TECHNOLOGICALLY ASSISTED READING FRAMEWORK FOR VISUALLY CHALLENGED

A Thesis Submitted to
Babu Banarasi Das University
for the

Degree of Doctor of Philosophy

in

Computer Science & Engineering

by

Avinash Verma

Under the Supervision of

Dr. Deepak Kumar Singh

Department of Computer Science & Engineering
School of Engineering

Babu Banarasi Das University
Lucknow 226028 (U.P.), India

June, 2017

Certificate of supervisor

This to certify that the thesis, titled ‘Technologically Assisted Reading Framework for Visually Challenged’ submitted by Avinash Verma for the award of Degree of Doctor of Philosophy by Babu Banarasi Das University, Lucknow is a record of authentic work carried out by him under my supervision. To the best of my knowledge, the matter embodied in this thesis is the original work of the candidate and has not been submitted elsewhere for the award of any other degree or diploma.

Dr. Deepak Kumar Singh

Associate Professor

Computer Science & Engineering

Integral University

Kursi Road, Lucknow.

Date:

Declaration by the Candidate

I, here by, declare that the work presented in the thesis, entitled ‘Technologically Assisted Reading Framework for Visually Challenged’ in fulfillment of the requirements for the award of Degree of Doctor of Philosophy of Babu Banarasi Das University, Lucknow is an authentic record of my own research work carried out under the supervision of Dr. Deepak Kumar Singh.

I also declare that the work embodied in the present thesis is my original work and has not been submitted by me for any other Degree or Diploma of any university or institution.

Date

Avinash Verma

ACKNOWLEDGEMENTS

This dissertation is an individual work that has been benefited in various ways form different people. I think that it would be simple to them but will not be easy to thank them enough.

I would like to express my acknowledgements to all those who have been there for me throughout this research work. I would like to express my gratitude towards almighty God the mother father of this universe for giving me strength and positive thinking to work on this research topic.

I would like extend my gratitude towards my enthusiastic and supportive research guide Dr. Deepak Kumar Singh for encouragement and motivation throughout the work. With his deep knowledge of the subject he has time and again helped and given his valuable suggestion which have helped to improve my work. I would also like to extend my gratitude towards the RDC committee for providing there valuable comments and suggestion in order to improve and motivate me in my work.

I would also like to thank my parents who sacrificed their comfort for my betterment and for being my strength throughout this work. I would like to thank my wife Nisha and my daughter Shivanshi for their unconditional love and for always being there for me whenever I required them.

I would like to extend my gratitude to my colleague and friends for their belief and support throughout the work.

Avinash Verma

PREFACE

Technologically Assisted reading framework for visually challenged is assistive technology for visually challenged person that can help then in reading the printed books, typed documents, journals, magazines and newspapers, and computer displays of emails, web pages, and railway/airline information and bookings etc like normal persons. It can provide the visually challenged person the capability to read with the help of Optical Character recognition and Text to Speech Conversion technology in which the visually challenged person captures the image of the surrounding having text. This captured image is the input to the system and then this image is passed to an Optical Character Recognition which recognizes the text information present in the input image and returns an output ASCII or the Text file. The text file is then passed to a Text to Speech converter that converts the text file into corresponding speech signals using various mechanisms and this speech file is communicated to the visually challenged person. Hence the visually challenged person instead of reading gets to hear the textual information captured by him. While working on the framework we came across the issues that can affect the accuracy of the framework or even can lead to the communication of inappropriate textual information. These issues are the presence of Blur and different lighting conditions or the noise in the captured image. As the image is captured by visually challenged person while on the move and the person not being an expert in capturing image, the image is bound to be affected with some blur and different contrast effect or noise. In order to overcome this problem we have introduced two extra steps in the framework that are deblurring step and preprocessing step before Optical Character Recognition. The deblurring step is responsible for the restoring the blurred input image so that it can give an accurate recognition otherwise blurred image can lead to false or no recognition of the textual information at all. This can affect the efficiency of the overall framework. For deblurring we have developed new deblurring algorithm using blind Deconvolution technique by iterative estimation of Point Spread Function and evaluation using Average Word Confidence metrics developed by us. After the deblurring process we get the restored deblurred image but due to the Deconvolution process the noise gets amplified hence the resultant deblurred image need to be preprocessed before the next stage Optical Character recognition. In the preprocessing step we perform the morphological operations such as thresholding and Binarization in order to convert the image into gray scale image and then Binarization converts the image input binary Image. The binary image thus obtained is send to an Optical Character Recognition for the recognition of text. Optical Character Recognition segments the preprocessed image into text and non-text boxes and then the text in the text boxes is used further. The text in the text boxes is segment into line then into words and lastly into individual character. After Segmentation next step is the recognition step in which the individual character is recognized after segmentation and ASCII or text file is generated as the output of the Optical Character recognition. Text to Speech engine performs the preprocessing operations on the text file. So that the text is corrected for the spelling mistakes, abbreviations and acronyms are handled in text normalization. Various Morphological operations of text to speech are performed for proper pronunciation of word in Linguistic and syntactic analysis. After preprocessing step the next step is speech generation which involves various types of intermediate conversion. Last step is the voice rendering step to get the speech output that is communicated to the visually challenged person in order to get the essence of the textual information captured by him.

During the development of technologically assisted reading framework there have been several findings that are as follows

- Development of Reading framework for visually challenged using Optical Character Recognition and Text to Speech synthesis.
- Finding the issues of blur and different lighting condition that affected the accuracy and the efficiency of the overall framework.
- Developed new algorithm for the text deblurring using Blind image Deconvolution technique using the single blurred input image without having any information about the Point Spread Function or the Blurring process.
- Improved the existing model by introducing two steps in the existing model that were deblurring step and preprocessing step before Optical character recognition.
- Development of Robust reading framework for visually challenged person that can handle the ideal as well as unexpected input such as blurred, skewed and noisy input image.

Based on the above findings this reading framework finds its application in the reading for visually challenged in which they can read printed books, typed documents, journals, magazines and newspapers, warning signs, notice boards and computer displays of emails, web pages, and railway/airline information and bookings etc like normal persons.

The directions for the future work are development of new technique for blur removal Computational Neural Networks (CNN) and then integrate the overall system. As a future work we would like to develop the assistive technology for the vision, writing and movement of the visually challenged person.

List of Research paper containing the result of the thesis:

1. Avinash Verma, Deepak Kumar Singh, 2017. Text Deblurring Using OCR Word Confidence. Published by Internation Journel of Image, Graphics and Signal Processing (IJIGSP) Vol.9 No.1, pp.33-40, 2017 DOI: 10.5815/ijigsp.2017.01.05.
2. Avinash Verma, Deepak Kumar Singh, Nitesh Kumar Singh, 2017. Review on Assistive Reading framework for Visually Challenged. Published in Computational Intelligence in Data Mining pp 533-541 part of Advances in Intelligent Systems and Computing (AISC), vol.556. SPRINGER, Singapore. DOI: 10.1007/978-981-10-3874-7_50.
3. Dr. Deepak Kumar Singh, Avinash Verma, 2014. Reading Framework for Visually Challenged. Presented and published in the proceedings of DigiGov-2014 National Conference on DIGITAL INDIA eMpowering by Computer Society of INDIA (CSI) Lucknow Chapter.
4. Avinash Verma, Dr. Deepak Kumar Singh, 2017. Robust Assistive Reading Framework for Visually Challenged. Communicated to International Journal of Image, Graphics and Signal Processing (IJIGSP).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The World Blind Union (WBU) [1] on behalf of blind and partially sighted persons of the world estimates 258 million blind and visually challenged persons in about 190 countries, of this about one third of them are in India. Conventionally it is accepted that the visually challenged people, especially the born blind, can do very limited types of jobs. The reasons for this restricted range of jobs is twofold – firstly, society finds that they are not suitable for certain types of jobs because of the need to protect the worker and the people around the person, the second reason being that the visually challenged could acquire very limited knowledge and skills. Most of the effort of the visually challenged individual and society goes to make out that the person lives in the world comfortably with the help of devices like Braille, canes and dogs etc. The support cost for the visually challenged is considerably high. The existing Assistive technologies [2][3] have helped the visually challenged in different ways at different point in time. However most of these technologies tried to provide minimal support so that the challenged person is able to survive rather than lead a fuller life. However the current ICT technologies have the means and capabilities of restoring the fullness to life by making the challenged person to live life like any one of us. The moment we provide the capabilities that would make the visually challenged to behave like normal human being, the set of problems associated with the challenged would disappear. The person can learn any discipline or activity and contribute to the society and have the same quality of life as anyone else. The dream of providing the near-normal capabilities to the visually challenged through the use of Information Technology is worth taking up with all the associated technological and cognitive challenges. The technologically assisted capabilities [4][5][6][7] should create equal opportunities to the visually challenged in their study and work in their chosen areas, and help them to have better quality of life. The technologically assisted capabilities include reading, seeing, writing, and moving around in a predefined environments like schools, laboratories, roads, buses, offices, railway stations, airports etc, which the visually challenged person comes across for learning and carrying out the work in their daily life.

1.1 CHALLENGES FACED BY THE VISUALLY CHALLENGED

- Reading from conventional books, documents, especially when they contain diagrams, graphs, tables etc. This is important if the visually challenged have to read books and documents to acquire domain knowledge in sciences, social sciences, engineering and medicine etc. This opportunity does not exist for them.
- Seeing and interpreting everyday objects, images, warnings, etc. within the house and outside (like tables, chairs, books, etc. within the houses; obstacles, pavements, bus stops, ditches on the road; warning signs on the road, traffic

signal status; numbers or names on the doors or houses, road and place names; train arrival, departure information displays; and objects that are normally found in the house, super markets, and in work environments, This capability is essential for the person to move around and understand the surroundings without the help of others. This will enhance the mobility of the person and self-confidence.

- Writing notes, documents, reports etc. without any regard to the size and shape of the paper. The capability includes electronic recording (e-writing) and manual writing.
- Moving and working around in the houses, workspaces, and public places etc,. The work may include manual work such as working in garden, using tools such as screw drivers, saws etc., fixing electrical connections, resetting MCBs etc, white collar work such as working with files, and intellectual work such as reading, working in the labs, participating in the discussions, making presentations, teaching etc.
- Learning about new objects such as tower bolt, new concepts like height, color which can be perceived through vision (A visually challenged has to imagine a 12 m height tree or how a tower bolt looks like and how it works) is a challenging job. The strategy of converting the visual information to a stream of audio information requires formation of concepts and their representation through words. Considerable cognitive challenges are to be solved with so that the communication is effective. The other difficulty is that the visual information is essentially parallel in nature (when we see an apple we get the information on the

size, color, shape, texture, freshness simultaneously) and the audio information or any other symbolic information is sequential in nature and the challenge lies in developing a protocol for this communication.

Once the visually challenged person overcomes these limitations, he/she would find no difficulty in learning what he/she wants and working where he/she wants. Thus it would dramatically improve the quality of life for this visually challenged group. In this thesis we have worked on the reading challenges faced by visually challenged person and tried to overcome this problem with the help of development of a reading framework for visually challenged which uses the current technologies such as optical character recognitions [2] and text to speech synthesis [2]. With the help of these technologies visually challenged person can capture the image of the text information with surrounding and then the acquired text information will be communicated to him as speech. Hence the visually challenged person will hear whatever text information he captures.

1.2 SCOPE OF THE WORK, APPROACH AND CHALLENGES

The work proposes to take up the following work to provide substantial capabilities to the visually challenged persons. The work is ambitious and needs to address considerable technological and cognitive challenges, and has to integrate diverse existing technologies and the new technologies to be developed. The major elements of the reading framework for visually challenged are as under:

Visually challenged should be able to 'read' normal printed books, typed documents, journals, magazines and newspapers, and computer displays of emails, web pages, and railway/airline information and bookings etc like normal persons. The documents may contain text besides pictures, diagrams, tables, special symbols and yet the person should be able to grasp the essence of these artifacts.

Approach

Technologically assisted reading framework for visually challenged will help the visually challenged person to read normal printed books, typed documents, journals, magazines, newspapers and computer displays of emails, web pages etc like normal persons. It is a system based on image processing and pattern recognition using which visually challenged person carries or wears a portable camera as a digitizing device and uses computer as a processing device [11]. The camera captures the image of the textual information to be read along with the surrounding data. An Optical Character Recognition (OCR) system segregates the image into text and non-text boxes and then the OCR converts the text from the text boxes to ASCII or Text file. The Text file is converted to voice by a Text to Speech (TTS) converter. Thus visually challenged person would 'hear' the text information that has been captured. So this technology can help the visually challenged to read the captured textual information independently which will be communicated to him as voice signal output of a Text to Speech converter.

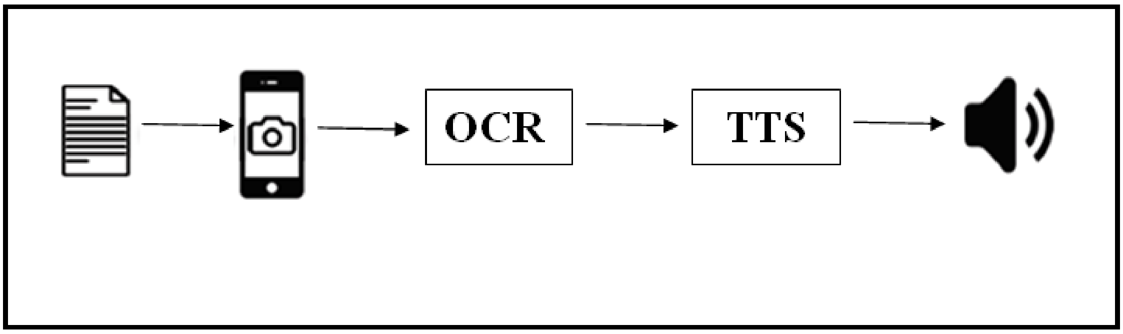


Figure 1.1. Overall System Overview

Challenges

The above process involves the following tasks:

- Digitization of the page.
- Appropriate framing of the page including adaption of different page sizes and autofocus.
- Elimination of the effect of blurring from captured Image.
- Elimination of effects of contrast changes and noise removal.
- Identification of text.
- Conversion of text to voice.
- Integrating the system into a workable and portable system

1.3 ASSISTIVE TECHNOLOGY FOR VISUALLY CHALLENGED

The Assistive Technology for visually challenged is conceived to provide total support to the visually challenged in an integrated manner. The assistance provided by this system should make the challenged person independent of the support from others and allow the person to use conventional resources used by the common man without recourse to special devices like Braille, walking sticks etc. No doubt that all this would be succeeded by the next generation technologies that would interface directly with respective activity centers in the brain, and repair and replace eyes and associated nervous systems. Till those technologies appear on the scene, one could use the currently available assistive technologies to good effect. The major challenges that are faced by the visually challenged are reading from conventional documents, seeing and interpreting images, warnings, announcements, objects in the study and work environments, writing documents, reports etc., and moving around in the houses, workspaces, roads, public places etc,. Once the challenged person overcomes these limitations, he/she would find no difficulty in learning what they want and working where they want. Thus it would dramatically improve the quality of life of this target group.

The broad goal of the work is to develop assistive technologies for the visually challenged that would provide them with the capabilities that are similar to those of the normal people. The objective of the work covers developing the specific technologies of reading in an integrated manner. The ‘reading technology’ assists the visually challenged to ‘read’ normal printed books, typed documents, journals, magazines and newspapers, computer screens, railway/airlines announcement screens (referred to as targets). The information on these screens may contain text, equations, computer programs, boxed text, tables, pictures, special symbols, etc. The information from the targets may simply be read by the technology device or may be read in an interpretive manner depending upon the context.

1.4 OPTICAL CHARACTER RECOGNITION (OCR)

Most of the reading frameworks for visually challenged [11] till date has used Optical Character Recognition in one way or the other. OCR is image processing technology [12] in which we extract the text present in the image and then output it as a Text file. Input to the Optical Character Recognition module is the acquired Image ideally it should be an uncompressed Bitmap Image. With the advent in the technology today we can extract the text from even compressed Images. On the input image we have to perform various pre-processing operations such as binarization [13], blur removal, skew correction, and noise removal step [14] in order to increase the accuracy of the OCR and to prepare the image for further processing. Then text and non-text segmentation [15][16][17] is performed to isolate text information from graphics, segmented text is then segmentation into various lines and is known as Line Segmentation. From each line we segment different words and then these words are segmented as individual characters known as character level segmentation. Character recognition [19] of the individual character is performed and output will be the Text file containing the recognized text.

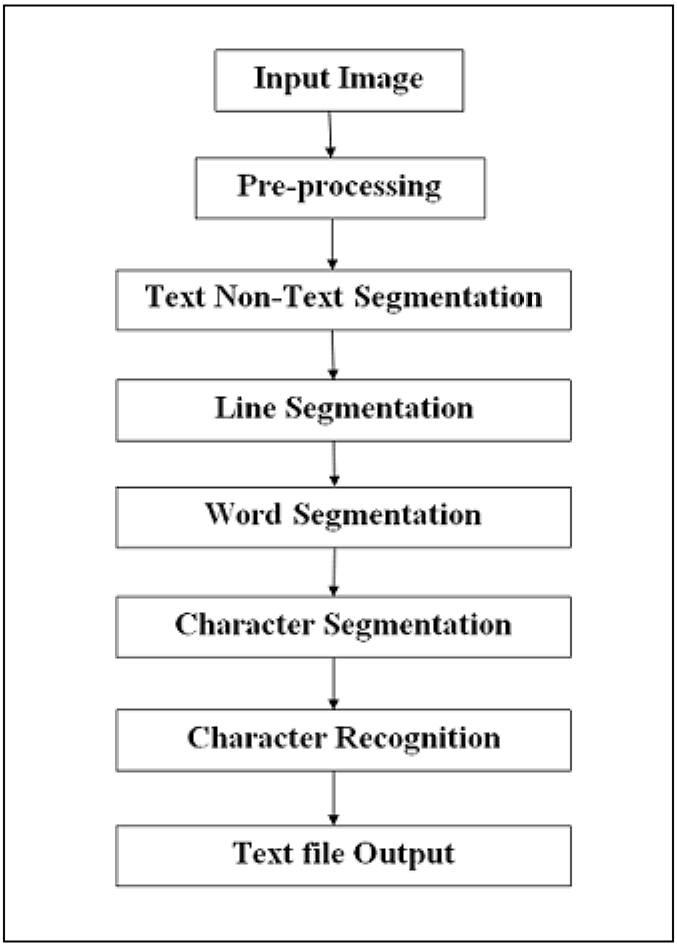


Figure 1.2 Architecture of Conventional Optical Character Recognition.

The main limitation associated with OCR Engine is with the input images affected with any kind of blur or noise. Blur in the image is caused due to the movement of the capturing device or the object to be captured at the time of Image Capture. This problem was not addressed by the previous Reading Framework for visually challenged persons. We have taken this problem into consideration and have come up with one of the efficient cost effective technique to remove the Blur in the image using blind Deconvolution Algorithm with the help of Average word confidence

Metrics of the optical Character recognition. Thus we remove the blur present in the image and then it send for Optical Character Recognition. Thus the recognition rate of the text increases and the overall reliability of the System increase drastically. As in the previous system the blurred input image always lead to the misrecognition and thus the wrongly detected text was communicated as voice to the visually challenged person which is not acceptable in any circumstances.

