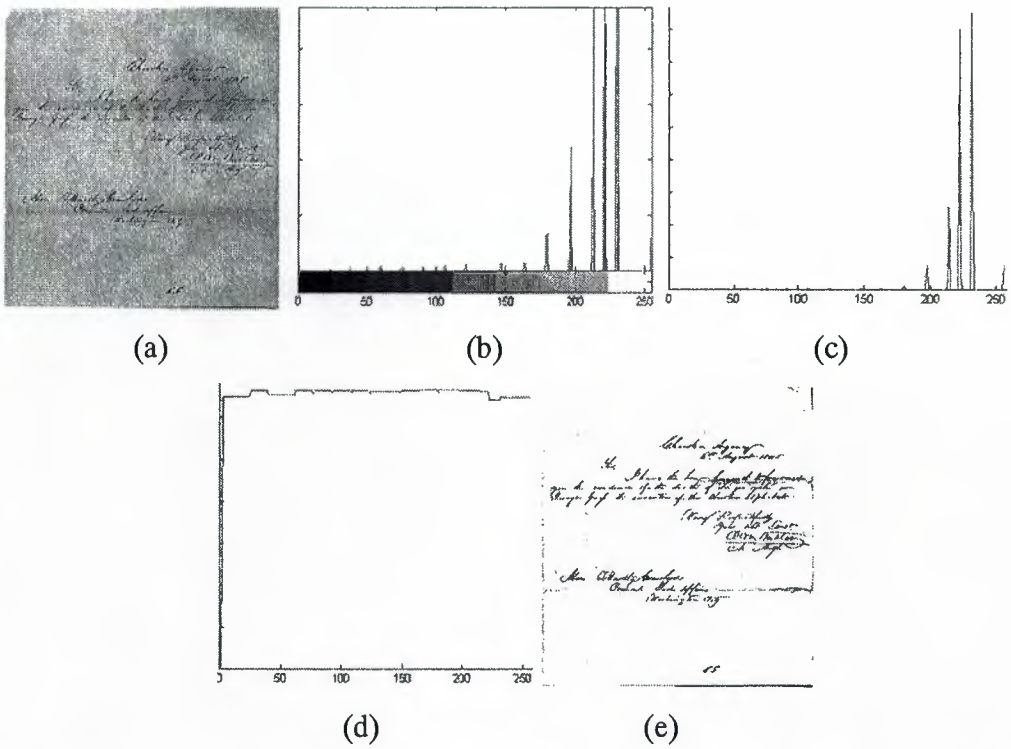


Figure 2.5 – Albuquerque et al. Thresholding Operations (a) Original Image, (b) Corresponding Gray-level Histogram, (c) Gaussian Distribution of Histogram, (d) Summation Graph of Two Classes with Maximum Argument point *T=179* and (e) Binarized Image.

Operations of Albuquerque et al. method can be seen in Figure 2.5.

These six global methods which were explained above, are also selected to perform comparison with proposed method in Chapter 3 and 4, because of their popularity in document binarization which almost every research in document binarization comprises the comparison at least three of these methods. Recently proposed method Albuquerque et al. Entropy Method was proposed as the optimum in entropy based methods, thus it was also included to these six methods.

2.3.7 Advantages and Disadvantages of Global Binarization Methods

Global binarization methods have some disadvantages besides their apparent advantages of binarizing images with various degrees of success depending on the type of image. The main advantages of global methods can be listed as faster execution time and less noise in resultant images. However, depending on the characteristics of the images,

global methods can over or under threshold which causes some loss of relevant information.

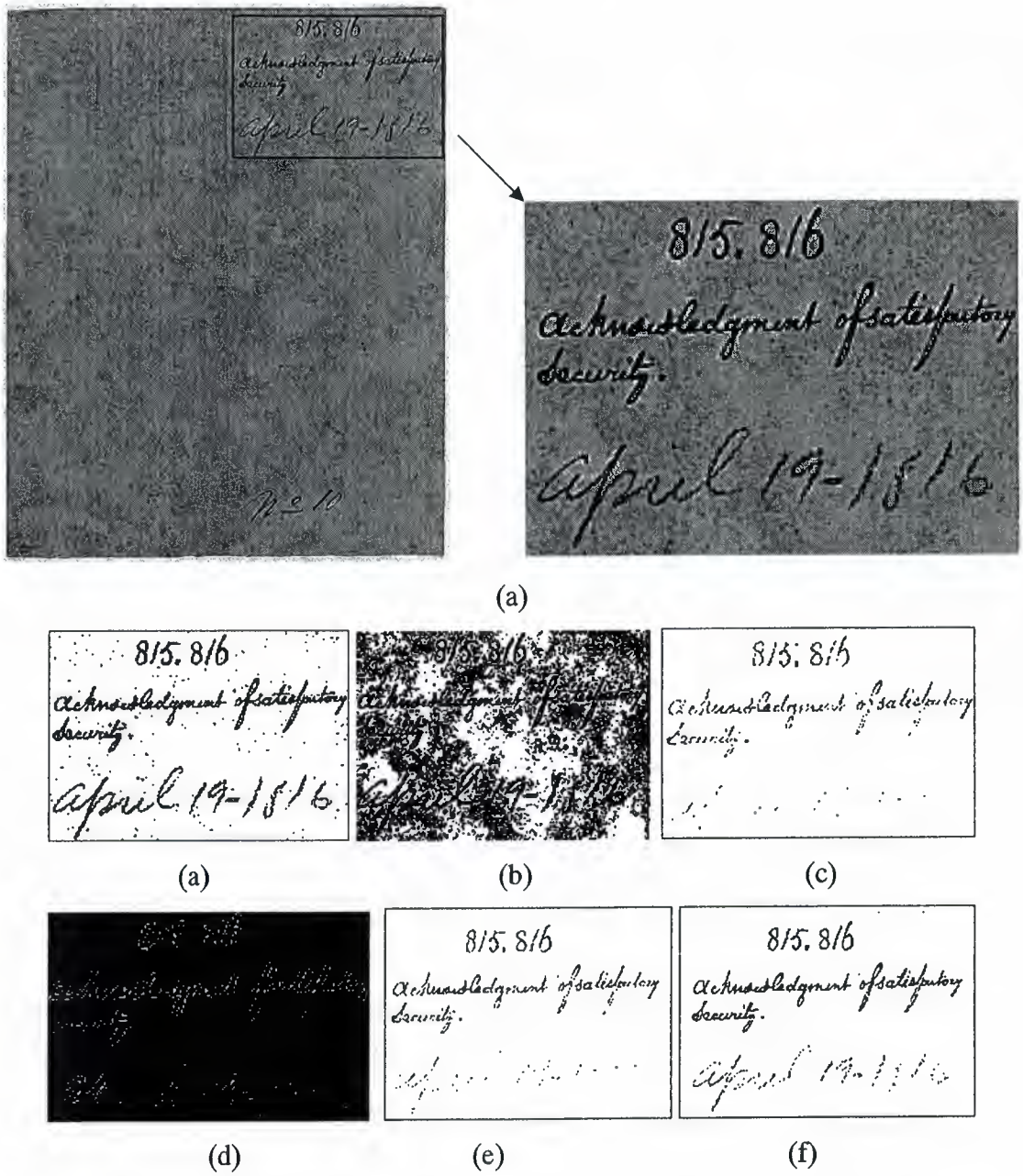


Figure 2.6 – Effects of Irrelevant Layers on Global Methods, (a) Original and Partial Image, (b) Kittler and Illingworth Method: produced some noises with clear characters, (b) Otsu Method: detect all irrelevant data as object, (c) Yanni and Horne Method: some loss of information and without noise, (d) Ramesh et al Method: almost all pixels are detected as object, (e) Albuquerque et al. Method: similar as Yanni Method, loss of information without noise and (f) Kapur et al. Method: little loss of information without noise.

In less degraded images which are generally comprise uniform and low-level of noise, global methods produce more efficient binarization than local methods. However, non-uniform backgrounds, high-level of noise, and another irrelevant layers such as smears, shadows ... etc. , may cause some global methods to produce complete loss of information or similar as local methods, the detection of these layers and noise as objects. In Figure 2.6, correctly and incorrectly thresholded images by global methods are demonstrated to show the effects of irrelevant layers in binarization process.

2.4 Local Binarization Methods

Local thresholding methods use different threshold values for segments within the image and several methods had been proposed to determine this local values. Most popular and conventional methods are Niblack Method [30], Sauvola et al. Method [47], Mean-Gradient Method of Leedham et al. [51], Pattern Averaging Thresholding (PAT) [58], Adaptive Logical Method [49], Bernsen Method [31], and Water Flow Model [54]. In next subsections, these methods will be explained in details and in Section 2.4.8, general advantages and disadvantages of local binarization methods will be presented.

2.4.1 Niblack Thresholding Method

Niblack method [30] is based on the varying threshold over the image, based on the local mean and local standard deviation. The threshold at pixel *(x, y)* is calculated as:

$$T(x,y) = \mu(x,y) + k.s(x,y), \tag{2.27}$$

where $\mu(x,y)$ and $s(x,y)$ are the sample and standard deviation values, respectively, in a local neighborhood of (x,y) and k is the value to adjust object boundary.

Basically, Niblack Method uses local mean and deviation that provides approximation of mean level by the amount of local deviation. However, increment or decrement of this mean value depends on constant k, where if $k>0$, mean value will be approximated to upper boundary, if $k<0$, mean value will be approximated to lower boundary and if $k=0$, mean value will become threshold point.

Operations of Niblack Method can be seen in Figure 2.7.

2.4.2 Sauvola et al. Thresholding Method

Sauvola et al. method [47] is an improvement of Niblack thresholding method aimed at producing better results with degraded documents. It can be described as:

$$T(i,j) = \mu(i,j) + \{1 + k.[1 - (\sigma(i,j)/R)]\} \tag{2.28}$$

where $\mu(i,j)$ and $\sigma(i,j)$ is the mean and variance value respectively in a local neighborhood of (i,j) with k and $R=128$.

Sauvola et al. improved Niblack method to add more adaptive local deviation to the mean value. However, the effect of local deviation is mostly eliminated by $\sigma(i,j)/R$ and threshold points mostly similar as mean value of defined segment.

Operations of Sauvola Method can be seen in Figure 2.8.

2.4.3 Mean-Gradient Thresholding Method

Mean-Gradient Method [51] was proposed by Leedham et al. and is the improved variant of Niblack's Method and based on local mean and local mean-gradient values. The gradient and the mean gradient of intensity images were defined as shown in Equation (2.29) and (2.30) respectively.

$$\nabla I(x,y) = [\partial I(x,y)/\ \partial x, \partial I(x,y)/\ \partial y] \tag{2.29}$$

$$G = \left(\sum_{x=0}^{i-1} \sum_{y=0}^{j-1} [\partial I(x,y)/\ \partial x, \partial I(x,y)/\ \partial y] \right) / x * y \tag{2.30}$$

Then, pre-condition was added to improve threshold selection:

$$\begin{aligned} &if\ \ c \geq R \quad T(x,y) = \mu(x,y) + k.G(x,y), \\ &else\ \ T(x,y) = 0.5m(x,y) \end{aligned} \tag{2.31}$$

where $k=-1.5$, $R=40$, $\mu(x,y)$ and $G(x,y)$ are the local mean and local-mean gradient of a kernel respectively and local contrast c was defined as shown in Equation (2.32):

$$c = Z_{\max} - Z_{\min} \tag{2.32}$$

where Z_{max} and Z_{min} represents maximum and minimum pixel values in a kernel.

Similar as Bernsen method, local contrast c is used as a initial criteria for threshold selection which if predetermined value R is smaller than c, threshold value determined as half value of mean of segment, otherwise threshold value determined by using Niblack's variant method which uses local gradient G instead of local deviation $\sigma(i,j)$.

Operations of Mean-Gradient Method can be seen in Figure 2.9.

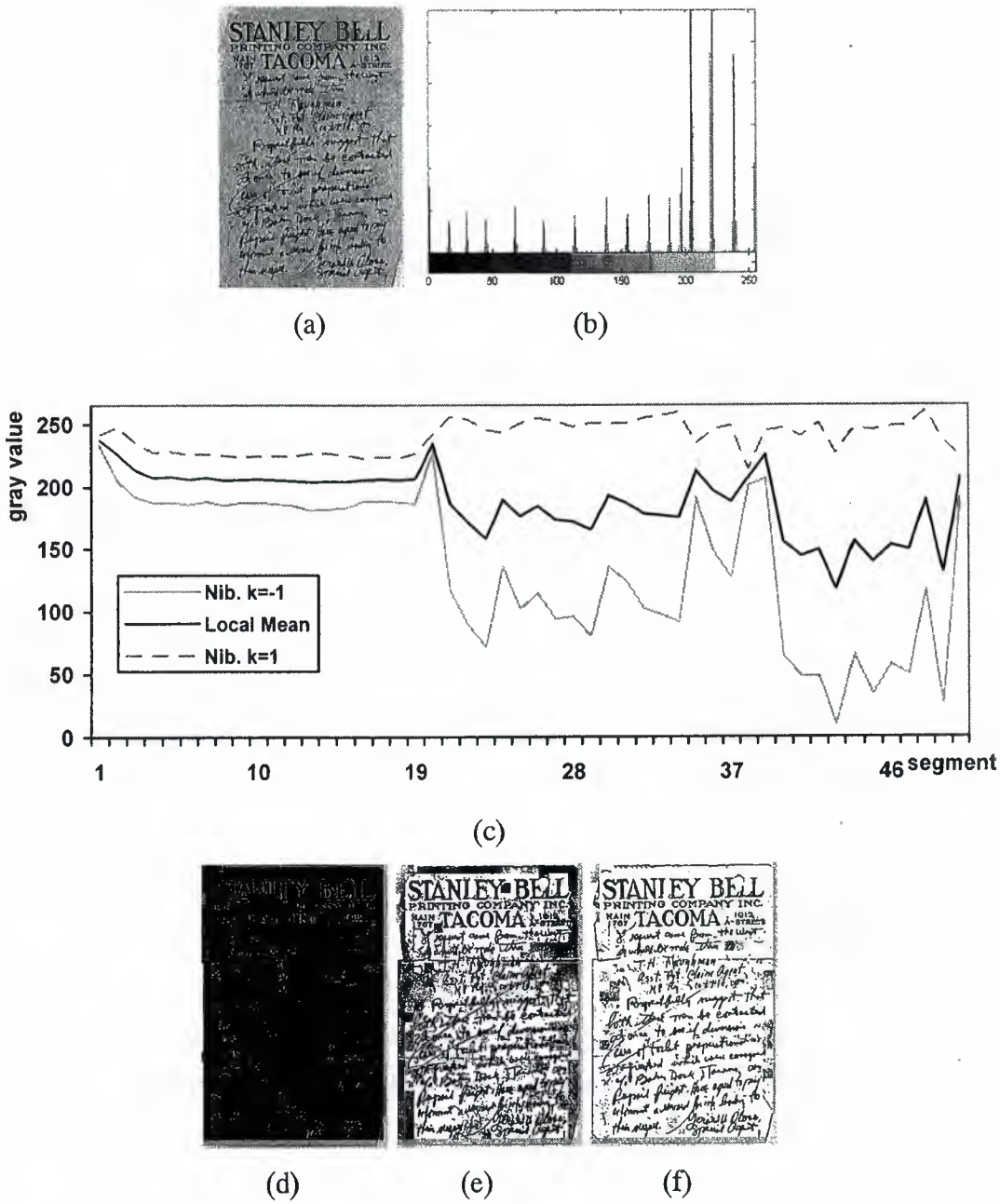
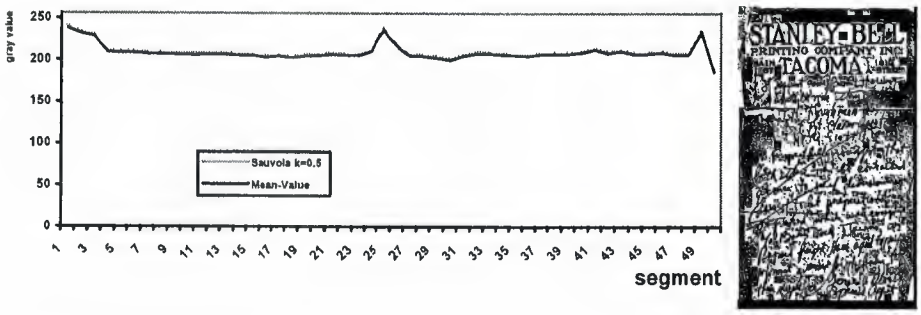


Figure 2.7 – Niblack Thresholding Operations and Examples of Approximation of Local Mean Values, (a) Original Image, (b) Corresponding Gray-level Histogram, (c) Threshold Values with different varieties of k, (d) Niblack Method with k=1, (e) Niblack Method with k=0 and (d) Niblack Method with k=-1.

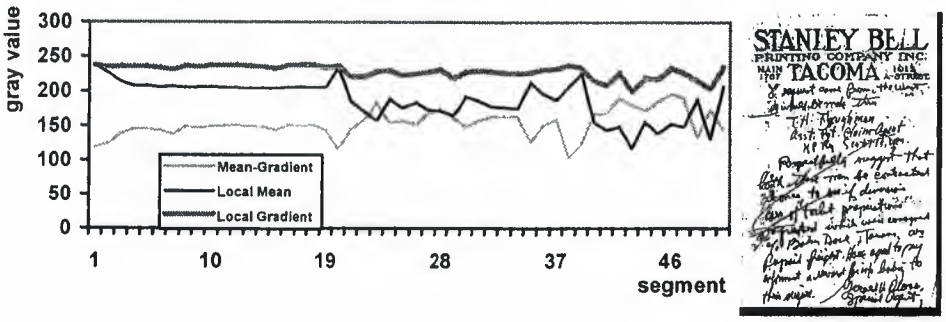
2.4.4 Pattern Averaging Thresholding (PAT) Method

PAT [58] is based on averaging pixel values within segments of a pattern, thus yielding one average pixel value per segment. The output pattern would contain averaged segment values and is defined as in Equation 2.33.



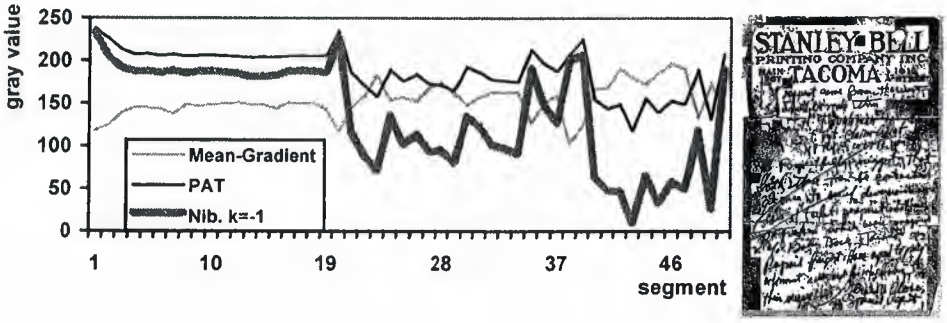
(a) (b)

Figure 2.8 – Sauvola et al. Thresholding Operations and Examples of Approximation of Local Mean Values, (a) Threshold Value of Sauvola et al. method with different k=0.5 and mean values, (b) Binarized image : Sauvola et al. Method, 12x12 kernel, k=0.5, R=128.



(a) (b)

Figure 2.9 – Mean-Gradient Thresholding Operations (a) Mean-Gradient Threshold, Local Mean and Local Gradient values, (b) Binarized image : Mean-Gradient Method, 15x15 kernel, k=-1.5, R=40.



(a) (b)

Figure 2.10 – Pattern Averaging Thresholding Operations (a) Mean-Gradient, Niblack and PAT values (b) Binarized image: PAT Method, 15x15 kernel.

$$I[x,y]=\begin{cases}255, if(\mu(x,y)\geq P[x,y])\\0, else.\end{cases} \tag{2.33}$$

where P *and* $\mu(x,y)$ denotes the original pixel value of an image and the PAT point of segment respectively; and I is the thresholded image.

Niblack, Sauvola and Mean-Gradient method uses local gradients or local deviations to approximate mean value of image to select optimum threshold for each segment. Instead of mean approximation, PAT directly uses mean value of segment as threshold point. Operations of PAT Method can be seen in Figure 2.10.

2.4.5 Adaptive Logical Thresholding (ALT)

ALT method [49] is the improvement of Kamel and Zhao logical level technique [39]. It is based on the idea of comparing the gray level of processed pixel or its smoothed gray level with some local averages in the neighborhoods about a few other neighboring pixels (Equation 2.34).

$$b(x,y)=\begin{cases}1 & if \vee_{i=0}^{3}\left[L(P_i)\wedge L(P_i^{'})\wedge L(P_{i+1})\wedge L(P_{i+1}^{'})\right]\\0 & otherwise\end{cases}, \tag{2.34}$$

where P_i , $P_i^{'}$, P_{i+1}, $P_{i+1}^{'}$ are four points of centered windows and $P_i^{'}$ and $L(P)$ were described as shown in Eq. 2.35.

$$P_i^{'}=P_{(i+4)\bmod 8,} for \quad i=0,....,7, L(P)=ave(P)-g(x,y)>T \tag{2.35}$$

where T is a predetermined parameter and $ave(P)$ was defined as:

$$ave(P)=\sum_{-sw\leq i\leq sw}\sum_{-sw\leq j\leq sw} f(P_x-i,P_y-j)/(2xSW+1)^2 \tag{2.36}$$

where P_x, P_y are the coordinates of P and $g(x,y)=f(x,y)$ and SW is a predetermined stroke width.

Yang and Yan improve this method to choose automatically SW and T which are stroke width and global parameter respectively by their developed algorithms.

2.4.6 Bernsen Method

Bernsen method [31] divides image into defined rxr segments than finds maximum Z_{max} and minimum Z_{min} gray level values within segment. After that, it measures local contrast c, by using these values and it compared by the predefined value d to determine if corresponding segment belongs to foreground or background layer. The formulae of Bernsen Method can be seen in Equation 2.37 and 2.38:

$$T(x,y) = (Z_{low} + Z_{high})/2 \tag{2.37}$$

$$c(x,y) = (Z_{low} + Z_{high}) < d \tag{2.38}$$

Thus, if local contrast is smaller than predefined value *d*, segment determined as foreground or background directly. Otherwise Equation 2.37 is used to determine threshold point for segment. Operations of Bernsen Method can be seen in Figure 2.11.

2.4.7 Water Flow Model

Kim, Jung and Park [54] proposed Water Flow Model (WFM) which is based on the property of water that always flows down to lower regions. In their method, it was assumed that lower gray levels denote characters and higher gray levels represent backgrounds. Rainfall fills the water in lower regions and it makes a lot of ponds in image terrain. After that, the final result is obtained by thresholding the amount of filled water using Otsu's thresholding method. The amount of water was defined as shown in Equation 2.39.

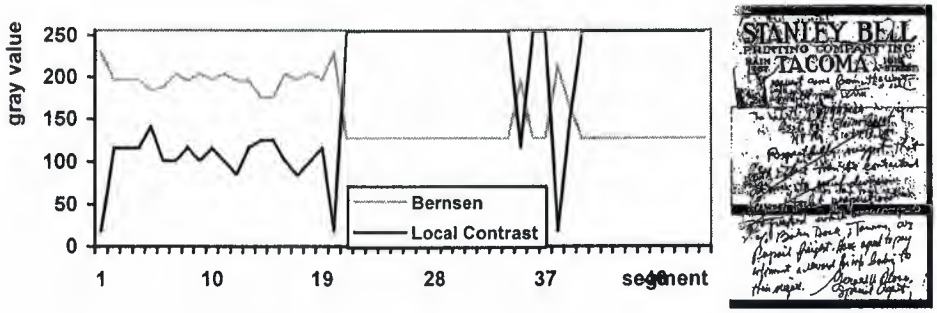
$$w_0 = \left\lceil \sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left(f_{top} - f(x,y)\right) \Big/ M.N \right\rceil \tag{2.39}$$

where M and N represent horizontal and vertical maximum distances of a terrain respectively, $f(x,y)$ denotes height of terrain and f_{top} denotes maximum level of terrain.

These explained seven local methods are also selected to perform comparison beside six global methods with proposed method in Chapter 3 and 4, because of their popularity and effective binarization in document binarization. Also recently proposed methods Adaptive Logical Method and Water Flow Model are included into comparison while they were mentioned by authors as the superior of local methods.

2.4.8 Advantages and Disadvantages of Local Methods

Local methods generally produce images with less loss of relevant information than global methods, however, the kernel size can be considered the main disadvantages of local methods; as additional noise may be added to images using small kernel sizes, whereas, large kernel sizes behave like global methods and thus can over or under threshold the image.



(a) (b)

Figure 2.11 – Bernsen Thresholding Operations (a) Local Contrast and Calculated Bernsen Threshold values for 15x15segments, (b) Binarized image : Bernsen Method, *15x15* kernel, *d=15*.

(a) (b) (c)

(d) (e) (f)

(g)

Figure 2.12 – Binarization of Figure 2.6 Image by Local Methods, (a) Niblack Method (15x15): most of the irrelevant layers detected as object, (b) Sauvola et al. (12x12): irrelevant layers and objects are mixed up together, (c) Mean-Gradient Method (15x15) :some of the irrelevant layers detected as object, (d) PAT (15x15): similar as Sauvola method, (e) ALT (8x8) : correctly detection of characters with some noise, (f) Bernsen Method (15x15): initial criteria affects some kernels to the detect irrelevant layer as object and (g) Water Flow Model (5x5 with w=17): some loss of characters with noise.

Beside these, if kernel does not consist of any interest object within, small changes in pixel values can affect the local methods to detect these pixels as objects.

In Figure 2.12, correctly and incorrectly thresholded images by local methods are demonstrated to show the effects of kernels in binarization process.

2.5 Application Areas of Image Binarization

Binarization methods can either be used for low level process of any system to prepare preliminary data or lonely as a system to produce required results, and simplicity and efficiency of image binarization methods make them popular almost for every field of science and industry. However, the main application areas of image binarization are pattern recognition, biometrics, medical imaging and document analysis and understanding. In next subsections, the aims of binarization in these fields and performed researched will be presented briefly.

2.5.1 Image Binarization in Pattern Recognition

Pattern Recognition applications mostly require low level processing of data to prepare effective input or training data for recognition systems and also they can be used to reduce the amount of unnecessary data. In oceanic applications, recognition of plankton images [60-62] and classification of sea-ices [63], use binarization as an initial process to prepare effective data for recognition or classification systems. In industrial applications, such as recognition of coin patterns, thresholding is used to reduce the amount of data sent to neural network that provides increment of recognition rate and decrement of time of training [64-68].

2.5.2 Image Binarization in Biometrics

In Biometrics, relevant data varies to applications. In fingerprint recognition, thresholding applications are mostly used to enhance and binarize the images to provide optimal effective and clean data for recognition systems. Thus, several researches recently had been performed for this purpose [69-72]. Also, thresholding applications had been used in face skin detection and retinal vessel recognition in personal authentication and security fields of biometrics [73-74]. In spite of these researches, binarization does not become popular tool for biometrics because little loss of

information may cause any biometrical system not to produce effective results for security or other fields that requires high level of accuracy.

2.5.3 Image Binarization in Medical Imaging

In medical imaging, the aim is to extract or enhance the region of interest within image. Thus, binarization techniques are frequently used to achieve this aim. However, application dependency of proposed thresholding methods, are mostly cause the use of different methods in different medical applications. Any method that produces optimum results for Magnetic Resonance Images may not produce effective results for X-Ray images. In spite of this disadvantage, thresholding methods are used almost in every field of medical imaging. Segmentation and enhancement of microscopic images [75-76], radiotherapy imaging applications [77], enhancement of x-ray, MRI and ultrasound images and detection of tumor cells [78-88] are the some of main applications of thresholding in medical imaging.

2.5.4 Image Binarization in Document Analysis and Understanding

Document analysis and understanding is the most popular application area of image binarization where effective separation of characters that belongs to foreground are required while discarding irrelevant information of background layer. However, variety of layers such as smears, smudges, shadows, acquisition noises, illumination effects, within documents make challenging task for binarization methods. Development of new techniques and comparison of conventional or recently proposed methods are still performing by researchers and this also makes another challenging task which is the determination of optimum method for document images. Therefore, similar as other application areas, application-dependency of methods makes some methods optimum for some images, and another for other images.

2.6 Summary

Binarization is low level process of image processing that is used for segmentation or enhancement of the region of interest within images. Several methods that are classified into local and global groups had been proposed. However, each method brings some disadvantages beside their apparent advantages.

In this chapter, basic definitions of image binarization, chronological development, detailed explanation about selected thirteen methods and application areas were presented.

In the next chapter, proposed global Mass-Difference Thresholding method, performed experiments and preliminary results will be explained.

CHAPTER 3
MASS – DIFFERENCE THRESHOLDING

3.1 Overview

In this chapter, basic characteristics of document images and proposed Mass-Difference (MD) Thresholding Method will be presented which was designed to extract dominant pixels within document that are categorized as 'text/foreground' by using maximum brightness point as the background limit of document and mean value as 'fuzzy layer'. Also performed experiments and preliminary experiments will be introduced.

3.2 Proposed Method

The proposed Mass-Difference (MD) method is a global single-stage thresholding method that finds a single threshold value using the global maxima (luminance value) and the mass average (mean of the intensities) of an image. It was designed by considering the characteristics of document images which generally consists darker foreground information than background.

3.2.1 Basic Characteristics of Document Images

Investigation of the characteristics of document images and their corresponding histograms showed that, pixel values of region of interest which is the text in document images, always smaller than the mean value of image, however the distance of this region from the mean point is not constant. Thus, shifting the mean position to a point that separates background and foreground could provide efficient separation of this region. Luminance value of image is the brightest pixel value within image and it defines the maximum limit of background in either uniform or non-uniform conditions. Thus considering the deviation of this limit from the mean point of image and symmetrically shifting the mean point by the amount of this deviation provides exact separation of text from background.

3.2.2 Mathematical Expression of the Proposed Method

A previously developed basic global thresholding method that resembles MD method is background-symmetry algorithm, which assumes a distinct and dominant peak for the background that is symmetric about its maximum, and defined as:

$$\theta = \max p - (p\% - \max p), \tag{3.1}$$

where *maxp* is the maximum peak value in the histogram and *p%* is the non-object pixel side of that maximum. However, in background-symmetry algorithm, thresholding methods may need some adaptation if brightness histogram has been changed.

MD thresholding method is different from the background-symmetry algorithm in that the maximum value is defined as the highest pixel value within the image, whereas in background-symmetry the maximum peak is found by searching for the maximum value in the histogram.

MD uses the deviations between mass average (mean value) and global maxima (luminance value). The Mass of an image, which yields center of "fuzzy area" of image, can be defined as:

$$\mu = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} I(x,y), \tag{3.2}$$

where μ represents the mass (mean) of image, n and m denotes the x and y dimensions of image respectively, and I represents the original grayscale image. The Global Maxima or maximum brightness / luminance L of an image which determines background limit of document is defined in Equation (3.3) as a maximum function of original grayscale image.

$$L = F_{\max}(I). \tag{3.3}$$

After the calculation of mass and global maxima, the Local Deviation (D) of the Mass *(μ)* from the Global Maxima *(L)* is calculated to determine the amount of shifting "fuzzy layer" to the left of the image histogram and it was defined as:

$$D = L - \mu. \tag{3.4}$$

Finally, mass is shifted to left by the amount of local deviation to determine The Total Deviation (T) which represents the Optimum Threshold Value and it is defined as the difference of local deviation and the mass of the image (Equation 3.5). Absolute value of the difference of local deviation and the mass of the image is considered to avoid the negative threshold values in the cases of the smaller mass of the images than the local deviation. Figures 3.1-3.4 show the Mass-Difference process in details.

$$\mathbf{T} = |\mu - D|. \tag{3.5}$$

Equation 3.5 can also be written as follows by replacing D:

$$T = |\mu - (L - \mu)|. \tag{3.6}$$

An MD enhanced image (*MI*) is obtained using:

$$MI[x,\ y] = \begin{cases} 0, & if \quad I[x,y] \le T \\ 255, & else. \end{cases} \tag{3.7}$$

Various experiments had been performed to test the behavior of the proposed method in different conditions. Figure 3.5 shows the MD experiment on non-uniform background of a clean image and Figure 3.6 shows the proposed method in artificially created text with bimodal histogram and different text color respectively which any of the other considered global and local methods could not produce effective separation for this image. Figure 3.7 shows the binarized images of Figure 3.6 (a) and (b) by using other global methods and Figure 3.8 shows the binarized images of Figure 3.6 (a) and (b) by using local methods.

If any noise occur within image which represented only by a pixel with a value of maximum gray level *255*, background can be determined as non-uniform and the bit affects the amount of shifting value, thus the behavior of proposed method also tested in this condition that can be seen in Figure 3.9.

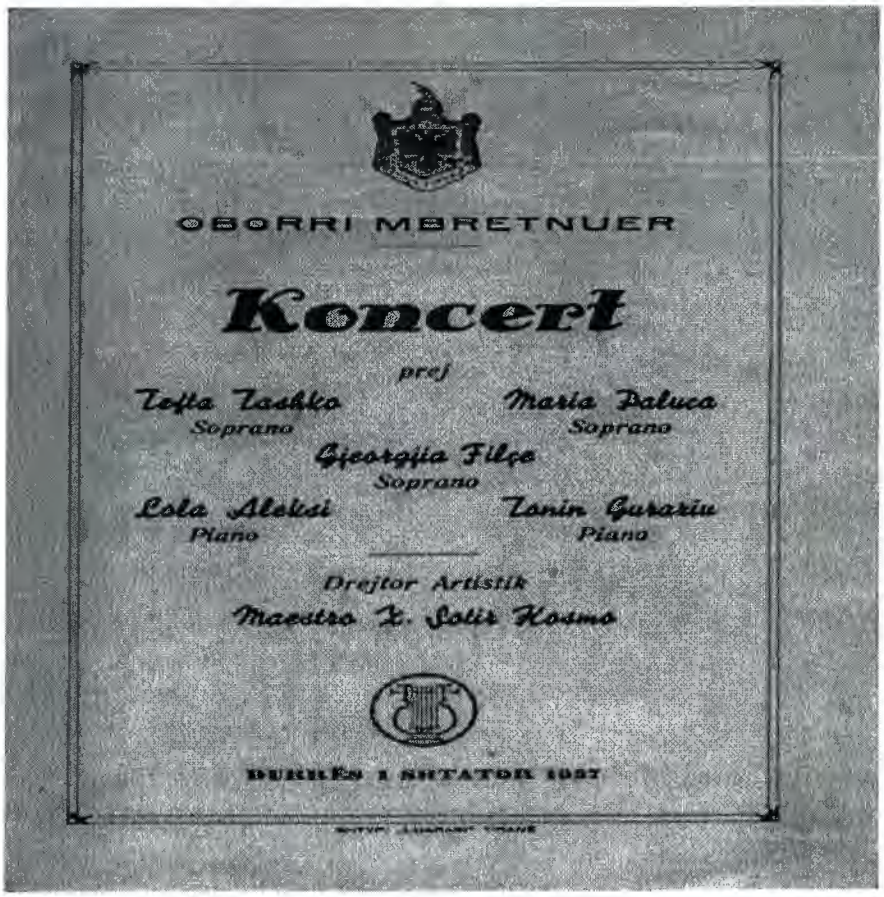


Figure 3.1 Example Image.

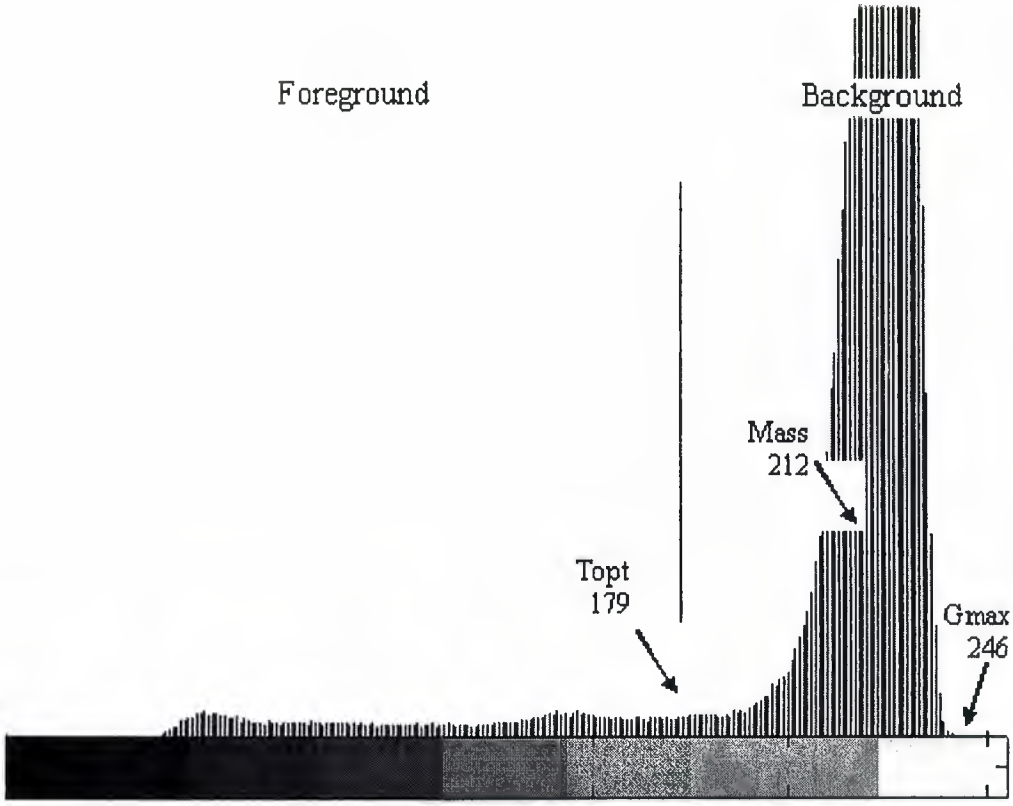


Figure 3.2 Corresponding Histogram and MD Operations on Image Fig 3.1.

