

## **INFORMATION TO USERS**

**This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.**

**The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.**

**In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.**

**Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.**

**Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.**

**ProQuest Information and Learning  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
800-521-0600**

**UMI<sup>®</sup>**





**Université d'Ottawa • University of Ottawa**



**Restoration of Documents**  
**With**  
**Show-through Distortion**

*by*

**Anita Pathak**

A thesis presented to the University of Ottawa in partial  
fulfillment of the requirement for the degree of  
Masters of Science  
in Systems Science

(Faculty of Administration)  
University of Ottawa, Ottawa, Canada



**National Library  
of Canada**

**Acquisitions and  
Bibliographic Services**

**395 Wellington Street  
Ottawa ON K1A 0N4  
Canada**

**Bibliothèque nationale  
du Canada**

**Acquisitions et  
services bibliographiques**

**395, rue Wellington  
Ottawa ON K1A 0N4  
Canada**

*Your file Votre référence*

*Our file Notre référence*

**The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.**

**The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.**

**L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.**

**L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.**

0-612-66099-0

**Canada**

## Summary

This thesis addresses the problem of show-through in old manuscripts and documents. After removing the show-through these documents can be stored in the digital libraries and will be accessible by everybody via Internet. The solution to this problem utilizes the image restoration technique. To apply this, the documents are scanned from each side and stored in a digital form. The two sides of document are called *recto* and *verso* images. To approach an apposite solution a basic model of the degradation is defined for these two images and then this model is used to eliminate the show-through. The technique for the enhancement is divided into two parts – first, achieving the most precise registration technique and secondly, to compute a method of restoration using the model of degradation. Four different techniques for registration are compared to select the most suitable method for this problem. This comparison brings forth the affine transformation approach to register the image. An approach for the restoration of the registered images is defined to eliminate the show-through from the document and save it in a digital library.

## **Acknowledgements**

I would like to sincerely thank my thesis supervisor, Dr. Eric Dubois, for his guidance and encouragement throughout this project.

I would like to thank the School of Information Technology and Engineering, University of Ottawa, for providing the research facilities in the Vision, Image, Video and Audio (VIVA) lab.

I also would like to thank my fellow students and staff members for assistance in various endeavors.

# Contents

Summary .....	i
Acknowledgements.....	ii
Contents .....	iii
List of Figures .....	v
List of Tables .....	vi
Chapter 1 Introduction.....	1
Chapter 2 Theoretical Background.....	6
2.1 Image .....	6
2.2 Digital Image .....	7
2.3 Gray Scale.....	8
2.4 Image Acquisition.....	9
2.4.1 Scanner.....	9
2.5 Image Processing.....	11
2.6 Image Enhancement .....	11
2.6.1 Point Operation .....	13
2.6.2 Spatial Operation .....	19
2.6.3 Transform Operation.....	19
2.7 Geometric Distortion .....	19
2.7.1 Translation .....	20
2.7.2 Rotation.....	20
2.7.3 Scaling.....	20
2.7.4 Shear .....	21
2.8 Image Registration.....	21
2.8.1 Pixel Transformation .....	24
2.8.2 Affine Transformation .....	25
2.8.3 Discrete Fourier Transform.....	26
2.9 Registration.....	28
2.9.1 Cross Correlation Method.....	28
2.9.2 Least Squares Method.....	29
2.10 Restoration.....	29
2.10.1 Spatial Image Restoration Techniques.....	30
2.10.2 Point and Spectral Restoration Techniques .....	31

2.10.3 Geometric Transformation.....	31
Chapter 3 Show-Through Correction Methods .....	32
3.1 Image Enhancement.....	32
3.1 Adobe® Photo Deluxe™ Business Edition.....	33
3.1.2 Applying a Threshold.....	34
3.2 Model of Degradation.....	35
3.2.1 Registration .....	38
3.2.2 Restoration .....	40
3.3 Image Registration.....	40
3.3.1 Pixel Translation approach.....	42
3.3.2 Pixel transformation ( translation and rotation) .....	43
3.3.3 Affine Transformation.....	44
3.3.4 Affine Transformation with Six Parameters .....	45
3.4 Restoration .....	46
3.4.1 Comparing Pixel Intensity.....	46
3.4.2 Selecting a Threshold .....	47
3.4.3 Neighborhood Processing .....	48
3.4.4 Comparing the Ratio of Pixel Intensity.....	48
Chapter 4 Experimental Results .....	50
4.1 Image Enhancement.....	53
4.1.1 Adobe Photo deluxe .....	55
4.1.2 Applying a Threshold .....	57
4.2 Registration.....	62
4.2.1 Pixel Translation Approach.....	62
4.2.2 Pixel Transformation with Rotation.....	64
4.2.3 Affine Transformation.....	66
4.2.4 Affine Transformation with six Parameters .....	67
4.3 Restoration .....	69
4.3.1 Comparing Pixel Intensity.....	69
4.3.2 Neighborhood approach .....	71
4.3.3 Comparing the ratio of Pixel Intensity .....	73
4.4 Discussion.....	76
Chapter 5 Conclusion .....	78
References.....	80

## List of Figures

Figure 1. 1: Sample of an old degraded document with high show-through.....	2
Figure 1. 2: The histogram of the degraded image with high show-through.....	3
Figure 1. 3: The histogram of the image with out show-through .....	4
Figure 2. 1: A continuous image divided into a discrete space .....	7
Figure 2. 2: A scanning mechanism (taken from [4] ).....	9
Figure 2. 3: Contrast stretching.....	13
Figure 2. 4: An example of binary thresholding.....	14
Figure 2. 5: A clear text document and its histogram. ....	17
Figure 2. 6: A degraded text document and its histogram. ....	18
Figure 2. 7: An image degradation and restoration model.....	30
Figure 3. 1: The text document image restored using the Adobe® Photo Deluxe™ Business Edition.....	33
Figure 3. 2: Image enhancement using threshold .....	35
Figure 3. 3: A basic model of <i>recto</i> and <i>verso</i> image .....	36
Figure 3. 4: Selecting a region of interest from the <i>recto</i> and flipped <i>verso</i> images .....	41
Figure 3. 5: A histogram of the degraded <i>recto</i> image .....	47
Figure 4. 1: The ideal image of the sample 1.....	51
Figure 4. 2: The ideal image of sample 2.....	52
Figure 4. 3: The histogram of the ideal image .....	53
Figure 4. 4: Sample 1 - The <i>recto</i> image with show-through .....	54
Figure 4. 5: Sample 2 – The degraded <i>recto</i> image .....	55
Figure 4. 6: The restored image using Adobe® Photo Deluxe™ Business Edition .....	56
Figure 4. 7: Another sample of the restored image using Adobe® Photo .....	57
Figure 4. 8: The histogram of the degraded <i>recto</i> image.....	58
Figure 4. 9: The histogram of restored image using threshold approach.....	59
Figure 4. 10: The restored image using threshold value 0.6.....	60
Figure 4. 11: The restored image using threshold value 0.45 .....	61
Figure 4. 12: The registered <i>recto</i> and flipped <i>verso</i> images before transformation. ....	63
Figure 4. 13: The registered image using pixel transformation approach. ....	64
Figure 4. 14: The registered image using pixel translation with rotation. ....	65
Figure 4. 15: The registered image using affine transformation.....	67
Figure 4. 16: The registered image using affine transformation approach. ....	68

Figure 4. 17: The restored image obtained by comparing pixel intensity .....	70
Figure 4. 18: The histogram of the restored image .....	71
Figure 4. 19: The restored <i>recto</i> image using neighborhood approach. ....	72
Figure 4. 20: The histogram of the restored image using Neighborhood approach .....	73
Figure 4. 21: The restored sample1 using option I .....	74
Figure 4. 22: The sample 1 restored using the comparison of intensity ratio .....	75
Figure 4. 23: The histogram of the restored image .....	76

## List of Tables

Table 2. 1: Image enhancement operations.....	12
Table 4. 1: Table for comparing the results of the restored image using three approaches.....	76

## **Chapter 1**

## **Introduction**

A collection of rare books and manuscripts, including early copies of works by Aristotle, Dante, Euclid, Homer, and Virgil, are available in the Vatican Library in Rome. And a beautiful work on medicine, religion and science written by the scholars, called Vedas, is preserved in India. The histories of civilizations are stored in libraries and museums. Around the world, there is a treasure of excellent literature, which cannot be accessed by most of the people in the world because of time and travel cost. By improving access to scientific, educational and historical documents and information, digital libraries can create powerful opportunities for revamping education, improving knowledge and providing historical background.

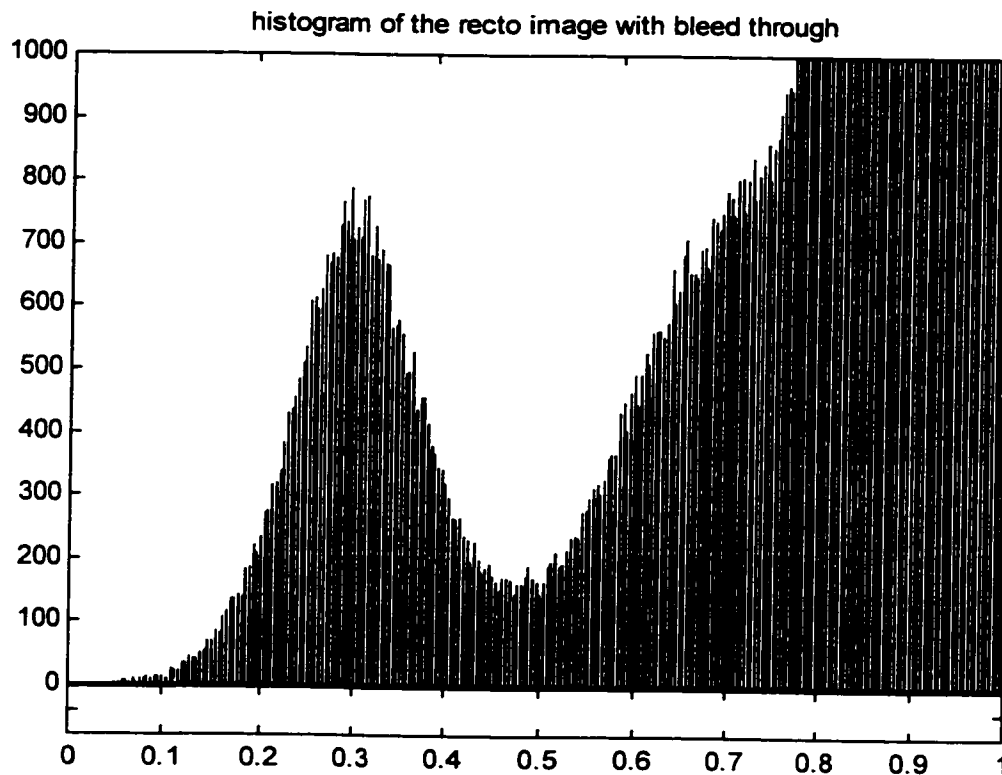
In the past few years the technology advances in two areas, computers and communication networks have helped in creating the Internet today. This advancement has provided the opportunity to make this literature available to the people from all around the world. By developing image databases of these manuscripts and documents, and making them available on the Internet it will be easier for the people to access them. Another significant advantage of image databases is we can save the documents from further degradation by preserving them and at the same time make them accessible in perpetuity in the digital library.

So far we are able to access images of historical artwork such as old paintings, sculptures, prints etc. through the digital libraries. Creating a large volume of text data would be very challenging, as it would require a lot of manpower if typed in manually. Hence we need to find an efficient approach to store the documents in digital library and make them available. One approach is to convert the document into a digital image and save the image form in the digital library, that can save a lot of time and work. One major

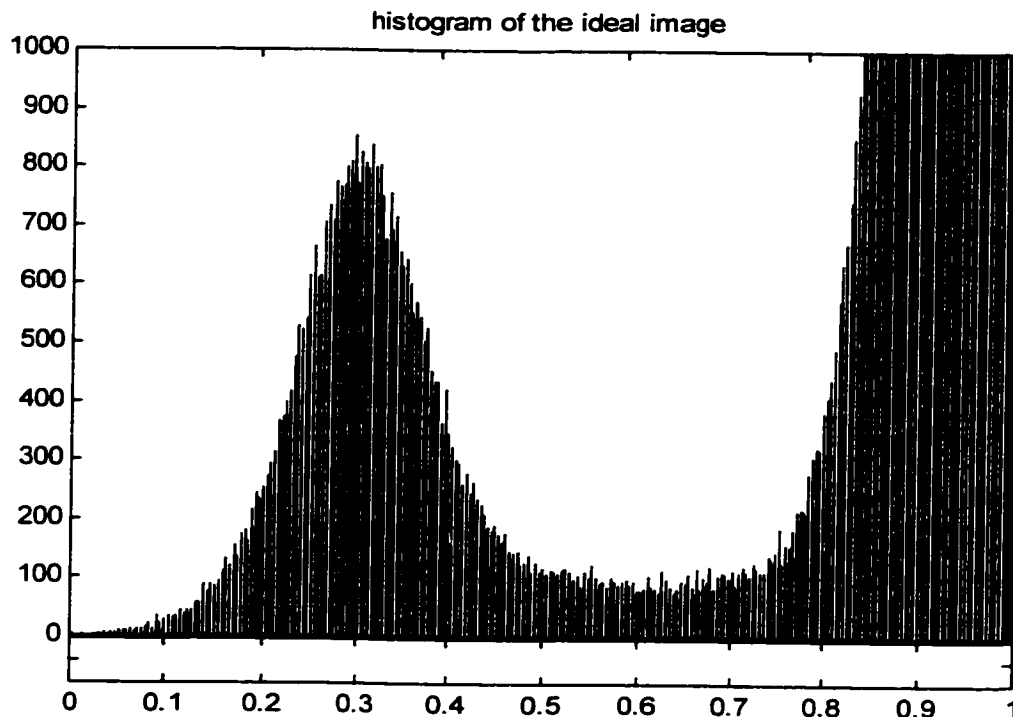


show-through. This approach works with only one side of the document at a time and only for a pale show-through.

The histogram of a degraded document shown in the Figure 1.2 elaborates the problem of high show-through. As shown in the histogram, the pixel intensity of some of the desired signal overlap with the value of show-through and that makes it difficult to separate the two. Due to this, applying a threshold to the histogram of the image does not give very efficient results either. If a threshold is chosen close to the show-through intensity, it will leave some of the show-through in the image and if the threshold is close to the signal value, it will remove some of the desired text signal. Another possibility is that, if the written *recto* is lighter than the show-through it is difficult to select a threshold to remove a show-through darker than the written *recto*. Therefore no general approach for removing the show-through exists yet.



**Figure 1. 2:** The histogram of the degraded image with high show-through.



**Figure 1. 3:** The histogram of the image with out show-through

The objective of this thesis is to find an approach to restore both sides of the document images by removing the show-through. Once we have a clear restored image of the document, we can make those documents available on-line and keep the document untouched by preserving them.

The novel technique proposed in this thesis allows both sides of the document to be restored at the same time. It is preferable to use a digital scanner to get the image of the document, but if the document is in too delicate condition to put on the scanner, then a digital camera can be used to get the image. Once the digital images of both sides of the document are obtained using any device either camera or scanner, the same restoration technique can be applied for both images, with some minor modifications. Getting both sides restored at the same time is an additional advantage of the technique introduced

here. In this thesis the image of the front of the document is called as *recto* and the back is called as *verso*. Some theoretical approaches are used to define the model of the scanned images and then these techniques are tested to find the best results for the restoration of image.

After getting the *recto* and *verso* image of the document, the *verso* image is flipped horizontally. The flipped *verso* image is similar to the *recto* but is different in intensity. Then the *recto* image and flipped *verso* images are registered very precisely using different approaches, e.g. pixel transformation method, affine transformation and a built-in MATLAB function. Once we get the perfectly registered image the restoration technique is applied to remove the show-through.

This thesis is organized in five chapters including this one. Chapter 2 presents the theoretical background of the digital images, image enhancement and some enhancement techniques, histogram, threshold, and a brief survey on different registration approaches. The theory about image restoration techniques is also included in chapter 2.

Chapter 3 presents the theoretical discussion about the different approaches for the restoration of degraded documents and the basis of the technique proposed. The model for the image restoration is set up and the various methods of registration are described which are related to the problem discussed here. The chapter is concluded with the discussion of restoration approach used for the technique proposed in this thesis.

The chapter 4 consists of the results obtained from the registration and restoration techniques discussed in the chapter 3. The result of each registration and the restoration techniques are discussed in detail. The subjective and objective approaches of selecting the most efficient technique are applied to achieve the best solution for restoration of the degraded documents.

In chapter 5 the conclusion of the results in this thesis is discussed. The chapter is concluded with the elaboration of the future aspects of working in this area.

## Chapter 2

## Theoretical Background

This thesis addresses quality improvement of documents degraded by show-through. The concept of quality improvement falls under the category of image enhancement and image restoration. The technique proposed in this thesis simultaneously uses the *recto* and *verso* images of a document for removing the show-through. These images are registered before enhancement, following which the restoration is carried out. The text documents considered in this thesis are either hand written documents, pictorial information, or sketches.

This chapter provides the theoretical background about the images, digital devices used to capture the digital images, followed by image enhancement and some of its techniques. Then the issues central to this thesis, namely the methods of registration are described and a brief survey of the different registration techniques is given. There are many techniques for registration depending upon the type of images, objective and application. However, only those which are similar to the techniques proposed in this thesis are discussed here. Finally, different approaches used for the restoration are discussed.

