Study on Potential of Binary Artificial Bee Colony (bABC) algorithm for Optimizing the Structuring Element towards Degraded Document Image Restoration

A Thesis Submitted in Partial Fulfillment for the Degree of Master of Technology in Printing Engineering and Graphic Communication

By

Sebanti Sengupta

Examination Roll No. M4PRI1601

Registration No.129436 of 14-15

Jadavpur University

Under the guidance of

Dr. Arpitam Chatterjee

Department of Printing Engineering

Faculty of Engineering and Technology

Jadavpur University

Kolkata-700032

JADAVPUR UNIVERSITY

FACULTY OF ENGINEERING AND TECHNOLOGY

Certificate of Recommendation

We hereby recommend that the thesis titled, **"Study on Potential of Binary Artificial Bee Colony (bABC) algorithm for Optimizing the Structuring Element towards Degraded Document Image Restoration**" has been carried out by **SEBANTI SENGUPTA** (Registration No. 129436 of 14-15) may be accepted in partial fulfillment of the requirement for the degree of "Master of Technology in Printing Engineering and Graphic Communication" of Jadavpur University.

……………………………………………..

Dr. Arpitam Chatterjee

Assistant Professor,

Printing Engineering Department,

Jadavpur University

……………………………………………..

Dr. Swati Bandyopadhyay

Head of the Department,

Printing Engineering Department,

Jadavpur University

………………………………………………..

Prof. Sivaji Bandyopadhyay

Dean,

Faculty of Engineering and Technology,

Jadavpur University

JADAVPUR UNIVERSITY

FACULTY OF ENGINEERING AND TECHNOLOGY

Certificate of Approval

The foregoing thesis is hereby approved as a credible study of an engineering subject, carried out and presented in a manner satisfactory to warrant its acceptance as a prerequisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusions drawn therein but approve the thesis only for the purpose for which it has been submitted.

……………………………………   **1**……………………………………..

Signature of the Examiner(s)   **Signature of the Supervisors**

**Only in case the thesis is approved*

JADAVPUR UNIVERSITY

Declaration of Originality and Compliance of Academic Ethics

I hereby declare that this thesis contains literature survey and original research work by the undersigned candidate, as part of the Master of Technology in Printing Engineering and Graphic Communication studies.

All information in this document has been obtained and presented in accordance with academic rules and ethical conduct.

I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name	:	**SEBANTI SENGUPTA**
Examination Roll No	:	**M4PRI1601**
Registration No.	:	**129436 of 14-15**
Thesis Title	:	**Study on Potential of Binary Artificial Bee Colony (BABC) algorithm for Optimizing the Structuring Element towards Degraded Document Image Restoration**
Signature with Date	:	

ACKNOWLEDGEMENTS

I would like to take this opportunity, to humbly, express my gratitude for the innumerable gesture of help, cooperation and inspiration that I have received from my teachers and well wishers during this thesis work. This project is far one among the most significant accomplishments in my life and it would not have been possible without them.

I feel honored to express my profound regard and deep sense of gratitude to my project guide **Dr. Arpitam Chatterjee,** Assistant Professor, Printing Engineering Department, Kolkata for his invaluable help and inspiring guidance that facilitated me to carry out this work. I am highly obliged and grateful to him for giving me the opportunity of the project work.

I would like to thank **Dr. Swati Bandyopadhyay,** Head of the Department, Printing Engineering, Jadavpur University for providing the necessary facilities for carrying out this thesis. My sincere thanks to all the faculty and staff in Printing Engineering Department for their valuable support.

I take this opportunity to express my gratitude to my friends and colleagues who had been an inspiration and reliable source throughout.

I wish to thank all my family members especially my father and mother for their continuous moral support without which this accomplishment would not have been possible. Last but not least I would like to thank almighty God for his blessings.

Sebanti Sengupta

Abstract

Resorting to restoration techniques for heritage documents becomes an increasingly urgent need. In fact, these valuable resources for human being are subject to several types of degradations limiting their use. A proposed solution to this problem is the application of restoration techniques on the digital copy of the degraded original document. This would improve human readability and allow further application of image processing techniques. Hence identifying a typology of different types of image degradation is of primary concern for use in restoration techniques. In this thesis, we have used binary Artificial Bee Colony (bABC) algorithm for optimizing the structuring element for restoration of degraded document images. The original ABC algorithm is mainly used for problems in continuous domain, however many engineering practices have been described as a combinatorial optimization problem. Binary artificial bee colony (bABC) algorithm can be used to optimize the 0-1 integer programming problem. For the better optimization, many optimization algorithms combination has become a hotspot which can produce a new hybrid optimization algorithm. Such algorithm can be useful to obtain optimized structuring element which is of binary nature to produce targeted goal when the objective function is properly designed as per problem under consideration.

In this thesis, at first an extensive review on different document restoration methods was conducted. The techniques reveal the scope of employing search based optimization techniques for the stated task. This thesis explores the scope of bABC restoration purpose by means of optimizing the structuring element. Randomly generated structuring elements were applied in document images using different parameters and were compared. The properties and behavior of the results were critically studied. Results have shown that bABC algorithm can be used for generating optimized structuring element towards degraded document image restoration, and thus found to be a suitable alternative method of restoration. Simulation results using psychovisual technique showed that the proposed bABC method is capable of restoring degraded document images.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1. Introduction

Millions of valuable historical documents that are stored on the shelves of libraries all over the world are ready to facilitate access to the information they contain. The first major problem faced is degradation, which renders the visual quality of the document very poor, and in most cases, difficult to decipher. Depending on the degree of degradation, the problem of removing them could be resolved by image enhancement and restoration [1]. Processing of documents to extract their content in an automated fashion is an essential task in all types of organizations for varied applications. Literature survey reveals that the documents under processing are subjected to the following steps:

1) The Pre-Processing Stage that enhances the quality of the input image.

2) The feature extraction stage that captures the distinctive characteristics of the document under processing.

3) The classification stage that identifies the document; accordingly group to certain classes & helps in their efficient recognition [2].

Document images are traditionally acquired by flat-bed scanners for optical character recognition (OCR) processing. The prevalence of digital cameras in recent years makes it much more common to shoot document images rather than to scan them. Such trend brings about big challenges to traditional OCR techniques. Unlike scanned document images that are acquired in flat-bed style with good lighting, document images are usually distorted due to non planar page shape or perspective distortion with uneven lighting [3]. Generally, traditional OCR applications cannot handle such deterioration. A document image will firstly be binarized and segmented by OCR applications in order to extract text regions; then characters will be separated for recognition. Uneven lighting often bring difficulties for image binarization because multiple binarization thresholds need to be identified for different image pixels to adapt different luminance on the page; otherwise, noise might arise on the binarized image.

There are several approaches that can solve binarization problem [4]. In this thesis, we have used binary Artificial Bee Colony (bABC) algorithm for optimizing the structuring element for restoration of degraded document images. The original ABC algorithm is mainly used for continuous optimization, however many engineering practices have been described as a combinatorial optimization problem. Binary artificial bee colony (bABC) algorithm can be used to optimize the 0-1 integer programming problem. Artificial bee colony (ABC) optimization is a relatively new population-based, stochastic optimization technique. ABC was developed to optimize unconstrained problems within continuous-valued domains. The performances of these binary ABC algorithms are compared on a benchmark of unconstrained optimization problems. For the better optimization, many optimization algorithms combination has become a hotspot which can produce a new hybrid optimization algorithm. In ABC algorithm, artificial bees consist of three roles: employed bees, onlookers and scouts. Employed bees correspond with the specific food sources. Their tasks are to forage in the corresponding food sources. Onlookers watch the dance of employed bees within the hive to choose the excellence food source. Scouts search for food sources randomly [5].

In this thesis, at first an extensive review on document restoration using bABC method was conducted. Then, randomly generated structuring elements were applied in document images using different parameters and were then compared. The properties and behavior of the results were critically studied. Results have shown that bABC algorithm can be used for generating optimized structuring element towards degraded document image restoration, and thus found to be a suitable alternative method of restoration. The result of the presented method was psychovisually compared with some of the standard techniques and was proven to be a suitable method of restoration.

1.2. Literature Survey

Author(s)	Year	Paper Name	Technique	Results
Changsong Liu · Yu Zhang · Baokang Wang ·Xiaoqing Ding [6]	2014	Restoring camera-captured distorted document images	A segmentation-based approach and TPS (Thin Plate Spline) algorithm are introduced in order to restore distorted document image. In order to be applicable in arbitrary warping cases, local shape of different positions on an image must be estimated with	It proves its effectiveness on arbitrary warping cases, and the processing speed is acceptable.

			least constraint. Therefore, baselines and vertical directions are estimated locally to reflect local shape. TPS interpolation method is used in order to keep neighborhood relationships.	
Jayant Kumar, Peng Ye and David Doermann [7]	2013	A dataset for quality assessment of camera captured document images	The dataset will be useful to researchers working on the purposive evaluation of quality estimation methods for predicting the OCR (Optical Character Recognition) quality of document images. The dataset has a total of 525 (175 _ 3) OCR-text files from three popular OCR engines. Furthermore, it also obtained character level accuracy for each OCR-text file.	It presented results of three recent methods on estimating the OCR quality of images based on output obtained from Fine Reader. Using two different evaluation measures it has been compared and also discussed the advantages of three quality estimation approaches.

O. Imocha Singh, *et al.* [8]	2012	Local contrast and mean based thresholding technique in image binarization	Image binarization using local minimum and maximum pixel value within a local window $w \times w$.	It finds a new way of Image binarization using a local threshold value and their comparison with other algorithms.
Xujun Peng, Huaigu Cao, Krishna Subramanian, Rohit Prasad, and Prem Natarajan [9]	2011	Automated Image Quality Assessment For Camera-Captured OCR	An OCR based document image quality assessment algorithm that can evaluate the degradation degree of the camera-captured document images according to the statistic features.	Experimental results show that this system has high regression accuracy and achieves the reliable classification results.
Chien-Hsing Chou *et al.* [10]	2010	A binarization method with learning built rules for document images produced by cameras.	Document images with non-uniform brightness require binarization methods with delicate local thresholds that must be determined according to various conditions. For this purpose, a region-based binarization method is established. Binarization based on information provided by a	It produces favorable results with respect to the thresholding actions judged in terms of the visual quality of images, and also in terms of the OCR performance.

			region is effective and robust, provided SVM method is used to construct decision functions from the information provided by training samples.	
Shijian Lu, Ben M. Chen *, C.C. Ko [11]	2006	A partition approach for the restoration of camera images of planar and curled document	A computationally efficient technique to restore the camera images of document lying over a planar or smoothly curved surface. The restoration is carried out through image partition, which divides camera documents into multiple quadrilateral patches through the exploitation of the identified VSBs(Vertical Stroke Boundary)and the fitted x lines and base lines of text. Compared with the reported document	The restoration process is fast and easy for implementation. The document restoration technique may open a new channel for document capture and understanding. Furthermore, it may be applied to some other portable devices such as the mobile phone and the personal digital assistant (PDA) with the increase of sensor resolution. As a result, these camera-sensor embedded devices may need only to store and transmit recognized ASCII

			restoration techniques, this method needs no auxiliary hardware, no camera calibration, no complicated 3D reconstruction and the only thing required is a single document image captured by a common digital camera.	text instead of the document images with a huge size.
Sarika Jain, Pankaj Parihar [12]	2015	Remove noise and reduce blurry effect from degraded document images	In this method, Wiener Filter is used to remove noise from the images. It also includes histogram equalization and de-blurring techniques.	Various parameters are evaluated in this algorithm for studying percentage of improvement and calculation of execution time for the final output.
Rupinder Kaur, Jaspreet Kaur [13]	2014	A novel image restoration algorithm for digitized degraded historical documents	The aim is to segment the text from the degraded document images. That means the foreground and background should be separated. The resultant image will be having the background as	Objects inside the boundaries will be extracted and then the feature enhancement will be performed on the extracted text. The boundary removal will be applied on the image before or after the feature

			white and text as black. The basic and most important part for this procedure is to determine a threshold for binarization.	extraction, wherever it will produce the best results. The performance parameters considered will be Elapsed Time, Peak Signal to Noise Ratio (PSNR), Normalized Absolute Error (NAE) and Mean Square Error (MSE) and Root Mean Square Error (RMSE).

1.3. Scope of the thesis

In this thesis, research has been carried out to study various restoration techniques of degraded document images. The results of each technique have been studied rigorously and further analysis was carried out. Focus has been given to achieve the best solution of restoration. A binary artificial bee colony optimization algorithm has been realized to optimize the structuring element for restoration. The algorithm has been experimentally implemented. The scope may be specifically stated as below:

1. Formulating objective function including various parameters used for restoration.
2. Optimizing structuring element towards generating the best solution.

1.4. Chapter Outline

Chapter 1 provides the introduction and scope of the thesis. The chapter defines preliminary concepts related to document restoration techniques in literature survey section. Finally it introduces briefly about the concept of binary artificial bee colony optimization technique.

Chapter 2 provides an extensive study on the existing techniques. The methods are implemented and the results of those established techniques are shown here which were used for comparison in later chapters.

Chapter 3 provides the concept of binary artificial bee colony optimization algorithm. Survey on different swarm intelligence techniques are carried out. The results of bABC are obtained.

Chapter 4 provides the objective function formulation for this thesis. Different parameters are applied to the algorithm according to the behavior of the experimental results few of the parameters are chosen as an objective function.

Chapter 5 contains the theoretical analysis of the binary artificial bee colony optimization technique. Simulation results show that the proposed technique is capable of finding the best optimized structuring element for restoration of the document.

Chapter 6 contains the conclusion of the thesis. Major findings and future scope of the work are also presented.

1.5. References

[1] Rachid Hedjam, Mohamed Cheriet, "Historical document image restoration using multispectral imaging system," in International Conference on Pattern Recognition (ICPR), Vol 46, Issue 8, August 2013, pp 2297–2312.

[2] Shazia Akram, Dr. Mehraj-Ud-Din Dar, and Aasia Quyoum, "Document Image Processing - A Review," in International Journal of Computer Applications (pp. 0975 – 8887) Volume 10– No.5, November 2010.

[3] J. Kumar, F. Chen, and D. Doermann, "Sharpness estimation of document and scene images," in International Conference on Pattern Recognition (ICPR), 2012, pp. 3292–3295.

[4] F. Chen, S. Carter, L. Denoue, and J. Kumar, "SmartDCap: semiautomatic capture of higher quality document images from a smartphone," in International conference on Intelligent user interfaces (IUI), 2013, pp. 287–296.

[5] Dervis Karaboga, Beyza Gorkemli, Celal Ozturk, Nurhan Karaboga, "A comprehensive survey: artificial bee colony (ABC) algorithm and applications," in Artif Intell Rev, 11 March 2012, pp. 37.

[6] "Restoring camera-captured distorted document images", Changsong Liu · Yu Zhang · Baokang Wang Xiaoqing Ding, IJDAR DOI 10.1007/s10032-014-0233-8, pp. 2-9 (2014).

[7] "A Dataset for Quality Assessment of Camera Captured Document Images ", Jayant Kumar, Peng Ye and David Doermann, University of Maryland, College Park, USA,pp. 2-3 (2013).

[8] "Local Contrast and Mean based Thresholding Technique in Image Binarization", O. Imocha Singh, *et al.*, International Journal of Computer Applications (0975 – 8887) Volume 51– No.6, August 2012.

[9] "Automated Image Quality Assessment For Camera-Captured OCR", Xujun Peng, Huaigu Cao, Krishna Subramanian, Rohit Prasad, and Prem Natarajan,18th IEEE International Conference on Image Processing, pp. 2-4, 2011.

[10] "A binarization method with learning built rules for document images produced by cameras", Chien-Hsing Chou *et al.,* Pattern Recognition 43 (2010) pp. 1518–1530.

[11] "A partition approach for the restoration of camera images of planar and curled document", Shijian Lu, Ben M. Chen *, C.C. Ko, S. Lu *et al.* / Image and Vision Computing 24 (2006) 837–848.

[12] "Remove Noise and Reduce Blurry Effect From Degraded Document Images", Sarika Jain and Pankaj Parihar, International Journal of Engineering Research and General Science Volume 3, Issue 1, January February, 2015, ISSN 2091-2730.

[13] "A Novel image restoration algorithm for digitized degraded historical documents", Rupinder Kaur and Jaspreet Kaur, International Journal of Science, Engineering and Technology Research, Volume 3, Issue 9, September 2014, pp. 2275-2277.

CHAPTER 2

REVIEW OF ESTABLISHED TECHNIQUES

2.1. Introduction

Document images that are degraded can be restored by different techniques. Different techniques have been studied. The documents that are degraded can be captured by digital camera or can also be scanned by scanner. Such trend brings about big challenges to traditional OCR techniques. Unlike scanned document images that are acquired in flat-bed style with good lighting, document images are usually distorted due to non planar page shape or perspective distortion with uneven lighting [1].

Generally, traditional OCR applications cannot handle such deterioration. A document image will firstly be binarized and segmented by OCR applications in order to extract text regions; then characters will be separated for recognition. Uneven lighting often brings about difficulties for image binarization because multiple binarization thresholds need to be identified for different image pixels to adapt different luminance on the page; otherwise, noise might arise on the binarized image. There are several approaches that solve binarization problem with uneven lighting. Distortion reduces accuracy of the segmentation process. In traditional cases, texts in flat documents are well aligned horizontally or vertically. Therefore, it is reasonable to use methods based on projection to perform segmentation because the projection's peak-and-valley shape can effectively reflect the distribution of text lines and characters in certain direction. In distorted cases, however, text lines are usually slant or curved. As a result, horizontal or vertical projection would no longer provide effective evidence and the traditional segmentation methods may result in wrongly segmented lines and characters [2].

The segmentation may be incorrect; character recognition would not probably have high accuracy. It mainly focuses on the problem of distortion in document images. The approaches try to recover the flat distorted document. Therefore, the restored image can be segmented and recognized by traditional applications.

With the increasing quality of cameras on mobile devices, imaging document pages as an alternative to scanning is becoming more feasible. However, document images may suffer from degradations arising from the image acquisition process. One of the most frequently occurring distortions that affect captured image quality is blur. When taking a photo, there are different causes of blur. Figure 1 shows examples of (a) out-of-focus blur (b) blur due to the motion of camera, and (c) blur due to limited depth of field which occurs when content is at different distances. This is especially apparent in close ups and with imaging devices that have a large aperture. Small, high-resolution cameras in smart phones are more susceptible to these distortions due to their relatively large apertures, and their light-weight and single hand usage, which make them difficult to hold steady [4]. In the presence of such distortions, the ability to automatically assess the quality of captured images is also becoming increasingly desirable. The required quality of a document image is usually constrained by the applications and usually with respect to human perception or machine readability.

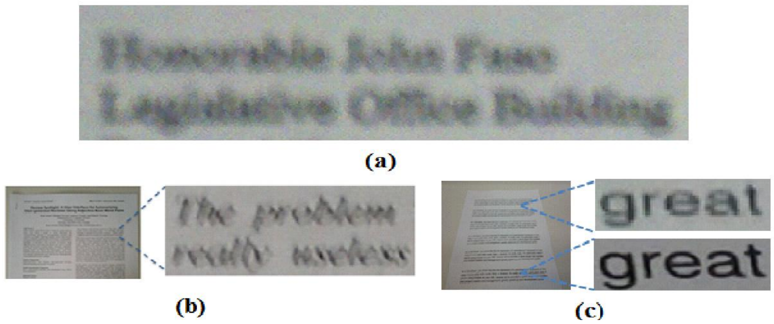


Fig. 1: (a) Out-of-focus blur (b) Motion-blur caused by hand-shake (c) Blur due to limited depth of field when content (characters) are at different distances [4].

2.2. Some of the established methods

There were several methods as shown in chapter 1. This section provides brief descriptions about those techniques.

2.2.1. Restoration of Camera-Captured Distorted Document Images using TPS (Thin Plate Spline) Algorithm:

It presents a segmentation-based approach. The approach performs text line segmentation first. Text lines' curve is the most obvious evidence to estimate a document image's warping shape. Therefore, line segmentation step is useful for shape estimation step. In order to process more complex cases, a cluster algorithm is performed in order to classify connected components into text lines correctly. After line segmentation, the restoration problem can be regarded as estimating baselines of all text lines and straightening them. Besides baselines, this approach also estimates and rectifies slant of characters.