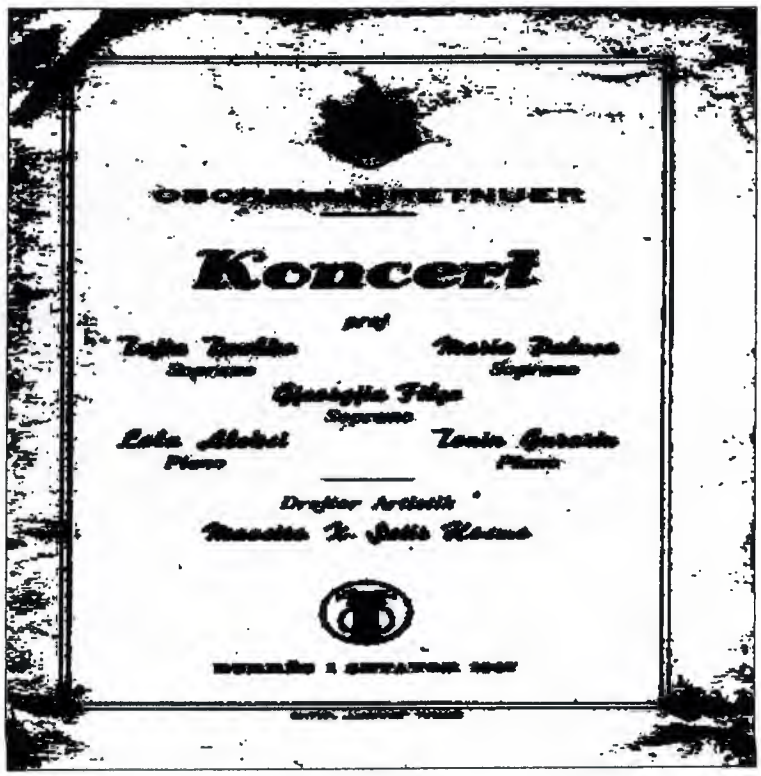


Figure 3.3 Thresholded Example of Image by using Mass Value.

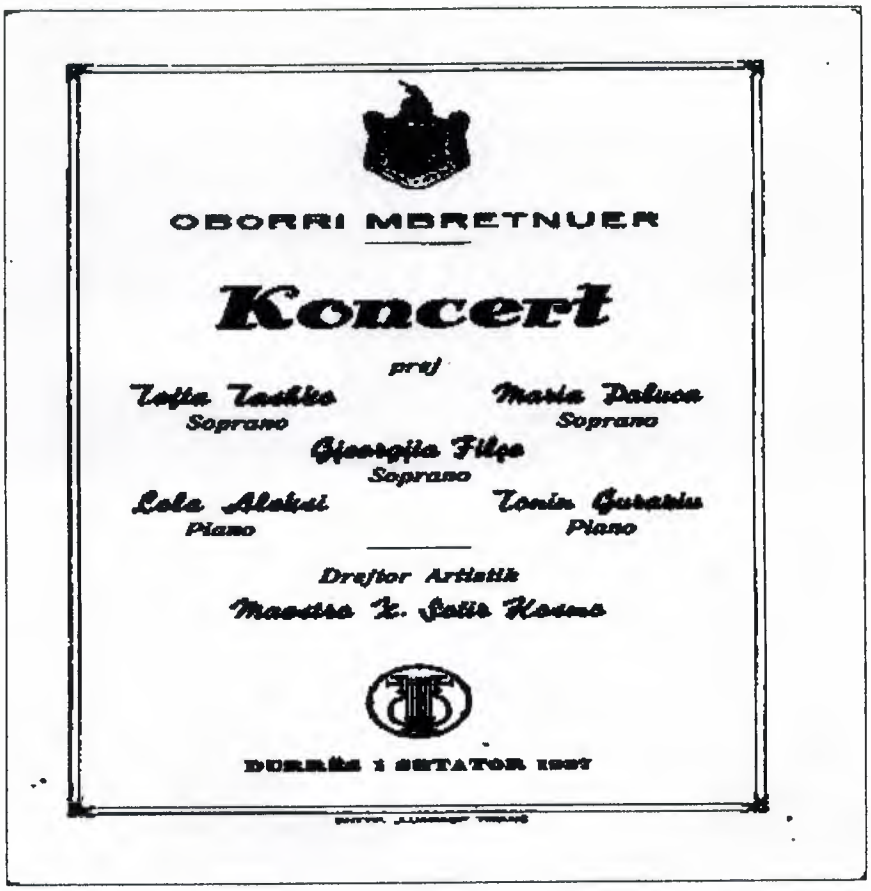
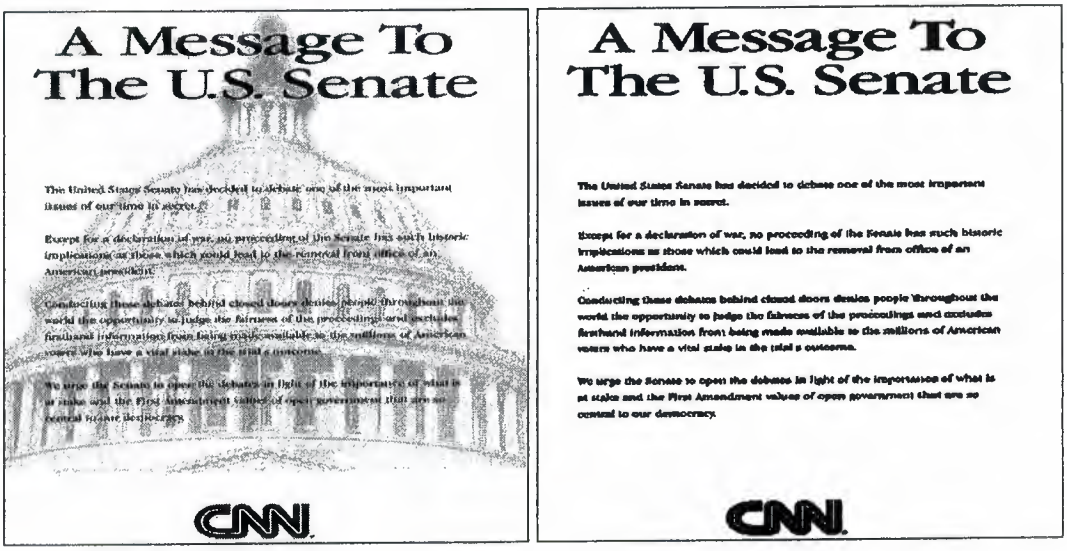


Figure 3.4 – Thresholded Example of Image by MD.



(a) (b)

Figure 3.5 Thresholding Example using proposed method (a) Original non-uniform background image and (b) Correctly thresholded image by MD Thresholding.

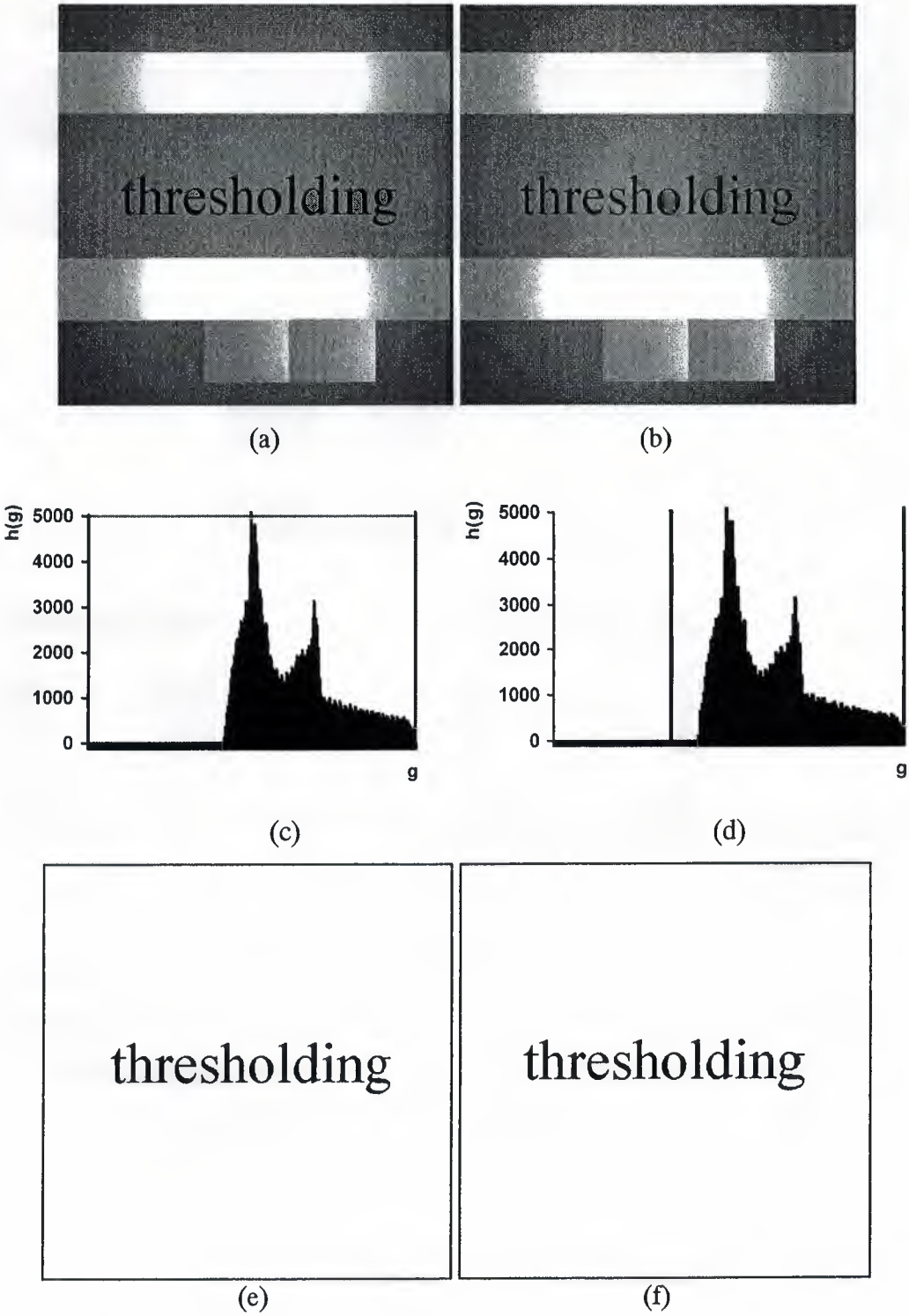


Figure 3.6 Testing of Proposed Method in Bimodal Images, (a)-(b) Artificially created document non-uniform background and illumination with darker and lighter text respectively, (c)-(d) Corresponding histograms of (a) and (b) respectively and (e)-(f) Binarized images using proposed method of (a) and (b).

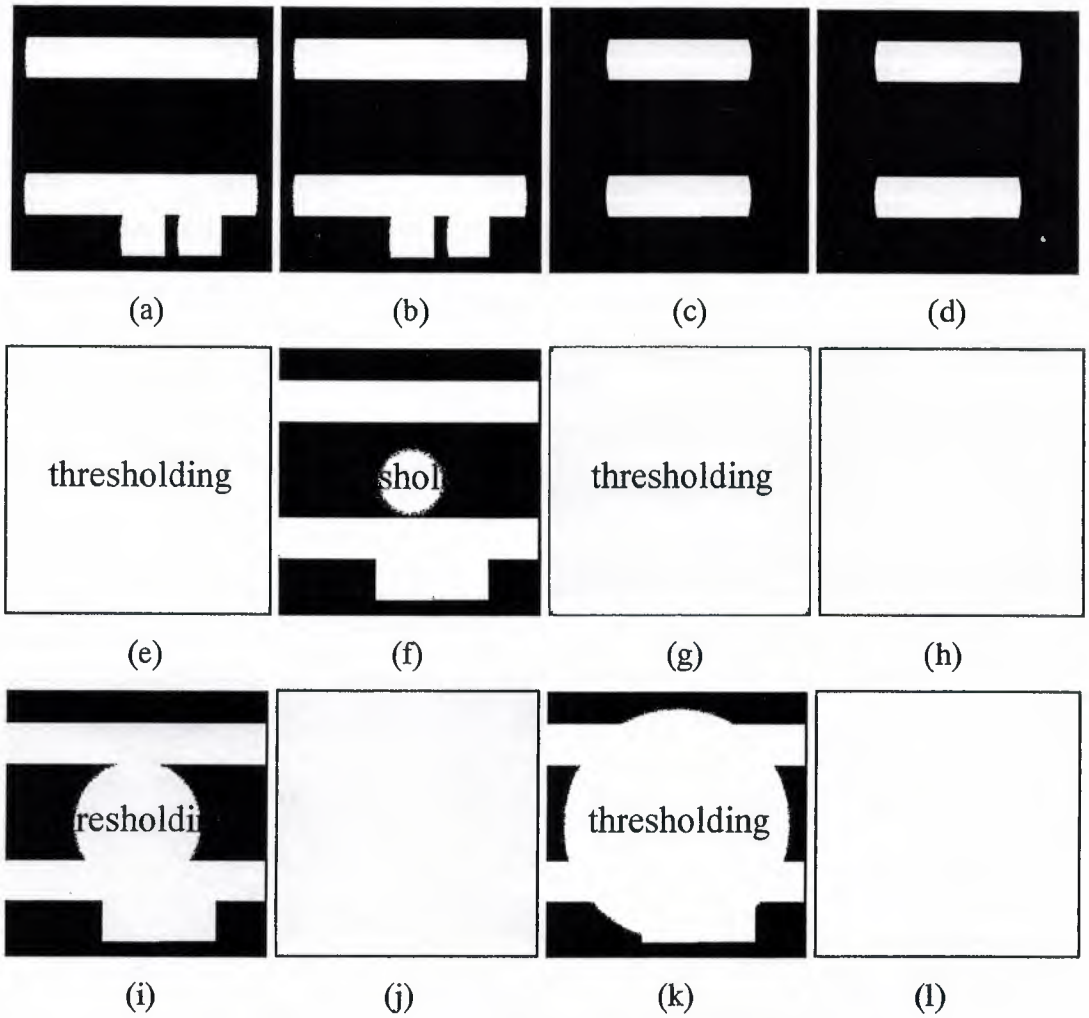


Figure 3.7 Binarization of Fig. 3.6 (a) and (b) images by global methods, (a)-(b) Binarization by Otsu Method of Fig. 3.6 (a) and (b) respectively, (c)-(d) Binarization by Kittler and Illingworth Method of Fig. 3.6 (a) and (b) respectively, (e)-(f) Binarization by Yanni and Horne Method of Fig. 3.6 (a) and (b) respectively, (g)-(h) Binarization by Ramesh et al. Method of Fig. 3.6 (a) and (b) respectively, (i)-(j) Binarization by Kapur et al. Method of Fig. 3.6 (a) and (b) respectively and (k)-(l) Binarization by Albuquerque et al. Method of Fig. 3.6 (a) and (b) respectively.

3.3 Preliminary Experiments on Proposed Method

Two preliminary experiments had been performed to test the success and efficiency of proposed method. In next subsection, these two preliminary experiments will be explained briefly.

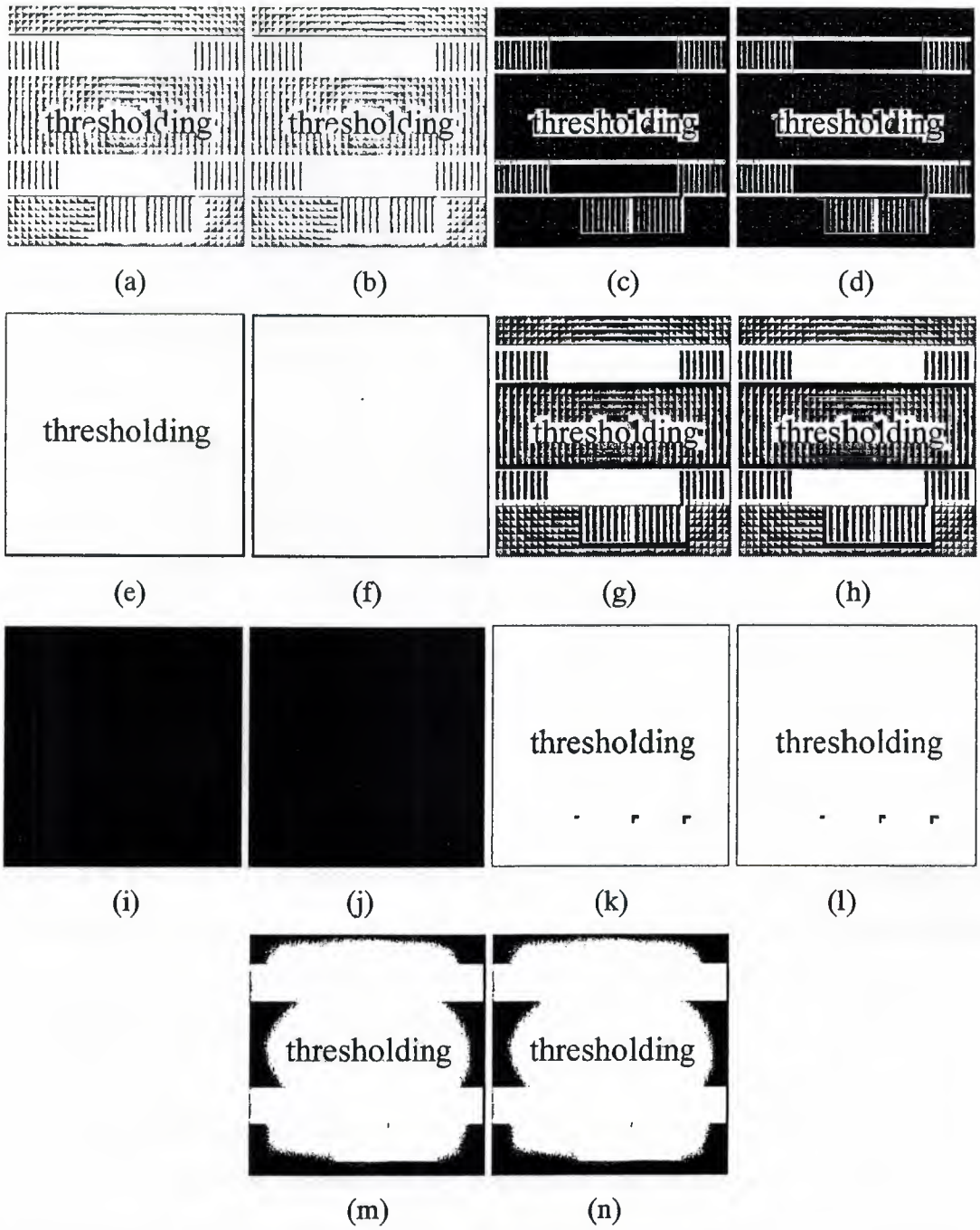


Figure 3.8 Binarization of Fig. 3.6 (a) and (b) images by local methods, (a)-(b) Binarization by Niblack Method of Fig. 3.6 (a) and (b) respectively, (c)-(d) Binarization by Sauvola et al. Method of Fig. 3.6 (a) and (b) respectively, (e)-(f) Binarization by Mean Gradient Method of Fig. 3.6 (a) and (b) respectively, (g)-(h) Binarization by PAT method of Fig. 3.6 (a) and (b) respectively, (i)-(j) Binarization by Bernsen Method of Fig. 3.6 (a) and (b) respectively, (k)-(l) Binarization by ALT Method of Fig. 3.6 (a) and (b) respectively and (k)-(l) Binarization by Water Flow Model of Fig. 3.6 (a) and (b) respectively.

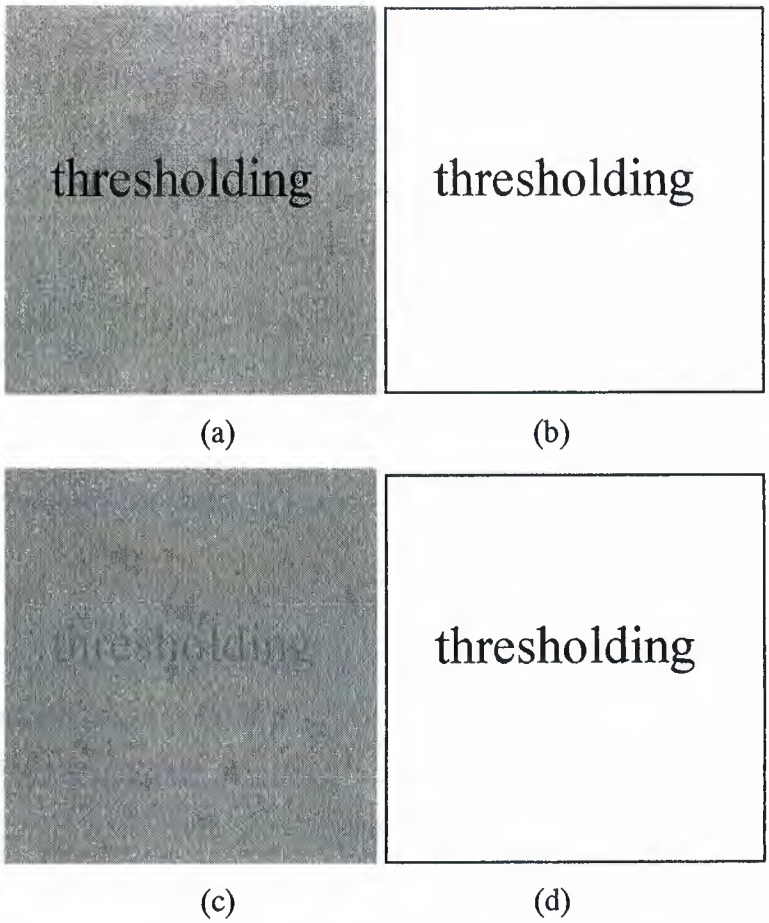


Figure 3.9 Test of MD Behavior with Single Noisy Luminance value, (a) Artificially Created Text Document with darker text : $\mu=218$, $L= 255$, (b) Binarized image of (a) by MD Method, (c) Artificially Created Text Document with lighter text : $\mu=222$, $L= 255$, (d) Binarized image of (c) by MD Method.

3.3.1 Preliminary Experiment I[1]

Preliminary Experiment I comprises the binarization of 30 document images by four global methods, Otsu, Kittler and Illingworth, Kapur et al., Quadratic Integral Ratio, beside MD, and a local method Parker.

The efficiency of these methods was determined by the recognition rate of words in the thresholded or segmented document images. Visual inspection of the documents was determined the number of recognized and readable words in a thresholded document. This was implemented by 15 independents persons. The total number of the words in the original document and the number of recognized and readable words after thresholding were used to determine the recognition rate.

[1]Preliminary Experiment I was published in IEEE International Conference on Industrial Technology, India, on December 2006 as "Novel Thresholding Method for Document Analysis".

Thirty images that comprises totally 1205 words, were divided into three groups as Clean, Degraded and Highly Degraded documents and results showed that MD was superior method which was followed by Otsu and Kapur et al. Methods. Table 3.1 shows the obtained results by implementation of 30 document images.

Table 3.1 Recognition Rates of Preliminary Experiment I

Method	Clean	Degraded	Highly Degraded	Total	Recognition Rate
Otsu	300/325	335/450	290/430	925/1205	77 %
Kapur et al.	289/325	325/450	175/430	789/1205	65 %
Parker	252/325	250/450	235/430	737/1205	61 %
Kittler and Illingworth	199/325	60/450	35/430	294/1205	24 %
QIR	215/325	90/450	80/430	385/1205	32 %
MD	**308/325**	**350/450**	**300/430**	**958/1205**	**80 %**

3.3.2 Preliminary Experiment II[2]

Preliminary Experiment II comprises the binarization of 55 images by using Otsu Method, Kittler and Illingworth method as global methods, and Niblack Method as local method. The implementations were carried out using two image sets of documents. The first set comprises 50 historical documents which contain a total of 2021 words with different contrasts and brightness. The second set comprises 5 created words which contain a total of 45 characters with different backgrounds and different color to test the occurrence of the characters after thresholding.

Similar as Preliminary Experiment I, the efficiency of these methods was determined by the recognition rate of words in the thresholded or segmented document images by 15 independents persons for two sets separately.

A general comparison is then performed by combining the results using the two sets of documents. General results are categorized as: *recognized* or *unrecognized* words in Set 1, and *clear* or *unclear* characters in Set 2.

Table 3.2 Recognition Rates of Words in Set 1 of Preliminary Experiment II

Thresholding Method	Total Words	Recognized Words	Unrecognized Words	Recognition Rate
Otsu	2021	1657	364	81.98 %
Kittler and Illingworth	2021	1045	976	51.70 %
Niblack	2021	1699	322	84.06 %
Proposed Method	2021	1730	291	85.60 %

[2]Preliminary Experiment II was published in 11th Panhellenic Conference on Informatics, Patras, Greece, on May 2007, as "A Novel Thresholding Method for Text Separation and Document Enhancement".

Results showed that MD was superior method which was followed by Niblack method. Table 3.2 and 3.3 shows the obtained results by implementation of Set I and Set II of Preliminary Experiments II.

Table 3.3 Recognition Rates of Characters in Set 2 of Preliminary Experiment II

Thresholding Method	Total Characters	Clear Characters	Unclear Characters	Recognition Rate
Otsu	45	40	5	88.88 %
Kittler and Illingworth	45	45	0	100 %
Niblack	45	45	0	100 %
Proposed Method	45	45	0	100 %

3.4 Summary

In this chapter, basic characteristics of document images, proposed Mass-Difference (MD) Thresholding Method and performed experiments in different conditions, were presented. Also preliminary experiments and comparisons were explained briefly.

Next chapter explains performed comparative evaluation of binarization methods which were considered in Chapter 2 and proposed method, for document image enhancement. Also document image database and evaluation criteria will be explained.

CHAPTER 4
COMPARATIVE EVALUATION OF THRESHOLDING METHODS FOR DOCUMENT IMAGE BINARIZATION[1]

4.1 Overview

Digitized document analysis has recently become more significant with the advances in digital archiving and electronic libraries. Scanned document images, specially historical and handwritten documents, generally carry various levels of noise because of the age, paper, pen and pencil influences on the documents. Efficient enhancement of scanned paper-based documents is usually required prior to further processing. The efficiency of document image enhancement depends on the efficient separation and classification of background and foreground layers. Thus, the initial purpose of document analysis techniques is the effective preparation and separation of various layers in documents in order to provide sufficient and clear data for recognition systems and human readers. In ideal conditions, optimistic separations of relevant data which belong to foreground layer of texts and characters is required while discarding unnecessary information of background layer. However, document images generally carry various levels of noise which causes pessimistic separation of the layers. Such variety in noise levels creates a challenging task for thresholding methods and may prevent efficient separation and enhancement of document images.

In this chapter, performed experiments on document image binarization, design of experiments, evaluation criteria and obtained results will be explained in details by performing comparative evaluation of considered methods.

4.2 Recent Comparisons

With the existence of many global and local thresholding methods, deciding upon an optimum method for document image binarisation is a challenging task; because efficiencies of the existing thresholding methods are usually application-dependent where one method's performance appears superior when using a certain type of document, but the test on a different type of document. The solution to this problem would be in creating and using a comprehensive multi-applications document image

[1] The content of this chapter was submitted to Journal of Image and Vision Computing, Elsevier Science, on September 2007.

database that accounts for different types of documents, such as historical documents, degraded documents, artificially created words, and handwritten documents.

Several comparisons have been previously performed [51, 89-92] in order to evaluate existing thresholding methods and deciding upon an optimum thresholding method for document binarisation in particular.

Trier and Taxt [89] compared fifteen known thresholding methods which included eleven local thresholding methods and four global thresholding methods. The local methods were: Bernsen [31], Chow and Kaneko [23], Eikvil et al. [36], Mardia and Hainsworth [93], Niblack [30], Taxt et al. [35], Yanowitz and Bruckstein [34], Parker [94], White and Rohrer Dynamic thresholding [26], White and Rohrer Integrated Function Algorithm [26], and Trier and Taxt method [95]. Whereas, the global methods were: Abutaleb [33], Kapur et al. [28], Kittler and Illingworth [19], and Otsu [21]. Yanowitz and Bruckstein [34], and White and Rohrer [26] used post-processing in their methods to remove ghost objects. Similar post-processing was also adopted by Bernsen [31], Eikvil et al. [36], Niblack [30], and Parker [94] in order to compare and evaluate modified versions of the two methods in [34,26]. The evaluation strategy in [89] was based on the visual inspection of binarized images of cable and hydro images, where an expert would visually inspect the images and evaluate the results according to the five criteria which were defined as: the broken line structures, broken symbols, text ...etc., blurring of lines, symbols and text, loss of complete objects, and noise in homogenous areas. In their evaluation, Trier and Taxt [89] concluded that thresholding methods with post-processing produced better results than others, and the optimum thresholding methods in local and global groups were modified Niblack method, and Otsu method, respectively. However, in their general evaluation and considering the overall performance of these methods, the modified Niblack method was considered superior to others while Otsu method was ranked as 13th in the general evaluation. They also concluded that Kittler and Illingworth method had the least successful overall result.

In a different work, Trier and Jain [90] performed a similar comparison using the same methods as in [89] but differing in that; their evaluation was based on the recognition of the characters within binarized images by an Optical Character Recognition (OCR) module. The results of this comparison and evaluation suggested similar ranking of the thresholding methods as found in [89]; where modified Niblack was proposed again as superior to the other methods in overall performance. However,

Trier and Jain noticed that the pre-processed Niblack method has a disadvantage with the size of neighborhood, where the size should be small enough to preserve local details and large enough to suppress noise.

A more recent comparison of thresholding methods was performed by Leedham et al. [51]. They compared an improved version of Niblack method, Yanowitz and Bruckstein method, and Solihin and Leedham Quadratic Integral Ratio [48] method with their proposed locally adaptive thresholding methods: Mean-Gradient technique and Background subtraction. Their evaluation was based on recall, precision and accuracy, as suggested in [96], and was performed on 40 images which were: equally distributed historical handwritten images, cheque images, form images, and newspaper images. Consequently, Leedham et al. concluded that there is no single efficient algorithm for all types of images.

Sezgin and Sankur [91] performed another comparison of 40 selected thresholding methods using two different image databases: degraded document images and nondestructive testing images. The degraded document images database consisted of 40 created document images with different fonts, sizes and typefaces. In order to obtain degradation of these documents, poor quality photocopied and faxed documents were used, and blur masks and speckle noise were added to these images. The evaluation procedure in [91] was based on five performance criteria which were: misclassification error, edge mismatch, relative foreground area error, modified Hausdroff distance and region non-uniformity. The ranks of these criteria were averaged for each image and used to measure their performance. Sezgin and Sankur noticed that, Kittler and Illingworth Minimum Error thresholding method was superior to all other global and local thresholding methods in their evaluation using both image databases. Using document image database, Kittler and Illingworth method was followed by Sauvola and Pietikainen method; where the later method did not perform very well when using the nondestructive database.

Recently, He et al. [92] performed a comparison of thresholding methods using historical documents as their database. They used a single global thresholding value of 165 for the database to simulate global effect and compared the results to Niblack, and Sauvola and Pietaksinen methods. They also modified these methods to choose constant k and kernel sizes sw automatically and called their modifications as Adaptive Niblack and Adaptive Sauvola methods. The evaluation was carried out on a set 4435 images

and was performed on only one typed word with dimensions 23 to 25 pixels on either grey or yellow background with black or red typing. The binarized images were then fed into OCR to determine the recognition rate of characters. He et al. concluded that, Niblack and Adaptive Niblack had slightly better performance than others.

The above described works on comparative evaluation of thresholding methods have attempted to suggest an optimum thresholding method that can be efficiently used for document image binarisation. However, the results of these different evaluations suggested different methods as being optimum; which is anticipated as the image databases differ from one evaluation to another; where one evaluation uses historical documents, others use created words, or artificially degraded document scans. Another problem is the insufficient number of document images used in some of these evaluations [89-91], which affects the significance of the evaluation outcome. In addition, using a large number of images that have similar noise and layer characteristics [92], does not provide an effective evaluation. Moreover, the use of visual inspection as in [89], without any computed analysis, as the only or main criteria for evaluation may not provide a robust evaluation outcome. On the other hand, the use of OCR module with some historical documents is not possible due to old different fonts that can not be recognized by the available OCR modules. Finally, there is a lack of clear categorisation of thresholding methods into adaptive local methods and global methods when performing the evaluations. Such clear categorisation would greatly aid in providing a more objective comparison and in suggesting an overall optimum thresholding method or a category-based optimum thresholding method. The work presented within this chapter aims to solve these problems.

4.3 Design of Experiments

In order to provide an efficient evaluation, large document image database, that includes historical documents, degraded documents, artificially created words, and handwritten within bright, low contrast images, is required. Clean document images can be used for evaluation, however, in the case of historical documents, extreme conditions such as shadows, non-uniform illumination, low contrast, signal dependent noise, smear and smudge may cause pessimistic foreground separation, thus document image database that used in experiments, include these degraded document images in order to perform a more effective evaluation.

The evaluation procedure should also be capable of performing accurate comparison of considered methods, thus, the use of both visual inspection and computed analysis based on two metrics which we derived using the Peak Signal-to-Noise Ratio (PSNR) of the enhanced images are proposed. The following subsections present the document image database and the evaluation procedure.

4.3.1 Document Image Database

The database consists of 174 historical documents, handwritten documents, and specially created text; which are organized into three sets:

4.3.1.1. Image Set I

This set comprises 115 scanned historical documents which contain a total of 10291 words with different contrasts and brightness. This set has been classified, according to the images' corresponding grey level histograms [2]; into three groups, namely Bright, Low-Contrast and Dark images. Classification was performed by considering grey level histograms of the images.

The Bright image group consists of 77 degraded images which have different types of noise, shadows and smudge. The Low-Contrast image group consists of 33 marginally-degraded images which also have different types of noise, shadows and smudge. The Dark image group has two images which contain additional noise. Examples of these historical document images and their corresponding grey level histograms can be seen in Figure 4.1 – 4.3 and other examples of Bright and Low Contrast images of Set I can be seen in Figure 4.4.

4.3.1.2. Image Set II

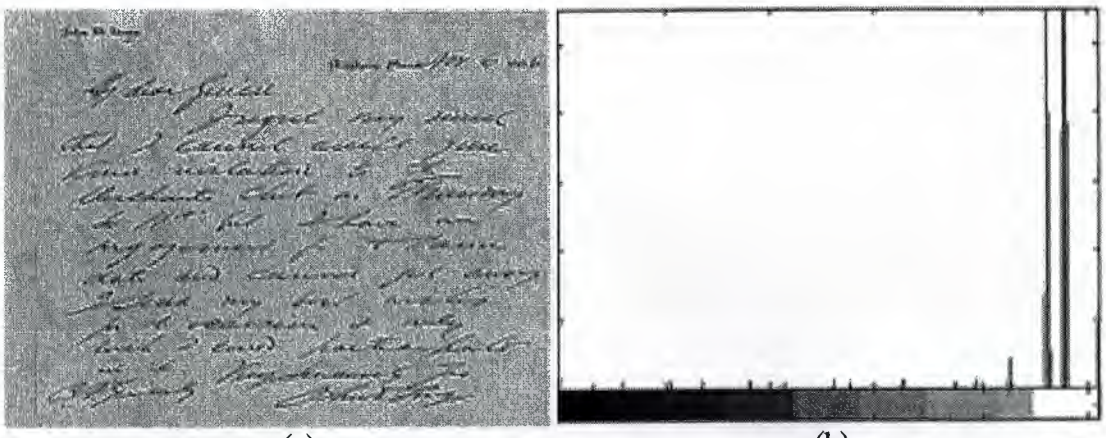
This set comprises six specially created words which contain a total of 72 characters with different backgrounds and different grey scales (non-uniform text). Examples of this set can be seen in Fig. 4.5.

4.3.1.3. Image Set III

This set comprises 54 scanned handwritten documents. The documents were prepared in by scanning the handwriting of nine different persons, using three different writing tools

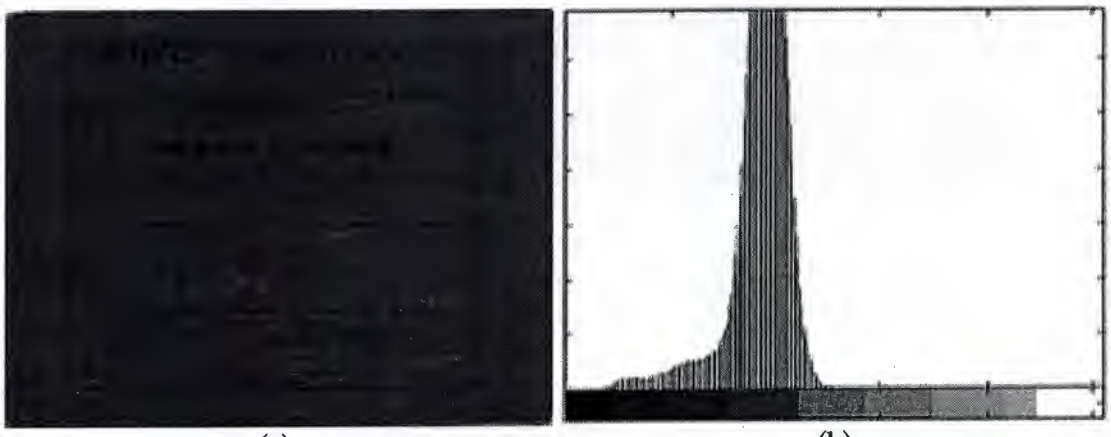
(pen, pencil and board-marker) on two different paper types (white paper and yellow envelope paper). Examples of the handwritten documents can be seen in Fig. 4.6.

In summary, the above three sets, which contain multi-application document images with different levels of noise and contrast, non-uniform illumination (background), signal-dependent noise, smears and non-uniform foreground (text) will be used for the implementation of the considered thresholding methods.



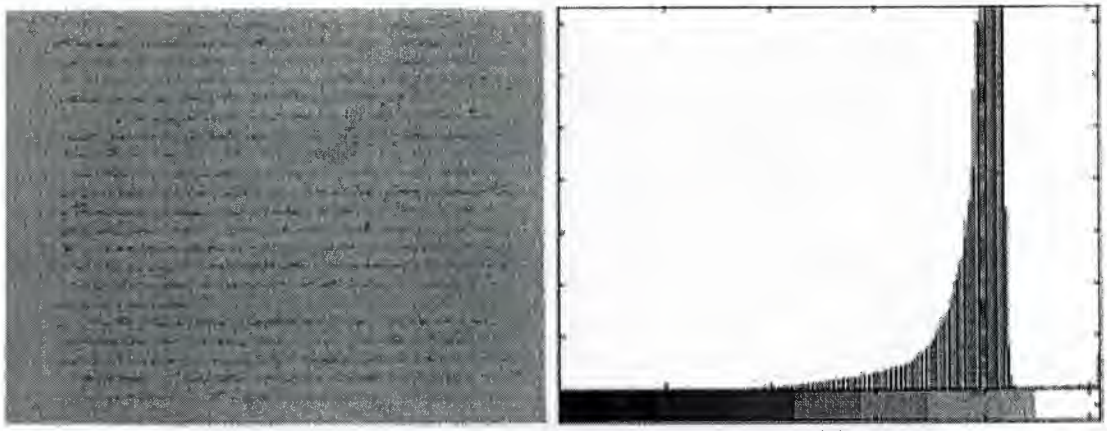
(a) (b)

Figure 4.1 – Image Set I Image (a) Bright Image and (b) Its Corresponding Gray Level Histogram.