1.5 TEXT TO SPEECH (TTS) ENGINE

After the Optical character recognitions of the Image we get the text file as output. Now text has to be converted to speech or Braille so that the visually challenged person can understand it. Text to Speech Engine [26][27][28] converts the text output from OCR engine to its corresponding speech output. Text to speech Engine initially works by performing the preprocessing operation on the Input text file. This is done in order to increase the efficiency of the text to speech generation. After this step speech generation is performed for the preprocessed text. The preprocessing step [26] prepares the input text file for further processing the operations like text analysis, text normalization and then translating into a phonetic or some other linguistic representation are performed. During Preprocessing the spell checking is performed this helps in correcting some misrecognized text during the Optical Character Recognition on the text based on the punctuation marks the formatting of the paragraph is done. The abbreviations and acronyms are handled in the Text Normalization [26] which enhances the speech output. This helps in communicating the meaningful speech to the visually challenged person. Morphological operations for the proper pronunciation of the word are performed in the Linguistic analysis followed by syntactic analysis to facilitate in handling ambiguities in the Input text. Now the speech generation process involves various steps such as phonetic analysis that is useful in finding the phone level within the word. Each phone level has the information about the sound tagged with it to be produced. Grapheme to phoneme conversion is the next step based on the dictionary. This is followed by the prosodic analysis [30] which attaches the pitch and the duration information for the speech conversion. Speech synthesis is the last step which involves voice rendering to get the speech from the text to speech Engine.

1.6 THESIS LAYOUT

In the thesis a robust framework has been developed which will communicate the text captured by visually challenged person in the form of speech. This will help him to get the knowledge about the surrounding. The presented framework overcome two problems that were not addressed before, first if the captured image is affected with the blur and second if the image is affected with the any kind of noise then the overall system accuracy was at stake. We have not only introduced new steps in our approach but also came up with the successful solution to this problem.

The thesis starts with introducing the Technologically Assisted reading framework for visually challenged and technologies used in the framework. The challenges faced by the visually challenged person in their day to day life are briefly explained along with the scope of the reading framework for the visually challenged person. The opening chapter also covers the reading framework for visually challenged person and details of Optical Character Recognition and text to speech engine.

Chapter 2 is the literature review on the assistive reading framework for visually challenged this chapter deals with the study of the existing system. It contains the description of various applications proposed for visually challenged along with the review of Optical Character Recognition Algorithms and Text to Speech Engine. In Optical Character Recognition the architecture of Tesseract OCR engine is discussed and in Text to Speech Engine the various TTS Engine developed are discussed along with conclusion.

Chapter 3 deals with the Image degradation and Restoration Model in this blur model and types of blur along with the noise model is discussed in image degradation and classical Image restoration and Blind image restoration techniques are discussed in Image restoration model in detail along its types.

Chapter 4 Text Deblurring using OCR word confidence covers the deblurring process using the OCR word confidence metrics. The related work in text deblurring is described along with proposed algorithm with the estimation of point spread function for motion blur. In the estimation of motion blur both blur length and the blur angle are iteratively estimated based on the value of average word confidence. The result of the algorithm is represented using figures and the table containing the results. Finally the result analysis is followed by conclusion and future work.

Chapter 5 Robust Assistive reading framework for visually challenged in this chapter we proposed and successfully came up with a new framework which is more reliable and efficient than previous system as it contains two more steps than the previous system that are deblurring step and preprocessing step before OCR. This chapter contains the detailed proposed approach algorithm, working of the deblurring step and preprocessing step on the resultant image before OCR. These two steps are followed by the Optical Character Recognition and the Text to Speech Conversion in detail.

Chapter 6 contains the conclusion of the technologically assisted reading framework for visually challenged how it works what are its advantages over the previous work done. In this section we have compared our approach with the previous approach and with the help of the table we have shown how our approach is better than the previous approach. Future works are also described so that in the near future there can be further research in this area of assisted technology for visually challenged person and this could increase the quality of life of the visually challenged person and can make them independent in their life.

CHAPTER 2

REVIEW ON ASSISTIVE READING FRAMEWORK FOR VISUALLY CHALLENGED

Objective of this chapter is to review various approaches used for providing the assistive reading framework for the visually challenged persons. A lot of research is being done for visually challenged to help them in living a quality life. This thesis studies the existing assistive technology for visually challenged and proposes a robust reading framework for visually challenged. The presented reading framework will help the visually challenged person to read normal printed books, typed documents, journals, magazines, newspapers and computer displays of emails, road signs, web pages etc on the go like normal persons. It is a system based on image processing and pattern recognition using which visually challenged person carries or wears a portable camera as a digitizing device and uses a portable computer as a processing device. The camera captures the image of the text to be read along with the relevant image and data. An Optical Character Recognition (OCR) system segregates the image into text and non-text boxes and then the OCR converts the text from the text boxes to ASCII or Text file. The Text file is converted to voice by a Text to Speech (TTS) converter. Thus blind person would 'hear' the text information that has been captured. So this technology can help the visually challenged to read the captured textual information independently which will be communicated as voice signal output of a text to speech converter.

2.1 REVIEW OF READING FRAMEWORK FOR VISUALLY CHALLENGED

Most of the technology proposed for visually challenged person is based on image acquisition and then extraction of the text from the image with the help of Optical character Recognition technology also known OCR. Then text file output of the OCR engine is converted to speech with the help of text to speech (TTS) synthesis technology. Thus the visually challenged person gets to listen whatever the text is being captured by him. The Fig.2.1 shows the details of conventional system.

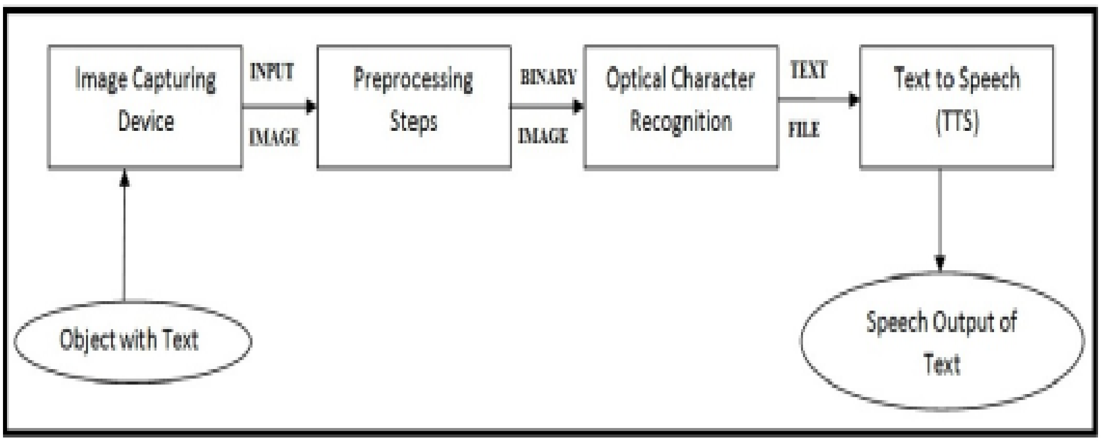


Figure 2.1 Architecture of Reading Framework for Visually Challenged

2.2 IMAGE ACQUISITION BY VISUALLY CHALLENGED

In Image acquisition visually challenged person has to capture the image of document, book, warning sign, computer display, magazine, journals and newspaper this is done with the help of capturing device such as portable mobile camera or any kind of specially prepared device that can capture the image of the document to read. As the Image will be captured by a visually challenged person the image will contain the document text as well as the surrounding. Input image will also be affected with the effect of noise caused due to different lighting condition, effect of shadow, skew and blur. This will make the localization of the text a challenging problem. Various text extraction techniques have been proposed for the text localization in natural scene [16] which can localize the text present in a captured image. The Geometric rectification of captured image [18] will require removing the effects of skew which makes the recognition of text difficult. As in the case discussed in this thesis the image is captured with a portable camera there is lot of possibility that the image captured is affected with the effect of blur. The blur encountered will be the motion blur as the image is captured on the move. Motion blur is caused due to the movement of the scene or the capturing device at the time of image capture. In all the previous proposed system none of the system has taken care of the blur in the acquired image. An extra step is required before the OCR that should take care of the motion blur and various noise removals after the image acquisition process and before the OCR. This can increase the accuracy of the OCR and this will increase the overall efficiency and reliability of the system on the whole.

2.3 PREPROCESSING STEPS INVOLVED BEFORE OCR

As the image is acquired by the visually challenged person it will be associated with a higher degree of complexities such as effect of shadow, different lighting conditions, effect of skew and blur. If the input image is affected by the skew and the blur then taking care of the blur in the image is essentially an important step. Blur in the image is mainly due to the motion of either the object or the capturing device mainly camera at the time of image capture this type of blur is known as motion blur.

$$\mathbf{g}(\mathbf{x},\mathbf{y}) = \mathbf{f}(\mathbf{x},\mathbf{y}) * \mathbf{h}(\mathbf{x},\mathbf{y}) + \mathbf{n}(\mathbf{x},\mathbf{y}) \tag{2.1}$$

In expression (2.1) Blurred Image can be modeled as the convolution of the ideal image $f(x,y)$ with the blur kernel or the Point Spread Function $h(x,y)$ to get the blurred Image $g(x,y)$ with additive noise $n(x,y)$ which is induced at the time of image capture. Motion blur kernel also known as Point Spread Function (PSF) is the function of two parameters which is length of blur L and angle of blur θ. Therefore, motion deblurring tries to estimate the blur kernel $h(x,y)$ and then it performs the deconvolution operation which is the inverse of convolution operation to recovers Original image $f(x,y)$ also known as latent Image from the blurred image. When the text document to be captured translates with a relative velocity V in respect to the camera, the blur length L in pixels is $L = VT_{exposure}$ where, $T_{exposure}$ is the time duration of the exposure. The expression for motion blur is given as,

$$\boldsymbol{h(x,y)} = \begin{cases} \dfrac{\mathbf{1}}{\boldsymbol{L}}, & \boldsymbol{if}\ \mathbf{0} \le |\boldsymbol{x}| \le \boldsymbol{L}\cos\boldsymbol{\theta}\,;\boldsymbol{y} = \boldsymbol{L}\sin\boldsymbol{\theta} \\ \mathbf{0}, & \boldsymbol{otherwise} \end{cases}$$

When the angle of blur θ = 0, it is called horizontal motion blur. In deblurring the original image $f(x, y)$ requires the deconvolution of the PSF $h(x, y)$ with the observed image $g(x, y)$. In most of the cases the PSF is known prior to the deblurring this type of deblurring is done using the classical well known techniques such as inverse filtering, wiener filtering , least square filtering, recursive Kalman filtering are available. The Blur model is covered in detail in chapter 3 along with the different deblurring techniques. If the deblurred image or the resultant image $f(x, y)$ after deblurring operation is affected with the skew then the image is de-skewed by finding the angle of skew and then the image is rotated to get the image which is free from the effect of skew. To make the recognition of the text form the image an effective process using an OCR we require various standard preprocessing steps before applying the OCR engine.

2.4 VARIOUS APPLICATIONS PROPOSED FOR VISUALLY CHALLENGED

In mid 90’s Xerox launched a device called “Reading Edge”, which scans printed materials and read it out load to a user. It also provides a Braille interface so that a blind person may read the content in Braille. The Reading Edge devices incorporate a scanner, speech generation software system, and a Braille keypad for editing. Users could adjust the reading speed and have a choice of nine speaking voices. While the device was a handy aid for a blind person, still its usage required significant effort. A book has to be laid in proper orientation for it to be scanned. Also the unit, due to its size and weight can’t be freely carried. One of the Android applications proposed is R-Map [33] it uses the camera to capture the Image, Optical character Recognition and text to speech engine are used to provide read out loud service. Mobile phone has lower processing power than a desktops or Notebooks. However to provide a processing power that can work in real life environment in mobile phone is difficult. Mobile camera may lead to various problems in the image like Skew, Blur, curved base lines and auto focus mechanism etc. there by causing the best available OCR to fail. Tesseract OCR [41] engine is used which provides text segmentation and recognition. The recognized text is send to Text To Speech (TTS) engine for further text to speech synthesis. Due to limited mobile screen size it is hard for a visually challenge person to take the image of the long printed material. Some Indian researchers has proposed a Reading System for Visually Challenged [34] that uses a scanner to scan the document image which is given as a Input to the OCR module which performs the text and non-text segmentation and recognition and generates a text file as output which is converted to the speech by Text to Speech Module. The difficulty with this is a visually challenged person cannot put the document into the scanner in proper orientation. Chucai Yi, et al. proposed a portable Camera-Based text and product Label reading for hand held Objects for Blind persons [35], this help blind person to read the text label and product packaging from hand held object in their day to day life. A motion based Region of Interest (ROI) [35] is proposed by asking the user to shake the object. This method extracts the moving object from its complex background then text extraction and recognition is performed on the object. The recognized text is outputted to the blind as voice using text to speech mechanism. Trinetra [36] by Patrick E.Lanigan, et al. is a cost-effective assistive technologies developed for visually challenged person to make them independent in their life. The objective of the Trinetra system is quality improvement of the life of visually challenged with the help of different networked devices to support them in navigation, shopping and transportation. Trinetra uses barcode-based solution consisting of combination of off-the-shelf components, like Internet and Bluetooth-enabled mobile phone, text-to-speech software and a portable barcode reader.

Table 2.1 Comparison between different applications proposed for Visually Challenged

Application Name	Function	Technology Used	Disadvantages
Reading Edge	Scans printed materials and read it	Scanner, Speech	A book has to be laid in proper orientation for it to be

	out load to a user	generation software system	scanned. Also the unit, due to its size and weight can't be freely carried
R-Map[33]	It an android application uses the captured Image, to provide read out loud service	Mobile Camera, OCR and TTS	Mobile phone has lower processing power than desktops. Mobile Image are effected with Skew, Blur, curved base lines and auto focus mechanism
Camera-Based text and product Label reading[35]	This method extracts the moving object from its complex background then recognized text in the object is outputted to the blind as voice.	A motion based Region of Interest (ROI) , OCR and TTS	This application is for only specific purpose that is for shopping only.
Trinetra[36]	With the help of networked devices to support Visually challenged person in navigation, shopping and transportation	Internet and Bluetooth-enabled mobile phone, text-to-speech software and a portable barcode reader	This application requires barcode reader which can be used for only specific purpose such as Shopping.

2.5 REVIEW OF OPTICAL CHARACTER RECOGNITION (OCR) ALGORITHMS

Most of the above used reading frame work used Optical Character Recognition. OCR is image processing technology in which the text in the image is extracted and converted into a text file. Input to the Optical Character Recognition module is the acquired bitmap image. On the input image we have to perform various pre-processing operations such as Binarization, Blur removal, skew correction [10] and noise removal step. Then text and non-text segmentation is performed to isolate text from graphics, segmented text is then segmentation into various lines. From each line we segment different words and then these words are segmented as individual characters. Character recognition of the individual character is performed and output will be the Text file containing the recognized text. Tesseract [41] is an open-source OCR engine that was developed at HP between 1984 and 1994. Tesseract was a PhD research project in HP Labs, Bristol, and became famous as commercial OCR engines due to failure of previous OCR engines on best quality print. Tesseract work concentrated more on improving rejection efficiency than on base-level accuracy. Tesseract was released as open source software by HP in 2005. After Tesseract overpowered different commercial OCR engines of that time when it was sent for the Annual Test of OCR Accuracy at UNLV in 1995.

2.5.1 Architecture of Tesseract OCR Engine

Tesseract [41] assumes binary image as input with polygonal text regions defined optionally. Initially the outline of the components is extracted with Connected Component Analysis. This was expensive step at that time but it had the advantage of recognizing the inverse text easily like black-on-white text. At this stage, outlines are collected together, purely by nesting into Blobs. Blobs are organized into text lines, and the lines and regions are analyzed for fixed pitch or proportional text. On Text lines word level segmentation is performed according to the type of character spacing. Fixed pitch text is broken immediately by character cells. Proportional text is broken into words using definite spaces and fuzzy spaces. Recognition is a two-pass process. In the first pass, each word was recognized. Satisfactory Recognized words are passed to an adaptive classifier as training data. The adaptive classifier then gets a chance to more accurately recognize text lower down the page. Since the adaptive classifier may have learned few useful words too late to make a contribution near the top of the page, Second pass is run over the page for better result, in which words that were not recognized well enough are recognized again. Fig.2.2 shows the overall architecture of Tesseract OCR engine.

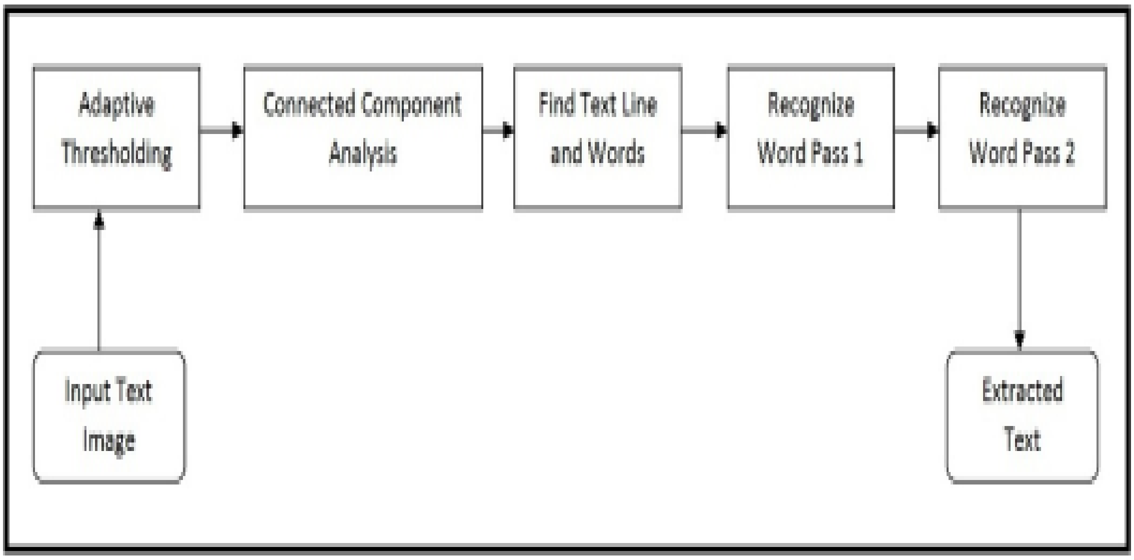


Figure 2.2 Overall Architecture of Tesseract OCR Engine.

2.6 REVIEW OF TEXT TO SPEECH ENGINE

After the Optical character recognitions of the Image we get the text file as output. Now text has to be converted to speech so that the visually challenged person can understand it. Text to Speech Engine converts the text output from OCR engine to its corresponding speech output.

2.6.1 Architecture of Text to Speech Engine

Text to speech Engine [50] [51] works by initially performing the pre-processing on the text and then the speech generation is performed. The preprocessing module prepares the input text for further processing the operations like text analysis, text normalization and then translating into a phonetic or some other linguistic representation are performed. In preprocessing the spell checking is performed on the text based on the punctuation marks the formatting of the paragraph is done. In text normalization the abbreviations and acronyms are handled. Text normalization enhances the speech output. In Linguistic analysis morphological operations for proper pronunciation of word and syntactic analysis to facilitate in handling ambiguities in written text is performed. Speech generation process consists of phonetic analysis which is used to find the phone level within the word. Each phone is tagged with the information about the sound to produce and how to be produced. This is followed by Grapheme to phoneme conversion based on dictionary. Next step is prosody analysis which attaches the pitch and duration information. Then after the speech synthesis that is rendering of voice we get the speech out form the text to speech Engine. The Fig.2.3 shows the overall architecture of Text to Speech Engine.