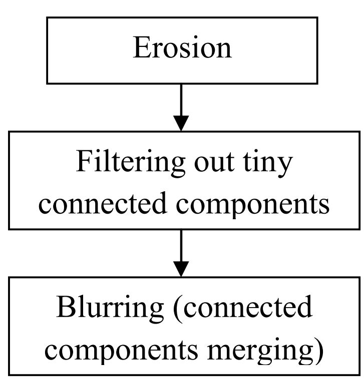


Fig. 2: Flowchart of preprocessing for connected components [5]

A global interpolation method is adopted in restoration step. Such method avoids the problem of grid piecing in others' work. The restoration result is then naturally smooth and continual. After distortion shape is estimated, the image should be restored. A thin-plate spline (TPS) method is used to perform interpolation. Thin-plate spline [6] is an interpolation method to calculate coordinate mappings from R2 to R2. Given several key points' mapping from original image to final image, TPS algorithm can give all points' new coordinate in final image in the same time keeping the given key points' mapping relationship.

TPS interpolation assumes that the image is a soft "thin plate," which means the moving of a point will bend all the plate as one pulls a point of a piece of cloth. Therefore, TPS algorithm is suitable to deal with warped images if enough key points' new and original coordinates are confirmed. In addition, because it is an interpolation algorithm on the whole image, TPS can keep the neighborhood relationship of the original image in the new image. It is important to determine key points and their mapping for TPS. Once they are determined, calculate a pixel mapping relationship between the original image and the new image was calculated, and then each pixel's gray level in the new image was used to get an interpolation result. Such key points are on baselines and mean lines in distorted document images. Their mapping are the position where they should been in if the document were flat [5].

2.2.2. A Dataset for Quality Assessment of Camera Captured Document Images:

It deals with the approaches on the creation of dataset for quality assessment of document images. More specifically it focus on datasets for estimating OCR quality of document images. Many datasets for assessing the quality of scanned document images have been discussed. One of the early works on predicting OCR accuracy was done by Blando *et. al.* [7]. It used two sets of test data in the experiments. The first set was a subset of ISRI's Sample 2 database [8] consisting of 460 pages. Each page was digitized at 300 dpi using a Fujitsu M3096M+ scanner. The second set consisted of 200 pages selected from 100 magazines that had the largest paid circulation in the U.S. in 1992. For each magazine, they selected two pages at random and each page was digitized (300 dpi) using a Fujitsu M3096G scanner. The images were binarized using a fixed threshold of 127 out of 255. They used a total of six OCR systems for processing their data sets and collected character accuracy for each image. In their evaluation, each character insertion, deletion, or substitution required to correct the generated OCR text was counted as an error. The character accuracy in their work is defined as equation (1):

$$CharacterAccuracy = \frac{n - number\,of\ errors}{n} \qquad (1)$$

where, n is the total number of characters in the ground-truth text [8].

Cannon *et. al.* [9] focused on the quality of type-written document images and applied it for selecting the optimal restoration approach. It used five quality measures that assess the severity of background speckle, touching characters, and broken characters. It used a dataset of 139 document images with 300 dpi resolution. Omni Page Pro v8.0 was used to perform OCR and the character error rate of the corpus was found to be 20.27%. It further formed a sub-corpus of 41 documents having OCR character error rates between 20% and 50% to perform analysis on highly degraded images. It also created a small corpus of documents spanning a range of gradually decreasing quality by repeatedly photocopying a page from a book (a total of 9 versions). Each successive copy was degraded with background speckle, widened stroke widths, touching characters and other common attributes of lower quality document images [10].

2.2.3. Local Contrast and Mean-based Thresholding Technique in Image Binarization:

Thresholding takes a major role in binarization. Determination of thresholding is major issue in binarization. Based on the technique of determination of threshold there are many techniques of binarization. It can be approached broadly in two ways like global and local technique. In this method only local techniques are considered. Local technique is normally local window block size dependent both in result and time complexity. Many techniques which associates local mean use integral image as a prior process to determine the local mean. It reduces time complexity. The technique of TR Singh *et. al.* [11] uses local mean only while that of Sauvola *et al.*'s [12] and Niblack's [13] technique use standard deviation also. Bernsen [14] uses local contrast pixel values to determine the threshold value. Local contrast is the difference of the local maximum and minimum pixel values. This technique of binarization is carried out with a local thresholding technique which uses local contrast and mean. It is expressed as below equation (2):

$$T(x,y)=k\left[m(x,y)+\left(I_{\max}-I_{\min}\right)\left(1-I(x,y)\right)\right] \quad (2)$$

where, $k \in (0,1)$ is a bias constant, $m(x,y)$ is local mean, and $I_{\max}$ $I_{\min}$ are the local maximum and minimum pixel values within the local window of size $w \times w$ and $I(x,y)$ is the concerned pixel value.

In this technique local maximum and local minimum pixel values are associated. As a result of using local minimum and maximum this technique is not suitable for using integral image which is suitable for finding local mean with no time. Hence it works like the other local techniques whose computational time is local window size dependent [15].

2.2.4. Automated Image Quality Assessment For Camera-Captured OCR:

In this approach Regression and Classification methods are used. To evaluate the quality of the document images according to their features, an OCR based method which predicts the N-WER (Normalized Word Error Rate which is defined by equation (3) of each document image where high WER indicates the low image quality. Assuming that there are n high quality document images and a total number of m degraded document images which are generated from high quality images in training data set X, each sample image is associated with a N-WER ϵ_i, which measures the document image's quality. Given the feature $f_i = [\bar{D}, \bar{R}]^T$ of a document image I_i, the goal of quality assessment is to find a hypothesis $H: f_i \rightarrow \epsilon_i$, which maps the feature f_i of the document image to an N-WER ϵ_i. Firstly, the N-WER for each document image on both training set and testing set is calculated in equation (3):

$$\epsilon_i = \frac{wer_i - wer_j}{1 - wer_j} \quad (i \in [1 \dots m], j \in [1 \dots n]) \qquad (3)$$

where, wer_i is the WER for degraded image I_i and wer_j is the WER for its corresponding original high quality document image I_j. The N-WER uses the WER of the original high quality document as the baseline and measures the increase of the WER which is caused by image degradations. Unlike other assessment methods that categorize the degraded images as several groups (typically five groups with bad, poor, fair, good and excellent quality) and classify the test image as one of these groups, SVR (Support Vector Regression) is used to calculate the N-WER of each document and take different actions (accept/reject) based on N-WER. In the training phase, considering the problem of approximating the data set $X = \{(f_1, \epsilon_1), \dots, (f_m, \epsilon_m)\}$, with a linear function $H(f) = < w, f > + b$, the optimal SVR function is calculated by minimizing the function as in equation (4):

$$\phi(w, \xi) = \frac{1}{2} \, || \, w \, ||^2 + \, C \, \sum_i (\xi_i^- + \, \xi_i^+) \qquad (4)$$

where, C is a pre-defined value, and ξ^-, ξ^+ are slack variables representing upper and lower constraints on the outputs of the system [16]. The non-linear regression of SVR is simply carried out by mapping the data into a high dimensional space where linear regression can be performed. In testing phase, an N-WER $€_i$ for each test document image is predicted by using the trained SVR (Support Vector Regression) [17].

2.2.5. A Binarization method with learning built rules for document images produced by cameras:

The approach involves the computation of thresholds using Otsu's method. Given gray scale image, Otsu's method is applied to a histogram of the image to set a threshold $thrl$ as equation (5):

$$thrl = argmax \sum_{i=1}^{2} \pi_i \, (v) (\mu_i(v) - \mu_T)^2, \qquad (5)$$

where, $\pi_1(v) = \Sigma_{m \le v} p_{m,} \mu_1(v) = \Sigma_{m \le v} m p_{m,} \pi_2(v) = \Sigma_{m > v} p_{m,} \mu_2(v) = \Sigma_{m > v} m p_{m,}$ $\mu_T = \sum_{m=1}^{256} m p_m$ and p_m is the proportion of pixels in the image whose gray scale is m for $m = 1,2, \dots, 256$. It then designates a pixel p as black, if p's gray scale is less than or equal to$thrl$; or white otherwise.

When the background and foreground intensities are well separated, Otsu's method [18] yields good binarized results. However, if the image intensities are inseparable, the resulting threshold value is unsuitable. An immediate solution is to divide an image into several regions and apply a thresholding method to each region separately. However, the image should not be divided according to the layout structure of the document, since the whole image falls within the same text region of the layout structure, while the brightness varies extensively over the region. Instead, the image is divided into equal-sized regions [19].

2.2.6. A partition approach for the restoration of camera images of planar and curled document:

This technique restores the camera image of planar documents through the exploitation of the vertical stroke boundary (VSB) and the x lines and base lines of text as labeled with (1) and (2) in Fig. 3. In this method the situations where text line shape can be modeled using a cubic polynomial curve is been focused. The advantages of this technique are it needs no auxiliary hardware, no camera calibration, and the only thing required is a single document image captured using a common digital camera.

This method restores camera documents through the image partition, which divides camera images into multiple quadrilateral patches where text can be approximated to lie on a planar surface. The document partition is implemented through the exploitation of the VSBs and the x lines and base lines of text [20]. For each partitioned patch, a target rectangle is constructed where rectangle height is commonly determined as the average size of the identified VSBs and the width is estimated through a character categorization process that classifies characters into different categories with different aspect ratios. With the partitioned image patches and the constructed target rectangles, global distortion are finally removed through the local rectification of partitioned image patches one by one.

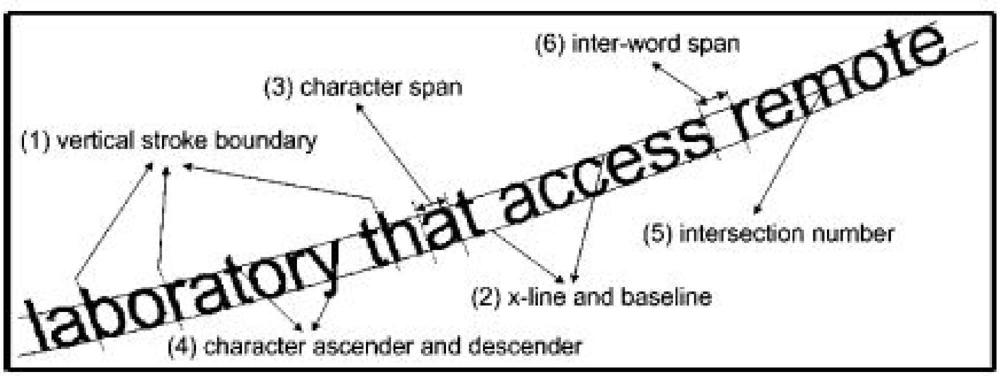


Fig 3: Text line definitions [21].

The algorithm assumes that document text is scanned horizontally and so the flat part of text lines lies on a horizontal straight line. Therefore, it cannot handle the camera images of documents lying over a smoothly curved surface. It requires that vertical text direction should be roughly parallel to the surface ruling and so it cannot restore documents curled in arbitrary direction [21].

2.2.7. Remove Noise and Reduce Blurry Effect From Degraded Document Images:

The segmentation of text from badly degraded document images in a very challenging task due to the high inter/intravariation between the document background and the foreground text of different document images. A novel document image binarization technique that addresses these issues by using adaptive image contrast in reported in [22]. The adaptive image contrast is a combination of the local image contrast and the local image gradient that is tolerant to text and background variation which are caused by different types of document degradations. This methodology can be used for efficient filtering of historical and degraded document images. It consists of many different steps as stated below;

The steps of the algorithm [22] are as follows:

1. Image Acquisition.
2. Applying FFT Algorithm
3. Converting input image into uint(8)
4. Applying Wiener Filter for Denoising
5. Getting the Output Image.

2.2.8. A Novel image restoration algorithm for digitized degraded historical documents:

Historical documents are often degraded by ink bleed through, stains, smudge, smear, cracks, watermarks, dust marks etc. These various types of degradations make it difficult to read the text. As historical documents contain very useful information, this data has to be preserved. The degraded historical documents are available in higher amounts. Archeologist need to restore the historical documents manually or digitally. It is very time consuming task to rewrite the larger number of documents. In the base paper, authors have used Particle Swarm Optimization (PSO) algorithm for image quality enhancements.

The aim is to segment the text from the degraded document images. That means the foreground and background should be separated. For this purpose a method known as binarization is been used. The resultant image will be having the background as white and text as black. This method is helpful because now the text is more legible and also it requires very less memory for storage. The basic and most important part for this procedure is to determine a threshold for binarization. Various steps followed in [23] are;

1) Image to grayscale conversion,

2) A Gaussian filter to remove the noise,

3) Image dilation to estimate the background,

4) Estimated background subtraction from grayscale image,

5) Global thresholding (OTSU) for modification,

6) Sauvola *et al.* threshold for local area or window,

7) Calculate parameters (PSNR, MSE, RMSE, and NAE),

8) Comparison with existing technique.

2.3. Results of established method

The techniques presented in Section 2.2 were applied and the results are noted down. This section contains a brief of the results.

2.3.1. Restoration of camera-captured distorted document Images using TPS (Thin Plate Spline) Algorithm:

For the basic evaluation, 23 English document images are selected and classified into different categories according to the degree they are distorted. Seven images are greatly distorted, while six images just have slight distortion on page borders. There are also five images with planar perspective distortion and five images that are basically flat for comparison. The below table shows the statistical results of the recognition rate. For samples that are greatly distorted or with planar distortion, the algorithm acquires large enhancement on recognition rate; for those just distorted slightly, the algorithm can rectify distortions on the edge and because characters there are just small part of a whole page, the enhancement is not that big; and for flat document images that can be processed well by traditional OCR applications, the algorithm will not bring about too much influence. As a whole, the restoration will bring positive effect to camera-based OCR shown in Table 2.1 below.

Table 2.1: Comparison of recognition rate before and after restoration [5].

Sample Type	No. of images	No. of characters	R.R. before restoration (%)	R.R. after restoration (%)
Greatly distored	7	14,730	81.30	97.39
Slightly distored	6	6,298	90.33	91.76
Planar distored	5	4,384	81.30	96.44
Flat	5	3,983	99.27	99.02
Total	23	29,395	85.67	96.27

2.3.2. A Dataset for Quality Assessment of Camera Captured Document Images:

Figure 4 summarizes the results of Spearman rank correlation (SROCC) values for 25 sets in test data. The SROCC for each set using the scores is computed by each method against the OCR accuracy. The top of the bars in figure 6 indicate observation median and the line segments represent the 75th and 25th percentile.

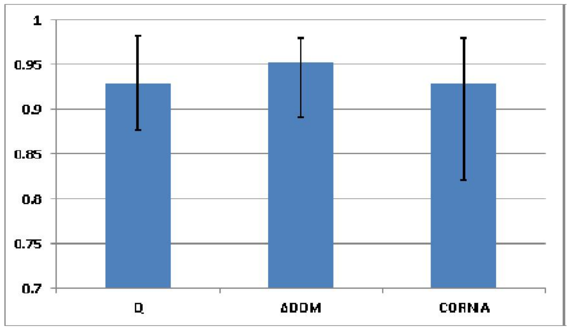


Fig 4: Median Spearman rank correlation for 25 sets in the dataset. The upper and lower end of line segments represent the 75th and 25th percentile respectively [10].

Of the three methods, ΔDoM performed consistently well on all the sets, while Q and CORNIA showed relatively higher variation in results on different sets. A higher SROCC value indicates the method's ability to rank images for a particular document, and can be used to select the image with best OCR accuracy.

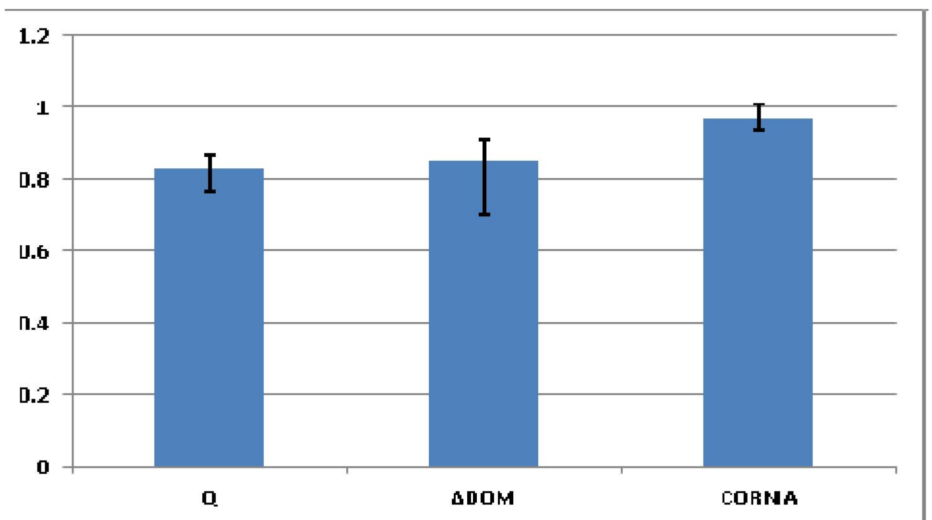


Fig. 5. Median Pearson linear correlation for 25 sets in the dataset. The upper and lower end of line segments represent the 75th and 25th percentile respectively [10].

Figure 5 shows the box-plot for Pearson correlation scores for 25 sets. Similar to previous plot, the bar shows the 75th and 25th percentile of scores. A good correlation score is needed for applications such as determining whether a captured image is good enough to keep or should be retaken. CORNIA performed better than other two approaches on modeling the linear relationship between two variables. When the goal of quality estimation is to predict the true quality score of images with different underlying content, CORNIA (or other supervised methods) usually outperforms unsupervised approaches [10].

2.3.3. Local Contrast and Mean-based Thresholding Technique in Image Binarization:

In this technique image binarization by adaptive thresholding technique using local contrast and mean is tested comparing with other techniques on many categories of images like medical image, scanned document images, document images and non document images.

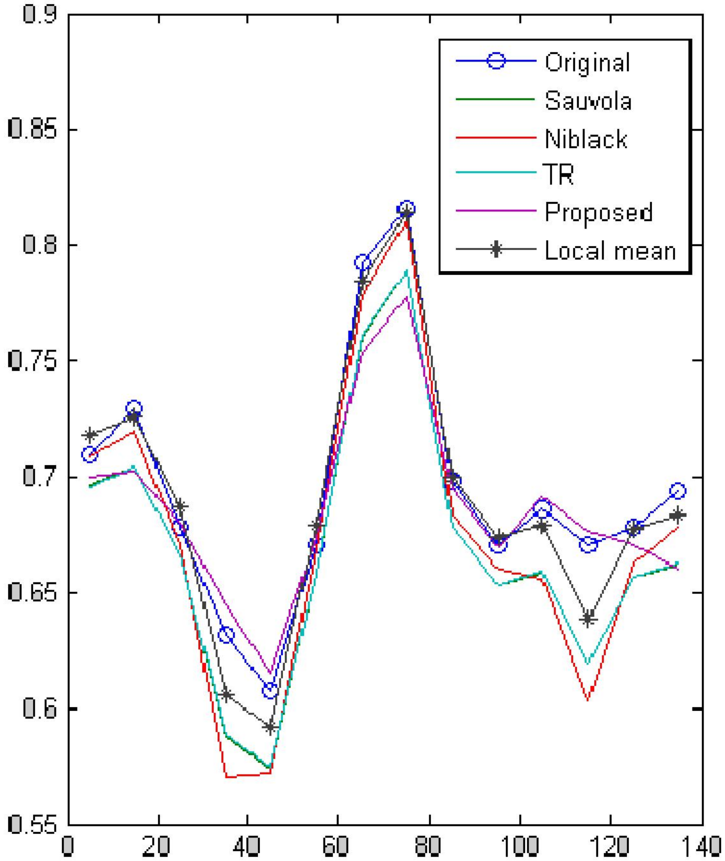


Fig 6: Nature of threshold values of different techniques with a window of size w=5 [15].

Figure 6 shows the graphical representation of different levels of the threshold values of the different techniques. From the experimental result it is observed that TR Singh and Sauvola *et al.* give similar result. But their results depend on the local window size. Sauvola *et al.* requires a window of size w=15 to get the high quality result while Niblack required a window size w=31 for a quality result. Even the widow size is less, the technique give better result for large character size while the other detects only the character boundary. This technique can give better result even at a small window size say w=3. Since the other local technique remove the background, continuous contrast image, cannot be binarized while this technique can binarize efficiently like global one [15].