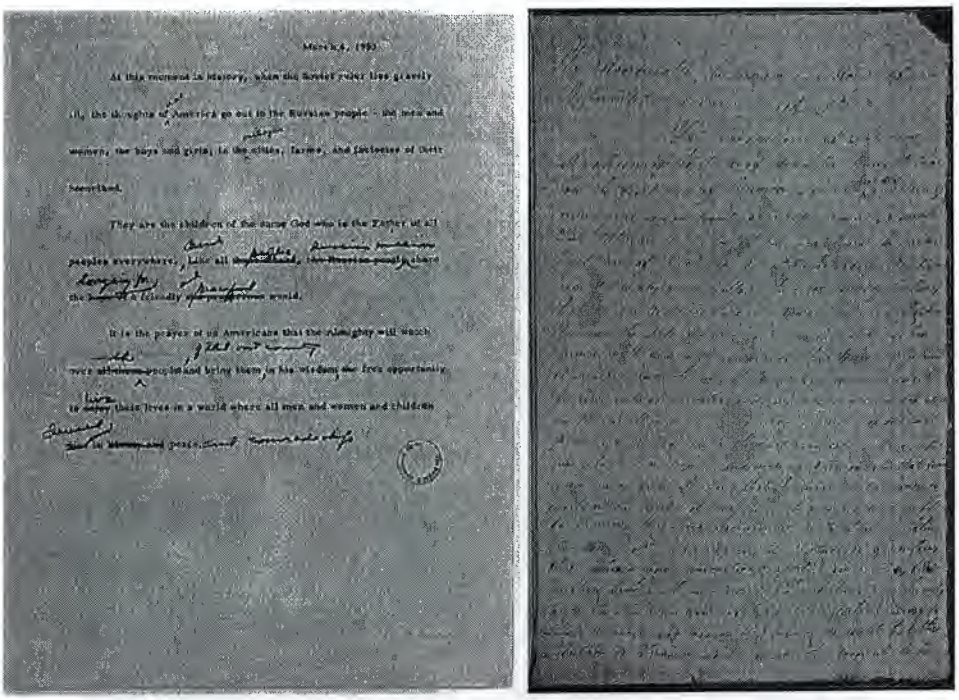
(a) (b)

Figure 4.2 – Image Set I Image (a) Dark Image and (b) Its Corresponding Gray Level Histogram.



(a) (b)

Figure 4.3 – Image Set I Image (a) Low Contrast Image and (b) Its Corresponding Gray Level Histogram.



(a) (b)

Figure 4.4 – Example Image of Set I, (a) Bright Image and (b) Low Contrast Image

4.3.2 Evaluation Procedure

In order to evaluate the obtained results when applying the considered methods using the proposed database, visual inspection of the enhanced documents was used, in addition to three metrics which were derived using the Peak Signal-to-Noise Ratio (PSNR) of the enhanced images.

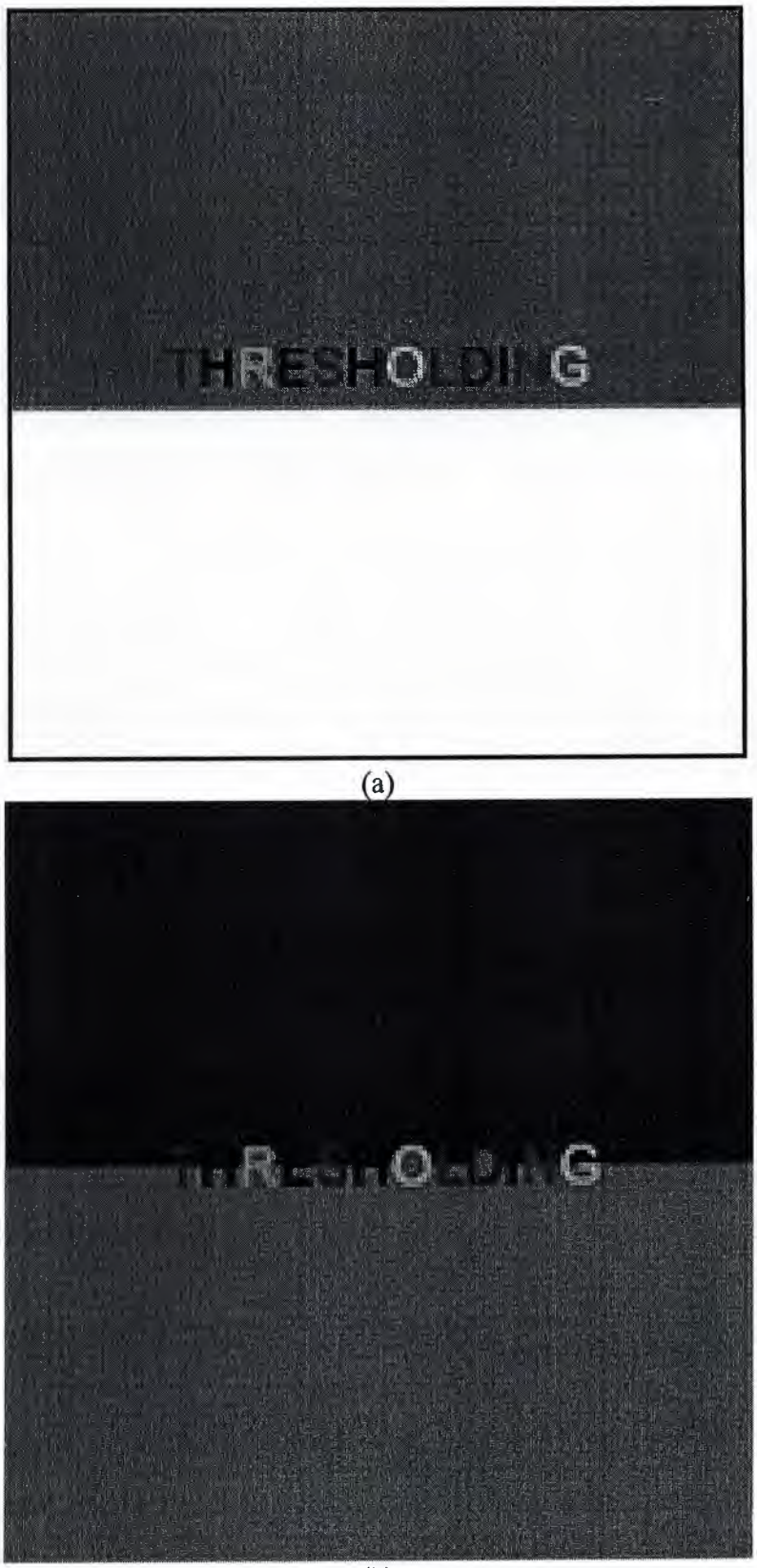


Figure 4.5 – Example Images of Test Set II (a) Bright Non-Uniform Background and Text and (b) Dark Non-Uniform Background and Text.

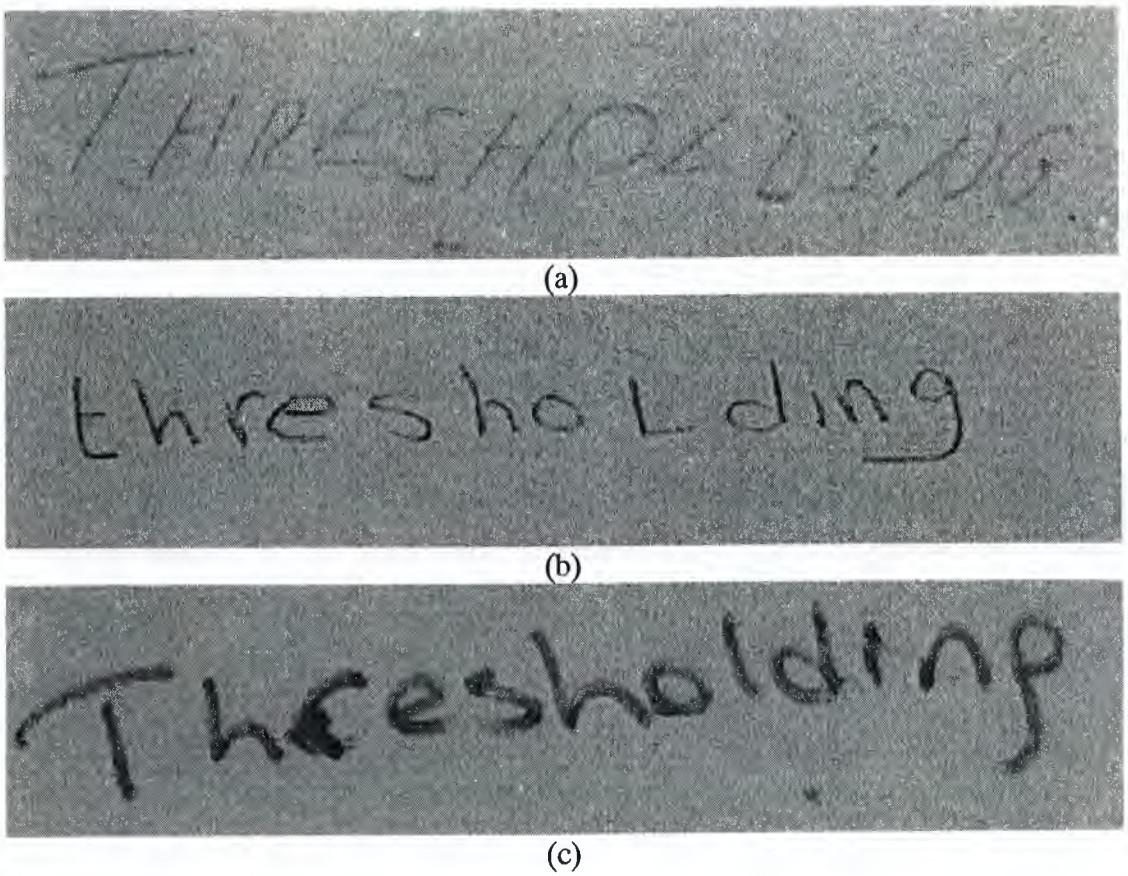


Figure 4.6 – Examples of Test Set III (a) Pencil Handwriting on Yellow Envelope Paper, (b) Pen Handwriting on Yellow Envelope Paper and (c) White Board Marker Handwriting on Yellow Envelope Paper.

4.3.2.1 Visual Inspection

Visual inspection of all enhanced documents in the three sets was performed by 15 independent human analyzers, who were asked to consider the clarity and readability of the words within Set I documents, noise occurrence and continuity of characters to determine clear characters within Set II documents, and clearly recognized readable characters within handwritten words in Set III documents. The general results of visual inspection were categorized as: recognized or unrecognized words for Set I, clear or unclear characters for Set II, and recognized or unrecognized characters for Set III. We believe that this method of evaluation is necessary as one of the objectives is to provide clearly enhanced document images that can improve human readability of degraded documents.

4.3.2.2 Computed Noise Analysis

Computed noise analysis of all enhanced images in Set I is also applied as the second part in the evaluation procedure. The application of this noise analysis was considered unnecessary with images from Set II and III, as satisfactory and clear evaluation results can be easily obtained by only using visual inspection with these two sets, which is not the case with Set I.

Using PSNR of enhanced images three metric parameters for each of the 14 binarization methods used in experiments was derived. These parameters are the Average PSNR Accuracy Rate (APAR), the Average PSNR Deviation (APD) of binarized images, and the Combined Performance Rate (CPR) of a thresholding method. PSNR (in dB) is defined as:

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right), \tag{4.1}$$

MSE is the mean squared error as defined as:

$$MSE = \frac{1}{M \times N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(Y_{i,j} - \hat{X}_{i,j}\right)^2, \tag{4.2}$$

where *(M x N)* is the size of the image, *Yi,j* and $\hat{X}_{i,j}$ represent the pixel values at location *(i, j)* of original and enhanced image, respectively.

The average PSNR accuracy rate (APAR) for a particular method is calculated by considering the maximum PSNR value obtained using the 14 thresholding methods, and the PSNR value obtained using only that particular method for a test image. The higher the APAR value is, the more efficient the thresholding method is.

The average PSNR deviation (APD) for a particular method is calculated by taking the difference between the maximum PSNR value obtained using the 14 thresholding methods, and the PSNR value obtained using only that particular method for a test image; the difference is then divided by the total number of test images. The lower the APD value is, the more efficient the thresholding method is. APAR and APD are defined as follows:

$$APAR_m = \left(\sum_{i=1}^{x}\left((PSNR_{mi} * 100)/\max(PSNR_i)\right)\right)/x, \tag{4.3}$$

$$APD_m = \left(\sum_{i=1}^{x}\left(\max(PSNR_i) - (PSNR_{mi})\right)\right)/x, \tag{4.4}$$

where $APAR_m$ is the average PSNR accuracy rate for method m, APD_m is the average PSNR deviation for method m, $PSNR_{mi}$ denotes the obtained PSNR value of enhanced image i using enhancement method m, $max(PSNR_i)$ denotes the maximum PSNR value of the enhanced image i obtained using the fourteen methods, and x is the total number of test images.

The Combined Performance Rate (CPR) is the third proposed parameter. CPR indicates the final performance of a thresholding method, where the higher CPR is the more superior the method is. CPR is defined as:

$$CPR_m = RW * (APAR/100) \tag{4.5}$$

where CPR_m is the Combined Performance Rate for method m, RW is the number of recognized words in Set I document images using visual inspection, and $APAR$ is the average PSNR accuracy rate for all document images in Set I.

Considering only PSNR values of the binarized enhanced images when evaluating the result is not always effective when comparing various thresholding methods and using such a diverse document image database. This is because a particular method may produce high PSNR values for a significant number of images, while producing low PSNR values for the rest of the images, thus making it difficult to determine stability and efficiency of the methods. Therefore, APAR, APD, and CPR metrics were derived, which provide a uniform indication of the efficiency of the compared methods.

4.4 Results and Comparisons

Experiments involved software implementation of the fourteen methods to binarize all documents in database. The C-programming language was used in implementation. The methods were: Otsu, Kittler and Illingworth, Yanni and Horne, Ramesh et al., Mass-Difference (MD), Kapur et al., and Albuquerque et al. entropy thresholding as global methods, and Niblack, Sauvola et al., Mean-Gradient, Pattern Average Thresholding (PAT), Adaptive Logical Thresholding (ALT), Water Flow Model (WFM), and Bernsen thresholding as local methods.

Throughout the experiments, the parameters and kernel sizes of the locally adaptive methods were chosen to provide the optimum performance of the methods; as suggested by previous works and comparisons. In their evaluation, Trier and Jain [90] noticed that *15x15* kernels with $k=-0.2$ produces optimum results and this was considered by Sezgin and Sankur [91] in their evaluation. Also Trier and Jain suggest *15x15* kernel size for

Bernsen method with *l=15*. Sauvola method is the improvement of Niblack method by adapting standard deviation thus *15x15* kernel size with suggested values *k=0.5* and *R=40* were used [91]. While comparing some methods and proposing Mean-Gradient Method, Leedham et al. [51] notice that *15x15* kernel size is suitable for their method with *k=-1.5* and *R=40*. After some experiments kernel sizes for PAT, ALT and WFM were decided as *15x15*, *8x8* and *5x5* respectively with *w=17* in WFM.

Table 4.1 shows the kernel sizes and chosen parameters for the implemented local thresholding methods.

Table 4.1 - Kernel Sizes And Parameters for Locally Adaptive Methods

Parameters	Niblack	Sauvola	Gradient	PAT	ALT	WFM	Bernsen
Kernel Size	15x15	12x12	15x15	15x15	8x8	5x5	15x15
k[2]	-1	0.5	-1.5	-	-	-	15
R[2]	-	128	40	-	-	-	-
w[2]	-	-	-	-	-	17	-

4.4.1 Image Set I Experiments

The first evaluation was performed on Set I images by visual inspection and noise analysis, with different results obtained for the different groups (*Bright*, *Low-Contrast* and *Dark images*) in this set.

4.4.1.1 Bright Images Group

After visual inspection of *Bright* images group (see Table 4.2), it was clear that global methods produce better results than local methods; with MD thresholding method being superior to the other global methods; followed by Kapur and Albuquerque et al. methods. However, when applying computed analysis using two PSNR-derived metrics (see Table 4.3), WFM followed by MD and ALT thresholding methods are seen to be superior to the other methods.

4.4.1.2 Low-Contrast Images Group

In *Low-Contrast* images group of Set I, Sauvola method produced optimum results using the visual inspection criterion; followed by Kittler and Illingworth and Kapur et al. methods (see Table 4.4). However, after performing noise analysis, Sauvola et al.

[2] *k* and *R* values are constants and *w* is the amount of water in Water Flow Model.

method seemed to add additional noise to binarized image, leaving ALT followed by WFM and MD as the superior methods (see Table 4.3).

Table 4.2 - Visual Inspection Results For *Bright* Images Group of Set I

Thresholding Method	Category[3]	Rank	Total Words	Recognized Words	Unrecognized Words	Recognition Rate
Otsu	G	5	8300	7533	767	90.75%
Kit. and Illing.	G	12	8300	4691	3609	56.51%
Yanni and Horne	G	14	8300	2932	5368	35.32%
Ramesh et al.	G	6	8300	7397	903	89.12%
MD	**G**	**1**	**8300**	**8097**	**203**	**97.55%**
Kapur et al.	G	2	8300	7939	361	95.65%
Albuquerque et al.	G	3	8300	7939	361	95.65%
Niblack	L	7	8300	7352	948	88.57%
Sauvola et al.	L	8	8300	7239	1061	87.21%
Mean-Gradient	L	11	8300	6518	1782	78.53%
PAT	L	10	8300	6833	1467	82.32%
ALT	L	9	8300	7194	1106	86.67%
WFM	L	4	8300	7917	383	95.38%
Bernsen	L	13	8300	3608	4692	43.43%

Table 4. 3 - APD and APAR Results for All Set I Groups

Thresholding Method	C[4]	Bright Group			Low-Contrast Group			Dark Group		
		APD	APAR	R[4]	APD	APAR	R[4]	APD	APAR	R[4]
Otsu	G	2.952	86.561	6	2.769	81.189	7	4.254	71.16	3
Kit. and Illing.	G	8.238	65.369	12	3.904	75.767	9	4.678	67.79	6
Yanni and Horne	G	6.699	69.500	9	6.144	62.723	11	-	-	-
Ramesh et al.	G	8.338	61.828	13	9.456	47.783	14	-	-	-
MD	G	1.029	95.744	2	1.106	89.176	3	4.705	67.62	7
Kapur et al.	G	1.801	91.834	4	2.719	79.268	6	8.714	41.39	10
Albuquerque et al.	G	4.967	76.438	8	3.065	77.585	8	-	-	-
Niblack	L	3.680	84.664	7	1.756	85.878	4	4.369	70.09	4
Sauvola et al.	L	9.539	59.093	14	6.327	65.064	12	**0.241**	**98.08**	**1**
Mean-Gradient	L	2.786	87.639	5	2.322	82.678	5	-	-	-
PAT	L	8.151	65.177	11	4.416	73.464	10	3.420	77.14	2
ALT	L	1.101	95.319	3	**0.865**	**90.144**	**1**	4.725	67.49	8
WFM	L	**0.365**	**98.519**	**1**	1.082	88.799	2	4.620	68.28	5
Bernsen	L	7.333	68.003	10	6.595	64.142	13	7.629	50	9

4.4.1.3 Dark Images Group

In *Dark* images group of Set I, MD, Niblack, ALT and WFM methods shared a similar rank producing optimum results using the visual inspection criterion. These were followed by Kittler and Illingworth and Sauvola methods (see Table 4.5). However, when applying computed analysis (see Table 4.3), Sauvola followed by PAT and Otsu thresholding methods are seen to be superior to the other methods.

[3] G: Global, L: Local.
[4] C: Category, R: Rank.

4.4.1.4 Combined Groups Results

Set I, by considering all results, it is clear that, MD is superior in visual inspection (See Table 4.6) and WFM is superior in noise analysis (see Table 4.7).

This results were supported by Final Performance Evaluation, which optimum results were achieved by WFM and MD respectively (see Table 4.8). Examples of obtained resutls for Set I can be seen in Figure 4.7-4.9.

Table 4.4 - Visual Inspection Results For *Low-Contrast* Images Group of Set I

Thresholding Method	Category	Rank	Total Words	Recognized Words	Unrecognized Words	Recognition Rate
Otsu	G	7	1901	1216	685	63.96%
Kit. and Illing.	G	2	1901	1431	470	75.27%
Yanni and Horne	G	13	1901	95	1806	4.99%
Ramesh et al.	G	12	1901	114	1787	5.99%
MD	G	6	1901	1235	666	64.96%
Kapur et al.	G	3	1901	1388	513	73.01%
Albuquerque et al.	G	11	1901	154	1747	8.10%
Niblack	L	9	1901	1102	799	57.96%
Sauvola et al.	**L**	**1**	**1901**	**1510**	**391**	**79.43%**
Mean-Gradient	L	10	1901	912	989	47.97%
PAT	L	5	1901	1349	552	70.96%
ALT	L	8	1901	1140	761	59.96%
WFM	L	4	1901	1359	542	71.48%
Bernsen	L	6	1901	1235	666	64.96%

Table 4.5 - Visual Inspection Results For *Dark* Images Group of Set I

Thresholding Method	Category	Rank	Total Words	Recognized Words	Unrecognized Words	Recognition Rate
Otsu	G	5	90	55	35	61.11%
Kit. and Illing.	G	2	90	85	5	94.44%
Yanni and Horne	G	-	90	0	90	-
Ramesh et al.	G	-	90	0	90	-
MD	**G**	**1**	**90**	**88**	**2**	**97.77%**
Kapur et al.	G	-	90	55	35	61.11%
Albuquerque et al.	G	-	90	0	90	-
Niblack	**L**	**1**	**90**	**88**	**2**	**97.77%**
Sauvola et al.	L	3	90	75	15	83.33%
Mean-Gradient	L	-	90	0	90	-
PAT	L	4	90	58	32	64.44%
ALT	**L**	**1**	**90**	**88**	**2**	**97.77%**
WFM	**L**	**1**	**90**	**88**	**2**	**97.77%**
Bernsen	L	6	90	10	80	11.11%

MD was designed to extract dominant pixels within document which are categorized as 'text/foreground' and this was achieved by using maximum brightness point as the background limit of document and mean value as 'fuzzy layer'. Difference of maximum point to mean value shifts 'fuzzy layer' to left on image histogram to detect details of foreground while discarding relatively deviated pixels within image.

Water Flow Model simulates water which falls down to lower regions of 3D terrain of images. This yields effective extraction of dominant pixel values which are 'text/foreground' of image and unnecessary information eliminated. Considering global and local properties of images makes difference between WFM and other local methods such as Niblack, Sauvola and Bernsen which considers local information within kernels. Thus, behavior of WFM mostly likes global methods which are more effective for noise removal.

These simple and effective properties of MD and WFM, provides efficient achievement of desired goals in Set I document images binarization. MD achieved optimum recognition of images by visual inspection and WFM achieved optimum clearance of noises during binarization process. Thus, relatively similar results of WFM and MD make them optimum for binarization of documents either in bright, low contrast or dark images.

Table 4.6 - General Average Visual Inspection Results for All Groups In Set I

Thresholding Method	Category	Rank	Total Words	Recognized Words	Unrecognized Words	Recognition Rate
Otsu	G	7	10291	8504	1787	82.63%
Kit. and Illing.	G	12	10291	6207	4084	60.31%
Yanni and Horne	G	14	10291	3027	7264	29.41%
Ramesh et al.	G	10	10291	7511	2780	72.98%
MD	**G**	**1**	**10291**	**9420**	**871**	**91.53%**
Kapur et al.	G	2	10291	9382	909	91.16%
Albuquerque et al.	G	9	10291	8093	2198	78.64%
Niblack	L	6	10291	8542	1749	83%
Sauvola et al.	L	4	10291	8824	1467	85.74%
Mean-Gradient	L	11	10291	7507	2784	72.94%
PAT	L	5	10291	8601	1690	83.57%
ALT	L	8	10291	8422	1869	81.83%
WFM	L	3	10291	9364	927	90.99%
Bernsen	L	13	10291	4853	5438	47.15%

4.4.2 Image Set II Experiments

In Set 2, due to the uniform illumination of the background within the images in this set, global methods generally causes more loss of information that local ones.

Visual inspection and recognition of characters within binarized images showed (see Table 4.9) that Sauvola method is the optimum for created text documents and followed by PAT, ALT, Niblack and WFM methods. Examples of obtained resutls for Set II can be seen in Figure 4.10.

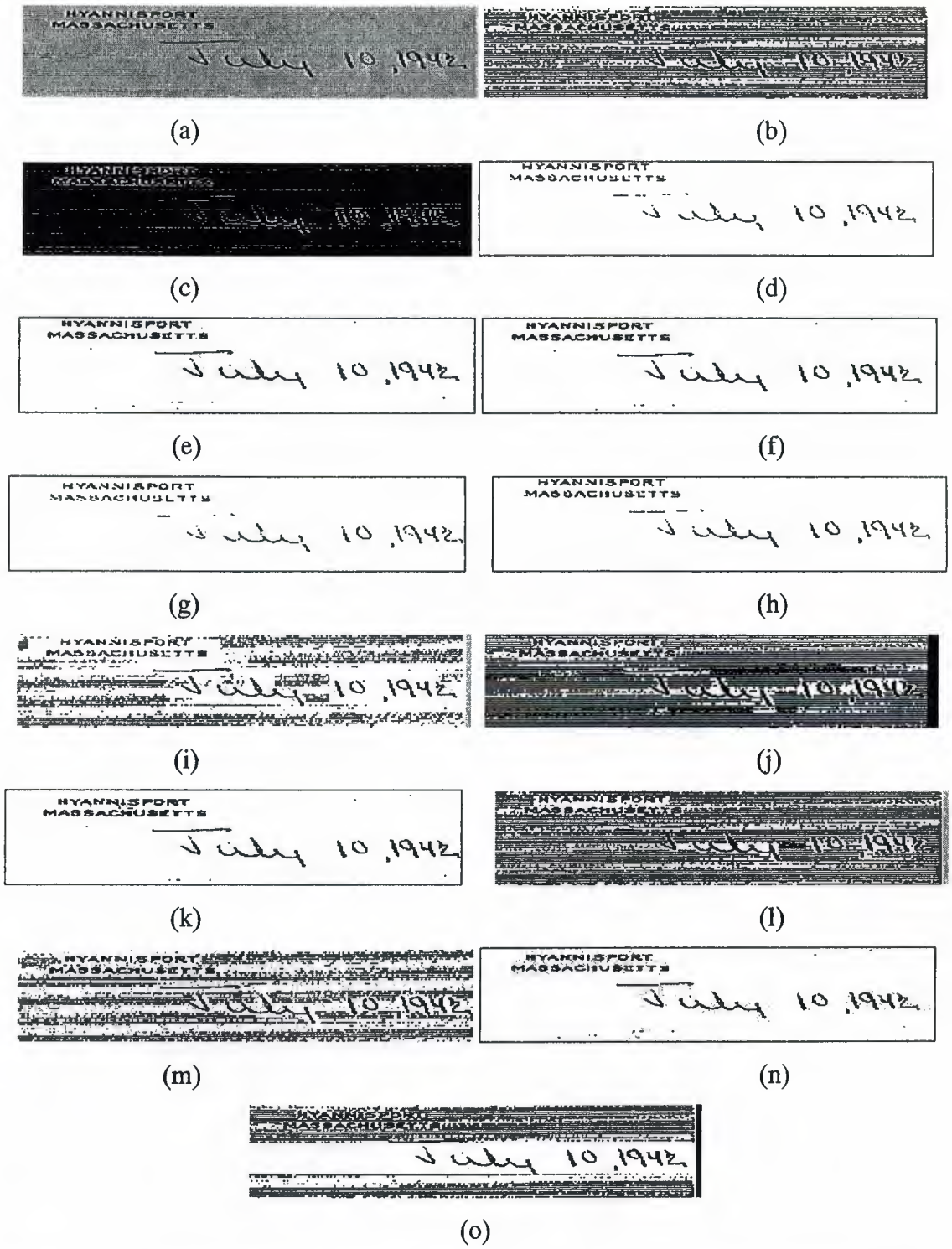


Figure 4.7 - Partial Result of Bright Image of Test Set I (a) Original Image, (b) Otsu Method, (c) Kittler and Illingworth Method, (d) Yanni and Horne Method, (e) Ramesh et al. Method, (f) MD Method, (g) Kapur et al. Entropy, (h) Albuquerque et al. Entropy, (i) Niblack Method, (j) Sauvola et al. Method, (k) Mean-Gradient Method, (l) PAT Method, (m) ALT Method, (n) Water Flow Model and (o) Bernsen Method.

The aim of Set II was the performance observation of binarization methods in extreme non-uniform conditions for both background and characters. Although any of methods could not achieve more than 60% of cleared characters, it was seen that, local methods produced better results than global methods.

Sauvola et al. method was the improvement of Niblack method by adapting standard deviation to lower the threshold value. This property provides optimum binarization of specially created extreme non-uniform texts and background by lowering threshold value to detect more characters within documents. However, this property of Sauvola et al. method causes extreme noise addition to Set I images and to decrease the recognition of texts within document by visual inspection.

Table 4.7 - General APD And APAR Results For All Groups in Set I

Thresholding Method	Category	Set I		
		APD	APAR	Rank
Otsu	G	2.943863	86.37631	6
Kittler and Illingworth	G	6.922652	70.1334	9
Yanni and Horne	G	6.821943	67.14557	10
Ramesh et al.	G	8.968075	57.24856	14
MD	G	1.120711	95.0755	3
Kapur et al.	G	2.198561	88.99776	4
Albuquerque et al.	G	4.572706	77.16784	8
Niblack	L	3.148074	86.42465	7
Sauvola et al.	L	8.489479	62.98983	13
Mean-Gradient	L	2.86563	86.30933	5
PAT	L	7.019589	69.32447	11
ALT	L	1.098199	95.07637	2
WFM	**L**	**0.653009**	**96.9005**	**1**
Bernsen	L	7.173098	68.03938	12

Table 4.8 - Final Performance Results For Set I

Thresholding Method	Category	Set I	
		CPR	Rank
Otsu	G	7344	6
Kittler and Illingworth	G	4352	12
Yanni and Horne	G	4877	11
Ramesh et al.	G	4299	13
MD	**G**	**8955**	**2**
Kapur et al.	G	8349	3
Albuquerque et al.	G	6244	8
Niblack	L	7381	5
Sauvola et al.	L	5557	10
Mean-Gradient	L	6478	7
PAT	L	5962	9
ALT	L	8006	4
WFM	**L**	**9073**	**1**
Bernsen	L	3301	14

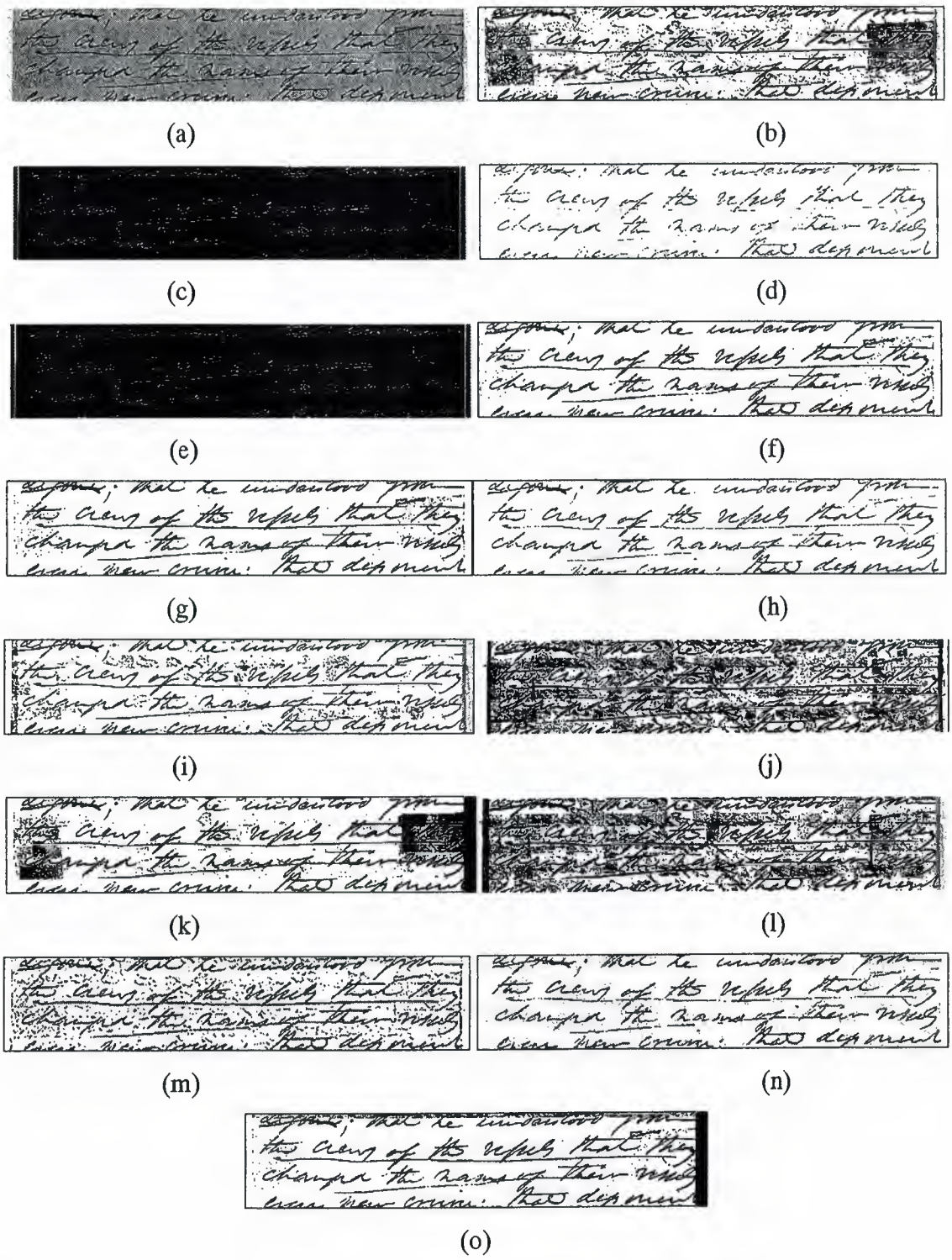


Figure 4.8 - Example Result of Low Contrast Image of Test Set I (a) Original Image, (b) Otsu Method, (c) Kittler and Illingworth Method, (d) Yanni and Horne Method, (e) Ramesh et al. Method, (f) MD Method, (g) Kapur et al. Method, (h) Albuquerque et al. Entropy, (i) Niblack Method, (j) Sauvola et al. Method, (k) Mean-Gradient Method, (l) PAT Method, (m) ALT Method, (n) Water Flow Model and (o) Bernsen Method.

Also PAT, ALT and Niblack methods performed better results than other methods by considering local information of kernels but they could not achieve the recognition rate as Sauvola method.

Table 4.9 - General Visual Inspection Results for Set II

Thresholding Method	Category	R	Total Characters	Clear Characters	Unclear Characters	Recognition Rate
Otsu	G	8	72	13	59	18.05%
Kit. and Illing.	G	10	72	10	62	13.88%
Yanni and Horne	G	12	72	6	66	8.33%
Ramesh et al.	G	11	72	9	63	12.50%
MD	G	6	72	24	48	33.33%
Kapur et al.	G	5	72	28	44	38.88%
Albuquerque et al.	G	11	72	9	63	12.50%
Niblack	L	3	72	36	36	50%
Sauvola et al.	**L**	**1**	**72**	**43**	**29**	**59.72%**
Mean-Gradient	L	7	72	22	50	30.55%
PAT	**L**	**2**	**72**	**40**	**32**	**55.55%**
ALT	**L**	**2**	**72**	**40**	**32**	**55.55%**
WFM	L	4	72	29	43	40.27%
Bernsen	L	9	72	12	60	16.66%

4.4.3 Image Set III Experiments

Set III consists of 6 different competitive image groups which were Pencil on White Paper (WP), Pen on WP, White Board Marker (WBM) on WP, Pencil on Yellow Envelope Paper (YP), Pen on YP, and WBM on YP. White Paper is used to performance evaluation of thresholding methods on uniform background by using different writing materials as pen, pencil and white board marker which causes different illuminations and stroke widths of characters. Yellow envelope paper is used to performance evaluation of thresholding methods on non-uniform and noisy background which yellow envelope papers consists.

For uniform background (white paper), it was observed that global methods, Otsu, Kittler and Illingworth and MD, and local methods PAT and WFM performed stable results for all kind of writing materials.

For non-uniform background (yellow envelope paper), it was observed that almost same methods produced better results than others, however, global methods except Otsu, are failed to binarize Pencil on Yellow Envelope Paper because converting RGB to Gray-scale causes extreme harmonization of text and background at low level contrast point. Thus, Otsu, MD, PAT, Niblack and WFM achieved better results than other methods.