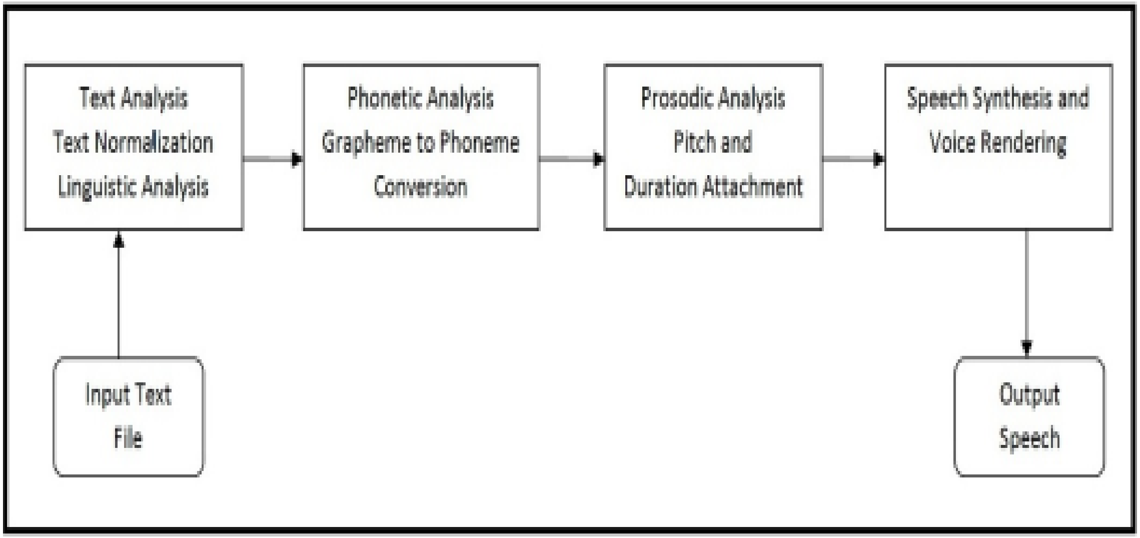


Figure 2.3 Overall Architecture of Text to Speech Engine.

2.6.2 Various Text to Speech (TTS) Engines developed

Considerable research has been done for the task of text to speech conversion. A screen reader in Indian languages has been designed by Speech and Vision Lab in Language Technologies research centre at IIIT Hyderabad named Reading aid for visually impaired (RAVI) [42]. The screen reading software (RAVI) is an assistive technology to help visually impaired people to use or access the computer and Internet. RAVI uses text-to-speech technology to speak out the contents on the current screen (display) of a computer system. This is a PC based software system. For wide adoption of the envisaged system support must be provided for Indian languages. Due to presence of large number of Indian languages and scripts this task is going to be huge. Effort have been made to convert Indian text to speech at various institutes, like ‘ACHARYA’ at IIT Madras, TTS for multiple Indian languages at CDAC, KGP-Talk at IIT Kharagpur, TTS for oriya language at utkal university etc. Today many mobile phone company provide Text to speech facility. Google has introduced TTS for android based phones available in English, Spanish, French and Italian languages. All these research identify host of research issues and a new system can be build upon all this existing work to develop capability to generate speech from multiple Indian language texts.

2.7 CONCLUSION

Based on the review [39] it is evident that to assist a visually challenged person the image captured from the capturing device can be used to acquire the knowledge by converting the text in the image to speech with the help of OCR and TTS engine. But a lot of work is still to be done when the input image is affected with the effect of skew, blur and different types of noises added due to the image capturing on the move. In order to increase the efficiency of the system we need to handle the following areas:

- Adaption of different text sizes and auto focus.

- Stabilization of page image against vibrations (in cases where the person is holding the book and reading).
- Elimination of the effect of blur in the image.
- Elimination of effects of shadows and the contrast changes.
- Identification of text.
- Conversion of text to voice.
- Integrating the system into a workable and portable system.

CHAPTER 3

IMAGE DEGRADATION AND RESTORATION MODEL

3.1 MODEL OF IMAGE DEGRADATION

The degraded image $g(x, y)$ is obtained by applying the degradation operation H over the input image $f(x, y)$ along with the additive noise $n(x, y)$.

$$\boldsymbol{g(x, y) = H[f(x, y)] + n(x, y)} \qquad \textbf{(3.1)}$$

Image restoration model is a process of estimating $f(x, y)$ for the given degraded Image $g(x, y)$ using the known or estimated value of H that is degradation Operator. The Operator H can be linear or non-linear.

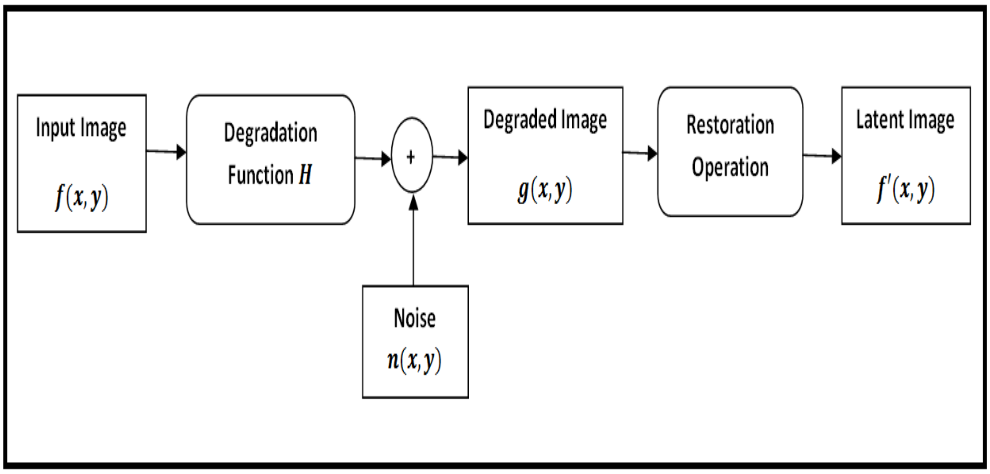


Figure 3.1 Overall degradation and restoration Model

Mostly, the degradation operator H it is assumed to be linear which satisfies the principles of superposition and homogeneity. The operator H is also considered to be space invariant or position invariant. An operator is said to be space invariant if the response at any point depends on the value at that point but not on the position of the point and is defined mathematically as,

$$\boldsymbol{H[f(x-a, y-b)] = g(x-a, y-b)} \quad \textbf{(3.2)}$$

For all $f(x, y)$ and any a and b. The overall model of degradation and restoration operations is shown in Fig. 1.

Impulse function $f(x, y)$ is expressed as,

$$\boldsymbol{f(x, y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(a, b)\delta(x-a, y-b)dadb} \quad \textbf{(3.3)}$$

Substituting (3) in (1) we get the blurred image $g(x, y)$ as

$$\boldsymbol{g(x,y) = H\left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(a,b)\delta(x-a, y-b)dadb\right] + n(x,y)} \quad \textbf{(3.4)}$$

As $f(a, b)$ is independent of x and y, using Linearity, the $g(x, y)$ can be expressed as,

$$\boldsymbol{g(x,y) = \left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(a,b)h(x,a,y,b)dadb\right] + n(x,y)} \quad \textbf{(3.5)}$$

Where $h(x, a, y, b) = H[\delta(x - a, y - b)]$ is called point spread function (PSF) in the optics Since H is spatial invariant

So the expression of H can be rewritten as,

$$\boldsymbol{H = [\delta(x-a, y-b)] = h(x-a, y-b)} \quad \textbf{(3.6)}$$

And the blurred image is given as,

$$\boldsymbol{g(x,y) = \left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(a,b)h(x-a, y-b)dadb\right] + n(x,y)} \quad \textbf{(3.7)}$$

As in two dimension image processing terms, the continuous convolution integral is of this form with integral limits form - ∞ to + ∞. Hence the expression is called the convolution integral in the continuous variable.

3.2 DISCRETE CONVOLUTION MODEL FOR DEGRADED IMAGE

Generally Image can be constructed in two ways either discrete or continuous. The continuous image can be converted to a discrete image by sampling, quantization and coding. The discrete model of the degraded image caused by the blur and the additive noise added can be expressed as,

$$\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y}) = \sum_{\boldsymbol{a}=\boldsymbol{1}}^{\boldsymbol{m}}\sum_{\boldsymbol{b}=\boldsymbol{1}}^{\boldsymbol{n}} \boldsymbol{f}(\boldsymbol{a},\boldsymbol{b})\boldsymbol{h}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{a},\boldsymbol{b}) + \boldsymbol{n}(\boldsymbol{x},\boldsymbol{y})$$

Where $f(x,y)$ represents the original image of size $m \times n$ and $h(x,y)$ represents the PSF of size $p \times q$. In the above equation $n(x,y)$ is taken as additive noise introduced by the system and is assumed to be zero mean white Gaussian noise.

Using spatial invariant property of PSF, the $g(x,y)$ can be described as,

$$\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y}) = \sum_{\boldsymbol{a}=\boldsymbol{1}}^{\boldsymbol{m}}\sum_{\boldsymbol{b}=\boldsymbol{1}}^{\boldsymbol{n}} \boldsymbol{f}(\boldsymbol{a},\boldsymbol{b})\boldsymbol{h}(\boldsymbol{x}-\boldsymbol{a},\boldsymbol{y}-\boldsymbol{b}) + \boldsymbol{n}(\boldsymbol{x},\boldsymbol{y}) \quad \textbf{(3.8)}$$

$$\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x},\boldsymbol{y}) * \boldsymbol{h}(\boldsymbol{x},\boldsymbol{y}) + \boldsymbol{n}(\boldsymbol{x},\boldsymbol{y}) \quad \textbf{(3.9)}$$

Where * denote the two dimensional linear convolution.

3.3 BLUR MODEL

Blur [59] [60] are considered as low pass filters which tends to smooth the abrupt changes in the gray level of the image. Blur also decreases the contrast of the image by averaging the pixels. Blur in the image is the degradation on the image which needs to be restored in order to extract the any information from it. Any image may be affected with blur in various ways including motion blur, de-focus blur, atmospheric turbulence blur etc. Blurred Image $g(x, y)$ is modeled as the convolution of the actual Image $f(x, y)$ with the degradation operator in this case it is the point spread function $h(x, y)$ and the additive noise $n(x, y)$. The mathematical equation is shown in equation (3.9). Different types of blur along with their point spread function $h(x, y)$ are as follows:

3.3.1 Motion Blur

Motion blur in the image is caused due to either the motion of the image capturing device or the object to be captured. Image capturing device can be a scanner or portable camera. Motion Blur [63][64] is governed by two parameter that is the length of motion (L) and the angle of motion (θ) [66]. When the text document to be captured translates with a relative velocity V in respect to the camera, the blur length L in pixels is $L = VT_{exposure}$ where, $T_{exposure}$ is the time duration of the exposure. The expression for motion blur is given as,

$$h(x,y) = \begin{cases} \frac{1}{L}, & \boldsymbol{if\ 0 \leq |x| \leq L\cos\theta\,;\, y = L\sin\theta} \\ \mathbf{0}, & \boldsymbol{otherwise} \end{cases}$$

When the angle of blur θ = 0, it is called horizontal motion blur. Point spread function can be represented in discrete as,

$$h(m,n,L) = \begin{cases} \frac{1}{L}, & \boldsymbol{if\ m = 0,\ |n| \leq \left|\frac{(L-1)}{2}\right|} \\ \frac{1}{2L}\left\{(L-1) - 2\left|\frac{L-1}{2}\right|\right\}, & \boldsymbol{if\ m = 0,}\ \ \boldsymbol{|n| \leq \left\lceil\frac{L-1}{2}\right\rceil} \\ \mathbf{0}, & \boldsymbol{elsewhere} \end{cases}$$

Figure 3.2 and 3.3 shows the PSF of Motion Blur and its spectrum in spatial domain and frequency domain respectively.

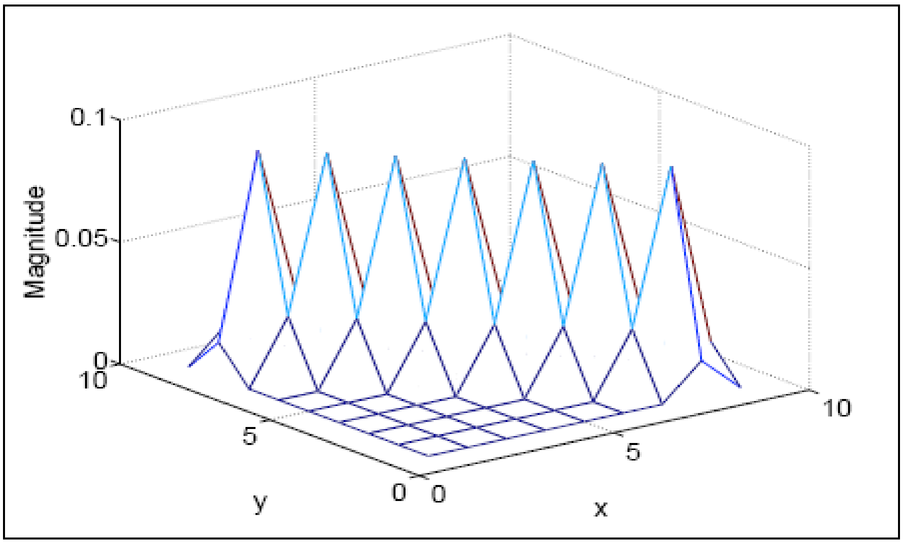


Figure 3.2 Point Spread Function of Motion Blur

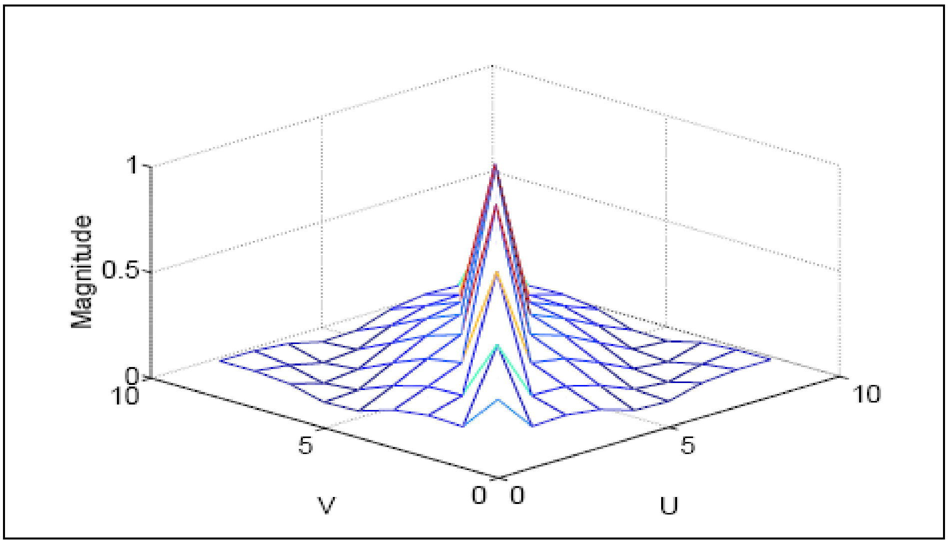


Figure 3.3 Frequency Response for PSF of Motion Blur

3.3.2 De-Focus Blur

De- focus blur is also known as out-of-focus blur in most of the places it is caused when a three dimensional image scene is captured and projected onto a two dimensional plane. This causes some part of the scene not to be focused properly. The camera equipped with circular aperture, the image of the point source is the small disk known as center of confusion. De-focus strength caused depends on two parameters that are focal length of the camera and the distance between the object to be captured and the camera. The Point spread function PSF is mathematically given as

$$h(x,y) \quad = \begin{cases} \dfrac{1}{\pi R^2}, & if \quad \sqrt{|x^2+y^2|} \leq |R| \\ \\ 0, & elsewhere \end{cases}$$

where R is the radius. The PSF of De-focus blur and Frequency Response are shown in the figure3.4 and figure 3.5 respectively.

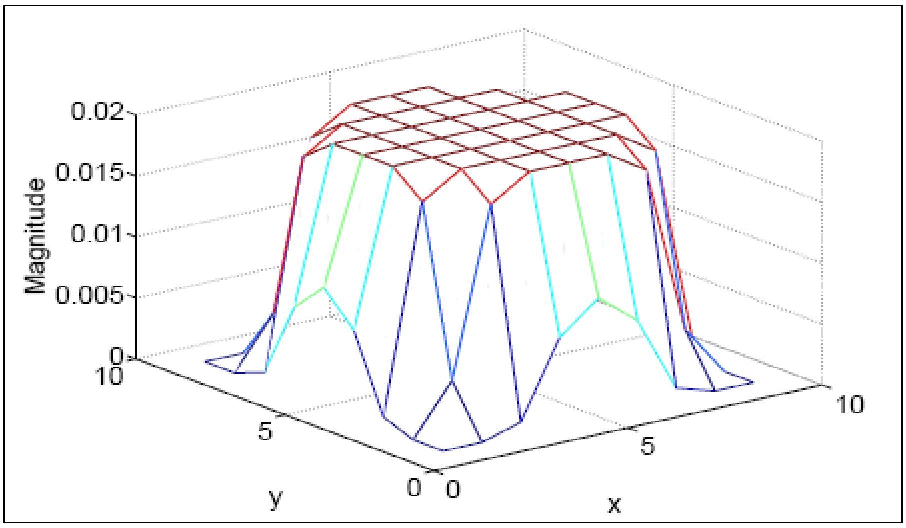


Figure 3.4 Point Spread Function of De-Focus Blur

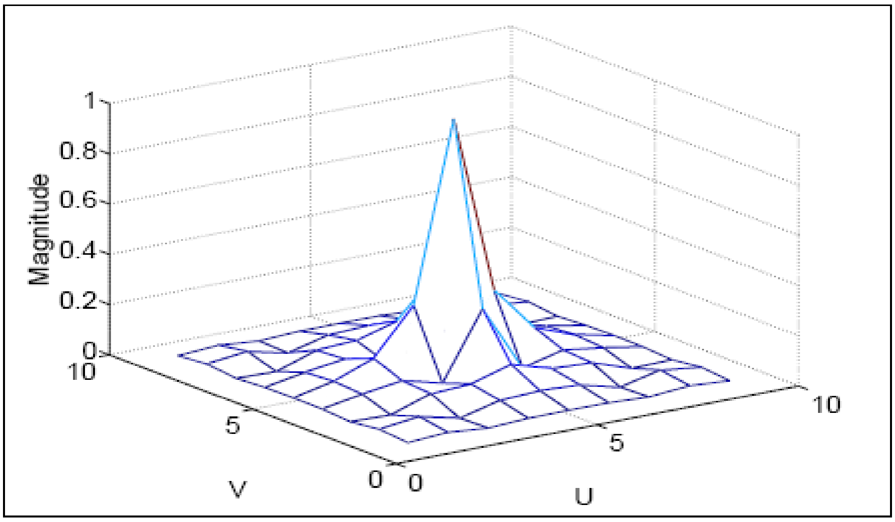


Figure 3.5 Frequency Response for PSF of De-Focus Blur

3.3.3 Atmospheric turbulence Blur

This type of blur is caused when there is degradation in the captured image due to the atmospheric turbulence of the earth. This type of blur is mostly found in the remote sensing applications which are mostly caused due to the wind velocity. Change in the wind velocity leads to the change in the refractive index of the layers of the earth which tends to distort the image to be captured. This type of blur efficiently modeled as a Gaussian PSF with standard deviation σ . Hence the PSF for the atmospheric turbulence Blur is given as

$$h(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

where standard deviation σ decides the strength or severity of the blur. The spatial domain and frequency domain plot of the turbulence blur are shown in the figure 3.6 and 3.7 respectively. As this thesis is mostly concentrated on the visually challenged person and the image captured by visually challenged person hence this type of blur is not discussed in this thesis, as it is out of the scope of the current research work.