### 2.1 Image

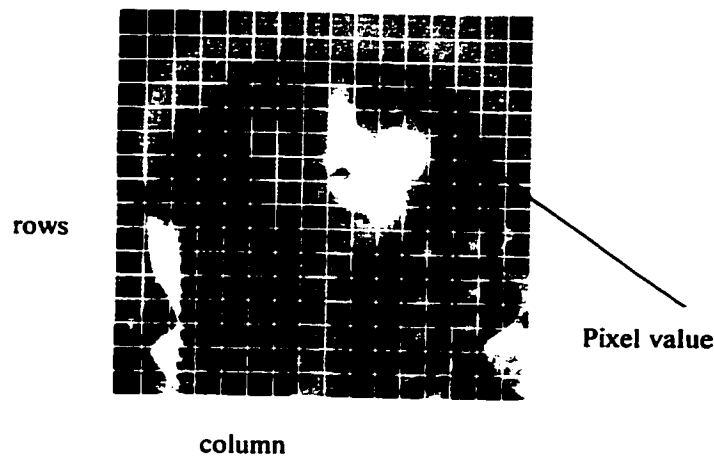
An image in everyday life is perceived as the light reflected from an object. “Theoretically, the term image refers to a two-dimensional light-intensity function, denoted by  $f(x,y)$ . The value or amplitude of function ‘ $f$ ’ at spatial co-ordinate  $(x,y)$  gives the intensity of the image at that point. The intensity of light is always positive and therefore can be represented as

$$0 < f(x,y) < \infty \quad \dots\dots\dots (2. 1)$$

## 2.2 Digital Image

Usually, the digital form of an image is used for computer processing. A digital image  $f[m,n]$ , described in a 2D discrete space, is derived from an analog image  $f(x,y)$  in a 2D continuous space through a *sampling* process that is frequently referred to as digitization [1].

The 2D continuous image  $f(x,y)$  is divided into  $N$  rows and  $M$  columns (see Figure 2.1). The intersection of a row and a column is termed a *pixel*. The value assigned to the integer coordinates  $[m,n]$  with  $\{m=0,1,2,\dots,M-1\}$  and  $\{n=0,1,2,\dots,N-1\}$  is  $f[m,n]$ .



**Figure 2. 1:** A continuous image divided into a discrete space

Suppose that a continuous image  $f(x,y)$  is approximated by equally spaced samples arranged in the form of an  $M \times N$  array as shown below:

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(M,0) & f(M,1) & \dots & f(M,N-1) \end{bmatrix} \dots (2. 2)$$

Here  $f(x,y)$  represents the digital image and each array element is referred to as an image element, picture element or a pixel. In this thesis, we will use the term pixel to refer to

the samples of the digital image. An image can have pixel values of several data types. A binary image has just two possible values often assigned to black and white; a gray level image is usually denoted with positive integer values from 0 to a maximum. A pixel can also have a real, complex or a negative value. For example, a thermal image with negative temperature has a negative value. In this thesis, we use a gray level image that has pixel values ranging from 0 to 255.

In digital processing of images, an image function  $f(x,y)$  is digitized in both spatial coordinates and amplitude. The spatial digitization is referred to as image sampling, and amplitude digitization is referred to as gray level quantization [2].

### 2.3 Gray Scale

The value of pixel intensity level of a monochromatic image is defined on a gray scale [3]. An image with two levels of gray scale is called binary. If a system has two binary bits, it represents four levels on the gray scale. Some ranges of gray scale are given below:

Gray scale	No. of levels	Range
$2^1$	2	0 – 1
$2^3$	8	0 – 7
$2^4$	16	0 – 15
$2^8$	256	0 – 255

The lowest value of the range represents a black pixel intensity and highest value represent the white pixel intensity. Increasing the number of gray levels increases the resolution and level of details in the image.

## **2.4 Image Acquisition**

An imaging sensor is required to acquire a digital image. These sensors are sensitive to certain bands in the electromagnetic energy spectrum such as X-ray, ultraviolet, infrared or visible bands, and produce an electrical signal output proportional to the energy. The imaging sensor can be digital or analog. Analog sensor generates a digitized signal as an output, *e.g.* digital camera, scanner, photocopier etc., the latter generates an analog signal as an output that can be digitized using an Analog to Digital (A/D) converter. In this thesis, a digital scanner is used to scan the document and obtain the image in its digitized form, *i.e.*, a signal matrix.

### **2.4.1 Scanner**

A scanner is a device that captures an image of the object such as a photograph, a transparency or any printed document, and converts it into a digital light-intensity map for computer processing. It converts the electrical signal into a two dimensional map of pixels, and each pixel represents the intensity measurement corresponding to the reflectance (for a solid surface) or the transmittance (for transparencies) of the object at that location [4].

**Figure 2. 2: A scanning mechanism (taken from [4] )**

As shown in *Figure 2.2*, a light source illuminates an original document. A horizontal strip of reflected light, called the raster, is captured by the CCD and is separated into color or black and white information. The photoreceptors, such as photodiodes, are located next to the CCD. These photodiodes measure the incident light and convert the measured value to an analog voltage. This analog voltage is converted to a digital value using an A/D converter to create the digital image for computer processing. The amount of light that is available during the scanning process corresponds to scanner responses with higher intensity of light leading to higher speed of the scanning. During each exposure, the scanner carriage is mechanically moved a small distance and this distance determines the sampling rate in y-direction. For example a 1/800 inch movement is equal to 800 dots per inch (dpi) in the y-direction. The width of copyboard divided by the number of elements in an array determines the sampling rate in x-direction.

There are many types of scanners available – drum scanner, flat bed scanner, and film scanner. In the drum scanner, as the drum spins, light from the image enters the lens and is recorded in a series of fine lines. It provides highest scanning resolution up to 12,500 dpi, but is very expensive. Most computer graphics professionals use the flat bed scanner. It resembles with a photocopier and can scan text documents, photographs, and transparencies. The document is placed on a flat bed facing down and it is scanned line by line.

The film scanner is used, directly to scan the negative of the film. A CCD consisting of three RGB line sensors is placed opposite to a fluorescent lamp. Images are scanned by moving the film between the CCD and the lamp. In this manner RGB sensors scan the image line by line.

## 2.5 Image Processing

The term image processing denotes the manipulation and analysis of images [5]. This can be divided into the following types:

- (a) *Digitization and compression*: This technique is used for converting an image to a discrete signal for computer processing and compressing it to economize on storage capacity or communication bandwidth for transmitting purpose.
- (b) *Enhancement, restoration and reconstruction*: The image signals are sometimes degraded by noise, low contrast or blurring. To obtain the original image or to improve it for analysis purposes, image enhancement, restoration or reconstruction techniques are used depending on the objective.
- (c) *Matching and recognition*: When images taken by different sensors or at different times are to be compared, we use the matching or registration techniques. This analysis includes segmentation of an image, measuring the properties of different parts and obtaining a relationship between the parts and comparing. The resulting descriptions are examined using certain models.

The ultimate goal of the above techniques is to help an observer translate the contents of an image into useful information.

## 2.6 Image Enhancement

After a digital image is obtained, it needs to be processed to improve its quality and make it suitable for a specific application [1]. This technique is called image enhancement.

The image enhancement operations generally fall into two categories:

- local operations, and
- global operations

A local operation includes those techniques that alter the value of a pixel based upon the values of its immediate neighbors. Global operation, on the other hand, refers to those

techniques that alter the value of a pixel by an identical amount regardless of the pixel's location in the image space.

Image enhancement refers to accentuation or sharpening of the image features, such as boundaries, edges or contrast to make graphical display more useful for analysis or display. This includes gray level manipulation, noise reduction, edge sharpening, filtering, interpolation, magnification, and pseudocoloring etc. There are two categories of image enhancement: spatial domain, which is based on the direct manipulation of the pixels in the image matrix, and frequency domain, which is based on the Fourier transform of the image. The various techniques of image enhancement are based on the above two approaches, either used individually or as combination of the two. These techniques are divided into four groups:

- (i) point operations
- (ii) spatial operations
- (iii) transform operations, and
- (iv) pseudocoloring

The above group of techniques can be further divided into the following approaches:

<b>Point operation</b>	<b>Spatial operation</b>	<b>Transform operation</b>
Contrast stretching	Noise smoothing	Linear filtering
Noise clipping	Median filtering	Root filtering
Window slicing	Unsharp masking	Homomorphic filtering
Histogram modeling	Low pass, band pass High pass filtering	
	Zooming	

**Table 2. 1: Image enhancement operations**

### 2.6.1 Point Operation

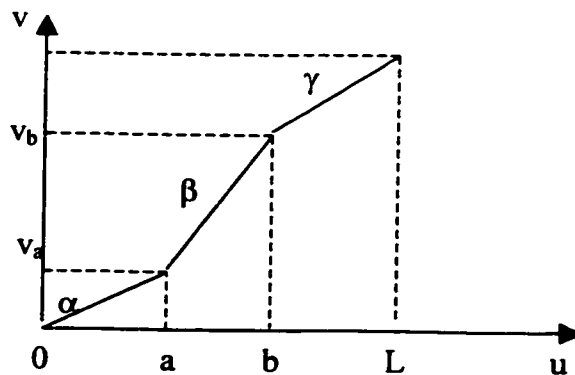
Point operation is a type of spatial domain approach. In the point operations, a given gray level  $u \in [0, L]$  is mapped into another gray level  $v \in [0, L]$ , according to the transform,  $v=f(u)$ , where  $f$  is a function of mapping. Several types of point operations are listed in Table (2.1). In this thesis, only a few point operations are used for the enhancement of images, therefore not all the approaches of point operation are discussed here.

#### 2.6.1.1 Contrast Stretching

The poor or non-uniform lighting conditions, nonlinearity or small dynamic range of the imaging sensor causes a low contrast of the image. To improve the contrast of an image, a contrast stretching transformation is performed as below:

$$V = \begin{cases} \alpha u & 0 \leq u < a \\ \beta(u - a) + v_a & a \leq u < b \\ \gamma(u - b) + v_b & b \leq u < L \end{cases} \dots\dots\dots(2.3)$$

where  $a$  to  $b$  is the region of stretch, (*please refer Figure 2.3*), within this region the slope of the transformation is chosen greater than unity in the region of stretch.



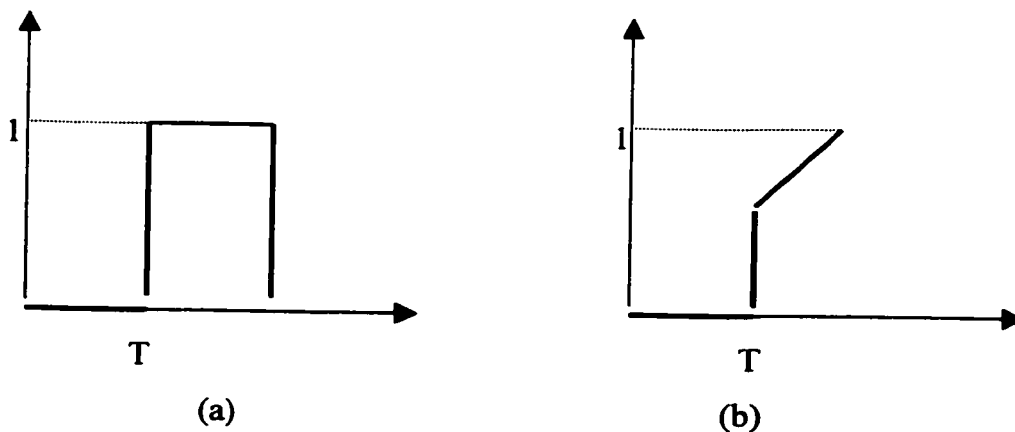
**Figure 2. 3:** Contrast stretching

### 2.6.1.2 Clipping and Thresholding

The clipping is another kind of point operation that is very similar to the contrast stretching [1]. In the equation 2.3, given above, if  $\alpha = \gamma = 0$ , it will clip any pixel with a value less than a and greater than b. This technique is helpful in the problems of noise reduction. If the range of the required pixel values is known, the rest of the pixels can be set to the highest or the lowest level. Any noise within this range is eliminated using this approach.

Thresholding is a special case of clipping, where the pixel intensity within the range (a, b) is set to a level, say t, and rest of the pixel values are set to the background. If the background has the same pixel values, e.g. white, then the thresholding is known as binary.

As shown in the *Figure 2.4(a)*, the gray level below a certain value is set to zero and above that it is set to the high level value L. Because the transformed image is defined by the two levels, it is called binary image. In *Figure 2.4 (b)* only the gray level below a certain value is set to zero.



**Figure 2. 4:** An example of binary thresholding

Thresholding is a type of point operation and a special case of clipping. It is a particularly useful region-approach technique for images containing an object with a contrasted background. In the threshold technique for image segmentation, all pixels at or above the threshold level are assigned to the object. If the object differs from their background by some property other than the gray level (texture, etc.), that property is first transformed to the gray level using some operation. Then the gray level thresholding is used to segment the image. In this thesis, text documents with contrast background are used.

The simplest and most widely used segmentation method is thresholding. It involves setting background values for pixels below a threshold value  $T$  and using a different set of values for the foreground. If the input image is  $f(x,y)$  and the thresholded image is  $g(x,y)$ , the equation for the thresholding operator is given by:

$$g(x,y) = \begin{cases} 0 & \text{if } f(x,y) < T \\ 1 & \text{otherwise} \end{cases} \quad \dots\dots\dots (2. 4)$$

### 2.6.1.3 Window Slicing

Window slicing is another type of clipping approach. This transformation allows us to segment the feature of a certain gray level from the rest of the region. If  $L$  is the highest intensity level of an image and the intensity level of required features are between  $a$  and  $b$ , then:

$$v = \begin{cases} L & a \leq u \leq b \\ 0 & \text{otherwise} \end{cases} \quad \text{without background} \quad \dots\dots\dots(2. 5)$$

$$v = \begin{cases} L & a \leq u \leq b \\ u & \text{otherwise} \end{cases} \quad \text{with background} \quad \dots\dots\dots (2. 6)$$

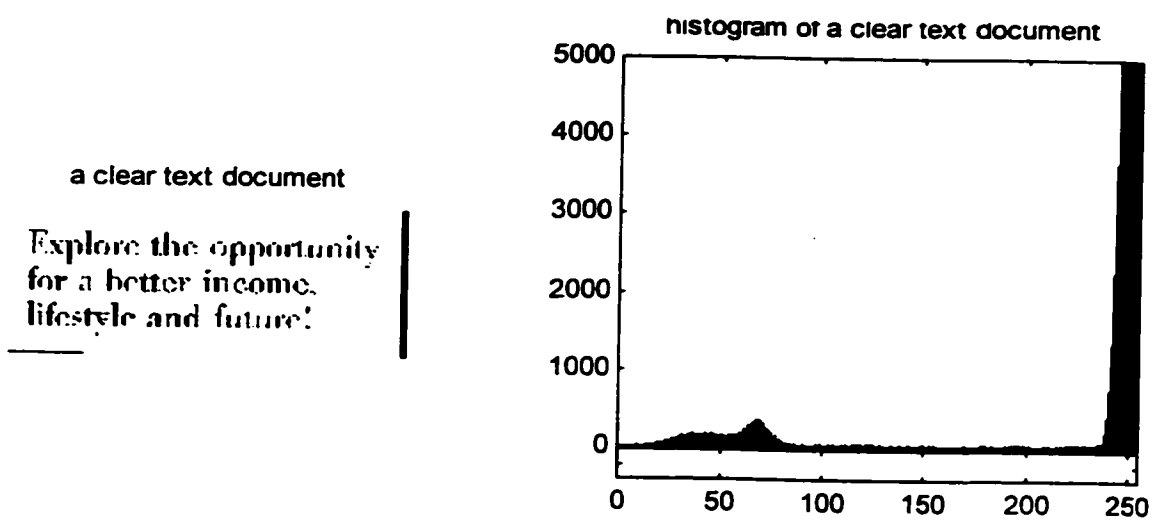
This approach is commonly used for remote sensing images.

#### 2.6.1.4 Histogram

Histogram is a function, which returns the relative frequency of the pixel values in a given image. The gray-level histogram is plotted on two axes: x-axis represents the gray level values and the y-axis represents the relative frequencies. In other words, a histogram is a list which contains as many elements as quantization levels. The number of pixels in each element show the corresponding gray level value stored. The calculation of a histogram is very straightforward. First, the entire list of the gray levels is set to zero. Then, the pixels of image are scanned, the gray value is taken as the index and the corresponding element of the list is incremented. The histogram gives a simple visual indication as to whether or not an image is properly scaled within the available range of gray levels. For example, if an image has been digitized between the gray level 0 to 255 and it has more brightness range, then the gray levels will be clipped at 0 and/or 255, producing spikes at both the ends of the histogram.