2.3.4. Automated Image Quality Assessment For Camera-Captured OCR:

Prior to feature extraction and training, the registration and rectification procedure were taken for each degraded document image to avoid the WER increase from geometric distortion. The WER was calculated using BBN's OCR engine for each document and normalized. The average WER and normalized WER for the original high quality document images on training set (DEV) and its corresponding four degraded subsets (DEV1 - DEV4) are shown in table 2.2.

Table 2.2: WER, N-WER and SNR of training data sets [17].

	DEV	DEV1	DEV2	DEV3	DEV4
WER (wer)	0.270	0.367	0.836	1.016	1.026
N-WER (ϵ)	N/A	0.133	0.775	1.022	1.036
SNR	N/A	6.51	7.04	7.35	7.14

As a reference, the mean SNR of each subset was listed in the last row of the table. It can be seen from above table that the DEV1 set which was captured by using auto-focus setup of the camera has the best quality and the lowest N WER. The DEV2 set which suffered slight blur has higher N-WER and the two subsets DEV3 and DEV4 have the highest N-WER due to the severe blur. However, the SNR has the poor ability to assess the quality of the document image, particularly for subset DEV1 which had the lowest SNR but had the best quality and the best OCR results. Similarly, the WER, N-WER and SNR of the testing data sets were shown in table 2.3.

Table 2.3: WER, N-WER and SNR of testing data sets [17].

	TEST	TEST1	TEST2	TEST3	TEST4
WER (wer)	0.262	0.369	0.776	1.002	1.001
N-WER (ϵ)	N/A	0.145	0.696	1.003	1.001
SNR	N/A	6.50	7.13	7.32	7.18

In table 2.4, the equal error rates (EER) and the corresponding optimal threshold T0 which was based on the predicted N-WER for different classification tests are shown. From the table 3, it can be observed that the system had the best performance and the optimal threshold T0 was consistent with the pre-defined threshold when T = 0.3. By analyzing the distribution of the N-WER, it can be seen that threshold T = 0.3 best describes the difference between good and bad OCR results of the data sets, even though some document images have "good" quality visually but have bad OCR and document analysis capability [17].

Table 2.4: Threshold & EER for Classification Results [17].

Pre-defined T	EER (%)	Optimal T' on test
T = 0.1	12.3	T' = 0.265
T = 0.3	9.7	T' = 0.307
T = 0.5	11.6	T' = 0.325
T = 0.7	14.3	T' = 0.340

2.3.5. A Binarization method with learning built rules for document images produced by cameras:

To apply SVM to the binarization problem, first the problem is mapped to the SVM setting. This is done by dividing each image into $k \times k$ regions, which constitute the training samples. First, for certain regions, there may be more than one appropriate binarization action which is divided into 9 regions from A to I. Among the regions, A and C carry a single label, T_{min}, D to I carry labels T_{Otsu} and T_{min}, and B carries 0 and T_{min} as detailed in table 2.5.

Table 2.5: The labels of regions A to I and the values of T_{Otsu} and T_{min} for each region [19].

Region	Label	T_{Otsu}	T_{min}
A	Tmin	138	106
B	0, Tmin	151	106
C	Tmin	147	106
D	TOtsu, Tmin	106	97
E	TOtsu, Tmin	109	97
F	TOtsu, Tmin	107	99
G	TOtsu, Tmin	97	97
H	TOtsu, Tmin	100	99
I	TOtsu, Tmin	99	99

Table 2.6 shows the recall rate, precision rate, and F1 score of the OCR results. All three measures suggest that dividing an image into 3×3 regions produces the best results.

Table 2.6: The OCR accuracy derived by applying the simple binarization scheme to the 122 images, sub-divided into kXk regions for $k = 2,3,\ldots,10$ [19].

Number of regions	Recall rate	Precision rate	F_1 score
2×2	93.73	93.76	93.74
3×3	**93.80**	**95.15**	**94.46**
4×4	92.77	94.82	93.78
5×5	92.58	92.75	92.65
6×6	91.86	93.85	92.84
7×7	92.80	93.84	93.31
8×8	92.06	93.47	92.74
9×9	91.86	94.11	92.97
10×10	91.29	93.54	92.40

2.3.6. A partition approach for the restoration of camera images of planar and curled document:

In this case the character recognition rate is used to evaluate the restoration performance since OCR is one of the main applications and the subsequent document indexing and retrieval depend heavily on OCR results as well. Recognition experiments are conducted using the OCR software Fine Reader. Ninety distorted camera images were first input to recognition system one by one and the recognition rates were recorded. Distorted camera documents were then restored using the presented method. The restored document images were again fed to the Fine Reader.

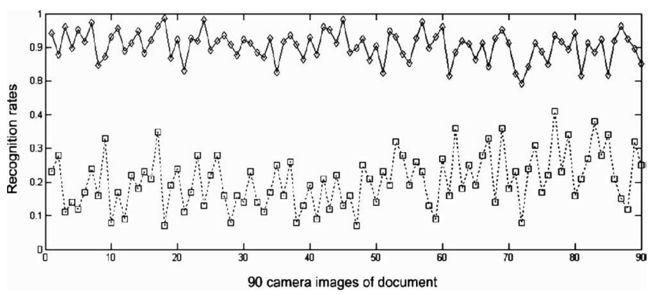


Fig 7: Recognition rate before and after restoration [21].

Fig. 7 gives the recognition rates of 90 camera documents before and after the restoration operation. The dotted line labeled with square in Fig. 7 gives the recognition rates of 90 camera documents before the restoration. The lower recognition rates (130% on average) can be expected since generic OCR systems can not deal with camera documents with perspective and geometric distortions. More specifically, most of text regions are falsely recognized as graphic due to the large skew angle. The solid curve labeled with diamond in Fig. 7 gives the recognition rates of document images restored using this method. The average recognition rate reaches over 90%. The recognition rate is so high because generic OCR systems are tolerant of slight variation of aspect ratio of character images. Therefore, restored document text can be recognized correctly even though the aspect ratios of some target rectangle are not accurately estimate [21].

2.3.7. Remove Noise and Reduce Blurry Effect From Degraded Document Images:

In this case the evaluation was carried using the following parameters.

I Size of I/O image. By calculating the row and column pixels, it is used to find the total size of original and restored image.

II MSE is **Mean Square Error**, $f(i,j)$ is pixel value of output image, $F(i,j)$ is pixel value of input image. Given by Formula:

$$MSE = ((no\_pixels\_in\_output\_imageno\_pixels\_in\_input\_image).\hat{}2)./((Size\_Of\_Image).\hat{}2)$$

III PSNR (peak signal to noise ratio) is used to measure the quality of restored image compared to the original image. Larger is the value, better will be the quality of image. It is calculated using equation as follow:

$PSNR = 20\ log10(\ 255\ /\ MSE)$ where MSE defined in II refers to mean square error.

The quality of the image is higher if the PSNR value of the image is high. Since PSNR is inversely proportional to MSE value of the image, the higher the PSNR value is, the lower the MSE value will be. Therefore better the image quality is, lower the MSE value will be.

IV Time calculation: It is used as MATLAB command CLOCK to calculate time for the code to be executed, CLOCK is inbuilt command to show the real time, the command is used twice to calculate time consuming parameter [22]. The evaluation parameters are shown in Table 2.7 below:

Table 2.7: Evaluation parameters [22].

S. NO.	NO. OF PIXELS OF I/P IMAGES	NO. OF PIXELS OF O/P IMAGES	SIZE OF I/O IMAGES	MSE	PSNR	TIME TAKEN(*10-3)
1	528721	560153	565820	0.0031	49.1715	1.919000
2	1283804	1317755	1320370	0.0066117	55.8623	2.964000
3	318176	332435	332478	0.0018	51.4189	1.373000
4	476790	501494	502095	0.0024	50.2258	1.747000
5	650976	673892	674866	0.0012	53.4470	2.028000

2.3.8. A Novel image restoration algorithm for digitized degraded historical documents

The performance parameters considered in this paper are Elapsed Time, Peak Signal to Noise Ratio (PSNR), Normalized Absolute Error (NAE) and Mean Square Error (MSE) and Root Mean Square Error (RMSE) [23].

2.4. Observation and selection of fundamental steps

Based on the literature review and the results shown in the previous sections, it can be observed that the potential of optimization techniques can be employed for the stated problem. The restoration includes the morphological operations where the structuring element can be engaged. Such structuring elements although have standard formats but the idea here is to get optimized image-wise structuring elements. Finally based on the literature survey in this thesis we present the following steps for restoration of degraded document images.

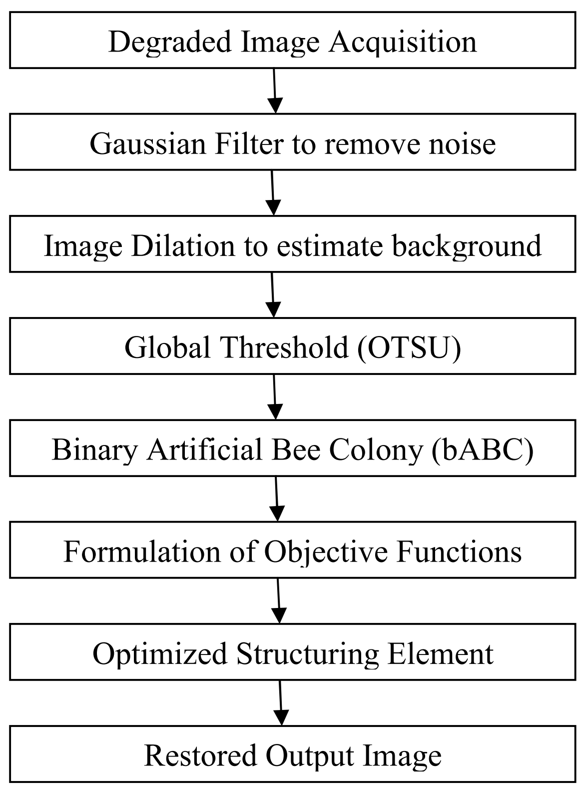
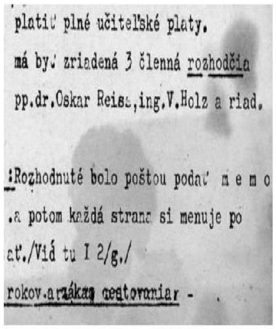


Fig 8: Flowchart of the process.

The presented method is described in details in the following chapters.

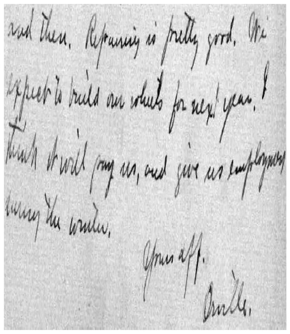
The following images have been considered for testing purpose.



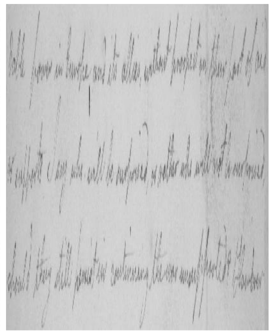
platiť plné učiteľské platy.
má byť zriadená 3 členná rozhodčia
pp.dr.Oskar Reiss,ing.V.Holz a riad.

Rozhodnuté bolo poštou podať m e m o
.a potom každá strana si menuje po
ať./Viď tu I 2/g./
rokov.arašaa ogstovanier -

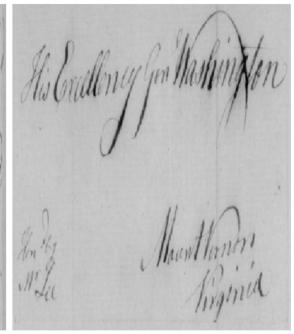
Sample 1



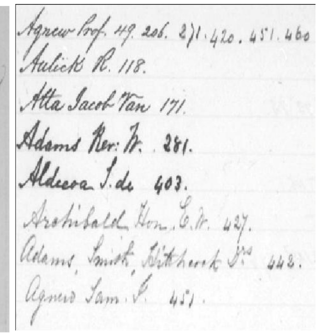
Sample 2



Sample 3



Sample 4



Sample 5

2.5. Conclusions

Each of the above methods was carried out and the results are jotted down and then compared. Analyses of each of the techniques are presented and conclusion is drawn towards the possible method that can be adopted for degraded document image restoration. In this thesis, a new algorithm is being developed for the restoration of the digitized degraded document images like, images of historical documents.

2.6. References

[1] J. Kumar, F. Chen, and D. Doermann, "Sharpness estimation of document and scene images," in International Conference on Pattern Recognition (ICPR), 2012, pp. 3292–3295.

[2] F. Chen, S. Carter, L. Denoue, and J. Kumar, "SmartDCap: semiautomatic capture of higher quality document images from a smartphone", in International conference on Intelligent user interfaces (IUI), 2013, pp. 287–296.

[3] J. Kumar, R. Bala, H. Ding, and P. Emmett, "Mobile Video Capture of Multi-page Documents," in IEEE International Workshop on Mobile Vision (IWMV), 2013, pp. 35–40.

[4] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," ACM Trans. Graph., vol. 25, pp. 787–794, July 2006.

[5] "Restoring camera-captured distorted document images", Changsong Liu · Yu Zhang · Baokang Wang Xiaoqing Ding, IJDAR DOI 10.1007/s10032-014-0233-8,pp. 2-9 (2014).

[6] Bookstein, Fred L., "Principal warps: thin-plate splines and the decomposition of deformations", IEEE Trans. Pattern Anal. Mach. Intell. 11, 567–585 (1989).

[7] L. Blando, J. Kanai, and T. Nartker, "Prediction of OCR accuracy using simple image features," in International Conference on Document Analysis and Recognition, vol. 1, 1995, pp. 319–322.

[8] S. V. Rice, J. Kanai, and T. A. Nartker, “The third annual test of OCR accuracy,” TR 94-03, ISRI, University of Nevada, Las Vegas, 1994.

[9] M. Cannon, J. Hochberg, and P. Kelly, “Quality assessment and restoration of typewritten document images,” International Journal on Document Analysis and Recognition, vol. 2, no. 2-3, pp. 80–89, 1999.

[10] “A Dataset for Quality Assessment of Camera Captured Document Images ”, Jayant Kumar, Peng Ye and David Doermann, University of Maryland, College Park, USA,pp. 2-3 (2013).

[11] T.Romen Singh, Sudipta Roy, O.Imocha Singh, Tejmani Sinam and Kh.Manglem Singh, “A New local Adaptive Thresholding Technique in Binarization”, IJCSI-Vol 8, issue 6 No. 2 pp. 271-277 (Nov, 2011).

[12] J. Sauvola *et al.* and M. Pietikainen, “Adaptive document image binarization,” Pattern Recognition 33(2), pp. 225–236, 2000.

[13] Niblack, W.: “An introduction to digital image processing” (Prentice- Hall, Englewood Cliffs, NJ, 1986), pp. 115–116.

[14] Bernsen, J.: “Dynamic thresholding of gray-level images”. Proc. 8th Int. Conf. on Pattern Recognition, Paris, 1986, pp. 1251–1255.

[15] “Local Contrast and Mean based Thresholding Technique in Image Binarization”, O. Imocha Singh, *et al.*, International Journal of Computer Applications (0975 – 8887) Volume 51– No.6, August 2012.

[16] Alex J. Smola and Bernhard Schölkopf, “A tutorial on support vector regression,” Statistics and Computing, vol. 14, pp. 199–222, 2004.

[17] “Automated Image Quality Assessment For Camera-Captured OCR”, Xujun Peng, Huaigu Cao, Krishna Subramanian, Rohit Prasad, and Prem Natarajan,18th IEEE International Conference on Image Processing, pp. 2-4, 2011.

[18] N. Otsu, “A thresholding selection method from gray-scale histogram”, IEEE Transactions on System, Man, and Cybernetics 9 (1979) 62–66.

[19] “A binarization method with learning built rules for document images produced by cameras”, Chien-Hsing Chou *et al.*, Pattern Recognition 43 (2010) pp. 1518–1530.

[20] S.J. Lu, B.M. Chen, C.C. Ko, Perspective rectification of document images using fuzzy set and morphological operations, Image and Vision Computing 23 (2005) 541–553.

[21] “A partition approach for the restoration of camera images of planar and curled document”, Shijian Lu, Ben M. Chen *, C.C. Ko, S. Lu *et al.* / Image and Vision Computing 24 (2006) 837–848.

[22] “Remove Noise and Reduce Blurry Effect From Degraded Document Images”, Sarika Jain and Pankaj Parihar, International Journal of Engineering Research and General Science Volume 3, Issue 1, January February, 2015, ISSN 2091-2730.

[23]” A Novel image restoration algorithm for digitized degraded historical documents”,
Rupinder Kaur and Jaspreet Kaur, International Journal of Science, Engineering and Technology Research, Volume 3, Issue 9, September 2014, pp. 2275-2277.

CHAPTER 3

BINARY ARTIFICIAL BEE COLONY (bABC)

3.1. Introduction

Artificial bee colony (ABC) [1] optimization is a relatively new population-based, stochastic optimization technique. ABC was developed to optimize unconstrained problems within continuous-valued domains. Artificial bee colony (ABC) algorithm is one of the recently swarm intelligence based algorithms for continuous optimization. It had been applied to many scientific and engineering problems for its ease of use and efficiency [2]. However, the original ABC technique could not be used in a binary search space directly. Therefore it is not possible to use the original ABC algorithm directly to optimize problems in binary domain. In our thesis, we have used the binary version of ABC, called bABC, which is particularly designed for binary optimization problems. bABC uses a new differential expression, which employs a measure of optimized structuring element for restoring the degraded document images.

The original ABC algorithm is mainly used for continuous optimization, however many engineering practices have been described as a combinatorial optimization problem. Binary artificial bee colony (bABC) algorithm can be used to optimize the 0-1 integer programming problem. In ABC algorithm, artificial bee hive consists of three types of bees: employed bees, onlookers and scouts. Employed bees correspond with the specific food sources. Their tasks are to forage in the corresponding food sources. Onlookers watch the waggle dance of employed bees within the hive to choose the excellence food source. Scouts search for food sources randomly [3]. In this thesis, theoretical and practical analysis had been carried out about operating scheme of the bABC algorithm, for generating optimized structuring element towards degraded document image restoration, and was proven to be a suitable method of restoration.

3.2. Survey on different swarm intelligence techniques

The term swarm is used for an aggregation of animals such as fish schools, bird flocks and insect colonies such as ant, termites and bee colonies performing collective behaviour. The individual agents of a swarm behave without supervision and each of these agents has a stochastic behaviour due to her perception in the neighbourhood. Local rules, without any relation to the global pattern, and interactions between self-organized agents lead to the emergence of collective intelligence called swarm intelligence. Swarm Intelligence appears in biological swarms of certain insect species. It gives rise to complex and often intelligent behavior through complex interaction of thousands of autonomous swarm members. Interaction is based on primitive instincts with no supervision. The end result is accomplishment of very complex forms of social behavior and fulfillment of a number of optimization and other tasks. The main principle behind these interactions is called stigmergy, or communication through the environment. An example is pheromone laying on trails followed by ants [4]. Pheromone is a potent form of hormone that can be sensed by ants as they travel along trails. It attracts ants and therefore ants tend to follow trails that have high pheromone concentrations. This causes an autocatalytic reaction, i.e., one that is accelerated itself. Ants attracted by the pheromone will lay more of the same on the same trail, causing even more ants to be attracted. Another form of stigmergy alters the environment in such a manner as to promote further similar action by the agents. This process is dubbed task-related stigmergy. An example is sand grain laying by termites when constructing nests. In the initial stages of construction, termites lay sand grains at random locations. This stimulates further laying by other members of the swarm, until a single heap of sand grains randomly reaches a critical mass that is larger than its neighboring heaps. At that point, most termites are attracted to that specific heap, thereby selecting that specific site for construction of their nest [4].

Swarm intelligence has become a research interest to many research scientists of related fields in recent years. The swarm intelligence is defined as ‘‘any attempt to design algorithms or distributed problem solving devices inspired by the collective behavior of social insect colonies and other animal societies...’’ by Bonabeau *et. al* [5]. However, the term swarm is used in a general manner to refer to any restrained collection of interacting agents or individuals. The classical example of a swarm is bees swarming around their hive; nevertheless the metaphor can easily be extended to other systems with a similar architecture. For instance, an ant colony can be thought of as a swarm whose individual agents are ants; a flock of birds is a swarm of birds; an immune system is a swarm of cells as well as a crowd is a swarm of people [6].