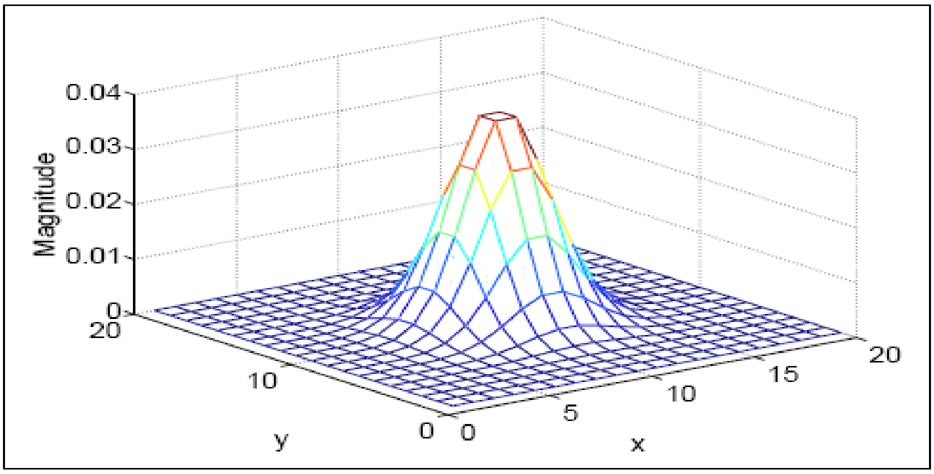


Figure 3.6 Point Spread Function of Atmospheric turbulence Blur

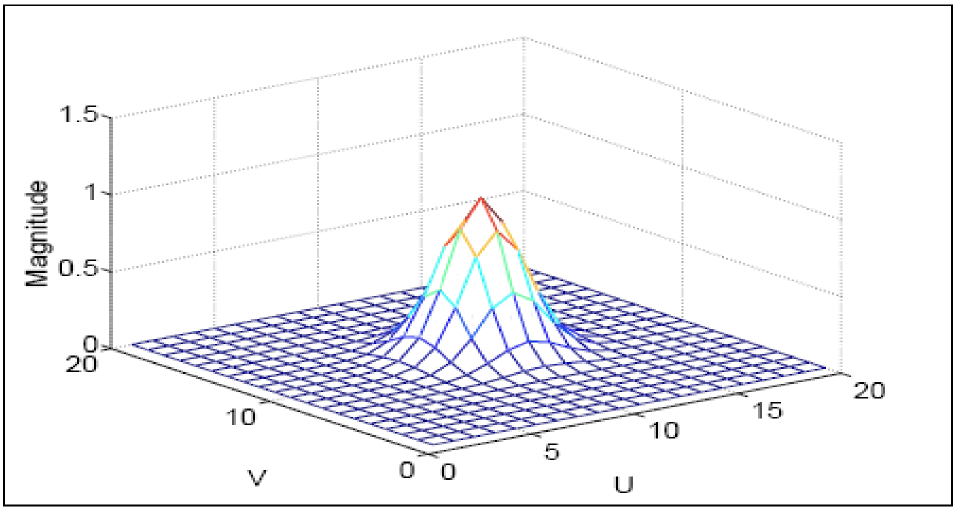


Figure 3.7 Frequency Response for PSF of Atmospheric turbulence Blur

3.4 NOISE MODEL FOR GAUSSIAN NOISE

Image capture and transmission is also affected with the additive noise [66] which is induced to the image. Basically the most common types of noise are impulsive and Gaussian noise, which affect the image at the time of capturing due to the noisy sensors which induce the noise in the image. As the image has to be transmitted through different channel it gets affected by the noise due to channel error. There are many noise models, but this thesis only discusses the Gaussian noise which is the most common in practical applications. Gaussian noise is a random variable and is expressed as,

$$n_{gaussian}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad \textbf{(3.10)}$$

It is characterized by its variance σ^2. The noisy image $f_G\ (x, y)$ is the addition of the original image $f(x, y)$ with the noise term. It is given as,

$$\boldsymbol{f_G(x, y) = f(x, y) + n_{gaussian}(x, y)} \qquad \textbf{(3.11)}$$

3.5 MODEL OF IMAGE RESTORATION

Image captured with help of capturing device is the degraded version of the actual image due to the imperfections added at the time of image production. In applications where image processing and machine vision are used there is need for correction in these degradations for the applications to work efficiently. There are different degradation which are caused due to nature that are blur, noise, geometric degradations, illuminations and different lighting conditions. In this thesis an effort to remove the degradation of blur has been made, which is induced in the captured image by visually challenged person in order to increase the efficiency of the overall system. Image restoration has been hot research topic for the researchers due to enormous applications. Research on image restoration started way back in 1950s with astronomical imaging by the researchers of United States and Soviet Union were involved in producing images of the earth and the solar system. The images were the degraded versions of the original images caused due to atmospheric turbulence. It is surprising to know that till today the image restoration is used in astronomical imaging for recovering the meaningful information from the degraded images. Image restoration [66] [70] [71] finds its applications in various fields which include medical imaging. It has been used to remove film-grain noise in X-ray images, angiography images and additive noise in magnetic resonance images. It has applications to quantitative auto radiography (QAR) in which image is obtained by exposing X-ray film to a radioactive specimen. Image restoration has been successfully applied to these paradigms, but still there is a scope for improvement in quality and resolution.

Image restoration finds application in media, video coding, post-processing and printing applications. It is used for recovering the video and pictures from the old scratched and deteriorated films. Image restoration also finds its application in defense field where a camera mounted on the missile takes distorted images due to pressure difference. Hence image restoration is important to recover the image. Due to a lot of applications image restoration has been an important topic for research in the technology world.

Image restoration is the phenomenon of restoring the degraded Image $g(x, y)$ with the help of Point Spread Function (PSF) $h(x, y)$ to obtain the restored image$f(x, y)$. If the value of PSF can be found then the restored image $f(x, y)$ can be easily obtained from the degraded image $g(x, y)$. If the PSF is known prior to restoration that type of restoration is known as Non- blind Restoration or the classical restoration. In real life problems the PSF is unknown and no or very less information is available of the original image. In order to get the original image from the degraded image without any information about the blurring process is known as

Blind Deconvolution. In real life situation some information about the original image is required in order to restore the degraded image. There are many areas where there is no prior information about the original image like in remote sensing application where the image are new and they are never been imaged before. There are situations in which the quality or the intensity of the capturing device cannot be increased. In such cases it is very hard to estimate the PSF using classical image restoration in that situation Blind Deconvolution method is used. In Blind Deconvolution method we estimate the parameters of the PSF first and then we restore the image using any of the classical image restoration technique.

3.5.1 Classical Image restoration Technique

In classical image restoration technique we have the information about the PSF before the restoration process. In order to recover the Original image $f(x, y)$ from the degraded Image $g(x, y)$ there are many different techniques. The restored Image $f(x, y)$ is given as

$$\boldsymbol{f(x,y) = g(x,y) * h(x,y)}$$

$$\boldsymbol{f(x,y) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} g(k,l)\, h(x-k\,, y-l)}$$

Where * denote Deconvolution operation which is the inverse of convolution operation.

In frequency domain it is represented as

$$\boldsymbol{F(u,v) = G(u,v)\; H(u,v)}$$

Where $F(u,v)$ represents the restored image in spectral domain, $G(u,v)$ represents degraded image and $H(u,v)$ represents PSF in frequency domain.

3.5.2 Inverse Filtering

This filter uses the inverse operation as the name suggests that is it takes the inverse of the point spread function as an impulse response. This type of filtering it is tough to be implemented in the image domain. An estimation of the latent or the restored image is found by dividing the transform of the degraded Image $\mathrm{G(u,v)}$ by the degradation function. This division is performed on the individual image basis using the value of the degraded image.

$$\boldsymbol{\hat{F}(u,v) = F(u,v) + \frac{N(u,v)}{H(u,v)}} \qquad \textbf{(3.12)}$$

This filtering requires the estimation of the PSF prior to the restoration process. It is best suited when the noise is almost negligible or absent. Examining the equation (3) it is clear that if the noise is absent, the second term of the equation vanishes. Thus the restored image is identical to the original image. When the value of $H(u,v)$ is zero for certain frequencies. At that point the noise associated gets amplified. Thus the result of this type of filtering is dominated by amplified frequency. It can be used when the noise associated with image is negligible or absent.

3.5.3 Least square Filtering

Inverse filtering has the constraints that it is sensitive to noise as the noise gets amplified and the restored image is not the clear image as the original image. The most commonly used least square filtering is the Wiener filter is based on the mean square error (MSE) criteria.

Wiener Filter: This filtering is spatial invariant that uses the minimum mean square error criteria. It chooses that PSF that minimizes the MSE between the restored image $\hat{F}(u,v)$ and the Original image $F(x,y)$. MSE defined is minimized. This minimization solution is called Wiener Filter in which the restored Image $\hat{F}(u,v)$ in frequency domain is given as

$$\boldsymbol{\hat{F}(u,v)} = \left[\frac{\boldsymbol{\hat{H}(u,v)}}{|\boldsymbol{H(u,v)}|^2 + \frac{\boldsymbol{S_n(u,v)}}{\boldsymbol{S_f(u,v)}}}\right]\boldsymbol{G(u,v)} \qquad \textbf{(3.13)}$$

Where $S_n(u,v)$ and $S_f(u,v)$ are the power spectrum of the noise and the original image respectively. The restored image is obtained by the inverse Fourier Transform of $\hat{F}(u,v)$ in the spatial domain. If the equation (3.5) is analysed then it is observed that if there is no noise addition to the degraded image then the noise power spectrum will vanish away. Hence the wiener filter reduces to an inverse filter. When the power spectrum of the original image $S_f(u,v)$ is not known the equation (3.5) changes to

$$\boldsymbol{\hat{F}(u,v)} = \left[\frac{\boldsymbol{1}}{\boldsymbol{H(u,v)}}\frac{|\boldsymbol{H(u,v)}|^2}{\boldsymbol{H(u,v)}|^2+\boldsymbol{K}}\right]\boldsymbol{G(u,v)} \qquad \textbf{(3.14)}$$

3.5.4 Lucy- Richardson Algorithm

This is the algorithm which has been independently proposed by Lucy and Richardson [66][73]. This is an iterative algorithm which works by maximizing a Poisson statistics image model likelihood function. The L-R algorithm performs the series of steps

1. An initial approximation of the restored image $\widehat{f_0}$ is made. Typically, the Observed image $g(x,y)$ is taken as$\widehat{f_0}$.

2. The approximation is convolved with the PSF as,

$$\varphi_n = h * \widehat{f_n}$$

3. A correction factor is calculated depending on the ratio of the blurred image and output of the last step as,

$$\emptyset_n = \overleftarrow{h} * \frac{g}{\varphi_n}$$

Where $\overleftarrow{h}$ is the PSF in reverse order.

4. The new restored image is given by

$$\widehat{f_{n+1}} = \widehat{f_n} \cdot \emptyset_n$$

Where **.** denotes the pixel by pixel multiplication in spatial domain

Steps 2-4 are iteratively performed till an acceptable image quality is obtained.

3.6 BLIND IMAGE RESTORATION TECHNIQUE

It has been already discussed that the blind image restoration is performed when the information about the degradation process is missing and there is no any prior knowledge of the PSF. Some important features of Blind Image Deconvolution [66] [75] algorithms are as follows

1. Both the Actual image and the PSF are irreducible. That is both the signals cannot be expressed as the convolution of two or more component signals.
2. It is obvious that the restored image is not the same as the actual image. Hence it can be said that restored image is the scale shifted version of the actual image.

$$\boldsymbol{\hat{f}(x,y) = K\, f(x - D_x\,, y - D_y)} \qquad \textbf{(3.15)}$$

   Where $K, D_x\,, D_y$ are real constants and $\hat{f}(x,y)$ is the restored image.

3. Solution to this blind image restoration is not unique.
4. There is always a poor compromise between the complexity, convergence and Portability of the algorithms used for blind Deconvolution.

The details of various blind deconvolution techniques for image restoration are as follows:

3.6.1 Motion Blur Identification Using Statistical Measure

It is known that motion blur is caused due to the movement of the object or the capturing device at the time of image capture. It can be estimated using statistical measures based on two parameters that are blur length and the blur angle. Moghaddam and Jamzad [66] proposed a method which uses two dimensional cepstrum of the blurred image $g(x, y)$ to identify the blur length and angle of motion blur. The cepstrum is defined as,

$$\boldsymbol{C}\big(\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y})\big) = \boldsymbol{F}^{-1}(\mathbf{log}\,|\boldsymbol{F}\big(\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y})\big)|) \qquad \textbf{(3.16)}$$

The line segment connecting the origin to the first negative peak in the cepstrum can be used to estimate blur length and angle. Length of the line segment gives the blur length and the blur angle is estimated by finding inverse tangent of the slope of the line segment. For noisy situations the bispectrum is suggested. The discrete bispectrum of the i^{th} segment of the blurred image $B_i(k, l)$ for one dimensional case $(l = 0)$ can be given as

$$\boldsymbol{B_i}(\boldsymbol{k},\mathbf{0}) = \left|\boldsymbol{F_i}(\boldsymbol{k})\boldsymbol{H}(\boldsymbol{k}) + \boldsymbol{W_i}(\boldsymbol{k})^{\mathbf{2}}\right|[\,\boldsymbol{F_i}(\mathbf{0}) + \boldsymbol{W_i}(\mathbf{0})] \qquad \textbf{(3.17)}$$

Where $F_i(k)$ and $W_i(k)$ are the Fourier transforms of i^{th} segment of the original image and noise respectively. $H(k)$ denotes the Fourier transform of the PSF.

3.6.2 Radon Transform

This transform is used to estimate the blur angle. The radon transform [76] is given as

$$\boldsymbol{R_f(x,\theta) = \int_{-\infty}^{\infty} f(x\cos\theta - y\sin\theta\,, x\sin\theta - \ y\cos\theta)\, dy} \qquad \textbf{(3.18)}$$

Radon transform is given as integration of f over the limits -∞ to ∞ over a line of distance x to the origin and angle θ to the y-axis. In order to estimate the blur angle θ is varied from 0 to 360 degree and the angle corresponding to the maximum value of R gives the blurring angle.

This method can be used for the images having square size with black background. This method does not give accurate results for the rectangular images. If the background is not black and additional objects close to the boundary of the object then this leads to additional lines in the spectral domain.

3.6.3 Hough Transform

This transform can also be used for the estimation of the blur angle using the spectrum of the blurred image. Hough transform [77] [78] estimates the angle based on the orientation of the lines in the spectrum treating it as an image. The Hough transform divides the parameter space into accumulator cells. The curve r for a given point (x, y) in Hough transform can be expressed as,

$$r = xcos\,\varphi + ysin\,\varphi \qquad (3.19)$$

For each point in the image, the corresponding curve is entered in the accumulator by incrementing the count in each cell along the curve. Hough transform returns the accumulator array in which there are values, the maximum value corresponds to blur direction. Hence blur direction and the blur angle can be estimated using Hough Transform. This transform produces best result when there is a single line in the spectrum which is practically rarely possible due to the presence of noise which leads to more than single line in the spectrum.

3.7 CONCLUSION

In this chapter the Image degradation model is explained using mathematical modeling. It is evident from the above literature that the Blurred Image $g(x, y)$ is a degraded Image which is the convolution of the actual Image $f(x, y)$ and the Point spread Function $h(x, y)$ with additive noise $n(x, y)$. One of the aim of this thesis is to recover the actual Image $f(x, y)$ from the blurred Image $g(x, y)$ with the help of point spread function $h(x, y)$. The Point Spread function as discussed need to be known or estimated in order to restore the blurred Image. In Technologically Assisted Reading framework for Visually Challenged the image is capture by the visually challenged person while moving and there is no prior information about the Point spread function of the blur nor there any information about the blurring process. More over there is only single input Image to the system which adds to the difficulty. Hence in our case we have to follow the blind deconvolution Image restoration technique to estimate the Point spread function $h(x, y)$.

In Technologically Assisted reading framework for visually challenged person the blur caused will be mostly due to the motion of object or the capturing device at the time of image capture. Hence the blur in our case will be motion blur. The motion blur in turn depends on two parameters that are Blur length l and Blur angle θ. So the task is now to estimate the Point spread function of the motion blur, to fulfill it some criteria is needed based on which for different value of Blur length l and Blur angle θ the best or the optimal Point Spread function $h(x, y)$ which will be most efficient in obtaining the original or restored image $f(x, y)$ can be found. Restoring the exact Image as the original image $f(x, y)$ is not possible. Hence the idea is to get the best equivalent match of the Original Image which can serve the purpose and increase the reliability and efficiency of the reading framework. In the next chapter an algorithm has been proposed to estimate the Point spread function $h(x, y)$ for different values of Blur length l and Blur angle θ and evaluated all the values of Point spread function $h(x, y)$ with the help of Average Word Confidence of the OCR.

CHAPTER 4

TEXT DEBLURRING USING OCR WORD CONFIDENCE

Text deblurring using OCR word confidence technique is based on estimating the blur kernel or the Point Spread Function (PSF) of the motion blur using Blind Deconvolution image restoration method. Motion blur is either due to the movement of the camera or the object at the time of image capture. The Point Spread Function of the motion blur is governed by two parameters length of the motion and the angle of the motion. In this approach the PSF will be estimated for the motion blur iteratively for different values of the length and angle of motion [80]. The Deconvolution operation with the blurred image is performed with every estimated PSF to get the non- blurred or the latent image. Latent image obtained is then feed to an optical character recognition so that the text in that image can be recognized. The Average Word Confidence [80] for the recognized text is calculated. Thus for every estimated Point Spread Function and the obtained latent image the value of Average Word Confidence is calculated. Hypothesis is that the Point Spread Function with the highest Average Word Confidence value is the optimal Point Spread Function which can be used to deblur the given blurred textual image. In this method there is no prior information about the PSF and only single image is used as an input to the system. This method has been tested with the naturally blurred image taken manually and through the internet as well as artificially blurred image for the evaluation of the results. The implementation of the proposed algorithm has been done in Matlab. The resultant image is also shown in this chapter.

4.1 DEBLURRING USING BLIND DECONVOLUTION OPERATION

The problem of Image restoration removing the blur, is known as Deblurring. In Deblurring the original image $f(x,y)$ requires the deconvolution of the PSF $h(x,y)$ with the observed image $g(x,y)$. In most of the cases the PSF is known prior to the deblurring this type of deblurring is done using the classical well known techniques such as inverse filtering, wiener filtering , least square filtering, recursive Kalman filtering is available. PSF is unknown in most of the cases and very little information is available about the original image. To recover the original image $f(x,y)$ from the observed image $g(x,y)$ using partial or no information about the blurring process this phenomenon is known as Blind Deconvolution.

Blind Deconvolution algorithm can be broadly classified in two types. In first type PSF is identified first and then utilized to deblur the image using any of the classical Deblurring techniques. In second type algorithm we estimate the PSF and restore the image simultaneously. A number of methods exist to remove the blur from the observed image $g(x, y)$ using a linear filter. The restored image $f(x, y)$ from a given blurred image is given by

$$\boldsymbol{f(x, y) = g(x, y) * h(x, y)} \qquad \textbf{(4.1)}$$

Where * denotes the Deconvolution which represents the inverse of the convolution. In the frequency domain, this can be expressed as

$$\boldsymbol{f(u, v) = g(u, v)\ H(u, v)} \qquad \textbf{(4.2)}$$

Where $f(u, v)$ denotes the estimated image in spectral domain. $g(u, v)$ And $H(u, v)$ are the blurred image and PSF in frequency domain respectively.

Motion Deblurring with single-input image is more complicated than that with two or more input images because multiple blurred images always provide more information in solving the problem. In this approach, the main focus is on Blind image Deconvolution in which the PSF $H(x, y)$ is estimated and then the deconvolution operation is performed of this PSF with the blurred Image $g(x, y)$ to get the original image$f(x, y)$. As deblurring is ill poised problem it is difficult to recover the original image due to additive noise. In the process of motion blur, the deblurred image loses much high-frequency information. The traditional methods always give undesirable deblurring results because of the effect of the additive noise on the single image based motion deblurring. The actual problem is to find the PSF which will give the best text recognition result on the input blurred textual Image. There are many algorithms available for deblurring the images but for textual image there are almost few algorithms that too using single input image.

4.2 RELATED WORK

Deblurring Text image is a problem which is being researched due to its wide application in optical character recognition. The recognition of a blurred text is still a big problem for the most efficient optical character Recognition (OCR) engine. Text Deblurring can increase the efficiency of OCR. Hence various Deblurring techniques have been proposed so far are either based on blur kernel estimation also known as blind deconvolution method and non-blind deconvolution method. In blind convolution method no information about the blur kernel is known. In non-blind deconvolution method we have some knowledge about the blur kernel.