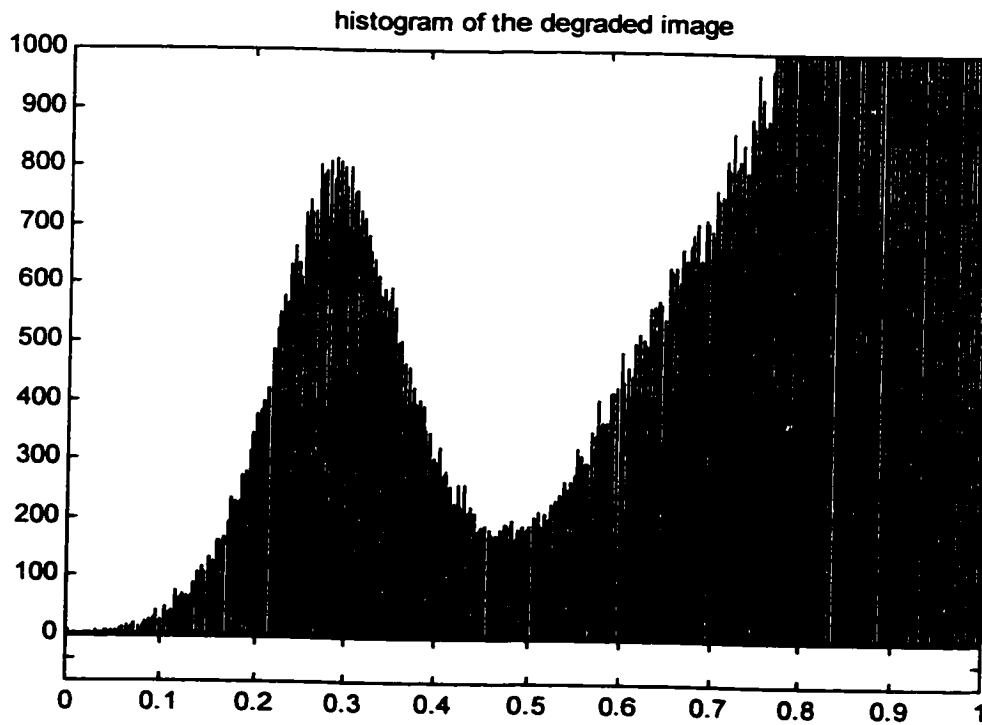
Two different images can have the same histogram but an image always has a unique histogram. A quick study of the histogram can give details about the problems with the image.

The image of a clear text document and its histogram is shown in *Figure 2.5*. It is clear from the image that the pixel intensity of the text, *i.e.*, the foreground, are condensed in one region close to the low level and pixel intensity of the background and close to the highest gray level. In another example of histogram (*please refer to Figure 2.6*), which is a sample of a degraded document, the gray level values of the foreground is spread out all over the gray scale. This shows the overlapping of the text and the show-through from the back of the document



**Figure 2. 5:** A clear text document and its histogram.

Around the world  
there is a treasure  
of excellent literature  
but not accessible  
to all the people



**Figure 2. 6:** A degraded text document and its histogram.

### 2.6.2 Spatial Operation

The spatial operations of image enhancement are performed on the local neighborhood of input pixels. The image is convolved with a finite impulse response filter called a *spatial mask, template, window* or *filter*. This technique is used to improve the image by removing blur, sharpening the edges, and replacing the pixel by the average of its neighboring pixels. Many types of spatial mask exist for the processing, *e.g.* Sobel edge detector, Laplacian edge enhancement operator, Robert's gradient operator, Low Pass filter, High Pass filter etc.

### 2.6.3 Transform Operation

This operation is applied in the frequency domain and in this technique the Fourier transform of the image is multiplied by the filter transfer function [5]. The inverse transform of the result gives an enhanced image. In practice, spatial masks are used more often due to their simplicity of implementation and speed of the operation. There are several types of transfer operation, *e.g.* homomorphic filtering, Lee filter, Butterworth filter, etc.

## 2.7 Geometric Distortion

When a document image is obtained from the scanner there are some distortion and misalignment between the *recto* and *verso* images. There are mainly three types of geometric distortions- translation, rotation and scaling [6]. Here all transforms are described in 2-dimensional Cartesian co-ordinates. Let the initial point  $(x_1, y_1)$  in image  $f_1$  be transformed on the image  $f_2$  at point  $(x_2, y_2)$ . The geometric transformation between two images can be described by two equations,  $x_2 = h_1(x_1, y_1)$  and  $y_2 = h_2(x_1, y_1)$ , where  $h_1$  and  $h_2$  are the transformation functions.

### 2.7.1 Translation

This transformation includes sliding the image up, down, left, and right. The x-values, *i.e.*, columns, give the left and right movements and the y-values, *i.e.*, rows, give the up and down movements. If the amount of translation is  $a_0$  and  $a_1$  then the transformed coordinate can be represented as below:

$$x_2 = a_0 + x_1 \quad \dots\dots\dots(2.7)$$

$$y_2 = a_1 + x_1 \quad \dots\dots\dots(2.8)$$

### 2.7.2 Rotation

The rotation transformation spins the image around the center point. The rotation by an angle  $\theta$  can be given as:

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \dots\dots\dots(2.9)$$

### 2.7.3 Scaling

This transformation is used to enlarge or shrink the image. To scale the x-direction x-values are used and similarly to scale the y-direction the y-values are used. The transformation is defined as :

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \dots\dots\dots(2.10)$$

The values of  $S_x$  and  $S_y$  are selected less than 1 to shrink the image, and more than 1 to enlarge the image.

### 2.7.4 Shear

The shear causes a distortion of pixels along one axis, proportional to their location in other axis. The shear acted along the x-axis, say  $SH_x$  can be represented as below.

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \dots\dots\dots(2. 11)$$

And the shear acted along the y-axis, say  $SH_y$  , can be represented as below.

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \dots\dots\dots(2. 12)$$

This type of distortion does not occur very often. The skew is a special case of shear when the distortion occurred along both axes.

The above discussed distortions can be combined into two equations each for x and y co-ordinates as follows:

$$x_2 = a_0 + S_x(\cos\theta x_1 - \sin\theta y_1) \quad \dots\dots\dots(2. 13)$$

$$y_2 = a_1 + S_y (\sin\theta x_1 + \cos\theta y_1) \quad \dots\dots\dots(2. 14)$$

Where  $a_0$  and  $a_1$  represent the translation only,  $S_x$  and  $S_y$  are scaling factors in x and y-directions respectively, and  $\theta$  is the angel of rotation.

## 2.8 Image Registration

Image registration is the process of geometrically aligning two images of the same object or scene. By definition the image registration can be defined as mapping between two images both spatially and with respect to intensity [6]. If we have two images  $I_1$  and  $I_2$

defined as two 2-D array of a given size, then the mapping between the two can be expressed as

$$I_2(x,y) = g(I_1(f(x,y))) \quad \dots\dots\dots(2. 15)$$

Where  $f$  is defined as a 2-D spatial co-ordinate transformation and  $g$  is a 1-D intensity or radiometric transformation.

The image registration has many applications in the various areas and substantial work has been done in this area [7]. Some examples of these applications include:

- *Medical image analysis*: detection of significant difference between biomedical images obtained from different subject or from the same subject at different time.
- *Pattern recognition*: object recognition, segmentation, motion detection, stereomapping and character recognition, and document restoration.
- *Remote sensing*: oceangraphy oil and mineral exploration, pollution and urban studies, military operation and forestry.

When we register two images there can be distortion, transformation, rotation or scaling differences. To map one image on to another those differences need to be removed using some techniques. Generally there are two methods that can be used to register images:

- Area based matching
- Feature based matching

The area based matching works by extracting the small area from one image and scanning the other image in order to find a similar area. Areas are matched by maximizing the cross correlation function. This method is very successful, but sometimes suffers from a number of limitations. The images need be matched have to be radiometrically very similar, *i.e.*, from the same sensor. Also the camera angle should be known in order to ensure the image has minimal differences in scaling and azimuth. For

these reasons, area based matching is in general, not applicable to multi source image matching.

The feature based matching, on the other hand, can be used to register same sensor or multi-sensor images. Usually the images that need to be registered are taken by the same image sensor therefore the research into feature based matching is less developed.

There are many approaches to solve the problem of image registration. One approach seeks to minimize the mean square difference between the pixel intensity values in a pair of images through the use of a gradient. Most approaches, however, consist of the following three steps. The first is to identify a set of points that correspond to the same point in the image. These match points are also referred to as corresponding points, control points, conjugate points and tie points [8]. The second is to compute the registration mapping function, which is the function that assigns to each point  $(x,y)$  in the reference image and is correspondent to the same control point in the data image. The last step is to identify the corresponding points in the two images using cross correlation technique or least square method.

Most of the registration approaches defined so far are problem based. Therefore before selecting a registration approach it is important to preprocess the image, *i.e.*, assess the data and select the correct way to proceed. In this thesis only those approaches, which are directly related to the objective are discussed.

The methods of registration can be broken down into four major components. For a particular application a different combination of choices are selected from the following components:

- a feature space,
- a search space,
- a search strategy and
- similarity matrix.

A feature space is used to define the feature of information in the image that is used for the matching. For example, for motion detection, the edge is defined as a feature to be used for the matching. The search space is the class of transformation such as FFT, affine etc. that is used to align the images. The search strategy decides how the feature selected in one space is use to find an optimal matching using the selected search space. The similarity matric component determines the relative matric for each test, for example, the least-squares method, the cross-correlation method etc.

### 2.8.1 Pixel Transformation

In this technique, the translation approach is used to register the two images. To perform this technique one image is transformed horizontally, one column at a time, and then vertically, one row at a time. During each transformation the Squared error between the pixel intensities of the transformed and reference image is calculated to identify the optimum match. The Squared error is minimized at the optimum point of matching. Let  $(x,y)$  be the pixel co-ordinate of a point in image  $f(x,y)$ . The image  $f(x,y)$  is translated by  $a_1$  horizontally to map the image  $f'$ . The new translated y co-ordinates of the image are defined by equation 2.16.

$$y' = a_1 + y \quad \dots\dots\dots(2. 16)$$

Similarly when the image is translated vertically by the points  $a_2$  the new x co-ordinate is defined by the equation 2.17.

$$x' = a_2 + x \quad \dots\dots\dots(2. 17)$$

These new co-ordinates, where the least square value of the difference in the pixels of two images is minimum, are the optimum matching point of the two images. Once the optimum point is found, the image is transformed by that point value. Then the transformed image is rotated by a small angle at a time and again the image is translated

to find the optimum matching point. This approach is very simple and gives reasonably good results.

### 2.8.2 Affine Transformation

Affine transformation is the most common general transformation. It can tolerate more complicated distortions while still maintaining some good mathematical properties. The affine transformation is composed of the Cartesian operation of translation, scaling, and rotation [9]. If we consider the two images  $f_1$  and  $f_2$ , the co-ordinates of image  $f_1(x_1, y_1)$  are mapped on the co-ordinate  $f_2(x_2, y_2)$ , the three operation are used to register the images as described in section 2.7.

The affine transformation is a global transformation, since the overall geometric relationships between points do not change, *i.e.*, a triangle in one image maps into a triangle in the second image. For a perspective type of transformation an affine transformation has six parameters, say,  $a_0, a_1, a_2, a_3, a_4,$  and  $a_5$ . If a point  $(x_1, y_1)$  of one image maps to the  $(x_2, y_2)$  of the second image then the co-ordinate  $(x_2, y_2)$  can be represented as follows:

$$x_2 = a_0 + a_2 x_1 + a_4 y_1 \quad \dots\dots\dots(2. 18)$$

$$y_2 = a_1 + a_3 x_1 + a_5 y_1 \quad \dots\dots\dots(2. 19)$$

In translation the origin is moved but axes remains parallel. In the above example the image is moved by  $a_0$  pixels in x direction and by  $a_1$  pixels in y direction, therefore only for the translation  $a_2 = a_5 = 1, a_3 = a_4 = 0$ . Similarly for the image rotated by an angle  $\theta$  the value of parameters will be  $a_0 = a_1 = 0, a_2 = a_5 = \cos\theta,$  and  $a_3 = -a_4 = \sin\theta$ . Thus we get the relation as given by the equation 2.9. For translation-zoom-rotation ( $a_0, a_1, a_2 = a_5, a_3 = -a_4$ ) all six parameters are used and for translation-zoom 3 parameters ( $a_0, a_1, a_2 = a_5 = 1$ ) are calculated. Using above six parameters any model of registration can be easily obtained.

### 2.8.3 Discrete Fourier Transform

The Discrete Fourier transform (DFT) method uses the Fourier domain approach to match the images that are rotated, translated and scaled with respect to each other [1]. In this technique the search for the optimal match is made according to the information in the frequency domain, unlike other search methods based on the spatial domain. This method relies on the properties of the Fourier transform such as shift, rotation, and scaling. Before defining the transformation a few terms used in the Fourier transform are explained first, which will make it easier to understand the transformation technique.

If  $f(x,y)$  is an image function its DFT and inverse DFT can be defined by the equation 2.20 and 2.21.

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux+vy)} \quad \dots\dots\dots (2. 20)$$

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux+vy)} \quad \dots\dots\dots (2. 21)$$

It is clear from the above definitions that DFT is a complex function; each value has a real part, say  $R(u, v)$  and an imaginary part, say  $I(u,v)$  at each frequency  $(u,v)$ . In terms of real and imaginary parts the DFT function can be written as:

$$F(u, v) = R(u, v) + i I(u,v) \quad \dots\dots\dots(2. 22)$$

Where  $i = \sqrt{-1}$ . Another alternative to express this term is the exponential form, as given below:

$$F(u, v) = |F(u, v)| e^{i\phi(u,v)} \quad \dots\dots\dots(2. 23)$$

where  $|F(u, v)|$  is the magnitude of the FFT and  $e^{i\phi(u,v)}$  is the phase angle. Let  $f_1$  and  $f_2$  be two images that differ only by displacement  $(x_0, y_0)$ , i.e.,

$$f_2(x, y) = f_1(x-x_0, y-y_0) \dots\dots\dots(2. 24)$$

Their corresponding Fourier transforms  $F_1$  and  $F_2$  are related by

$$F_2(u, v) = e^{-j2\pi(ux + vy)} * F_1(u, v) \dots\dots\dots(2. 25)$$

For images which are shifted with respect to each other, the Fourier registration method entails determining the location of the peak of the inverse Fourier transform of the cross power spectrum phase [7]. Since the phase difference for every frequency contributes equally, the location of the peak will not change if there is a noise, because noise is limited to a narrow bandwidth. Therefore this Fourier transform technique is well suited for shifted images with a noise of small frequency range. But if the images have significant white noise, which spreads all over the frequency range, it is difficult to find the accurate peak.

The rotational movement without translation can be represented in the same way using the phase correlation. Rotating an image rotates the Fourier transform of that image by the same angle. If  $f_2(x, y)$  is a rotated replica of  $f_1(x, y)$  with rotation  $\phi$  then,

$$f_2(x, y) = f_1(x \cos\phi_0 + y \sin\phi_0, -x \sin\phi_0 + y \cos\phi_0) \dots\dots\dots(2. 26)$$

From the 2-D similarity theorem, it follows that a rotation of  $f(x,y)$  through an angle also rotate the spectrum of  $f(x,y)$  by the same amount. Therefore the corresponding Fourier transform are related by

$$F_2(u, v) = F_1(u \cos\phi_0 + v \sin\phi_0, -u \sin\phi_0 + v \cos\phi_0) \dots\dots\dots(2. 27)$$

If  $f_1$  is a scaled replica of  $f_2$  with scale factors (a, b) for the horizontal and vertical directions, according to the Fourier scale property, the Fourier transforms of  $f_1$  and  $f_2$  are related by

$$F_2(u, v) = \{1/|ab|\} F_1(u/a, v/b) \dots\dots\dots(2. 28)$$

The Fourier Transform are very efficient for large value of translation, rotation or scaling.

## 2.9 Registration

A very important part of designing a registration method is to find the best matching point. There are many criteria to calculate the similarity measures for the matrices of both the images. The typical similarity measures are – Cross Correlation method and Least Square method. The choice of similarity metric is one of the most important part of registration and it gives the measure of the degree of similarity between the two images. For the given search space of transformation, the similarity metric is used to find the parameters of the final registration

### 2.9.1 Cross Correlation Method

The cross correlation is the basic statistical approach to registration. This can be used with or without prefiltering. For matching an image  $f_1$  with another image  $f_2$ , the two-dimensional normalized cross correlation function measures the similarity for each translation as below:

$$C(x_0, y_0) = \frac{\sum_x \sum_y f_1(x, y) f_2(x - x_0, y - y_0)}{\sqrt{\sum_x \sum_y f_2^2(x - x_0, y - y_0)}} \dots\dots\dots(2. 29)$$

If the two images match exactly for a translation of  $(x_0, y_0)$ , the cross correlation will have the peak value at  $C(x_0, y_0)$ . Thus, by computing C over all possible translations, the best matching value of translation can be calculated.

## 2.9.2 Least Squares Method

This is a very simple approach to calculate the similarity measures. This approach is better suited for the images with curves. For two matrices  $f_1$  and  $f_2$  the squared error can be calculated as given by the equation 2.30.

$$L(x,y) = \sum_x \sum_y (f_1(x,y) - f_2(x-x_0, y-y_0))^2 \quad \dots\dots\dots(2.30)$$

This measure decreases with the degree of similarity. For the best matching translation the value of squared error should be minimum.