A survey was conducted on Swarm intelligence classified based on different natural insects and can be represented as the Table 3.1.

Table 3.1: Some of the popular swarm intelligence algorithms

S. No.	**Swarm Intelligence Algorithms**			
	Name of Algorithm	**Year of Development**	**Based on Technique**	**Brief of the Technique**
1	Altruism	Foster KR, Wenseleers T (2006) [7]	Hamilton’s rule of kin selection	Altruistic process increases fitness of a randomly chosen chromosome-recipient, and decreases fitness of a randomly chosen chromosome with an altruistic allele-donor. The increase/decrease is unproportional to an absolute value of the difference of their affiliated groups. Results

				suggest that an altruistic behavior in the initial stage was created by a random genetic drift in a subgroup and initial division of a population into subgroups is crucial for its evolutionary (meta) stability.
2	Ant Colony Optimization	Marco Dorgio (1992) [8]	Behavior of ants searching for food.	At first, the ants wander randomly. When an ant finds a source of food, it walks back to the colony leaving "markers" (pheromones) that show the path has food. When other ants come across the markers, they are likely to follow the path with a certain probability. If they do, they then populate the path with their own markers as they bring the food back. As more ants find the path, it gets stronger until there are a couple streams of ants traveling to various food sources near the colony.
3	Artificial Bee Colony	Karaboga (2005) [9]	Honey Bee colony for searching of food sources	The colony consists of three groups of bees: employed bees, onlookers and scouts. It is assumed that there is only one artificial employed bee for each food source. Employed bees go to their food source and come back to hive and dance on this area. The employed bee whose food source has been abandoned becomes a scout and starts to search for finding a new food source. Onlookers watch the dances of employed bees and choose food sources depending on dances.
4	Artificial Immune System	De Castro & Von Zuben's (2002) [10]	Abstract Structure and function of	The clonal selection theory has been used as inspiration for the development of AIS that perform computational optimisation and pattern recognition tasks. In particular, inspiration has been taken from the antigen driven affinity

			immune system	maturation process of B-cells, with its associated hypermutation mechanism. These AIS also often utilise the idea of memory cells to retain good solutions to the problem being solved.
5	Particle Swarm Optimization	Kennedy & Eberhart (1995) [11]	Inspired by Swarm	A basic variant of the PSO algorithm works by having a population (called a swarm) of candidate solutions (called particles). These particles are moved around in the search-space according to a few simple formulae. The movements of the particles are guided by their own best known position in the search-space as well as the entire swarm's best known position. When improved positions are being discovered these will then come to guide the movements of the swarm. The process is repeated and by doing so it is hoped, but not guaranteed, that a satisfactory solution will eventually be discovered.
6	Charged System Search	Kaveh A. & Talatahari S. (2010) [12]	Based on some principles from physics	In order to find the pareto-optimal results, the problem has been transformed into single-objective optimization problem by using weighted sum method. The algorithm has been applied to the 30 bus 6 generator (IEEE) test system by including transmission line losses. Also, economic power dispatch problem with prohibited operating zone which considers ramp

			and mechanics	rate limit, has been solved by CSS. B loss matrix has been used for the computation of the transmission line losses. In the study, the pareto-optimal solutions obtained for different weight values (*w*) have been compared with the solution values obtained by the other methods in literature.
7	Cuckoo Search	Yang Xin-She & Deb Suash (2009) [13]	Mimics the brooding behavior of some cuckoo species	Each egg in a nest represents a solution, and a cuckoo egg represents a new solution. The aim is to use the new and potentially better solutions (cuckoos) to replace a not-so-good solution in the nests. In the simplest form, each nest has one egg. The algorithm can be extended to more complicated cases in which each nest has multiple eggs representing a set of solutions.
8	Firefly Algorithm	Yang Xin-She (2008) [14]	Inspired by The flashing behavior of fireflies.	The primary purpose for a firefly's flash is to act as a signal system to attract other fireflies. Assuming all fireflies are unisexual, so that any individual firefly will be attracted to all other fireflies; Attractiveness is proportional to their brightness, and for any two fireflies, the less bright one will be attracted by (and thus move towards) the brighter one; however, the intensity (apparent brightness) decrease as their mutual distance increases; If there are no fireflies brighter than a given firefly, it will move randomly. The brightness should be associated with the objective function.
9	Intelligent Water Drops	Shah-Hosseini hamed	Inspired by natural rivers and	Almost every IWD algorithm is composed of two parts: a graph that plays the role of distributed memory on which soils of different edges are preserved, and the moving part of the

		(2009) [15]	how they find almost optimal paths to their destination.	IWD algorithm, which is a few number of Intelligent water drops. These Intelligent Water Drops (IWDs) both compete and cooperate to find better solutions and by changing soils of the graph, the paths to better solutions become more reachable. It is mentioned that the IWD-based algorithms need at least two IWDs to work.
10	River formation Dynamics (RFD)	Pablo Rabanal, Ismael Rodríguez, Fernando Rubio (2007) [16]	Based on copying how water forms rivers by eroding the ground and depositing sediments.	It is a Gradient version of ACO. The idea is to imitate the movement of drops on the edges between given nodes thus performing a search based on their height, which is modified through the mechanism of soil erosion and sediment deposition. In this way decreasing gradients are constructed, and these are followed by subsequent drops to compose new gradients and reinforce the best ones.
11	Gravitational Search Algorithm (GSA)	Rashedi, Nezamabadipour & Saryazdi (2009) [17]	Based on law of gravity and the notion of mass interaction.	The searcher agents are a collection of masses which interact with each other based on the Newtonian gravity and the laws of motion. GSA is reported to exclude the distance between masses in its formula, whereas mass and distance are both integral parts of the law of gravity.

Studying the above survey on various swarm intelligence techniques, we can say that two algorithms, namely, PSO and ABC have shown better result than others. After comparing both of them, we have found that

ABC has performed better over PSO in terms of global optimization [18]. Also, ABC algorithm produces better results on multimodal and multivariable problems than other algorithms, and is thus considered to be a base in our thesis. ABC also comes in binary form for 0-1 problems in binary domain and is known as Binary artificial bee colony (*bABC*). It is described in details in the next section.

3.3. Binary Artificial Bee Colony Techniques

Artificial bee colony (ABC) is a new direction in swarm intelligence that mimics the foraging behavior in honeybee swarms [6]. It has been successfully employed to solve complex engineering problems in diverse fields [19]. The comparative studies between the performance of ABC and other natural behavior inspired optimization techniques, for example, particle swarm optimization (PSO) algorithms [6] have shown that ABC has higher efficiency. ABC also includes comparatively fewer control parameters [20].

The classical ABC algorithm includes three groups of bees: employed bees, onlookers, and scouts. The optimization begins with a virtual hive that consists of equal number of employed bees and onlookers and proceeds as following steps:

Step 1: Initialization: optimization is initialized with an initial population consisting of N randomly generated initial solutions.

Step 2: Initialization of employed bees: generally half of the total solutions in the initial population are considered as employed bees, that is, $SN = N/2$ where SN denotes the number of employed bees in the hive. Each employed bee is associated with a random solution $x_{i1|i1} = 1,2,\ldots\ldots,SN$ from the initial population. The fitness of individual solution $f_i t_{i1}$ is evaluated.

Step 3: Generation of new solutions by employed bees: each of the employed bees performs search in its vicinity and generates new solution $v_{i1}j_1$ following Eq. (1).

$$v_{i1}j_1 = x_{i1}j_1 + \eta_{i1}j_1 \; (x_{i1}j_1 - x_kj_1) \qquad (1)$$

where $j1$ is the index for the dimension of the optimization problem, $\eta i1j1$ is a random number generated in $[-1,1]$, and k is a randomly selected number in $\{1,2,\ldots\ldots,SN\}$. The fitness of the new solution is evaluated and retained in case it is better than its current fitness.

Step 4: Selection by onlookers: the bees other than employed bees are considered as onlookers. The SN numbers of onlookers are probabilistically assigned to the employed bees using the probability (ρr) calculated as Eq. (2).

$\rho r_{i1} = f_i t_{i1} / \sum_{i1=1}^{SN} f_i t_{i1}$ (2)

Step 5: Generation of new solution by the onlookers: the onlookers assigned to individual employed bees further search in the vicinity of its employed bee and generate new solutions by Eq. (1). The fitness of the new solutions generated by the onlookers is evaluated. In case onlookers find any solution with higher fitness, it is retained over the solution of respective employed bee.

Step 6: Generation of new solutions by scout bees: the solutions associated with the employed bees that could not get any onlookers during step 4 are abandoned and associated bees are considered as scout bees. The scout bees perform random search and generate new solutions by Eq. (3) followed by fitness evaluation.

$x_{i1j1} = x_{j1}min + r\,(x_{j1}max - x_{j1}min)$ (3)

where r is a random number in $[0, 1]$, and x_{j1}^{max} and x_{j1}^{min} are the upper and lower bound of the j_1th dimension in the problem space.

Step 7: Selection of the optimized solution: the solutions found by the scout bees are compared in terms of fitness with the solution with highest fitness obtained at the end of step 5. The solution with best fitness is stored as optimized solution.

Steps 3–7 are repeated until the stopping criterion is satisfied.

3.3.1. ABC for binary domain

In binary search space, the continuous domain ABC may be modified as bABC where the modification is mainly in the update dynamics of ABC. The search in case of bABC is initiated with randomly generated binary solutions x_{i1j1}. In binary space, $_{jl}$ is the index for bit position in the binary solution. Thus, the difference calculated in Eq. (1) becomes the Hamming distance between solutions which leaves only three possibilities shown in Eq. (4).

$$x_{i1j1} - x_{kj1} = \begin{cases} 1, & if\ \ x_{i1j1} = 1 \text{ and } x_{kj1} = 0 \\ 0, & if\ \ x_{i1j1} = x_{kj1} \text{ 0 or 1} \\ -1, & if\ \ x_{i1j1} = 0 \text{ and } x_{kj1} = 1 \end{cases} \quad (4)$$

With the values obtained from Eq. (4), vi1 j1 can be calculated using equation (1) but that still results in updated values in continuous space instead of required binary space. A probabilistic approach is adopted to handle this where the update of a bit from its current position is done probabilistically. The sigmoid transform of the continuous valued v_{i1j1} is calculated as Eq. (5) to obtain an intermediate variable v_{i1j1}. The sigmoid transform also squashes the range of v_{i1j1} within $[0, 1]$. Values of v_{i1j1} more than 0.5 indicate higher probability of a bit to change its current state. Finally, the binary value for the bit is decided by comparing the current v_{i1j1}values against a uniformly generated random number τ in $[0, 1]$ as Eq. (6).

$$v_{i1j1} = 1/(1 + \exp(-v_{i1j1})) \quad (5)$$

$$v_{i1j1} = \begin{cases} 1, & if\ V_{i1j1} > \tau \\ 0, & otherwise \end{cases} \quad (6)$$

The same formulation can be employed for Eq. (3) associated with scout bees where the upper bound and lower bound become "1" and "0", respectively, in binary space [21].

3.4. Conclusions

The aim of preserving historical handwritten documents is to restore the degraded text containing information. But generally global threshold fails to restore the text adequately. Adaptive (local) thresholding is required for preserving the text in these documents. Optimization can be a possible solution to this and since the problem is in binary domain the optimization algorithms that can be used for binary problems need to be employed. This chapter provides a brief survey on various swarm intelligence based optimization algorithms and choose ABC to employ for the thesis problem since it has better performance in terms of global optimization and also includes comparatively fewer control parameters**.** To deal with the current problem the binary development of classical ABC also been shown in this chapter which will be used to optimize problem specific objective function.

3.5. References

[1] G. Pampara, A. P. Engelbrecht, "Binary artificial bee colony optimization", in Proceedings of the IEEE International Joint Conference on Neural Networks, pp. 1942-1948, 1995.

[2] M. Tuba, N. Bacanin, and N. Stanarevic, "Guided artificial bee colony algorithm," in Proceedings of the 5th European conference on European computing conference, 2011, pp. 398–403.

[3] Dervis Karaboga, Beyza Gorkemli, Celal Ozturk, Nurhan Karaboga, "A comprehensive survey: artificial bee colony (ABC)algorithm and applications," in Artif Intell Rev, 11 March 2012, pp. 37.

[4] I. Kassabalidis,M.A.El-Sharkawi, R.J.Marks Ii,P. Arabshahi,A.A.Gray, "Swarm Intelligence for Routing in Communication Networks".

[5] Bonabeau, E. (1996), “Marginally stable swarms are flexible and efficient”, Journal de Physique I, 6, 309–324.

[6] Karaboga, D., Basturk, B.: “On the performance of artificial bee colony (ABC) algorithm”, Appl. SoftComput. 8(1), 687–697 (2008).

[7] Jiří Pospíchal, Vladimír Kvasnička, “A Study of Altruism by Genetic Algorithm”, in Springer-Verlag Berlin Heidelberg, Autonome Mobile Systeme, 1997, pp. 507-508.

[8] Wiktor K. Macura, “Ant Colony Algorithm”, in MathWorld-A Wolfram Web Resource.

[9] Wu, L. (2011). “Magnetic Resonance Brain Image Classification by an Improved Artificial Bee Colony Algorithm”, Progress in Electromagnetics Research – Pier 116: 65–79.

[10] V. Cutello, G. Narzisi, G. Nicosia, M. Pavone, “Clonal Selection Algorithms: A Comparative Case Study using Effective Mutation Potentials”, optIA versus CLONALG, 4th Int. Conference on Artificial Immune Systems, ICARIS 2005, August 14-17, 2005, Banff, Canada. Springer, LNCS 3627:13-28, 2005.

[11] Kennedy, J. (1997). “The particle swarm: social adaptation of knowledge”. Proceedings of IEEE International Conference on Evolutionary Computation. pp. 303–308.

[12] Serdar Özyön, Hasan Temurtaş, Burhanettin Durmuş, Gültekin Kuvat, “Charged system search algorithm for emission constrained economic power dispatch problem”, in Energy and Exergy Modelling of Advance Energy Systems, Volume 46, Issue 1, October 2012, Pages 420–430.

[13] X.-S. Yang; S. Deb (December 2009), “Cuckoo search via Lévy flights”, World Congress on Nature & Biologically Inspired Computing (NaBIC 2009). IEEE Publications. pp. 210–214.

[14] Lukasik, S.; Zak, S. (2009), “Firefly algorithm for continuous constrained optimization task”. ICCCI 2009, Lecture Notes in Artificial Intelligence (Eds. N. T. Ngugen, R. Kowalczyk, S. M. Chen) 5796. pp. 97–100.

[15] Shah-Hosseini, H. (2008). “Intelligent water drops algorithm: a new optimization method for solving the multiple knapsack problem”. Int. Journal of Intelligent Computing and Cybernetics 1 (2): 193–212.

[16] G. Redlarski, A. Pałkowski, M. Dąbkowski, "Using River Formation Dynamics Algorithm in Mobile Robot Navigation", Solid State Phenomena, Vol. 198, pp. 138-143, 2013.

[17] Esmat Rashedi, Hossein Nezamabadi-pour, Saeid Saryazdi, "GSA: A Gravitational SearchAlgorithm", Information Sciences, Volume 179, Issue 13, 13 June 2009, Pages 2232–2248, Special Section on High Order Fuzzy Sets.

[18] Dervis Karaboga, Bahriye Akay, "A comparative study of Artificial Bee Colony algorithm", Erciyes University, The Department of Computer Engineering, Melikgazi, 38039 Kayseri, Turkey, Applied Mathematics and Computation 214 (2009) 108–132.

[19] Pan, Q.K., Fatih Tasgetiren, M., Suganthan, P.N., Chua, T.J.: A discrete artificial bee colony algorithm for the lot-streaming flow shop scheduling problem. Inform. Sci. 181(12), 2455–2468 (2011).

[20] Karaboga, D., Akay, B.: A comparative study of artificial bee colony algorithm. Appl. Math. Comput. 214(1), 108–132 (2009).

[21] Arpitam Chatterjee, Bipan Tudu, Kanai Ch. Paul," Binary grayscale halftone pattern generation using binary artificial bee colony (bABC)",in SIViP DOI 10.1007/s11760-012-0388-z, Received: 26 April 2012 / Revised: 19 September 2012 / Accepted: 25 September 2012 © Springer-Verlag London 2012.

CHAPTER 4

OBJECTIVE FUCTION FORMULATION

4.1. Introduction

The concept of objective function is used in optimization. It denotes the function that can be minimized or maximized. It is used in many areas as in curve fitting, regression, economics, linear programming, etc. The objective function indicates how much each variable contributes to the value to be optimized in the problem. In mathematics, computer science and operations research, mathematical optimization is the selection of a best element (with regard to some criteria) from some set of available alternatives [1].

In the simplest case, an optimization problem consists of maximizing or minimizing areal function by systematically choosing input values from within an allowed set and computing the value of the function. The generalization of optimization theory and techniques to other formulations comprises a large area of applied mathematics. More generally, optimization includes finding "best available" values of some objective function given a defined domain (or a set of constraints), including a variety of different types of objective functions and different types of domains [2].

This chapter explores the suitability of different mathematical functions that are used for image quality evaluation as objective function to achieve the desired results. The different such mathematical functions called image quality metrics (IQM) are described and tested for the problem in this thesis. Although number of different such metrics are available this thesis concentrates on few of the most commonly used such metrics which can be adopted for comparison of restored binary image against the original grayscale image. Here the suitability of each such metrics as objective function is tested for standard sample images and compared against the established techniques. Such comparison may help to finalize the composition of optimization function that to be subjected to ABC.

4.2. Image quality evaluation metrics

Reliable assessment of image quality plays an important role in meeting the promised quality of service (QoS) and in improving the end user's quality of experience (QoE). There is a growing interest to develop objective quality assessment algorithms that can predict perceived image quality automatically. These methods are highly useful in various image processing applications, such as image compression, transmission, restoration, enhancement, and display. For example, the quality metric can be used to evaluate and control the performance of individual system components in image/video processing and transmission systems [3]. The following section describes each of the popular metrics.

4.2.1. Mean-squared error (MSE):

In statistics, the **mean squared error (MSE)** or mean squared deviation (MSD) of an estimator measures the average of the squares of the errors or deviations, that is, the difference between the estimator and what is estimated. MSE is a risk function, corresponding to the expected value of the squared error loss or quadratic loss. The difference occurs because of randomness or because the estimator doesn't account for information that could produce a more accurate estimate [4].

The MSE, as described in equation (1), is the second moment (about the origin) of the error, and thus incorporates both the variance of the estimator and its bias. For an unbiased estimator, the MSE is the variance of the estimator. Like the variance, MSE has the same units of measurement as the square of the quantity being estimated. In an analogy to standard deviation, taking the square root of MSE yields the root-mean-square error or root-mean-square deviation (RMSE or RMSD), which has the same units as the quantity being estimated; for an unbiased estimator, the RMSE is the square root of the variance, known as the standard deviation [5]. If $\hat{\boldsymbol{Y}}$ is a vector of η predictions, and $\boldsymbol{Y}$ is the vector of observed values corresponding to the inputs to the function which generated the predictions, then the MSE of the predictor can be estimated by equation(1):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{\boldsymbol{Y}}_i - \boldsymbol{Y}_i)^2 \qquad (1)$$

where, the MSE is the mean ($\frac{1}{n}\sum_{i=1}^{n}$) of the square of the errors ($(\hat{\boldsymbol{Y}}_i - \boldsymbol{Y}_i)^2$).

4.2.2. Peak signal-to-noise ratio (PSNR):

Peak signal-to-noise ratio, often abbreviated **PSNR**, is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale.

PSNR is most commonly used to measure the quality of reconstruction of lossy compression codecs (e.g., for image compression). The signal in this case is the original data, and the noise is the error introduced by compression. When comparing compression codecs, PSNR is an approximation to human perception of reconstruction quality. Although a higher PSNR generally indicates that the reconstruction is of higher quality, in some cases it may not. One has to be extremely careful with the range of validity of this metric; it is only conclusively valid when it is used to compare results from the same codec (or codec type) and same content [6].