Blur can be estimated by the estimation of the blur kernel or the PSF $H(x, y)$ .PSF for the motion blur depends on the pixel length of the blur (L) and the angle of the blur(θ). Based on these two parameters the blur kernel is estimated and then used for deblurring process. Blur Kernel estimation is iterative process in which the blur kernel is found which can help us to restore the original image$f(x, y)$. But as it is known that after deblurring we cannot get the exact deblurred image as the original image due to the additive noise $n(x, y)$ which is induced due to the deconvolution process. So the image is that much deblurred that the desired information which is text information can be extracted. Various kernel estimation technique have been proposed for the estimation of the motion blur kernel for deblurring in which Taeg

Sang Cho et. al. [76] proposed a method to estimate blur caused by the camera shakes using edge analysis and by constructing randon transform of the blur Kernel. Shamik Tiwari [81] in his paper compared the entire blind deconvolution algorithm to estimate the PSF and then estimated motion blur parameter can be used in a standard non blind deconvolution algorithm. Lu Fang [83] proposed method by decomposing a blur kernel into three individual descriptors trajectory, intensity and point spread function. So that it can be optimized separately. Nimali RajaKaruna [84] proposed a method for deblurring for visually impaired people it used a 3-axis accelerometers and gyroscopes of the smart phone camera used for image capture to get the motion vector and the heuristics method is developed to determine the optimal motion vector to deblur. Jing Wang [85] proposed a blind motion deblurring approach that works by kernel optimization using edge mask is used. An alternative iterative method is introduced to perform kernel optimization under a multiscale scheme. Total-variation based algorithm is proposed to recover the latent image via non-blind deconvolution. Long Mai [86] addressed the problems of fussing multiple kernels estimated using different methods into a more accurate one that give better deblurring result. Jinshan Pan [87] proposed Lo-regularized prior based on intensity and gradient for Deblurring using distinct property of the text image.

Deblurring can also be performed using non-blind deconvolution technique in which we have more than one blurred image of the input. We also have some knowledge about the PSF. So it helps us in the estimation of the blur Kernel which can then be used to get the latent Image. Another solution to deblurring is through the neural networks in this approach we train a neural network using a back propagation algorithm [89] [90] in which we require both the ideal image as well as the blurred image of the ideal image. We train the neural network with blur image and ideal image pairs to find the relationship between the blur Image to the ideal Image.

Convolution Neural Network (CNN) [91] is also one of the solutions for deblurring using blind Deconvolution method. CNN is based on the concept of deep learning and is powerful machine learning technique [92]. CNN is trained using large collections of diverse images. From diverse image it can learn rich features representations. Pre-trained CNN can also be used. It is a multi layered structure which can help in extracting different features in each layer. In deblurring the blurred images are used to input and they are mapped into its corresponding ideal image to learn various features relationship between blurred and ideal image in the training phase. Thus after training lots of diverse blurred image into its ideal image. The CNN can then deblur an input image based on the learning from the trained samples.

4.3 PROPOSED APPROACH

In our proposed approach [80] we have followed the blind Deconvolution Image restoration method for the estimation of the blur kernel or the PSF due the reason that we have only single blurred input image and we do not have any prior information about the PSF or the blurring process. As we know that Motion blur is caused by motion of the camera at the time of the image acquisition. So the PSF of motion blur is a function of two parameters that are the length of the blur (L) and angle of blur (θ).

In this research an iterative method to estimate the length (L) and angle of blur (θ) has been devised. For each PSF the latent image $f'(x, y)$ has been calculated by performing the deconvolution operation between the input blurred image $g(x, y)$ and the iteratively estimated PSF $h(x, y)$, different values of length (L) and angle of blur (θ) were used iteratively. Then the Optical Character Recognition has been performed on the obtained Latent image $f'(x, y)$ to calculate the cumulative sum of the confidence of each word recognized and then calculating the Average Word Confidence (AWC) metric by dividing the cumulative sum of word confidence by number of words.

$$\boldsymbol{AWC = \frac{[\sum_{i=0}^{n}(Word\ confidence\ of\ each\ word)]}{n}}$$

Where n is the total number of words in the text obtained after OCR result. The value of the blur length (L) and blur angle (θ) used for the estimation of PSF $h(x, y)$ along with the corresponding obtained Latent image $f'(x, y)$ OCR Average Word Confidence (AWC) value are tabulated in a two dimensional array. The maximum value of average word confidence is found in the table and denoted as AWC_m. Then the value of the length (L_m) and angle of blur (θ_m) corresponding to the maximum average word confidence (AWC_m) result of the OCR is found from the table. Hence the value of the length (L_m) and angle of blur (θ_m) is the value for which the estimated PSF $h(x, y)$ at which the deblurring using the deconvolution method gives the optimal text recognition result using average word confidence metrics of OCR. Thus PSF is estimated using blur length and blur angle iteratively using the OCR average word confidence metrics. Average word confidence gives the criteria on which we can say that the estimated PSF is the optimal for deconvolution using any filter. Thus average word confidence for the best recognition will be the highest and the Blur length and the Blur angle corresponding to it will be optimized PSF to get the latent image with higher text recognition rate than other PSF.

4.4 ESTIMATING POINT SPREAD FUNCTION (PSF) FOR MOTION BLUR

Motion blur is caused by the motion of the camera at the time of image capture. Point spread function of the motion blur depends on the length (L) and angle of blur(θ). Estimation of PSF is blind deconvolution problem. In the current approach the focus is to find out the length of blur (L) for various images and estimate the length of the blur to be between 0 to 25 pixels based on our practical requirements. This value can be increased or decreased based on the requirement. The value of the angle for a blur (θ) is taken between 0 to 180 degrees counter clockwise. As the value of θ is taken in the broad range it can easily estimate the angle in the textual input image.

This means if length of blur $L = 9$ and angle is θ =0 degree then the blurred Observed image $g(x, y)$ will produce a translated image of Original image $f(x, y)$ by 9 pixels. The blur kernel or PSF is iteratively generated for the different value of length (L) and angle of blur (θ).

For each estimated PSF we apply Deconvolution operation which is the inverse of convolution operation with the blurred image $g(x, y)$ to obtain a latent image $f(x, y)$. From the latent image obtained we perform the OCR operation to find the value of Word Confidence of each word.

4.5 OCR EVALUATION OF THE LATENT IMAGE

OCR is the image processing technology which helps to recognize the text in the image. OCR operation requires to converts true color or grayscale input images to a binary image which contains only two values that is 0 which represents black pixel or 1 which represent white pixel, before the recognition process. It uses the Otsu's thresholding technique for the conversion. Then segmentation of text and non-text is the next step for the OCR. Obtained text is segmented into line and then individual words. Words are recognized and a text file containing the recognized word is the output of OCR.

In OCR evaluation of the latent image obtained after the deconvolution operation of the blurred input image and the estimated PSF. The OCR on the Latent Image is applied and from the OCR result the value of word confidence of the recognized words is calculated that indicates the confidence level with which the word is recognized. Word Confidence is normalized value between 0 and 1. Where 0 represents the lowest confidence value this indicates that the recognized word is likely to be incorrect and 1 is the highest value of confidence which indicate that the recognition is most likely to be correct.

The word confidence of Individual word is calculated and then the Average Word Confidence (AWC) is calculated which the mathematical average of the all the Word Confidence obtained. Then this AWC value will be used as a metrics for the estimation of the Optimal PSF for the Deblurring Process.

4.6 OCR WORD CONFIDENCE METRICS

Word confidence attribute of the OCR result is used. Word confidence is determined based on character level confidence. Character confidence gives the normalized value of how effectively the character is recognized. Better the character confidence of recognition better is the word confidence of recognition. In addition the word confidence is affected by the dictionary based verification. If a word is found in the dictionary, it increases the word confidence value of that word. The longer the word, the higher will be the confidence value if it is found in the dictionary. For example if a long word of around 15 characters is found in dictionary it is pretty sure that the word is correct and will yield a higher word confidence, while on wrongly detected character a match against the dictionary by mistake is unlikely to occur. Short words like 'add' or 'odd' will both be found in dictionary. Therefore for smaller words there is a probability that we can get the dictionary match. Hence to overcome this problem words with 2 or less characters are not checked against the dictionary. The word confidence is normalized to an interval of 0.00 to 1.00 where 1.00 is the best and 0.00 is the worst word confidence. This OCR word confidence metrics is used for the evaluation of the deblurring result for the estimation of the blur kernel or the PSF.

4.7 ALGORITHM OF THE PROPOSED APPROACH

The stepwise algorithm of the proposed approach is as follows:

Step 1: Input a blurred image $g(x, y)$.

Step 2: Repeat step 3 to 8 for length (L) =1 to 25 pixels

Step 3: Repeat step 4 to 8 for angle (θ) = 0 to 180 degree

Step 4: Create a blur kernel of motion blur using length L and angle θ as $H(x, y)$.

Step 5: Apply Deconvolution operation between $g(x, y)$ and $H(x, y)$ to get the latent image$f'(x, y)$.

Step 6: Apply OCR on latent image $f'(x, y)$ obtained in step 5.

Step 7: Find the sum of word confidence of every word in the latent image $f'(x, y)$ then find Average word confidence AWC by dividing sum of word confidence by number of words n.

$$AWC = \frac{[\sum_{i=0}^{n}(Word\ confidence\ of\ each\ word)]}{n}$$

Step 8: Store the value of length(L), angle(θ) and AWC obtained in a 2D-array R or a table.

[End of angle for loop]

[End of length for loop]

Step 9: Find the value of length L_m and angle θ_m corresponding to largest value of Average word confidence AWC_m .

Step 10: Return length L_m and angle θ_m obtained in step 9.

Step 11: END.

4.8 RESULT ANALYSIS

The input blurred image is taken manually which is affected with the motion blur. The motion blur is caused due to the movement of the capturing device at the time of image capture which simulates the image captured while moving for visually challenged person. On the input blurred image in Figure 4.1 the PSF is estimated for different value of blur length l and blur angle θ iteratively. Latent image for different values of blur length l and blur angle θ along with the individual word confidence are shown in Figure 4.2, Figure 4.3 and Figure 4.4.

Figure. 4.1. Blurred Input Image $g(x, y)$

Figure 4.2. Latent Image $f'(x, y)$ with AWC= 0.508564115 for PSF with L=17 and θ= 10.

Figure 4.3 Latent Image $f'(x, y)$ with AWC= 0.756988287 for PSF with L=20 and θ= 14.

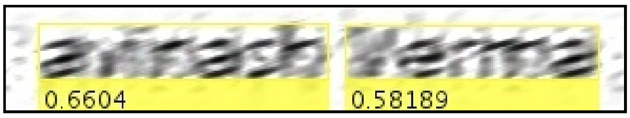


Figure 4.4 Latent Image $f'(x, y)$ with AWC= 0.621142268 for PSF with L=25 and θ= 13.

Table 4.1 Results of the Average Word Confidence Metrics

Top 20 results of Largest Average Word Confidence for Estimated PSF		
Estimated PSF for the Blur		**Average Word Confidence (AWC)**
Length (L)	**Angle(θ)**	
20	14	**0.756988287**
21	14	**0.745501578**

21	13	**0.739337564**
19	12	**0.734167099**
20	13	**0.725162864**
22	14	**0.714256167**
19	13	**0.713874578**
21	12	**0.713807285**
20	12	**0.712606966**
22	12	**0.706631124**
25	14	**0.698771179**
23	14	**0.696232617**
21	11	**0.691197157**
19	11	**0.688717008**
20	11	**0.678997517**
25	8	**0.67813015**
22	11	**0.653479457**
19	10	**0.640951991**
24	14	**0.635393023**
23	11	**0.632594824**

The Table 4.1 shows the values of blur length L and blur angle θ and average word confidence (AWC) for the blurred input image in Figure 4.1. The table contains the top 20 values of AWC and its corresponding values of blur length L and blur angle θ.

In the proposed work the PSF has been estimated for the blurred image $g(x, y)$ using blind deconvolution method iteratively by calculating the average word confidence (AWC) metrics that helps to estimate the optimal PSF. For every estimated PSF value we apply the deconvolution operation between the blurred image $g(x, y)$ and the PSF to get the latent image $f'(x, y)$ then result of OCR Average word confidence AWC is calculated. For latent Image highest Word Confidence means that its word recognition rate is higher than the others. Hence it can be the optimal PSF for recovering the text information using the OCR efficiently. Based on the Proposed Algorithm we have calculated AWC for different values of Blur Length L and Blur Angle θ. In Figure 4.1 the Blurred Input Image $g(x, y)$ containing the text is shown this is the Input to our Algorithm. The next Figure 4.2 shows the Output latent Image $f'(x, y)$ after the Deconvolution operation of $g(x, y)$ and the estimated PSF with Blur Length $L = 17$ and Angle $\theta = 10$ degree. The individual Word Confidences are 0.39378 and 0.62334 respectively. Hence the AWC evaluates to 0.508564115 this value represents the Confidence for the recognition of text in the image. The individual Word Confidence value is shown in the figures with the help of Yellow boxes. Similarly Figure 4.3 shows the Latent image $f'(x, y)$ obtained for the PSF with Blur Length $L = 20$ and Angle $\theta = 14$ degree with AWC = 0.756999287. Figure 4.4 also shows the Latent image $f'(x, y)$ obtained for the PSF with Blur Length $L = 25$ and Angle $\theta = 13$ degree with AWC = 0.621142268.

It is quite evident from the figures that the based on the value of AWC we can estimate the value of blur length L and angle θ for which the text recognition efficiency will be the highest. The highest value of average word confidence AWC_m= 0.756988287 for the length of blur L_m = 20 pixels and angle of blur θ_m= 14 degree are obtained from the Table 4.1. This is the value for the PSF or the blur kernel estimated using our method which will get the best OCR output for the input blurred image $g(x, y)$ using blind deconvolution method. The highest value of the average word confidence AWC_m means that the text document is best recognized with that PSF.

The Table 4.1 also shows that the blur length L varies from 19 to 25 pixels and blur angle θ varies in the range 8 to 14 degree. The value of Average Word Confidence (AWC) varies in the range 0.756988287 to 0.632594824. These values of AWC are acceptable but the highest value will give the best text recognition result so the value blur length L and blur angle θ corresponding to it gives the optimal PSF to apply deconvolution operation with the blurred input image to get the latent image which can be feed to an optical character recognition Engine. Then the optical character recognition engine will give better text recognition for a blurred input image with text, which was one of the biggest problems with the existing system.

Table 4.2 Comparison of our Algorithm with Lucy Richardson Algorithm for deblurring

Methods	Our Algorithm	Lucy Richardson Algorithm
INPUT IMAGE		
L=17 θ=10		
L=20 θ=14		
L=25 θ=13		

In the above Table 4.2 it is evident that the proposed algorithm gives better result with text deblurring than the Lucy Richardson Algorithm. In table the input to both the algorithm is blurred text image. After applying our algorithm for text deblurring using OCR word confidence we have estimated the value of the PSF $h(x, y)$ which is formed with the combination of blur length L =20 and Angle θ =14 with AWC_m= 0.756988287. For the same PSF our result is better than the result of Lucy Richardson algorithm for the optical character recognition. We have also shown the resultant image for different PSF. The resultant images shows that for the blurred image having text our algorithm gives better results for the optical character recognitions.

4.9 CONCLUSION

The proposed method is applied successful on the blurred image taken manually and through the internet. The method is also tested on artificially blurred images to validate our result. This method works successfully for the image having text and it cannot be used for deblurring any other type of image as it is only meant for the image having text in it, as OCR word confidence is used. This method was proposed to overcome a problem that was to deblur a text document image and to increase the accuracy of Recognition. The existing system was heavily dependent on OCR for text segmentation and recognition. But when the acquired image was blurred Image then most of the Efficient OCR engines performance was not satisfactory. In this research issue of motion blur has been tried to be resolved, which is the most common type of blur. The blur kernels or the PSF were estimated and evaluated with the average word confidence value. Based on AWC we have estimated the optimal PSF for the Input Image with text. The method used is based on the estimation of PSF which is an iterative process and another solution to the deblurring text is through artificial neural network. Training a neural network is a time taking process and requires a lot of computation. As a part of future work we would like combine this proposed approach and a Neural Network based approach together to get more optimized solution to Deblurring textual images.

CHAPTER 5

ROBUST ASSISTIVE READING FRAMEWORK FOR VISUALLY CHALLENGED

The Main objective of robust assistive reading framework [94] is to communicate the textual information present in the image captured by the visually challenged person as speech, So that the visually challenged person can acquire knowledge about the surrounding. This framework can help Visually Challenged person to read books, magazine, road sign, warnings, instructions and various displays as well by taking their image along with the surrounding. The inherent problem with the conventional approaches was if the acquired image is affected with the issues of different lighting conditions, noise and issue of skew and blur, the final output of TTS will not be useful. As the visually challenged person may not be trained photographer or may be the photo is clicked by a autonomous device the issue of lighting, noise and blur will surely be there. Most of the time image captured by the visually challenged person or the device helping him, will be affected with the motion blur as the person will capture the image on the go or when he is moving. As we all know that the motion blur is caused when there is a movement of either the captured object or the capturing device. It is also well known that if the input image is affected with any type of blur then it will fail even the best of the available OCR. The OCR output in that case will be far away from the expected output. This will lead to ambiguous or false detection of the text information. This ambiguous text information will be then converted to speech with the help of Text to Speech engine. This speech if communicated to the visually challenged person will create confusion instead of providing the correct information to the visually challenged person.

In the current approach a new method has been introduced which includes two more processes that are deblurring using blind deconvolution method to remove the anomalies of blur. The second process is pre-processing operation to remove the effect of noise after the deblurring operation as it is known that noise gets amplified during deblurring step. Thus it prepares the image for efficient result of the framework for visually challenged. The proposed approach is implemented in Matlab with the image captured manually and taken from the internet and the result along with the OCR text file and corresponding output Speech shows that our framework is better than the previous framework.

5.1 PROPOSED APPROACH

In the proposed approach input to the system is the image captured by the visually challenged person containing the textual information. Our aim is to communicate the text information in the image to the visually challenged person as Speech, so that he can acquire the knowledge about its surrounding. In order to fulfill the aim an overall architecture of robust reading framework for visually challenged person has been developed. In this framework the deblurring step is performed on the input image$f(x, y)$. This step is responsible for the blur kernel or the point spread function estimation in the image and then performing the deblurring operation on the input image $f(x, y)$ to get the deblurred Image$f^{'}(x, y)$. Deblurring is performed using blind deconvolution method using OCR performance [9]. After deblurring we get the deblurred image $f^{'}(x, y)$ then we perform the preprocessing steps in order to prepare the image $f^{'}(x, y)$ for better optical character recognition the steps involved are noise removal, thresholding and then binarization operations to obtain the perfect binary black and white image $f^{''}(x, y)$ which can be sent to the OCR for the extraction of the text information in the image $f^{''}(x, y)$ and convert the text information into corresponding output text file. The final step is conversion of output text file to speech or voice output with the help of Text to Speech. Hence the output will be the speech which will be communicated to visually challenged person so that he can get the knowledge of the communicated text. The Output Speech is feed to the Visually Challenged person so that he/she could listen whatever text is there in the image.