## 2.10 Restoration

The area of image restoration deals with recovering image information that has been degraded. In other words *image restoration* means the removal or reduction of degradations that were incurred when the image was obtained. This degradation can be blurring due to the optical systems, image motion, noise due to the electronic and photometric sources, and unwanted information in the image such as show-through. To design a digital image restoration system it is necessary to quantitatively characterize the image degradation effects of the physical imaging system, the image digitizer, and the image display and then undo the degradation model to obtain the restored image. There are two basic approaches to the modeling of image degradation effects: *a priori* modeling and a *posteriori* modeling. The difference in the two modeling methods is in the manner in which information about the degradation is gathered. For an *a priori* modeling the response of imaging system, digitizer and display are measured on an arbitrary image field and for a *posteriori* modeling the model is developed for the image degradation based on the image to be restored. The restoration techniques can be divided as follows:

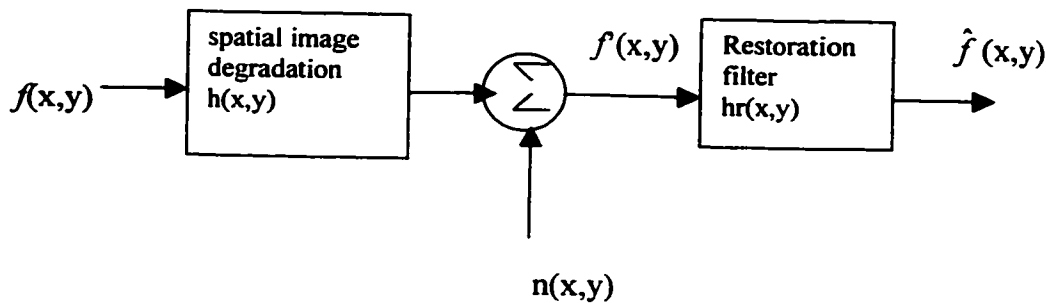
- Spatial image restoration techniques
- Point and spectral restoration techniques
- Geometric transformation

### 2.10.1 Spatial Image Restoration Techniques

This is the most common technique used for image restoration. It is used to compensate for image blur and additive noise effect caused by the sensor amplifier. The additive noise is independent of the image field. It is assumed that image blurring is modeled as a superposition of the operation with impulse response  $h(x,y)$ , (please refer to Figure 2.7). The model of a degraded image can be represented as follows:

$$f'(x,y) = f(x,y) * h(x,y) + n(x,y) \quad \dots\dots\dots(2.31)$$

where  $f'$  is the degraded image,  $f$  is the original image or the ideal image,  $h$  is the point spread function and  $n$  is an additive noise.



**Figure 2. 7:** An image degradation and restoration model

The objective of the restoration is to produce an array of samples  $\hat{f}(x,y)$  that are estimates of points on an ideal input image field  $f(x,y)$  obtained by a perfect digitizer. To produce a digital restoration model it is necessary to relate the physical image sample to the ideal image points. There are various methods of restoration, e.g. Inverse filter, Wiener convolution filter, Wiener deconvolution filter, recursive filter, least mean square filter etc.

### **2.10.2 Point and Spectral Restoration Techniques**

A post-processing correction of the sensor signal and pre-processing correction of the display signal can be reduced using this technique. The source degradation can be separated into the spatial and point effects.

### **2.10.3 Geometric Transformation**

Geometric transformations are different from the above discussed restoration techniques. This transformation modifies the spatial relationship between pixels in an image. The Geometric transformation are also called as *rubber sheet transformation*, because they can be compared with a printing on a rubber sheet and then stretching the sheet according to some predefined set of rules. There are two types of Geometric transformations: (1) a spatial transformation, which defines the rearrangement of the pixel on the image plane, and (2) a gray level interpolation, which deals with assigning a gray level to pixel based on some spatial analysis.

## Chapter 3

## Show-Through Correction Methods

In this chapter, various techniques applied for the restoration of *recto* and *verso* images of a document degraded due to the show-through are discussed. First, two simple approaches of image enhancement are discussed which use Adobe® Photo Deluxe<sup>(TM)</sup> Business Edition and applying a threshold for restoration of the image. This is followed by a discussion of the technique suggested for the restoration of documents with high show-through; that is registering the images and then applying a restoration method to obtain the clear *recto* and *verso* images. The explanation includes four methods of registration followed by three approaches used for restoration.

### 3.1 Image Enhancement

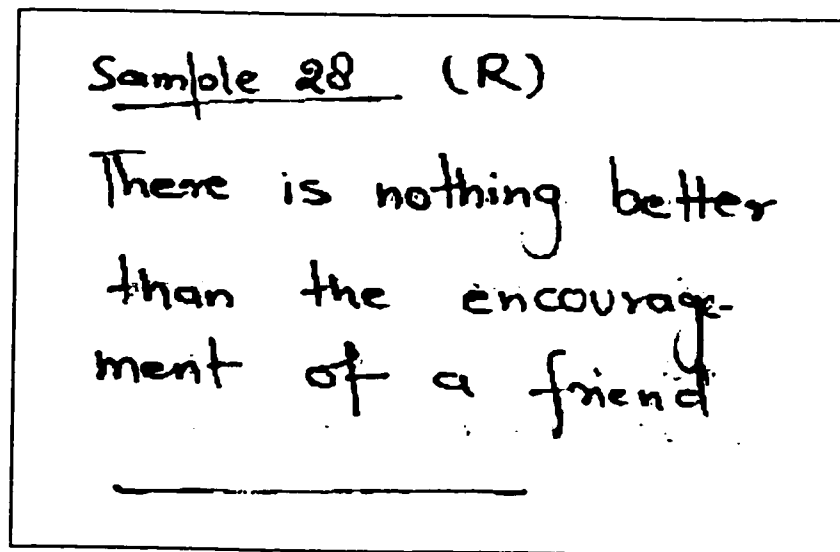
As discussed in section 2.6, the image enhancement technique is used to improve the visibility of pictorial information or to convert an image to a form that is better suited to the human visual system. For the old text documents which are degraded due to the show-through, we use this approach to remove the show-through from the document images. So far, no specific approach has been suggested for this type of problem. From the various approaches used for the enhancement, we chose two simple techniques of image enhancement:

- using Adobe® Photo Deluxe<sup>(TM)</sup> Business Edition software package, and
- applying threshold technique.

### 3.1 Adobe® Photo Deluxe<sup>(TM)</sup> Business Edition

The Adobe® Photo Deluxe<sup>(TM)</sup> Business Edition is one of a variety of advanced software packages available for image processing. This package is provided with the HP Scanner used in the VIVA lab. It allows one to import, organize and share the images, and also provides simple tools to edit photographs like a professional - from removing unwanted objects and people in the photographs to removing wrinkles, repairing damage and applying special effects.

In this approach, each side of the document is scanned and processed independently to remove the show-through from the *recto* and *verso* images. Using this software, the pale show-through can be removed totally. When processing documents with dark show-through it eliminates some of the desired signal and also modifies the texture of the paper, *i.e.* the desired details of the background is eliminated. This approach is therefore not very efficient for the documents with high show-through. It is used to restore each *recto* and *verso* individually. A sample image restored using the Adobe® Photo Deluxe<sup>(TM)</sup> Business Edition software package is shown below in Figure 3.1.



**Figure 3. 1:** The text document image restored using the Adobe® Photo Deluxe<sup>(TM)</sup> Business Edition.

### **3.1.2 Applying a Threshold**

By studying the histogram of *recto* and *verso* images (*please refer to Figure 2.6*), we conclude that the high show-through overlaps with the intensity of the desired signal. A certain value of threshold, say  $t$ , can be selected on the basis of the histogram of the degraded image with a pale show-through, as the overlapping of the desired signal and the show-through is not very significant. Any pixel with a higher value than the threshold is set to a high intensity level, *please refer to equation 2.6*.

This approach is good for a pale show-through but for the images degraded with high show-through, due to the wide overlapping of desired signal and show-through, it is difficult to choose a threshold value (*please refer to Figure 3. 2*). If a threshold is chosen close to the show-through intensity, it will leave some of the show-through in the image. For the threshold close to the signal value, it will remove some of the desired text signal. This approach does not give a very good result for the high show-through documents as can be seen in Figure 3.2, although it is better than using Adobe® Photo Deluxe™ Business Edition software package.

enhanced recto image using 0.48 threshold

Sample 28 (R)

There is nothing better  
than the encourage-  
ment of a friend

Figure 3. 2: Image enhancement using threshold

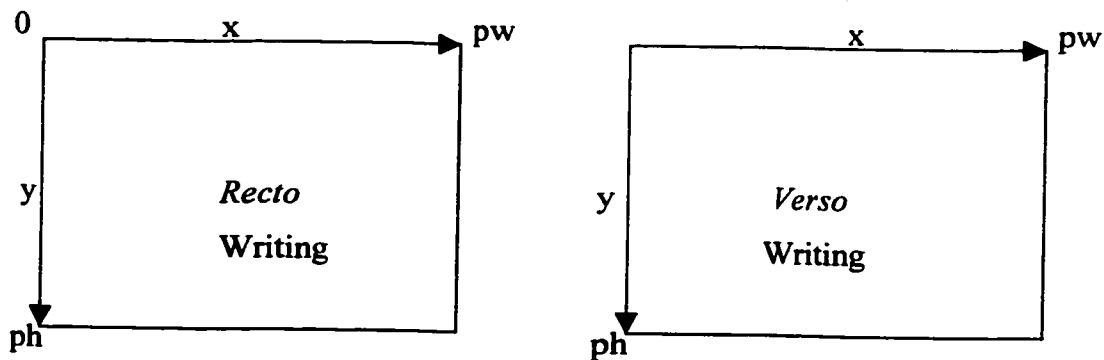
### 3.2 Model of Degradation

To perform a better enhancement, it is necessary to find a more appropriate technique. To define that technique a model of degradation is needed, which provides the basis for the enhancement approach. Let us assume that the degraded model of the *recto* image is  $f_r(x,y)$  and of the *verso* image is  $f_v(x,y)$ . As the light intensity value is always positive, and both the *recto* and *verso* images are scaled between 0 to 1, therefore  $0 \leq f_r(x,y) \leq 1$  and  $0 \leq f_v(x,y) \leq 1$ . For this model we assume that lower level 0 represents the white and higher level 1 represents the black on the gray scale of 256 gray levels. To remove a show-through, the restoration is carried out in the following three steps:

- a left–right flip of the degraded *verso* image,  $f_v(x,y)$  is obtained.
- the *recto* image and flipped *verso* image are registered precisely.
- show-through is removed using the same algorithm on the registered *recto* and flipped *verso* image.

In this thesis, the technique for the removal of show-through is based on the pixel by pixel comparison of the registered *recto* and flipped *verso* images to cancel the undesired signal. Most imaging devices induce geometric (spatial) distortion. Therefore, it is necessary to achieve a good registration before doing any significant work on the show-through removal algorithm. In the process of registration, we are basically attempting to register the desired *recto* image with the show-through on the *verso* image and the desired *verso* image with the show-through on the *recto* image.

The *recto* and *verso* images are sampled on a rectangular lattice with origin at upper left corner (please refer to Figure 3. 3). The width and height of the image can be represented by  $pw$  and  $ph$  respectively.



**Figure 3. 3:** A basic model of *recto* and *verso* image

We are representing ideal images as *recto* writing, *i.e.*  $f_{wr}(x,y)$ , and *verso* writing, *i.e.*  $f_{vw}(x,y)$ , which are only the writing with no background. A reflection operator  $\mathfrak{R}$  is used

to flip the *verso* image from left to right. The reflection operator  $\mathfrak{R}$  is a linear operator that does a left-right flip of an image and can be defined as below:

$$g = \mathfrak{R} f_v : g(x,y) = f_v(pw-x, y) \quad \dots\dots\dots(3. 1)$$

As defined before  $f_v$  is the degraded *verso* image and  $pw$  is the width of the image. The degraded *recto* image combines the background, the ideal *recto* image and the reflected ideal *verso* image in the following manner.

$$f_r(x, y) = C(f_{br}(x, y), f_{wr}(x, y), \alpha \mathfrak{R} f_{wv}(x, y)) \quad \dots\dots\dots(3. 2)$$

where  $f_{br}$  is the ideal *recto* background,  $f_{wr}$  is the *recto* writing,  $\mathfrak{R} f_{wv}$  is the flipped *verso* writing, and the  $C$  represents the combination operator. In the *recto* image, the *verso* writing appears as a show-through. In other words, the intensity of *verso* writing is higher than the show-through on the *recto* image due to the *verso* writing. Therefore, the flipped *verso* writing is multiplied by a coefficient  $\alpha$ , where  $\alpha < 1$ . The measured model of *recto* image defined in equation 3. 2 can be simplified as:

$$f_r(x, y) = \max(f_{br}(x, y), f_{wr}(x, y), \alpha \mathfrak{R} f_{wv}(x, y)) \quad \dots\dots\dots(3. 3)$$

In the same manner, the *verso* image can be given as follows.

$$f_v(x, y) = C(f_{bv}(x, y), f_{wv}(x, y), \alpha \mathfrak{R} f_{wr}(x, y)) \quad \dots\dots\dots(3. 4)$$

where  $f_{bv}$  is the ideal *verso* background,  $f_{wv}$  is the *verso* writing and  $\alpha \mathfrak{R} f_{wr}$  is the flipped *recto* writing.

A *verso* image with no geometric distortion and perfect registration can be written as follows.

$$f_v^l(x, y) = C(f_{hv}(x, y), f_{wv}(x, y), \alpha \mathcal{R}f_{wr}(x, y)) \dots\dots\dots(3. 5)$$

A transformation operator  $A_p$  can be applied to get the registered image with no geometric distortion, where  $A_p$  is a transformation operator that changes the x and y coordinates according to a transformation. We assume that any geometric distortion, as well as lack of registration in scanning the *verso* image with respect to the *recto* image in the measured image, can be modeled by such a transformation operator. When this transformation is applied on the ideal *verso* image with respect to the *recto* image, it gives the perfect match with the *recto* image, i.e.,  $A_p f_v^l = f_r$  where  $A_p$  is to be determined. For the given sample of the degraded *recto* image  $f_r(x,y)$  and degraded *verso* image  $f_v(x,y)$ , the restored images can be estimated as below:

$$\hat{f}_r = C(f_{hr}(x, y), f_{wr}(x, y), 0) \dots\dots\dots(3. 6)$$

$$\hat{f}_v = C(f_{hr}(x, y), f_{wv}(x, y), 0) \dots\dots\dots(3. 7)$$

where  $\hat{f}_r$  is the restored *recto* image and  $\hat{f}_v$  is the restored *verso* image, with no show-through. From the above model the actual problem now can be divided into two steps as follow:

- to estimate  $f_v^l(x,y)$  from  $f_v(x,y)$ , which is registration of the flipped *verso* and
- to use  $f_r(x,y)$  &  $f_v^l(x,y)$  to estimate  $\hat{f}_r$  and  $\hat{f}_v$  which is restoration, i.e. eliminating the show-through form both *recto* and *verso* images.

### 3.2.1 Registration

Let the model for the *recto*, *verso* and registered *verso* image can be re-written as simple linear combination model for  $C$ .

$$f_r(x,y) = f_{br}(x,y) + f_{wr}(x,y) + \alpha \mathfrak{R} f_{wv}(x,y) \quad \dots\dots\dots(3. 8)$$

$$f_v(x,y) = f_{bv}(x,y) + f_{wv}(x,y) + \alpha \mathfrak{R} f_{wr}(x,y) \quad \dots\dots\dots(3. 9)$$

$$f_v^l(x,y) = A_p f_{bv}(x,y) + A_p f_{wv}(x,y) + \alpha A_p \mathfrak{R} f_{wr}(x,y) \quad \dots\dots\dots(3. 10)$$

To register an image, we want to estimate a transformation operator, say  $A_p$ , and apply that on the image. In this thesis we are taking *recto* image as the reference image and the flipped *verso* image is mapped on the *recto* image. The reflection operator  $\mathfrak{R}$  is applied on the *verso* image and this flipped *verso* image is mapped on the *recto* image using transformation operator  $A_p$  (please refer to equation 3.11).

$$\mathfrak{R} A_q f_v = \mathfrak{R} A_p f_{bv} + \mathfrak{R} A_p f_{wv} + 2 \mathfrak{R} A_p \mathfrak{R} f_{wr} \quad \dots\dots\dots(3. 11)$$

If the registration is very precise then the subtraction of the flipped and transformed *verso* image from the *recto* image will be the minimum (please refer to equation 3.12)

$$f_r - \mathfrak{R} A_q f_v = (f_{br} - \mathfrak{R} A_p f_{bv}) + (\alpha \mathfrak{R} - \mathfrak{R} A_p) f_{wv} + (I - 2 \mathfrak{R} A_p \mathfrak{R}) f_{wr} \quad \dots\dots\dots(3. 12)$$

The cost function for the above equation can be defined as below.

$$\begin{aligned} \| f_r - \mathfrak{R} A_q f_v \|^2 = & \| f_{br} - \mathfrak{R} A_p f_{bv} \|^2 + \| (\alpha \mathfrak{R} - \mathfrak{R} A_p) f_{wv} \|^2 + \| (I - 2 \mathfrak{R} A_p \mathfrak{R}) f_{wr} \|^2 \\ & + 2 \langle f_{br} - \mathfrak{R} A_p f_{bv}, (\alpha \mathfrak{R} - \mathfrak{R} A_p) f_{wv} \rangle \\ & + 2 \langle f_{br} - \mathfrak{R} A_p f_{bv}, (I - 2 \mathfrak{R} A_p \mathfrak{R}) f_{wr} \rangle \\ & + 2 \langle (\alpha \mathfrak{R} - \mathfrak{R} A_p) f_{wv}, (I - 2 \mathfrak{R} A_p \mathfrak{R}) f_{wr} \rangle \end{aligned} \quad \dots\dots\dots(3. 13)$$

We can safely assume that the background of *recto*  $f_{br}$  and *verso*  $f_{bv}$  are unrelated to each other, and similarly the written *recto*  $f_{wr}$  and *verso*  $f_{wv}$  are unrelated to each other.

Therefore, the first norm involving background and the last three inner products, should be relatively small and insensitive to  $A_q$ . For precisely registered images the value of cost function will be minimum.