PSNR, as shown in equation (2), is most easily defined via the mean squared error (MSE). Given a noise-free m×n monochrome image I and its noisy approximation K, MSE is defined as:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2$$

$$PSNR = 10.log_{10}\left(\frac{MAX_I^2}{MSE}\right)$$

$$= 20.log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right)$$

$$=20.log_{10}(MAX_I) - 10.log_{10}(MSE) \qquad (2)$$

Here, MAX_I is the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, this is 255. More generally, when samples are represented using linear PCM with B bits per sample, MAX_I is $2B - 1$. For color images with three RGB values per pixel, the definition of PSNR is the same except the MSE is the sum over all squared value differences divided by image size and by three [7].

4.2.3. Structural similarity index (SSIM):

The **structural similarity (SSIM)** index is a method for predicting the perceived quality of digital television and cinematic pictures, as well as other kinds of digital images and videos. SSIM is used for measuring the similarity between two images. The SSIM index is a full reference metric; in other words, the measurement or prediction of image quality is based on an initial uncompressed or distortion-free image as reference. SSIM is designed to improve on traditional methods such as peak signal-to-noise ratio (PSNR) and mean squared error (MSE), which have proven to be inconsistent with human visual perception.

The SSIM, as shown in equation (3), index is calculated on various windows of an image. The measure between two windows x and y of common size $N \times N$ is:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \qquad (3)$$

with

- μ_x the average of x;
- μ_y the average of y;
- σ_x^2 the variance of x;

- σ_y^2 the variance of y;
- σ_{xy} the covariance of x and y;
- $c_1 = (k_1 L)^2,\ c_2 = (k_2 L)^2$ two variables to stabilize the division with weak denominator;
- L the dynamic range of the pixel-values (typically this is $2^{\#bits\ per\ pixel} - 1$);
- $k_1 = 0.01$ and $k_2 = 0.03$ by default.

In order to evaluate the image quality this formula is usually applied only on luma, although it may also be applied on color (e.g., RGB) values or chromatic (e.g. YC_bC_r) values. The resultant SSIM index is a decimal value between -1 and 1, and value 1 is only reachable in the case of two identical sets of data. Typically it is calculated on window sizes of 8×8. The window can be displaced pixel-by-pixel on the image but the authors propose to use only a subgroup of the possible windows to reduce the complexity of the calculation [8].

4.2.4. Visual information fidelity (VIF):

Measurement of visual quality is becoming increasingly important in many image processing applications, such as acquisition, compression, communication, restoration, enhancement and reproduction. The goal of quality assessment (QA) methods is to assess the quality of images in a perceptually consistent manner and in close agreement with subjective human judgments. Traditionally, researchers have focused mainly on measuring visual quality of images by modeling the salient features of the human visual system (HVS) using the so-called full reference (FR) QA paradigm. FRQA algorithms typically measures the distance between the test and reference signals in some perceptual space using HVS models [9]. With the source and the distortion models, the **visual information fidelity (VIF)** criterion can be derived [10].

Thus, assuming that the model parameters G, the reference and distorted image information is given by equation (4):

$$I(\vec{C}^N ; \vec{E}^N \mid s^N) = \frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{M} log_2(1 + \frac{s_i^2 \lambda_k}{\sigma_n^2})$$

$$I(\vec{C}^N ; \vec{F}^N \mid s^N) = \frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{M} log_2(1 + \frac{g_i^2 s_i^2 \lambda_k}{\sigma_v^2 + \sigma_n^2}) \tag{4}$$

where, *i* is an index to a spatiotemporal block and n,v,k are the Eigen values. The reason for this conditioning is to tune the GSM model for a particular reference signal because it measures the quality of a particular reference test pair and not the ‘quality’ of a distortion channel for the whole ensemble of natural signals [11].

4.2.5. Image fidelity criterion (IFC):

Given a statistical model for the source and the distortion (channel), the obvious information fidelity criterion is the mutual information between the source and the distorted images. First the mutual information for one sub band is derived and later for multiple sub bands are generalized. Let $CN = (C1; C2; ::: ; CN)$ denote N elements from C. The underlying RF U is uncorrelated (and hence C is an RF with conditionally independent elements given S), and that the distortion model parameters G and $2¾\,V$ are known as priori. Let $DN = (D1; D2; :::\ DN)$ denote the corresponding N elements from D. The mutual information between these is denoted as $I(CN; DN)$. Due to the non-linear dependence among the CN by way of S, it is much easier to analyze the mutual information assuming S is known. This conditioning resembles the GSM model for the particular reference image, and thus models the source more specifically. Thus the image fidelity criterion is the conditional mutual information $I(CN; DNjSN = sN)$, where $SN = (S1;\ S2; :::\ ;\ SN)$ are the corresponding N elements of S, and SN denotes a realization of SN. The RF's V_k are also independent of each other. The **image fidelity criterion (IFC)** can be obtained by summing up all sub bands as shown in equation (5):

$$IFC = \sum_{\kappa \in subbands} I(C^{N_k,k}\ ;\ D^{N_k,k} \mid s^{N_k,k}) \qquad (5)$$

where, $C^{N_k,k}$ denotes N_k coefficients from the RF C_k of the k^{th} subband, and similarly for $D^{N_k,k}$ and $s^{N_k,k}$. The above is the image fidelity criterion that quantifies the statistical information which is shared between the source and the distorted images. An attractive feature of this criterion is that like MSE and some other mathematical fidelity metrics, it does not involve parameters associated with display device physics, data from visual psychology experiments, viewing configuration information, or stabilizing constants that dictates the accuracy of HVS based FR QA methods. The IFC does not require training data, however some implementation parameters will arise once it is implemented [12].

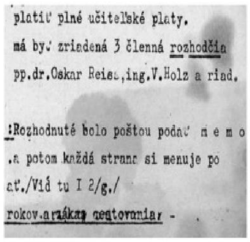
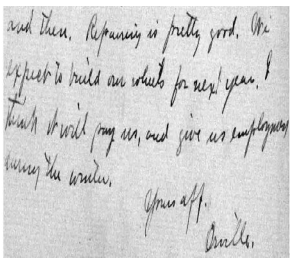
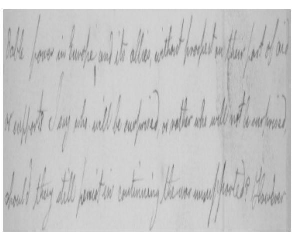
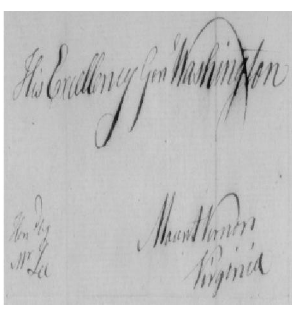
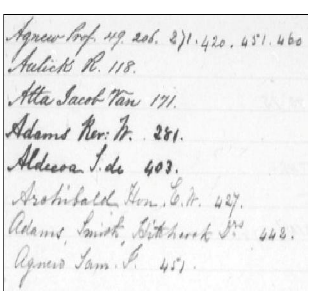
4.3. Quality Metrics as Objective Function

Quality Metrics are the measurement of the metrics as how they are effective as an objective function. A metric is a verifiable measure stated in either quantitative or qualitative terms. Objective functions based on the quality measurement are considered to give better performance.

This section explores the suitability of the above described quality metrics as objective function. Each of them was subjected to ABC for optimization of structuring element which will be used for morphological operations to restore the degraded images. The results with some of the standard test images are shown here. Also a comparison with the result of existing techniques as discussed in previous chapter has been presented for assessing the possibility of each quality metrics towards objective function.

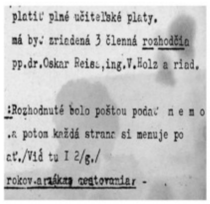
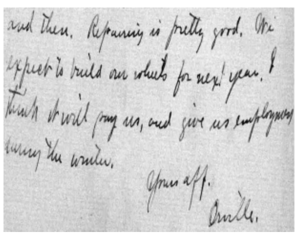
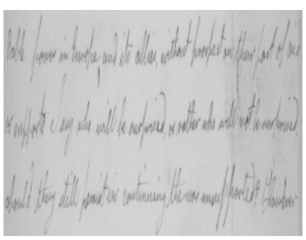
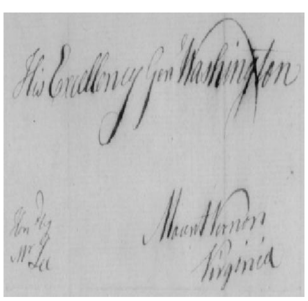
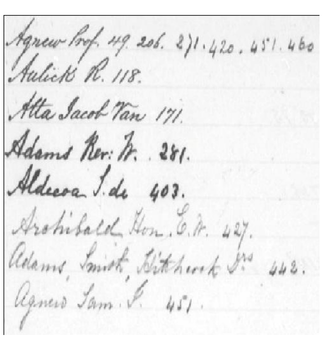
4.3.1. Result using MSE:

In our thesis, we have used MSE to optimize the structuring element. The value for some of our sample images using the existing method and bABC method are tabularized below:

Sample Images	Existing Process MSE value	ABC Method MSE value	Optimized Structuring Element
	0.065431	0.02003	0 1 0 1 0 1 1 0 0
	0.033653	0.004063	0 1 0 1 0 0 0 0 1
	0.058599	0.000823	0 1 0 1 0 0 0 0 1
	0.064837	0.000881	0 0 1 1 1 0 0 1 1
	0.016832	0.003169	0 1 0 1 0 1 1 0 0

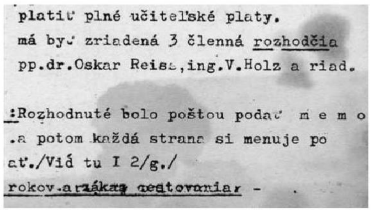
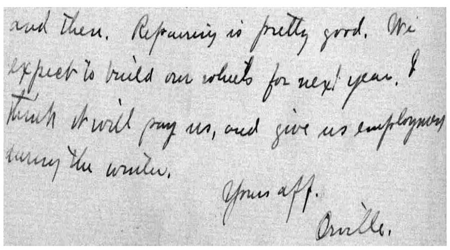
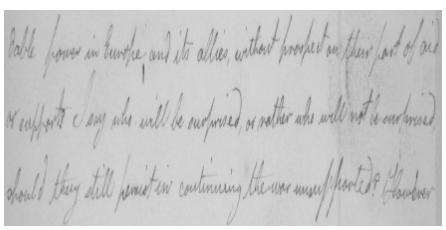
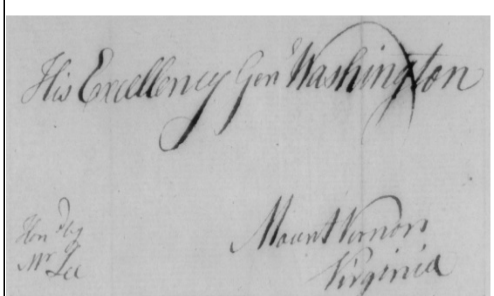
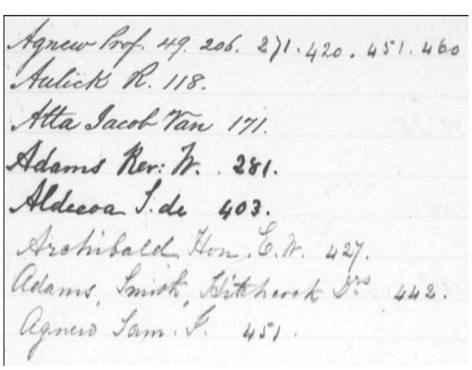
4.3.2. Result using PSNR:

In our thesis, we have used PSNR to optimize the structuring element. The value for some of our sample images using the existing method and bABC method are tabularized below:

Sample Images	Existing Process PSNR value	ABC Method PSNR value	Optimized Structuring Element
 platič plné učiteľské platy. má byť zriadená 3 členná rozhodčia pp.dr.Oskar Reis.,ing.V.Holz a rind. Rozhodnuté bolo poštou podn... n e m o .a potom každá strana si menuje po .Viď tu I 2/g./	59.198138	66.501162	0 1 0 1 0 0 0 0 1
 and then. Repairing is pretty good. We expect to build our wheels for next year. I think it will pay us, and give us employment during the winter. Yours aff. Orville.	62.867454	69.824151	0 1 0 0 1 1 1 1 1
	60.262912	77.373889	1 1 0 1 1 1 0 0 1
 His Excellency Gen Washington Mount Vernon Virginia	59.640789	78.454822	1 0 1 1 1 1 1 1 1
 Agnew Prof. 49. 206. 271. 420. 451. 460 Aulick R. 118. Alta Jacob Van 171. Adams Rev: W. 281. Archibald Hon. E.W. 427. Agnew Sam. J. 451.	65.841834	72.918651	0 1 0 1 0 0 0 0 1

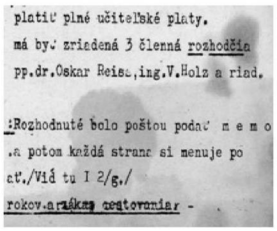
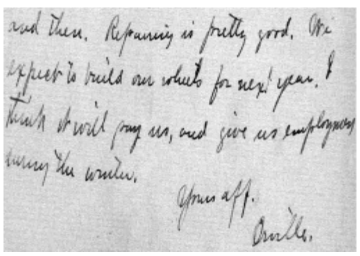
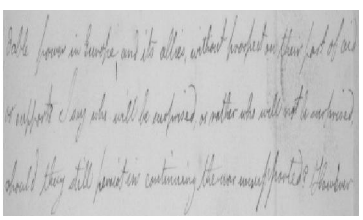
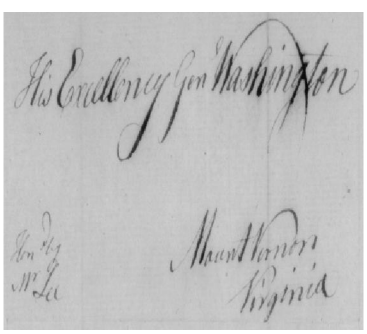
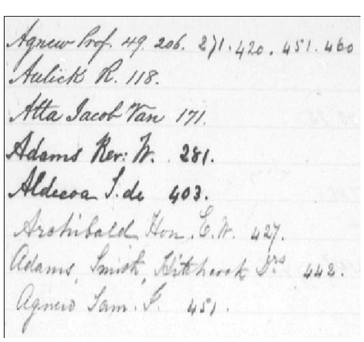
4.3.3. Result using SSIM:

In our thesis, we have used SSIM to optimize the structuring element. The value for some of our sample images using the existing method and bABC method are tabularized below:

Sample Images	Existing Process SSIM value	ABC Method SSIM value	Optimized Structuring Element
 platiť plné učiteľské platy. má byť zriadená 3 členná rozhodčia pp.dr.Oskar Reiss,ing.V.Holz a riad. :Rozhodnuté bolo poštou podať m e m o .a potom každá strana si menuje po ať./Viď tu I 2/g./ rokov.a rázkag nestovaniar -	0.992857	0.998279	1 0 1 1 1 1 1 1 1
 and then. Repairing is pretty good. We expect to build our wheels for next year. I think it will pay us, and give us employment during the winter. Yours aff. Orville.	0.996901	0.999595	1 1 0 0 0 0 1 1 1
	0.916021	0.999916	0 0 0 0 0 0 0 0 0
 His Excellency Gen Washington Hon by Mr Lee Mount Vernon Virginia	0.992862	0.999908	1 1 0 0 0 0 1 1 1
 Agnew Prof. 49. 206. 271. 420. 451. 460 Aulick R. 118. Alta Jacob Van 171. Adams Rev: W. 281. Aldecoa J. de 403. Archibald Hon. E.W. 427. Adams, Smith, Hitchcock Drs 442. Agnew Sam. J. 451.	0.999167	0.99959	0 1 0 1 0 0 0 0 1

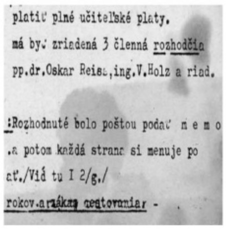
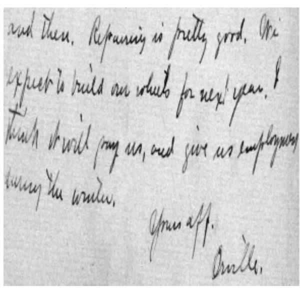
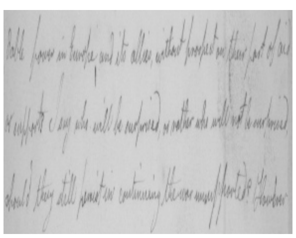
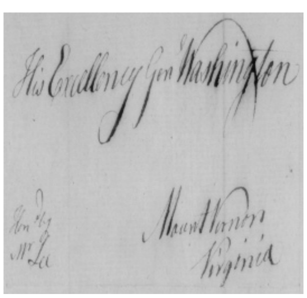
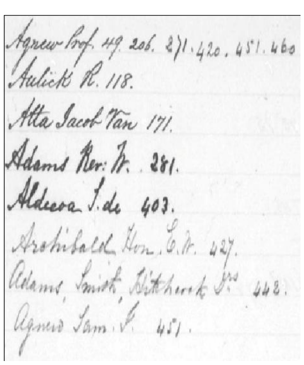
4.3.4. Result using VIF:

In our thesis, we have used VIF to optimize the structuring element. The value for some of our sample images using the existing method and bABC method are tabulárized below:

Sample Images	Existing Process VIF value	ABC Method VIF value	Optimized Structuring Element
 platiť plné učiteľské platy. má byť zriadená 3 členná rozhodčia pp.dr.Oskar Reiss,ing.V.Holz a riad. Rozhodnuté bolo poštou podať m e m o .a potom každá strana si menuje po ať./Viď tu I 2/g./ rokov.a rašťam cestovania. -	0.295590	0.271956	0 1 0 1 1 0 1 1 0
 and then. Repairing is pretty good. We expect to build our wheels for next year. I think it will pay us, and give us employment during the winter. Yours aff. Orville.	0.502038	0.520796	1 0 1 1 1 1 1 1 1
	0.399839	0.367465	1 0 1 1 1 1 1 1 1
 His Excellency Gen Washington Mount Vernon Virginia	0.542419	0.596276	1 0 1 1 1 1 1 1 1
 Agnew Prof. 49. 206. 271. 420. 451. 460 Aulick R. 118. Atta Jacob Van 171. Adams Rev: W. 281. Aldecoa J. de 403. Archibald Hon. E.M. 427. Agnew Sam. J. 451.	0.558641	0.559157	1 1 0 0 0 0 1 1 1

4.3.5. Result using IFC:

In our thesis, we have used IFC to optimize the structuring element. The value for some of our sample images using the existing method and bABC method are tabularized below:

Sample Images	Existing Process IFC value	ABC Method IFC value	Optimized Structuring Element
 platiť plné učiteľské platy. má byť zriadená 3 členná rozhodčia pp.dr.Oskar Reise,ing.V.Holz a riad. :Rozhodnuté bolo poštou podať n e m o .a potom každá strana si menuje po cť./Viď tu I 2/g./	1.425765	2.211746	0 1 0 1 0 0 0 0 1
 and then. Repairing is pretty good. We expect to build our wheels for next year. I think it will pay us, and give us employment during the winter. Yours aff. Orville.	1.695204	2.56389	0 0 1 0 1 0 0 0 1
	1.608719	2.56128	0 0 1 1 1 0 0 1 1
 His Excellency Genl Washington Mount Vernon Virginia	1.442371	2.525893	0 0 1 1 1 0 0 1 1
 Agnew Prof. 49. 206. 271. 420. 451. 460 Aulick R. 118. Adams Rev: W. 281. Aldecoa J. de 403.	1.712294	2.490761	1 1 0 0 0 0 1 1 1

4.4. Mathematical formulation of objective function

The objective function in our thesis is chosen from the parameters as described in previous section. We can calculate the efficiency of each of them as mentioned in equation (6).

$$Efficiency(\eta)\% = \frac{|bABC\ value - Original\ value|}{Original\ value}\ X\ 100 \qquad (6)$$

Each of the sample images efficiency for the parameters are calculated as follows:

For Sample 1:

MSE: $\frac{|0.020030-0.065431|}{0.065431} * 100 \cong 69.40\%$

PSNR: $\frac{|66.501162-59.198138|}{59.198138} * 100 \cong 12.34\%$

SSIM: $\frac{|0.998279-0.992857|}{0.992857} * 100 \cong 0.55\%$

VIF: $\frac{|0.271956-0.295590|}{0.295590} * 100 \cong 8.01\%$

IFC: $\frac{|2.211746-1.425765|}{1.425765} * 100 \cong 55.13\%$

For Sample 2:

MSE: $\frac{|0.004063-0.033653|}{0.033653} * 100 \cong 87.93\%$

PSNR: $\frac{|69.824151-62.867454|}{62.867454} * 100 \cong 11.07\%$

SSIM: $\frac{|0.999595-0.996901|}{0.996901} * 100 \cong 0.27\%$

VIF: $\frac{|0.520796-0.502038|}{0.502038} * 100 \cong 3.74\%$

IFC: $\frac{|2.563890-1.695204|}{1.695204} * 100 \cong 51.24\%$

For Sample 3:

MSE: $\frac{|0.000823-0.058599|}{0.058599} * 100 \cong 98.60\%$

PSNR: $\frac{|77.373889-60.262912|}{60.262912} * 100 \cong 28.40\%$

SSIM: $\frac{|0.999916-0.916021|}{0.916021} * 100 \cong 9.16\%$

VIF: $\frac{|0.367465-0.399839|}{0.399839} * 100 \cong 8.10\%$

IFC: $\frac{|2.561280-1.608719|}{1.608719} * 100 \cong 59.21\%$

For Sample 4:

MSE $\frac{|0.000881-0.064837|}{0.064837} * 100 \cong 98.64\%$

PSNR: $\frac{|78.454822-59.640789|}{59.640789} * 100 \cong 31.56\%$

SSIM: $\frac{|0.999908-0.992862|}{0.992862} * 100 \cong 0.71\%$

VIF: $\frac{|0.596276-0.542419|}{0.542419} * 100 \cong 9.93\%$

IFC: $\frac{|2.525893-1.442371|}{1.442371} * 100 \cong 75.12\%$

For Sample 5:

MSE: $\frac{|0.003169-0.01632|}{0.01632} * 100 \cong 81.17\%$

PSNR: $\frac{|72.918651-65.841834|}{65.841834} * 100 \cong 10.75\%$

SSIM: $\frac{|0.999590-0.999167|}{0.999167} * 100 \cong 0.04\%$

VIF: $\frac{|0.559157-0.558641|}{0.558641} * 100 \cong 0.09\%$

IFC: $\frac{|2.490761-1.712294|}{1.712294} * 100 \cong 45.46\%$

Based on the study, we can possibly see that among the five parameters, MSE, PSNR and IFC shows notable significant results. Again, from section 4.2; we can tell that MSE and PSNR are highly correlated, so we are taking PSNR. Thus, PSNR and IFC can be chosen as our objective functions. Final objective function in our thesis can be formulated as equation (7):

$$\phi = \omega_1 PSNR_{ij} + \omega_2 IFC_{ij} \quad (7)$$

where, ij corresponds to original and restored image, ω_1, ω_2 are the weight function which can be tuned depending on the contribution of respective metrics. However, $\omega_1 + \omega_2$ should be equal to 1.