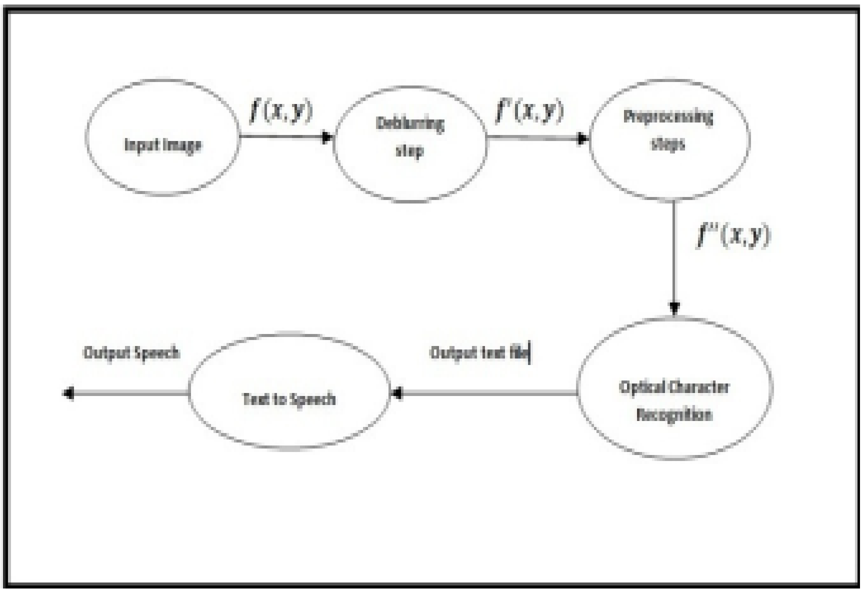


Figure 5.1 Overall Architecture of Robust Reading Framework for Visually Challenged

5.2 ALGORITHM

The step wise algorithm of the proposed robust reading framework for visually challenged approach is as follows:

i. Input to the System is Image Captured by Visually Challenged person $f(x, y)$.

ii. Blur Removal is the most important operation in which we first Estimate the Blur Kernel $H(x, y)$.Blur Kernel is estimated with the help of two parameters that are Blur Length L and Blur Angle θ.

iii. Iteratively the Blur Kernel $H(x, y)$ for different values of Blur Length L and Blur Angle θ is created. Then using Deconvolution operation we extract the Latent Image $f^{'}(x, y)$ for the Input Image $f(x, y)$ and the Current computed Blur Kernel $H(x, y)$.

iv. In Order to find the best Blur Kernel $H(x, y)$ which can help in the text information extraction each resultant Latent Image $f^{'}(x, y)$ are evaluated by calculating the Average Word Confidence AWC with the help of OCR.

v. Highest value of AWC_{max} indicates that Latent Image $f^{'}(x, y)$ is best recognized with the corresponding Blur Kernel $H(x, y)$ with two parameters that are Blur Length L_{max} and Blur Angle θ_{max}.

vi. Thus De-blur using Blind Deconvolution, the Input Image $f(x, y)$ with the Blur Kernel $H(x, y)$ with Average Word Confidence AWC_{max} for Blur Length L_{max} and Blur Angle θ_{max} to get the Deblurred Image or the Actual Latent Image $f^{'}(x, y)$ that can be used for further processing.

vii. The Preprocessing Operations are applied to the Deblurred Image or the Actual Latent Image $f^{'}(x, y)$ before OCR to obtain the Image $f^{''}(x, y)$ which is used for further processing. Preprocessing Operations increase the efficiency of the overall system.

viii. The Image $f^{''}(x, y)$ thus obtained is feed to an Optical Character Recognition (OCR) Engine for Text recognition. Output of OCR engine is the text file containing the text in the Image.

ix. Text file Output is send to a Text to Speech (TTS) Engine which performs the preprocessing operation on the text file and then converts it to the corresponding Speech output.

5.3 INPUT IMAGE CAPTURED BY VISUALLY CHALLENGED

Image of the text can be taken either with the help of a scanner or with the help of a portable camera or smart phone camera. In today's era with the advent of technology we have high resolution portable camera with many features such as autofocus, wide angle and high picture clarity it is easy to capture text with clarity. If the text in the captured image is clear then its segmentation and recognition [2] will be efficient and overall accuracy of the system will be high. But for a Visually Challenged person to capture the image with the help of a portable camera and with clarity will be a herculean task. We have assumed that the image captured by the Visually Challenged will be affected with the issues of Blur [2] and skew [3]. Problem associated with the Scanner is that for Visually Challenged person, it is difficult to place a text document in a proper way as the document need to be placed in the scanner in proper orientation, for the scanner to scan efficiently. A lot of research is done on the image acquired by portable camera for efficient optical character recognition [4]. This research is trying to solve the problem of blurred textual image that too affected with a Uniform Motion Blur. Motion blur is the most common type of blur in our case as this caused due to the movement of object or the capturing device at the time of Image capture.

5.4 DEBLURRING STEP ON INPUT IMAGE

Input image to the system $f(x, y)$ is image captured by the visually challenged person with the motive to acquire the text information present in the image as speech, so that they can act accordingly. As already discussed that the image if effected by motion blur that is caused by the movement of the capturing device or the object at the time of image capture then this system accuracy will be at stake as the OCR is going to fail miserably and the next step that is TTS engine will have nothing as input or incorrect text for speech generation. To overcome this problem Deblurring step is being implemented. Average Word Confidence AWC metrics [9] is used for the deblurring process. AWC Value lies between 0 and 1 it is calculated as

$$\boldsymbol{AWC = \frac{\sum Individual\ Word\ Confidences}{Total\ number\ of\ words}}$$

In this Deblurring step, iteratively based on two parameters of motion blur that are Blur Length L and Blur Angle θ estimation of the optimal PSF is tried that will give the best recognition rate based on AWC value of the OCR. AWC Is the mathematical average of individual word confidence. Word confidence [9] is the normalized sum of character level confidence. Character confidence is the normalized measure of the how effectively the character is recognized. Higher the character confidence of recognition Higher is the word confidence of recognition and vice-versa. The word confidence is also affected by the dictionary based verification. If a word is found in the dictionary, it increases the word confidence value of that word. The longer the word, the higher will be the confidence value if it is found in the dictionary. For example if a long word of around 15 characters is found in dictionary it is pretty sure that the word is correct and will yield a higher word confidence, while on wrongly detected character a match against the dictionary by mistake is unlikely to occur. Short words like 'add' or 'odd' will both be found in dictionary. Therefore for smaller words there is a probability that we can get the dictionary match. Hence to overcome this problem words with 2 or less characters are not checked against the dictionary. The word confidence is normalized to an interval of 0.00 to 1.00 where 1.00 is the best and 0.00 is the worst word confidence.

5.5 WORKING OF DEBLURRING STEP

For different value of Blur Length $L\ (L_1 L_2 L_3 L_4 \ldots\ldots L_n)$ and different value of Blur angle $\theta(\theta_1 \theta_2 \theta_3 \theta_4 \ldots\ldots \theta_n)$ we make the PSF $H(x, y)$ and then for every PSF we perform the deconvolution operation which is inverse operation of the blur to get the latent Image $f^{'}(x, y)$ and the using OCR we calculate the AWC value for the latent Image. The value of Blur length, Blur Angle and AWC are stored. The highest value of AWC_{max} is identified and the corresponding values of the Blur length L_{max} and Blur Angle θ_{max} are calculated. These values constitute the best PSF $H(x, y)$ which can deblur the given input Image $f(x, y)$ to get the maximum recognition rate of the text information in the Image. In the next step deconvolution operation is performed on the Input Image $f(x, y)$ with the obtained value of Blur kernel or the PSF $H(x, y)$ to obtain the latent Image $f^{'}(x, y)$ which is given as

$$\boldsymbol{f'(x,y) = f(x,y) ** H(x,y)}$$

where ** is the Deconvolution operation which helps in deblurring and is the inverse of Blurring operation. The resultant Image as known as Latent image $f^{'}(x, y)$ as it will be a distorted version of actual after deblurring step as the noise in the image get amplified during this process. Latent image $f^{'}(x, y)$ will be used for the further processing. Figure 5.2 shows Input Image $f(x, y)$ to this system is Blurred Image. Now we estimate the PSF for which we will get the recognition of the text. Hence for every estimated PSF we perform the deblurring operation and calculate AWC value for them. The Highest value indicates that the PSF used will give the best recognition result for the Blurred Input Image. Figure 5.3 shows the Deblurred Input Image $f'(x, y)$ with PSF obtained for Blur Length L =18 pixel, Blur Angle θ =24 degree $AWC =$ 0.586364 the yellow boxes with values in the image indicate the word confidences of the individual words and Figure 5.4 shows the Deblurred Input Image $f'(x, y)$ with PSF obtained for Blur Length L =17 pixel, Blur Angle θ =19 degree $AWC =$ 0.646726. For different values of Blur Length L and Blur Angle θ the highest value of AWC_{max}= 0.646726. Hence the corresponding values of the Blur length L_{max} = 17 pixels and Blur Angle θ_{max} = 19 degree are found. The Latent image thus obtained $f'(x, y)$ thus obtained in shown in Figure 5.5 that is used for further processing in next steps.

Figure 5.2 Blurred Input Image $f(x, y)$

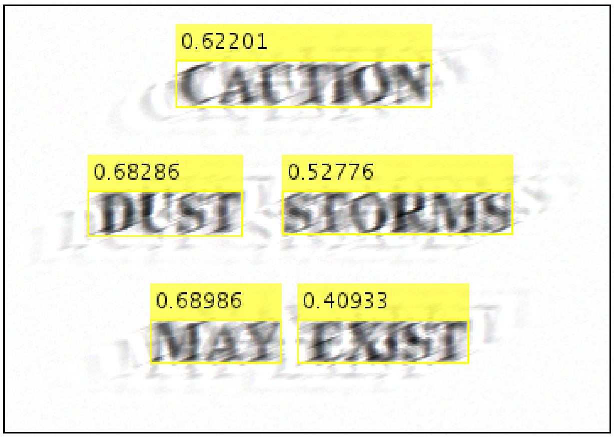


Figure 5.3 Deblurred Input Image $f'(x, y)$ with PSF obtained for Blur Length L =18 pixel, Blur Angle θ =24 degree $AWC = 0.586364$

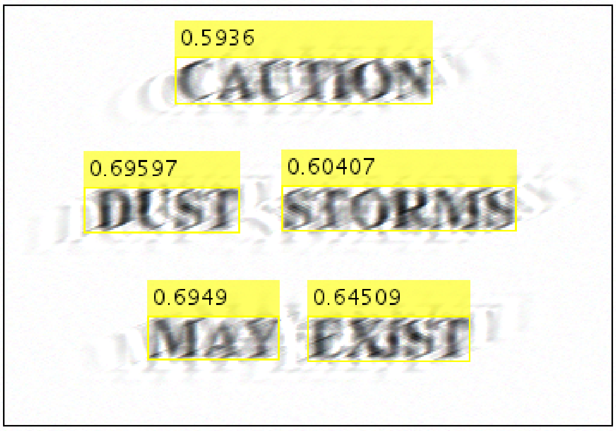


Figure 5.4 Deblurred Input Image $f'(x, y)$ with PSF obtained for Blur Length L =17 pixel, Blur Angle θ =19 degree $AWC = 0.646726$

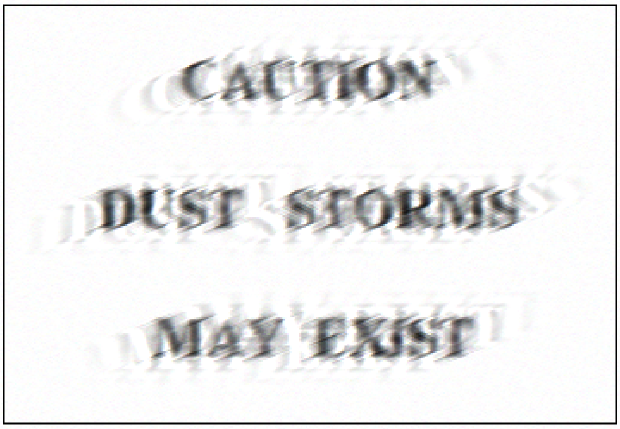


Figure 5.5 Latent Image $f'(x, y)$ with Optimal PSF obtained for Blur Length L =17 pixel, Blur Angle θ =19 degree $AWC = 0.646726$

Table 5.1 Results of the Average Word Confidence Metrics with Non Zero Values

Results Of Average Word Confidence for Estimated PSF In Decreasing Order		
Estimated PSF for the Blur		**Average Word Confidence (AWC)**
Length (L)	**Angle(θ)**	
17	19	**0.646726**
19	17	**0.637206**
17	21	**0.636685**
18	18	**0.635106**
18	22	**0.624655**
18	21	**0.623174**
18	20	**0.619045**
23	23	**0.617842**
18	19	**0.600523**
19	21	**0.597347**
19	18	**0.594126**
18	24	**0.586364**
20	22	**0.582763**
18	23	**0.563801**
21	22	**0.558978**
16	23	**0.558913**
20	19	**0.558427**
17	22	**0.557654**
20	21	**0.556653**

20	23	**0.555898**
20	18	**0.551193**
21	21	**0.549634**
21	20	**0.543814**
17	20	**0.5413**
20	17	**0.534714**
20	25	**0.53085**
19	23	**0.501604**
19	24	**0.489104**
20	24	**0.482753**
18	25	**0.475628**
22	23	**0.45861**
17	17	**0.448429**
17	23	**0.442411**
17	16	**0.427585**
19	25	**0.399794**
17	18	**0.377004**
19	19	**0.373256**
20	20	**0.371853**
19	22	**0.368714**
23	21	**0.36798**
19	20	**0.367736**
16	21	**0.365479**

18	16	**0.363268**
19	16	**0.357037**
16	22	**0.353944**
20	16	**0.347211**
15	19	**0.332066**
15	22	**0.325607**
16	19	**0.322062**
15	20	**0.305288**
21	23	**0.303825**
18	17	**0.297701**
25	25	**0.291613**
15	21	**0.283951**
16	18	**0.279449**
22	20	**0.272963**
21	24	**0.267132**
17	24	**0.249571**
22	22	**0.241624**
17	25	**0.240239**
16	25	**0.237284**
21	25	**0.235572**
16	16	**0.235418**
22	25	**0.229653**
22	24	**0.224419**

16	24	**0.221309**
15	23	**0.221071**
23	20	**0.213367**
15	18	**0.202676**
15	17	**0.19**
25	22	**0.12594**
24	24	**0.121792**
25	23	**0.120894**
23	25	**0.118905**
25	24	**0.117938**
23	24	**0.117733**
24	21	**0.117175**
24	25	**0.116435**
16	20	**0.115595**
24	23	**0.11544**
25	21	**0.104723**
25	20	**0.104677**
24	20	**0.056409**
15	25	**0.034019**
24	22	**0.030051**
15	24	**0.021606**

In the table 5.1 it contains the value of Average Word Confidences (AWC) for the input image in the Figure 5.2. The value of AWC is in the descending order that is the highest value is on the top and then the next highest value below it. The table shows that highest value of AWC_{max}= 0.646726 for the corresponding values of the Blur length L_{max} = 17 pixels and Blur Angle θ_{max} = 19 degree. Hence the PSF obtained for this value of Blur length L_{max} and Blur Angle θ_{max} will give the best recognition result of the OCR for the given blurred Image.

5.6 PREPROCESSING STEP ON RESULTANT IMAGE BEFORE OCR

Resultant Image $f^{'}(x, y)$ obtained from the Deblurring step is a mostly degraded image. As it is known that during the Deblurring process to obtain the latent image to the extent that its text information can be extracted with higher degree of Word confidence the deconvolution operation with the estimated Blur Kernel $H(x, y)$ is performed. During this process the noise gets amplified in order to increase the efficiency of the system we need the preprocessing steps. On the Image $f^{'}(x, y)$ the thresholding operation [2] is applied in order to increase the speed of processing. After thresholding step next operation is the binarization. Binarization [2] operation converts the image into perfect black and white pixels. In the image each and every pixel has either two values that are 0 or 1. Now the resultant Image of the previous stage $f^{'}(x, y)$ now gets converted to binary image having only black and white pixels. This binary image thus obtained $f^{''}(x, y)$ is the input to the next stage OCR engine which will give the best possible recognition result.

5.7 IMPLEMENTATION OF OPTICAL CHARACTER RECOGNITION

After the preprocessing steps the Resultant Image $f''(x, y)$ is feed to the OCR engine. OCR engine used in this research is Tesseract [20]. It is an open source OCR engine developed by HP between 1984 and 1994. Tesseract [20] works on improving the rejection efficiency than on base-level accuracy. The input image to the Tesseract is a binary Image $f''(x, y)$ obtained in the previous preprocessing step. Connected component analysis is performed and the resultant Image $f''(x, y)$ outlines of the components are stored. This step had the advantage of being simple to detect the inverse text and to recognize it as easily as black and white text. At this stage the outlines are gathered together by nesting the blobs. Blobs are organized into text lines, and the lines and regions are analyzed for fixed pitch or proportional text. Text lines need to be segmented into words this is done according to the type of character spacing. Fixed pitch text is chopped immediately by character cells. Proportional text is broken into words using definite spaces and fuzzy spaces. After the segmentation process next step is Recognition. Recognition process then proceeds as a two-pass process. In the first pass, each word is recognized. Each satisfactory recognized word is passed to an adaptive classifier as training data in order to learn the words for further recognition. The adaptive classifier then gets a chance to more accurately recognize text lower down the page. Since the adaptive classifier may have learned something useful too late to make a contribution near the top of the page. A Second pass is run over the page, in which words that were not recognized well enough are recognized again. This increase the accuracy of the recognition as two passes of recognition are run on the same input text.

In Matlab implementation of the OCR we assign the result of the OCR evaluation to a variable result. Then the confidence value of the recognition is extracted using this result variable. This variable is a two dimensional array containing different attributes associated with it. The word confidence value can be extracted using the dot operator that is result.WordConfidences this will return the array of dimensions [number of words, 1] of word confidences where each entry gives the word confidences of the word recognized. In the same way the character confidences can be found of the individual characters using result.CharacterConfidence. Then the Value of Average word Confidence AWC is calculated by taking the sum of Individual word confidence and dividing this with the total number of words. A two dimension array is created during the deblurring process to store the different value of the blur Length l and the blur angle θ and the AWC. Then the programming is applied to find the maximum value of the AWC. The value of blur Length l and the blur angle θ corresponding to this maximum value of AWC are found with the help of programming and these highest values are named as AWC_{max} and the corresponding values of the Blur length L_{max} and Blur Angle θ_{max} . Then we create the PSF of motion blur $h(x, y)$ for the estimated value of Blur length L_{max} and Blur Angle θ_{max}. After that the deconvolution operation is applied between the blurred Image $f(x, y)$ and the estimated PSF $h(x, y)$ to obtain the latent image $f'(x, y)$ which is the image that will give the improved recognition as compared to the input blurred image. In order to enhance the efficiency of the OCR result few preprocessing steps before the OCR need to be applied. As to increase the rate of recognition which is desired that will be used to communicate the visually challenged person so not any scope is left for any wrong detection of the text information. After the preprocessing steps the image obtained will be a binarized image $f''(x,y)$ which will be feed to an OCR for text recognition and the output will be the array of recognized words these words need to be combined together to formed a string. All the words are concatenated using the strjoin function of Matlab the OCR with a space after each word. The String thus found is then saved as the text file. This text file is then passed to the next function of Text to Speech for voice conversion.

5.8 IMPLEMENTATION OF TEXT TO SPEECH CONVERSION

Optical Character Recognition extracts the text information present in the input image $f(x,y)$ in the form of text file. In order to communicate it to the visually challenged person this text file need to be convert to speech. The text file is initially preprocessed for the spelling check as during recognition process few words can be misrecognized they are corrected. Text normalization is also performed in order to handle the abbreviations and acronyms to enhance the speech output. Various morphological operations of text to speech are performed for proper pronunciation of word in linguistic and syntactic analysis. After preprocessing step the next step is speech generation which involves the phonetic analysis to find the phone level present in the words. Different phone levels have the information about the sound tagged with it to be produced. Graphemes are then converted to phonemes based on the dictionary. The Prosodic analysis is one of the important steps of speech generation in this step pitch and the duration information is attached for speech conversion. After this speech synthesis is done which involves voice rendering to get the speech for the text input. In addition to that there are various options available for visually challenged person which can help him in order to understand the communicated speech output.

In the implementation of the text to speech function named TTOS is created which synthesizes the text input. The audio format is mono, 16 bit, 16k Hz by default which can be varied based on the requirement. The available voice in the list can be selected based on our preference by default first voice is set form the List. The pace of the speech output can also be set that is -10 as the slowest and +10 as the fastest. In our case we have used the default value that is 0. The sampling rate need to be set for the speech in Kilo Hertz that are 8000, 11025, 12000, 16000, 22050, 24000, 32000, 44100 and 48000. The default value 1600 is used. The function also requires the Microsoft Win32 Speech API (SAPI) for the voice output.