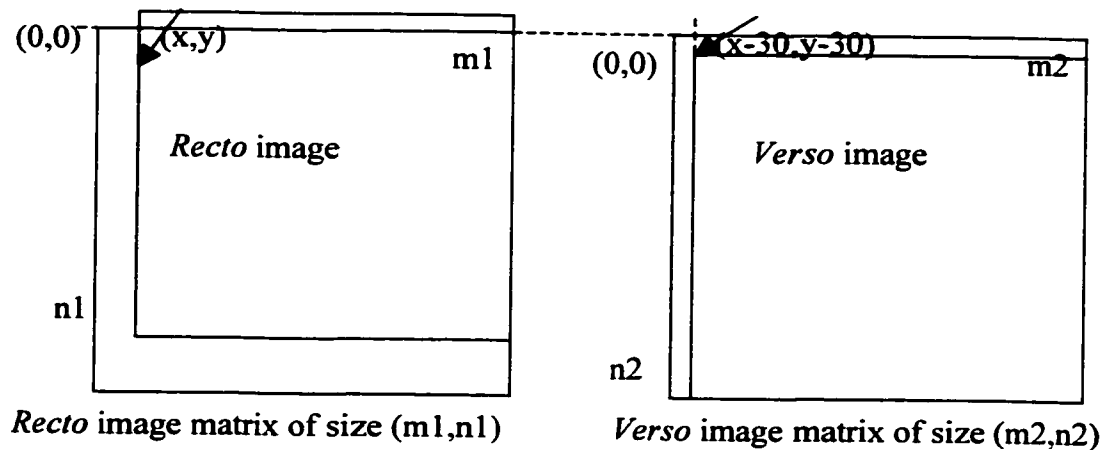
### 3.2.2 Restoration

Once the *recto* and *verso* images are registered precisely with each other an algorithm is applied to restore the images. The restoration approach used here is a spatial approach. The pixel on *recto* image,  $f_r(x,y)$  is compared with the same pixel on the registered flipped *verso* image  $f_v(x,y)$ . If the intensity of  $f_r(x,y)$  is less than the intensity of  $f_v(x,y)$ , then the signal on *recto* image is the desired signal, otherwise it is simply a show-through from the *verso*, which should be eliminated.

### 3.3 Image Registration

As described in section 2.8, the registration is used to map one image on to another for the same comparison. In this thesis we register the images in order to compare the two sides of a document and eliminate the unwanted signals, *i.e.*, the show-through coming from the other side. As shown in the equation 3.3, the *recto* image is a combination of the written *recto*, the show-through, which is the *verso* writing, and the background. The restored image model as defined in equation 3.7, is the combination of the writing *recto* and the background. On the other hand, from equation 3.10, we can conclude that the *verso* image is the combination of written *verso*, show-through coming from the *recto* and the background. Therefore both images have the same signal but with different intensities.

For the registration of *recto* and *verso* images, we assume that the horizontal and the vertical shift is within  $\pm 0.1$  inch and the rotation is within  $\pm 1.0^\circ$ . The resolution of the scanner is 300 dpi. Therefore, for  $\pm 0.1$  inch shift, we iterate 60 pixels down and 60 pixels right. The selection of the region from each image is shown in Figure 3.4.



**Figure 3. 4:** Selecting a region of interest from the *recto* and flipped *verso* images

During the scanning process, the document may have shifted to the right or left. Instead of mapping the *recto* left to the reference point  $(x,y)$ , we select a region of interest 30 pixels to the left from the reference point. Similarly, instead of sliding the image up to the reference point, a region of interest is selected 30 pixels up from the reference point. The *recto* image is considered as the reference image and *verso* is slid over the *recto* image to find the optimum matching point. The Mean Square Error (MSE) is calculated with each step of slide.

It is more likely that both sides of the document are scanned by the same scanner. This minimizes the optical distortion which occurs due to different image sensors. Therefore, the major cause of degradation of the images is geometric distortion. The document can be misaligned, *i.e.*, translated, rotated or sheared which might cause distortion. To register the *recto* and flipped *verso*, the following techniques are applied:

- Pixel translation approach.
- Pixel transformation (translation and rotation).
- Affine transformation (translation and rotation).
- Affine transformation (using six parameters).

### 3.3.1 Pixel Translation approach

This technique is used to register the image for the translation distortion only. First a window is chosen from the images (*please refer to Figure 3.4*). To find an optimum point for registration, the region selected from *verso* is isomorphically mapped to the selected region from the *recto* image. This approach is very simple and gives an acceptable result for the translation distortion. Here, we only use a sample of size 256x256 from both the images for the registration optimization. The disadvantage of this technique is that it does not include correction for the rotational distortion. A region of interest is selected from both *recto* and *verso* image to apply the algorithm. The algorithm used for pixel transformation is as follows:

#### ***Algorithm 1***

1. *select a sample of 256x256 from recto, say R1;*
2. *select a sample of 256x256 from verso, say V1;*
3. *for step= 1:1:30*
  - move V1 left by the value step;*
  - calculate the MSE between R1 and V1 ;*
  - end*
4. *find the point of left shift where MSE is minimum, say sl;*
5. *for step= 1:1:30*
  - move V1 right by the value step;*
  - calculate the MSE between R1 and V1;*
  - end*
6. *find the point of right shift where MSE is minimum, say sr;*
7. *compare sl with sr,*
  - if sl<sr*
    - move the verso image left by sl*
  - else*
    - move the verso image to right*
- end*
8. *again select the same sample of 256x256 from recto, say R2;*
9. *select the same sample of 256x256 from horizontally shifted verso, say V2;*
10. *repeat the step 5 to 7; except for moving V2 up and down instead of left and right.*
11. *find the value of step for minimum square error during each move, say sd for down and su for up movement*
12. *compare su and sd, if su<sd*

```

        move the verso image up by su
    or else
        move the verso image down by sd
end

```

### 3.3.2 Pixel transformation ( translation and rotation)

Sometimes when the document is being scanned there may be some rotational distortion because of the placing of the document on the scanner. Therefore, the approach explained in section 3.3.1, is not sufficient because it only corrects for the translation distortion between the images. A transformation is added for the rotational distortion in the above approach. The new algorithm for pixel transformation approach for the translation and rotation is described here.

#### **Algorithm 2**

1. *select a region of interest on the recto and verso images as given in algorithm 1;*
2. *calculate the optimum horizontal and vertical shift as given in algorithm 1*
3. *For a sample of 256 x256 pixels from recto (R1) and verso (V1)*
4. *for r1=0.1 : 0.1 : 1.0*
  - rotate image counterclockwise by r1*
  - calculate the MSE after each rotation between R1 and V1*
  - end*
5. *find the optimum rotation, say r1*
6. *for r2= -1.0 : 0.1 : 0*
  - rotate image clockwise by r2*
  - calculate the MSE after each shift between R1 and V1*
  - end*
7. *find the optimum rotation, say r2*
  - if r1 < r2*
    - rotate the verso image clockwise by r1*
  - or else*
    - rotate the verso image clockwise by r2*
  - end*

In this approach we are using a sample of small size from both recto and flipped verso images instead of the whole image. This might give a local correction for the geometric

transformation and therefore it is not an apposite way to calculate the optimum value for the transformation

### 3.3.3 Affine Transformation

The affine transformation registration is applied on the entire region of interest, unlike the pixel transformation approach. As described in section 2.8, the affine transformations are very efficient for small translations and rotations. To apply this, a region of interest is selected as shown in Figure 3. 4 and the algorithm 3 is applied to register the *recto* and *verso* images. In this approach the image is shifted to the location between the image pixels as the value of coefficients  $a_0, a_1, a_2, a_3, a_4,$  and  $a_5$  can be fractional. To apply the interpolation, an in-built interpolation function 'interp2' in Image Processing Toolbox of Matlab 5.3 is used. The 'interp2' function in the Matlab 5.3 provides three interpolation methods – 'nearest', 'bilinear' and 'bicubic'. For the most precise interpolation 'bicubic' method is used for this registration approach.

#### **Algorithm 3**

1. *select the region from verso image shifted by 30 pixels right and down;*
2. *for step=1:1:60*  
     *move verso image horizontally by the value step*  
     *calculate the mean square error between recto and transformed verso*  
     *end*
3. *measure the step value for minimum mean square error, say  $a_0$*
4. *for step=1:1:60*  
     *move verso image vertically by the value step*  
     *calculate the mean square error between recto and transformed verso*  
     *end*
5. *measure the shift for minimum mean square error, say  $a_1$*
6. *move verso image right by  $a_0$  and down by  $a_1$*
7. *for  $a=0.1 : 0.1: 1.0$  deg*  
      $a_5 = a_2 = \cos \theta,$   
      $a_5 = -a_3 = \sin \theta,$   
      $x' = a_0 + a_2 x + a_4;$   
      $y' = a_1 + a_3 x + a_5 y;$   
      $V'(x,y) = \text{interp2}(V(x', y'), \text{'cubic'});$   
     *calculate MSE between R and V'*

*end*

8. *select the value of  $a_2$ ,  $a_3$ ,  $a_4$ , and  $a_5$  for the minimum MSE;*
9. *transform the verso image using calculated affine coefficients;*

The affine transformation approach described above gives good results for the images that are translated and rotated. However, the above approach does not include the scaling, skew and shear distortion. Though these types of distortion give very low misalignment, but to get the most precise registration these distortions should also be taken into the consideration. To amend this problem the above approach is modified in the next section.

### **3.3.4 Affine Transformation with Six Parameters**

This approach is discussed briefly in section 2.8. It is based on using all six parameters individually to register the images,  $a_0$ ,  $a_1$ ,  $a_2$ ,  $a_3$ ,  $a_4$  and  $a_5$ . The parameters  $a_0$  and  $a_1$  are used for the translation and parameters  $a_2$ ,  $a_3$ ,  $a_4$  and  $a_5$  are used for rotation, scaling, skew or shear. For the purpose of matching, a MATLAB in-built interpolation function 'interp2' is used to obtain the registered flipped *verso* image. The constraints for parameter  $a_2$ ,  $a_3$ ,  $a_4$  and  $a_5$  are decided by the expected angle of rotation, which is assumed to be within  $\pm 1^\circ$ . The algorithm for this affine transformation approach is as follows.

#### **Algorithm 4**

1. *apply the algorithm 1 to calculate  $a_0$  and  $a_1$*
2. *move verso right by  $a_0$  and down by  $a_1$*
3. *for  $a_2=1.002 : -0.001 : 0.997$* 
  - $x' = a_0 + a_2 x + a_4 y$ ;*
  - $y' = a_1 + a_3 x + a_5 y$ ;*
  - calculate the new pixel intensity for the transformed co-ordinates*
  - calculate MSE between recto and transformed image*
- end*
4. *find the  $a_2$  for which MSE is minimum;*
5. *similarly repeat step 3 to find optimum  $a_3$ ,  $a_4$  and  $a_5$ ;*
6. *calculate the new co-ordinate  $x'$  and  $y'$*
7. *transformed the  $v$  to the new co-ordinate  $x'$  and  $y'$ ;*

### 3.4 Restoration

In this thesis, the model for images assumes that 0 represents the lowest intensity signal and 1 represents the highest intensity signal. When the *recto* image is matched with the *verso*, the show-through on the *recto* image is the written *verso* but higher in intensity and needs to be eliminated. On the other hand, the signal with low intensity on the *recto*, which is the written *recto*, is the desired signal. To restore these images, we apply three different approaches namely:

- Comparing pixel intensity
- Selecting a threshold
- Neighborhood Processing

#### 3.4.1 Comparing Pixel Intensity

In this approach, the pixel intensity for each pixel on *recto*, say  $r(x,y)$ , is compared with the pixel intensity of the same pixel on *verso*, say  $v(x,y)$ . If the intensity on *recto*,  $r(x,y)$ , is less than the intensity on *verso*  $v(x,y)$  then  $r(x,y)$  is desired. This eliminates the show-through and at the same time the intensity for the written *recto* is retained. This can be explained by algorithm 4 given as follows.

##### *Algorithm 5*

1. *read the recto image, say  $r(x,y)$ ;*
2. *read the registered flipped verso image  $v(x,y)$ ;*
3. *select a matrix of size(20, total number of column);*
4. *create a matrix  $b$  of background using the above sample by copying that(  $m/20 + 1$ ) times;*
5. *now  $b$  is a matrix of the background only*  
*if  $r(x,y) < v(x,y)$*   
*$b(x,y) = r(x,y)$*   
*end*
6. *the restored recto is  $b(x,y)$*

### 3.4.2 Selecting a Threshold

This step is an enhancement to the preceding approach. It is observed from the histogram of the degraded image that there is an overlap of show-through and written signals between the values 0.5 to 0.6. Therefore a change is made in the above algorithm when comparing the pixel intensity. Each time the pixel intensity of the pixels from *recto*  $r(x,y)$  and *verso*  $v(x,y)$  is compared, the pixel intensity of  $r(x,y)$  is also compared with the a threshold value  $h$ . If the  $r(x,y)$  is greater than  $h$  it is rejected. The value of threshold  $h$  is selected from the histogram (please refer to Figure 3.5). It is the lowest point in the valley in the histogram of a degraded *recto* image.

#### Algorithm 6

1. read the *recto* image, say  $r(x,y)$ ;
2. read the registered flipped *verso* image  $v(x,y)$ ;
3. select a matrix of size(20, total number of column);
4. create a matrix  $b$  of background using the above sample by copying that  $(m/20 + 1)$  times;
5. now  $b$  is a matrix of the background only  
if  $r(x,y) < v(x,y)$   
if  $r(x,y) < h$   
     $b(x,y) = r(x,y)$   
end  
end
6. the restored *recto* is  $b(x,y)$

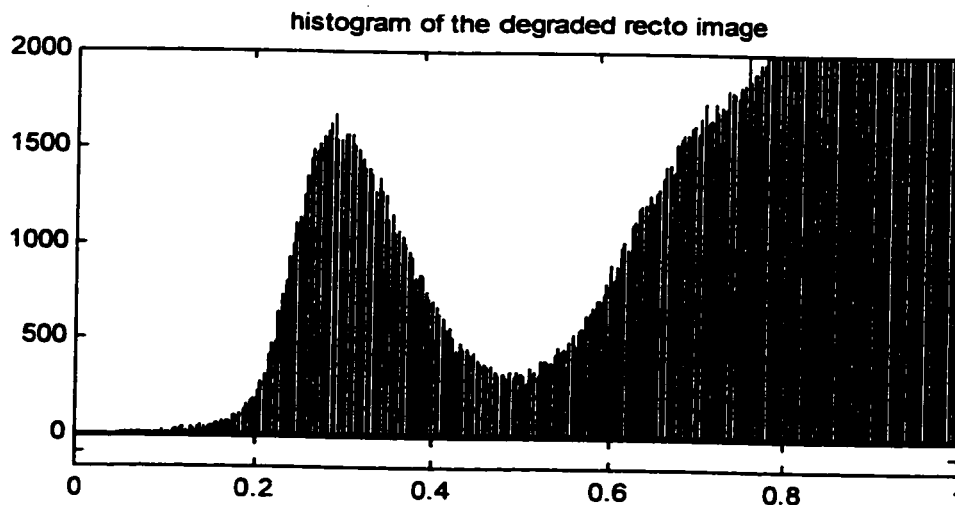


Figure 3. 5: A histogram of the degraded *recto* image

### 3.4.3 Neighborhood Processing

The above two approaches do not give good results unless the registration is very efficient. Using the technique described in section 3.4.1, some of the show-through still remains if the images are not matched very precisely; whereas the technique described in section 3.4.2 removes some of the written *recto* signals. A more efficient technique, which would give good result even with small misalignment, is required to restore the document images. A different approach with neighborhood block processing applied for the restoration is defined here.

#### *Algorithm 7*

1. *read the recto image, say  $r(x,y)$ ;*
2. *read the registered flipped verso image  $v(x,y)$ ;*
3. *select a matrix of size(20, total number of column);*
4. *create a matrix  $b$  of background using the above sample by copying that(  $m/20 + 1$ ) times;*
5. *now  $b$  is a matrix of the background only*
6. *let the size of the recto be  $m \times n$*
7.  *$w=3$  (the value of  $w$  varies form 3 to 9)*  
*for  $i=1:m$*   
*for  $j= 1:n$*   
*$a=$  sum of the block of  $w \times w$  neighborhood of the recto image*  
*$b=$  sum of the block of  $w \times w$  neighborhood of the verso image*  
*if  $a < b$*   
*$b(x,y) = r(x,y)$*   
*end*  
*end*  
*end*
8. *the restored recto is  $b(x,y)$*

### 3.4.4 Comparing the Ratio of Pixel Intensity

In this thesis the dark signal is represented by the low pixel value ( $f(x,y) = 0$ ) and the bright signal are represented as high pixel value ( $f(x,y) = 1$ ). This approach is based on the assumptions that for each position there is one of four possible cases.

- (i) There is no show-through, only the desired signal is present; the pixel value on the *recto* image is significantly lower than the flipped *verso* image
- (ii) There is no desired signal, only the show-through is present; the pixel value on the *recto* image is higher than the flipped *verso*
- (iii) There is no *recto* writing or any show-through, only the background is present; the pixel value on both the images is almost same.
- (iv) Both, desired signal and show-through are present; the overlapping of both the signals implies that pixel value of both the images are similar.

To eliminate the show-through the (ii) condition is considered. No processing is required for conditions (iii) and (iv) as the signal can be kept unchanged. Similarly the condition (i) occurs when the *recto* writing present and therefore the signal is desired. The algorithm for this approach is discussed below:

**Algorithm 8**

1. read the *recto* image, say  $r(x,y)$ ;
2. read the registered flipped *verso* image  $v(x,y)$ ;
3. select a matrix of size(20, total number of column);
4. create a matrix  $b$  of background using the above sample by copying that(  $m/20 + 1$ ) times;
5. now  $b$  is a matrix of the background only
6. calculate the ratio  $l$  of *recto* and flipped *verso*
7. let the size of the *recto* be  $m \times n$ 
  - for  $i=1:m$ 
    - for  $j= 1:n$ 
      - if  $r(x,y) < \text{threshold AND } l > 0.6$
      - $b(x,y) = r(x,y)$
      - end
    - end
  - end
8. the restored *recto* is  $b(x,y)$

In chapter 4 these approaches are tested and compared on several sample images with show-through.