4.5. Conclusion

In this chapter, the suitability of different image quality evaluation metrics as objective function been explored. The result of optimization is compared against established technique to assess the performance. The results shown suitability of PSNR and IFC. Hence, the objective function is formulated using those two metrics following the weighted summation formula.

4.6. References

[1] "The Nature of Mathematical Programming", Mathematical Programming Glossary, INFORMS Computing Society.

[2] Bonnans, J. Frédéric; Shapiro, Alexander (2000), "Perturbation analysis of optimization problems". Springer Series in Operations Research. New York: Springer-Verlag. pp. xviii+601.

[3] Tong Zhu and Lina Karam, "A no-reference objective image quality metric based on perceptually weighted local noise", in EURASIP Journal on Image and Video Processing, **DOI:** 10.1186/1687-5281-2014-5, © Zhu and Karam; licensee Springer. 2014.

[4] Lehmann, E. L.; Casella, George (1998). "Theory of Point Estimation" (2nd ed.). New York: Springer. ISBN 0-387-98502-6.MR 1639875.

[5] Wackerly, Dennis; Mendenhall, William; Scheaffer, Richard L. (2008), "Mathematical Statistics with Applications" (7 ed.). Belmont, CA, USA: Thomson Higher Education.

[6] Huynh-Thu, Q.; Ghanbari, M. (2008). "Scope of validity of PSNR in image/video quality assessment". Electronics Letters 44 (13): 800. doi:10.1049/el:20080522.

[7] Huynh-Thu, Q.; Ghanbari, M. (2008), "Scope of validity of PSNR in image/video quality assessment", Electronics Letters 44 (13): 800.

[8] Wang, Zhou; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. (2004-04-01). "Image quality assessment: from error visibility to structural similarity", IEEE Transactions on Image Processing 13 (4): 600–612. doi:10.1109/TIP.2003.819861. ISSN 1057-7149.

[9] Z. Wang, H. R. Sheikh, and A. C. Bovik, "Objective video quality assessment," in The Handbook of Video Databases: Design and Applications, B. Furht and O. Marques, Eds. CRC Press, 2003.

[10] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," IEEE Trans. Image Processing, Sept. 2004, Accepted.

[11] Hamid R. Sheikh and Alan C. Bovik, "A VISUAL INFORMATION FIDELITY APPROACH TO VIDEO QUALITY ASSESSMENT", at The University of Texas at Austin, USA.

[12] "An Information Fidelity Criterion for Image Quality Assessment Using Natural Scene Statistics", Hamid Rahim Sheikh, Student Member, IEEE, Alan C. Bovik, Fellow, IEEE, Gustavo de Veciana, Senior Member, IEEE, in INFORMATION FIDELITY CRITERION FOR IQA USING NSS.

CHAPTER 5

Restoration of degraded document images using bABC optimized structuring elements

5.1. Introduction

Optimization techniques play an important role to solve many complex engineering problems, the very purpose of which is to find the best ways so that a designer or a decision maker can derive a maximum benefit from the available resources [1]. The optimization problem is classified on the basis of nature of equations with respect to design variables. If the objective function and the constraints involving the design variable are linear then the optimization is termed as linear optimization problem. If even one of them is nonlinear it is classified as the non-linear optimization problem [2].

In this work the optimization of structuring element is explored towards restoration of degraded document images. The objective function as formulated in previous chapter based on the detailed study on the suitability of different image quality evaluation metrics are subjected to the optimization process. In this work, the binary version of artificial bee colony (ABC) been adopted for optimization. Since the structuring element is of binary nature, the optimization problems remains in binary domain, hence choosing the binary version of ABC instead of its original continuous domain version.

This chapter elucidates the fundamental concept of structuring element and morphological operations using them. Then the optimization of randomly generated initial structuring element by bABC is described to achieve the defined goal. The final results of optimization i.e. the restored images are also shown in this chapter. The comparison with standard techniques and analysis on the results are discussed at the end of this chapter as well.

5.2. Structuring element

In mathematical morphology, a **structuring element** (***strel***) is a shape, used to probe or interact with a given image, with the purpose of drawing conclusions on how this shape fits or misses the shapes in the image. It is typically used in morphological operations, such as dilation, erosion, opening, and closing, as well as the hit-or-miss transform.

According to Georges Matheron, knowledge about an object or an image depends on the manner in which it can be observed [3]. In particular, the choice of a certain *strel* for a particular morphological operation influences the information that one can obtain. There are two main characteristics that are directly related to *strel*, they are:

Shape: For example, the *strel* can be a "ball" or a line; convex or a ring, etc. By choosing a particular *strel*, one sets a way of differentiating some objects (or parts of objects) from others, according to their shape or spatial orientation.

Size: For example, one *strel* can be a $3X3$ square or a $21X21$ square. Setting the size of the structuring element is similar to setting the observation scale, and setting the criterion to differentiate image objects or features according to size.

Structuring elements are particular cases of binary images, usually being small and simple. In mathematical morphology, binary images are subsets of an Euclidean space $\Re^d$ or the integer grid $\mathcal{Z}^d$, for some dimension d. Here are some examples of widely used structuring elements (denoted by $\mathbb{B}$):

- Let $\mathbb{E} = \Re^2$; $\mathbb{B}$ is an open disk of radius r, centered at the origin.
- Let $\mathbb{E} = \mathcal{Z}^2$; $\mathbb{B}$ is a 3x3 square, that is, $\mathbb{B}$ ={(-1,-1), (-1,0), (-1,1), (0,-1), (0,0), (0,1), (1,-1), (1,0), (1,1)}.
- Let $\mathbb{E} = \mathcal{Z}^2$; $\mathbb{B}$ is the "cross" given by: $\mathbb{B} = \{(-1,0), (0,-1), (0,0), (0,1), (1,0)\}$.

In the discrete case, a structuring element can also be represented as a set of pixels on a grid, assuming the values 1 (if the pixel belongs to the structuring element) or 0 (otherwise).

The commonly performed morphological operation using *strel* are briefly discussed below:

Erosion: The erosion is one of the elementary operators of Mathematical Morphology, that is, it is one of the building blocks of a large class of operators. The key mechanism under the erosion operator is the local comparison of a shape, called structural element, with the object that will be transformed. If, when positioned at a given point, the structural element is included in the object then this point will appear in the result of the transformation, otherwise not [3]. For sets $\mathbb{A}$ and $\mathbb{B}$ in $\mathcal{Z}^2$ the erosion of $\mathbb{A}$ and $\mathbb{B}$, denoted $\mathbb{A} \ominus \mathbb{B}$, is defined as equation 1 [4]:

$$\mathbb{A} \ominus \mathbb{B} = \{z \mid (\mathbb{B})_z \subseteq \mathbb{A}\} \qquad (1)$$

This equation indicates that erosion of $\mathbb{A}$ by $\mathbb{B}$ is the set of all points z such that $\mathbb{B}$, translated by z, is contained in $\mathbb{A}$. The figure 1 below shows an object and the result of its erosion by a 3x3 cross structuring element [(0,1,0),(1,1,1),(0,1,0)].

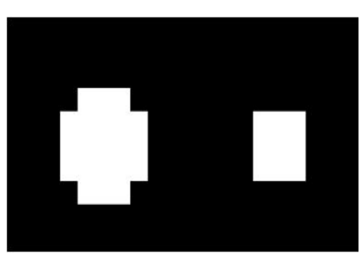


Fig 1: Object (left) and Result (right) of erosion [3].

Dilation: Dilation is one of the elementary operators of Mathematical Morphology, that is, it is a building block for a large class of operators. The key process in the dilation operator is the local comparison of a shape, called structural element, with the object to be transformed. When the structural element is positioned at a given point and it touches the object, then this point will appear in the result of the transformation, otherwise it will not [3]. With $\mathbb{A}$ and $\mathbb{B}$ as sets in $\mathcal{Z}^2$, the dilation of $\mathbb{A}$ by $\mathbb{B}$, denoted $\mathbb{A} \oplus \mathbb{B}$, is defined as in equation (2) [4]:

$$\mathbb{A} \oplus \mathbb{B} = \left\{ z \,\middle|\, (\hat{\mathbb{B}})_z \cap \mathbb{A} \neq \phi \right\} \qquad (2)$$

The figure 2 below shows an original object and the result of its dilation by a 3x3 square structuring element [(0,1,0),(1,1,1),(0,1,0)].

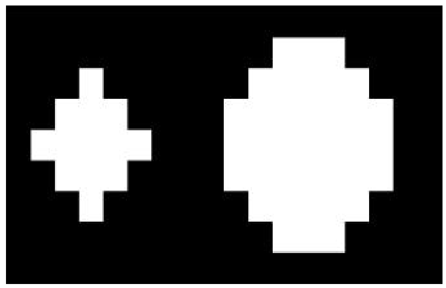


Fig 2: Object (left) and Result (right) of dilation [3].

Dilation and Erosion are duals of each other with respect to set complementation and reflection as shown in equation (3) [4]:

$$(\mathbb{A} \ominus \mathbb{B})^c = \mathbb{A}^c \oplus \hat{\mathbb{B}} \qquad (3)$$

Opening: This morphological operator can be expressed as a composition of erosion followed by a dilation, both by the same input structural element. As erosions and dilations, the key mechanism under the opening operator is the local comparison of a shape, the structural element, with the object that will be transformed. If, when positioned at a given point, the structural element is included in the object than the whole structural element will appear in the result of the transformation, otherwise none of its points will appear [3]. Opening generally smoothes the contour of an object, breaks narrow isthmuses, and eliminates thin protrusions. The opening of set $\mathbb{A}$ by structuring element $\mathbb{B}$, denoted $\mathbb{A} \circ \mathbb{B}$, is defined in equation (4) as [4]:

$$\mathbb{A} \circ \mathbb{B} = (\mathbb{A} \ominus \mathbb{B}) \oplus \mathbb{B} \qquad (4)$$

Thus, the opening $\mathbb{A}$ by $\mathbb{B}$ is the erosion of $\mathbb{A}$ by $\mathbb{B}$, followed by a dilation of the result by $\mathbb{B}$. The following figure 3 presents a binary image and the result of its opening by a small Euclidian disk.

Fig 3: Object (left) and Result (right) of opening [3].

Closing: This morphological operator can be expressed as composition of a dilation followed by erosion. As in the opening operator, the key mechanism under the closing operator is the local comparison of a shape, the structural element, with the object that will be transformed, if, when positioned at a given point, the structural element is included in the complement of the image, then the whole structural element will appear in the complement of the transformed image, otherwise none of its points will appear in this image [3]. Closing also tends to smooth sections of contours but, as opposed to opening, it generally fuses narrow breaks and long thin gulfs, eliminates small holes, and fills gaps in the contour. The closing of set $\mathbb{A}$ by structuring element$\mathbb{B}$, denoted by equation (5) as [4]:

$$\mathbb{A} \odot \mathbb{B} = (\mathbb{A} \oplus \mathbb{B}) \ominus \mathbb{B} \qquad (5)$$

Thus, the closing of $\mathbb{A}$ by $\mathbb{B}$ is simply the dilation of $\mathbb{A}$ by $\mathbb{B}$, followed by the erosion of the result by $\mathbb{B}$. The following figure 4 presents a binary image and the result of its closing by a small Euclidian disk.

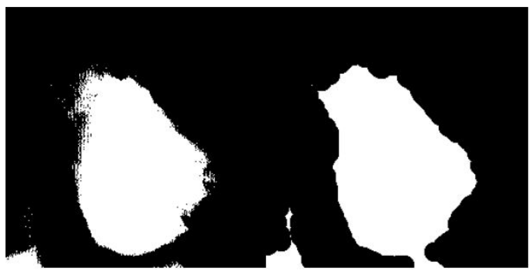


Fig 4: Object (left) and Result (right) of closing [3].

In our thesis, we have used Binary Artificial Bee Colony (bABC) algorithm to optimize the structuring element for restoration of degraded document images.

5.3. bABC based optimization

The Artificial Bee Colony (ABC) algorithm recently gained high popularity by providing a robust and efficient approach for solving continuous optimization problems. In order to apply ABC in discrete landscape, a binary version of artificial bee colony (bABC) algorithm is proposed in this thesis. Unlike the original ABC algorithm, the proposed bABC represents a food source as a discrete binary variable and applies discrete operators to change the foraging trajectories of the employed bees, onlookers and scouts in the probability that a coordinate will take on a zero or one value. The bABC is competitive to have significantly better performance than the other two successful discrete optimizer, namely, the genetic algorithm (GA) and particle swarm optimization (PSO) [5]. The implementation of novel binary ABC algorithm consists of initialization, modification of position by employed bees and onlooker bees and introduction of scout bees when there is no improvement in the fitness function. In this thesis, using binary artificial bee colony (bABC) algorithm, the optimized structuring element is generated and it is then used for restoring the degraded document images. The following section describes each of the steps of the bABC algorithm in details.

5.3.1. Phases of bABC:

The implementation of novel binary ABC algorithm consists of initialization, modification of position by employed bees and onlooker bees and introduction of scout bees for getting the optimized structuring element [5]. These phases are described below briefly:

- **Initialization phase:**

In the bABC model, structuring elements in a binary space exploited by bees represent possible solutions to a given binary optimization problem. At the initialization stage, a set of structuring elements positions are randomly selected by the bees. That is, the bee colony is initialized randomly over the entire search space [5]. In our thesis work, a set of initial binary solution of decided size was generated. The results presented in this chapter were initiated with 10 such solutions while each solution of 3x3 binary matrix.

- **Employed bee phase:**

After all the bees return to the hive with a certain amount of nectar (here it is the results obtained by the objective function), the first half (SN) that found the best structuring elements become "employed bees." At this stage, each employed bee x_i generates a new structuring element v_m in the neighborhood of its present position. Each of the employed bees performs search in its vicinity and generates new solution v_{m1n1} following eq. 6:

$$v_{m1n1} = x_{m1n1} + \eta_{m1n1} \, (x_{m1n1} - x_{kn1}) \qquad (6)$$

where $n1$ is the index for the dimension of the optimization problem, η_{m1n1} is a random number generated in $[-1, 1]$, and k is a randomly selected number in $\{1, 2, \ldots\ldots, SN\}$. The fitness of the new solution is evaluated and retained in case it is better than its current fitness. In order to restrict the bees' positions within the range [0,1], a normalization method and a threshold level has to be introduced to map all real valued numbers of v_{mn} to the range [0,1] [5],[6].

- **Onlooker bee phase:**

The remainder of the bees ("onlooker bees") watches the waggle dance to decide which of the employed bees should be followed. An onlooker bee selects a structuring element found by the employed bee x_m depending on its probability value. The SN numbers of onlookers are probabilistically assigned to the employed bees using the probability (ρr) is calculated as in equation (7).

$$\rho r_{m1} = f_m t_{m1} / \sum_{m1=1}^{SN} f_m t_{m1} \qquad (7)$$

Once the onlooker has selected the structuring elements found by the employed bee x_m, it produces a modification on the position x_m. As in the case of the employed bees, if the modified structuring element has a better or equal nectar amount(objective function value) than x_m, the modified structuring element will replace x_m and become a new member in the population. The onlookers assigned to individual employed bees further search in the vicinity of its employed bee and generate new solutions by Eq. (6). The fitness of the new solutions generated by the onlookers is evaluated. In case onlookers find any solution with higher fitness, it is retained over the solution of respective employed bee [5], [6].

- **Scout bee phase:**

In this stage, a random selection process carried out by the scout bees to explore new potential structuring elements. This is simulated as: if a position cannot be improved further through a predetermined number of cycles called "limit", then the structuring element is assumed to be abandoned, and the corresponding employed bee becomes a scout that randomly reinitializes in the fitness landscape. The solutions associated with the employed bees that could not get any onlookers are abandoned and associated bees are considered as scout bees. The scout bees perform random search and generate new solutions by equation (8) followed by fitness evaluation.

$$x_{m1n1} = x_{n1}^{min} + r\,(x_{n1}^{max} - x_{n1}^{min}) \qquad (8)$$

where r is a random number in $[0, 1]$, and x_{n1}^{max} and x_{n1}^{min} are the upper and lower bound of the ${n_1}^{th}$ dimension in the problem space. The solutions found by the scout bees are compared in terms of fitness with the solution with highest fitness obtained from Eq. (3). The solution with best fitness is stored as optimized solution [5],[6].

The optimization is performed using the previously developed objective function in section 4.4 with equal weightage to PSNR and IFC, again shown in equation (9):

$$\phi = \omega_1 PSNR_{mn} + \omega_2 IFC_{mn} \qquad (9)$$

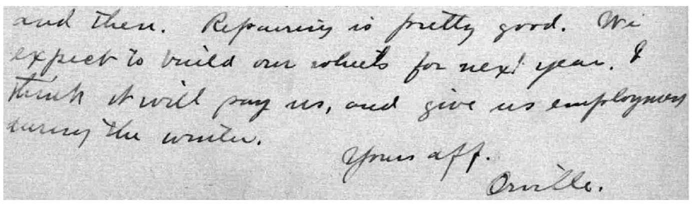
where, mn corresponds to original and restored image, ω_1, ω_2 are the weight functions which can be tuned depending on the contribution of respective metrics. However, $\omega_1 + \omega_2$ should be equal to 1.