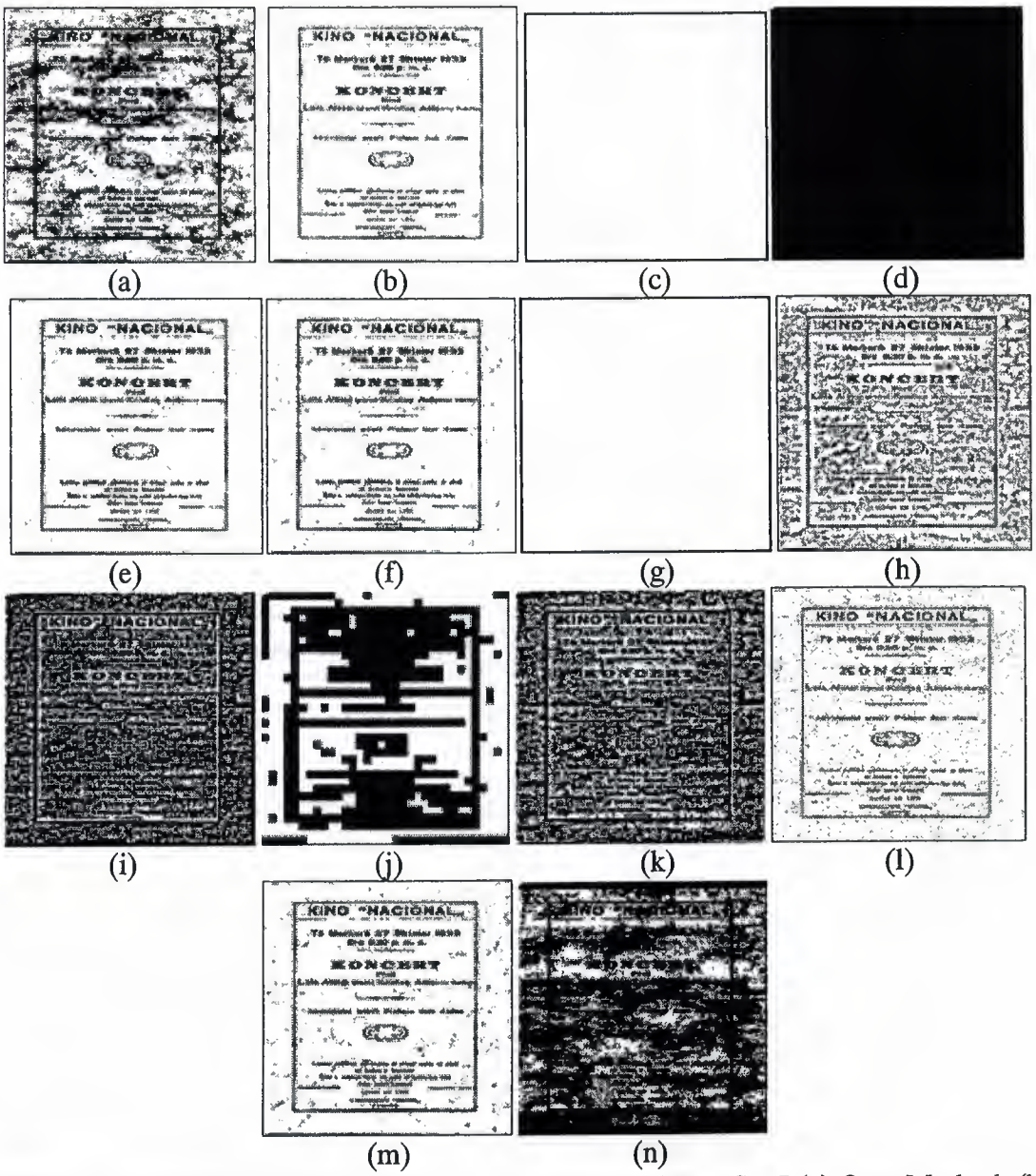


Figure 4.9 - Example Result of Fig 4.3 -Dark Group- of Test Set I (a) Otsu Method, (b) Kittler and Illingworth Method, (c) Yanni and Horne Method, (d) Ramesh et al. Method, (e) MD Method, (f) Kapur et al. Method, (g) Albuquerque et al. Entropy, (h) Niblack Method, (i) Sauvola et al. Method, (j) Mean-Gradient Method, (k) PAT Method, (l) ALT Method, (m) Water Flow Model and (n) Bernsen Method.

Evaluation of whole images in Set III shows that, Otsu which is superior one followed by WFM, PAT, MD and Sauvola et al. methods which recognition rates are higher than 90% (see Tables 10-11). Examples of obtained resutls for Set III can be seen in Figure 4.11.

Table 4.10 - Visual Inspection Results for Set III

Thresholding Method	Category	Recognized Characters[5]								
		Pen			Pencil			WB Marker		
		WP	YP	Av.[6]	WP	YP	Av.[6]	WP	YP	Av.[6]
Otsu	G	107	107	107	107	103	**105**	108	107	107.5
Kit. and Illing.	G	108	83	95.5	108	47	77.5	108	108	**108**
Yanni and Horne	G	22	15	18.5	55	5	30	50	65	57.5
Ramesh et al.	G	26	22	24	49	6	27.5	66	58	62
MD	G	108	106	107	108	59	83.5	108	108	**108**
Kapur et al.	G	95	85	90	40	32	36	10	5	7.5
Albuquerque et al.	G	11	26	18.5	12	8	10	5	2	3.5
Niblack	L	105	106	105.5	106	103	104.5	73	88	80.5
Sauvola et al.	L	95	97	96	89	100	94.5	103	105	104
Mean-Gradient	L	21	27	24	0	0	0	65	73	69
PAT	L	108	102	105	105	79	92	103	102	102.5
ALT	L	108	107	**107.5**	61	95	78	97	93	95
WFM	L	102	107	104.5	102	103	102.5	94	104	99
Bernsen	L	88	76	82	10	66	38	25	68	46.5

Table 4.11 - General Visual Inspection Results for Set III

Thresholding Method	Category	Rank	Total Characters	Clear Characters	Unclear Characters	Recognition Rate
Otsu	**G**	**1**	**648**	**639**	**9**	**98.61%**
Kit. and Illing.	G	8	648	472	176	72.83%
Yanni and Horne	G	12	648	212	436	32.71%
Ramesh et al.	G	11	648	227	421	35.03%
MD	G	4	648	597	51	91.12%
Kapur et al.	G	10	648	267	381	41.20%
Albuquerque et al.	G	14	648	64	584	9.87%
Niblack	L	6	648	581	67	89.66%
Sauvola et al.	L	5	648	589	59	90.89%
Mean-Gradient	L	13	648	186	462	28.70%
PAT	L	3	648	599	49	92.43%
ALT	L	7	648	561	87	86.57%
WFM	**L**	**2**	**648**	**612**	**36**	**94.44%**
Bernsen	L	9	648	333	315	51.38%

During the obtaining binarization results, processing time of each method were calculated. Table 4.12 shows the average processing time of each method.

4.5 Summary

In this chapter, performed experiments on document image binarization, design of experiments, evaluation criteria and obtained results were explained in details by performing comparative evaluation of considered methods.

[5] Out of 108 characters for each group (Pen, Pencil, WB Marker).
[6] Average character recognition rate.

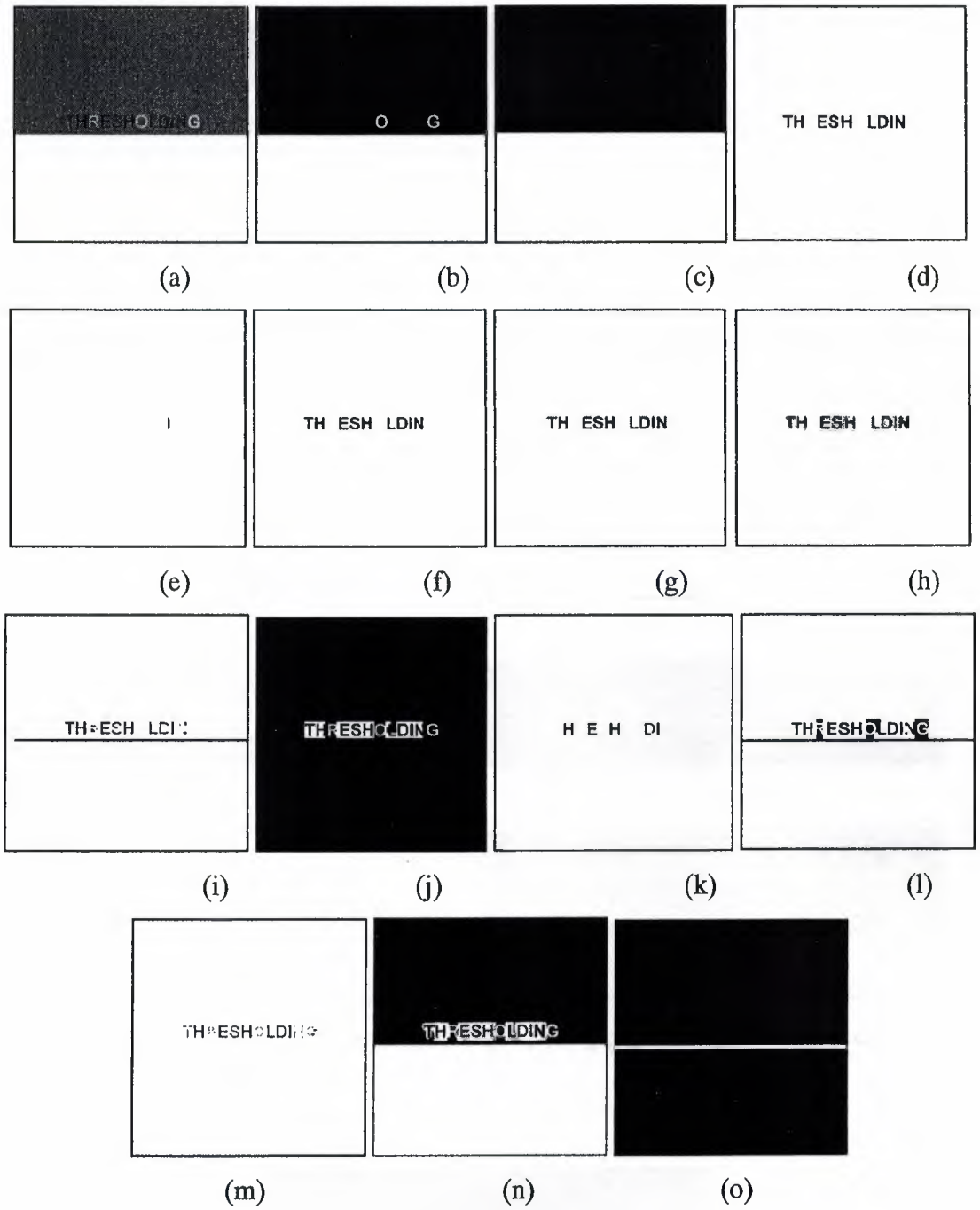


Figure 4.10 Example Result of Created Word Image of Test Set II (a) Original Image, (b) Otsu Method, (c) Kittler and Illingworth Method, (d) Yanni and Horne Method, (e) Ramesh et al. Method, (f) MD Method, (g) Kapur et al. Method, (h) Albuquerque et al. Entropy, (i) Niblack Method, (j) Sauvola et al. Method, (k) Mean-Gradient Method, (l) PAT Method, (m) ALT Method, (n) Water Flow Model and (o) Bernsen Method.

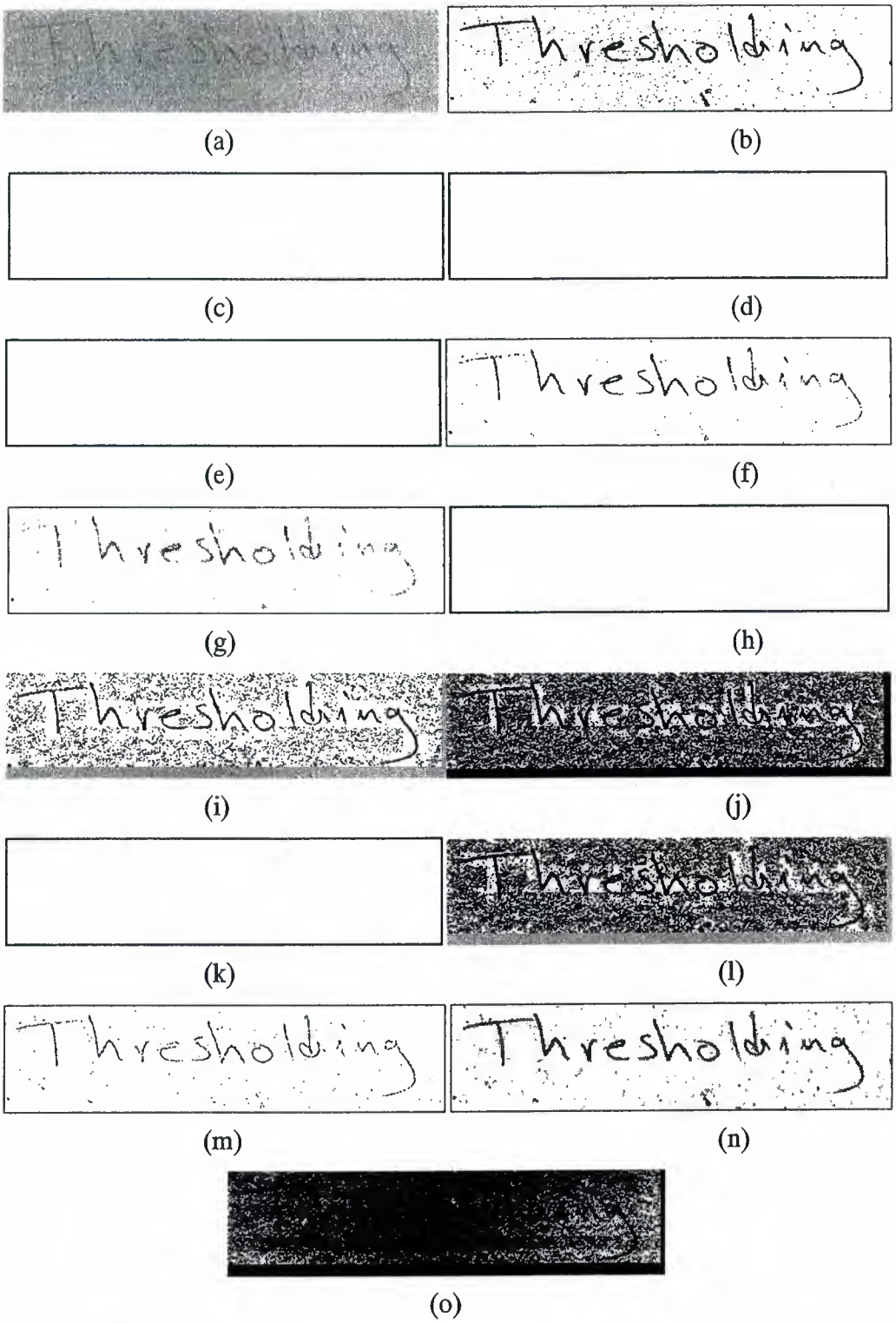


Figure 4.11 Example Result of Handwritten Image of Test Set III (a) Original Image, (b) Otsu Method, (c) Kittler and Illingworth Method, (d) Yanni and Horne Method, (e) Ramesh et al. Method, (f) MD Method, (g) Kapur et al. Method, (h) Albuquerque et al. Entropy, (i) Niblack Method, (j) Sauvola et al. Method, (k) Mean-Gradient Method, (l) PAT Method, (m) ALT Method, (n) Water Flow Model and (o) Bernsen Method.

Table 4.12- Average Processing Time of the Methods

Thresholding Method	Average Time[7]
Otsu	0.68 s.
Kittler and Illingworth	0.66 s.
Yanni and Horne	0.71 s.
Ramesh et al.	1.06 s.
MD	0.085 s.
Kapur et al.	0.75 s.
Albuquerque et al.	0.75 s.
Niblack	0.013 s.
Sauvola et al.	0.11 s.
Mean-Gradient	0.29 s.
PAT	0.062 s.
ALT	0.42 s.
WFM	4.3 s.
Bernsen	0.11 s.

[7] Pentium IV, 2 GHz. CPU and 512 MB RAM

CONCLUSION

In this thesis a novel single-stage global binarization method which is designed to enhance and separate dominant pixels of images was proposed. The method which was named Mass Difference (MD) uses luminance value and mean of intensity to shift mean value of an image to optimum position that is determined as a threshold point. Several experiments, such as luminance value test, non-uniform background test and whole document test, were performed to test the success and efficiency of the proposed method.

Preliminary experiments were performed to demonstrate the efficiency of MD method and the results of these experiments suggest that MD thresholding method provided superior results when compared to the other considered methods in two different preliminary experiments under different conditions.

Furthermore, thirteen known or recently developed thresholding methods which can be used for document image binarization were categorized into global and local thresholding groups and were implemented using the C language. The global thresholding group comprised: Otsu, Kittler and Illingworth, Yanni and Horne, Ramesh et al., Kapur et al., and Albuquerque et al. entropy methods. The local thresholding group comprised: Niblack, Sauvola et al., Mean-Gradient, Pattern Average Thresholding (PAT), Adaptive Logical Thresholding (ALT), Water Flow Model (WFM), and Bernsen methods.

A comprehensive multi-applications database was also created to comprise different challenging tasks such as historical documents, artificially created words and handwriting document images that simulate real life implementation in extreme and noisy conditions. The diversity of the document types within this database provides a larger database and multi-application document images with different levels of noise and contrast, non-uniform illumination (background), signal-dependent noise, smears, and non-uniform foreground (text). The use of multi-application database was suggested in order to provide a more objective evaluation and, consequently, the determination of an optimum thresholding method.

The strategy for evaluating the 14 methods was based on two criteria: visual inspection and computed noise analysis of the binarized images. The sole use of one of these criteria does not provide an objective conclusion on which method is optimum,

thus the need for using both criteria. For this purpose, three PSNR-derived metric parameters that were used for noise analysis were introduced; these parameters were called: the Average PSNR Accuracy Rate (APAR), the Average PSNR Deviation (APD) of binarized images, and the Combined Performance Rate (CPR) of a thresholding method. The later determines the optimum performance for a thresholding method and combines obtained results from both visual inspection and noise analysis criteria.

A comparative evaluation of the thresholding methods was performed using the multi-application image database and using the proposed evaluation strategy. The overall evaluation results of comparative evaluation suggest that MD thresholding method, which is a global method, provided superior results when compared to the other methods, and thus can be considered as the optimum global thresholding method. On the other hand, WFM thresholding method, which is a local thresholding method, showed superior performance, and thus can be considered as the optimum local thresholding method.

By comparing MD and WFM, it was shown that these two methods achieved similar results in historical documents. MD was superior in visual inspection that was used to measure human readability with a minimal noise addition and WFM was superior in noise removal with a little loss of information. When using the especially created word documents, the recognition rate of both these methods decreased, whereas using handwriting document images both methods produced similar successful results.

However, the determination of mask size of WFM is a serious drawback that sometimes causes a huge loss of information during binarization and computational cost. MD has faster execution time which makes marginal difference in computational cost when binarizing a huge collection of documents. Hence, MD thresholding method can be considered as the optimum threshold method.

Future work will focus on improving the performance of MD thresholding for applications to images with similar characteristics to artificially created words.

REFERENCES

[1] S. E. Umbaugh, Computer Imaging : Digital Image Analysis and Processing, CRC Press, USA, ISBN 0-8493-2919-1, 2005.

[2] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Prentice Hall, USA, ISBN: 9780131687288, 2008.

[3] E. Davies, *Machine Vision: Theory, Algorithms and Practicalities*, Academic Press, pp 26-27, 79-99, 1990.

[4] A. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, pp. 235, 989, 1989.

[5] J. K. Kim, J. M. Park, K. S. Song and H. W. Park, "Adaptive Mammographic Image Enhancement Using First Derivatives and Local Statistics", *IEEE Transactions on Medical Imaging*, vol.16, no.5, pp.495-502, 1997.

[6] I. Koren, A. L. and F. Taylor, "Enhancement via Fusion of Mammographic Features", *International Conference on Image Processing* (ICIP 98), vol.3, pp. 722-726, 1998.

[7] M. Wirth, M. Fraschini and J. Lyon, "Contrast Enhancement of Microcalcifications in Mammograms Using Morphological Enhancement and Non-flat Structuring Elements", *17th IEEE Symposium on Computer-Based Medical Systems* (CBMS'04), 2004.

[8] U. Numburi, G. P. Chatimavroudis, A. E. Uber and J. P. Kalafut, "Modelling of Contrast Enhancement for Cardiac Multi-Detector Row CT" *IEEE International Workshop on Imaging Systems and Techniques* (IEEE IST 2005), pp. 125-129, 2005.

[9] M. Jiang, Q. Ji and B. F. McEwen, "Enhancement of Microtubules in EM Tomography", *IEEE International Symposium on Biomedical Imaging*, pp. 1123-1126, 2004.

[10] C. H. Lo, Y. Guo and C. C. Lu, "A Binarization Approach for CT-MR Registration Using Normalized Mutual Information" *5th IASTED International Conference on Signal and Image Processing*, 2003.

[11] E. W. Abel, C. A. Wigderowitz and D. I. Rowley, "Frequency Domain Imaging of Cancellous Bone", *18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1182-1184, 1996.

[12] D. Sale, R. R. Schultz and R. J. Szczerba, "Super-Resolution Enhancement of Night Vision Image Sequences", *International Conference on Systems, Man, and Cybernetics,* vol. 3, pp.1633-1638, 2000.

[13] V. Areekul, U. Watchareeruetai, K. Suppasriwasuseth, and S. Tantaratana, "Separable Gabor filter realization for fast fingerprint enhancement", *IEEE International Conference on Image Processing*, (ICIP 2005), vol. 3, pp. 253-256, 2005.

[14] S. Wang and Y. Wang, "Fingerprint Enhancement in the Singular Point Area", *IEEE Signal Processing Letters*, vol. 11, no. 1, 2004.

[15] L. Jin, S. Satoh and M. Sakauchi, "A Novel Adaptive Image Enhancement Algorithm for Face Detection", *Proceedings of the 17th International Conference on Pattern Recognition* (ICPR'04), vol. 4, pp. 843-848, 2004.

[16] B. Lunden, K. Wester and G. Bax, "Satellite Image Enhancement For Rock Type Separation", *International Geoscience and Remote Sensing Symposium 'Remote Sensing: Global Monitoring for Earth Management'* (IGARSS '91), vol. 4, pp. 2047-2050, 1991.

[17] T. R. Randolph and M. J. T. Smith, "Enhancement of Fax Documents Using a Binary Angular Representation", *Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing*, pp.125-128, 2001.

[18] G. A. Ware, D. M. Chabries, R. W. Christiansen and C. E. Martin, "Multispectral Document Enhancement: Ancient Carbonized Scrolls", *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, (IGARSS 2000), vol. 6, pp. 2486-2488, 2000.

[19] J. Kittler and J. Illingworth, "Minimum Error Thresholding", *Pattern Recognition*, vol. 19, no 4, pp. 41-47, 1986.

[20] T.W. Riddler and S. Calvard, "Picture Thresholding Using an Iterative Selection Method", *IEEE Transactions on Sytems, Man and Cybernetics*, vol. 8, pp. 630-632, 1978.

[21] N. Otsu, "A Threshold Selection Method from Gray-Level Histogram", *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 9, pp. 62-66, 1979.

[22] Y. Nakagawa and A. Rosenfeld, "Some Experiments on Variable Thresholding", *Pattern Recognition*, vol. 11, no.3, pp.191-204, 1979.

[23] C.K. Chow and T. Kaneko, "Automatic, Boundary Detection of the Leftventricle from Cine-Angiograms", *Computational Biomedical Research*, vol. 5, pp. 388–410, 1972.

[24] T. Pun, "A New Method for Gray-level Picture Threshold Using the Entropy of the Histogram", *Signal Processing*, vol. 2, no.3, pp. 223-237, 1980.

[25] Y. Yasuda, M. Dubois and T. S. Huang, "Data Compression for Check Processing Machines", *Proceedings of IEEE*, vol.68, pp. 874-885, 1980.

[26] J.M. White and G.D. Rohrer, "Image Thresholding for Optical Character Recognition and Other Applications Requiring Character Image Extraction", *IBM J. Research and Development*, vol. 27, pp. 400-411, 1983.

[27] A. Rosenfeld and P. De la Torre, "Histogram Concavity Analysis as an Aid in Threshold Selection", *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 13, pp. 231-235, 1983.

[28] J.N. Kapur, P.K. Sahoo, and A.K.C. Wong, "A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram", *Computer Vision, Graphics, and Image Processing*, vol. 29, pp. 273-285, 1985.

[29] D. E. Lloyd, "Automatic Target Classification Using Moment Invariant of Image Shapes", Technical Report, RAE IDN AW126, UK, 1985.

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

[31] J. Bernsen, "Dynamic Thresholding of Gray-Level Images", *Proceedings of the 8th International Conference on Pattern Recognition*, pp. 1251-1255, 1986.

[32] P. W. Palumbo, P. Swaminathan, and S. N. Srihari, "Document Image Binarization: Evaluation of Algorithms", *Proc. SPIE*, vol. 697, pp. 278–286, 1986.

[33] S. Abutaleb, "Automatic Thresholding of Gray-Level Pictures Using Two Dimensional Entropy" *Computer Vision, Graphics, and Image Processing*, vol. 47, pp. 22-32, 1989.

[34] S.D. Yanowitz and A.M. Bruckstein, "A New Method for Image Segmentation", *Computer Vision, Graphics, and Image Processing*, vol. 46, pp. 82-95, 1989.

[35] T. Taxt, P.J. Flynn and A.K. Jain, "Segmentation of Document Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp.1322-1329, 1989.

[36] L. Eikvil, T. Taxt, and K. Moen, "A Fast Adaptive Method for Binarization of Document Images", *Proceedings of First International Conference on Documents Analysis and Recognition*, pp. 435-443, 1991.

[37] J.R. Parker, "Gray Level Thresholding in Badly Illuminated Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 813-819, 1991.

[38] C. H. Li and C. K. Lee, ''Minimum Cross-Entropy Thresholding,'' *Pattern Recognition* vol. 26, pp. 617-625, 1993.

[39] M. Kamel and A. Zhao, "Extraction of Binary Character/Graphics Images from Grayscale Document Images", *CVGIP: Graphical Models Image Processing*, vol.55, no.3, pp.203-217, 1993.

[40] M.K. Yanni and E. Horne, "A New Aproach to Dynamic Thresholding", *9th European Conference on Signal Processing,* EUSIPCO'94, vol. 1, pp.34-44, 1994.

[41] N. Ramesh, J.H. Yoo and I.K. Sethi, "Thresholding Based on Histogram Approximation", *IEE Vision, Image and Signal Processing*, vol. 142, no.5, 1995.

[42] J. C. Yen, F. J. Chang, and S. Chang, "A New Criterion for Automatic Multilevel Thresholding", *IEEE Transactions on Image Processing,* vol. 4, pp. 370-378, 1995.

[43] N. R. Pal, "On Minimum Cross-Entropy Thresholding", *Pattern Recognition,* vol. 29, no. 4, pp. 575-580, 1996.

[44] P. Sahoo, C. Wilkins, and J. Yeager, "Threshold Selection Using Renyi's Entropy", *Pattern Recognition, vol.* 30, pp. 71-84, 1997.

[45] M. Portes de Albuquerque et al., "Image Thresholding Using Tsallis Entropy", *Pattern Recognition Letters,* vol.25, pp.1059-1065, 2004.

[46] W. Oh and B. Lindquist, "Image Thresholding by Indicator Kriging", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 21, pp. 590-602, 1999.

[47] J. Sauvola and M. Pietikainen, "Adaptive Document Image Binarization", *Pattern Recognition*, vol.33, pp. 225-236, 2000.

[48] Y. Solihin and C. G. Leedham, "Integral Ratio: A New Class of Global Thresholding Techniques for Handwriting Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 21, pp. 761-768, 1999.

[49] Y.Yang and H. Yan, "An Adaptive Logical Method for Binarization of Degraded Document Images", *Pattern Recognition*, vol.33, pp. 787-807, 2000.

[50] C. Wolf and J.M. Jolian, "Extraction and Recognition of Artificial Text in Multimedia Documents", RFV RR, 2002.

[51] G. Leedham et al., "Comparison of Some Thresholding Algorithms for Text/Background Segmentation in Difficult Document Image", *Proceedings of Seventh International Conference on Document Analysis and Recognition (ICDAR'03)*, pp. 859-864, 2003.

[52] E. Badekas and N. Papamarkos, "A System for Document Binarization", *Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis*, pp. 909-914, 2003.

[53] M. Sezgin and B. Sankur, "Image Multi-Thresholding Based on Sample Moment Function", *International Conference on Image Processing*, vol. 2, pp. 415-418, 2003.

[54] I. K. Kim, D. W. Jung, and R. H. Park, "Document Image Binarization based on Topographic Analysis Using a Water Flow Model", *Pattern Recognition*, Elsevier Science, vol. 35, pp. 265-277, 2002.

[55] E. Kavallieratou, "A Binarization Algorithm Specialized on Document Images and Photos", *Proceedings of the Eight International Conference on Document Analysis and Recognition* (ICDAR'05), vol. 1, pp. 463-467, 2005.

[56] E. Kavallieratou and H. Antonopoulou, "Cleaning and Enhancing Historical Document Images", *Lecture Notes in Computer Scinece*, Springer, vol. 3708, pp. 681-688, 2005.

[57] Y. Chen and G. Leedham, "Decompose Algorithm for Thresholding Degraded Historical Document Images", *IEE Proc. Vis. Image Signal Process.*, vol. 152, no. 6, pp. 702-714, 2005.

[58] A. Khashman and B. Sekeroglu, "Enhancement of Unclear Patterns Using Pattern Average Thresholding", *Proceedings of the 3rd International Symposium on Electrical, Electronic & Computer Engineering*, vol. 1, pp. 253-257, 2006.

[59] W.S. Ng, C.K. Lee, "Comment on Using the Uniformity Measure for Performance Measure in Image Segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 9, pp. 933-934, 1996.

[60] T. Luo et al., "Recognizing Plankton Images From the Shadow Image Particle Profiling Evaluation Recorder", *IEEE Transactions on Systems, Man, And Cybernetics—Part B: Cybernetics*, vol. 34, no. 4, pp. 1753-1762, 2004.

[61] Feng Zhao et al., "Binary Plankton Image Classification Using Random Subspace", *IEEE International Conference on Image Processing*, vol. 1, pp. 357-360, 2005.

[62] T. Luo et al., "Learning to Recognize Plankton", *IEEE International Conference on Systems, Man and Cybernetics*, vol. 1, pp. 888-893, 2003.

[63] D. Haverkamp, L. E. Soh, and C. Tsatsoulis, "A Dynamic Local Thresholding Technique For Sea Ice Classification", *International Geoscience and Remote Sensing Symposium,* IGARSS, vol. 2, pp. 638-640, 1993.

[64] R. Bremananth, B. Balaji, M. Sankari and A. Chitra, "A New Approach to Coin Recognition using Neural Pattern Analysis", *IEEE Indicon Conference*, pp. 366-370, 2005.

[65] A. Khashman, B. Sekeroglu and K. Dimililer, "ICIS: A Novel Coin Identification System", *Lecture Notes in Control and Information Sciences*, vol. 345, Springer-Verlag, 2006.

[66] A. Khashman, B. Sekeroglu and K. Dimililer, "Intelligent Rotation-Invariant Coin Identification System", *WSEAS Transactions on Signal Processing*, Issue 5, vol. 2, pp. 781-786, 2006.

[67] A. Khashman, B. Sekeroglu and K. Dimililer, "Intelligent Coin Identification System", Proceedings of the *IEEE International Symposium on Intelligent Control* (ISIC'06), 2006.

[68] A. Khashman, B. Sekeroglu and K. Dimililer, "Rotated Coin Recognition Using Neural Network", *Advances in Soft Computing*, Springer, vol.41, pp. 290-297, 2007.

[69] V. Onnia, and M. Tico, "Adaptive Binarization Method for Fingerprint Images", *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, (ICASSP), vol. 4, pp. 3692-3695, 2002.

[70 P. Meenen, and R. Adhami, "Approaches to Image Binarization in Current Automated Fingerprint Identification Systems", *Proceedings of the Thirty-Seventh Southeastern Symposium on System Theory*, pp. 276-281, 2005.

[71] Y. Zhang and Q. Xiao, "An Optimized Approach for Fingerprint Binarization", *2006 International Joint Conference on Neural Networks*, pp. 391-395, 2006.

[72] J. Gao and M. Xie, "The Layered Segmentation, Gabor Filtering and Binarization Based on Orientation for Fingerprint Preprocessing", *The 8th International Conference on Signal Processing*, vol. 4, 2006.

[73] X. Jiang amd D. Mojon, "Adaptive Local Thresholding by Verification-Based Multithreshold Probing with Application to Vessel Detection in Retinal Images", *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 1, pp. 131-137, 2003.

[74] H. Zhu, F. H. Y. Chan, F. K. Lam, and K. Y. Lam, "Segmentation of Pathology Microscopic Images" *Proceedings of 19th IEEE/EMBS International Conference*, pp. 580-581, 1997.

[75] A. Ruggeri, S. Pajaro, and A. Vita, "Classification of Corneal Layers in Confocal Microscopy", *Proceedings of the 22. Annual EMBS International Conference*, pp. 1030-1032, 2000.

[76] E. Espinoza, G. Martinez, J. G. Frerichs, and T. Scheper, "Cell Cluster Segmentation based on Global and Local Thresholding for in-situ Microscopy", *3rd IEEE International Symposium on Biomedical Imaging: Macro to Nano*, pp. 542-545, 2006.

[77] T. Sund and K. Eilertsen, "An Algorithm for Fast Adaptive Image Binarization with Applications in Radiotherapy Imaging", *IEEE Transactions on Medical Imaging*, vol. 22, no. 1, pp.22-28, 2003.

[78] Y. Ebrahim, "Entropy based thresholding of cross-dissolved ultrasound images", *IEEE Canadian Conference on Electrical and Computer Engineering*, vol. 3, pp.1477-1480, 2003.

[79] N. Hiransakolwong, K. A. Hua, K. Vu, and P. S. Windyga, "Segmentation of Ultrasound Liver Images: An Automatic Approach", *2003 International Conference on Multimedia and Expo*, vol. 1, pp. 573-76, 2003.

[80] A.I. Konkachbaev, M.F. Casanova, J.H. Graham, and A.S. Elmaghraby, "Automated Recursive Segmentation of Large Neocortical Images Using Standard Deviation as Termination Criteria", *Proceedings of the 27th IEEE Annual Conference on Engineering in Medicine and Biology*, pp.2531-2534, 2005.

[81] H. J. Jeong et al., "Comparison of Thresholding Methods for Breast Tumor Cell Segmentation", *Proceedings of 7th International Workshop on Enterprise networking and Computing in Healthcare Industry*, pp. 392-395, 2005.

[82] J. Baudewig, P. Dechent, K.D. Merboldt, and J. Frahm, "Thresholding in Correlation Analyses of Magnetic Resonance Functional Neuroimaging", *Magnetic Resonance Imaging,* Elsevier Science, vol. 21, pp.1121-1130, 2003.

[83] C. Varelaa et al., "Computerized Detection of Breast Masses in Digitized Mammograms", *Computers in Biology and Medicine*, Elsevier Science, vol. 37, pp. 214-226, 2007.

[84] A. F. Obrist, A. Flisch, and J. Hofmann, "Point Cloud Reconstruction with Sub-Pixel Accuracy By Slice-Adaptive Thresholding of X-Ray Computed Tomography Images", *NDT&E International*, Elsevier Science, vol.37, pp. 373-380, 2004.

[85] Z. Zhoua, and Z. Ruana, "Multicontext Wavelet-Based Thresholding Segmentation of Brain Tissues in Magnetic Resonance Images", *Magnetic Resonance Imaging*, Elsevier Science, vol. 25, pp. 381-385, 2007.

[86] D. Y. Kim, and J. W. Park, "Connectivity-Based Local Adaptive Thresholding For Carotid Artery Segmentation Using MRA Images", *Image and Vision Computing*, Elsevier Science, vol. 23, pp. 1277-1287, 2005.

[87] N. Sang, H. Li, W. Peng, and T. Zhang, "Knowledge-Based Adaptive Thresholding Segmentation of Digital Subtraction Angiography Images", *Image and Vision Computing*, Elsevier Science, vol. 25 , pp. 1263-1270, 2007.

[88] R. Rodriguez, "A Strategy For Blood Vessels Segmentation based on the Threshold which Combines Statistical and Scale Space Filter Application to the Study of Angiogenesis", *Computer Methods and Programs in Biomedicine*, Elsevier Science, vol. 82, pp.1-9, 2006.

[89] O.D. Trier and T. Taxt, "Evaluation of Binarization Methods For Document Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 312-315, 1995.

[90] O.D. Trier and A.K. Jain, "Goal-directed Evaluation of Binarization Methods", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 1191-1201, 1995.

[91] M. Sezgin and B. Sankur, "Survey Over Image Thresholding Techniques and Quantitative Performance Evaluation", *Journal of Electronic Imaging*, vol.13, pp. 146-165, 2004.

[92] J.He et al., "A Comparison of Binarization Methods for Historical Archive Documents", *Proceedings of Eight International Conference on Documents Analysis and Recognition*, vol. 1, pp. 538-542, 2005.

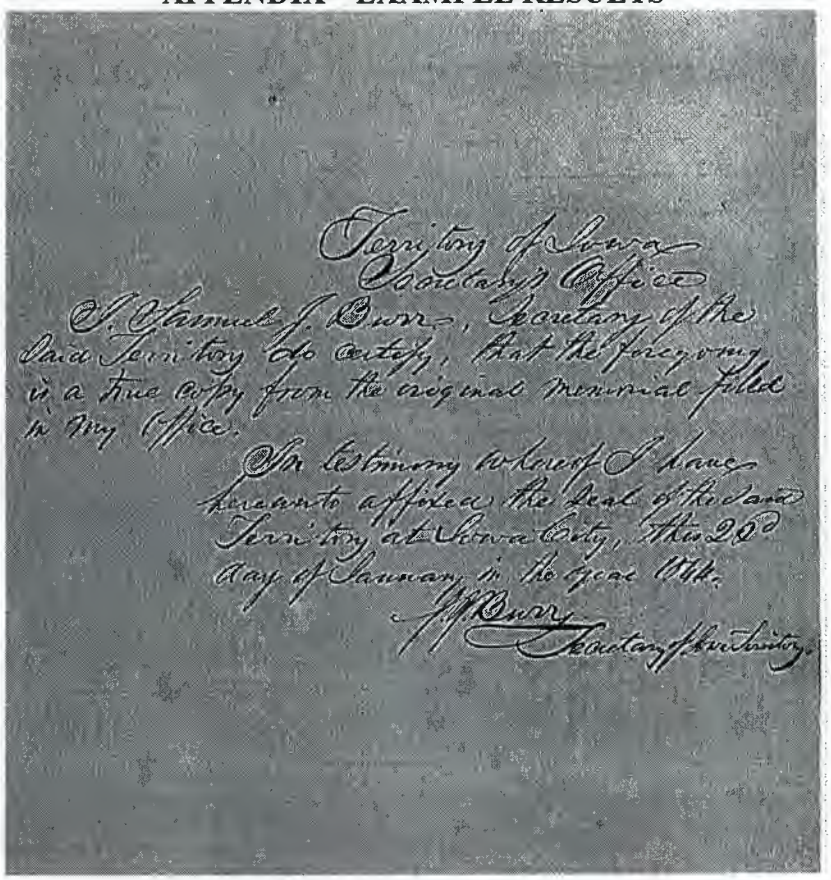
[93] K.V. Mardia and T.J. Hainsworth, "A Spatial Thresholding Method for Image Sgmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, pp. 919-927, 1988.

[94] J.R. Parker, "Gray Level Thresholding in Badly Illuminated Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.13, pp. 813-819, 1991.

[95] O.D. Trier and T. Taxt, "Improvement of 'Integrated Function Algorithm' for Binarization of Document Images", *Pattern Recognition Letters*, vol. 16, pp. 277-283, 1995.