5.9 CONCLUSION

Robust reading framework for visually challenged person can help them in getting the knowledge of the surrounding in the form of speech. In this framework we have overcome the inherent problem associate with all the previous frameworks that was issue of blur in the Input image captured by the visually challenged person. The input image $f(x, y)$ to the system is initially checked for the motion blur. If the blur is present in the image we have iteratively based on two parameters that are blur length l and Blur angle θ created the PSF that is the blur kernel. For every PSF in the range the optimal PSF that will give the best Deblur image for further processing based on the average word confidence AWC metrics of the OCR are found. This deblurring step was not present in all the previous systems. Without this step the overall framework was of no use if the image had any kind of Blur. As blurred image will give incorrect recognition of text and the incorrect text will be communicated as voice which will be of no use for visually challenged person.

CHAPTER 6

CONCLUSION AND FUTURE WORKS

In this research work "Technologically assisted reading framework for visually challenged", the aim was to develop assistive technology for the visually challenged person through which they can read the textual information captured by them by hearing the captured text. The solution proposed by us to this problem was to use the current technologies such as Optical Character Recognitions (OCR) and Text to Speech (TTS) conversion to overcome this problem. A reading framework is proposed which consisted of an image acquisition through portable camera on the move consisting of textual information and then with the help OCR technology the textual information in the image captured can be converted to the ASCII or Text file. This text file can be send as the input to Text to speech engine which converts it to speech output of the corresponding text. Hence the Visually challenged person gets communicated about the text information captured by him whether it is a warning sign, magazine, newspaper, various displays at railways or airports and computer display or email.

This framework will provide the visually challenged person with greater degree of Independence in her life. In the previous systems the main issue was the problem of blur in the input image, noise and the contrast changes or different lightning conditions. This research addresses few issues that were never addressed before by the previous proposed system. These issues are mainly responsible for the failure of the previous systems. Out of the above problem the blurred input image was the most challenging problem as if the input image was blurred then this will cause even the best possible OCR to fail, and if the text is not detected then it cannot be converted to speech so ultimately the overall framework was at stake.

In order to overcome this problem and to increase the efficiency and reliability of the framework two more steps in the overall architecture of the framework was introduced, namely Deblurring step and the Pre-processing step before OCR. These two steps were the milestone in improving the performance of the OCR and TTS. As the image was captured by visually challenged person that too while moving then the image was bound to be affected with the issue of blur that too motion blur. Blurred input image can lead to the failure of the entire system as the blurred input image will lead to failure of the OCR to detect the text and if the text is not detected then it cannot be converted to voice output with the help of Text to Speech conversion. This situation will lead to failure of the entire framework. In the deblurring step our task was to convert the blurred input image to Non-blurred image for further processing. Then the image degradation and restoration model was studied in chapter 3 in order to come up with a new algorithm for text deblurring. Blind deconvolution image restoration method was used because the input to framework was a single blurred image and there was no prior information about the Point Spread Function of the blur nor there was any information about the blurring process. A new text deblurring algorithm using OCR word confidence to estimate the Point Spread Function was proposed in chapter 4. The PSF of the motion blur $h(x, y)$ is a function of two parameters that are blur length l and blur angle θ. Iteratively for different values of blur length l and blur angle θ estimation of the PSF was tried that can give the better recognition result of OCR to increase the reliability of the framework. A criteria was required based on which we can say that estimated PSF is the optimal and better than the other. This led to the generation of new parameter for evaluation of Text deblurring that was average word confidence AWC metric given by us. The higher value of AWC indicate that for the given PSF $h(x, y)$ we get a better recognition rate than the other with PSF with the lower value of AWC. The result shows that our average word confidence can be successfully used for the estimation of PSF of the motion blur using blur length l and blur angle θ.

Once the text deblurring algorithm was developed and applied in deblurring the textual input image. Then next task was to come up with a more reliable and efficient framework that can be used as by visually challenged with confidence. In chapter 5 A robust reading framework for visually challenged person was proposed which can provide then with higher degree of independence in there day to day life. This framework implements the Deblurring step and the preprocessing step before OCR.

After the deblurring process on the input image the latent image $f'(x, y)$ is obtained. As it is known that deblurring is an ill-poised problem and during the deblurring process noise gets amplified and the latent image $f'(x, y)$ thus obtained will be not the same as the original image. Hence before the OCR step the preprocessing steps has been introduced that will prepare the image for efficient Optical Character Recognition. Preprocessing also speed up the processing of the Optical Character Recognitions. For Optical Character Recognitions Tesseract OCR engine was used which is an open source OCR engine developed by HP between 1984 and 1994. Tesseract works on improving the rejection efficiency than on base-level accuracy. It performs the segmentation of the text with the help of connected components as it is simple and has the ability of detecting the inverse text.

Text are segmented into lines then into word and then into individual characters using definite spaces and fuzzy spaces. After the segmentation of the text into individual characters they are passed for the recognition of character. Recognition is two-pass process in which the satisfactory recognized word is passed to an adaptive classifier for training. This training increases the efficiency of the recognition. A second pass is run over the text, in which words that were not recognized well enough are recognized again. This increase the accuracy of the recognition as two passes of recognition are run on the same input text. Hence the recognition step is an important step of the framework so it handled with special care so that the reading framework is effective as well as robust.

After the Optical Character Recognition of the deblurred and preprocessed input image we get the output text file. Now our task is to convert this text file into voice output so that it could be communicated to the visually challenged person. In text to speech generation two major steps are followed that are preprocessing step and the speech generation step. In the preprocessing step spell checking for misrecognized words, text normalization for abbreviations and acronyms and various morphological operations are performed in order to enhance the text input so that it could give better speech output. Then the speech generation is performed to generate the voice output of the Text to Speech engine.

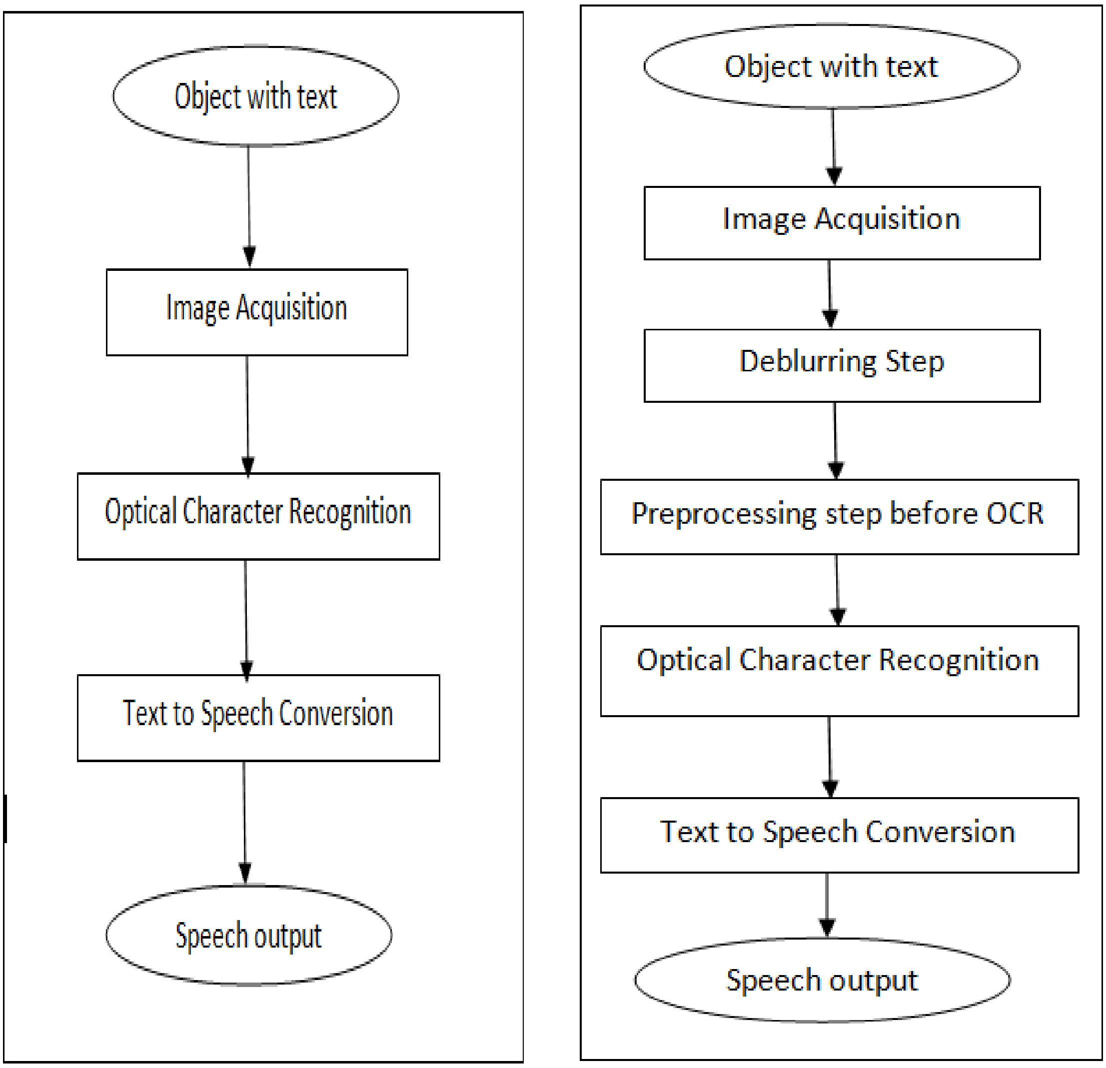
Text to Speech function was implemented in the programming using the Microsoft Win32 Speech API (SAPI) for the voice output. In this the values of various parameters that are set the voice available, pace and sampling rate of the speech output can be set. If these parameters are not set then the default values are used. Hence the Visually challenged person gets to hear the speech output of the captured textual information and in this way he can acquire the knowledge about the surrounding so that he can feel independent in his life.

6.1 COMPARISON WITH THE PREVIOUS APPROACHES

In the technologically Assisted reading framework for visually challenged we have improved the existing assistive technology for reading. In the previous approaches there were only the image capturing routine and then the Optical character

Recognition which generated the text file which was send to a Text to Speech engine for the generation of the speech which was communicated to the Visually Challenged person. The previous approaches assumed that the image will be an ideal image for Optical Character Recognition that will give the best recognition text output. But when the image is captured by the visually challenged person who is not an expert in capturing image as well as the person is having difficulty in viewing and the images are also captured on the move that person cannot capture the image with any clarity. Hence the captured image is bound to be affected with the issue of Blur, noise and different lightning conditions. These issues can fail the best possible Optical Character Recognition engine. These issues were not addressed in the previous approaches.

In Technologically assisted reading framework for visually challenged we not only addressed these issues but also have come up with two steps in our architecture to deal with the blur and the different contrast changes or noise using Deblurring step and preprocessing step before OCR. In our framework a new text deblurring algorithm using blind deconvolution method and average word confidence metric of the OCR. The figure 6.1 below shows the difference between the previous architecture in part (a) of the figure and part (b) shows our architecture for reading framework for Visually Challenged.



(a) **(b)**

Figure 6.1 Comparison between the Previous Architecture (a) and Our Architecture (b) for the reading framework for visually Challenged

Table 6.1 Shows the comparison of previous architecture and our Architecture using sample blurred input image.

	Previous Architecture	Our Architecture
Input Image		
Deblurring Step	Not Applicable	
Preprocessing Step	Not Applicable	 
OCR Output	No output	avinash Verma

From the table 6.1 clearly shows that for the blurred input image the previous approach did not had any measures to cope with them and hence the OCR generated the incorrect recognition result. This is not the desired result and then incorrect recognition is then passed to the next step Text to Speech conversion that will again give the undesired result. On the other hand our architecture has the measures to deblur the blurred input image and then the preprocessing step before OCR adds to the accuracy of the OCR recognition even after the input image is blurred. This shows that our framework is a robust reading framework for visually challenged as it can handle the blurred and noisy input image.

6.2 FUTURE WORK

In the future work for Technologically Assisted Reading framework for Visually Challenged research work on the text deblurring model and development of new techniques for the estimation of the Point Spread Function using blind Image restoration technique should be done because of few work in this field. The research work for various non- blind image restoration algorithms can be easily found but finding the algorithm for the blind image restoration technique that too for text deblurring is not an easy job. The artificial neural network can be thought of as the new solution to deblurring. As the part of future work some deblurring algorithm based on neural network need to be proposed. Recent study in the neural networks has popped a new concept that is Computational Neural Network (CNN). This CNN can be thought as the future solution for the development of new technique to deblurring the textual input image.

In this thesis the work was based on assistive technology for visually challenged for reading. As a part of future work the other aspects of visually challenged person's life need to be addressed such as vision, writing, movement, learning etc and assistive technology need to be developed for them. So that visually challenged person can live life independently without the support of others but with support of the technology. Vision of visually challenged person requires seeing and interpreting everyday objects. This capability is essential for the person to move around and understand the surroundings without the help of others. This will enhance the mobility of the person and self-confidence. Another important aspect that needs to be addressed is the writing independently for the visually challenged person in which they can write notes, documents, reports etc. without any regard to the size and shape of the paper. The capability includes electronic recording (e-writing) and manual writing. The Movement for visually challenged includes moving and working around in the houses, workspaces, and public places etc. Learning about new objects such as tower bolt, new concepts like height, color which can be perceived through vision. The strategy of converting the visual information to a stream of audio information requires formation of concepts and their representation through words. Considerable cognitive challenges are to be solved with so that the communication is effective. So a lot of work in the field of assistive technology can be done for the visually challenged person in the near future.

References

[1] World Blind Union *www.worldblindunion.org.*

[2] Marion A.Hersh, Michael A.Johnson Assistive technology for Visually Impaired and Blind people Book by Springer ISBN 978-1-84628-866.

[3] Sándor Tihamér BRASSAI, László BAKÓ, Lajos LOSONCZI, 2011. Assistive Technologies for Visually Impaired People. Acta Universitatis Sapientiae Electrical and Mechanical Engineering, 3 (2011) 39-50.

[4] Alves CCF, Monteiro GBM, Rabello S, Gasparetto MERF, Carvalho KM, 2009. Assistive technology applied to education of students with visual impairment. Rev Panam Salud Publica. 2009. 148–52.

[5] Mohd Helmy Abd Wahab, Amirul A. Talib, Herdawatie A. Kadir, Ayob Johari, A.Noraziah, Roslina M. Sidek, Ariffin A. Mutalib, 2011. Smart Cane: Assistive Cane for Visually-impaired People. *International Journal of Computer Science Issues,* Vol. 8, Issue 4, No 2, July 2011.

[6] Chaitali K. Lakde, Dr. Prakash S. Prasad, 2015. Review Paper on Navigation System for Visually Impaired People. *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 4, Issue 1, January 2015 DOI 10.17148/IJARCCE.2015.4134 166.

[7] R. Velázquez, 2010. Wearable Assistive Devices for the Blind. Chapter 17 in A. Lay-Ekuakille & S.C.Mukhopadhyay (Eds.), Wearable and Autonomous Biomedical Devices and Systems for Smart Environment: Issues and Characterization, LNEE 75, Springer, pp 331-349, 2010.

[8] Roberto Netoa, Nuno Fonsecaa 2014. Camera Reading For Blind People HCIST 2014 International Conference on Health and Social Care Information Systems and Technologies, Procedia Technology 16 (2014) 1200 – 1209 published by ELSEVIER.

[9] Abdel Ilah Nour Alshbatat, 2013. Automated Mobility and Orientation System for Blind or Partially Sighted People *INTERNATIONAL JOURNAL ON SMART SENSING AND INTELLIGENT SYSTEMS* VOL. 6, NO. 2, APRIL 2013

[10] Nihat Kocyigita, Pinar Sabuncu Artara 2015. A challenge: Teaching English to visually-impaired learners *Procedia - Social and Behavioral Sciences 199* (2015) Elsevier.

[11] Dr.Deepak Kumar Singh, Avinash Verma 2014. Reading framework for visually challenged" published at National Conference on Digital India and E-Governance DiGov-2014 by Computer society of India Lucknow Chapter.

[12] Sukhpreet Singh 2013.Optical Character Recognition Techniques :A Survey *CIS Journal of emerging trends in Computing and Information Science* Vol 4,June 6.

[13] Nabendu Chaki , Soharab Hossain Shaikh, Khalid Saeed, 2014. A Comprehensive Survey on Image Binarization Techniques. Chapter Exploring Image Binarization

Techniques Volume 560 of the series Studies in Computational Intelligence pp 5-15 Date: 15 May 2014 Springer India.

[14] C. Gonzalez Richard E. Woods Digital Image Processing Book Third Edition Rafael Interactive Pearson International Edition prepared by Pearson Education PEARSON Prentice Hall.

[15] Lukas Neumann and Jiri Matas. A method for text localization and recognition in real-world images published at the 10th Asian Conference on Computer Vision, Queenstown, New Zealand.

[16] Honggang Zhang,KailiZhao,Yi-ZheSong,Jun Guo 2013. A Text extraction from natural scene image:A survey" Neurocomputing 122 (2013) ELSEVIER.

[17] Rainer Lienhart, Member, IEEE, and Axel Wernicke 2002 Localizing and Segmenting Text in Images and Videos *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, VOL. 12, NO. 4, APRIL 2002.

[18] Jian Liang, Daniel DeManthon 2008. Geometric rectification of Camera-Captured Document Image IEEE Transaction on Pattern Analysis and Machine Intelligence Vol 30 No 4 April 2008.

[19] U. Pal, B.B. Chaudhuri Indian script character recognition: a survey Pattern Recognition 37 (2004) 1887 – 1899 published by ELSEVIER

[20] Taiebeh Askari Javara Hamid HassanpourVahid Abolghasemi, 2017. Automatic estimation and segmentation of partial blur in natural images. The Visual Computer February 2017, Volume 33, Issue 2, pp 151–161 DOI: 10.1007/s00371-015-1166-z. Springer.

[21] Faisal Shafait and Thomas M. Breuel, 2008. The Effect of Border Noise on the Performance of Projection-Based Page Segmentation Methods. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 33, NO. 4, APRIL 2011.

[22] Minako Sawaki and Norihiro Hagita, 1998. Text-Line Extraction and Character Recognition of Document Headlines With Graphical Designs Using Complementary Similarity Measure. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 20, NO. 10, OCTOBER 1998.

[23] Nanning Zheng, Qubo You, Gaofeng Meng, Jihua Zhu, Shaoyi Du, and Jianyi Liu, 2008. 50 Years of Image Processing and Pattern Recognition in China Published by the IEEE Computer Society IEEE INTELLIGENT SYSTEMS November/December 2008 IEEE.

[24] Tapas Kanungo, Member, Robert M. Haralick, 1992. An Automatic Closed-Loop Methodology for Generating Character Groundtruth for Scanned Documents. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 21, NO. 2, FEBRUARY 1999.

[25] Laurence Likforman-Sulem, Abderrazak Zahour, Bruno Taconet, 2006. Text Line Segmentation of Historical Documents: a Survey. submitted to Special Issue on Analysis of Historical Documents, *International Journal on Document Analysis and Recognition*, Springer, 2006.

[26] Suhas R. Mache,Manasi R. Baheti,C. Namrata Mahender 2015. Review on Text-To-Speech Synthesizer *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 4, Issue 8, August.

[27] Thierry Dutoit , 1997, An Introduction to Text-to-Speech Synthesis, BookText, Speech and Language Technology Volume 3 1997 ISBN: 978-1-4020-0369-1 (Print) 978-94-011-5730-8 (Online) DOI 10.1007/978-94-011-5730-8, Springer Netherlands.