5.4. Results and Discussions:

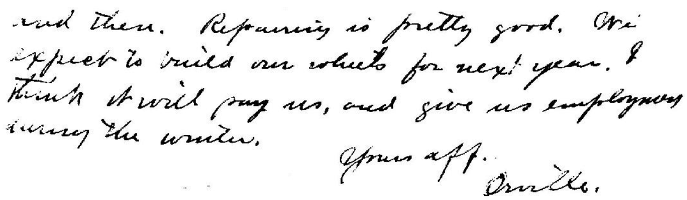
The bABC method was implemented to get the restored document images. The realization of this method is done in MATLAB software. Various parameters as the objective functions were applied to obtain optimized structuring elements. Then the restored image was compared with the ground truth images and the result was then compared with the ground truth standards. The restored results were also compared with some of the existing techniques such as Niblack [7], Sauvola *et al.* [8] and Otsu [9]. Some of the results are shown in the following figures from Fig. 5 to Fig. 9. The corresponding optimized structuring element is also mentioned along with the images.

The final results of bABC algorithm in the form of original images of some of the sample images, their corresponding restored images, optimized structuring elements along with the result images of Niblack [7], Sauvola *et al.* [8], Otsu's [9] methods and ground truth images, which will be used for comparison with our technique, are also shown as follows:

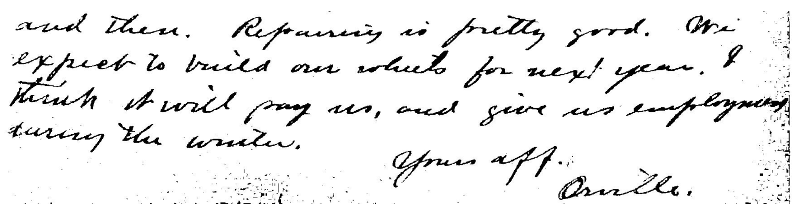
Result 1: Test Image 1



(a)



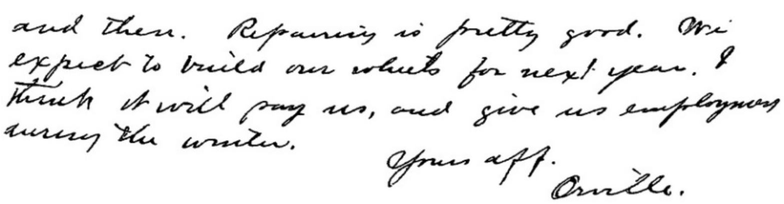
(f)



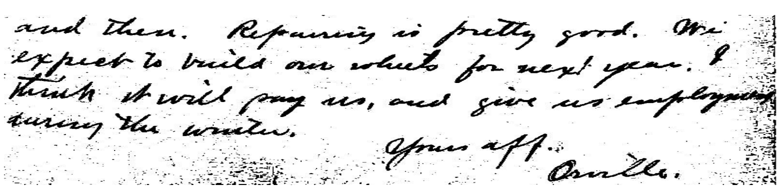
(b)

Optimized Structuring Element
0 1 0
0 1 1
1 1 1

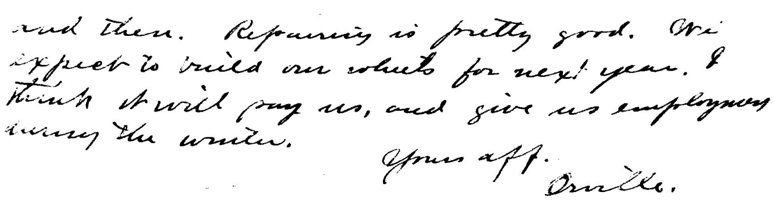
(g)



(c)



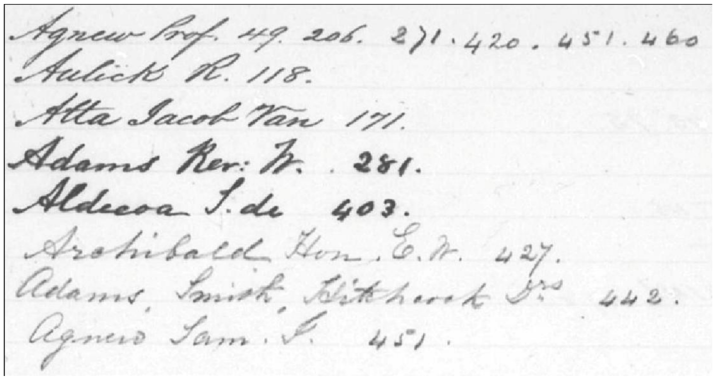
(d)



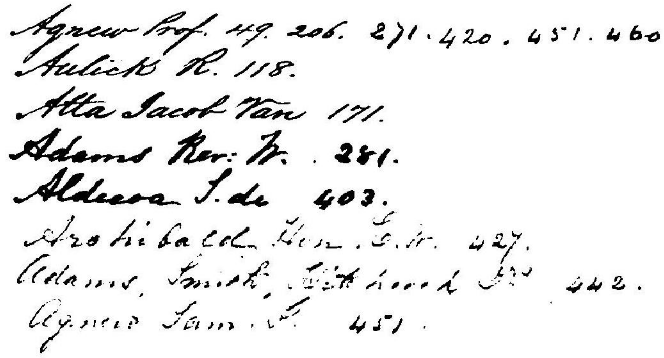
(e)

Fig 5: (a) Original, (b) bABC restored, (c) Ground Truth, (d) Niblack, (e) Sauvola *et al.*, (f) Otsu, (g) Optimized *strel*

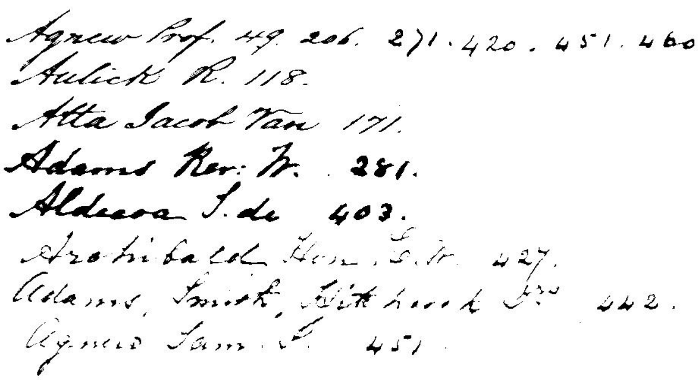
Result 2: Test Image 2



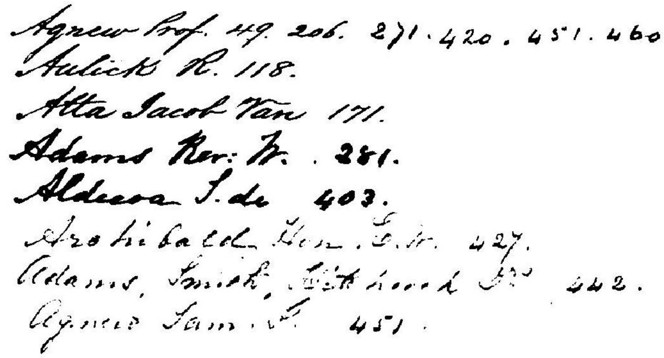
(a)



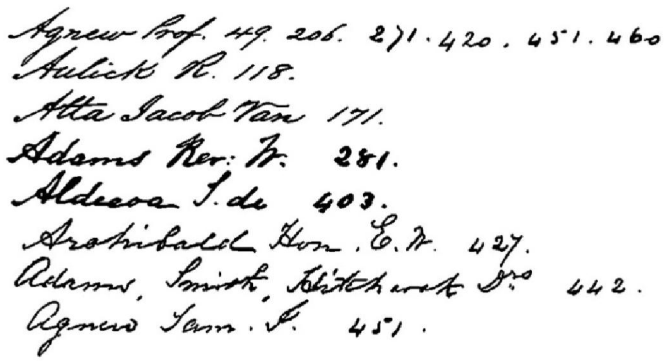
(e)



(b)



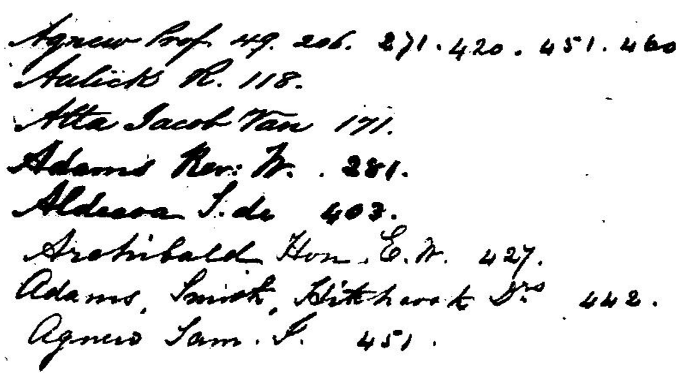
(f)



(c)

Optimized Structuring Element
1 1 0 1 1 1 0 0 1

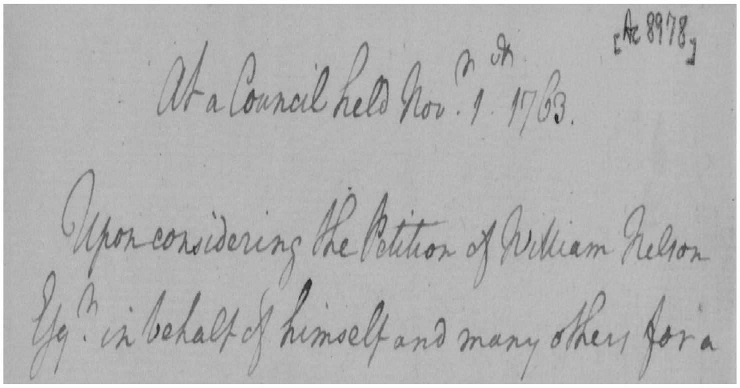
(g)



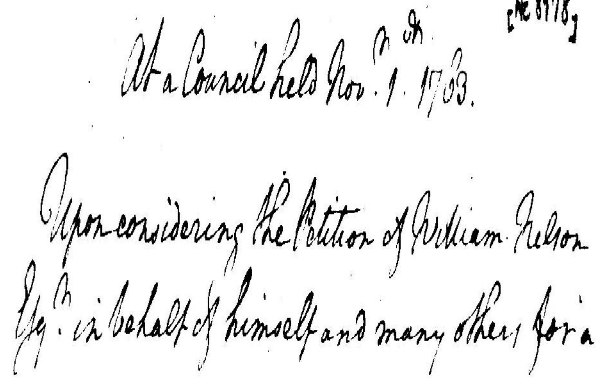
(d)

Fig 6: (a) Original, (b) bABC restored, (c) Ground Truth, (d) Niblack, (e) Sauvola *et al.*, (f) Otsu, (g) Optimized *strel*

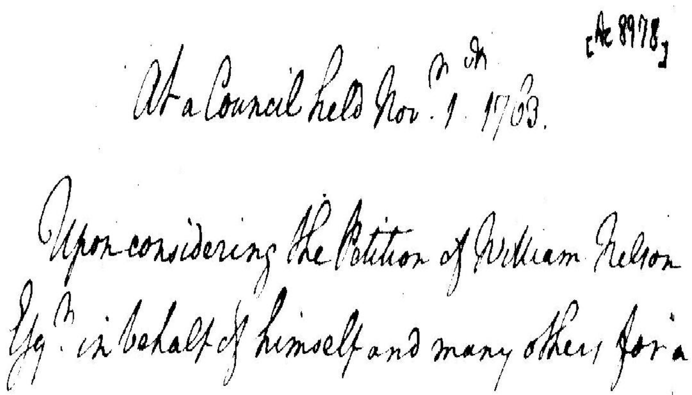
Result 3: Test Image 3



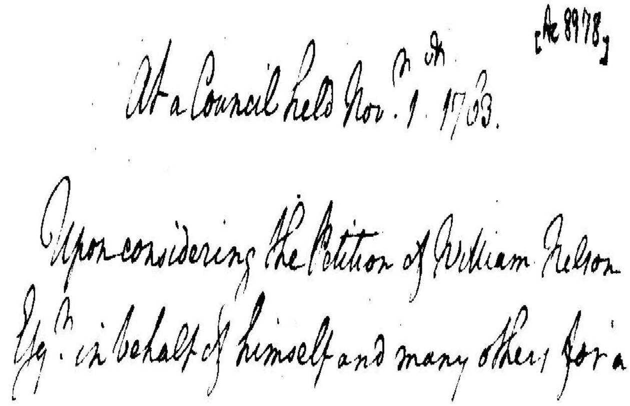
(a)



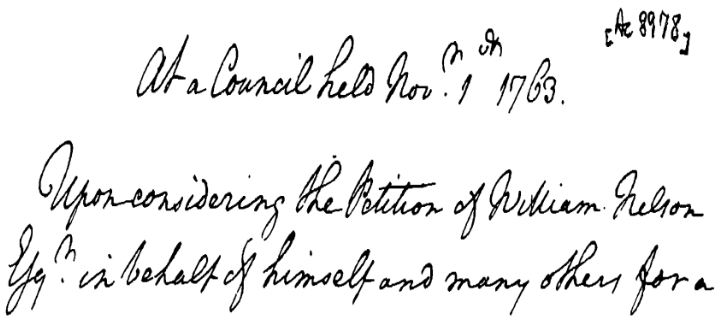
(e)



(b)



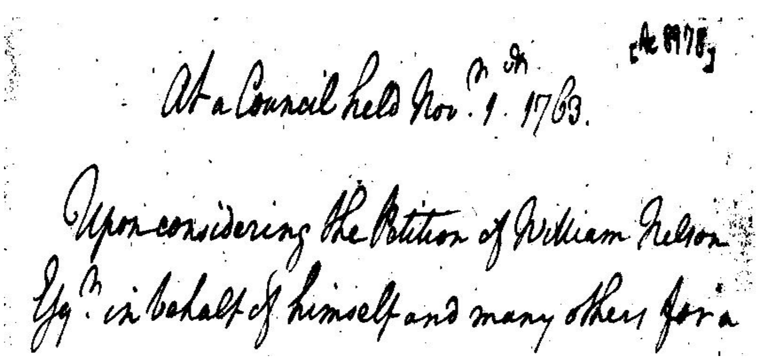
(f)



(c)

Optimized Structuring Element		
1	1	0
0	0	0
1	1	1

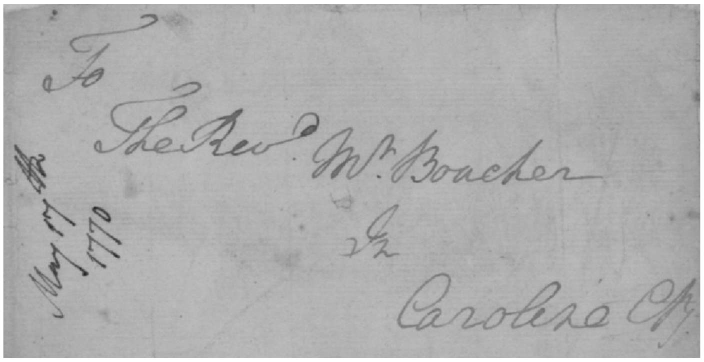
(g)



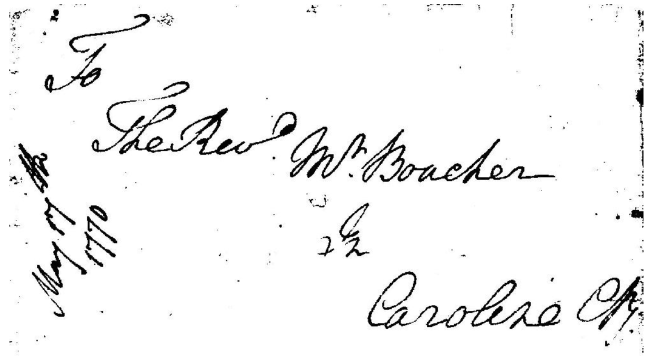
(d)

Fig 7: (a) Original, (b) bABC restored, (c) Ground Truth, (d) Niblack, (e) Sauvola *et al.*, (f) Otsu, (g) Optimized *strel*

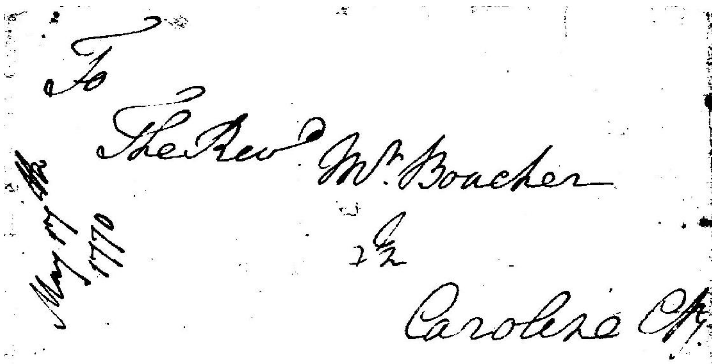
Result 4: Test Image 4



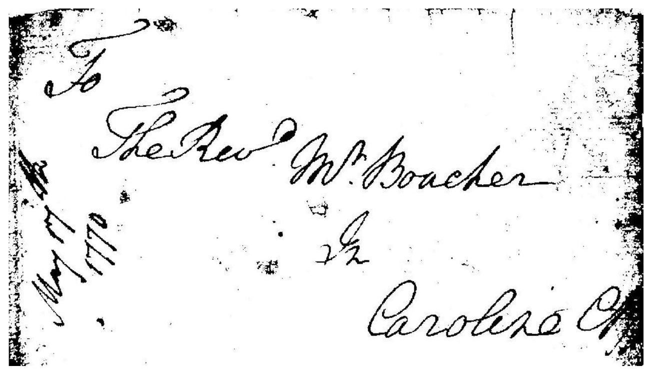
(a)



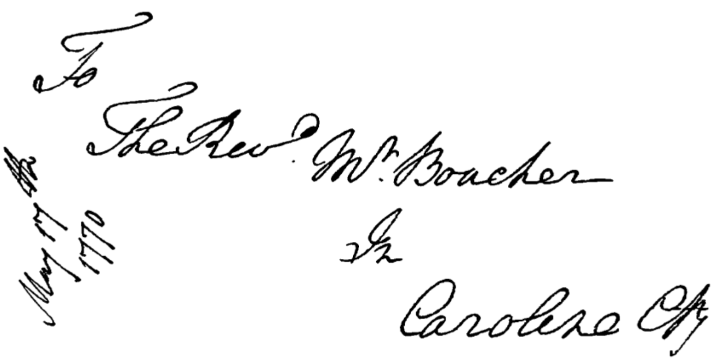
(e)



(b)



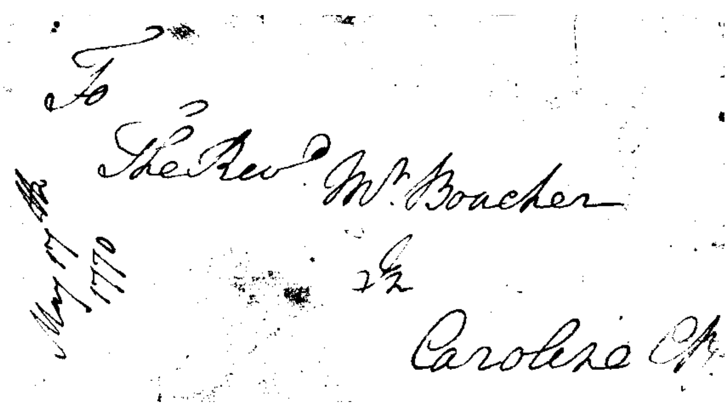
(f)



(c)

Optimized Structuring Element		
0	1	0
1	1	0
1	1	0

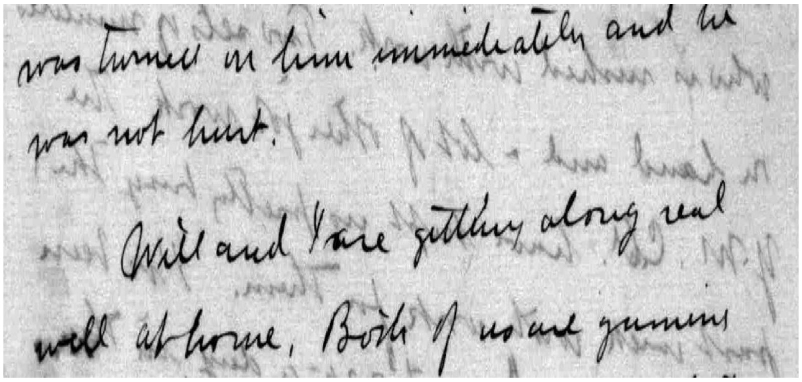
(g)