[96] M. Junker and R. Hoch, "On the Evaluation of Document Analysis Components by Recall, Precision and Accuracy", *Proceedings of Fifth International Conference on Documents Analysis and Recognition*, pp. 713-716, 1999.

APPENDIX – EXAMPLE RESULTS



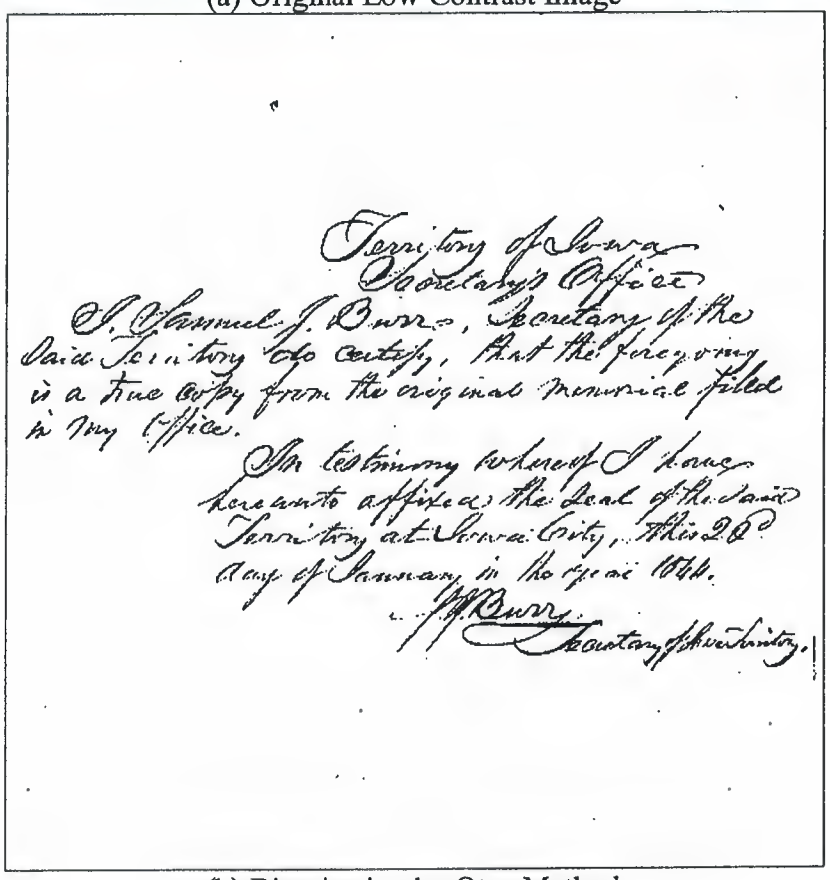
Territory of Iowa
Secretary's Office

I, Samuel J. Burr, Secretary of the said Territory do certify, that the foregoing is a true copy from the original memorial filed in my Office.

In testimony whereof I have hereunto affixed the Seal of the said Territory at Iowa City, this 20th day of January in the year 1844.

J. J. Burr
Secretary of the Territory.

(a) Original Low Contrast Image



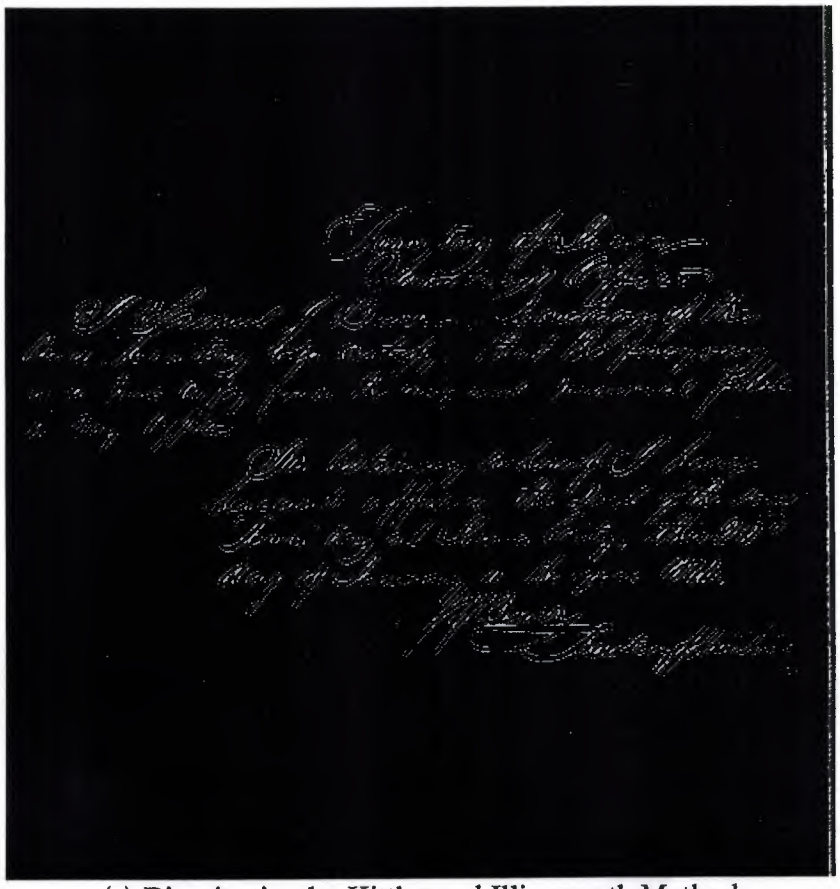
Territory of Iowa
Secretary's Office

I, Samuel J. Burr, Secretary of the said Territory do certify, that the foregoing is a true copy from the original memorial filed in my Office.

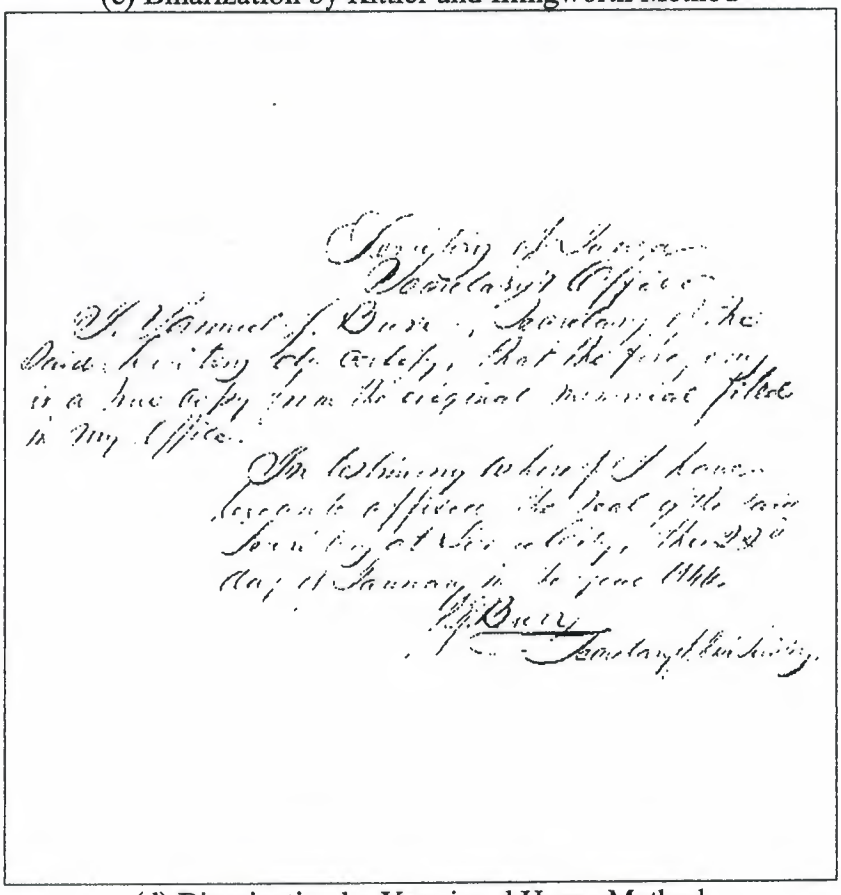
In testimony whereof I have hereunto affixed the Seal of the said Territory at Iowa City, this 20th day of January in the year 1844.

J. J. Burr
Secretary of the Territory.

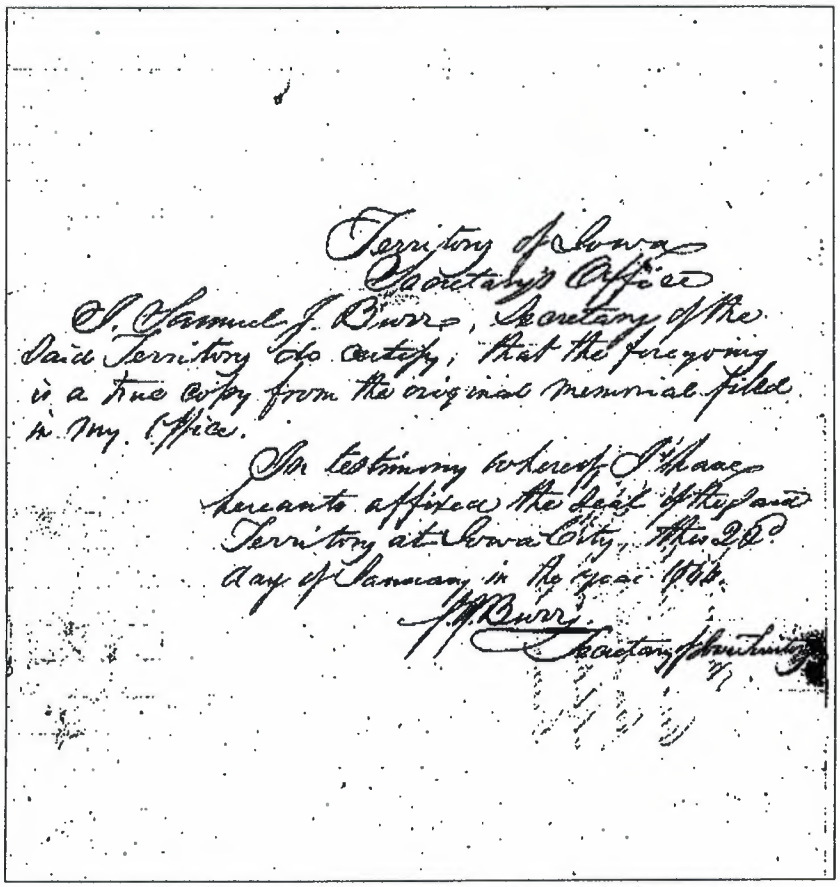
(b) Binarization by Otsu Method



(c) Binarization by Kittler and Illingworth Method

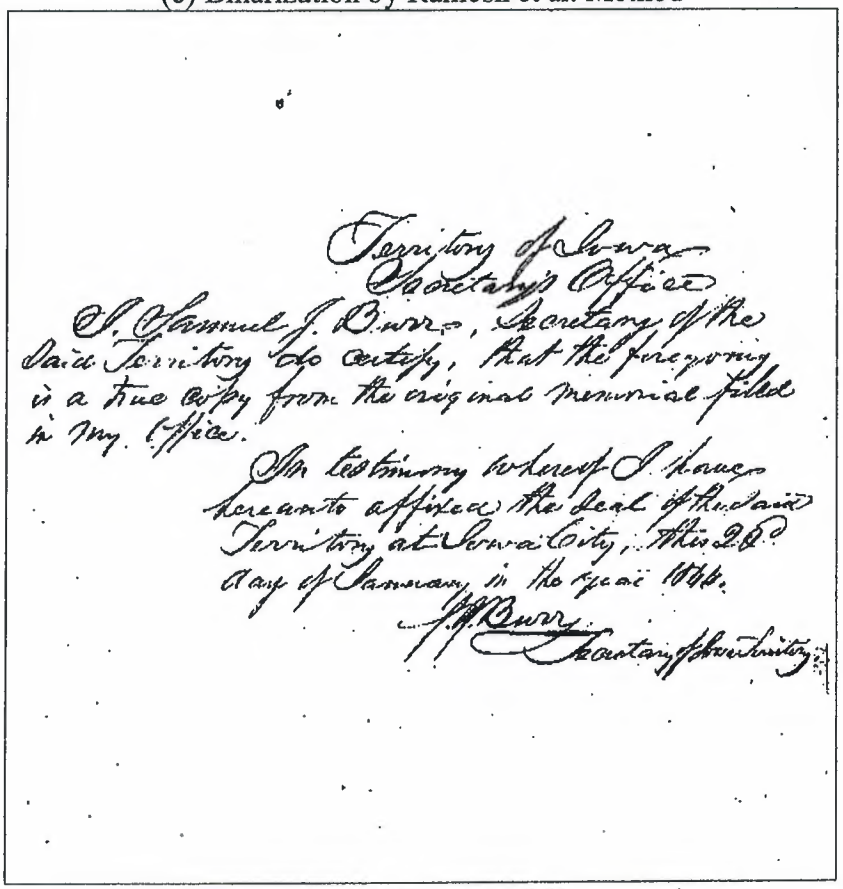


(d) Binarization by Yanni and Horne Method



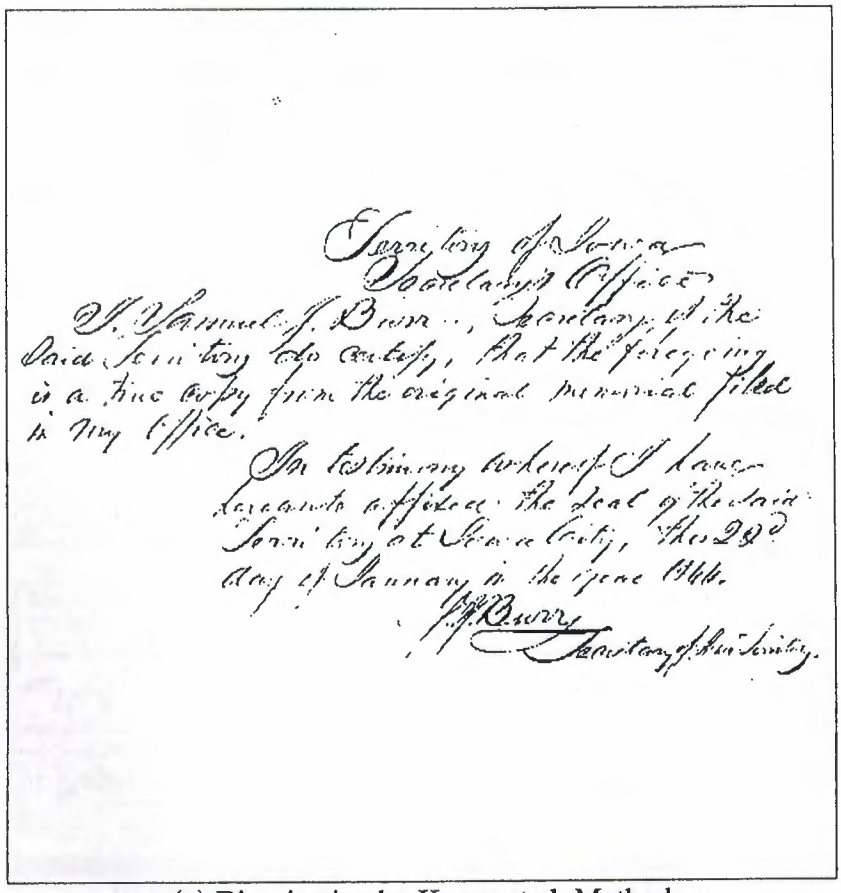
Territory of Iowa
Secretary's Office
I, Samuel J. Burr, Secretary of the
said Territory do certify, that the foregoing
is a true copy from the original memorial filed
in my Office.
In testimony whereof I have
hereunto affixed the seal of the said
Territory at Iowa City, this 28
day of January in the year 1846.
S. J. Burr
Secretary of Iowa Territory

(e) Binarization by Ramesh et al. Method

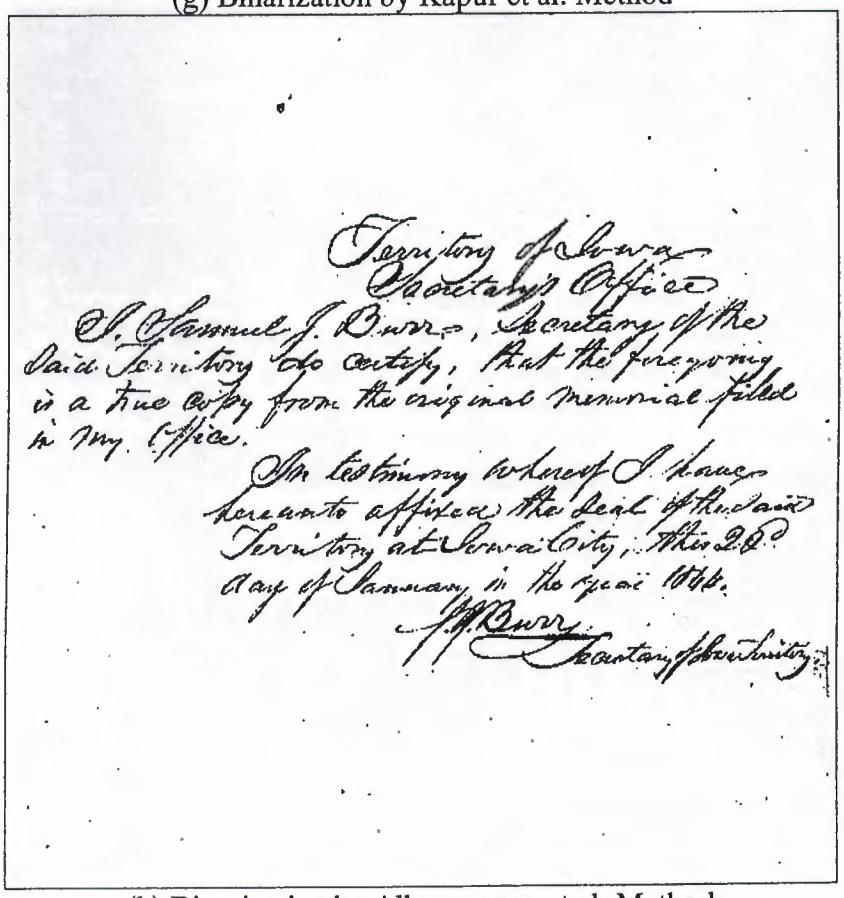


Territory of Iowa
Secretary's Office
I, Samuel J. Burr, Secretary of the
said Territory do certify, that the foregoing
is a true copy from the original memorial filed
in my Office.
In testimony whereof I have
hereunto affixed the seal of the said
Territory at Iowa City, this 28
day of January in the year 1846.
S. J. Burr
Secretary of Iowa Territory

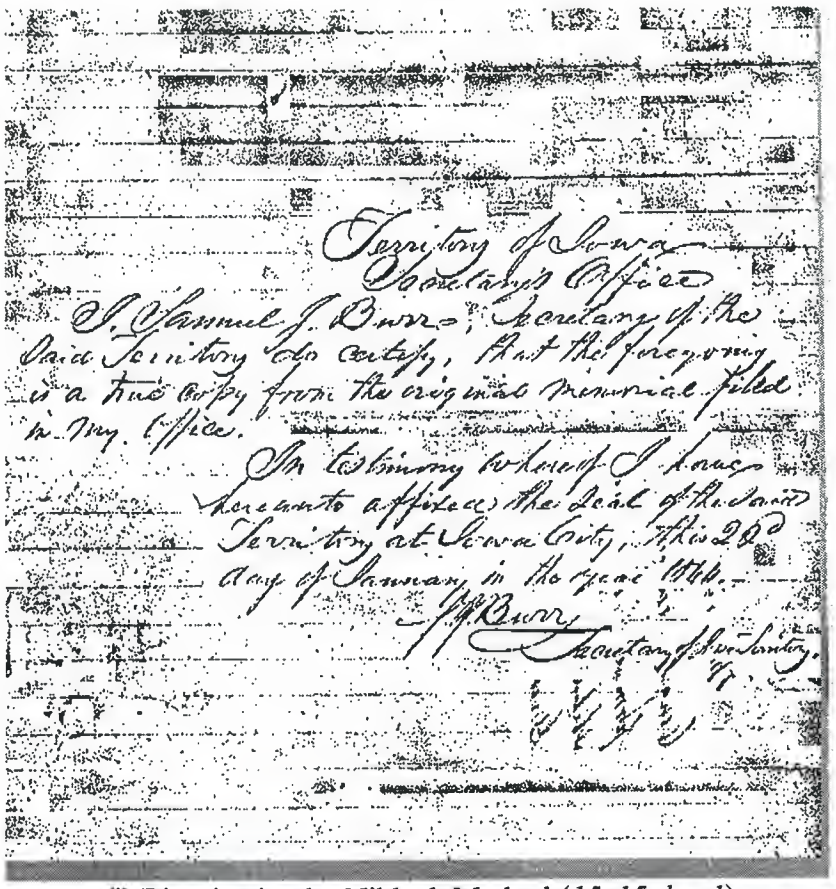
(f) Binarization by Proposed MD Method



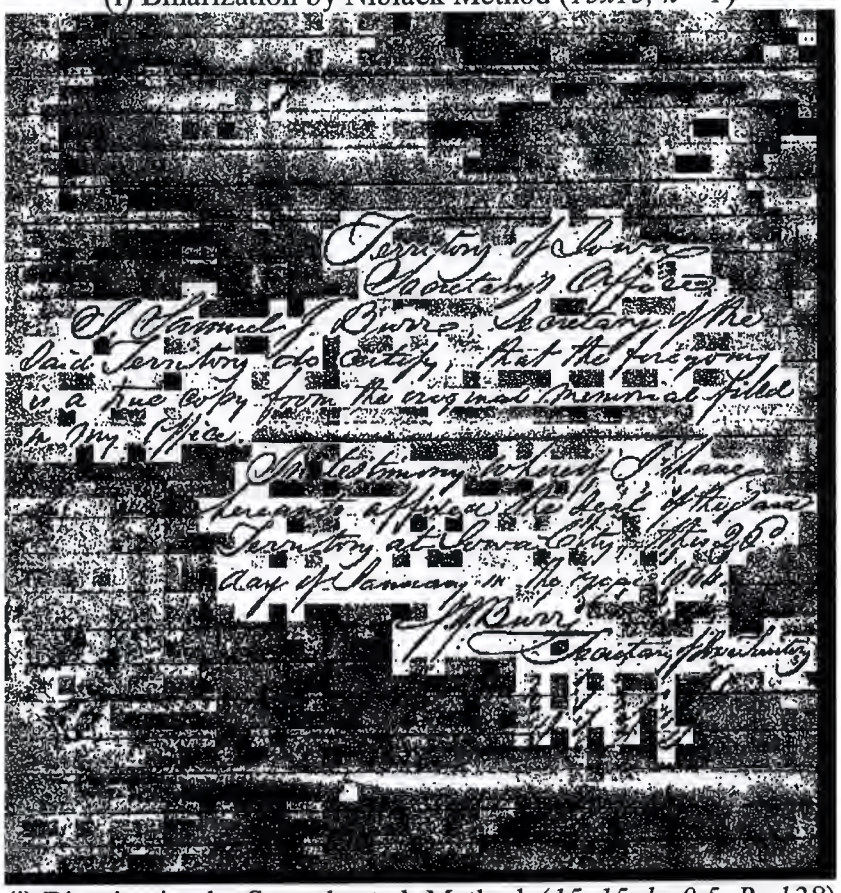
(g) Binarization by Kapur et al. Method



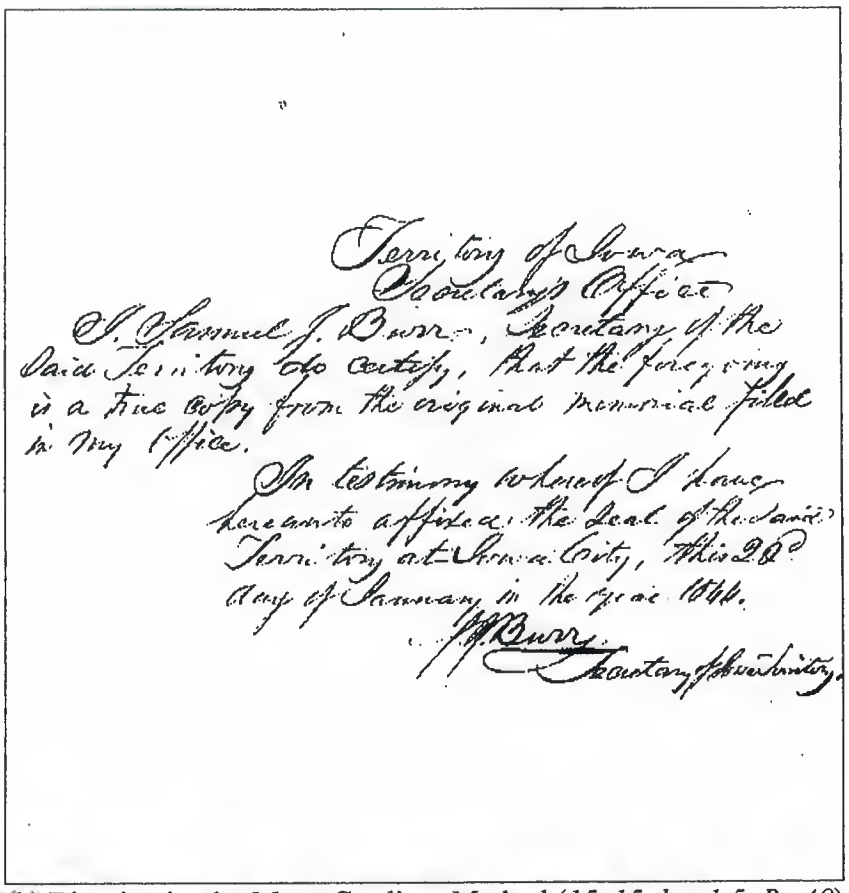
(h) Binarization by Albuquerque et al. Method



(i) Binarization by Niblack Method (*15x15, k=-1*)



(j) Binarization by Sauvola et al. Method (*15x15, k=0.5, R=128*)



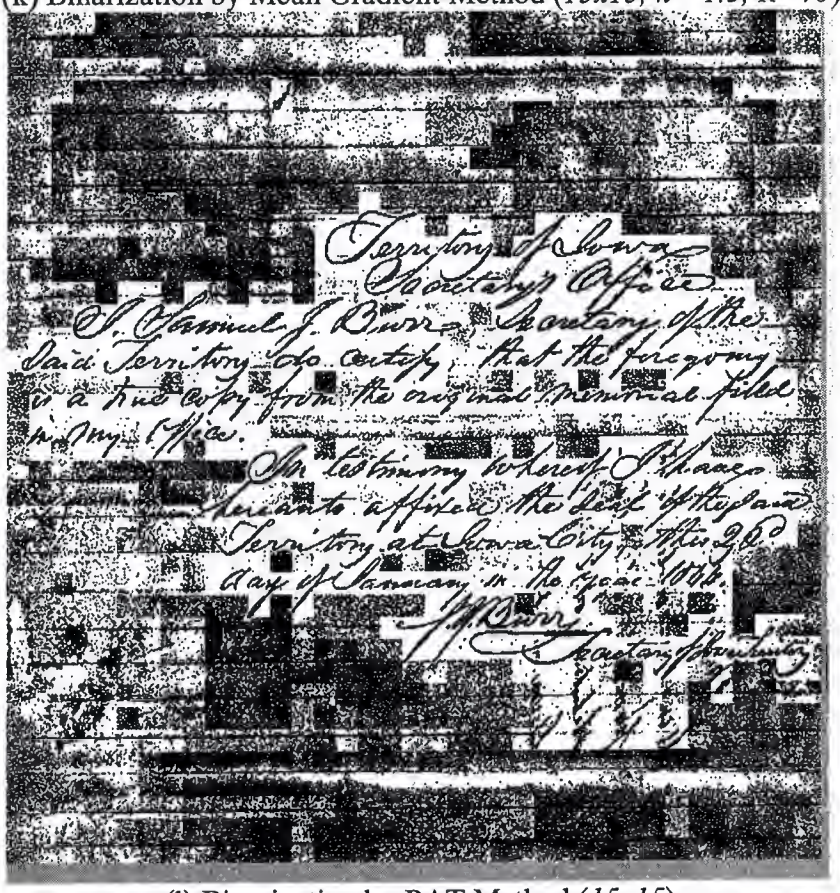
Territory of Iowa
Secretary's Office

I, Samuel J. Burr, Secretary of the
said Territory do certify, that the foregoing
is a true copy from the original Memorial filed
in my Office.

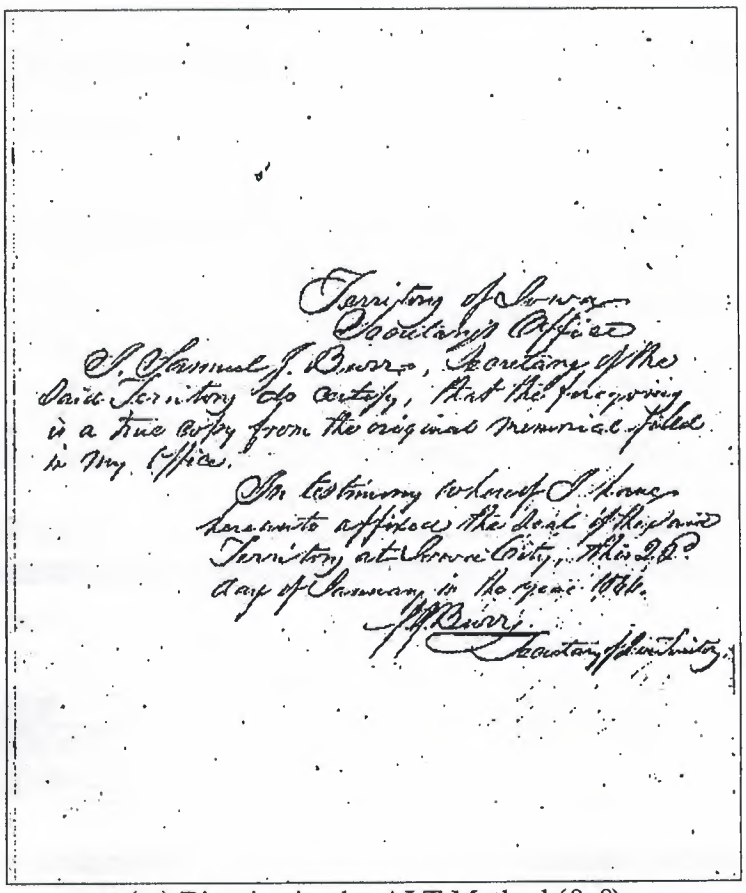
In testimony whereof I have
hereunto affixed the Seal of the said
Territory at Iowa City, this 28
day of January in the year 1844.

J. J. Burr
Secretary of Iowa Territory.

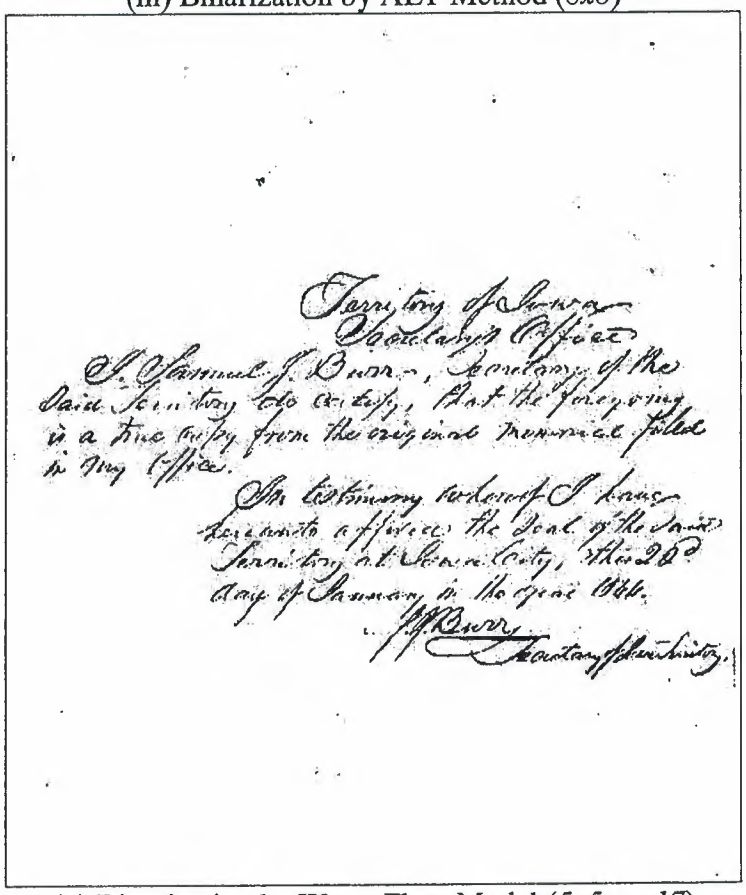
(k) Binarization by Mean Gradient Method (*15x15, k=-1.5, R=40*)



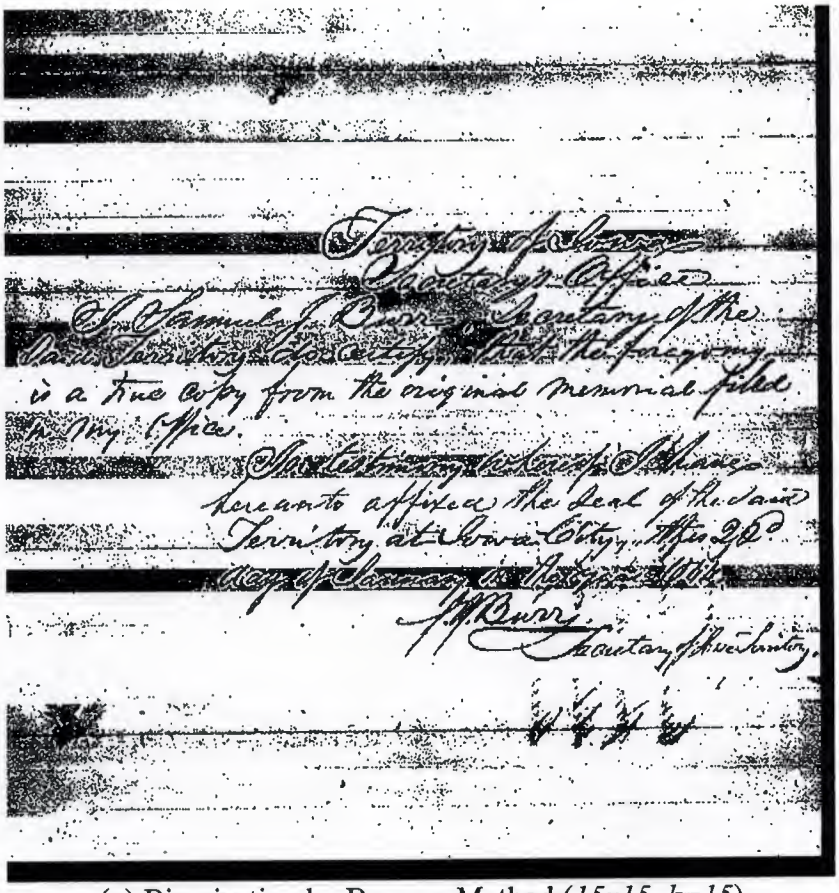
(l) Binarization by PAT Method (*15x15*)



(m) Binarization by ALT Method (*8x8*)

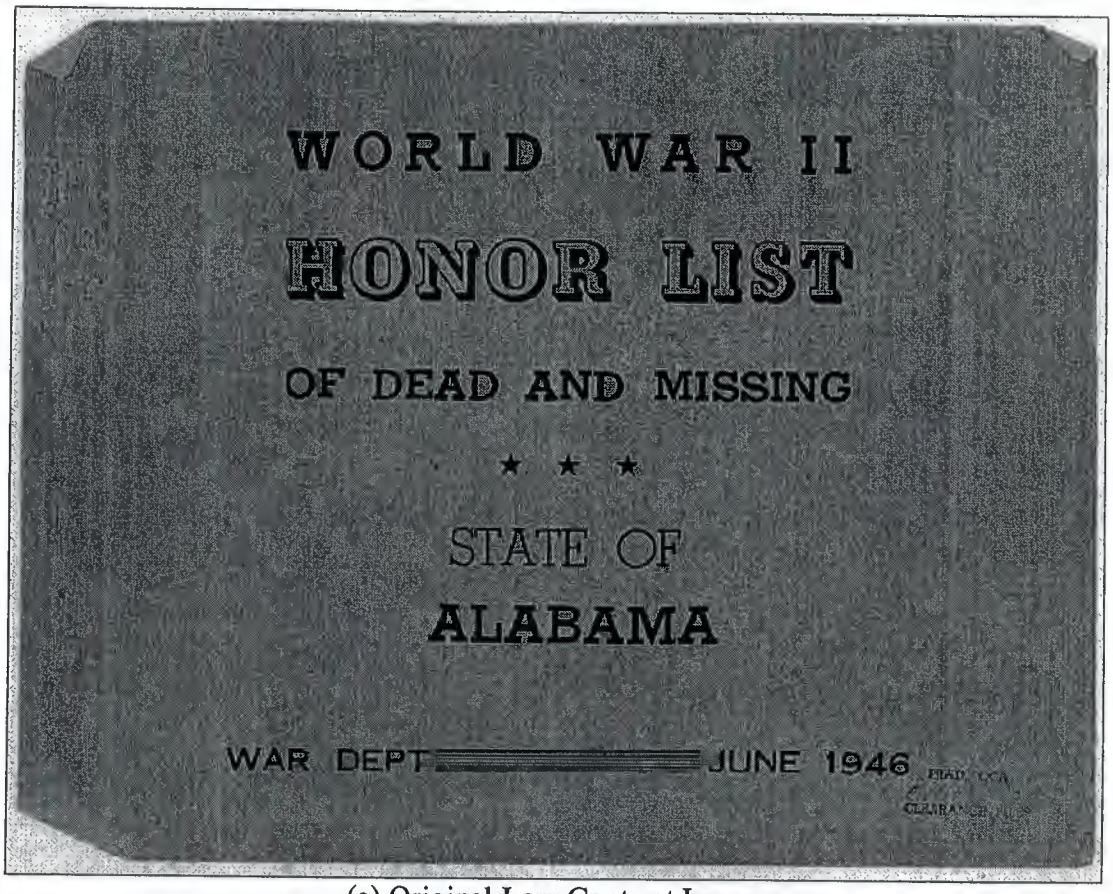


(n) Binarization by Water Flow Model (*5x5,w=17*)