## Chapter 4

## Experimental Results

The results obtained by using various approaches described in the previous chapter are discussed here in detail. The different techniques are implemented using MATLAB 5.3 software running on a Pentium MMX workstation in the VIVA lab. It is necessary to use the same image sample to provide an effective comparison of the different registration methods, thus two image samples are used for this purpose. To examine the efficacy of the restoration algorithms utilized in image restoration, different samples are used and the results are compared using subjective and objective methods. In the former method we compare the results visually with the ideal image, since the ideal image for each sample is available (*please refer to Figure 4.1 and 4.2*). In the latter method of comparison we perform two techniques, namely histogram comparison and calculation of Mean Square Error. To compare the histogram the histogram of the ideal image is available (*please refer to Figure 4.3*).

---

Oct 19, 1999.

Test sample. Recto

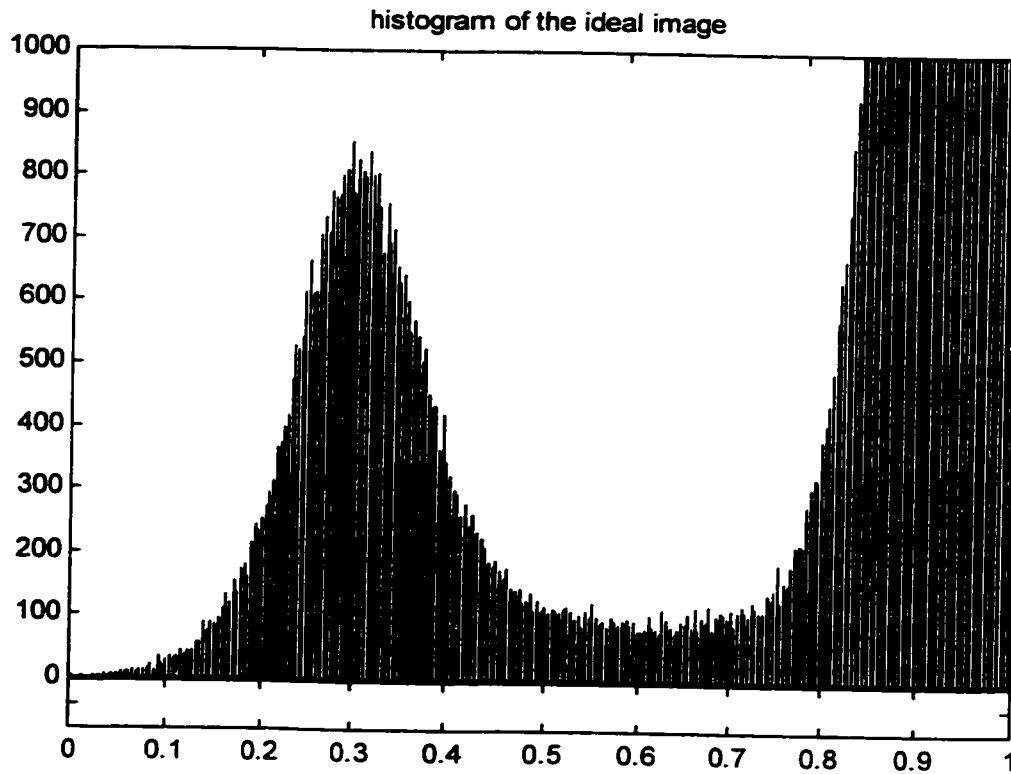
Image registration has few direct applications by itself, but it is a necessary step for computer analysis, tracking with multiple images, or motion analysis.

---

**Figure 4. 1:** The ideal image of the sample 1.

Around the world  
there is a treasure  
of excellent literature  
but not accessible  
to all the people.

**Figure 4. 2:** The ideal image of sample 2.



**Figure 4. 3:** The histogram of the ideal image

## **4.1 Image Enhancement**

As discussed in section 3.1, the image enhancement techniques are not well suited for the restoration of documents with high show-through. Two samples of degraded documents with a high show-through used here are shown in Figure 4.4 and 4.5.

Two enhancement techniques are used here: a built-in function of Adobe® Photo Deluxe™ Business Edition software package and applying a threshold.

Oct 19, 1999.

Test sample. Recto

Image registration has few direct applications by itself, but it is a necessary step for computer analysis, tracking with multiple images, or motion analysis.

Figure 4. 4: Sample 1 - The recto image with show-through

The recto image with bleed through

Around the world  
there is a treasure  
of excellent literature  
but not accessible  
to all the people

Figure 4. 5: Sample 2 – The degraded *recto* image

#### 4.1.1 Adobe Photo deluxe

The Adobe® Photo Deluxe™ Business Edition software is applied on the sample shown in the Figure 4.4 and 4.5. It is shown by the result that the entire show-through can not be removed (*please refer to Figure 4.6 a and 4.7*). If the image is processed further to remove the show-through, we start losing some of the desired signal (*please refer to Figure 4.6 b*). This process can be clearly shown in the Figure 4.6b, where the enhanced image has no show-through but the background is white and the gray scale is changed into a binary level. Also the further processing caused the elimination of the desired signal. This can be clearly shown at the edge of ‘s’ in the word ‘is’ and ‘e’ in the word ‘people’.

Around the world  
there is a treasure  
of excellent literature  
but not accessible  
to all the people

(a) sample 1

Around the world  
there is a treasure  
of excellent literature  
but not accessible  
to all the people

(b) sample 2

**Figure 4. 6:** The restored image using Adobe® Photo Deluxe™ Business Edition

---

Oct 19, 1999.

Test sample. Recto

Image registration has few direct applications by itself, but it is a necessary step for computer analysis, tracking with multiple images, or motion analysis.

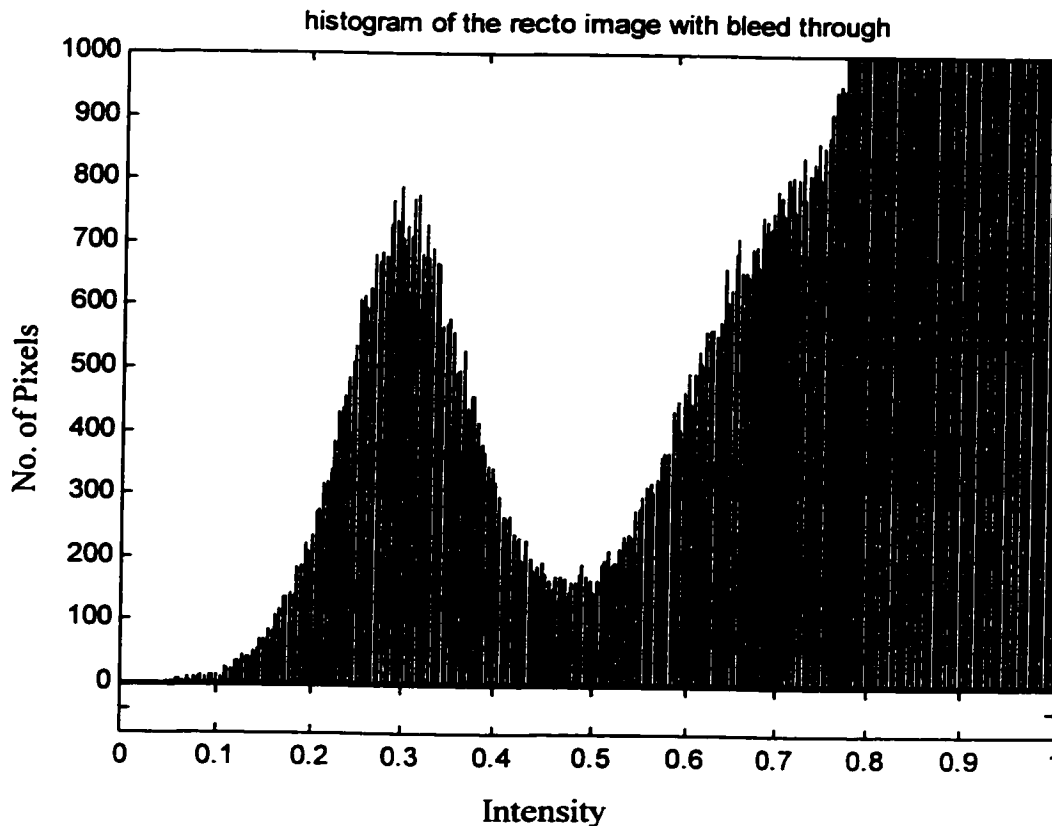
---

**Figure 4. 7:** Another sample of the restored image using Adobe® Photo Deluxe™ Business Edition

#### **4.1.2 Applying a Threshold**

In the histogram of the degraded image we see the overlapping of the desired signal and show-through (*please refer to Figure 4.8*). Thresholding distinguishes the pixels that have lower gray values, *i.e.* the desired signal, from the pixels that have higher gray value, *i.e.*, the show-through and the background. As described in section 3.1.2, a threshold value is chosen from the histogram as a lowest value in the valley between the two histogram modes. The valley represents the overlapping of the desired signal and

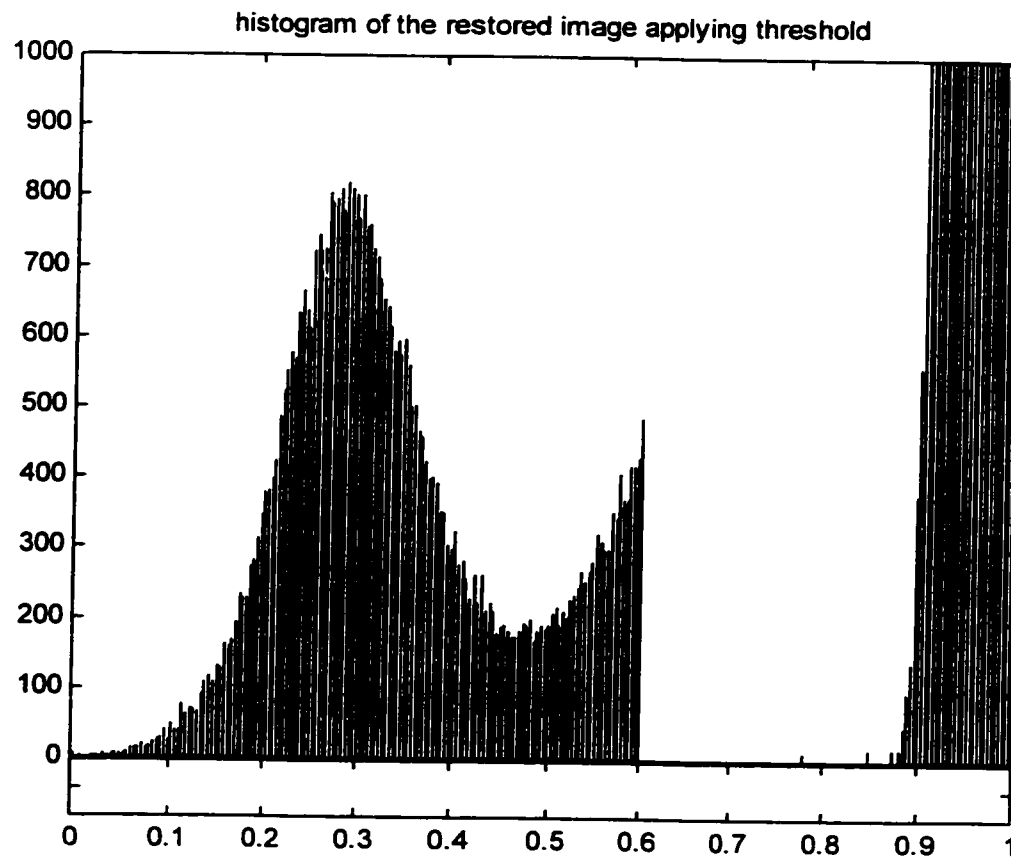
show-through from the *verso*. For the very high show-through the distribution of dark and bright pixels become more and more difficult. That makes the selection of the threshold value more difficult.



**Figure 4. 8:** The histogram of the degraded *recto* image

In this example reasonably accurate threshold value is chosen as 0.5, which is the lowest point in the valley of the histogram. The pixel values less than 0.5 is replaced by the 0 and the pixel values greater than 0.5 is replaced by the background. The result of the threshold approach is shown in Figure 4.10. As is evident, this approach does not give good results either. When the threshold value 0.6 is selected some of the show-through, with pixel intensity less than 0.6, is difficult to remove, however if we choose a lower threshold value, it eliminates some of the desired signal. The histogram of the restored image is shown in Figure 4.9, which is very similar to the histogram of the ideal image (*please refer to the Figure 1.3*). But in the histogram of the restored image a region between the 0.5 and the background pixel value is flat and some background information

is also removed. Therefore, both of the aforementioned enhancement techniques are not efficient enough to remove the high show-through. The low value of threshold causes the elimination of the desired signals (*please refer to Figure 4.11*).



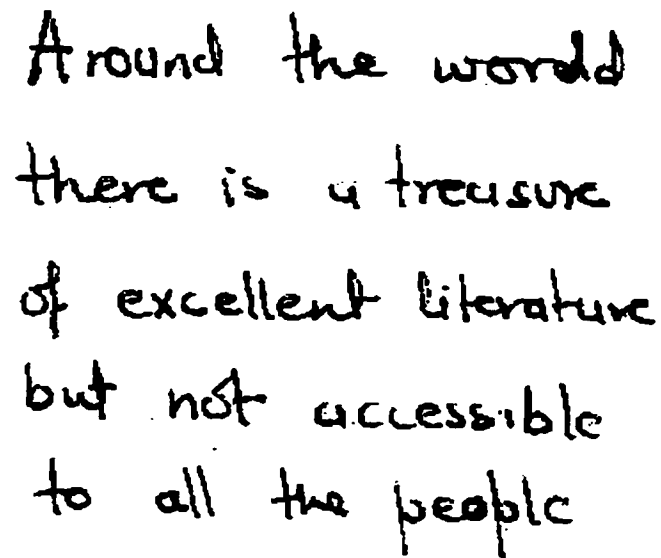
**Figure 4. 9:** The histogram of restored image using threshold approach

enhanced recto image using 0.6 threshold

Around the world  
there is a treasure  
of excellent literature  
but not accessible  
to all the people

Figure 4. 10: The restored image using threshold value 0.6.

enhanced recto image using 0.45 threshold



Around the world  
there is a treasure  
of excellent literature  
but not accessible  
to all the people

**Figure 4. 11:** The restored image using threshold value 0.45

The techniques of image enhancement, discussed above, are not very efficient for the images with high show-through. Therefore, a new approach was suggested in section 3.3. In this approach both the *recto* and *verso* images are registered precisely and then a restoration technique is applied on the registered images to get the improved image. The first step of this approach is to find the most efficient method of registration and the next step is to implement the restoration. The next section sets forth the comparative study of the results emanating from the registration techniques discussed in the section 3.3.

## 4.2 Registration

While scanning the documents there is some shift or distortion between the *recto* and *verso* images. When *verso* image is flipped and mapped to the *recto* there is a mismatching. To eliminate this different techniques are suggested in this thesis (*please refer section 3.3*) for the registration of the two images.

### 4.2.1 Pixel Translation Approach

As discussed in section 3.3.1, a sub-matrix is selected from the image to calculate the translation. Figure 4.13 shows the registration of *recto* and flipped *verso* images before applying the transformation.

The pixel transformation approach is discussed in detail in section 3.3.1. In this approach a small region is selected from the two images and the optimum value of the transformation between the *recto* and flipped *verso* is calculated. Then the flipped *verso* is transformed by that shift value and the average between both the images is obtained. This gives the mapping between the two images (*please refer to Figure 4.14*). By the visual inspection of the Figure 4.14, it is obvious that the area from which the sample was taken is well registered, whereas there is a not a precise matching in the rest of the region. For a sample of 256x256 the time taken in the calculating the optimum transformation is 356sec on a Pentium MMX 500MHz PC. For a region of 400x400 the time taken is more than 800 sec. Therefore, it will take a lot of time and memory space to process the entire image to calculate the translation. Obviously, this approach does not give very efficient transformation and consume a lot of time and memory space.