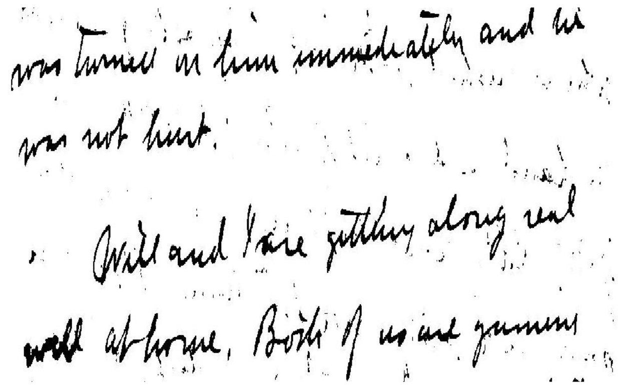
(d)

Fig 8: (a) Original, (b) bABC restored, (c) Ground Truth, (d) Niblack, (e) Sauvola *et al.*, (f) Otsu, (g) Optimized *strel*

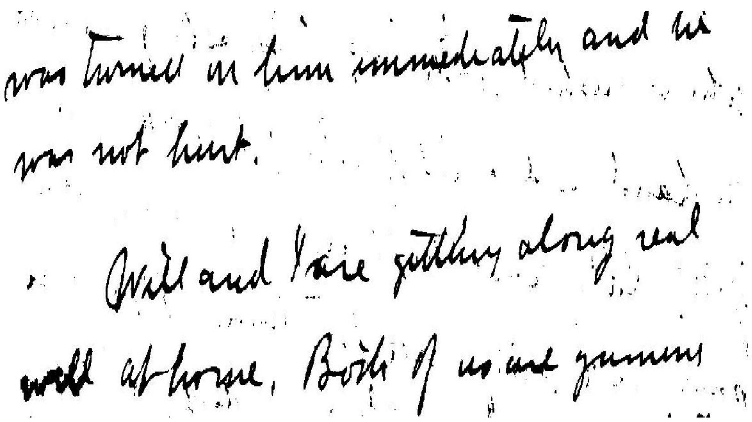
Result 5: Test Image 5



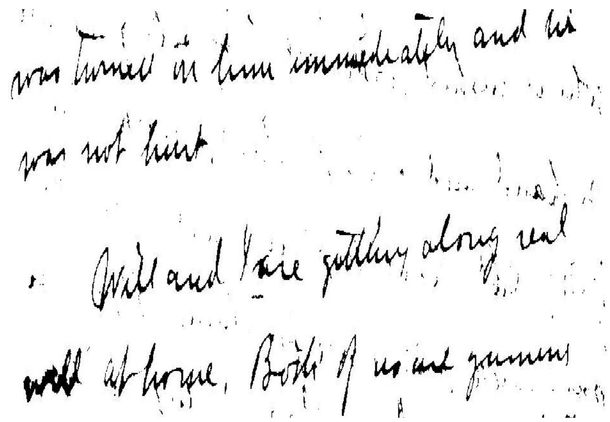
(a)



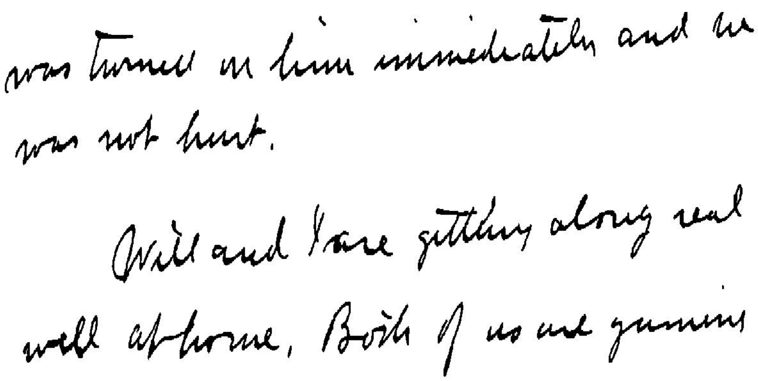
(e)



(b)



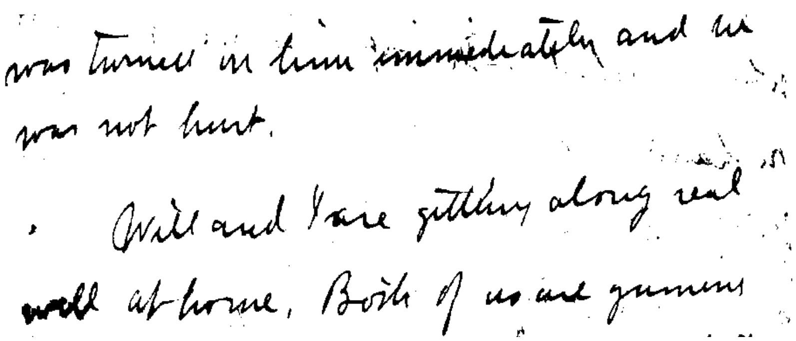
(f)



(c)

Optimized Structuring Element
0 1 0
0 1 1
1 1 1

(g)



(d)

Fig 9: (a) Original, (b) bABC restored, (c) Ground Truth, (d) Niblack, (e) Sauvola *et al.*, (f) Otsu, (g) Optimized *strel*

From our technique we can see visually that we got less background noise in our restored images. The visual comparison between the results of different algorithm and the optimized method was further performed using psycho-visual testing. This testing included the above test images and the results were subjected to a panel of 10 viewers with a fixed viewing distance. The viewers were asked to score the results in a scale ranging from 1 to 5, where 1 indicates poor and 5 indicates the best. The algorithm of generating the results was kept unknown to the viewers. The results were also presented randomly to avoid possible biasness in scoring. The viewers were requested to judge the results against the provided ground truth images in terms of clarity and presence of background noise. The scores of each of the techniques applied on individual images are tabulated in the following table as Table 5.1.

Table 5.1: Scores of each of the processes implemented in different test images.

Image Name	**Niblack**	**Sauvola *et al.***	**Otsu**	**Presented Method**
Test Image 1	4	3.75	4	4
Test Image 2	4.25	4	3.75	4
Test Image 3	4	4.5	4.25	5
Test Image 4	3.75	4.25	5	4.5
Test Image 5	4.5	5	4	4.5
Average Score	**4.1**	**4.3**	**4.2**	**4.4**

From the results we can see that our presented method shows more average score of the images than that of Niblack's, Otsu's and Sauvola *et al.*'s method. Also, it gives visually almost close results with the ground truth images. The optimized structuring elements formed by the bABC method are the one that produces better result of restoration. From the results, we can also visually observe that the presented bABC method obtained an obviously remarkable performance. Hence we got more efficient result by the application of binary Artificial Bee Colony (bABC) algorithm. Thus, we can say that our approach of restoration can be a possible competitive solution for restoring the degraded document images.

5.5. Conclusions

Since the original ABC algorithm cannot be directly applied to solve discrete problems, this thesis adopted a discrete artificial bee algorithm – bABC (binary Artificial Bee Colony), which works in binary space, while still maintains the major characteristics of the original ABC's expression. The presented method successfully demises various kinds of degraded documents enhancing textures with clear background. The

Binary Artificial Bee Colony (bABC) algorithm is used for optimizing the structuring element for providing the best solution of degraded document image restoration. From the simulation results, it is concluded that the performance of the presented algorithm is better than the existing techniques and hence gives suitable results in restoring the degraded document images.

5.6. References

[1] Iyengar, N.G.R and Gupta S.K "Programming methods in structural design",Edward Arnold Pub.Ltd.U.K.1980.

[2] Ali Hadidi, Sina Kazemzadeh Azad and Saeid Kazemzadeh Azad," Structural optimization using artificial bee colony algorithm" in 2nd International Conference on Engineering Optimization September 6 - 9, 2010, Lisbon, Portugal.

[3] Nick Efford. Digital Image Processing: A Practical Introduction Using JavaTM. Pearson Education, 2000.

[4] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", Prentice Hall, New Jersey 07458, pp. 523-531.

[5] Liu Wei, Chen Hanning, "BABC: A Binary Version of Artificial Bee Colony Algorithm for Discrete Optimization" in International Journal of Advancements in Computing Technology4(14):307-314 · August 2012 with 8 Reads.

[6] Arpitam Chatterjee, Bipan Tudu, Kanai Ch. Paul," Binary grayscale halftone pattern generation using binary artificial bee colony (bABC)",in SIViP DOI 10.1007/s11760-012-0388-z, Received: 26 April 2012 / Revised: 19 September 2012 / Accepted: 25 September 2012 © Springer-Verlag London 2012.

[7] Niblack, W.: "An introduction to digital image processing" (Prentice- Hall, Englewood Cliffs, NJ, 1986), pp. 115–116.

[8] J. Sauvola *et al.*, M. PietikaKinen, "Adaptive document image binarization", Machine Vision and Media Processing Group, Infotech Oulu, University of Oulu, 21 January 1999.

[9] N. Otsu, A thresholding selection method from gray-scale histogram, IEEE Transactions on System, Man, and Cybernetics 9 (1979) 62–66.

CHAPTER 6

CONCLUDING REMARKS

6.1. Introduction

Historical documents that are degraded are needed to be restored to preserve the heritage of the country. Extensive study has been carried out and we have come to know that many works has been done in this field. So, we have implemented the binary Artificial Bee Colony (bABC) technique to optimize the structuring element for restoring the degraded document images. Results have shown that bABC gives better performance and thus it can be a potential method of restoration.

6.2. Major Findings

The major findings of this work can be listed as below

- ❖ Optimization of structuring elements and using them towards restoration of degraded document images may be a potential alternative to the global or local thresholding based techniques that are normally used for this purpose.
- ❖ The binary formation of structuring element drives the choice to use the optimization that can work in binary domain.
- ❖ Swarm intelligence based search techniques may be used for this purpose as observed by the successful results obtained using bABC.
- ❖ The use of bABC optimized structuring element results in improved restoration results as compared to the standard techniques, namely, Niblack's, Sauvola *et al.*'s and Otsu's method.

6.3. Future scope of work

Some of the possible directions in which the presented work can further be extended are listed below;

- ❖ Since the objective function comprises of multiple parameters multi- objective optimization methods can be adopted.
- ❖ A more robust objective function that is contextually driven can be developed and used for optimization.
- ❖ This can be further enhanced by extending the scope of work area to suitability in restoration of color images.
- ❖ The presented work can be further implemented by other techniques in search heuristics.

6.4. Conclusion

Document restoration is an important objective for various degraded historical documents that preserve many valuable resources. The presented work is established to restore such heritages in our country. First of all an extensive review on different document restoration methods was conducted. The techniques reveal the scope of employing search based optimization techniques for the stated task. The presented method explores the scope of bABC restoration purpose by means of optimizing the structuring element. Randomly generated structuring elements were applied in document images using different parameters and were compared. The properties and behavior of the results were critically studied. Results have shown that bABC algorithm can be used for generating optimized structuring element towards degraded document image restoration, and thus found to be a suitable alternative method of restoration. Simulation results using psychovisual techniques showed that the proposed bABC method is capable of restoring degraded document images.

APPENDIX

1. Programme to generate established technique:

```
  clc;
 img=imread('../images/H06.jpg');
 %----------------Normal Image to Grayscale Conversion-------

 img=rgb2gray(img);
 figure('Name','Input Image'),imshow(img);

 %----------------Gaussian Filter to remove the noise---------

 G = imnoise(img,'Gaussian',0.04,0.003);
 figure('Name','Gaussian Filtered Image'), imshow(G);

 %----------------Dilation to estimate background-------------------

  s=[1   1   0
  0   0   0
  1   1   1];
  d=imdilate(G,s);
  figure('Name','Dilated Image'), imshow(d);

 %----------------Estimated background subtraction-----------------

[m n]=size(d);
bs=[m n];
    for i=1:m
        for j=1:n
            if(d(i,j)>= 210)
                bs(i,j)=1;
            else
```

```
                bs(i,j)=0;
            end
        end
    end
  figure('Name','Background Subtraction'),imshow(bs);
```

2. Programme of Niblack's method of Restoration:

```
clc;
clear;
x=imread('../images/HT05.jpg');
size(x);
% img=rgb2gray(x);
% [p q]=size(img);
% x=imresize(img,[500 800]);
x=imresize(x,[500 800]);
figure;
imshow(x);
title('original image');

z=rgb2hsv(x);       %extract the value part of hsv plane
v=z(:,:,3);
v=imadjust(v);
m = mean(v(:))
s=std(v(:))
k=-.4;
value=m+ k*s;
temp=v;
for p=1:1:500
    for q=1:1:800
        pixel=temp(p,q);
        if(pixel>value)
            temp(p,q)=1;
        else
            temp(p,q)=0;
```

```
        end
    end
end
figure;
imshow(temp);
title('result by niblack');
imwrite(temp,'E:\D drive\M.E\3rd and 4th Sem\Thesis\Thesis Documentation\result images\Niblack.jpg');
```

3. Programme of Sauvola *et al.* 's method of Restoration:

```
            clc;
         img=imread('../images/HT05.jpg');

         %--------------Normal Image to Grayscale Conversion-------

         img=rgb2gray(img);
         figure('Name','Input Image'),imshow(img);
         [i j]=size(img);

         level = graythresh(img);
         gt = im2bw(img,level);

         %figure('Name','Global Thresholding(OTSU)'),imshow(gt);

        %---------------Sauvola et al. Threshold-------------------
         %pixel = ( pixel > mean * ( 1 + k * ( standard_deviation / r - 1 ) ) ) ? object : background

         m=mean(gt);
         sd=std(gt);
         k=0.5;
         r=128;
         value=m.*(1+k.*(sd./r-1));
         [p q]=size(gt);
         st=[p q];
        for i=1:p
```

```
   for j=1:q
       pixel=gt(i,j);
       st(i,j)=gt(i,j);
       if(pixel>value)
           st(p,q)=1;
       else
           st(p,q)=0;
       end
    end
   end

 figure('Name','Sauvola et al.  Threshold'),imshow(st);
 imwrite(st,'E:\D drive\M.E\3rd and 4th Sem\Thesis\Thesis Documentation\result images\Sauvola et al. .jpg');
```

4. Programme of Otsu's method of Restoration:

```
   clc;
   img=imread('../images/HT05.jpg');

   %--------------Normal Image to Grayscale Conversion-------

   img=rgb2gray(img);
   figure('Name','Input Image'),imshow(img);
   [i j]=size(img);

%    %----------------Global Thresholding(OTSU) for modification-----------

  level = graythresh(img);
  gt = im2bw(img,level);
  figure('Name','Global Thresholding(OTSU)'),imshow(gt);
  imwrite(gt,'E:\D drive\M.E\3rd and 4th Sem\Thesis\Thesis Documentation\result images\Otsu.jpg');
```

5. Programme of bABC's method of Restoration:

```
clc;
%----------------Step 1: Initialization of random solutions N. Here N=10.-------------

initial(:,:,1)=[1    1      0
                1    1      1
                0    0      1];
initial(:,:,2)=[1     1     0
                0     0     0
                1     1     1];
initial(:,:,3)=[1     0     1
                1     1     1
                1     1     1];
initial(:,:,4)=[0     1     0
                1     0     0
                0     0     1];
initial(:,:,5)=[0     1     0
                1     1     0
                1     1     0];
initial(:,:,6)=[1     0     0
                0     1     0
                1     0     1];
initial(:,:,7)=[0     1     0
                0     1     1
                1     1     1];
initial(:,:,8)=[0     0     1
                1     1     0
                0     1     1];
initial(:,:,9)=[0     0     1
                0     1     0
                0     0     1];
```

```
initial(:,:,10)=[0     1     0
                1     0     1
                1     0     0];

 [a b c]=size(initial);

 %----------------Step 2: Initialization of Employed Bees.Half of N,i.e SN=N/2;SN=5.-------------

 n=c/2;% Numbers to be extracted
 N = c;    % Numbers from 1 to N will be permuted
 employedB = randperm(N);    % Permute numbers between 1 and N
 employedB = employedB(1:n);
 [m y]=size(employedB);

 for i=1:y
     emp=employedB(i);
     s=initial(:,:,emp);
     stele(i)=abc(s);%objective function evaluation
 end
 original_stele=stele;
 %highest_stele=max(stele);
 %s  --- employedB structuring element
 %[m n p]=size(s);

 %----------------Step 3,4 & 5: Generation of new solutions then Selection by onlooker bees & Generation of
new solutions by the onlooker bees-------------

 r=-1;
 R=1;
 SN=N/2;
 for i=1:y
     emp(i)=employedB(i);
 end
 x=initial(:,:,emp); %employeed bee associated with random solutions
```

```
sum_ol=0;
for j=1:y
     num_ol(j)=round((stele(j)/sum(stele))*SN); %step 4 probabilistic function
end
for j=1:y
     sum_ol=sum_ol+num_ol(j);
end

%checking if sum_ol=5 or not

if(sum_ol>SN || sum_ol<SN)
    while(sum_ol>SN || sum_ol<SN)
    for j=1:y
       num_ol(j)=round((stele(j)/sum(stele))*SN); %step 4 probabilistic function
    end
    sum_ol=0;
    for j=1:y
       sum_ol=sum_ol+num_ol(j);
    end
    if(sum_ol==SN)
        break;
    end
    end
end
%num_ol  %-- number of onlooker bees

for i=1:SN
    for j=1:num_ol(i)
       k=round(R+(SN-R).*rand(1,1)) %k is a randomly selected numberin {1, 2, . . . . . . , SN}
       nrand(:,:,i)=r + (R-r).*rand(1,1); %n is a random number generated in [-1,1];
       v(:,:,i)=fix(x(:,:,i)+nrand(:,:,i)*(x(:,:,i)-x(k)));  %step 3 generation of new solutions
       newstele(i)=abc(v); %step5-- generation of new solutions by onlooker bees
    end
end
```

```
%iterating 20 times

for iteration=1:20
  for i=1:SN
   if(newstele(i)>stele(i))
       max(i)=newstele(i);
       olstele(:,:,i)=v(:,:,i);
   else
       max(i)=stele(i);
       olstele(:,:,i)=x(:,:,i);
   end
  end
end

stele=max;

for i=1:SN
   if(stele(i)>original_stele(i))
       imax(i)=stele(i);
   else
       imax(i)=original_stele(i);
   end
end

%finding number of scout bees-- bees with no improvement result

count=0;
for i=1:SN
   if(imax(i)<=original_stele(i))
       count=count+1;
       scout(count)=i;
   end
end
scoutBees=employedB(scout);
%count %--no of scout bees
```

```
%scout %-- scout bees
%max %--better bee solution
%olstele %step 5 better solution of onlooker bee
%[e f g]=size(olstele)

%------------Step 6: Generation of new solutions by scout bees-----------

 lim1=0;
 lim2=1;
 rand_num=lim1 + (lim2-lim1).*rand(1,1);
 scoutB = randperm(count);
 scoutB = scoutB(1:count);
 scoutB = round(1+rand_num*(scoutB(1:count)));
 for i=1:count
    sb=scoutB(i);
    b=initial(:,:,sb);
    bstele(:,:,i)=b;
    scoutstele(i)=abc(b);
 end
 %scoutstele %---scout bee solutions
 %[p q l]=size(scoutstele)
 for i=1:count
    imax(scout(i))=scoutstele(i);
 end
 %imax  %---solutions by scout bees

%------------Step 7: Selection of the optimized solution-----------------

maximum=imax(1);
for i=1:SN
   if(max>imax(i))
      optimal=maximum;
      optimizedstele(:,:,i)=bstele(:,:,i);
   else
      optimal=imax(i);
```

```
        optimizedstele(:,:,i)=olstele(:,:,i);
    end
    maximum=optimal;
end
for opt=1:SN
    if(maximum==imax(opt))
        break;
    end
end
optimizedstele(:,:,opt);
% optimizedstele  %--optimized solution
%[u v w]=size(optimizedstele)

%------------Using Optimized structuring element in abc for restoration-----

restored=abcOutput(optimizedstele);
fprintf('Optimized Structuring Element::\n')
disp(optimizedstele(:,:,opt));
fprintf('Objective Function Value:: %f', restored);
```