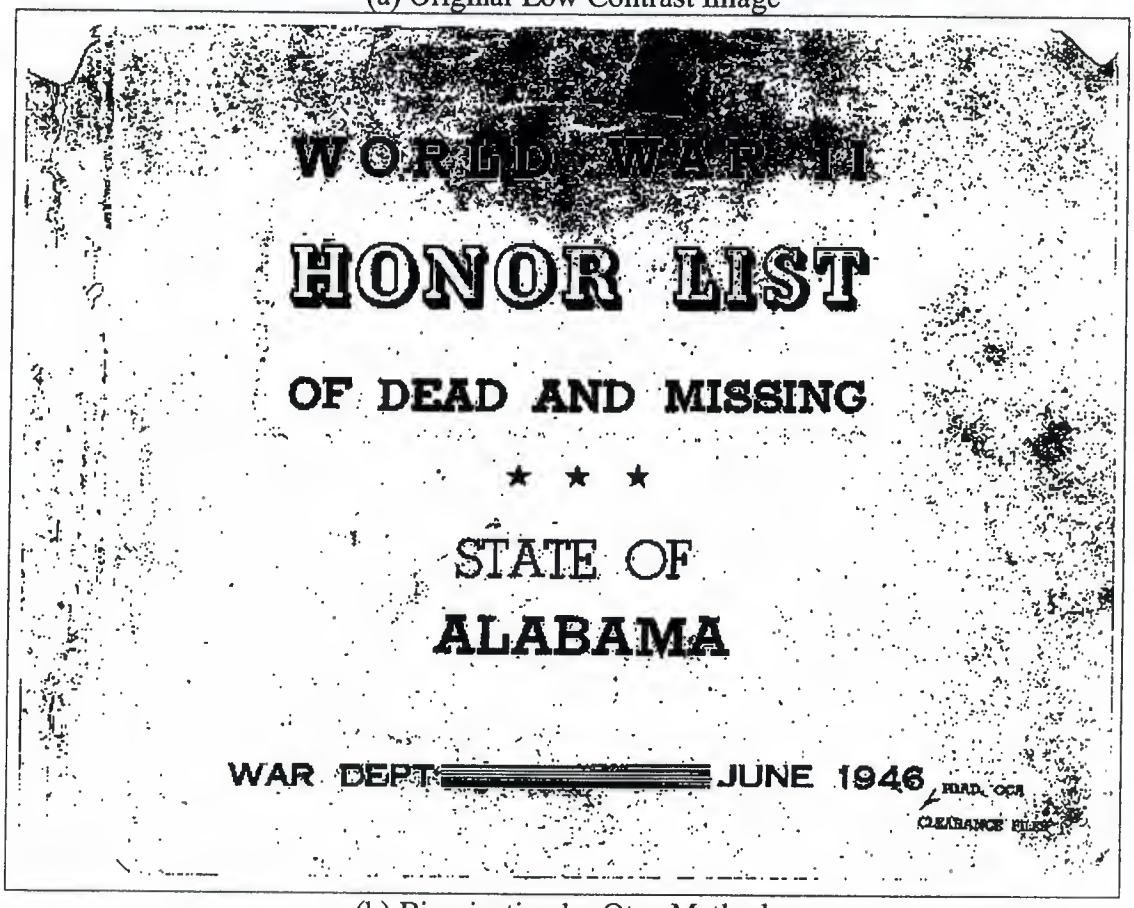


(o) Binarization by Bernsen Method (*15x15, k=15*)

Figure A1-1 Example Result of Bright Image of Test Set I

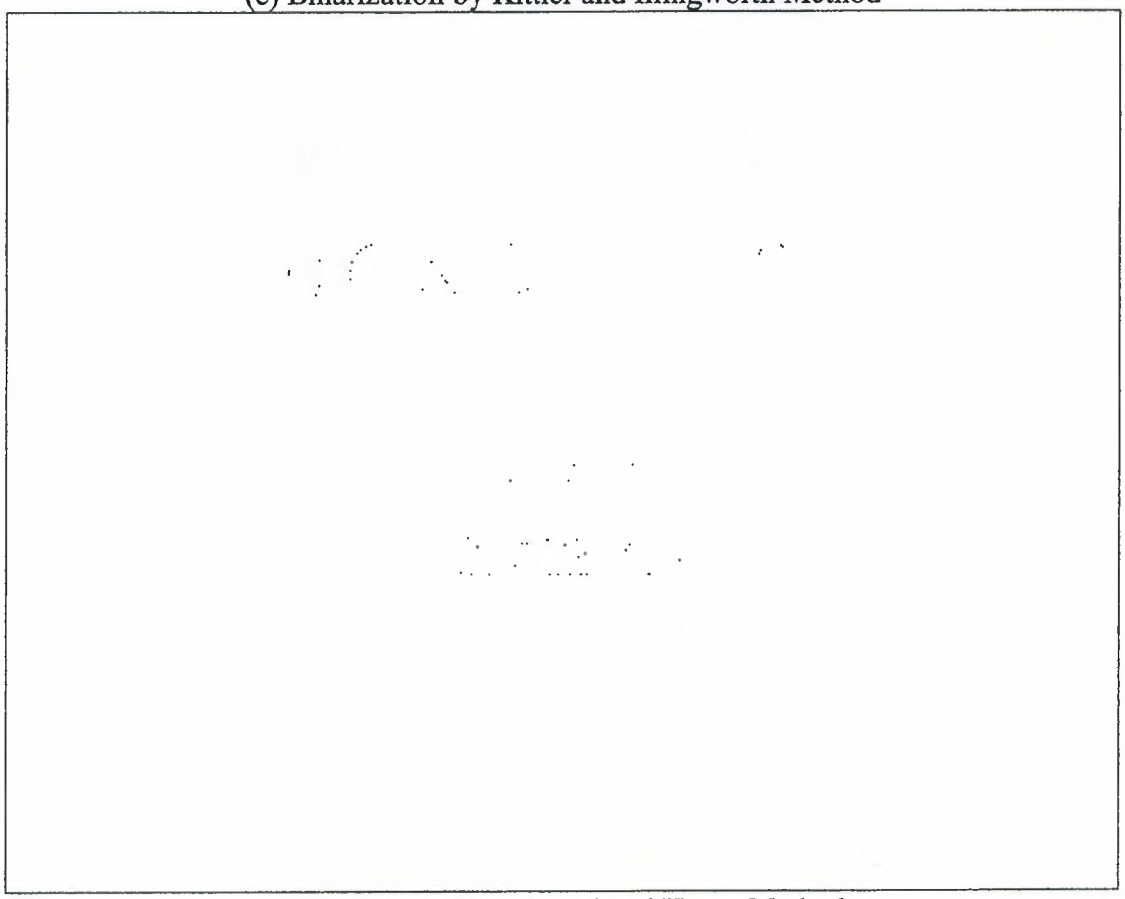


(a) Original Low Contrast Image

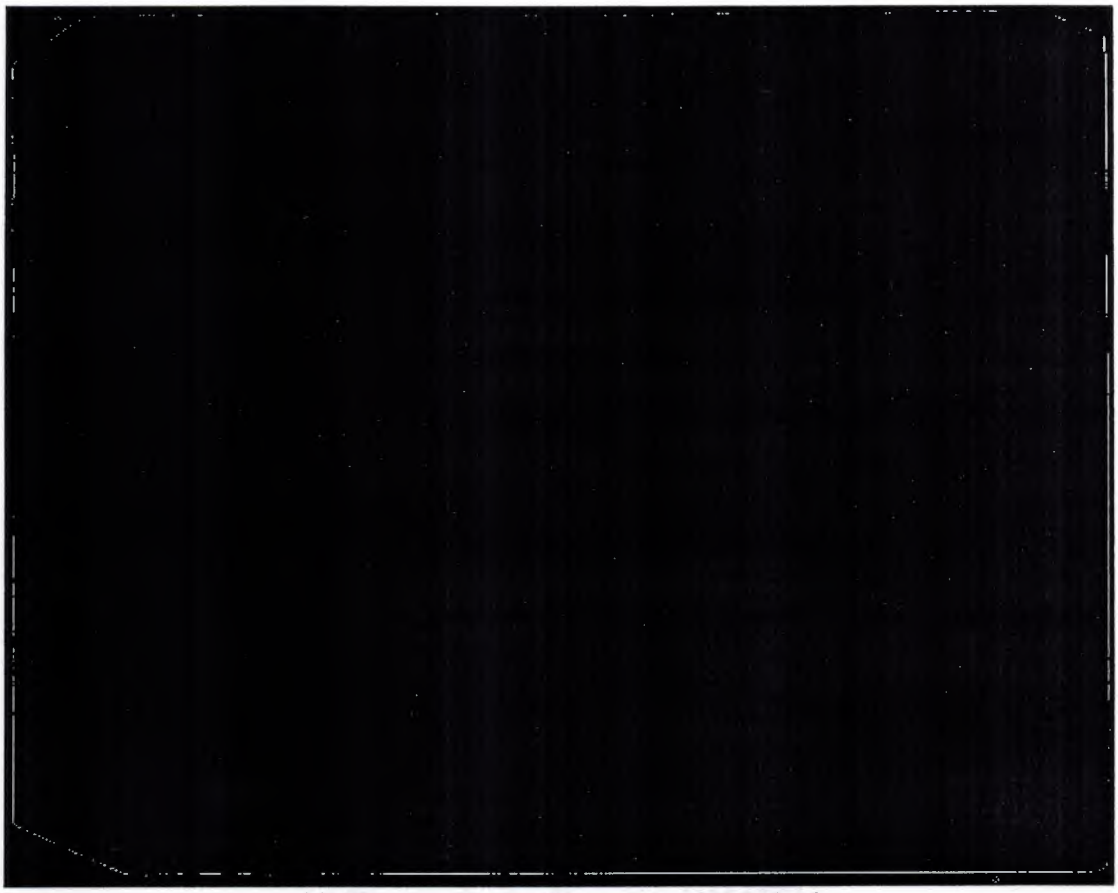


(b) Binarization by Otsu Method

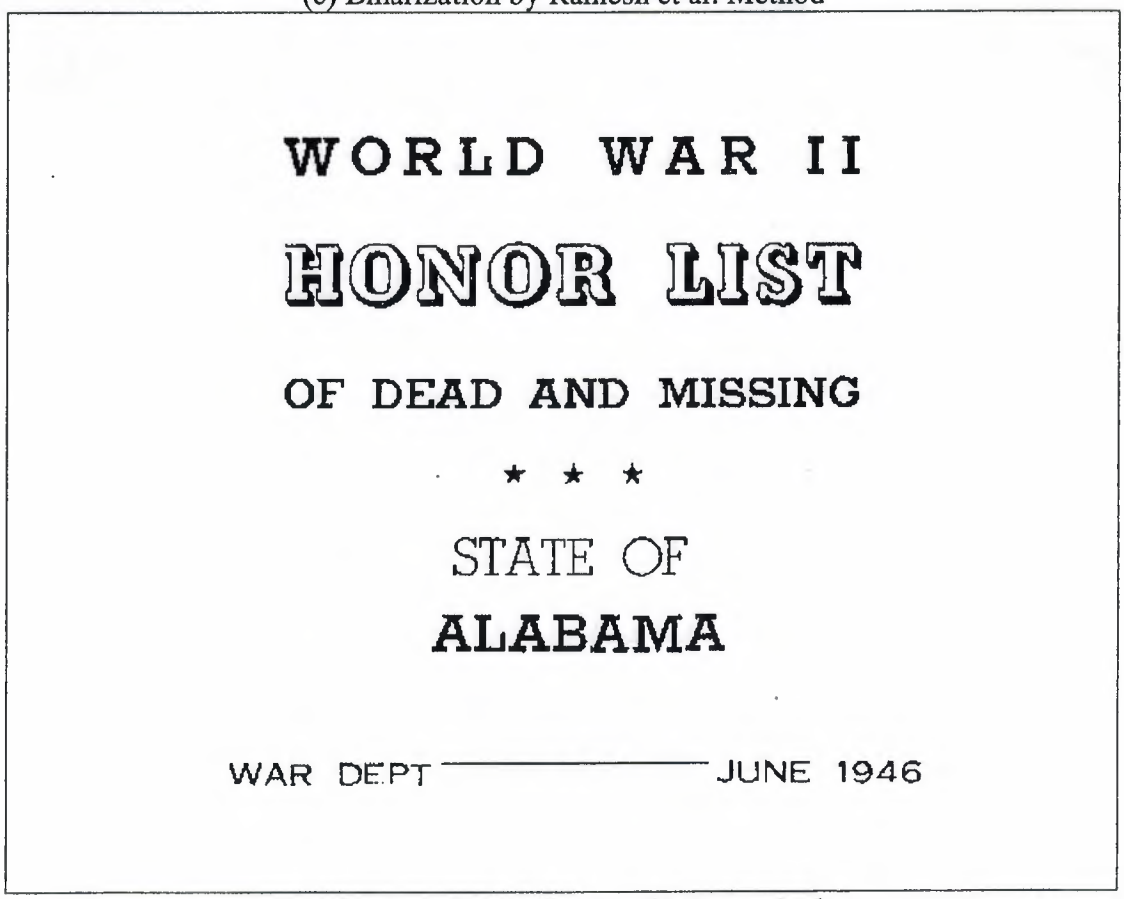
(c) Binarization by Kittler and Illingworth Method



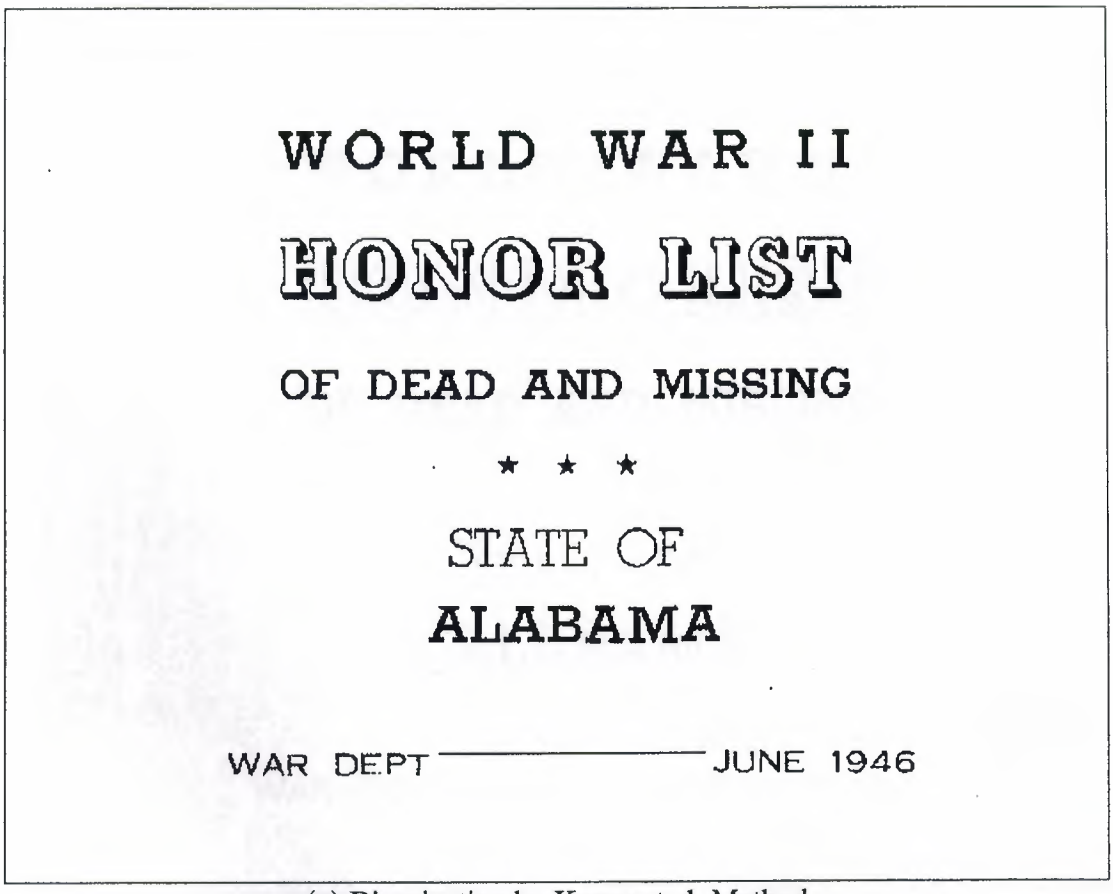
(d) Binarization by Yanni and Horne Method



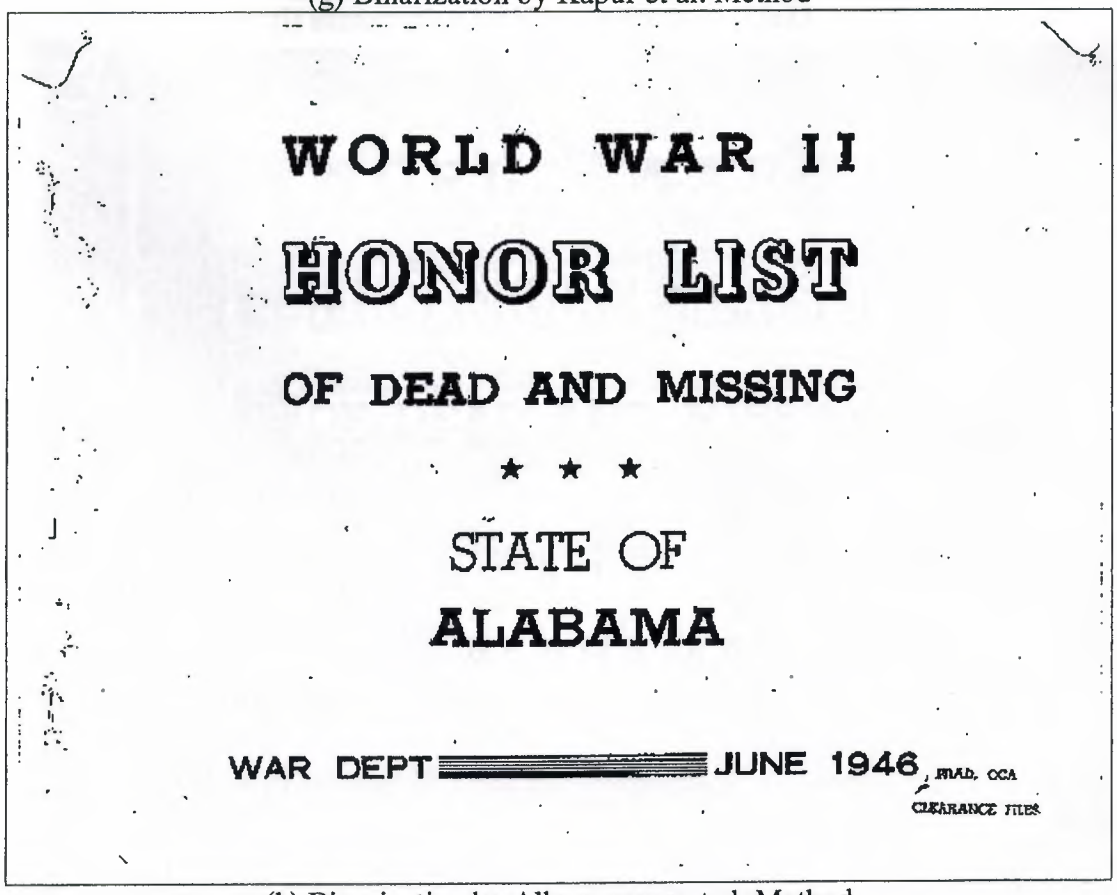
(e) Binarization by Ramesh et al. Method



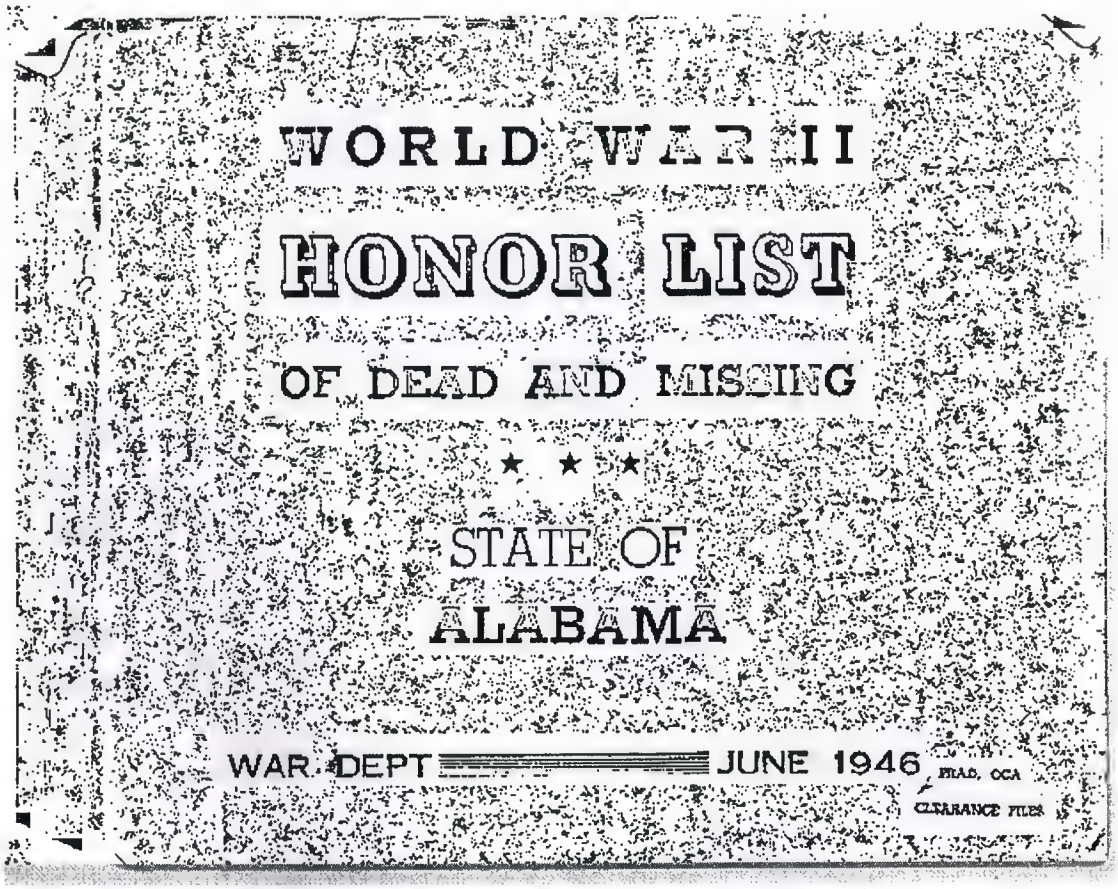
(f) Binarization by Proposed MD Method



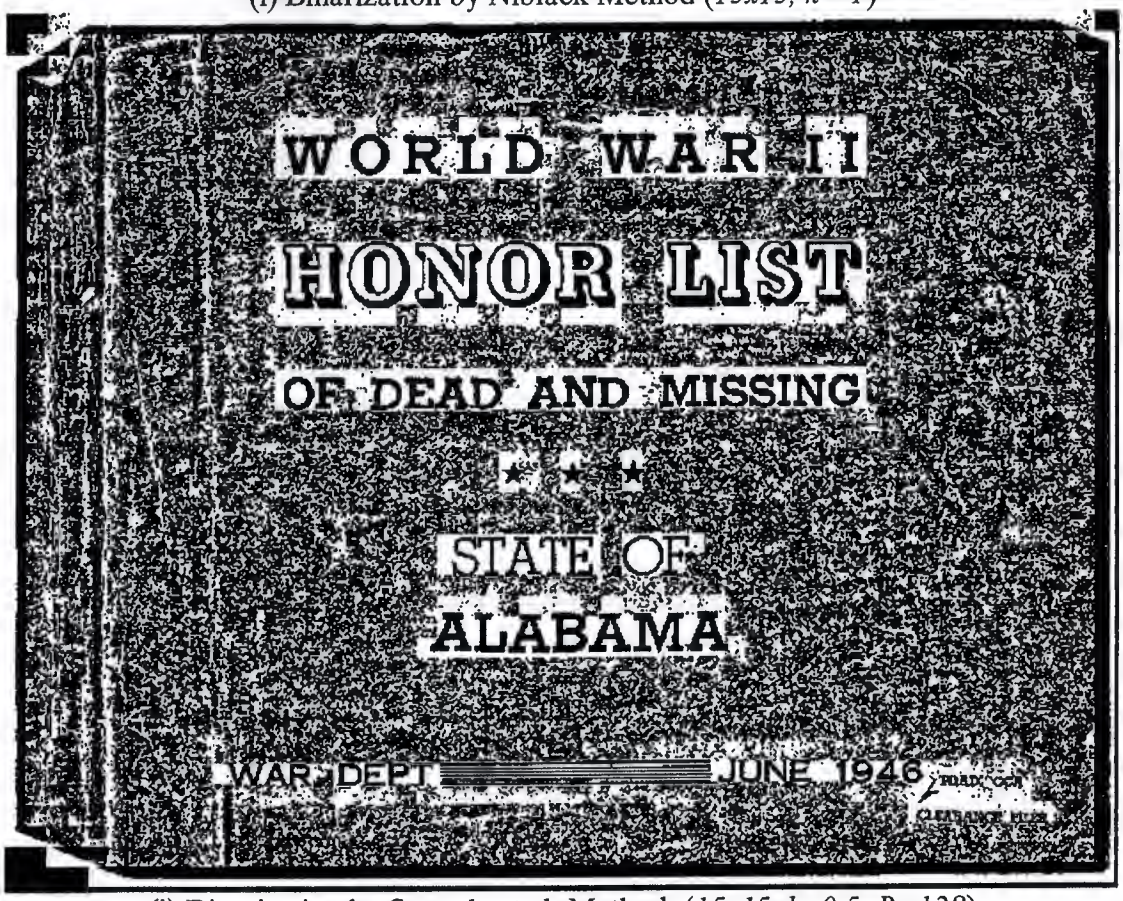
(g) Binarization by Kapur et al. Method



(h) Binarization by Albuquerque et al. Method

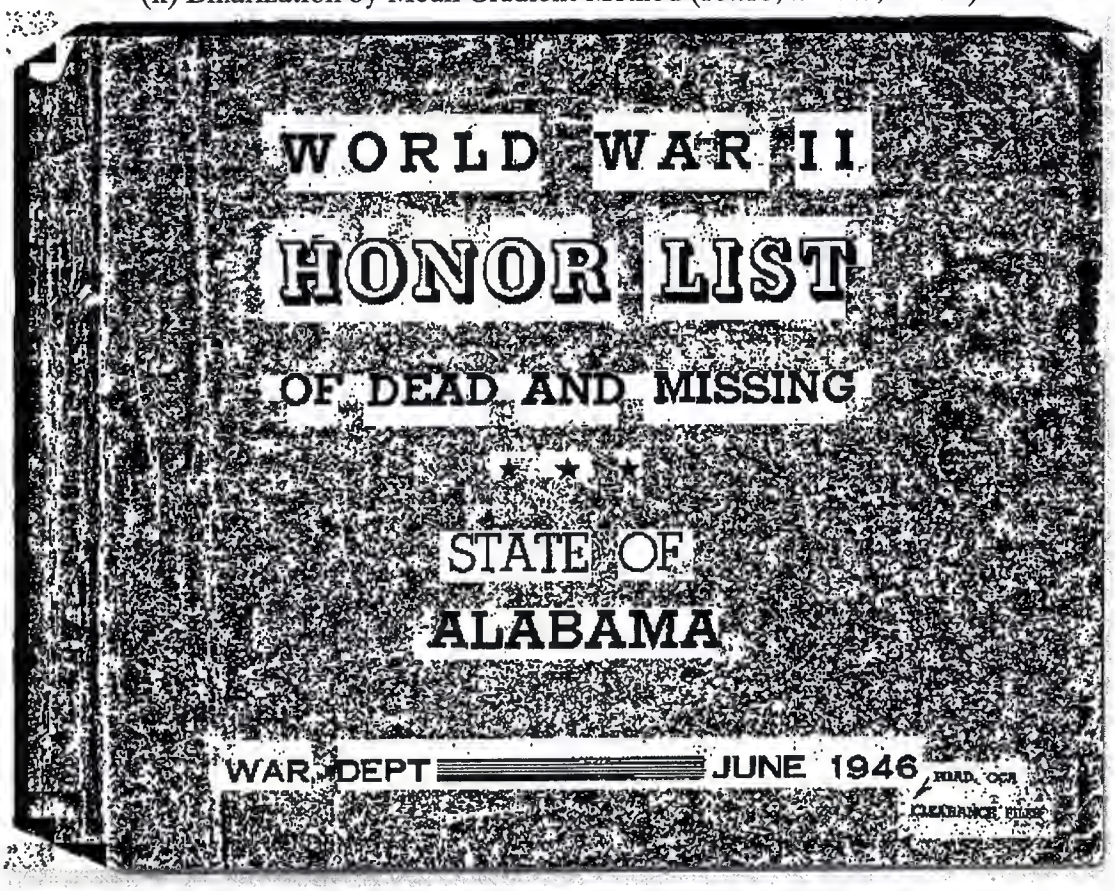


(i) Binarization by Niblack Method (*15x15, k=-1*)

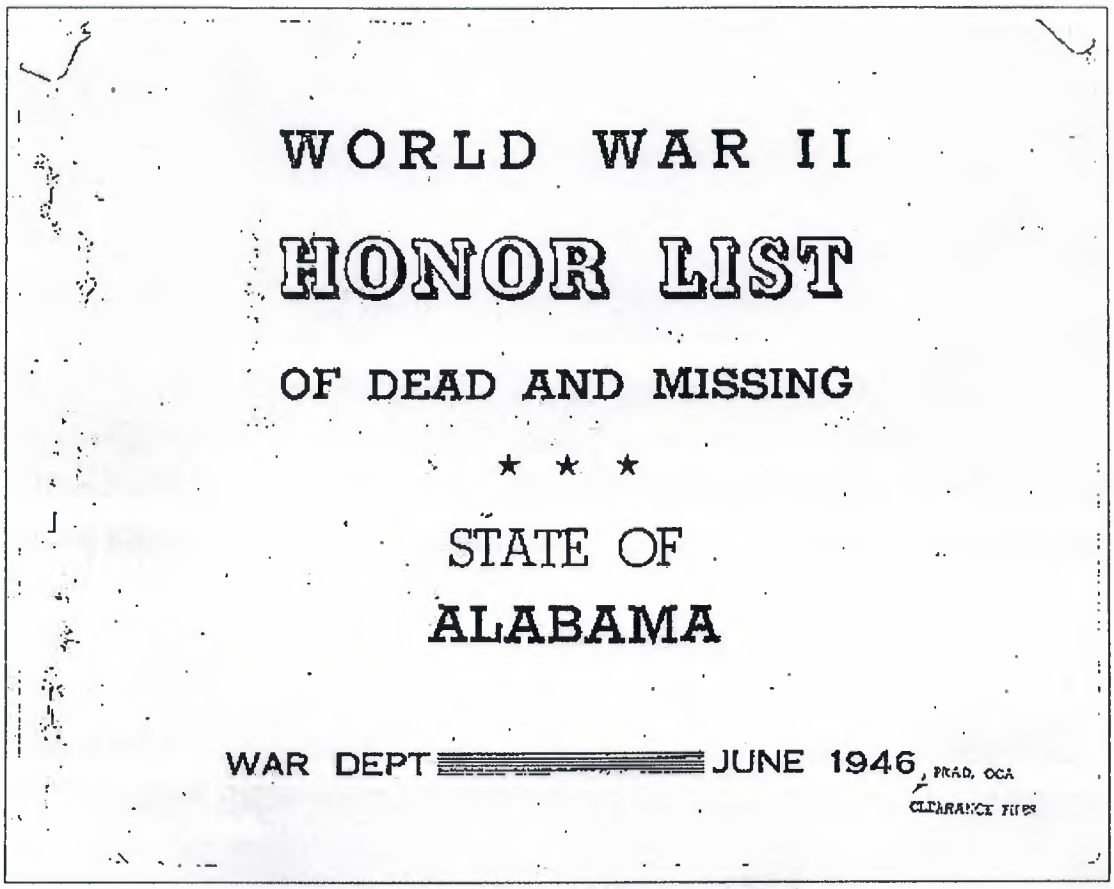


(j) Binarization by Sauvola et al. Method (*15x15, k=0.5, R=128*)

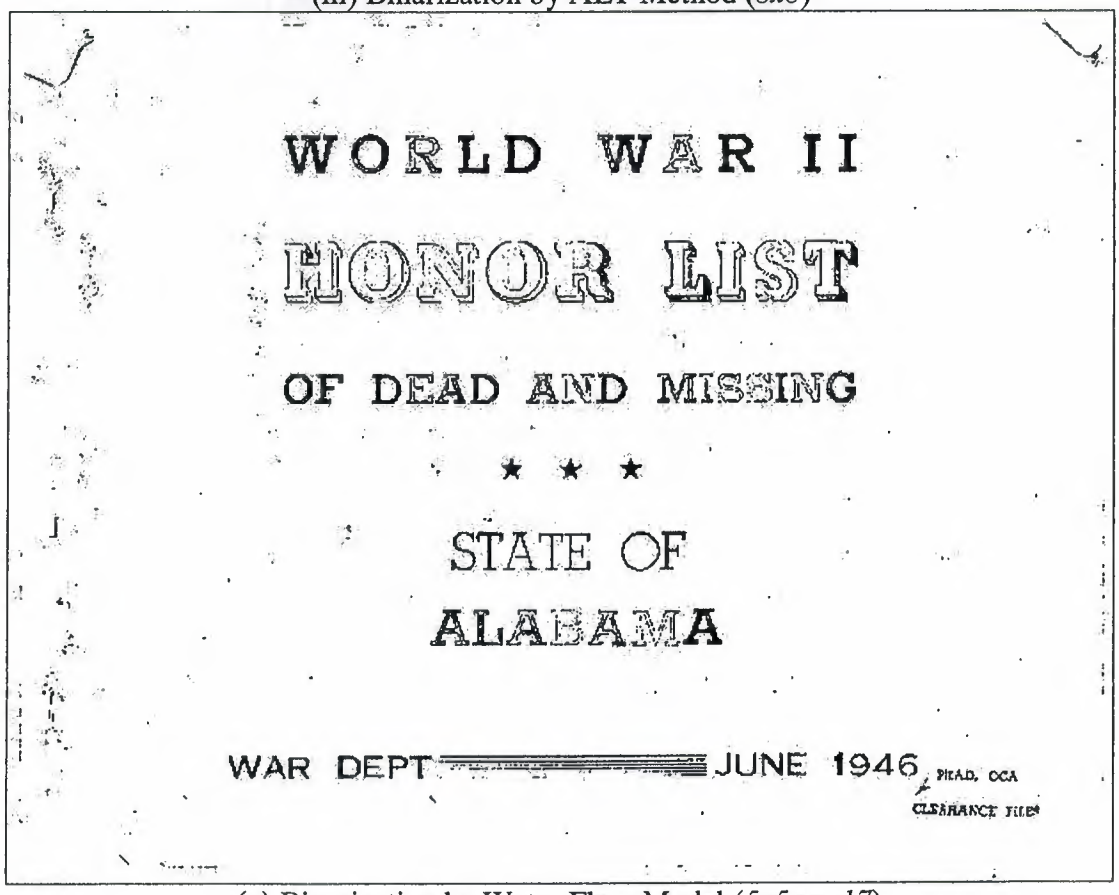
(k) Binarization by Mean Gradient Method (*15x15, k=-1.5, R=40*)



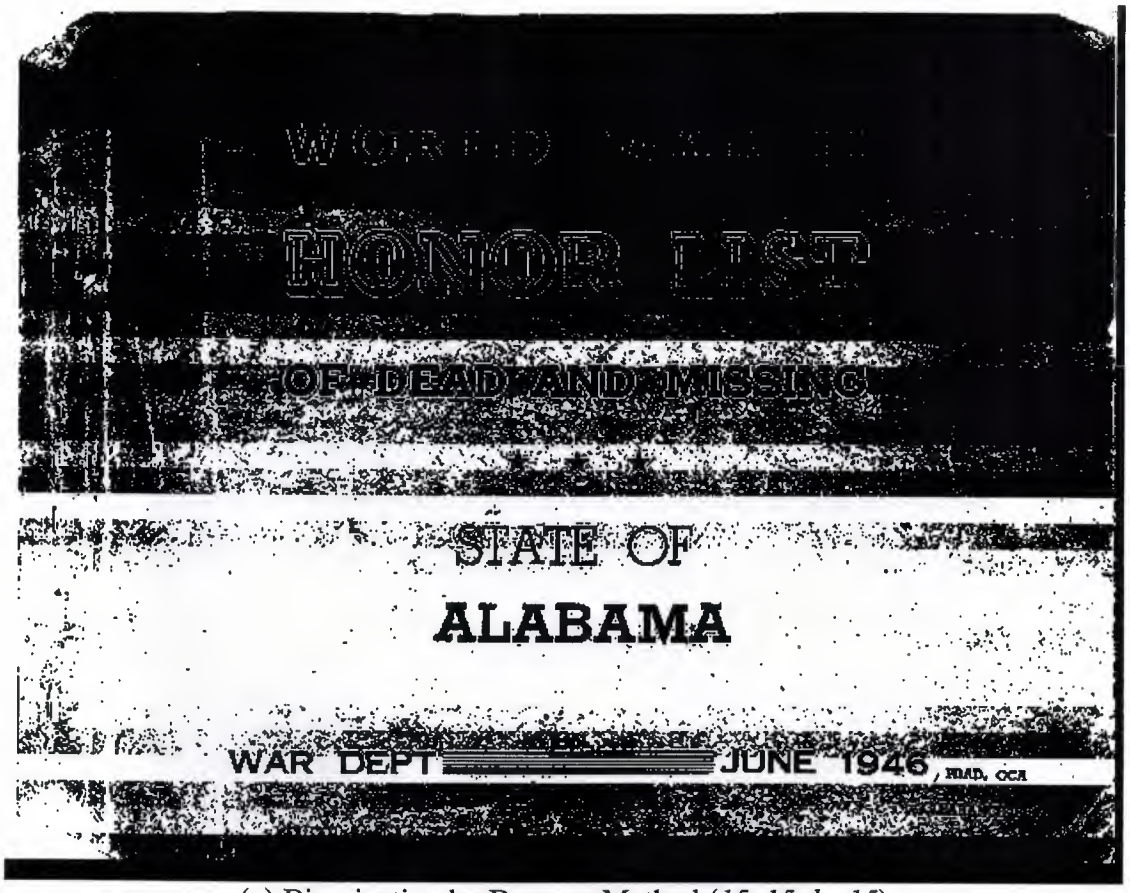
(l) Binarization by PAT Method (*15x15*)



(m) Binarization by ALT Method (*8x8*)



(n) Binarization by Water Flow Model (*5x5,w=17*)



(o) Binarization by Bernsen Method (*15x15, k=15*)