Registered recto and flipped verso images before transformation

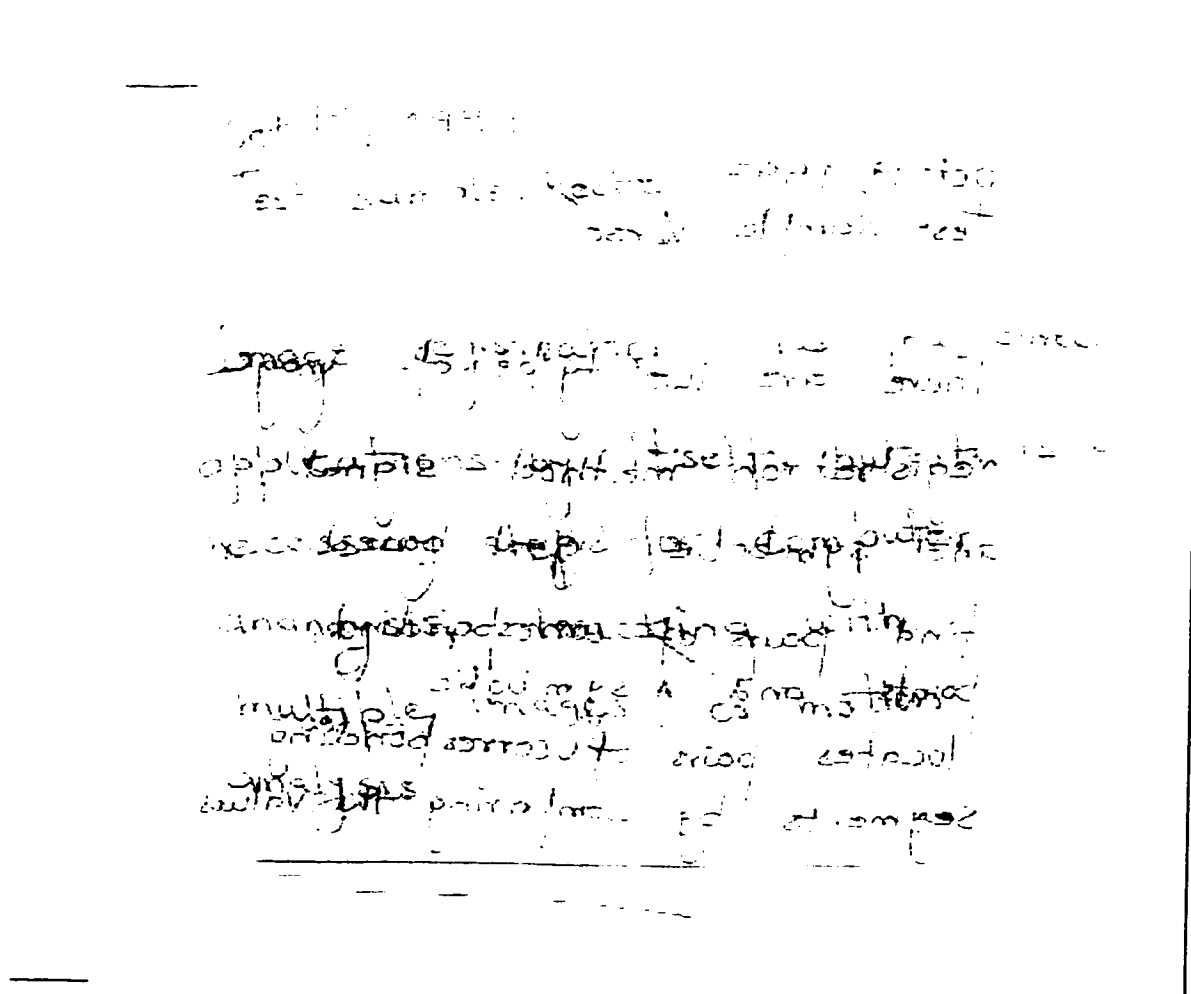


Figure 4. 12: The registered *recto* and flipped *verso* images before transformation.

registered image after the transformation

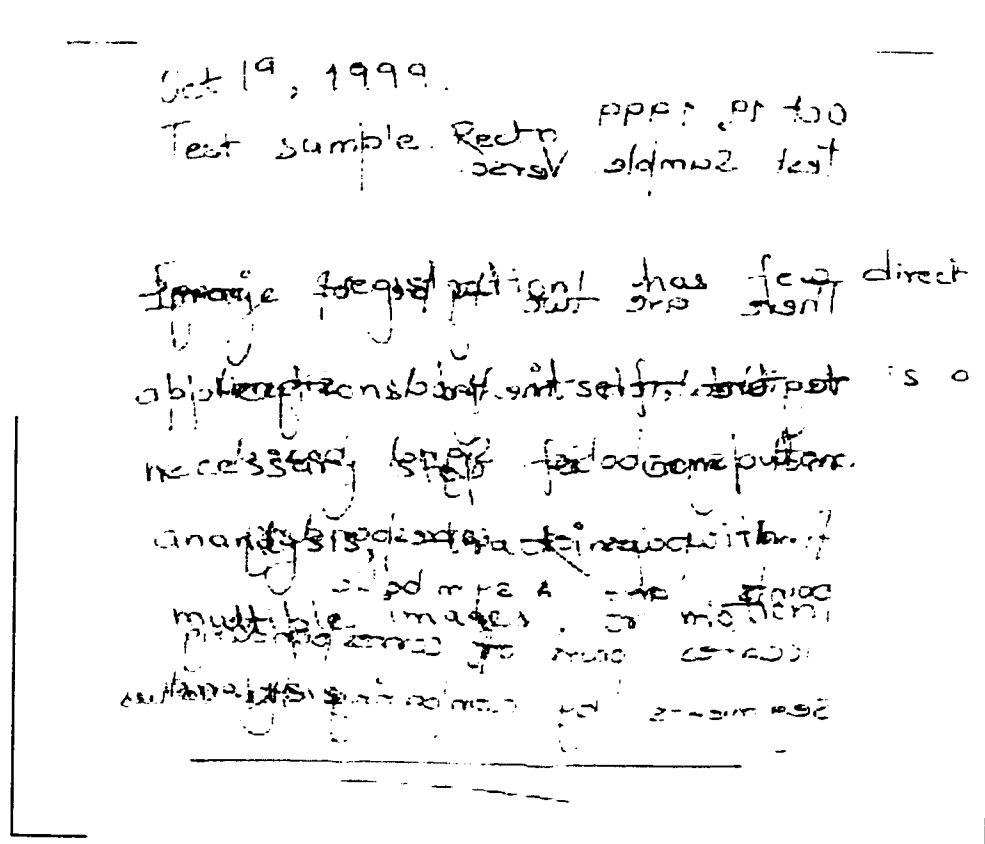
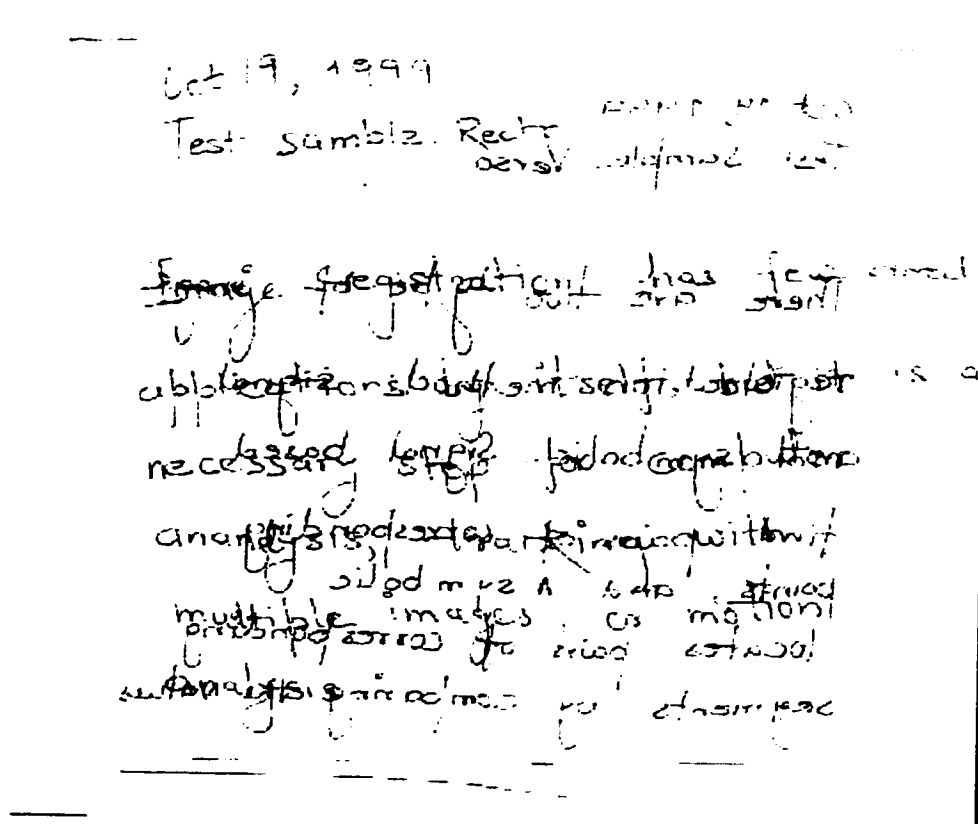


Figure 4. 13: The registered image using pixel transformation approach.

#### 4.2.2 Pixel Transformation with Rotation

In the above approach we did not consider the rotational distortion. There is a high probability of a rotational distortion during the scanning of the documents. Therefore, the correction for the rotational distortion is added to find more accurate registration. This registration approach is an extended part of the previous Pixel Transformation technique. In this approach after calculating the translation between the images, the image is rotated by an angle of  $1.0^\circ$  clockwise and anti-clockwise in increments of  $0.1^\circ$ . A small region is selected from the image to adjust for rotational distortion. It is assumed

that there is a small rotational distortion between the *recto* and flipped *verso* image. The flipped *verso* is rotated by an angle of  $1.0^\circ$  in steps of  $0.1^\circ$ .



**Figure 4. 14:** The registered image using pixel translation with rotation.

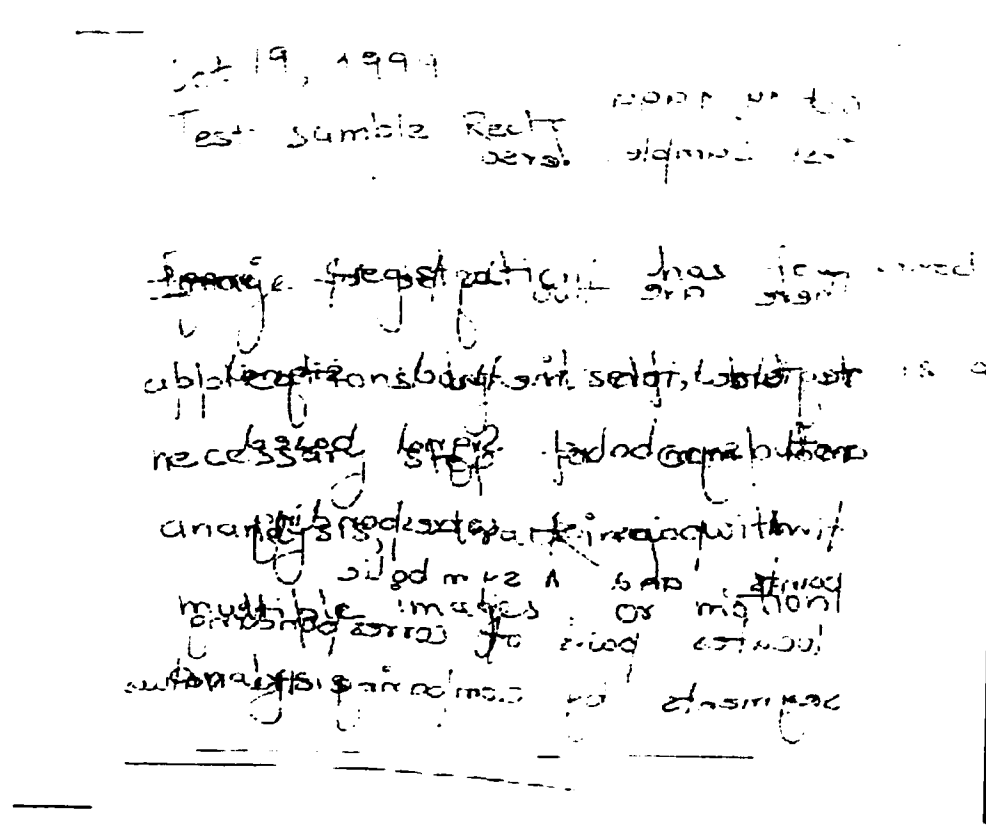
After each rotation the MSE is measured to get the optimum angle of rotation. The process of rotation adds some more calculation and hence more computation required which consumes more time. In other words, even with a small sample of size  $256 \times 256$  it takes almost 50% more time than the previous registration technique to find the optimum registration.

This approach provides reasonably good result for the registration, (*please refer Figure 4.14*). There are two factors which lead to a decline in the efficiency of this technique. First, the time taken for the calculation for each document is very high and second, because a small area is selected for the computation, it does not give the optimum value for the whole image. If the whole image is considered for calculating the transformation operators it take a lot of time and memory space. Due to these two reasons the Pixel Transformation approach is not feasible.

### **4.2.3 Affine Transformation**

As discussed in section 3.3.3, this is a very simple approach for image registration. The affine transformation discussed here includes both translation and rotation distortion. The scanned images may have both types of distortion but in small value, as the document is scanned at the same time and placed in the same position. Therefore, this technique is designed for  $\pm 0.1''$  translation and  $\pm 1.0^\circ$  of rotation. To calculate the optimum transformation the entire region of interest is considered. To avoid the calculation for left and right shift, or up and down shift, a region is initially selected as described in Figure 3.4. Then the translation is calculated in two directions only right and down. This reduces the computational requirements.

This approach provides a better result than the Pixel Transformation approach (*please refer to Figure 4.15*) as the entire region of interest is considered to calculate the optimum value of transformation for the whole image, which leads to good registration.



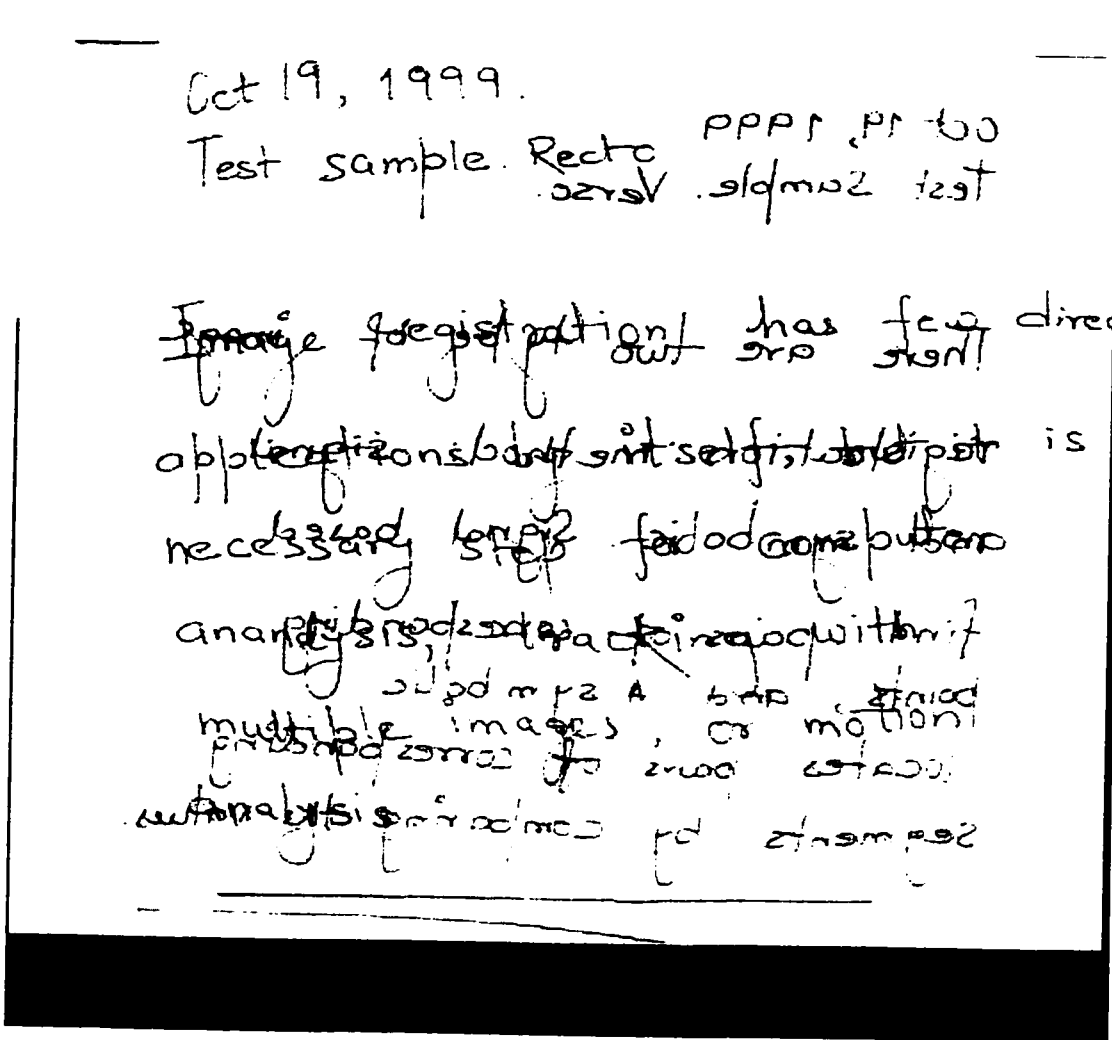
**Figure 4. 15:** The registered image using affine transformation

#### 4.2.4 Affine Transformation with six Parameters

This is more efficient approach of registration as it calculates the transformation for all type of distortions discussed in section 2.7. A matrix of six parameters is calculated to transform the image (*please refer to section 3.3.4*). In this technique the entire region of interest is considered to calculate the transformation. The matching pixels are calculated using interpolation, therefore it gives very precise values of the six parameters. The first two parameters  $a_0$  and  $a_1$  are used for the translation (*please refer to equation 2.33 and 2.34*). The other 4 parameters  $a_3$ ,  $a_4$ ,  $a_5$ , and  $a_6$  are used to eliminate the rotational, scaling

skew or shear distortion. This approach gives fast and most precise registration (*please refer to Figure 4.16*). For an image of size 957 x 957 the affine transformation approach takes 288 sec. This is significantly lower than the pixel transformation approach, which is 800 sec for an image of 400 x 400.

registered image after the transformation



**Figure 4. 16:** The registered image using affine transformation approach.

### 4.3 Restoration

After the registration of the *recto* and flipped *verso* images a restoration approach is applied on the registered images. To achieve the best restoration different techniques are applied (*please refer to section 3.4*). Before the restoration the best method of registration is selected from the above discussed techniques of registration. It is concluded, that the 'affine transformation with six parameters' technique is best to register the image more precisely. Therefore, all the samples used for the restoration are registered using this approach and then different restoration techniques are applied on them.