[28] Kaveri Kamble, Ramesh Kagalkar 2014. A Review: Translation of Text to Speech Conversion for Hindi Language *International Journal of Science and Research (IJSR)* Volume 3 Issue 11, November 2014

[29] Itunuoluwa Isewon,Jelili Oyelade,Olufunke Oladipupo 2014. Design and Implementation of Text To Speech Conversion for Visually Impaired People *International Journal of Applied Information Systems Foundation of Computer Science* FCS, New York, USA Volume 7– No. 2, April 2014.

[30] K. Sreenivasa Rao, 2012 Predicting Prosody from Text for Text-to-Speech Synthesis, Book Springer Briefs in Electrical and Computer Engineering ISBN: 978-1-4614-1337-0 (Print) 978-1-4614-1338-7 (Online) DOI: 10.1007/978-1-4614-1338-7 Springer New York.

[31] Text-to-Speech (TTS) Synthesis Reference Work Entry Encyclopedia of Biometrics pp 1332-1332 DOI 10.1007/978-0-387-73003-5 Springer US

[32] Richard W. Sproat, The Analysis of Text in Text-to-Speech Synthesis, Chapter Progress in Speech Synthesis pp 73-75Section Introduction. DOI 10.1007/978-1-4612-1894-4_6 Springer New York.

[33] Akbar S. Shaik, G Hossain 2010. Desgin and Development and performance Evaluation of Reconfigured Mobile Android Phone for People Who are blind or Visually Impaired In Proceedings of the 28th ACM International Conference on Design of Communication,New York, NY, USA, 2010. ACM. October 2010.

[34] Akshay Sharma, Abhishek Srivastava, Adhar Vashishth 2014.An Assistive Reading System for Visually Impaired using OCR and TTS *International Journal of Computer Applications*, Vol 95, No. 2 June 2014.

[35] Chucai Yi, Yingli Tian and Aries Arditi. Portable Camera-Based Assistive Text and Product Label Reading from Hand Held Objects for Blind Persons IEEE/ASME Transaction on Mechatronics.

[36] Patrick E.Lanigan, Aaron M. Paulos, Andrew W. Williams, Dan Rossi and Priya Narasimhan 2006 .Trinetra: Assistive Technology for Grocery Shopping for the Blind Published by Cylab in 2006.

[37] Bindu Philips and R.D. Sudhaker Samuel 2009. Human Machine Interface – A smart OCR for the visually challenged , *International Journal of Recent Trends in Engineering*, Vol 2, No. 3 November 2009.

[38] Akella S Narasimha Raju, G Satya Mohan Chowdary, Ch.Karunakar, 2013. RF Communication Based Assistive Device for Visually Impaired person. *International Journal of Research in Computer and Communication Technology*, Vol 2, Issue 11, November- 2013

[39] Avinash Verma, Deepak Kumar Singh, Nitesh Kumar Singh 2016. Review on Assistive Reading framework for Visually Challenged published in Computational Intelligence in Data Mining pp 533-541 part of Advances in Intelligent Systems and Computing (AISC), vol.556. SPRINGER, Singapore.DOI: 10.1007/978-981-10-3874-7_50.

[40] Maria M. Martins, Cristina P. Santos, Anselmo Frizera.Neto, Ramón Ceres 2011. Assistive mobility devices focusing on Smart Walkers: Classification and review Robotics and Autonomous Systems 2011 published by ELSEVIER.

[41] Ray Smith Google Inc. theraysmith@gmail.com 2007. An Overview of the Tesseract OCR Engine Google Inc OSCON 2007.

[42] www.ravi.iiit.ac.in.

[43] Mausumi Acharyya and Malay K. Kundu, 2002. Document Image Segmentation Using Wavelet Scale–Space Features. IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 12, NO. 12, DECEMBER 2002.

[44] R. Jayadevan, Satish R. Kolhe, Pradeep M. Patil, and Umapada Pal, 2011. Offline Recognition of Devanagari Script: A Survey. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 41, NO. 6, NOVEMBER 2011.

[45] Gian Marco Bo, Daniele D. Caviglia, Member, IEEE, and Maurizio Valle, 1998. An Analog VLSI Implementation of a Feature Extractor for RealTime Optical Character Recognition. *IEEE JOURNAL OF SOLID-STATE CIRCUITS*, VOL. 33, NO. 4, APRIL 1998.

[46] A.Conci, J. E. R. de Carvalho, T. W. Rauber, 2009. A Complete System for Vehicle Plate Localization, Segmentation and Recognition in Real Life Scene. IEEE LATIN AMERICA TRANSACTIONS, VOL. 7, NO. 5, SEPTEMBER 2009.

[47] Alessandro Vinciarelli and Jean-Marc Odobez, 2006 Application of Information Retrieval Technologies to Presentation Slides. IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 8, NO. 5, OCTOBER 2006.

[48] Jian Liang David DoermannHuiping Li, 2005. Camera-based analysis of text and documents: a survey. *International Journal of Document Analysis and Recognition (IJDAR)* July 2005, Volume 7, Issue 2, pp 84–104 Springer-Verlag Berlin/Heidelberg.

[49] Amandeep Kaur Renu DhirGurpreet Singh Lehal, 2017. A survey on camera-captured scene text detection and extraction: towards Gurmukhi script. *International Journal of Multimedia Information Retrieval* June 2017, Volume 6, Issue 2, pp 115–142 Springer-Verlag London 2017.

[50] Ragavi K., Radja P., Chithra S. ,2016, Portable Text to Speech Converter for the Visually Impaired. In: Suresh L., Panigrahi B. (eds) Proceedings of the International Conference on Soft Computing Systems. Advances in Intelligent Systems and Computing, vol 397. Springer, New Delhi

[51] Neumann L., Matas J. ,2012, A Real-Time Scene Text to Speech System. In: Fusiello A., Murino V., Cucchiara R. (eds) Computer Vision – ECCV 2012. Workshops and Demonstrations. ECCV 2012. Lecture Notes in Computer Science, vol 7585. Springer, Berlin, Heidelberg.

[52] Gore Megha B., Surwase S. V., Kailash. J. Karande 2014. Technical Aspects in Development of Marathi Calculator using Text to Speech Synthesizer 2014 IEEE Global Conference on Wireless Computing and Networking (GCWCN).

[53] Sangheeta Roy, Palaiahnakote Shivakumara, Partha Pratim Roy, Umapada Pal, Chew Lim Tan,Tong Lu 2015. Bayesian classifier for multi-oriented video text recognition system *Expert Systems with Applications 42* (2015) ELSEVIER

[54] Goto H., Hoda T. ,2014, Real-Time Text Tracking for Text-to-Speech Translation Camera for the Blind. In: Miesenberger K., Fels D., rchambault D., Penáz P., Zagler W. (eds) Computers Helping People with Special Needs. ICCHP 2014. Lecture Notes in Computer Science, vol 8547. Springer, Cham.

[55] Goto H., Liu K. , 2016, Scene Text Detection and Tracking for Wearable Text-to-Speech Translation Camera. In: Miesenberger K., Bühler C., Penaz P. (eds) Computers Helping People with Special Needs. ICCHP 2016. Lecture Notes in Computer Science, vol 9759. Springer, Cham

[56] Sorin L., Lemarié J., Aussenac-Gilles N., Mojahid M., Oriola B. ,2014, Communicating Text Structure to Blind People with Text-to-Speech. In: Miesenberger K., Fels D.,

Archambault D., Penáz P., Zagler W. (eds) Computers Helping People with Special Needs. ICCHP 2014. Lecture Notes in Computer Science, vol 8547. Springer, Cham

[57] Henry S. Baird, The State of the Art of Document Image Degradation Modelling Chapter Digital Document Processing Part of the series Advances in Pattern Recognition pp 261-279 Springer london.

[58] Anna Tonazzini Stefano VezzosiLuigi Bedini, 2003. Analysis and recognition of highly degraded printed characters. Document Analysis and Recognition April 2003, Volume 6, Issue 4, pp 236–247 Springer-Verlag Berlin/Heidelberg 2003.

[59] Shamik Tiwari,V. P. Shukla, S. R. Biradar, A. K. Singh, 2014. Blur parameters identification for simultaneous defocus and motion blur. CSI Transactions on ICT March 2014, Volume 2, Issue 1, pp 11–22 DOI: 10.1007/s40012-014-0039-3 Springer.

[60] Aizenberg I., Bregin T., Butakoff C., Karnaukhov V., Merzlyakov N., Milukova O. (2002) Type of Blur and Blur Parameters Identification Using Neural Network and Its Application to Image Restoration. In: Dorronsoro J.R. (eds) Artificial Neural Networks — ICANN 2002. ICANN 2002. Lecture Notes in Computer Science, vol 2415. Springer, Berlin, Heidelberg.

[61] Peleshko D., Rashkevych M., Klyuvak A., Ivanov Y. (2014) Partial Blur: Model, Detection, Deblurring. In: Zamojski W., Mazurkiewicz J., Sugier J., Walkowiak T., Kacprzyk J. (eds) Proceedings of the Ninth International Conference on Dependability and Complex Systems DepCoS-RELCOMEX. June 30 – July 4, 2014, Brunów, Poland. Advances in Intelligent Systems and Computing, vol 286. Springer, Cham.

[62] Flusser J., Suk T., Saic S. ,1996. Recognition of Images Degraded by Linear Motion Blur without Restoration. In: Kropatsch W., Klette R., Solina F., Albrecht R. (eds) Theoretical Foundations of Computer Vision.Computing Supplement, vol 11. Springer, Vienna.

[63] Oliveira J.P., Figueiredo M.A.T., Bioucas-Dias J.M. , 2007. Blind Estimation of Motion Blur Parameters for Image Deconvolution. In: Martí J., Benedí J.M., Mendonça A.M., Serrat J. (eds) Pattern Recognition and Image Analysis. IbPRIA 2007. Lecture Notes in Computer Science, vol 4478. Springer, Berlin, Heidelberg

[64] Brusius F., Schwanecke U., Barth P. ,2013. Blind Image Deconvolution of Linear Motion Blur. In: Csurka G., Kraus M., Mestetskiy L., Richard P., Braz J. (eds) Computer Vision, Imaging and Computer Graphics. Theory and Applications. VISIGRAPP 2011. Communications in Computer and Information Science, vol 274. Springer, Berlin, Heidelberg.

[65] Shiqian Wu, Shoulie Xie, Weisi Lin, Blind Measurement of Image Blur for Vision-Based Applications. Chapter Multimedia Analysis, Processing and Communications Volume 346 of the series Studies in Computational Intelligence pp 185-215 Springer Berlin Heidelberg.

[66] Ratnakar Dash 2012. Parameter Estimation for Image Restoration PhD thsis NIT Rourkela, Orrisa,India March 2012.

[67] Gwanggil Jeon, Sang-Jun Park, Abdellah Chehri, Junsang Cho, Jechang Jeong. Color Image Restoration Technique Using Gradient Edge Direction Detection. Chapter Intelligent Interactive Multimedia Systems and Services Volume 6 of the series Smart Innovation, Systems and Technologies pp 101-109 Springer Berlin Heidelberg.

[68] P. S. Considine, R. A. Gonsalves, 2006. Optical image enhancement and image restoration. Chapter Optical Data Processing Volume 23 of the series Topics in Applied Physics pp 53-88 Date: 27 June 2006 DOI: 10.1007/BFb0057983 Springer Berlin Heidelberg.

[69] Xia Y.S., Bin S.Q.,2014. A Novel Blind Image Restoration Algorithm Using A SVR-Based Noise Reduction Technique. In: Sun F., Hu D., Liu H. (eds) Foundations and Practical Applications of Cognitive Systems and Information Processing. Advances in Intelligent Systems and Computing, vol 215. Springer, Berlin, Heidelberg.

[70] Konstantinos N. Plataniotis, Anastasios N. Venetsanopoulos, Color Image Enhancement and Restoration. Chapter Color Image Processing and Applications Part of the series Digital Signal Processing pp 209-235 DOI 10.1007/978-3-662-04186-4_5 Springer, Berlin, Heidelberg.

[71] Liu Puyin Email authorLi Hongxing,2002. Image restoration techniques based on fuzzy neural networks. Science in China Series F: Information Sciences August 2002, Volume 45, Issue 4, pp 273–285 Springer.

[72] Nam-Yong Lee 2015. Block-iterative Richardson-Lucy methods for image deblurring *Lee EURASIP Journal on Image and Video Processing (2015)* Springer 2015:14 DOI 10.1186/s13640-015-0069-2.

[73] Yang HL., Chiao YH., Huang PH., Lai SH. (2011) Blind Image Deblurring with Modified Richardson-Lucy Deconvolution for Ringing Artifact Suppression. In: Ho YS. (eds) Advances in Image and Video Technology. PSIVT 2011. Lecture Notes in Computer Science, vol 7088. Springer, Berlin, Heidelberg.

[74] A. G. Al-BakkarS. S. IpsonG. J. PorterD. G. Gleed,1997. A parallel implementation of a modified Richardson-Lucy algorithm for image de-blurring. International Journal of Infrared and Millimeter Waves March 1997, Volume 18, Issue 3, pp 555–575. DOI: 10.1007/BF02677995 Springer.

[75] Subhasis Chaudhuri, Rajbabu Velmurugan, Renu Rameshan, 2014. Blind Deconvolution Methods: A Review. Chapter Blind Image Deconvolution pp 37-60 Date: 22 August 2014 DOI: 10.1007/978-3-319-10485-0_3 Springer.

[76] Taeg Sang Cho,Sylvain Paris,Berthold K. P. Horn,William T. Freeman 2011. Blur Kernel Estimation using the Radon Transform *Computer Vision and Pattern Recognition (CVPR), 2011* IEEE Conference on June 2011.

[77] Alan Bundy, Lincoln Wallen. Generalised Hough Transform. Chapter Catalogue of Artificial Intelligence Tools. DOI 10.1007/978-3-642-96868-6_84 Springer Berlin Heidelberg.

[78] Wayne NiblackDragutin Petkovic,1990. On improving the accuracy of the Hough transform. Machine Vision and Applications March 1990, Volume 3, Issue 2, pp 87–106 DOI: 10.1007/BF01212193 Springer.

[79] Wen-Ze Shao, Qi GeHai-Song DengZhi-Hui WeiHai-Bo Li, 2016. A Unified Optimization Perspective to Single/Multi-observation Blur-Kernel Estimation with Applications to Camera-Shake Deblurring and Nonparametric Blind Super-Resolution. Journal *of Mathematical Imaging and Vision* February 2016, Volume 54, Issue 2, pp 216–239 DOI: 10.1007/s10851-015-0598-4Springer.

[80] Avinash Verma, Deepak Kumar Singh 2017. Text Deblurring Using OCR Word Confidence, *International Journel of Image, Graphics and Signal Processing (IJIGSP),* Vol.9, No.1, pp. 33-40, 2017, DOI :10.5815/ijigsp.2017.01.05.

[81] Shamik Tiwari, V. P. Shukla, and A. K. Singh 2013. Review of Motion Blur Estimation Techniques *Journal of Image and Graphics* Vol. 1, No. 4, December 2013.

[82] Kishore R. Bhagat,Puran Gour 2013. Novel Approach to Estimate Motion Blur Kernel Parameters and Comparative Study of Restoration Techniques *International Journal of Computer Applications* (0975 – 8887) Volume 72– No.17, June 2013.

[83] Lu Fang, Haifeng Liu, Feng Wu, Xiaoyan Sun, Houqiang Li 2014. Separable Kernel for Image Deblurring IEEE Conference on Computer Vision and Pattern Recognition Pages: 2885 - 2892, DOI: 10.1109/CVPR.2014.369.

[84] Nimali Rajakaruna,Chamila Rathnayake,Kit Yan Chan,Iain Murray 2014 Image deblurring for navigation systems of vision impaired people using sensor fusion data. Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2014 IEEE Ninth International Conference on Year: 2014.

[85] Jing Wang, Ke Lu, Qian Wang, and Jie Jia 2012 Kernel Optimization for Blind Motion Deblurring with Image Edge Prior. *Hindawi Publishing Corporation Mathematical Problems in Engineering* Volume 2012, Article ID 639824, 10 pages doi:10.1155/2012/639824.

[86] Long Mai,Feng Liu 2015. Kernel fusion for better image deblurring. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Pages: 371 - 380, DOI: 10.1109/CVPR.2015.7298634.

[87] Jinshan Pan, Zhe Hu,Zhixun Su,Ming-Hsuan Yang 2014. Deblurring Text Images via L0-Regularized Intensity and Gradient Prior. IEEE Conference on Computer Vision and Pattern Recognition Pages: 2901 - 2908, DOI: 10.1109/CVPR.2014.371.

[88] Linyang He, Gang Li,and Jinghong Liu 2015. Joint Motion Deblurring and Superresolution from Single Blurry Image. *Hindawi Publishing Corporation Mathematical Problems in Engineering* Volume 2015, Article ID 965690.

[89] Neeraj Kumar, Rahul Nallamothu, Amit Sethi 2012. Neural Network Based Image Deblurring. *Neural Network Applications in Electrical Engineering (NEUREL)*, 2012 11th Symposium on Year: 2012 IEEE Pages: 219 - 222, DOI: 10.1109/NEUREL.2012.6420015.

[90] Jian Sun; Wenfei Cao; Zongben Xu; Jean Ponce 2015. Learning a convolutional neural network for non-uniform motion blur removal. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Year: 2015 Pages: 769-777, DOI: 10.1109/CVPR.2015.729867.

[91] Michal Hradis, Jan Kotera, Pavel Zemˇcík, Filip Šroubek 2015. Convolutional Neural Networks for Direct Text Deblurring. British Machine Vision Conference 2015 September.

[92] Li Xu, Jimmy SJ. Ren, Ce Liu, Jiaya Jia 2015. Deep Convolutional Neural Network for Image Deconvolution" conference on Advances in Neural Information Processing Systems 27 (NIPS 2014).

[93] Zohair Al-Ameen,Ghazali Bin Sulong,Md. Gapar Md. Johar 2013. Computer Forensics and Image Deblurring: An Inclusive Investigation. *IJMECS* Vol.5, No. 11, November 2013 PP.42-48, DOI: 10.5815/ijmecs.2013.11.06.

[94] Avinash Verma, Dr.Deepak Kumar Singh, 2017. Robust Assistive Reading Framework for Visually Challenged. communicated to *Internation Journel of Image, Graphics and Signal Processing* (IJIGSP).

[95] Zeng Fanfeng North China University Of Technology NCUT Beijing, China 2010. Application research of Voice control in Reading Assistive device for visually impaired persons. 2010 International Conference on Multimedia Information Networking and Security,2010 IEEE

[96] Tae Hyun Kim and Kyoung Mu Lee 2015. Generalized Video Deblurring for Dynamic Scenes. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Year: 2015 Pages: 5426 - 5434, DOI: 10.1109/CVPR.2015.7299181.

[97] Wei-Sheng Lai; Jian-Jiun Ding; Yen-Yu Lin; Yung-Yu Chuang 2015. Blur kernel estimation using normalized color-line priors. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Year: 2015 Pages: 64 - 72, DOI: 10.1109/CVPR.2015.7298601.

[98] J.Amudha, N.Pradeepa, R.Sudhakar 2012. A Survey on Digital Image Restoration. *ELSEVIER Procedia Engineering*, Volume 38, 2012, Pages 2378-2382.

[99] Haichao Zhang,David Wipf,Yanning Zhang 2013. Multi-image Blind Deblurring Using a Coupled Adaptive Sparse Prior. Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on Year: 2013 Pages: 1051 - 1058, DOI: 10.1109/CVPR.2013.140.

[100] Alex Krizhevsky , Ilya Sutskever , Geoffrey E. Hinton 2012. Imagenet classification with deep convolutional neural networks. *CiteSeer Document Advances in Neural Information Processing Systems 2012*.

[101] Oliver Whyt Josef SivicAndrew Zisserman, 2014. Deblurring Shaken and Partially Saturated Images. International Journal of Computer Vision November 2014, Volume 110, Issue 2, pp 185–201 DOI: 10.1007/s11263-014-0727-3 Springer.