Figure A1-2 Example Result of Low Contrast Image of Test Set I


KINEMA NASIONAL

Të Hënën 29 Gusht 1938 Orën 21

Në prenën e Patronazhës së Nënës e Artave
Ns. Sh. Princeshës RUHIJE

me rastin e festimeve të Dhjet Vjetorit të
Shqipërisë se Mbretnisë

jepet

KONCERT

prej

TEFTA TASHKO
Soprano

MARIA PALUCA
Soprano

KRISTAQ KOÇO
Basitone
(Student në Konservator)

LOLA ALEKSI
Piano

TONIN GURAZIU
Piano

Drejtor Artistik Prof. Z. SOTIR KOSMO

(a) Original Bright Image


KINEMA NASIONAL

Të Hënën 29 Gusht 1938 Orën 21

Në prenën e Patronazhës së Nënës e Artave
Ns. Sh. Princeshës RUHIJE

me rastin e festimeve të Dhjet Vjetorit të
Shqipërisë se Mbretnisë

jepet

KONCERT

prej

TEFTA TASHKO
Soprano

MARIA PALUCA
Soprano

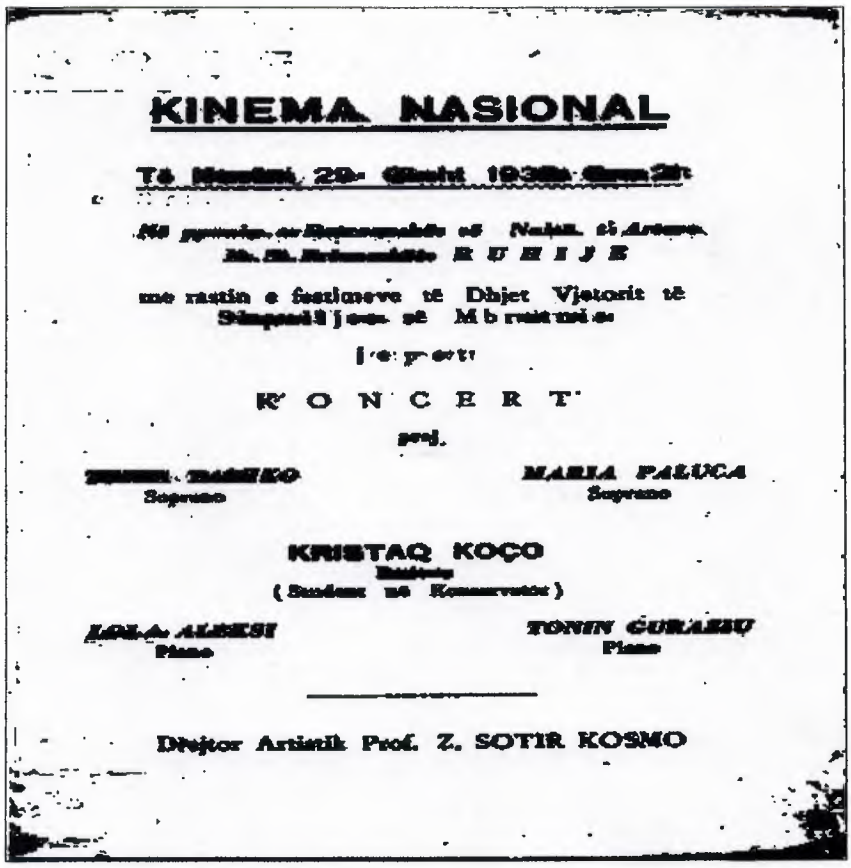
KRISTAQ KOÇO
Basitone
(Student në Konservator)

LOLA ALEKSI
Piano

TONIN GURAZIU
Piano

Drejtor Artistik Prof. Z. SOTIR KOSMO

(b) Binarization by Otsu Method



KINEMA NASIONAL

K O N C E R T

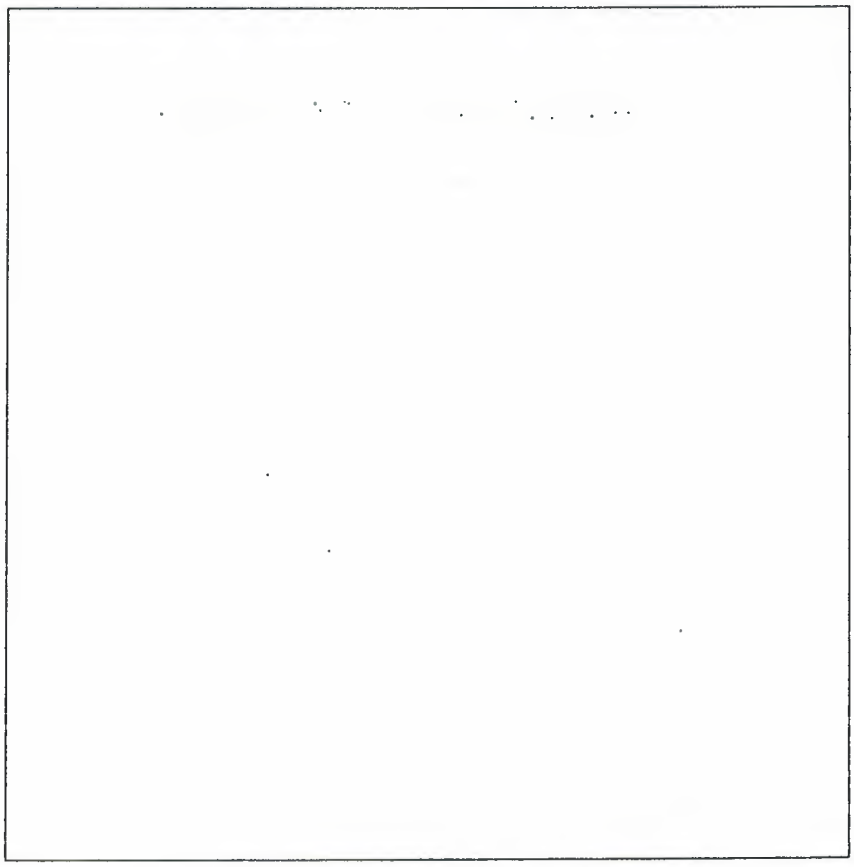
MARIA PALUCA
Soprano

KRISTAQ KOÇO

TONIN ÇURAKU
Piano

Drejtor Artistik Prof. Z. SOTIR KOSMO

(c) Binarization by Kittler and Illingworth Method



(d) Binarization by Yanni and Horne Method

(e) Binarization by Ramesh et al. Method

KINEMA NASIONAL

Të Hënë 29. Gusht 193

RUHIJE

me rastin e festimeve të Dhjet Vjetorit të

KONCERT

KRISTAQ KOÇO

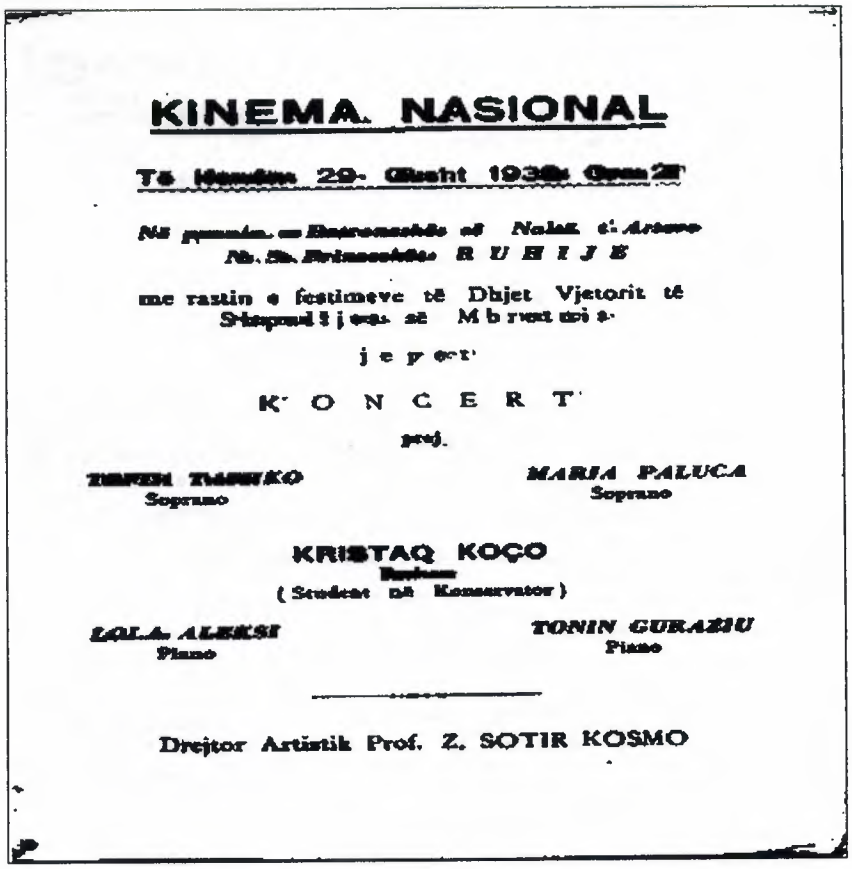
(Student në Konservator)

LOLA ALEKSI
Piano

TONIN GURAZIU
Piano

Drejtor Artistik Prof. Z. SOTIR KOSMO

(f) Binarization by Proposed MD Method



KINEMA NASIONAL

Të Hënë 29- Gusht 1938

RUHIJE

me rastin e festimeve të Dhjet Vjetorit të

KONCERT

MARIA PALUCA
Soprano

KRISTAQ KOÇO
(Student në Konservator)

LOLA ALEKSI
Piano

TONIN GURAZIU
Piano

Drejtor Artistik Prof. Z. SOTIR KOSMO

(g) Binarization by Kapur et al. Method

KINEMA NASIONAL

Të Hënë 29- Gusht 1938

RUHIJE

me rastin e festimeve të Dhjet Vjetorit të

KONCERT

MARIA PALUCA
Soprano

KRISTAQ KOÇO
(Student në Konservator)

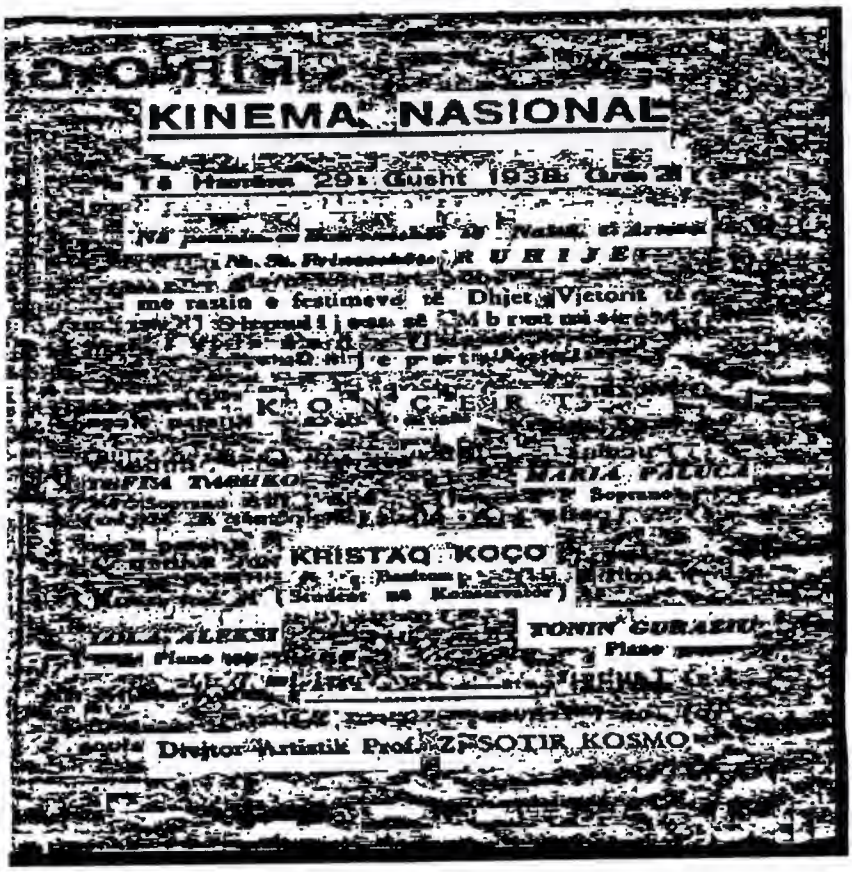
LOLA ALEKSI
Piano

TONIN GURAZIU
Piano

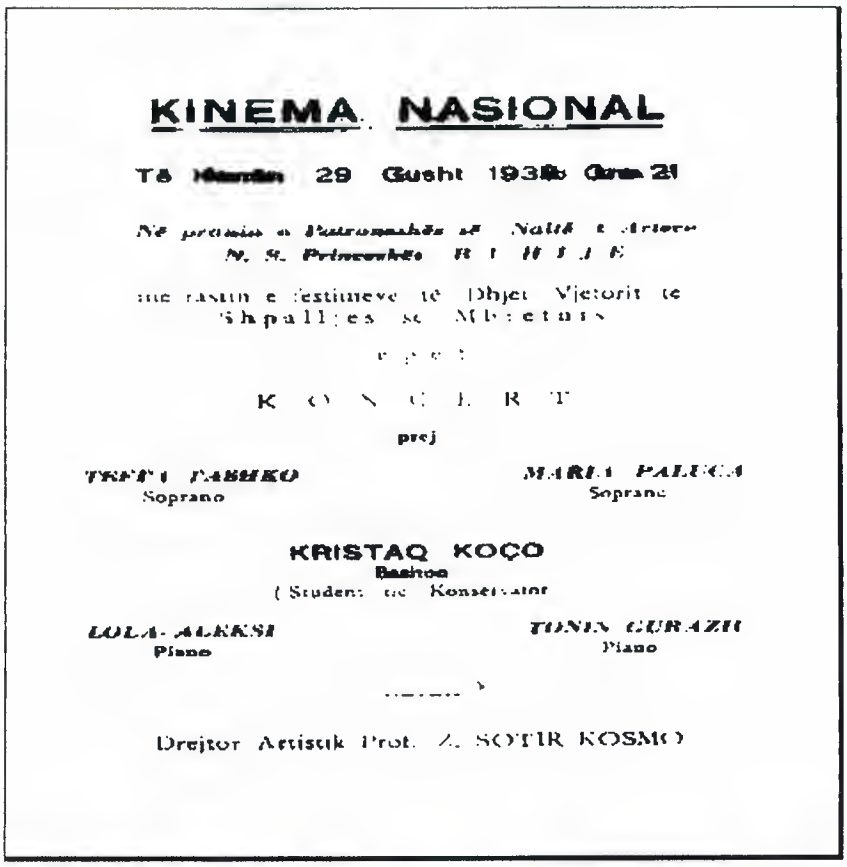
Drejtor Artistik Prof. Z. SOTIR KOSMO

(h) Binarization by Albuquerque et al. Method

(i) Binarization by Niblack Method (*15x15, k=-1*)



(j) Binarization by Sauvola et al. Method (*15x15, k=0.5, R=128*)



KINEMA NASIONAL

Të Hënën 29 Gusht 1938 Orës 21

Në prenisin e Patronazhës së Nënës e Artiste N. S. Princeshës R U H I J E

me rastin e festimeve të Dhjetë Vjetorit të Shpalljes së Mbretërisë

paraqet

K O N C E R T

prej

TEFTA FASHKO
Soprano

MARIA PALUCA
Soprano

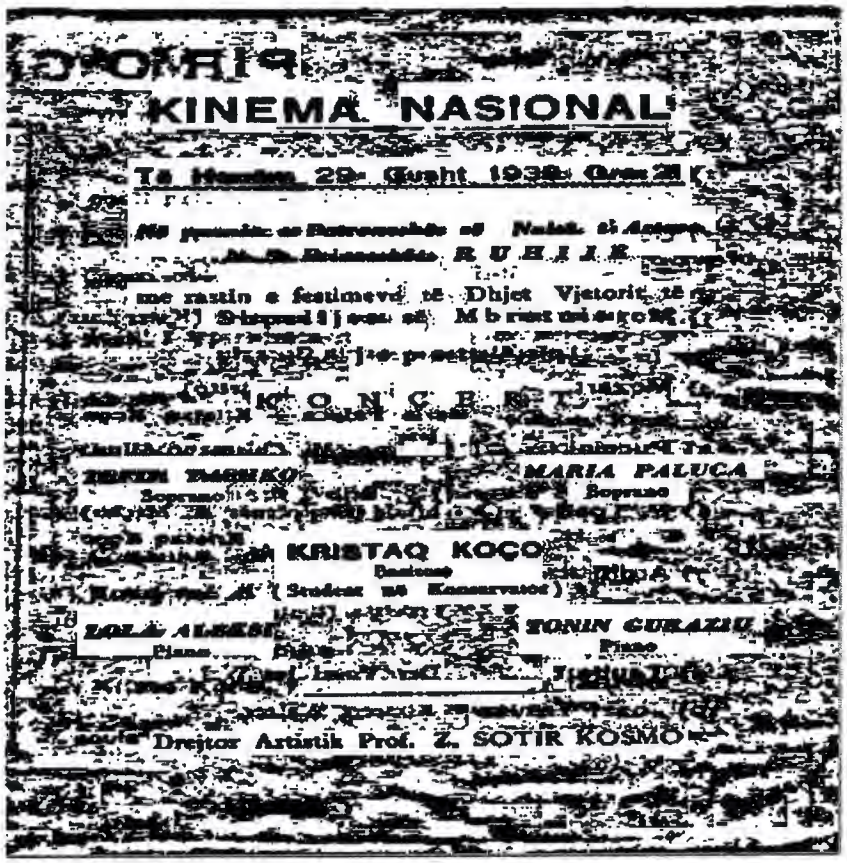
KRISTAQ KOÇO
Bashoo
(Student i Konservatorit)

LOLA ALEKSI
Piano

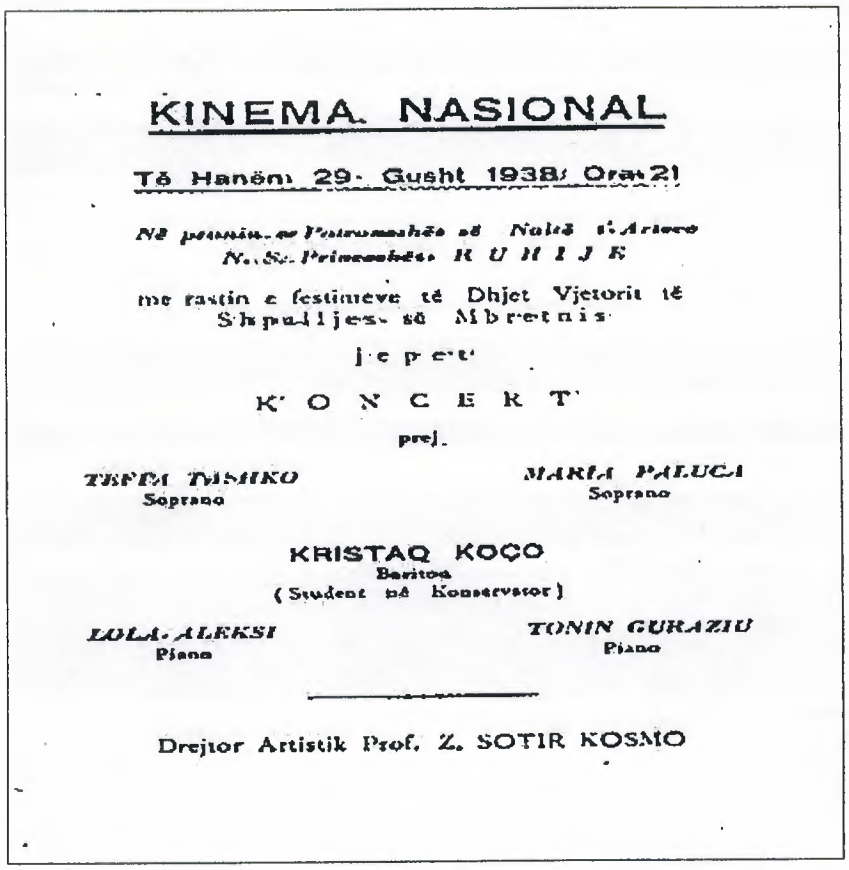
TONIN GURAZIU
Piano

Drejtor Artistik Prof. Z. SOTIR KOSMO

(k) Binarization by Mean Gradient Method (*15x15, k=-1.5, R=40*)



(l) Binarization by PAT Method (*15x15*)



KINEMA NASIONAL

Të Hanën: 29- Gusht 1938/ Ora 21

Në prenin e Pallatmashës së Nolit t'Artove
N. S. Princeshës R U H I J E

me rastin e festimeve të Dhjet Vjetorit të
Shpalljes së Mbretnis

jepet

K O N C E R T

prej.

TEFTA TASHKO
Soprano

MARIA PALUCA
Soprano

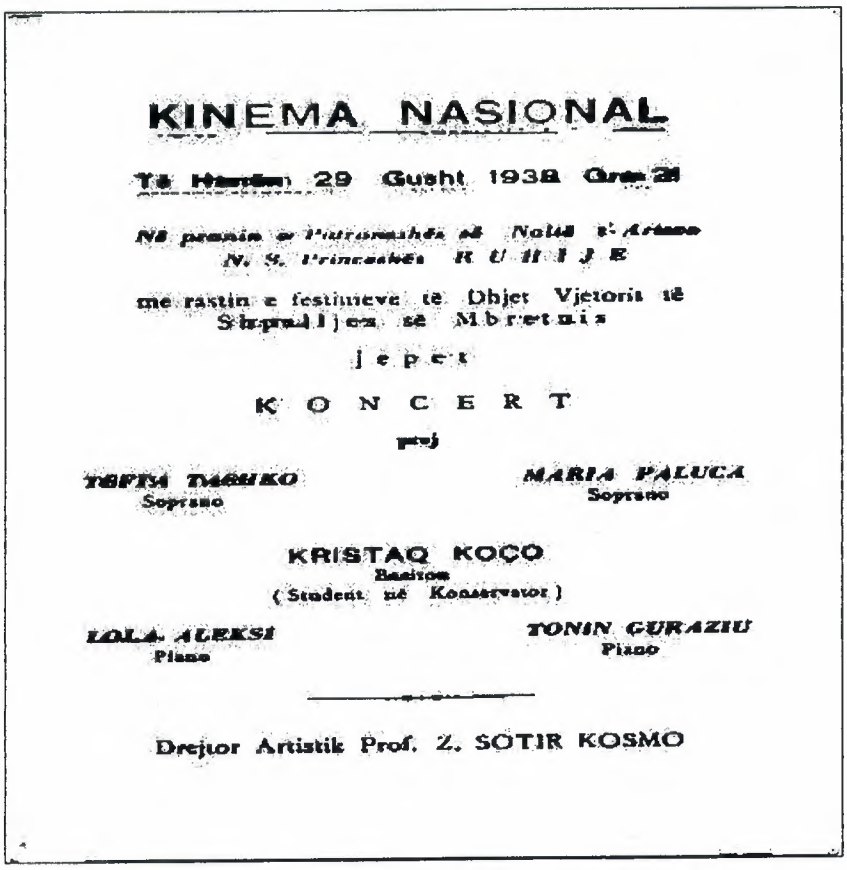
KRISTAQ KOÇO
Baritoni
(Student në Konservator)

LOLA ALEKSI
Piano

TONIN GURAZIU
Piano

Drejtor Artistik Prof. Z. SOTIR KOSMO

(m) Binarization by ALT Method (*8x8*)



KINEMA NASIONAL

Të Hanën 29 Gusht 1938 Ora 21

Në prenin e Pallatmashës së Nolit t'Artove
N. S. Princeshës R U H I J E

me rastin e festimeve të Dhjet Vjetorit të
Shpalljes së Mbretnis

jepet

K O N C E R T

prej

TEFTA TASHKO
Soprano

MARIA PALUCA
Soprano

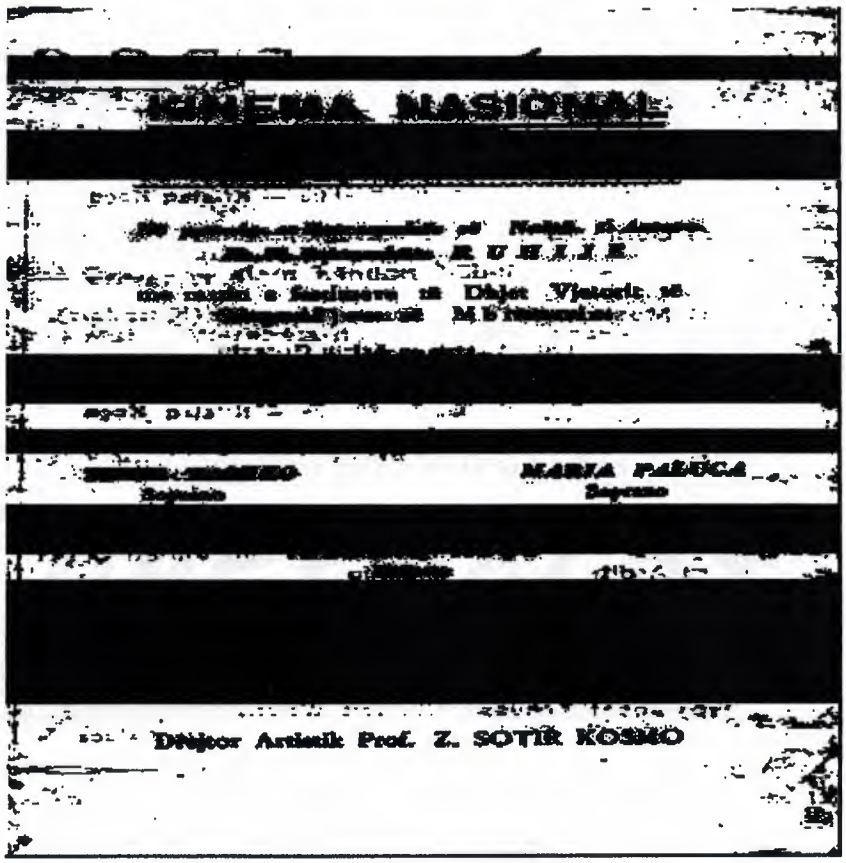
KRISTAQ KOÇO
Baritoni
(Student në Konservator)

LOLA ALEKSI
Piano

TONIN GURAZIU
Piano

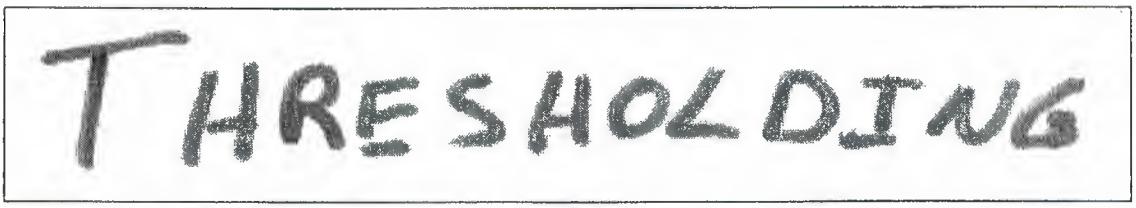
Drejtor Artistik Prof. Z. SOTIR KOSMO

(n) Binarization by Water Flow Model (*5x5,w=17*)

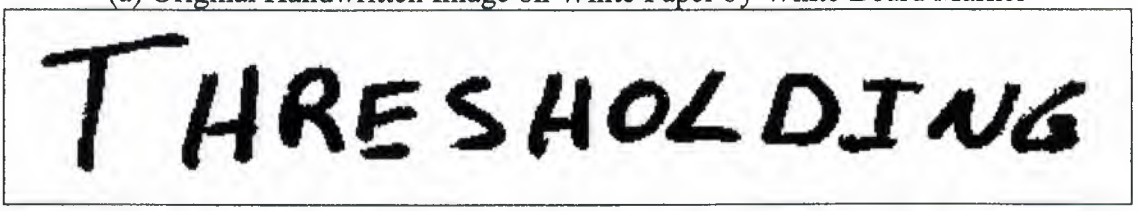


(o) Binarization by Bernsen Method (*15x15, k=15*)

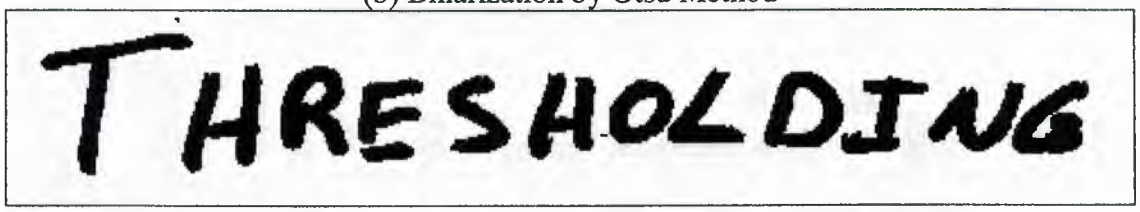
Figure A1-3 Example Result of Bright Image of Test Set I



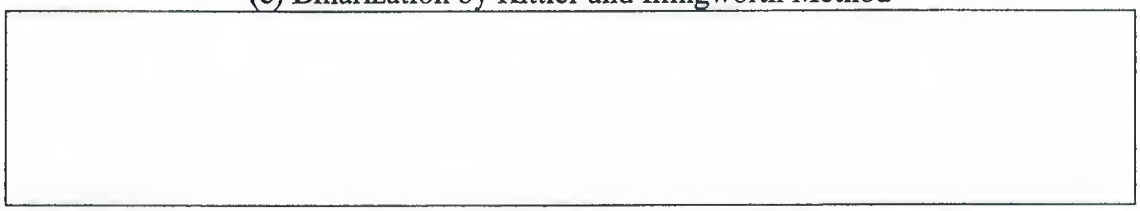
(a) Original Handwritten Image on White Paper by White Board Marker



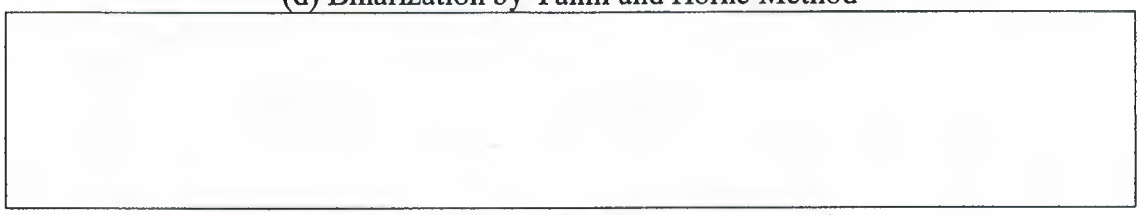
(b) Binarization by Otsu Method



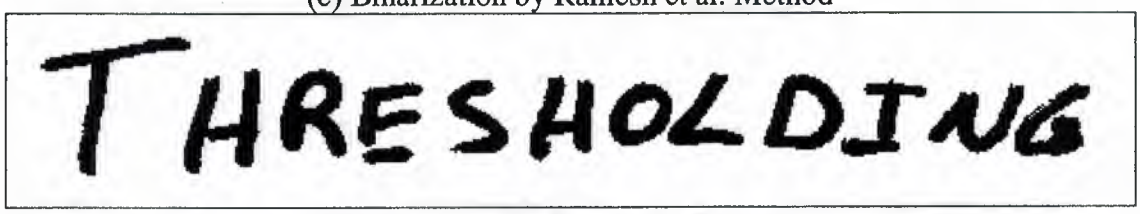
(c) Binarization by Kittler and Illingworth Method



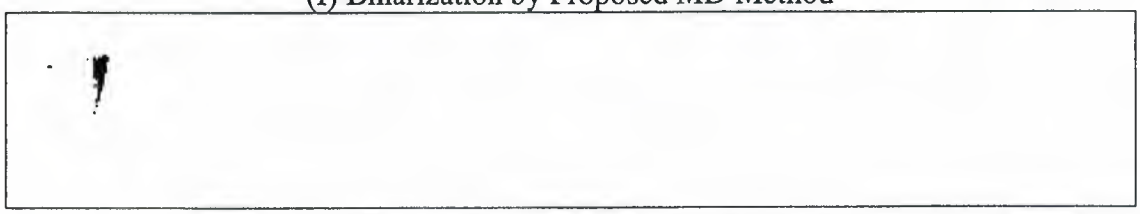
(d) Binarization by Yanni and Horne Method



(e) Binarization by Ramesh et al. Method

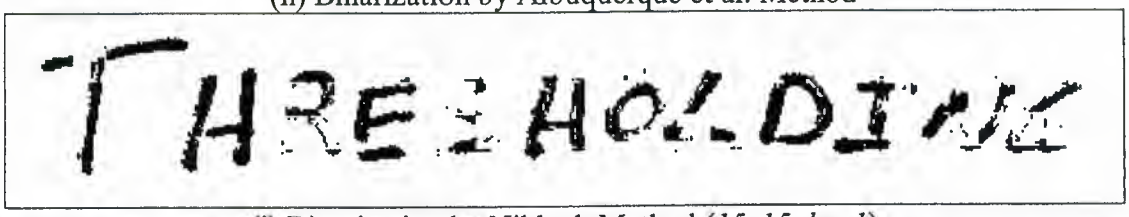


(f) Binarization by Proposed MD Method

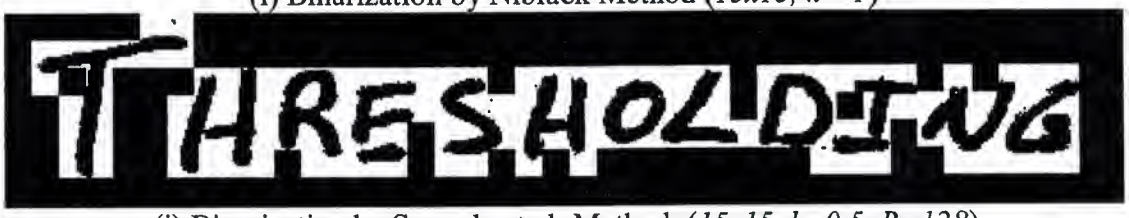


(g) Binarization by Kapur et al. Method

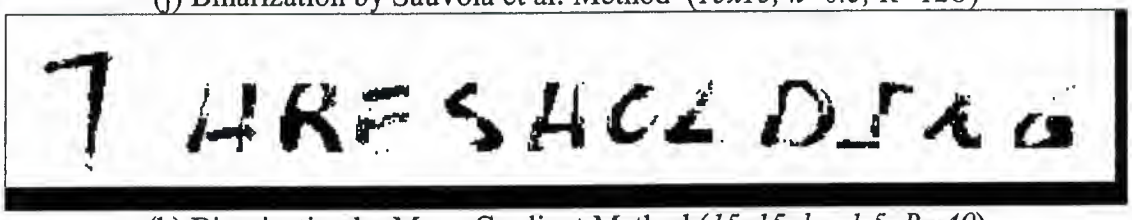
(h) Binarization by Albuquerque et al. Method



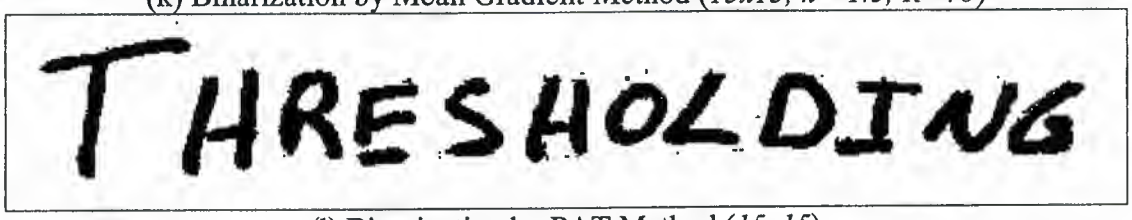
(i) Binarization by Niblack Method (*15x15, k=-1*)



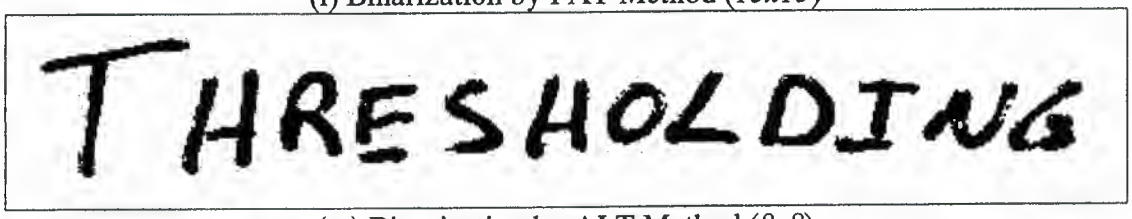
(j) Binarization by Sauvola et al. Method (*15x15, k=0.5, R=128*)



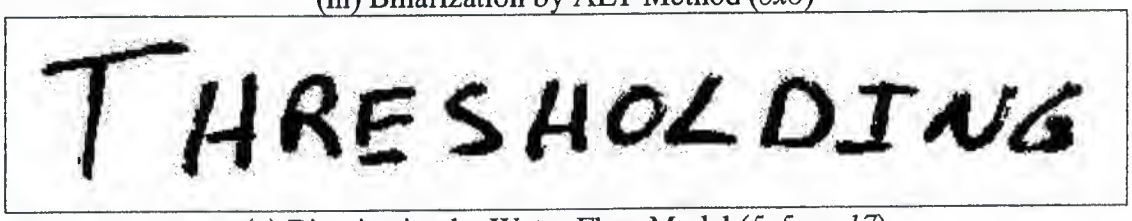
(k) Binarization by Mean Gradient Method (*15x15, k=-1.5, R=40*)



(l) Binarization by PAT Method (*15x15*)



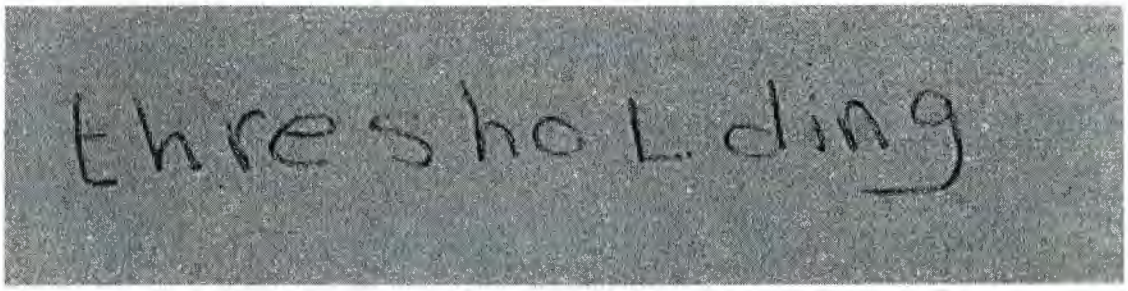
(m) Binarization by ALT Method (*8x8*)