#### 4.3.1 Comparing Pixel Intensity

This is a very simple approach applied on the registered image to restore the document, (*please refer to section 3.4.1*). In this technique the pixel intensity of the same pixel on both the images is compared and the show-through signal is eliminated. This approach eliminates some show-through and also some of the desired signal (*please refer to Figure 4.17*). If there is a slight mismatch in the mapping, the pixel in the flipped *verso* is writing but in the *recto* it is not the show-through, then it is possible that instead of eliminating the show-through the desired signal is eliminated. On the other hand some signal on the *recto* is show-through but not written *verso* and they are not eliminated. The histogram of the image gives peaks at the same point but shows some missing pixels in the image (*please refer to Figure 4.18*). The MSE (mean square error) between the ideal *recto* and the restored *recto* image is calculated as 0.0320.

This approach gives very good results with precisely registered matched images, but when the images are not mapped precisely the results are not satisfactory. To overcome the problem of the mismatching in registration another approach is suggested in the section 3.4.2

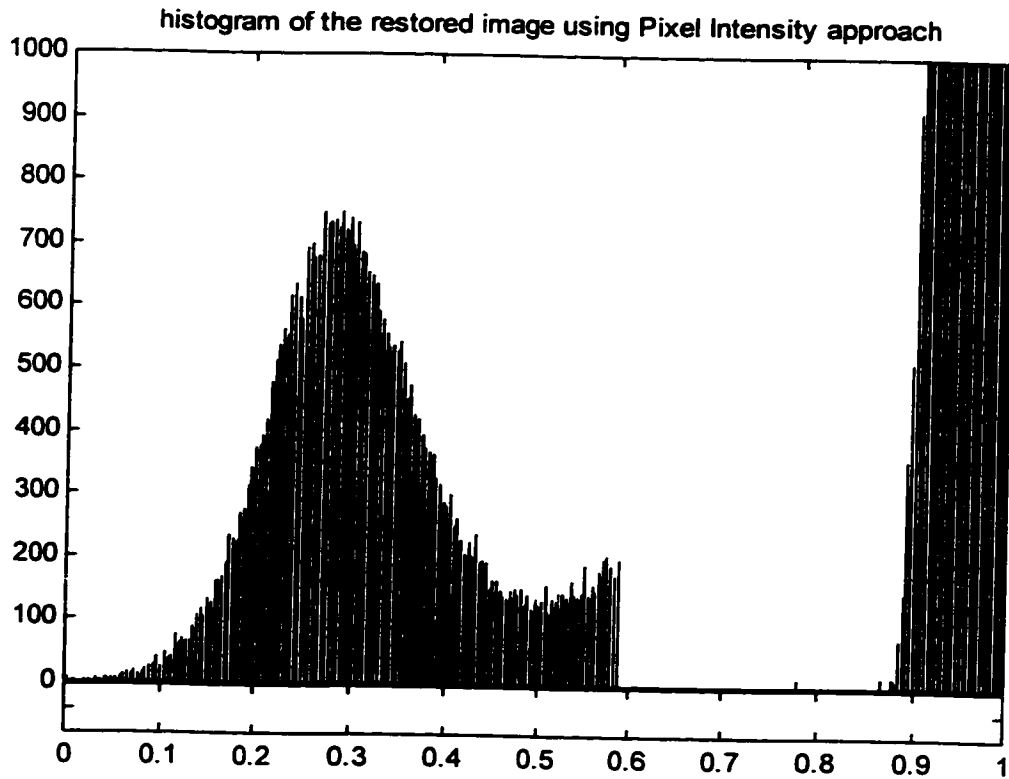
restored recto

Oct 19, 1999.

Test sample. Recto

Image registration has few direct applications by itself, but it is a necessary step for computer analysis, tracking with multiple images, or motion analysis.

**Figure 4. 17:** The restored image obtained by comparing pixel intensity

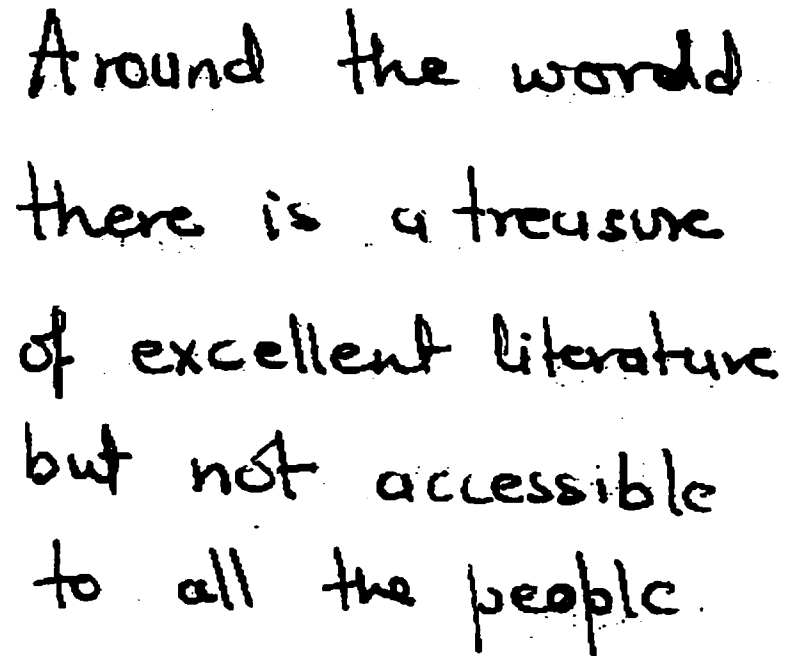


**Figure 4. 18:** The histogram of the restored image

### 4.3.2 Neighborhood approach

This technique is based on the neighborhood operator whose output depends on the pixel intensity of the neighboring pixel values as well. A neighborhood of 3x3 is selected each time for the comparison (*please refer to section 3.4.2*). This approach gives a good restored image of the sample 1 (*please refer to Figure 4.19*). The histogram is close to the histogram of the ideal image. The MSE for sample 2 is 0.0178.

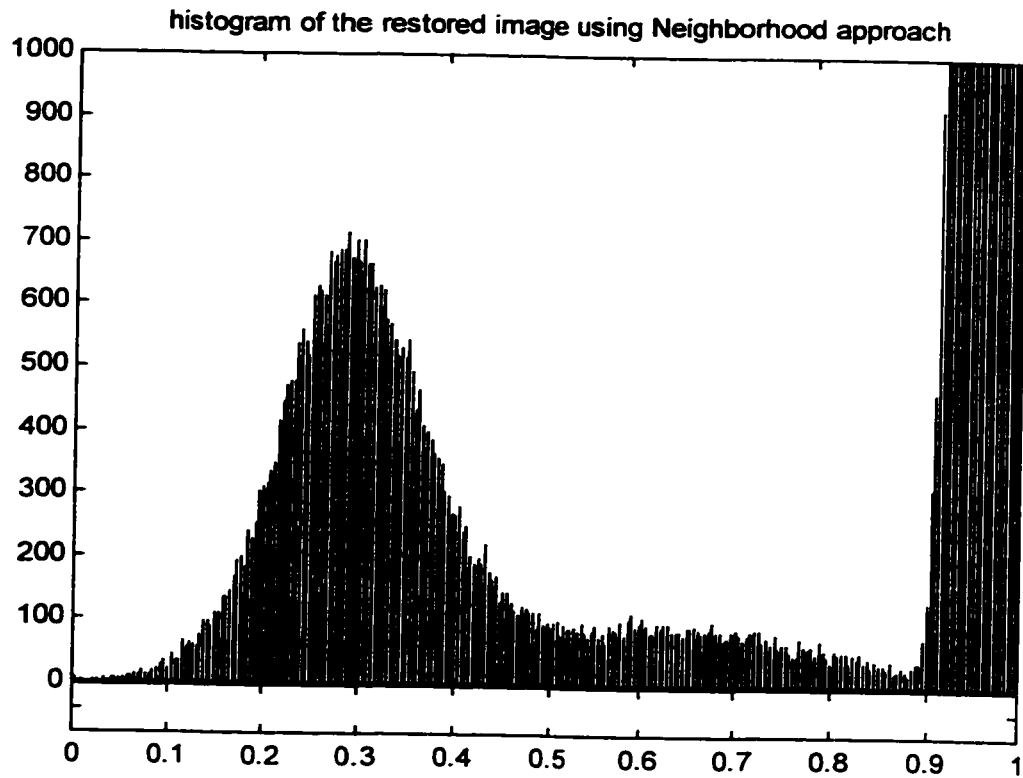
restored verso image using neighborhood approach



Around the world  
there is a treasure  
of excellent literature  
but not accessible  
to all the people.

**Figure 4. 19:** The restored *recto* image using neighborhood approach.

With the sample 2, based on the visual inspection the results are not as good, though the MSE is 0.0136. This implies that the technique is not applicable for all kind of documents. The histogram of the restored image is given in the Figure 4.20.



**Figure 4. 20:** The histogram of the restored image using Neighborhood approach

### 4.3.3 Comparing the ratio of Pixel Intensity

As described in section 3.4.3, this technique is based on the two possibilities when comparing the pixel intensities in the *recto* and flipped *verso*. This approach gives the most optimum results with both of the samples (*please refer to Figure 4.21 and 4.22*).

restored image using option I

Around the world  
there is a treasure  
of excellent literature  
but not accessible  
to all the people.

**Figure 4. 21:** The restored sample I using option I

The restored image

---

Oct 19, 1999.

Test sample. Recto

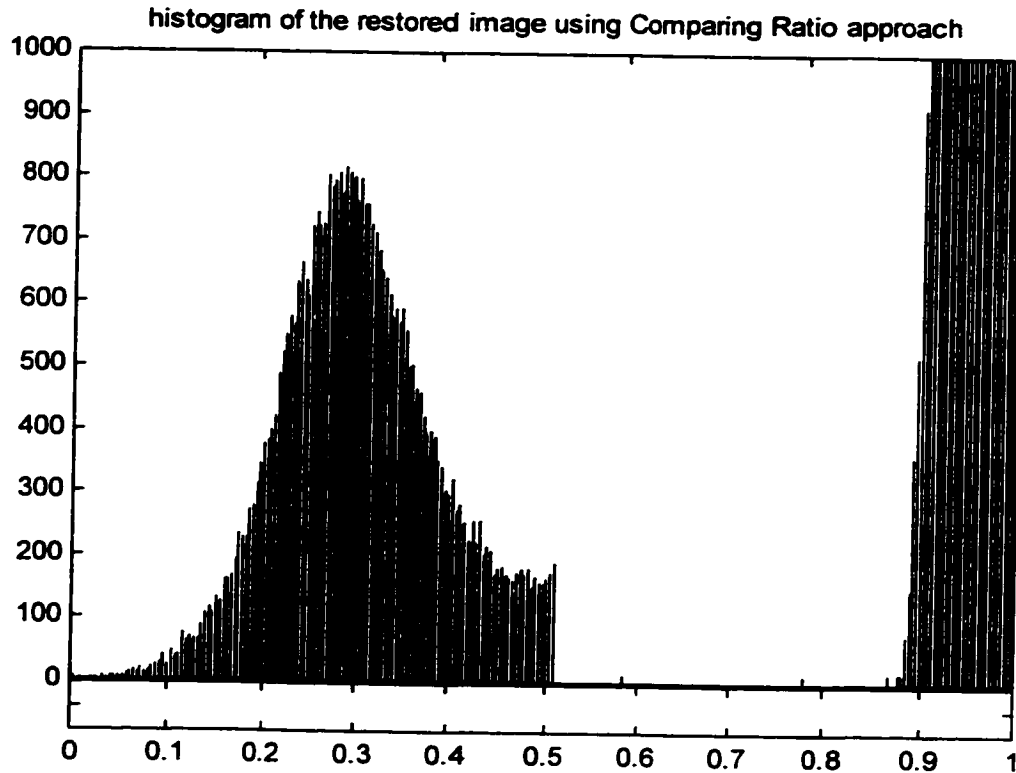
Image registration has few direct applications by itself, but it is a necessary step for computer analysis, tracking with multiple images, or motion analysis.

---

---

**Figure 4. 22:** The sample 1 restored using the comparison of intensity ratio

For the subjective comparison the histogram of the restored sample 2 is shown in Figure 4.23 which is similar to the histogram of the ideal image shown in Figure 4.3. The MSE calculated for both the samples are 0.0094 and 0.0178 respectively.



**Figure 4. 23:** The histogram of the restored image

#### 4.4 Discussion

The results of the restored images using three different approaches are given in the table 4.1. These results are obtained using sample 1 and sample 2. For both the samples the results are similar. As shown in the Table 4.1, the Mean Square error, MSE1, between the restored and ideal image in sample 1 and the Mean Square error, MSE2, form the sample 2.

S. No	Restoration Methods	Results	MSE1	MSE2	Time (sec)
1.	Comparing Pixel Intensity	Not good	0.0095	0.0170	10
2.	Neighborhood approach	Acceptable	0.0071	0.0136	134
3.	Comparing the ratio of Pixel Intensity	Very Good	0.0054	0.0092	12

**Table 4. 1:** Table for comparing the results of the restored image using three approaches

It is clear from the Table 4.1 that the third approach of comparing the ratio of pixel intensity gives the good result and minimum MSE with both the samples.

The time required in each restoration technique is shown in the last column of the Table 4.1. It is clear from the table, that neighborhood approach is the slowest approach when compared to the other two approaches. The first approach, of comparing the pixel intensity, is faster but does not give efficient result. The third approach, of comparing the pixel ratio, is fast and also gives the best result.

Although time comparison is not a critical factor in ascertaining the best approach because fast execution can be achieved using software other than Matlab.

## **Chapter 5**

## **Conclusion**

This concluding chapter outlines the results obtained from the research, and the conclusion drawn from this study. This chapter concludes with the future aspects and the work needed to be done to make the technique more appropriate for degraded documents restoration is suggested. The work presented in this thesis is applied on the gray images but in future the restoration can be applied to the color documents.

The objective of this thesis, to set up an approach for the restoration of the documents degraded due to the show-through, has been achieved. The technique developed here is efficient for the documents with writing, sketching, graphs or gray images.

The study was divided into two parts – the registration and the restoration technique. To test the validity of the registration techniques the objective test was applied, this generated the results for all the techniques discussed here. The most efficient technique was selected on the basis of visual inspection, computational requirements and the complexity of the method. Based on this analysis the affine transformation technique with the calculation of six parameters is able to adjust for all types of possible distortions during the scanning of the documents. This technique is much faster than the other three registration techniques discussed here and it also gives the most efficient registration required for the restoration of the documents. The registration effectiveness achieved by this technique is sufficient to compensate for the distortion generated while scanning the documents and therefore was applied for the restoration techniques.

The objective and subjective methods were applied to demonstrate the efficiency of each restoration technique. The final suggested restoration method, comparing the ratio of pixel intensity, was applied on different samples to prove the acceptance of the technique which should be used for the restoration of the documents. This approach resulted in visually better results than the other three approaches. The MSE calculated between the restored and the ideal image is lower and the pattern of the histogram of restored image is closest to the histogram of the corresponding ideal image. All these factors prove the efficacy of the restoration technique suggested here.

This study provides only the simple method for the restoration of degraded gray documents. This work excludes the color documents. The same comparison technique of restoration can be applied to the color documents. If the same algorithm of restoration is applied with the neighborhood operator, the result would be much better. Using neighborhood approach in MATLAB5.3 is very slow therefore some other environment, such as C, C++, can be used to achieve a faster restoration.

## References

1. Gonzalez, Rafael C. and Woods, Richard E., *Digital Image Processing*, Addison Wesley, 1998.
2. Castleman, Kenneth R., *Digital Image Processing*, Prentice Hall, New Jersey, 1996.
3. Rosenfeld, A and Kak, Avinash C., *Digital Picture Processing*, second edition, Academic Press, New York, 1982.
4. Gann, Robert G., *Desktop Scanner – Image Quality Evaluation*, Hewlett-Packard Professional books, Prentice Hall, 1999.
5. Jain, Anil K., *Fundamentals of Digital Image Processing*, Prentice Hall Englewood Cliffs, New Jersey, 1996.
6. Baxes, Gregory A., *Digital Image Processing*, John Wiley & Sons, Inc. USA, 1994.
7. Galbiati, Louis J. Jr., *Machine Vision and Digital Image Processing Fundamentals*, Prentice Hall, Englewood Cliffs, New Jersey, 1990.
8. Brown, Lisa G. , “Image registration Techniques”, *ACM Computing Survey*, Vol 24, No. 4, December 1992.
9. Lim, Jae S., *Two dimensional Signal and Image Processing*, Prentice Hall, Englewood Cliffs, New Jersey, 1990.
10. Andrews, H.C. and Hunt, B.C., *Digital image restoration*, Prentice-Hall, Englewood Cliffs, New Jersey, 1977.
11. Proakis, John G., *Digital Signal Processing: Principle Algorithms and Applications*, Prentice Hall, Englewood Cliffs, New Jersey, 1996.
12. M. D. Pratt, *Subpixel image registration of remotely sensed imagery*, *IBM J. Res. Development*, Vol. 38 No. 2, March 1994.
13. Dufax, Frédéric, and Konrad, Janusz, “Efficient, robust and fast global estimation for video coding”, *IEEE Transaction on Image Processing*, Vol. 9, No.3, pp 497 – 501, March 2000