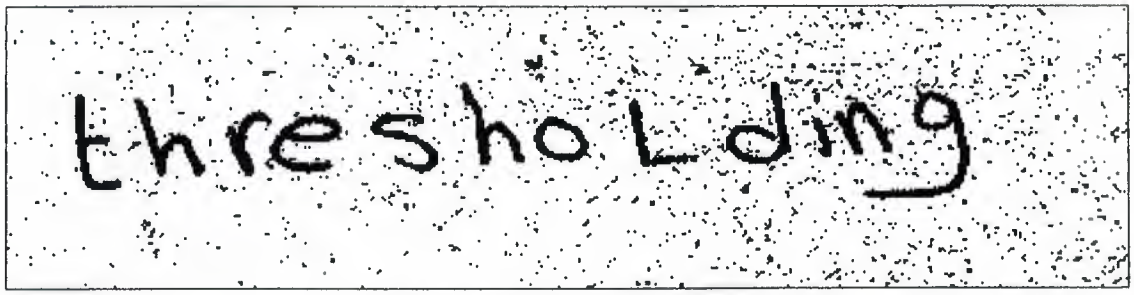
(n) Binarization by Water Flow Model (*5x5,w=17*)

(o) Binarization by Bernsen Method (*15x15, k=15*)

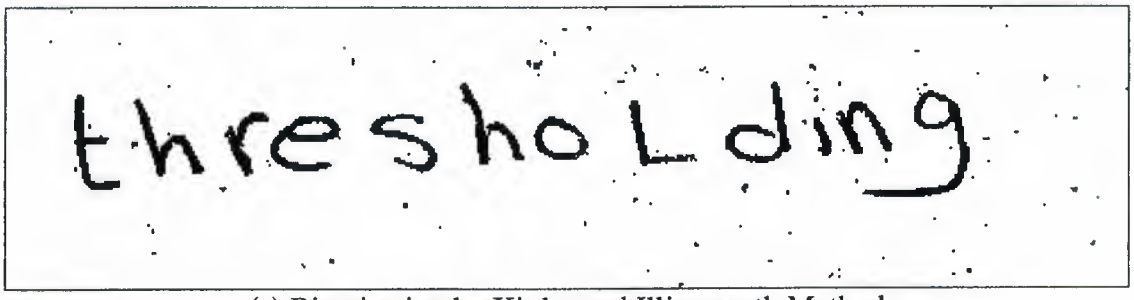
Figure A1-4 Example Result of White Board Marker on White Paper in Image Set III



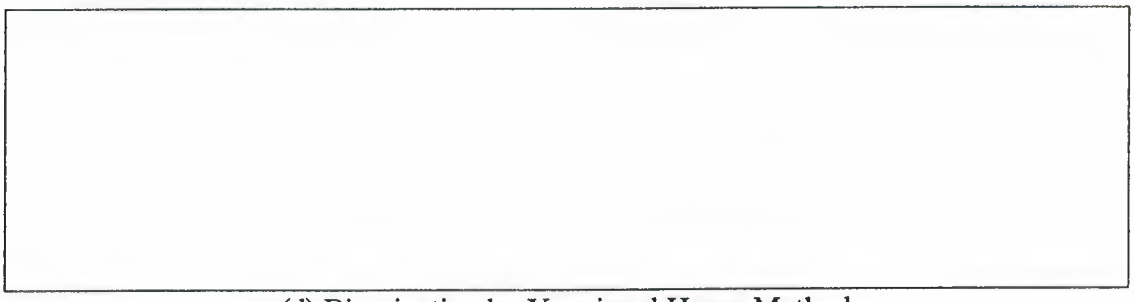
(a) Original Handwritten Image on Yellow Envelope Paper by Pen



(b) Binarization by Otsu Method

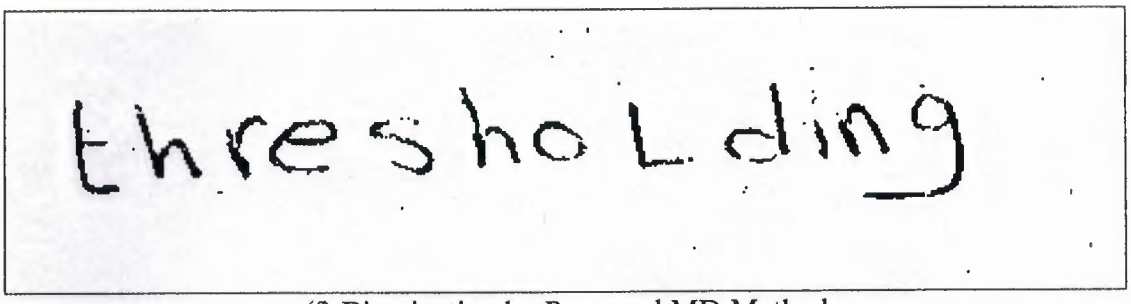


(c) Binarization by Kittler and Illingworth Method

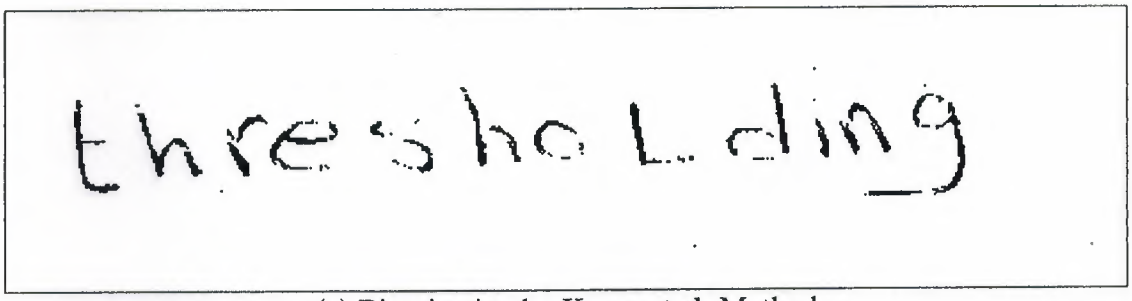


(d) Binarization by Yanni and Horne Method

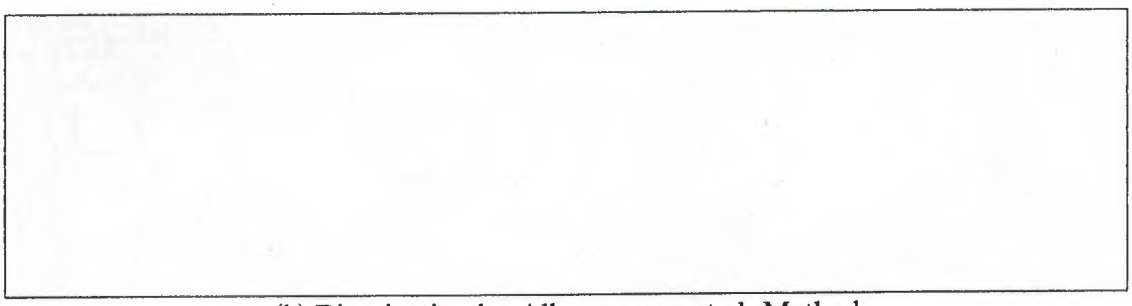
(e) Binarization by Ramesh et al. Method



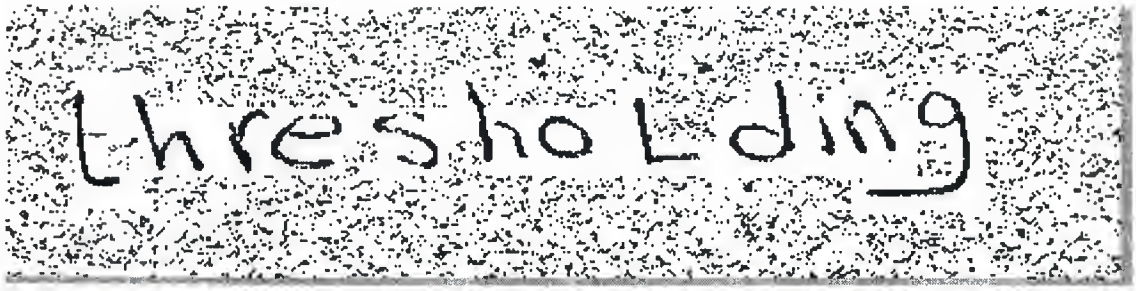
(f) Binarization by Proposed MD Method



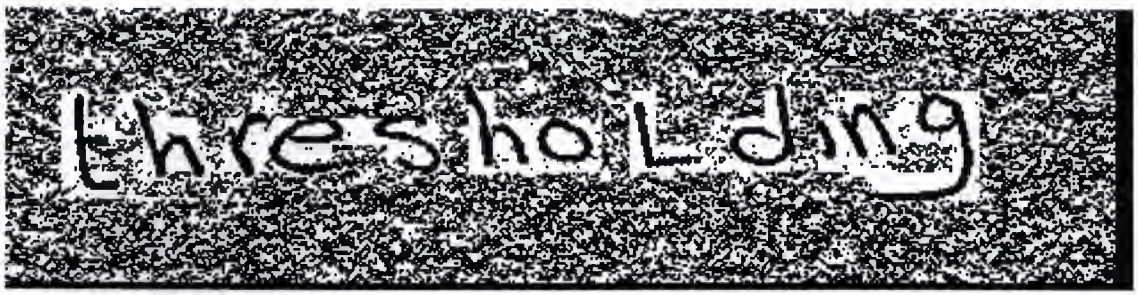
(g) Binarization by Kapur et al. Method



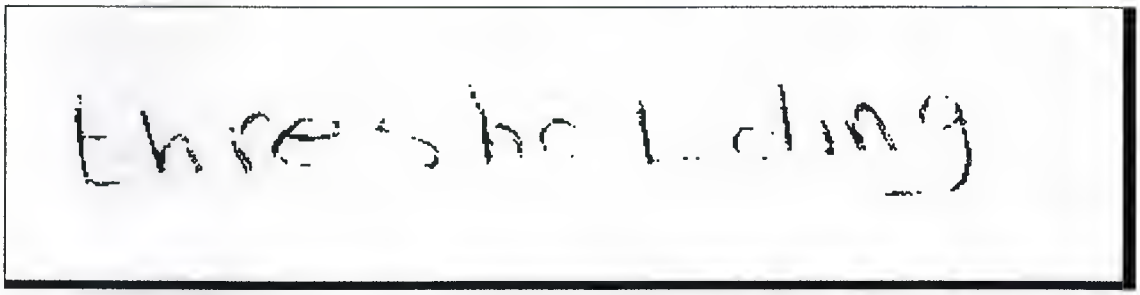
(h) Binarization by Albuquerque et al. Method



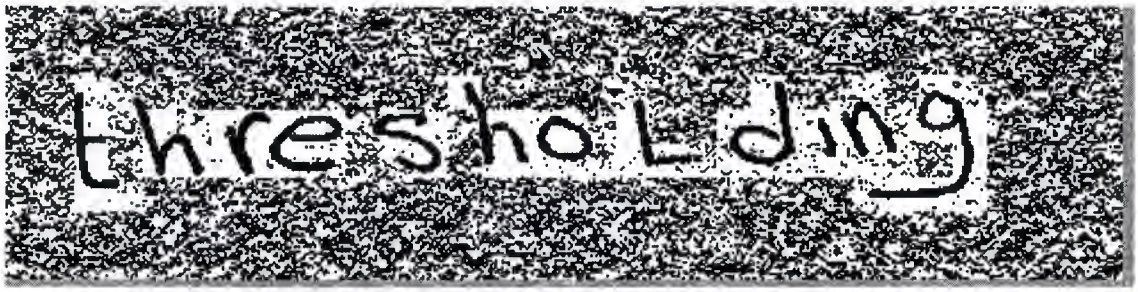
(i) Binarization by Niblack Method (*15x15, k=-1*)



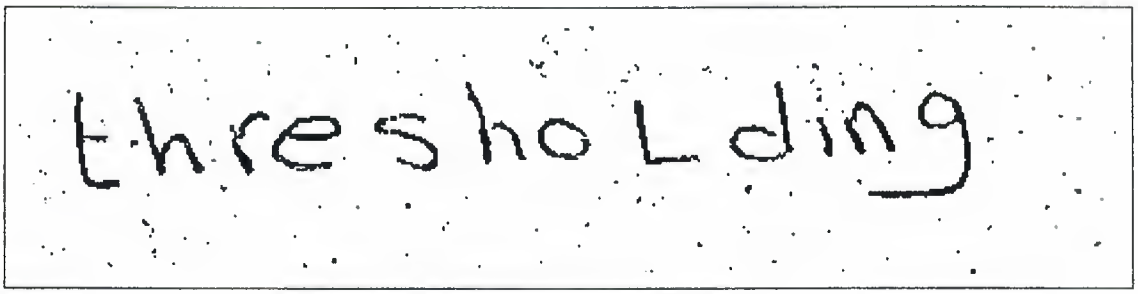
(j) Binarization by Sauvola et al. Method (*15x15, k=0.5, R=128*)



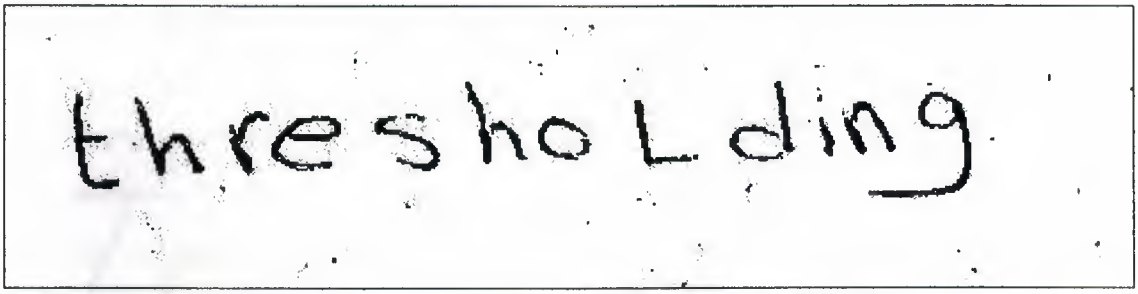
(k) Binarization by Mean Gradient Method (*15x15, k=-1.5, R=40*)



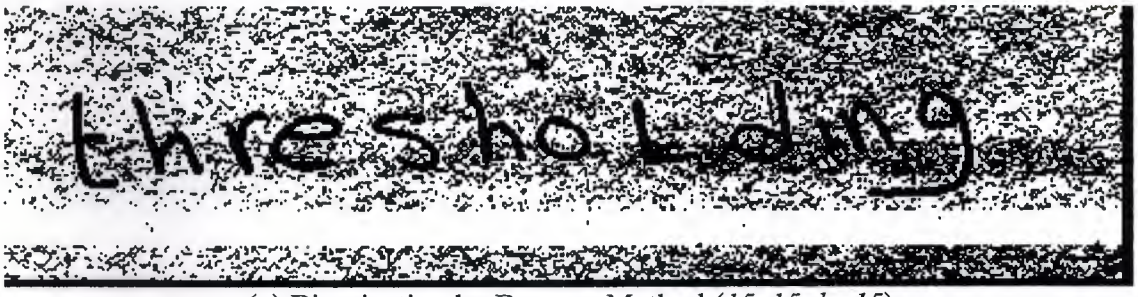
(l) Binarization by PAT Method (*15x15*)



(m) Binarization by ALT Method (*8x8*)

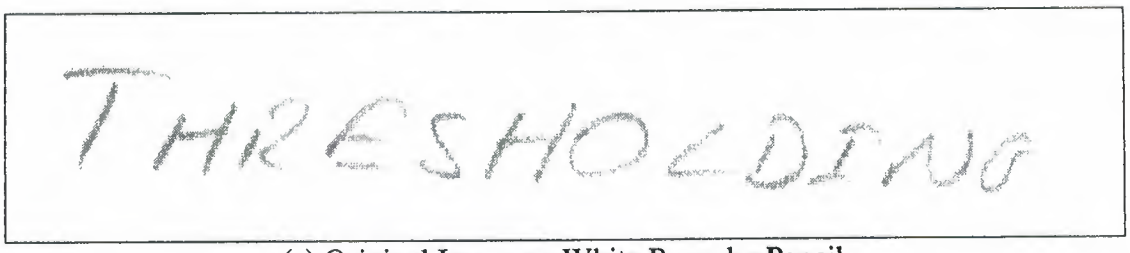


(n) Binarization by Water Flow Model (*5x5,w=17*)

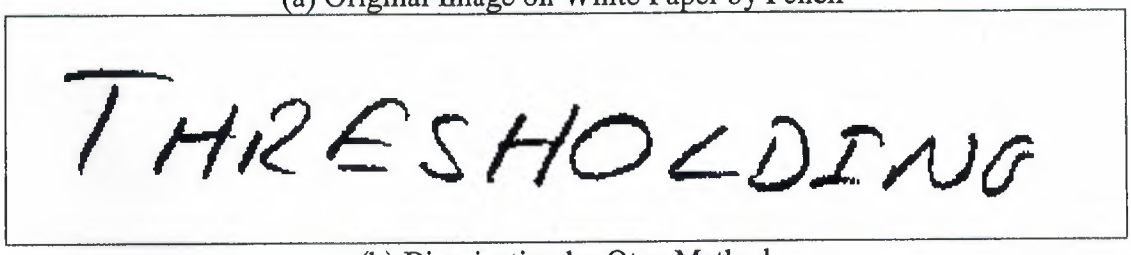


(o) Binarization by Bernsen Method (*15x15, k=15*)

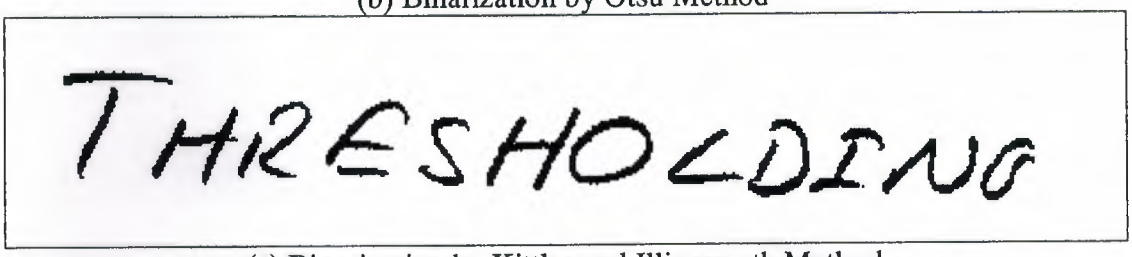
Figure A1-5 Example Result of Pen on Yellow Envelope Paper in Image Set III



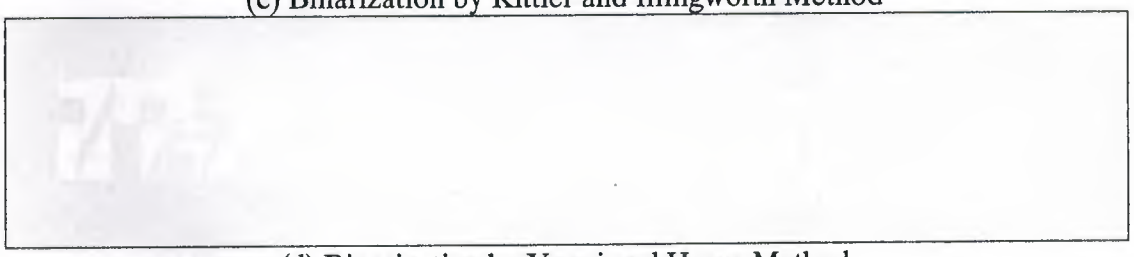
(a) Original Image on White Paper by Pencil



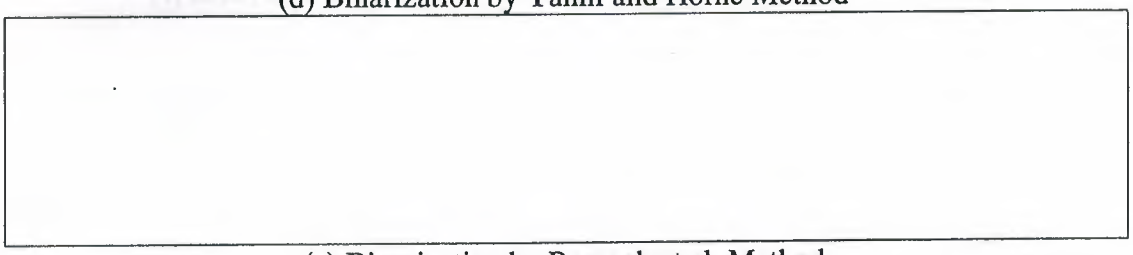
(b) Binarization by Otsu Method



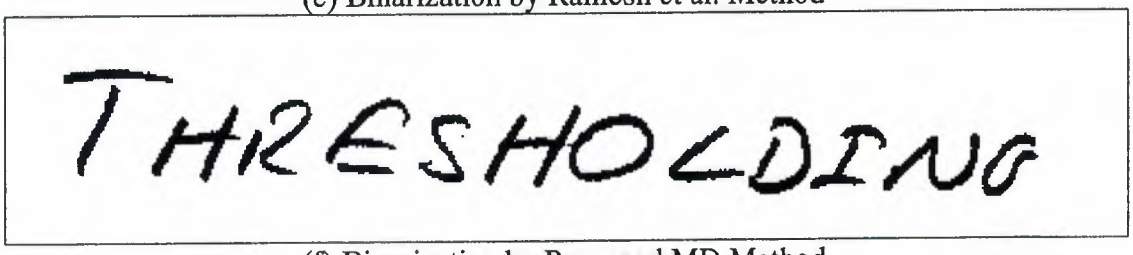
(c) Binarization by Kittler and Illingworth Method



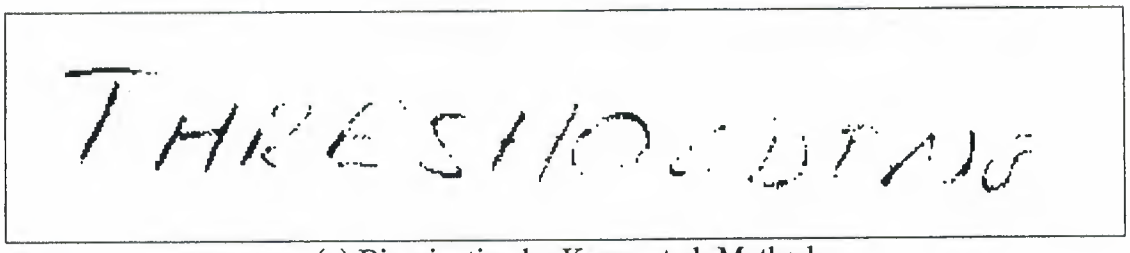
(d) Binarization by Yanni and Horne Method



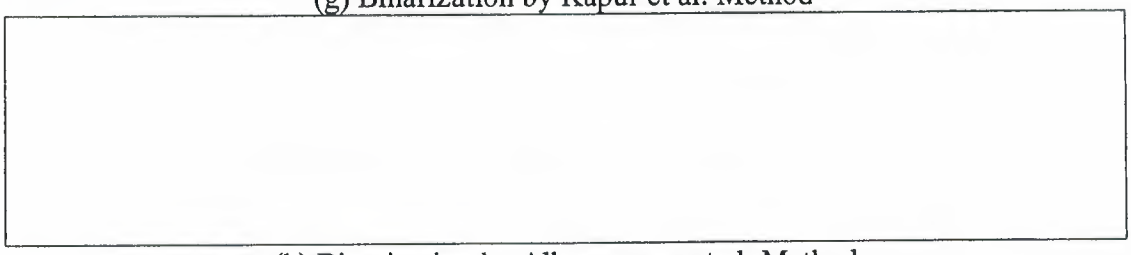
(e) Binarization by Ramesh et al. Method



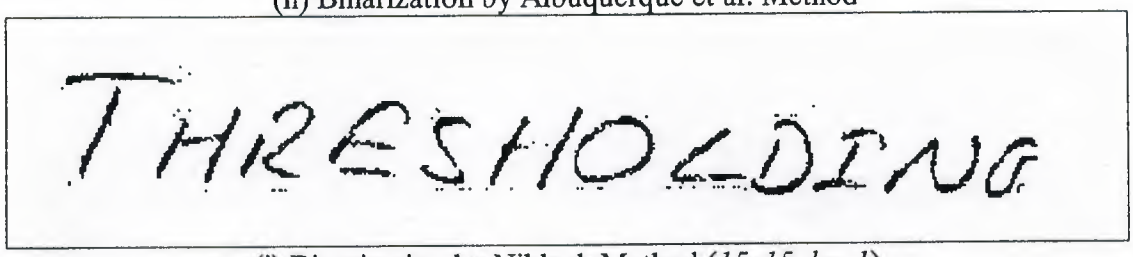
(f) Binarization by Proposed MD Method



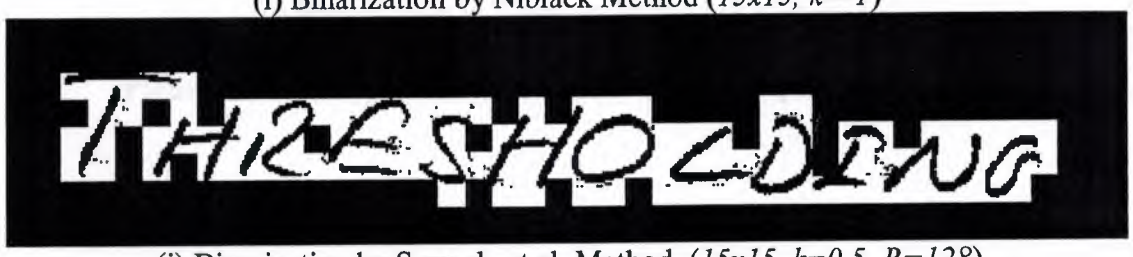
(g) Binarization by Kapur et al. Method



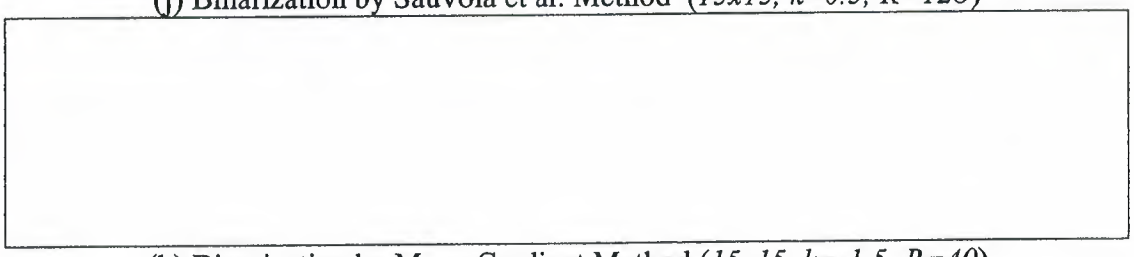
(h) Binarization by Albuquerque et al. Method



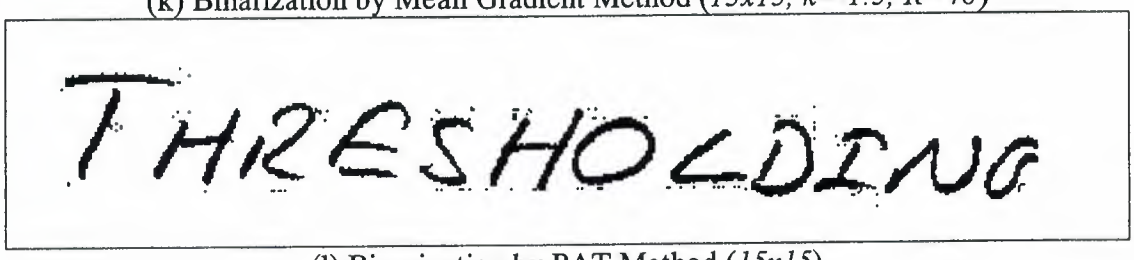
(i) Binarization by Niblack Method (*15x15, k=-1*)



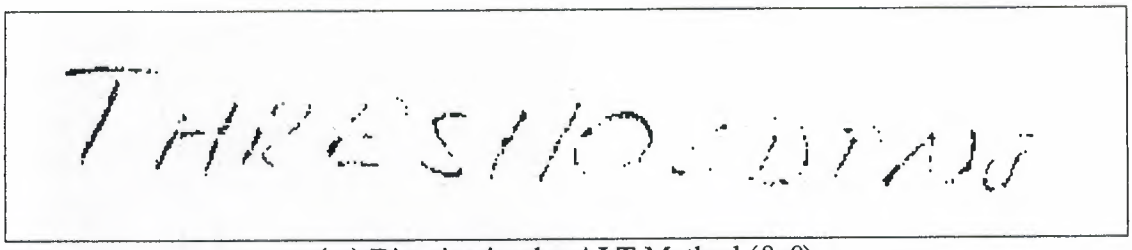
(j) Binarization by Sauvola et al. Method (*15x15, k=0.5, R=128*)



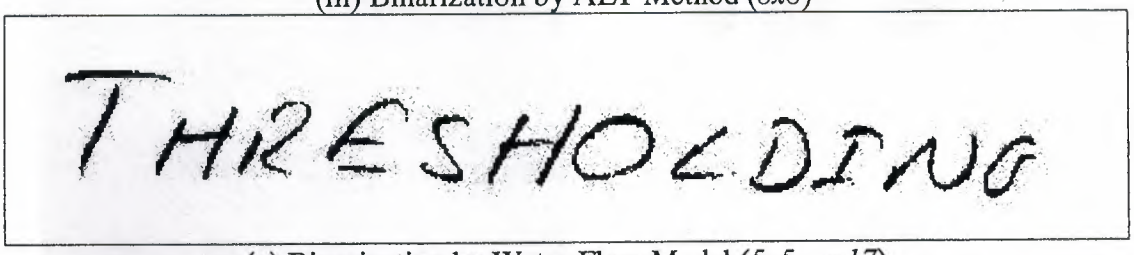
(k) Binarization by Mean Gradient Method (*15x15, k=-1.5, R=40*)



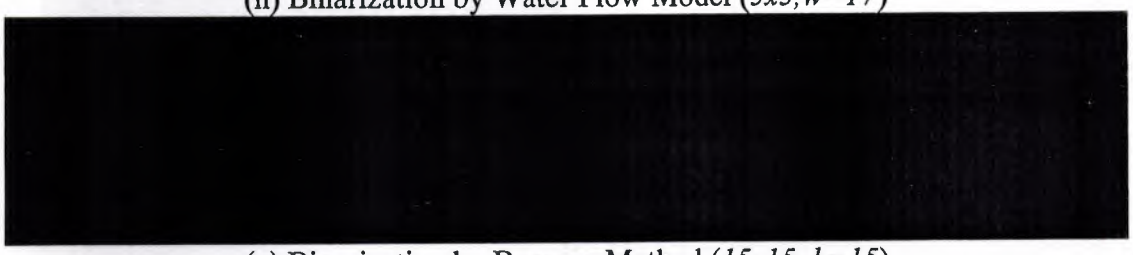
(l) Binarization by PAT Method (*15x15*)



(m) Binarization by ALT Method (*8x8*)

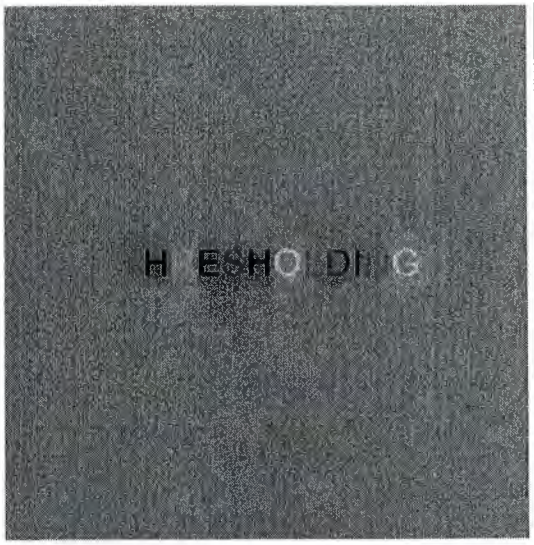


(n) Binarization by Water Flow Model (*5x5,w=17*)

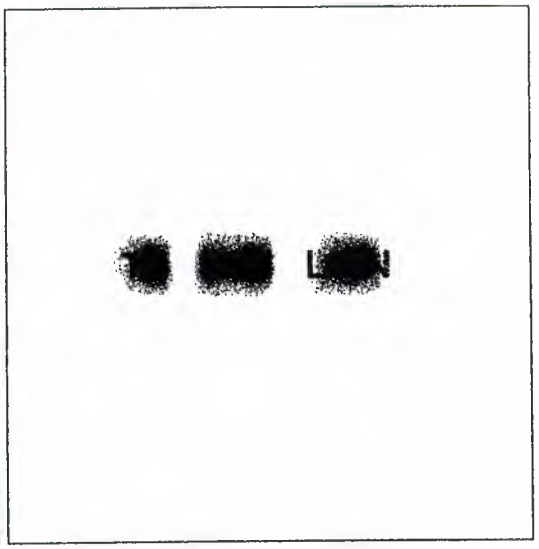


(o) Binarization by Bernsen Method (*15x15, k=15*)

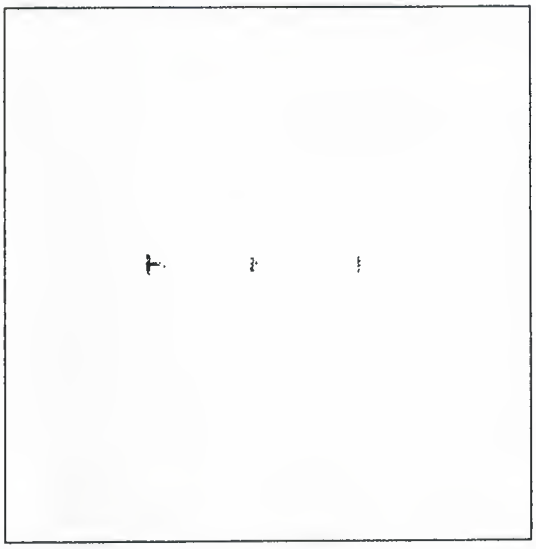
Figure A1-6 Example Result of Pencil on White Paper in Image Set III



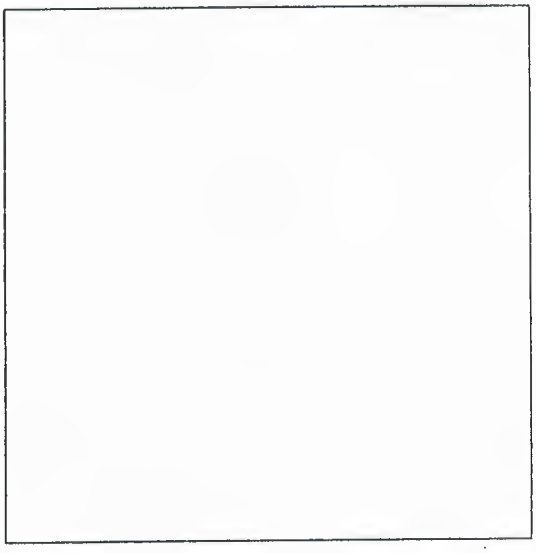
(a) Original Created Word



(b) Binarization by Otsu Method

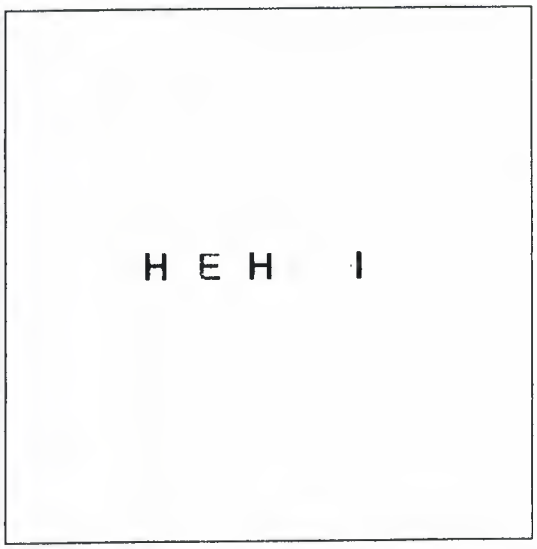


(c) Binarization by Kitt. and Illing. Method



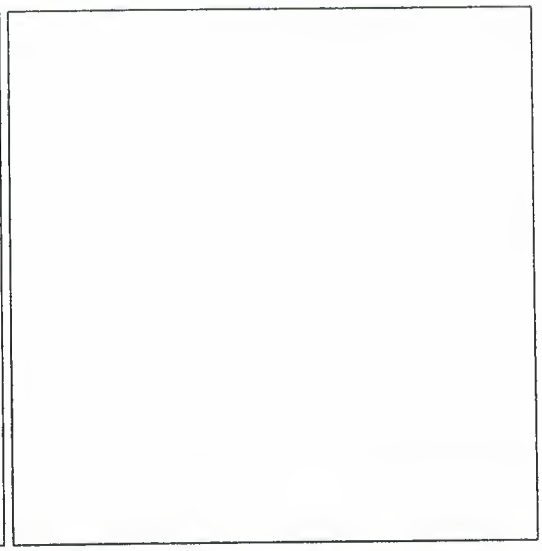
(d) Bin. by Yanni and Horne Method

(e) Binarization by Ramesh et al. Method



(f) Binarization by Proposed MD Method

(g) Binarization by Kapur et al. Method

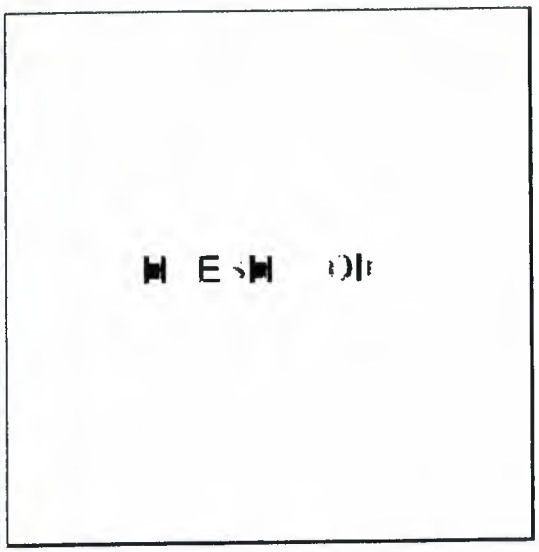


(h) Binarization by Albuquerque et al. Method

(i) Bin. by Niblack Method

(j) Binarization by Sauvola et al. Method

(k) Bin. by Mean Gradient Method



(l) Binarization by PAT Method

(m) Binarization by ALT Method

(n) Binarization by Water Flow Model

(o) Binarization by Bernsen Method

Figure A1-7 Example Result of Artificially Created Text in Image Set II